Le Frido 2017,
volume 2
Laurent Claessens

Plusieurs extensions et versions de ce livre.

1. La version courante, régulièrement mise à jour et qui deviendra petit à petit le Frido 2018. Téléchargeable sur

   `https://laurent.claessens-donadello.eu/pdf/lefrido.pdf`

2. La version la plus complète, contenant des exercices ainsi que de la mathématique de niveau recherche sur

   `https://laurent.claessens-donadello.eu/pdf/mazhe.pdf`

3. Et bien entendu les sources LaTeX sur github :

   `https://github.com/LaurentClaessens/mazhe`

Copyright 2011-2017 Laurent Claessens, Carlotta Donadello, Lilian Besson
Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the chapter entitled "GNU Free Documentation License".

(c) 2015 David Revoy pour les illustrations de couverture CC-BY,
`https://www.peppercarrot.com/`

ISBN : 979-10-97085-06-3

Thèmes

Ceci est une sorte d'index thématique.

1 : intégration
2 : suites et séries
3 : polynôme de Taylor
4 : séries de Fourier
5 : normes
6 : topologie produit
7 : espaces métriques, normés
8 : norme opérateur
9 : gaussienne
10 : produit de compact
11 : densité
12 : espaces de fonctions
13 : fonctions Lipschitz
14 : formule des accroissements finis
15 : différentiabilité
16 : points fixes
17 : théorème de Stokes, Green et compagnie
18 : permuter des limites
19 : applications continues et bornées
20 : inégalités
21 : connexité
22 : suite de Cauchy, espace complet
23 : application réciproque
24 : tribu, algèbre de parties, λ-systèmes et co.
25 : déduire la nullité d'une fonction depuis son intégrale
26 : mesure et intégrale
27 : équations différentielles
28 : injections
29 : logarithme
30 : inversion locale, fonction implicite
31 : convexité
32 : dualité
33 : opérations sur les distributions
34 : transformée de Fourier
35 : convolution
36 : méthode de Newton
37 : méthodes de calcul
38 : définie positive
39 : norme matricielle et rayon spectral
40 : série de matrices
41 : rang
42 : extension de corps et polynômes
43 : décomposition de matrices
44 : systèmes d'équations linéaires
45 : racines de polynôme et factorisation de polynômes
46 : formes bilinéaires et quadratiques
47 : arithmétique modulo, théorème de Bézout
48 : polynômes
49 : invariants de similitude
50 : diagonalisation
51 : endomorphismes cycliques
52 : déterminant
53 : polynôme d'endomorphismes
54 : exponentielle de matrice
55 : sous-groupes
56 : groupe symétrique
57 : action de groupe
58 : classification de groupes
59 : produit semi-direct de groupes
60 : théorie des représentations
61 : isométries
62 : caractérisation de distributions en probabilités
63 : théorème central limite
64 : lemme de transfert
65 : probabilités et espérances conditionnelles
66 : dénombrements
67 : enveloppes
68 : équations diophantiennes

Thème 1 : intégration

(1) L'existence d'une primitive pour toute fonction continue est le théorème 18.10.

(2) La définition d'une primitive est la définition 12.135.

(3) Primitive et intégrale, proposition 13.204.

(4) Intégrale impropre, définition 13.217.

Thème 2 : suites et séries De façon un peu contre-intuitive, les suites et séries sont surtout traitées dans le chapitre sur les espaces vectoriels. Section 10.5 pour les suites et 10.6 pour les séries.

(1) La définition de la somme d'une infinité de termes est donnée par la définition 10.43.

(2) La définition de la convergence absolue est la définition 10.45.

(3) Convergence d'une série implique convergence vers zéro du terme général, proposition 10.53.

(4) Quelque séries usuelles dans $\mathbb{R}$ dans la section 10.7.3.

Pour les sommes infinies l'associativité et la commutativité dans une série sont perdues. Néanmoins, il subsiste que

(1) si la série converge, on peut regrouper ses termes sans modifier la convergence ni la somme (associativité) ;

(2) si la série converge absolument, on peut modifier l'ordre des termes sans modifier la convergence ni la somme (commutativité, proposition 9.176).

Une somme indexée par un ensemble quelconque est la définition 9.178.

Thème 3 : polynôme de Taylor Énoncés :

(1) Énoncé : théorème 12.263.

(2) De classe C^2 sur $\mathbb{R}^n$, proposition 12.267.

(3) Avec un reste donné par un point dans $]x, a[$, proposition 12.271.

(4) Le polynôme de Taylor généralise à l'utilisation de toutes les dérivées disponibles le résultat de développement limité donné par la proposition 12.120.

(5) Pour les fonctions holomorphes, il y a le théorème 25.27 qui donne une série de Taylor sur un disque de convergence.

Utilisation :

(1) Il est utilisé pour justifier la méthode de Newton autour de l'équation (32.92).

(2) On utilise pas mal de Taylor dans les résultats liant extrema et différentielle/Hessienne. Par exemple la proposition 18.68.

Thème 4 : séries de Fourier

— Formule sommatoire de Poisson, proposition 27.9.

— Inégalité isopérimétrique, théorème 26.21.

— Fonction continue et périodique dont la série de Fourier ne converge pas, proposition 26.18.

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

(1) Si f est continue et périodique, convergence au sens de Cesaro, théorème de Fejèr 26.5.

(2) Convergence au sens $L^2\Big([0,2\pi]\Big)$ dans le théorème 24.57.

(3) Si f est continue, périodique et sa série de Fourier converge uniformément, théorème 26.13.

(4) Si f est périodique et la série des coefficients converge absolument pour tout x, proposition 26.14.

(5) Si f est périodique et de classe C^1, théorème 26.15.

Il est cependant faux de croire que la continuité et la périodicité suffisent à obtenir une convergence, comme le montre dans la proposition 26.18.

Thème 5 : normes

Définition définition 6.31.

Équivalence de norme (1) Définition de l'équivalence de norme 10.1.

(2) La proposition 10.2 sur l'équivalence des normes dans $\mathbb{R}^n$.

(3) Toutes les normes sur un espace vectoriel de dimension finie sont équivalentes, théorème 10.3.

(4) Montrer que le problème $a - b$ est stable dans l'exemple 32.26.

(5) La proposition 10.15 donnant $\rho(A) \leqslant \|A\|$ utilise l'équivalence de toutes les normes sur un espace vectoriel de dimension finie.

Norme opérateur voir le thème 8.

Thème 6 : topologie produit

(1) La définition de la topologie produit est 6.5.

(2) Pour les espaces vectoriels normés, le produit est donné par la définition 10.73.

(3) L'équivalence entre la topologie de la norme produit et la topologie produit est le lemme 10.75.

Thème 7 : espaces métriques, normés

(1) Le théorème-définition 6.19 donne la topologie sur un espace métrique.

(2) La définition de la convergence d'une suite est la définition 6.4.

(3) Dans un espace vectoriel normé, une application est continue si et seulement si elle est bornée, proposition 10.18.

Thème 8 : norme opérateur Pour la norme matricielle et le rayon spectral, voir le thème 39.

(1) Définition 10.5.

(2) Définition d'une norme d'algèbre 10.9.

(3) Pour des espaces vectoriels normée, être borné est équivalent à être continu : proposition 10.18.

(4) Le lemme à propos d'exponentielle de matrice 15.94 donne :

$$\|e^{tA}\| \leqslant P\big(|t|\big) \sum_{i=1}^{r} e^{t\,\mathrm{Re}(\lambda_i)}. \tag{0.1}$$

Thème 9 : gaussienne

(1) Le calcul de l'intégrale

$$\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi} \tag{0.2}$$

est fait de deux façons dans l'exemple 13.233. Les deux utilisent le théorème de Fubini 13.230.

(2) Le lemme 27.15 calcule la transformée de Fourier de $g_\epsilon(x) = e^{-\epsilon\|x\|^2}$ qui donne $\hat{g}_\epsilon(\xi) = \big(\frac{\pi}{\epsilon}\big)^{d/2} e^{-\|\xi\|^2/4\epsilon}$.

(3) Le lemme 27.18 donne une suite régularisante à base de gaussienne.

(4) Elle est utilisée pour régulariser une intégrale dans la preuve de la formule d'inversion de Fourier 27.20

Thème 10 : produit de compact

(1) Les produits d'espaces métriques compacts sont compacts ; c'est le théorème de Tykhonov. Nous verrons ce résultat dans les cas suivants.
— $\mathbb{R}$, lemme 6.122.
— Produit fini d'espaces métriques compacts, théorème 6.162.
— Produit dénombrable d'espaces métrique compacts, théorème 6.164.

Thème 11 : densité

(1) Densité des polynômes dans $C^0\big([0,1]\big)$, théorème de Bernstein 34.133.

(2) Densité de $\mathscr{D}(\mathbb{R}^d)$ dans $L^p(\mathbb{R}^d)$ pour $1 \leqslant p < \infty$, théorème 24.33.

(3) Densité de $\mathscr{S}(\mathbb{R}^d)$ dans l'espace de Sobolev $H^s(\mathbb{R}^d)$, proposition 29.17.

(4) Densité de $\mathscr{D}(\mathbb{R}^d)$ dans l'espace de Sobolev $H^s(\mathbb{R}^d)$, proposition 29.19.
Cela est utilisé pour le théorème de trace 29.21.

(5) Les applications étagées dans les applications mesurables (qui plus est avec limite croissante), théorème fondamental d'approximation 24.36.
(6) Les fonctions continues à support compact dans $L^2(I)$, théorème 24.37.
(7) Les polynômes trigonométriques sont denses dans $L^p(S^1)$ pour $1 \leqslant p < \infty$. Deux démonstrations indépendantes par le théorème 26.6 et le théorème 24.51.

Les densités sont bien entendu utilisées pour prouver des formules sur un espace en sachant qu'elles sont vraies sur une partie dense. Mais également pour étendre une application définie seulement sur une partie dense. C'est par exemple ce qui est fait pour définir la trace γ_0 sur les espaces de Sobolev $H^s(\mathbb{R}^d)$ en utilisant le théorème d'extension 18.112.

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

— Le lemme 18.1 le donne pour la racine carré.
— Le théorème 18.6 donne la densité des polynômes dans les fonctions continues sur un compact.
— Le théorème 18.4 est une généralisation qui donne la densité uniforme d'une sous-algèbre de $C(X, \mathbb{R})$ dès que X sépare les points.
— Le lemme 26.1 est une version pour les polynômes trigonométriques.
— Le lemme 18.1 est une cas particulier du théorème 18.6, mais nous en donnons une démonstration indépendante afin d'isoler la preuve de la généralisation 18.4. Une version pour les polynômes trigonométriques sera donnée dans le lemme 26.1.

Thème 12 : espaces de fonctions En ce qui concerne les densités, voir le thème 11.

Topologie Les espaces de fonctions sont souvent munis de topologies définies par des semi-normes.

(1) La topologie des semi-normes est la définition 6.190.
(2) La définition 28.3 donne les topologies sur $C^\infty(\Omega)$, $\mathscr{D}(K)$ et $\mathscr{D}(\Omega)$.
(3) La topologie $*$-faible sur $\mathscr{D}'(\Omega)$ est donnée par la définition 28.10.

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

(1) Définition de $L^2(\Omega, \mu)$, 24.53.
(2) Le produit scalaire $\langle f, g\rangle$ est donné en (24.187) et la base trigonométrique est (24.188).
(3) La densité des polynômes trigonométriques dans $L^p(S^2)$ est le théorème 24.51 ou le théorème 26.6, au choix.
(4) Une conséquence de cette densité est que le système trigonométrique est une base hilbertienne de L^2 par le lemme 24.56.

L'espace L^2 est discuté en analyse fonctionnelle, en 24.5 parce que l'étude de L^2 utilise entre autres l'inégalité de Hölder 24.22.

Le fait que L^2 soit une espace de Hilbert est utilisé dans la preuve du théorème de représentation de Riesz 24.60.

Thème 13 : fonctions Lipschitz

(1) Définition : 12.243.
(2) La notion de Lipschitz est utilisée pour définir la stabilité d'un problème, définition 32.25.

Thème 14 : formule des accroissements finis Il en existe plusieurs formes :

(1) Une version adaptée aux espaces de dimension finie est le théorème 12.241.
(2) Pour les fonctions $\mathbb{R} \to \mathbb{R}$ en le théorème 12.129.
(3) Une généralisation pour les intervalles non bornés : théorème 12.130.
(4) Espaces vectoriels normés, théorème 12.229

Thème 15 : différentiabilité

(1) La recherche d'extrema d'une fonction sur $\mathbb{R}^n$ passe par la seconde différentielle, proposition 18.68.

(2) Lien entre différentielle seconde (Hessienne) et convexité en la proposition 18.90 et le corollaire 18.92.

(3) La différentielle est liée aux dérivées partielles par les formules données au lemme 12.178

$$df_a(u) = \frac{\partial f}{\partial u}(a) = \frac{d}{dt}\Big[f(a+tu)\Big]_{t=0} = \sum_{i=1}^{m} u_i \frac{\partial f}{\partial x_i}(a) = \nabla f(a) \cdot u. \tag{0.3}$$

Thème 16 : points fixes

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

Théorème de Picard 18.23 donne un point fixe comme limite d'itérés d'une fonction Lipschitz. Il aura pour conséquence le théorème de Cauchy-Lipschitz 18.37, l'équation de Fredholm, théorème 18.28 et le théorème d'inversion locale dans le cas des espaces de Banach 18.47.

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

(a) Dans $\mathbb{R}^n$ en version C^∞ via le théorème de Stokes, proposition 18.31.

(b) Dans $\mathbb{R}^n$ en version continue, en s'appuyant sur le cas C^∞ et en faisant un passage à la limite, théorème 18.32.

(c) Dans $\mathbb{R}^2$ via l'homotopie, théorème 25.19. Oui, c'est très loin. Et c'est normal parce que ça va utiliser la formule de l'indice qui est de l'analyse complexe [1].

Théorème de Markov-Kakutani 18.35 qui donne un point fixe à une application continue d'un convexe fermé borné dans lui-même. Ce théorème donnera la mesure de Haar 18.36 sur les groupes compacts.

Théorème de Schauder 18.33 qui est une version valable en dimension infinie du théorème de Brouwer.

(2) Pour les équations différentielles

(a) Le théorème de Schauder a pour conséquence le théorème de Cauchy-Arzela 18.43 pour les équations différentielles.

(b) Le théorème de Schauder 18.33 permet de démontrer une version du théorème de Cauchy-Lipschitz (théorème 18.37) sans la condition Lipschitz, mais alors sans unicité de la solution. Notons que de ce point de vue nous sommes dans la même situation que la différence entre le théorème de Brouwer et celui de Picard : hors hypothèse de type «contraction», point d'unicité.

(3) En calcul numérique

— La convergence d'une méthode de point fixe est donnée par la proposition 32.47.
— La convergence quadratique de la méthode de Newton est donnée par le théorème 32.53.
— En calcul numérique, section 32.5
— Méthode de Newton comme méthode de point fixe, sous-section 32.6.2.

(4) D'autres utilisation de points fixes.

— Processus de Galton-Watson, théorème 36.47.
— Dans le théorème de Max-Milgram 23.56, le théorème de Picard est utilisé.

Thème 17 : théorème de Stokes, Green et compagnie

(1) Forme générale, théorème 14.48.

(2) Rotationnel et circulation, théorème 22.8.

Le théorème de Stokes peut être utilisé pour montrer le théorème de Brower, proposition 18.31.

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

Thème 18 : permuter des limites

(1) Les théorèmes sur les fonctions définies par des intégrales, section 14.10. Nous avons entre autres
   (a) $\partial_i \int_B f = \int_B \partial_i f$, avec B compact, proposition 14.72.
   (b) Si f est majorée par une fonction ne dépendant pas de x, nous avons le théorème 14.62.
   (c) Si l'intégrale est uniformément convergente, nous avons le théorème 14.63.
   (d) Pour dériver $\int_B g(t,z)dt$ avec B compact dans $\mathbb{R}$ et $g\colon \mathbb{R}\times\mathbb{C}\to\mathbb{C}$, il faut aller voir la proposition 25.17.

(2) Théorème de la convergence monotone, théorème 13.129.

(3) Le théorème de Fubini permet non seulement de permuter des intégrales, mais également des sommes parce que ces dernières peuvent être vues comme des intégrales sur $\mathbb{N}$ muni de la tribu des parties et de la mesure de comptage. Nous utilisons cette technique pour permute une somme et une intégrale dans l'équation (25.122).
   — le théorème de Fubini-Tonelli 13.227 demande que la fonction soit mesurable et positive ;
   — le théorème de Fubini 13.230 demande que la fonction soit intégrable (mais pas spécialement positive) ;
   — le corollaire 13.229 demande l'intégrabilité de la valeur absolue des intégrales partielles pour déduire que la fonction elle-même est intégrable.

Thème 19 : applications continues et bornées

(1) Une application linéaire non continue : exemple 10.20 de $e_k \mapsto ke_k$. Les dérivées partielles sont calculées en (23.131).

(2) Une autre application linéaire non continue en l'exemple 10.21 : la dérivation sur les polynômes.

(3) Une application linéaire est bornée si et seulement si elle est continue, proposition 10.18.

Thème 20 : inégalités

Inégalité de Jensen (1) Une version discrète pour $f\big(\sum_i \lambda_i x_i\big)$, la proposition 18.93.

(2) Une version intégrale pour $f\big(\int \alpha d\mu\big)$, la proposition 24.20.

(3) Une version pour l'espérance conditionnelle, la proposition 34.52.

Inégalité de Minkowsky (1) Pour une forme quadratique q sur $\mathbb{R}^n$ nous avons $\sqrt{q(x+y)} \leqslant \sqrt{q(x)} + \sqrt{q(y)}$. Proposition 9.254.

(2) Si $1 \leqslant p < \infty$ et si $f, g \in L^p(\Omega, \mathcal{A}, \mu)$ alors $\|f+g\|_p \leqslant \|f\|_p + \|g\|_p$. Proposition 24.26.

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

$$\Big[\int_X \Big(\int_Y |f(x,y)|d\nu(y)\Big)^p d\mu(x)\Big]^{1/p} \leqslant \int_Y \Big(\int_X |f(x,y)|^p d\mu(x)\Big)^{1/p} d\nu(y). \tag{0.4}$$

ou sous forme compacte

$$\Big\| x \mapsto \int_Y f(x,y)d\nu(y)\Big\|_p \leqslant \int_Y \|f_y\|_p d\nu(y) \tag{0.5}$$

Transformée de Fourier Pour tout $f \in L^1(\mathbb{R}^n)$ nous avons $\|\hat{f}\|_\infty \leqslant \|f\|_1$, lemme 27.6.

Inégalité des normes Inégalité de normes : si $f \in L^p$ et $g \in L^1$, alors $\|f * g\|_p \leqslant \|f\|_p \|g\|_1$, proposition 24.40.

Thème 21 : connexité

(1) Définition 6.84
(2) Le groupe $\mathrm{SL}(n, \mathbb{K})$ est connexe par arcs : proposition 12.333.
(3) Le groupe $\mathrm{GL}(n, \mathbb{C})$ est connexe par arcs : proposition 12.334.
(4) Le groupe $\mathrm{GL}(n, \mathbb{C})$ est connexe par arcs, proposition 12.334.
(5) Le groupe $\mathrm{GL}(n, \mathbb{R})$ a exactement deux composantes connexes par arcs, proposition 12.335.
(6) Le groupe $\mathrm{O}(n, \mathbb{R})$ n'est pas connexe, lemme 12.329.
(7) Les groupe $\mathrm{U}(n)$ et $\mathrm{SU}(n)$ sont connexes par arcs, lemme 12.330.
(8) Le groupe $\mathrm{SO}(n)$ est connexe mais ce n'est pas encore démontré, proposition 12.331.
(9) Connexité des formes quadratiques de signature donnée, proposition 18.101.

Thème 22 : suite de Cauchy, espace complet Nous parlons d'espaces topologiques complets. À ne pas confondre avec un espace mesuré complet, définition 13.50.

(1) La définition 6.39 donne la notion de suite de Cauchy dans un espace métrique.
(2) La définition 6.37 donne la notion de suite de τ-Cauchy dans un espace vectoriel topologique.
(3) Deux espaces métriques (avec une distance) peuvent être isomorphes en tant qu'espaces topologiques, mais ne pas avoir les mêmes suites de Cauchy, exemple 6.42.
(4) La proposition 6.52 donne l'équivalence entre les suites de Cauchy et les suites τ-Cauchy dans le cas des espaces vectoriels topologiques *normés*.
(5) L'exemple 6.42 est un exemple pire que simplement une suite de Cauchy qui ne converge pas. Le problème de convergence de cette suite n'est pas simplement que la limite n'est pas dans l'espace ; c'est que la suite de Cauchy donnée ne convergerait même pas dans $\mathbb{R}$.
(6) Le théorème 18.119 est un théorème de complétion d'un espace métrique.

Quelque espaces qui sont complets sont listés ci-dessous. Attention : la complétude est bien une propriété de la norme ; le même ensemble peut être complet pout une norme et pas pour une autre. Si on vous demande «est-ce que tel espace est complet ?» vous devez répondre en demandant «pour quelle topologie ?» ... sauf si la topologie est évidente dans le contexte.

(1) Les réels $\mathbb{R}$, théorème 6.57.
(2) La proposition 12.299 donne quelque espaces complets. Soit X un espace topologique métrique (Y, d) un espace espace métrique complet. Alors les espaces
   (a) $\left(C_b^0(X, Y), \|.\|_\infty\right)$
   (b) $\left(C_0^0(X, Y), \|.\|_\infty\right)$
   (c) $\left(C_0^k(X, Y), \|.\|_\infty\right)$
   sont complets.
(3) Le lemme 12.300 dit que $\left(C^0(A, B), \|.\|_\infty\right)$ est complet dès que A est compact et B est complet.
(4) L'espace $\mathscr{D}(K)$ est complet tant pour la topologie des semi-normes que pour la topologie métrique (qui sont les mêmes). C'est la proposition 28.7.
(5) L'espace $\mathscr{S}(\Omega)$ est complet et métrisable par la proposition 28.48.

La limite uniforme d'une suite de fonctions dérivables n'est pas spécialement dérivable. Même si les fonctions sont de classe C^∞, la limite n'est pas spécialement mieux que continue. En effet, le théorème de Stone-Weierstrass 18.6 nous dit que les polynômes (qui sont C^∞) sont denses dans les fonctions continues sur un compact pour la norme uniforme. Vous ne pouvez donc pas espérer que $\left(C^p(X, Y), \|.\|_\infty\right)$ soit complet en général.

Thème 23 : application réciproque

(1) Définition 6.187.
(2) Dans le cas des réels, des exemples sont donnés en 12.7.
(3) Continuité, proposition 6.189.
(4) Théorème de la bijection 12.41 (qui contient aussi de la continuité).
(5) Dérivabilité, proposition 12.42.

Thème 24 : tribu, algèbre de parties, λ-systèmes et co. Il existe des centaines de notions de mesures et de classes de parties.

(1) Le plus souvent lorsque nous parlons de mesure est que nous parlons de mesure positive, définition 13.23 sur un espace mesuré avec une tribu, définition 13.1.
(2) Une mesure extérieure est la définition 13.17
(3) Une algèbre de partie : définition 13.18. Une mesure sur une algèbre de parties : définition 13.20. L'intérêt est que si on connait une mesure sur une algèbre de parties, elle se prolonge en une mesure sur la tribu engendrée par le théorème de prolongement de Hahn 13.60.
(4) Un λ-système : définition 13.37.
(5) Une mesure complexe : définition 13.147.

Thème 25 : déduire la nullité d'une fonction depuis son intégrale Des résultats qui disent que si $\int f = 0$ c'est que $f = 0$ dans un sens ou dans un autre.

(1) Il y a le lemme 13.123 qui dit ça.
(2) Un lemme du genre dans L^2 existe aussi pour $\int f\varphi = 0$ pour tout φ. C'est le lemme 24.43.
(3) Et encore un pour L^p dans la proposition 24.65.
(4) Si $\int f\chi = 0$ pour tout χ à support compact alors $f = 0$ presque partout, proposition 24.1.
(5) La proposition 24.17 donne $f = 0$ dans L^p lorsque $\int fg = 0$ pour tout $g \in L^q$ lorsque l'espace est σ-fini.
(6) Une fonction $h \in C_c^\infty(I)$ admet une primitive dans $C_c^\infty(I)$ si et seulement si $\int_I h = 0$. Théorème 18.11.

Thème 26 : mesure et intégrale

(1) Mesure de Lebesgue, définition 13.172
(2) Intégrale associée à une mesure, définition 13.113
(3) Mesure à densité, définition 13.137.

Thème 27 : équations différentielles L'utilisation des théorèmes de point fixe pour l'existence de solutions à des équations différentielles est fait dans le chapitre sur les points fixes.

(1) Le théorème de Schauder a pour conséquence le théorème de Cauchy-Arzela 18.43 pour les équations différentielles.
(2) Le théorème de Schauder 18.33 permet de démontrer une version du théorème de Cauchy-Lipschitz (théorème 18.37) sans la condition Lipschitz
(3) Le théorème de Cauchy-Lipschitz 18.37 est utilisé à plusieurs endroits :
— Pour calculer la transformée de Fourier de $e^{-x^2/2}$ dans le lemme 27.15.
(4) Théorème de stabilité de Lyapunov 30.34.
(5) Le système proie prédateurs, Lokta-Voltera 30.35
(6) Équation de Schrödinger, théorème 30.41.
(7) L'équation $(x - x_0)^\alpha u = 0$ pour $u \in \mathscr{D}'(\mathbb{R})$, théorème 28.56.
(8) La proposition 30.37 donne un résultat sur $y'' + qy = 0$ à partir d'une hypothèse de croissance.
(9) Équation de Hill $y'' + qy = 0$, proposition 30.39.

Thème 28 : injections

(1) L'espace de Sobolev $H^1(I)$ s'injecte de façon compacte dans $C^0(\bar{I})$, proposition 29.8.
(2) L'espace de Sobolev $H^1(I)$ s'injecte de façon continue dans $L^2(I)$, proposition 29.8.
(3) L'espace $L^2(\Omega)$ s'injecte continument dans $\mathscr{D}'(\Omega)$ (les distributions), proposition 28.13.

Thème 29 : logarithme

(1) L'exemple 15.63 donne une série pour $\ln(1-x)$.

(2) La proposition 9.196 dit que toute matrice complexe admet un logarithme. En particulier une série explicite est donné pour le logarithme d'un bloc de Jordan.

(3) Le logarithme pour les réels strictement positifs est donné en la définition 15.57.

(4) Sur les complexes, le logarithme $\ln\colon \mathbb{C}^* \to \mathbb{C}$ est la définition 25.54. Attention : ce n'est pas la seule définition possible.

(5) La série harmonique diverge à vitesse logarithmique, et la série des inverses des nombres premiers, c'est encore plus lent : théorème 15.68.

Thème 30 : inversion locale, fonction implicite

Des énoncés (1) Inversion locale dans $\mathbb{R}^n$: théorème 18.46. Pour un Banach c'est le théorème 18.47.

(2) Fonction implicite dans un Banach : théorème 18.48.

Des utilisations (1) Utilisé pour montrer que le flot d'une équation différentielle est un C^p-difféomorphisme local, voir 30.30.

(2) Pour le théorème de Von Neumann 18.60.

Thème 31 : convexité L'essentiel des résultats sur les fonctions convexes sont dans la sections 18.9.

(1) Définition des fonctions convexes : 18.74 et 18.88 en dimension supérieure.

(2) En termes de différentielles, 18.89 pour la différentielle première et 18.92 pour la Hessienne.

(3) Une courbe paramétrée convexe est la définition 19.84.

(4) L'enveloppe convexe d'une courbe fermée simple et convexe : 19.86.

(5) Courbure et convexité d'une courbe paramétrée : section 19.12.4.

(6) Une courbe paramétrée convexe est localement le graphe d'une fonction convexe, lemme 19.85.

(7) La convexité est utilisée dans la méthode du gradient à pas optimal de la proposition 18.98.

Thème 32 : dualité Ne pas confondre dual algébrique et dual topologique d'un espace vectoriel.

Dual topologique et algébrique Ils sont définis par 7.40. Le dual algébrique est l'ensemble des formes linéaires, et le dual topologique ne considère que les formes linéaires continues (en dimension infinie, les applications linéaires ne sont pas toutes continues).

Topologie Une topologie possible sur le dual d'un espace vectoriel topologique est celle $*$-faible de la définition 6.196.

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

(1) Le théorème de représentation de Riesz 23.18 pour les espaces de Hilbert.

(2) La proposition 24.60 pour les espaces $L^p\big([0,1]\big)$ avec $1 < p < 2$.

(3) Le théorème de représentation de Riesz 24.64 pour les espaces L^p en général.

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

Thème 33 : opérations sur les distributions

(1) Convolution d'une distribution par une fonction, définition par l'équation (28.74).

(2) Dérivation d'une distribution, proposition-définition 28.15.

(3) Produit d'une distribution par une fonction, définition 28.14.

Thème 34 : transformée de Fourier

(1) Définition sur L^1, définition 27.2.

(2) La transformée de Fourier d'une fonction $L^1(\mathbb{R}^d)$ est continue, proposition 27.5.

(3) L'espace de Schwartz est stable par transformée de Fourier. L'application $\mathcal{F}\colon \mathscr{S}(\mathbb{R}^d) \to \mathscr{S}(\mathbb{R}^d)$ est une bijection linéaire et continue. Proposition 27.12

Thème 35 : convolution

(1) Définition pour $f, g \in L^1$, théorème 24.38.

(2) Inégalité de normes : si $f \in L^p$ et $g \in L^1$, alors $\|f * g\|_p \leqslant \|f\|_p \|g\|_1$, proposition 24.40.

(3) $\varphi \in L^1(\mathbb{R})$ et $\psi \in \mathscr{S}(\mathbb{R})$, alors $\varphi * \psi \in \mathscr{S}(\mathbb{R})$, proposition 24.87.

(4) Les suites régularisantes : $\lim_{n\to\infty} \rho_n * f = f$ dans la proposition 27.17.

(5) Convolution d'une distribution par une fonction, définition par l'équation (28.74).

Thème 36 : méthode de Newton

(1) Nous parlons un petit peu de méthode de Newton en dimension 1 dans 32.6.

(2) La méthode de Newton fonctionne bien avec les fonction convexes par la proposition 32.55.

(3) La méthode de Newton en dimension n est le théorème 32.61.

(4) Un intervalle de convergence autour de α s'obtient par majoration de $|g'|$, proposition 32.47.

(5) Un intervalle de convergence quadratique s'obtient par majoration de $|g''|$, théorème 32.53.

(6) En calcul numérique, section 32.6.

Thème 37 : méthodes de calcul

(1) Théorème de Rothstein-Trager 14.90.

(2) Algorithme des facteurs invariants 4.99.

(3) Méthode de Newton, théorème 32.61

(4) Calcul d'intégrale par suite équirépartie 26.8.

Thème 38 : définie positive

(1) Une application bilinéaire est définie positive lorsque $g(u,u) \geqslant 0$ et $g(u,u) = 0$ si et seulement si $u = 0$ est la définition 9.23.

(2) Un opérateur ou une matrice est défini positif si toutes ses valeurs propres sont positives, c'est la définition 9.167.

(3) Pour une matrice symétrique, définie positive implique $\langle Ax, x\rangle > 0$ pour tout x. C'est le lemme 9.170.

(4) Une application linéaire est définie positive si et seulement si sa matrice associée l'est. C'est la proposition 9.247.

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

Thème 39 : norme matricielle et rayon spectral

(1) Définition du rayon spectral 10.10.

(2) Lien entre norme matricielle et rayon spectral, le théorème 10.25 assure que $\|A\|_2 = \sqrt{\rho(A^tA)}$.

(3) Lien entre valeurs propres et norme opérateur : le lemme 10.26 pour les matrices symétriques strictement définies positives donne $\|A\|_2 = \lambda_{max}$.

(4) Pour toute norme algébrique nous avons $\rho(A) \leqslant \|A\|$, proposition 10.15.

(5) Dans le cadre du conditionnement de matrice. Voir en particulier la proposition 32.104 qui utilise le théorème 10.25.

(6) Rayon spectral et convergence de méthode itérative, proposition 32.140.

Une norme matricielle donne une topologie. Il y a donc également des liens entre rayon spectral et convergence de série. Dans cette optique, pour les séries de matrices, voir le thème 40.

Thème 40 : série de matrices

(1) Rayon spectral et norme opérateur : thème 39.

(2) Exponentielle de matrices : thème 54.

(3) Série entière de matrices : section 15.12.

(4) Pour la série $\sum_k A^k = (1-A)^{-1}$.
— Pour un espace de Banach : proposition 10.52.
— Pour les matrices nilpotentes : proposition 9.138.
— En lien avec le rayon spectral (si et seulement si $\rho(A) < 1$) dans la proposition 15.90.
— Le lemme 15.44 parle de la série entière $\sum_{n\in\mathbb{N}} z^{nk} = (1-z^k)^{-1}$.
Cette série est utilisée entre autres dans la proposition 32.160 pour prouver qu'une M-matrice irréductible vérifie $A^{-1} > 0$.

Thème 41 : rang

(1) Définition 7.22.

(2) Le théorème du rang, théorème 7.23

(3) Prouver que des matrices sont équivalentes et les mettre sous des formes canoniques, lemme 7.28 et son corollaire 7.29.

(4) Tout hyperplan de $\mathbb{M}(n,\mathbb{K})$ coupe $\mathrm{GL}(n,\mathbb{K})$, corollaire 7.29. Cela utilise la forme canonique sus-mentionnée.

(5) Le lien entre application duale et orthogonal de la proposition 7.45 utilise la notion de rang.

(6) Prouver les équivalences à être un endomorphisme cyclique du théorème 9.233 via le lemme 9.232.

Thème 42 : extension de corps et polynômes

(1) Définition d'une extension de corps 5.47.

(2) Pour l'extension du corps de base d'un espace vectoriel et les propriétés d'extension des applications linéaires, voir la section 9.13.

(3) Extension de corps de base et similitude d'application linéaire (ou de matrices, c'est la même chose), théorème 9.235.

(4) Extension de corps de base et cyclicité des applications linéaires, corollaire 9.234.

(5) À propos d'extensions de $\mathbb{Q}$, le lemme 5.111.

Thème 43 : décomposition de matrices

(1) Décomposition de Bruhat, théorème 12.352.

(2) Décomposition de Dunford, théorème 9.194.

(3) Décomposition polaire 12.346 et la proposition 18.56 pour la régularité.

Thème 44 : systèmes d'équations linéaires

— Algorithme des facteurs invariants 4.99.
— La méthode du gradient à pas optimal permet de résoudre par itérations $Ax = b$ lorsque A est symétrique strictement définie positive. Il s'agit de minimiser une fonction bien choisit. Propositions 18.97 pour l'existence et 18.98 pour la méthode.

Thème 45 : racines de polynôme et factorisation de polynômes

(1) Si $\mathbb{A}$ est une anneau, la proposition 4.119 factorise une racine.

(2) Si $\mathbb{A}$ est un anneau, la proposition 4.121 factorise une racine avec sa multiplicité.

(3) Si $\mathbb{A}$ est un anneau, le théorème 4.123 factorise plusieurs racines avec leurs multiplicités.

(4) Si $\mathbb{K}$ est un corps et α une racine dans une extension, le polynôme minimal de α divise tout polynôme annulateur par la proposition 5.65.

(5) Le théorème 5.68 annule un polynôme de degré n ayant $n+1$ racines distinctes.

(6) La proposition 5.114 nous annule un polynôme à plusieurs variables lorsqu'il a trop de racines.

(7) En analyse complexe, le principe des zéros isolés 25.20 annule en gros toute série entière possédant un zéro non isolé.

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

Thème 46 : formes bilinéaires et quadratiques

(1) Les formes bilinéaires ont été définies en 9.22.

(2) Forme quadratique, définition 9.242

Thème 47 : arithmétique modulo, théorème de Bézout

(1) Pour $\mathbb{Z}^*$ c'est le théorème 3.32.

(2) Théorème de Bézout dans un anneau principal : corollaire 4.77.

(3) Théorème de Bézout dans un anneau de polynômes : théorème 5.40.

(4) En parlant des racines de l'unité et des générateurs du groupe unitaire dans le lemme 3.99. Au passage nous y parlerons de solfège.

(5) La proposition 3.34 qui donne tout entier assez grand comme combinaisons de a et b à coefficients positifs est utilisée en chaînes de Markov, voir la définition 36.41 et ce qui suit.

(6) Calcul effectif du PGCD puis des coefficients de Bézout : sous-sections 3.5.4.1 et 3.5.4.2.

Thème 48 : polynômes

Coefficients dans un anneau commutatif (1) Les polynômes à coefficients dans un anneau commutatif sont à la section 4.10.

Coefficients dans un corps (1) Les polynômes à coefficients dans un corps sont à la section 5.3.

(2) Nous parlons de l'idéal des polynômes annulateurs dans le théorème 5.37.

(3) Le théorème 5.37 dit que $\mathbb{K}[X]$ est une anneau principal et que tous ses idéaux sont engendrés par un unique polynôme unitaire.

(4) Le polynôme minimal est irréductible, proposition 5.50.

Polynôme primitif (1) Un polynôme est irréductible sur A si et seulement si irréductible et primitif sur le corps des fractions, corollaire 4.124.

Thème 49 : invariants de similitude

(1) Théorème 9.221.

(2) Pour prouver que la similitude d'applications linéaires résiste à l'extension du corps de base, théorème 9.235.

(3) Pour prouver que la dimension du commutant d'un endomorphisme de E est de dimension au moins $\dim(E)$, lemme 9.232.

(4) Nous verrons dans la remarque 9.222 à propos des invariants de similitude que toute matrice est semblable à la matrice bloc-diagonale constituées des matrices compagnon (définition 9.216) de la suite des polynômes minimals.

Thème 50 : diagonalisation Des résultats qui parlent diagonalisation

(1) Définition d'un endomorphisme diagonalisable : 9.140.

(2) Conditions équivalentes au fait d'être diagonalisable en termes de polynôme minimal, y compris la décomposition en espaces propres : théorème 9.143.

(3) Diagonalisation simultanée 9.146, pseudo-diagonalisation simultanée 9.173.

(4) Diagonalisation d'exponentielle 9.197 utilisant Dunford.

(5) Décomposition polaire théorème 12.346. $M = SQ$, S est symétrique, réelle, définie positive, Q est orthogonale.

(6) Décomposition de Dunford 9.194. $u = s + n$ où s est diagonalisable et n est nilpotent, $[s, n] = 0$.

(7) Réduction de Jordan (bloc-diagonale) 9.224.

(8) L'algorithme des facteurs invariants 4.99 donne $U = PDQ$ avec P et Q inversibles, D diagonale, sans hypothèse sur U. De plus les éléments de D forment une chaîne d'éléments qui se divisent l'un l'autre.

Le théorème spectral et ses variantes :

(1) Théorème spectral, matrice symétrique, théorème 9.165. Via le lemme de Schur.

(2) Théorème spectral autoadjoint (c'est le même, mais vu sans matrices), théorème 9.261

(3) Théorème spectral hermitien, lemme 9.159.

(4) Théorème spectral, matrice normales, théorème 9.162.

Pour les résultats de décomposition dont une partie est diagonale, voir le thème 43 sur les décompositions.

Thème 51 : endomorphismes cycliques

(1) Définition 9.113.

(2) Son lien avec le commutant donné dans la proposition 9.230 et le théorème 9.233.

(3) Utilisation dans le théorème de Frobenius (invariants de similitude), théorème 9.221.

Thème 52 : déterminant

(1) Les n-formes alternées forment un espace de dimension 1, proposition 9.3.

(2) Déterminant d'une famille de vecteurs 9.6.

(3) Calcul d'un déterminant de taille 2×2 : équation (9.39).

(4) Déterminant d'un endomorphisme 9.9.

(5) Interprétations géométriques

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

      i. Liaison au produit vectoriel (orthogonalité) dans la proposition 9.58.

      ii. En particulier le lemme 9.59 nous dit comment un déterminant donne un vecteur orthogonal à une famille donnée de vecteurs.

   (b) Déterminant et aires, volumes

      i. Déterminant et mesure de Lebesgue : théorème 13.223.

      ii. Aire du parallélogramme : proposition 16.49.

      iii. Volume du parallélépipède avec le produit mixte et le déterminant 3×3, 16.50.

Tant que nous en sommes dans les interprétations géométrique, il faut lier déterminant, produit vectoriel, orthogonalité et mesure en notant que l'élément de volume lors de l'intégration en dimension 3 est donné par (14.117) : $dS = \|T_u \times T_v\|$ qui est la norme du produit vectoriel des vecteurs tangents à la paramétrisation.

Nous voyons dans l'équation (14.114) que l'élément de volume pour une partie de dimension n dans $\mathbb{R}^m$ (à l'occasion d'y intégrer une fonction) est donné par un déterminant mettant en jeu les vecteurs tangents de la paramétrisation.

(6) Le déterminant de Vandermonde est à la proposition 9.12. Il est utilisé à divers endroits :
   (a) Pour prouver que $\mathrm{Tr}(u^p) = 0$ pour tout p si et seulement si u est nilpotente (lemme 9.136).
   (b) Pour prouver qu'un endomorphisme possédant $\dim(E)$ valeurs propres distinctes est cyclique (proposition 9.230).

Thème 53 : polynôme d'endomorphismes

(1) Endomorphismes cycliques et commutant dans le cas diagonalisable, proposition 9.230.
(2) Racine carré d'une matrice hermitienne positive, proposition 12.341.
(3) Théorème de Burnside sur les sous groupes d'exposant fini de $\mathrm{GL}(n, \mathbb{C})$, théorème 9.238.
(4) Décomposition de Dunford, théorème 9.194.
(5) Algorithme des facteurs invariants 4.99.

Thème 54 : exponentielle de matrice

(1) Le lemme à propos d'exponentielle de matrice 15.94 donne :

$$\|e^{tA}\| \leqslant P\big(|t|\big) \sum_{i=1}^{r} e^{t\,\mathrm{Re}(\lambda_i)}. \tag{0.6}$$

(2) La proposition 9.197 : si $A \in \mathbb{M}(n, \mathbb{R})$ a un polynôme caractéristique scindé, alors A est diagonalisable si et seulement si e^A est diagonalisable.
(3) La section 15.12.3 parle des fonctions exponentielle et logarithme pour les matrices. Entre autres la dérivation et les séries.
(4) Pour résoudre des équations différentielles linéaires : sous-section 30.6.1.
(5) La proposition 9.196 dit que l'exponentielle est surjective sur $\mathrm{GL}(n, \mathbb{C})$.
(6) La proposition 10.89 : si u est un endomorphisme, alors $\exp(u)$ est un polynôme en u.
(7) Calcul effectif : sous-section 9.8.2.
(8) Proposition 12.338 : si $A \in \mathbb{M}(n, \mathbb{C})$ alors $e^{\mathrm{Tr}(A)} = \det(e^A)$.

Thème 55 : sous-groupes

(1) Théorème de Burnside sur les sous groupes d'exposant fini de $\mathrm{GL}(n, \mathbb{C})$, théorème 9.238.
(2) Sous-groupes compacts de $\mathrm{GL}(n, \mathbb{R})$, lemme 12.353 ou proposition 12.354.

Thème 56 : groupe symétrique

(1) Définition 3.90.
(2) La table des caractères du groupe symétrique S_4 est donné dans la section 17.5.
(3) Le groupe symétrique S_4 est le groupe des symétries affines du tétraèdre régulier, proposition 16.17.
(4) Le groupe alterné A_5 est l'unique groupe simple d'ordre 60, proposition 8.49.
(5) La proposition 8.41 donne la position du groupe alterné dans le groupe symétrique : A_n est un sous-groupe caractéristique de S_n et l'unique sous-groupe d'indice 2.

Thème 57 : action de groupe

(1) Définition d'une action de groupe sur un ensemble : 3.72.
(2) Action du groupe modulaire sur le demi-plan de Poincaré, théorème 21.93.
(3) La formule de Burnside (théorème 3.85) parle du nombre d'orbites pour l'action d'un groupe fini sur un ensemble fini.
(4) Des applications de la formule de Burnside : le jeu de la roulette et l'affaire du collier, 16.4.4.1 et 16.4.4.2.

Thème 58 : classification de groupes

(1) Structure des groupes d'ordre pq, théorème 8.26.
(2) Le groupe alterné est simple, théorème 8.46.
(3) Définition 8.6 d'un p-groupe.
(4) Théorème de Sylow 8.11.
(5) Théorème de Burnside sur les sous groupes d'exposant fini de $\mathrm{GL}(n, \mathbb{C})$, théorème 9.238.
(6) $(\mathbb{Z}/p\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z}$, corollaire 5.131.

Thème 59 : produit semi-direct de groupes

(1) Définition 3.93.
(2) Le corollaire 3.95 donne un critère pour prouver qu'un produit NH est un produit semi-direct.
(3) L'exemple 16.102 donne le groupe des isométries du carré comme un produit semi-direct.
(4) Le théorème 8.26 classifie les groupes d'ordre pq (p, q premiers distincts) à grands coups de produit semi-directs.
(5) Le théorème 16.6 donne les isométries de $\mathbb{R}^n$ par $\mathrm{Isom}(\mathbb{R}^n) = T(n) \times_\rho O(n)$ où $T(n)$ est le groupe des translations.
(6) La proposition 16.8 donne une décomposition du groupe orthogonal $\mathrm{O}(n) = \mathrm{SO}(n) \times_\rho C_2$ où $C_2 = \{\mathrm{Id}, R\}$ où R est de déterminant -1.
(7) La proposition 11.53 donne $\mathrm{Aff}(\mathbb{R}^n) = T(n) \times_\rho \mathrm{GL}(n, \mathbb{R})$ où $\mathrm{Aff}(\mathbb{R}^n)$ est le groupe des applications affines bijectives de $\mathbb{R}^n$.

Thème 60 : théorie des représentations

(1) Table des caractères du groupe diédral, section 17.6.
(2) Table des caractères du groupe symétrique S_4, section 17.5.

Thème 61 : isométries Ne pas confondre une isométrie d'un espace affine avec une isométrie d'un espace euclidien. Les isométrie d'un espace euclidien préservent le produit scalaire et fixent donc l'origine (lemme 9.28). Les isométrie des espaces affines par contre conservent les distances (définition 11.54) et peuvent donc déplacer l'origine de l'espace vectoriel sur lequel il est modelé ; typiquement les translations sont des isométries de l'espace affine mais pas de l'espace euclidien.

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

(1) Définition d'une isométrie pour une forme bilinéaire, 9.27. Pour une forme quadratique : définition 16.2.
(2) Définition du groupe orthogonal 9.48, et le spécial orthogonal $\mathrm{SO}(n)$ en la définition 9.51. Le groupe $\mathrm{SO}(2)$ est le groupe des rotations, par corollaire 16.81.
(3) Le lemme 16.113 donne à toute rotation une matrice de la forme connue. C'est autour de cela que nous définissons les angles.

(4) Le groupe orthogonal est le groupe des isométries de $\mathbb{R}^n$, proposition 9.50.

(5) Les isométries de l'espace euclidien sont affines, 16.4.

(6) Les isométries de l'espace euclidien comme produit semi-direct : $\mathrm{Isom}(\mathbb{R}^n) \simeq T(n) \times_\rho \mathrm{O}(n)$, théorème 16.6.

(7) Isométries du cube, section 10.3.

(8) Générateurs du groupe diédral, proposition 16.100.

Thème 62 : caractérisation de distributions en probabilités

(1) La probabilité conjointe est la définition 34.19.

(2) La fonction de répartition est la définition 34.53.

(3) La fonction caractéristique est la définition 34.55.

Thème 63 : théorème central limite

(1) Pour les processus de Poisson, théorème 38.5.

Thème 64 : lemme de transfert Il s'agit du résultat $\hat{f}' = i\xi\hat{f}$.

(1) Lemme 27.11 sur $\mathscr{S}(\mathbb{R}^d)$

(2) Lemme 29.12 pour L^2.

Thème 65 : probabilités et espérances conditionnelles Les deux définitions de base, sur lesquelles se basent toutes les choses conditionnelles sont :

— La probabilité conditionnelle d'un événement en sachant un autre : $P(A|B)$ de la définition 34.29.

— L'espérance conditionnelle d'une variable aléatoire sachant une tribu : $E(X|\mathcal{F})$ de la définition 34.31.

Les autres sont listées ci-dessous.

La probabilité conditionnelle d'un événement par rapport à un autre donnée dans la proposition 34.29 est le nombre

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{0.7}$$

La probabilité conditionnelle d'un événement vis-à-vis d'une variable aléatoire discrète est par la définition 34.47 la variable aléatoire donnée par

$$P(A|X)(\omega) = P(A|X = X(\omega)). \tag{0.8}$$

Dans le cas continu, c'est la définition 34.48 :

$$P(A|X) = P(A|\sigma(X)) = E(\mathbb{1}_A|\sigma(X)). \tag{0.9}$$

L'espérance conditionnelle d'une variable aléatoire par rapport à une tribu $E(X|\mathcal{F})$ est la variable aléatoire $\mathcal{F}$-mesurable telle que

$$\int_B E(X|\mathcal{F}) = \int_B X \tag{0.10}$$

pour tout $X \in \mathcal{F}$. Si $X \in L^2(\Omega, \mathcal{A}, P)$ alors $E(X|\mathcal{F}) = \mathtt{proj}_K(X)$ où K est le sous-ensemble de $L^2(\Omega, \mathcal{A}, P)$ des fonctions $\mathcal{F}$-mesurables (théorème 34.31). Cela au sens des projections orthogonales.

La probabilité conditionnelle d'un événement par rapport à une tribu est la variable aléatoire

$$P(A|\mathcal{F}) = E(\mathbb{1}_A|\mathcal{F}). \tag{0.11}$$

L'espérance conditionnelle d'une variable aléatoire par rapport à une autre de la définition 34.33 est une variation sur le thème :

$$E(X|Y) = E(X|\sigma(Y)), \tag{0.12}$$

Notons que partout, si X est une variable aléatoire, la notation «sachant X» est un raccourcis pour dire «sachant la tribu engendrée par X».

Thème 66 : dénombrements

— Coloriage de roulette (16.4.4.1) et composition de colliers (16.4.4.2).
— Nombres de Bell, théorème 15.69.
— Le dénombrement des solutions de l'équation $\alpha_1 n_1 + \dots \alpha_p n_p = n$ utilise des séries entières et des décomposition de fractions en éléments simples, théorème 15.47.

Thème 67 : enveloppes

(1) L'ellipse de John-Loewner donne un ellipsoïde de volume minimum autour d'un compact dans $\mathbb{R}^n$, théorème 18.100.
(2) Le cercle circonscrit à une courbe donne un cercle de rayon minimal contenant une courbe fermée simple, proposition 19.83.
(3) Enveloppe convexe du groupe orthogonal 12.351.
(4) Enveloppe convexe d'une courbe fermée plane comme intersection des demi-plans tangents, proposition 19.89.

Thème 68 : équations diophantiennes

(1) Équation $ax + by = c$ dans $\mathbb{N}$, équation (3.74).
(2) Dans 3.5.7, nous résolvons $ax + by = c$ en utilisant Bézout (théorème 3.32).
(3) L'exemple 4.89 donne une application de la pure notion de modulo pour $x^2 = 3y^2 + 8$. Pas de solutions.
(4) L'exemple 4.90 résout l'équation $x^2 + 2 = y^3$ en parlant de l'extension $\mathbb{Z}[i\sqrt{2}]$ et de stathme.
(5) Les propositions 4.94 et 4.96 parlent de triplets pythagoriciens.
(6) Le dénombrement des solutions de l'équation $\alpha_1 n_1 + \dots \alpha_p n_p = n$ utilise des séries entières et des décomposition de fractions en éléments simples, théorème 15.47.
(7) La proposition 2.18 donne une bijection $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ en résolvant dans $\mathbb{N}$ (entre autres) l'équation $k = y^2 + x$ pour k fixé.

Table des matières

Thématique **2**

Table des matières **19**

Index **46**

Liste des notations **68**

1 Introduction (Vol 1) **73**
1.1 Préface . . . 73
1.2 Auteurs, contributeurs, sources et remerciements . . . 73
1.2.1 Ceux qui ont travaillé sur le Frido . . . 73
1.2.2 Aide directe, mais pas volontairement sur le Frido . . . 74
1.2.3 Des gens qui ont fait un travail qui m'a bien servi . . . 74
1.3 Les choses qui doivent vous faire tiquer . . . 75
1.4 Les questions pour lesquelles je n'ai pas (encore) de réponse . . . 75
1.4.1 Mes questions d'algèbre, géométrie. . . . 75
1.4.2 Mes questions de probabilité et statistiques. . . . 77
1.4.3 Les preuves à relire par des experts . . . 77
1.5 Comment contribuer et aider (math) ? . . . 82
1.5.1 Niveau facile . . . 82
1.5.1.1 Géométrie . . . 83
1.5.2 Niveau moyen . . . 83
1.5.2.1 Questions d'analyse . . . 83
1.5.2.2 Questions de géométrie . . . 84
1.5.2.3 Questions de numérique . . . 84
1.5.3 Niveau difficile . . . 84
1.5.3.1 Questions d'analyse . . . 84
1.5.4 Question de numérique . . . 86
1.5.5 Pour ma culture générale . . . 86
1.6 Comment contribuer et aider (LaTeX et programmation) . . . 86
1.6.1 Niveau facile . . . 86
1.6.2 Niveau moyen . . . 86
1.6.3 Niveau difficile . . . 86
1.7 Sage est là pour vous aider . . . 87
1.7.1 Lancez-vous dans Sage . . . 87
1.7.2 Exemples de ce que Sage peut faire pour vous . . . 87

2 Construction des ensembles de nombres (Vol 1) **89**
2.1 Quelques éléments sur les ensembles . . . 89
2.1.1 Ensemble ordonné . . . 89
2.1.2 Lemme de Zorn . . . 90
2.1.3 Complémentaire . . . 90
2.1.4 Relations d'équivalence . . . 91

2.2 Les naturels . . . 92
2.2.1 La construction . . . 92
2.2.2 Quelque autres résultats . . . 92
2.3 Les entiers . . . 93
2.4 Quelques structures algébriques . . . 93
2.5 Les rationnels . . . 95
2.5.1 Suites de Cauchy dans les rationnels . . . 95
2.5.2 Insuffisance des rationnels . . . 97
2.6 Les réels . . . 100
2.6.1 L'ensemble . . . 100
2.6.2 Relation d'ordre . . . 102
2.7 Les complexes . . . 105
2.7.1 Définitions . . . 106

3 Théorie des groupes (Vol 1) 107
3.1 Groupes . . . 107
3.2 Sous groupe normal . . . 108
3.2.1 Classes de conjugaison . . . 109
3.3 Groupe dérivé . . . 109
3.4 Théorèmes d'isomorphismes . . . 110
3.5 Le groupe et anneau des entiers . . . 112
3.5.1 Division euclidienne . . . 112
3.5.2 Sous-groupes de $(\mathbb{Z}, +)$ . . . 113
3.5.3 PGCD, PPCM et Bézout . . . 113
3.5.4 Calcul effectif du PGCD et de Bézout . . . 116
3.5.4.1 Algorithme d'Euclide pour le PGCD . . . 116
3.5.4.2 Algorithme étendu : calcul effectif des coefficients de Bézout . . . 117
3.5.5 Décomposition en facteurs premiers . . . 118
3.5.6 Écriture des fractions . . . 119
3.5.7 Équation diophantienne linéaire à deux inconnues . . . 120
3.5.8 Quotients . . . 121
3.6 Indice d'un sous-groupe et ordre des éléments . . . 122
3.7 Suite de composition . . . 125
3.8 Groupes résolubles . . . 128
3.9 Action de groupes . . . 129
3.10 Groupe des permutations . . . 134
3.11 Produit semi-direct de groupes . . . 134
3.12 Groupe de torsion . . . 136
3.13 Famille presque nulle . . . 136
3.14 Le groupe des racines de l'unité . . . 136
3.15 Fonction indicatrice d'Euler (première partie) . . . 138
3.15.1 Introduction par les racines de l'unité . . . 139
3.15.2 Générateurs . . . 140

4 Anneaux (Vol 1) 141
4.1 Généralités . . . 141
4.2 Binôme de Newton et morphisme de Frobenius . . . 143
4.3 Idéal dans un anneau . . . 144
4.3.1 Résultats supplémentaires sur l'anneau des entiers . . . 145
4.4 Caractéristique . . . 145
4.5 Modules . . . 146
4.6 Anneau intègre . . . 148
4.6.1 Caractéristique d'un anneau intègre . . . 149
4.6.2 Divisibilité et classes d'association . . . 150

4.6.3 PGCD et PPCM . . . 150
4.6.4 Anneaux intègres et corps . . . 151
4.6.5 Corps des fractions . . . 152
4.7 Anneau factoriel . . . 152
4.8 Anneau principal . . . 153
4.8.1 Bézout . . . 154
4.8.2 Anneau noetherien . . . 156
4.9 Anneau euclidien . . . 156
4.9.1 Équations diophantiennes . . . 158
4.9.2 Triplets pythagoriciens et équation de Fermat pour $n = 4$ . . . 159
4.9.3 Lignes et colonnes de matrices . . . 162
4.9.4 Algorithme des facteurs invariants . . . 163
4.10 Polynômes à coefficients dans un anneau commutatif . . . 165
4.10.1 Division euclidienne . . . 166
4.10.2 Polynôme primitif . . . 167
4.10.3 Racines des polynômes . . . 167
4.10.4 Quelques identités . . . 169

5 Corps (Vol 1) 171
5.1 Généralités . . . 171
5.1.1 Corps ordonnés . . . 171
5.1.2 Automorphismes de $\mathbb{R}$ et $\mathbb{C}$ . . . 171
5.1.3 Corps premier . . . 173
5.1.4 Petit théorème de Fermat . . . 173
5.2 Théorème des deux carrés . . . 174
5.2.1 Un peu de structure dans $\mathbb{Z}[i]$ . . . 174
5.2.2 Résultats chinois . . . 178
5.3 Polynômes à coefficients dans un corps . . . 180
5.3.1 Irréductibilité . . . 180
5.3.2 Idéaux . . . 181
5.3.3 Bézout . . . 183
5.3.4 Lemme et théorème de Gauss . . . 183
5.4 Extension de corps . . . 184
5.4.1 Polynôme minimal . . . 184
5.4.2 Racines de polynômes . . . 187
5.4.3 Corps de rupture . . . 189
5.4.4 Corps de décomposition . . . 191
5.4.5 Clôture algébrique . . . 192
5.4.6 Extensions séparables . . . 192
5.4.7 Idéal maximum . . . 196
5.5 Espaces de polynômes . . . 198
5.5.1 Polynômes symétriques, alternés ou semi-symétriques . . . 200
5.5.2 Polynôme symétrique élémentaire . . . 200
5.5.3 Relations coefficients racines . . . 202
5.6 Polynômes cyclotomiques . . . 203
5.6.1 Définitions et propriétés . . . 203
5.6.2 Nombres premiers . . . 207
5.6.3 Théorème de Wedderburn . . . 209
5.7 Corps finis . . . 211
5.7.1 Existence, unicité . . . 211
5.7.2 Symboles de Legendre et carrés . . . 213
5.7.3 Théorème de Chevalley-Warning . . . 219
5.7.4 Contenu d'un polynôme . . . 221
5.7.5 Théorème de l'élément primitif . . . 222

5.7.6 Construction de $\mathbb{F}_q$ . . . 226
5.7.6.1 La version du faignant . . . 226
5.7.6.2 La version plus élaborée . . . 227
5.7.7 Exemple : étude de $\mathbb{F}_{16}$ . . . 229
5.7.8 Polynômes irréductibles sur $\mathbb{F}_q$ . . . 231
5.8 Constructions à la règle et au compas . . . 233
5.8.1 Quelque constructions . . . 233
5.8.2 Nombres constructibles . . . 235
5.8.3 Polygones constructibles . . . 237
5.8.4 Matrices . . . 242
5.9 Minuscule morceau sur la théorie de Galois . . . 242
5.9.1 Théorème de d'Alembert . . . 243

6 Topologie générale (Vol 1) 245
6.1 Topologie en général . . . 245
6.1.1 Définitions basiques . . . 245
6.1.2 Topologie produit . . . 246
6.1.3 Séparabilité . . . 246
6.1.4 Topologie métrique . . . 247
6.1.5 Espace vectoriel topologique . . . 249
6.2 Espace vectoriel normé . . . 249
6.2.1 Introduction : norme et valeur absolue . . . 249
6.2.2 Norme . . . 250
6.2.3 Quelque exemples . . . 251
6.3 Suites de Cauchy, métrique et espaces complet . . . 252
6.3.1 Généralités . . . 252
6.3.2 Cas d'équivalence . . . 254
6.3.3 Espace topologique normé . . . 256
6.3.4 Espace métrique . . . 256
6.4 Topologie sur l'ensemble des réels . . . 257
6.5 Base de topologie . . . 259
6.6 Limite et continuité . . . 260
6.7 Compacité . . . 261
6.7.1 Quelque propriétés . . . 261
6.8 Topologie induite . . . 263
6.9 Connexité . . . 264
6.10 Un peu de topologie réelle . . . 266
6.10.1 Suites numériques . . . 266
6.10.2 Maximum, supremum et compagnie . . . 268
6.10.2.1 ...et quelque exemples . . . 268
6.10.3 Intervalles . . . 273
6.10.3.1 Connexité . . . 273
6.10.3.2 Compacité . . . 274
6.10.4 Image d'un compact . . . 276
6.10.5 Connexité par arcs . . . 278
6.10.6 Topologie de la droite réelle complétée . . . 278
6.10.7 Limite pointée ou épointée ? . . . 279
6.11 Espaces métriques . . . 279
6.11.1 Espaces métrisables . . . 279
6.11.2 Fonctions continues . . . 279
6.11.3 Caractérisations séquentielles . . . 280
6.11.4 Distance entre un point et un ensemble . . . 282
6.11.5 Compacité . . . 283
6.11.6 Ensembles enchaînés . . . 288

6.11.7 Produit fini d'espaces métriques . . . 289
6.12 Ensembles nulle part denses . . . 290
6.13 Encore de la topologie réelle . . . 291
6.13.1 Ouverts et fermés . . . 291
6.13.2 Intérieur, adhérence et frontière . . . 291
6.13.3 Point d'accumulation, point isolé . . . 293
6.13.4 Limite de suite . . . 293
6.14 Application réciproque . . . 294
6.15 Topologie des semi-normes . . . 295
6.15.1 Espace dual . . . 297
6.15.2 Espace $C^k(\mathbb{R}, E')$ . . . 297
6.16 Espaces de Baire . . . 298

7 Espaces vectoriels (Vol 1) 301
7.1 Parties libres, génératrices, bases et dimension . . . 301
7.2 Applications linéaires . . . 305
7.2.1 Définition . . . 305
7.2.2 Rang . . . 306
7.2.3 Matrices équivalentes et similaires . . . 307
7.3 Matrice d'une application linéaire . . . 309
7.3.1 Écriture dans une base . . . 309
7.3.2 Changement de base : vecteurs de base . . . 310
7.3.3 Changement de base : coordonnées . . . 311
7.3.4 Changement de base : matrice d'une application linéaire . . . 311
7.3.5 Changement de base : matrice d'une forme bilinéaire . . . 312
7.4 Dualité . . . 312
7.4.1 Orthogonal . . . 312
7.4.2 Transposée . . . 313
7.4.3 Transposée : sans le dual . . . 315
7.4.4 Polynômes de Lagrange . . . 316
7.4.5 Dual de $\mathbb{M}(n, \mathbb{K})$ . . . 316

8 Classification de certains groupes (Vol 1) 319
8.1 Théorèmes de Sylow . . . 319
8.2 Un peu de classification de groupes . . . 323
8.2.1 Groupe monogène . . . 323
8.2.2 Automorphismes du groupe $\mathbb{Z}/n\mathbb{Z}$ . . . 324
8.2.3 Groupes abéliens finis . . . 326
8.2.4 Groupes d'ordre pq . . . 328
8.2.5 Fonction indicatrice d'Euler . . . 331
8.3 Le groupe symétrique . . . 332
8.3.1 Décomposition en cycles . . . 332
8.3.2 Le groupe alterné . . . 334
8.3.3 Sous-groupes normaux . . . 338
8.3.4 Indice . . . 340

9 Suite des espaces vectoriels (Vol 1) 343
9.1 Déterminants . . . 343
9.1.1 Formes multilinéaires alternées . . . 343
9.1.2 Déterminant d'une famille de vecteurs . . . 345
9.1.3 Déterminant d'un endomorphisme . . . 347
9.1.4 Déterminant d'une matrice . . . 349
9.1.5 Déterminant de Vandermonde . . . 349
9.1.6 Déterminant de Gram . . . 352

9.1.7 Déterminant de Cauchy . . . 352
9.1.8 Matrice de Sylvester . . . 352
9.1.9 Théorème de Kronecker . . . 355
9.2 Produit scalaire . . . 357
9.2.1 Projection et angles . . . 359
9.2.2 Procédé de Gram-Schmidt . . . 361
9.3 Hermitien, orthogonal, adjoint . . . 362
9.3.1 Opérateur orthogonal, matrice orthogonale . . . 363
9.4 Directions conservées . . . 364
9.4.1 Généralités . . . 364
9.4.2 Comment trouver la matrice d'une symétrie donnée ? . . . 365
9.4.2.1 Symétrie par rapport à un plan . . . 365
9.4.2.2 Symétrie par rapport à une droite . . . 366
9.4.2.3 En résumé . . . 367
9.5 Produit vectoriel . . . 367
9.5.1 Produit vectoriel . . . 368
9.5.2 Produit mixte . . . 369
9.6 Topologie . . . 371
9.6.1 Boules et sphères . . . 371
9.6.2 Ouverts, fermés, intérieur et adhérence . . . 372
9.6.3 Point isolé, point d'accumulation . . . 379
9.7 Valeur propre et vecteur propre . . . 380
9.8 Polynôme d'endomorphismes . . . 380
9.8.1 Polynômes d'endomorphismes . . . 381
9.8.2 Calcul effectif de l'exponentielle d'une matrice . . . 383
9.8.3 Polynôme minimal et minimal ponctuel . . . 383
9.8.4 Polynôme caractéristique . . . 389
9.9 Diagonalisation et trigonalisation . . . 393
9.9.1 Matrices semblables . . . 393
9.9.2 Endomorphismes nilpotents . . . 394
9.9.3 Endomorphismes diagonalisables . . . 397
9.9.4 Diagonalisation : cas complexe, pas toujours . . . 400
9.9.5 Trigonalisation : généralités . . . 400
9.9.6 Trigonalisation : cas complexe . . . 401
9.9.7 Diagonalisation : cas complexe, ce qu'on a . . . 403
9.9.8 Diagonalisation : cas réel . . . 405
9.9.9 Pseudo-réduction simultanée . . . 408
9.10 Sommes de familles infinies . . . 409
9.10.1 Convergence commutative . . . 409
9.11 Fonctions . . . 413
9.12 Sous espaces caractéristiques . . . 414
9.12.1 Théorèmes de décomposition . . . 415
9.12.2 Diverses conséquences . . . 417
9.12.3 Diagonalisabilité d'exponentielle . . . 418
9.12.4 Valeurs singulières . . . 420
9.13 Extension du corps de base . . . 420
9.13.1 Extension des applications linéaires . . . 420
9.13.2 Projections . . . 422
9.13.3 Rang, polynôme minimal, polynôme caractéristique . . . 425
9.14 Frobenius et Jordan . . . 426
9.14.1 Matrice compagnon . . . 426
9.14.2 Réduction de Frobenius . . . 427
9.14.3 Forme normale de Jordan . . . 430

9.15 Commutant et endomorphismes cycliques . . . 431
9.15.1 Endomorphisme cyclique . . . 431
9.15.2 Commutant : cas diagonalisable . . . 432
9.15.3 Commutant : cas général . . . 434
9.16 Théorème de Burnside . . . 437
9.16.1 Théorème de Lie-Kolchin . . . 438
9.17 Formes bilinéaires et quadratiques . . . 441
9.17.1 Généralités . . . 441
9.17.2 Topologie . . . 441
9.17.3 Matrice associée . . . 441
9.17.4 Dégénérescence . . . 443
9.17.5 Inégalité de Minkowski . . . 444
9.17.6 Ellipsoïde . . . 445
9.18 Théorème spectral auto-adjoint . . . 446
9.19 Mini introduction au produit tensoriel . . . 449
9.19.1 Définitions . . . 449
9.19.2 Application d'opérateurs . . . 449
9.20 Méthode de Gauss pour résoudre des systèmes d'équations linéaires . . . 450

10 Espaces vectoriels normés (Vol 1) 451
10.1 Équivalence des normes . . . 451
10.1.1 En dimension finie . . . 451
10.1.2 Contre-exemple en dimension infinie . . . 452
10.2 Norme opérateur . . . 453
10.2.1 Norme d'algèbre . . . 455
10.2.2 Application linéaire continue et bornée . . . 459
10.2.3 Normes de matrices et d'applications linéaires . . . 461
10.3 Isométriques du cube . . . 462
10.4 Chiffrement RSA . . . 464
10.4.1 Mise en place par Bob . . . 464
10.4.2 Chiffrement . . . 464
10.4.3 Déchiffrement . . . 465
10.4.4 Une imprudence à ne pas commettre . . . 465
10.4.5 Problèmes calculatoires . . . 466
10.4.6 La solidité de RSA . . . 466
10.4.7 Note non mathématique pour doucher l'enthousiasme . . . 466
10.5 Suites . . . 466
10.5.1 Limites, convergence . . . 466
10.5.2 Critère de Cauchy . . . 467
10.5.3 Approximation . . . 468
10.6 Séries . . . 469
10.7 Série réelle . . . 472
10.7.1 Critères de convergence absolue . . . 472
10.7.2 Critères de convergence simple . . . 474
10.7.2.1 Critère d'Abel . . . 474
10.7.3 Quelques séries usuelles . . . 475
10.7.4 Moyenne de Cesaro . . . 476
10.7.5 Écriture décimale d'un nombre . . . 476
10.8 Produit fini d'espaces vectoriels normés . . . 480
10.8.1 Norme . . . 480
10.8.2 Suites . . . 482
10.8.3 Continuité du produit de matrices . . . 483
10.9 Applications multilinéaires . . . 484
10.10Exponentielle de matrice . . . 486

10.11 Mini introduction aux nombres p-adiques . . . 487
10.11.1 La flèche d'Achille . . . 487
10.11.2 La tortue et Achille . . . 487
10.11.3 Dans les nombres p-adiques, c'est vrai . . . 488

11 Espaces affines (Vol 2) 491
11.1 Repères affines . . . 491
11.2 Classification affine des conique . . . 492
11.3 Applications affines . . . 494
11.4 Isomorphismes . . . 495
11.5 Sous espaces affines . . . 496
11.6 Barycentre . . . 497
11.6.1 Sous-espaces affines . . . 498
11.6.2 Enveloppe convexe . . . 500
11.6.3 Applications affines et barycentre . . . 503
11.7 Repères, coordonnées cartésiennes et barycentriques . . . 504
11.7.1 Équation de droite . . . 506
11.7.2 Associativité, coordonnées barycentriques dans un triangle . . . 506
11.8 Structure du groupe des applications affines . . . 507
11.9 Isométries . . . 509

12 Analyse réelle (Vol 2) 511
12.1 Intervalles . . . 511
12.2 Application réciproque . . . 512
12.2.1 Définitions . . . 512
12.2.2 Graphe de la fonction réciproque . . . 513
12.3 Fonctions continues . . . 514
12.3.1 La fonction la moins continue du monde . . . 516
12.3.2 Approche topologique . . . 516
12.3.3 Continuité de la racine carré, invitation à la topologie induite . . . 518
12.3.4 Limites en des nombres . . . 520
12.3.5 Limites quand tout va bien . . . 520
12.3.6 Discussion avec mon ordinateur . . . 525
12.3.7 Limites et prolongement . . . 525
12.3.8 Théorèmes des valeurs intermédiaires . . . 526
12.3.9 Théorème de la bijection . . . 526
12.3.10 Prolongement par continuité . . . 529
12.4 Espace des fonctions continues . . . 530
12.5 Uniforme continuité . . . 535
12.6 Compacité . . . 537
12.7 Limite et continuité . . . 539
12.7.1 Définition . . . 539
12.7.2 Propriétés de base . . . 540
12.7.3 Limites de fonctions . . . 540
12.7.4 Règles simples de calcul . . . 546
12.8 Limites à plusieurs variables . . . 546
12.8.1 Règle de l'étau . . . 547
12.8.2 Méthode des chemins . . . 548
12.8.3 Méthode des coordonnées polaires . . . 550
12.8.4 Méthode du développement asymptotique . . . 554
12.9 Dérivée : exemples introductifs . . . 554
12.9.1 La vitesse . . . 554
12.9.2 La tangente à une courbe . . . 555
12.9.3 L'aire en dessous d'une courbe . . . 556

12.10 Continuité et dérivabilité . . . 557
12.11 Dérivation de fonctions d'une variable réelle . . . 558
12.11.1 Exemples . . . 558
12.11.1.1 La fonction $f(x) = x$ . . . 558
12.11.1.2 La fonction $f(x) = x^2$ . . . 558
12.11.1.3 La fonction $f(x) = \sqrt{x}$ . . . 559
12.11.2 Calcul de la dérivée . . . 559
12.11.3 Interprétation géométrique : tangente . . . 560
12.11.4 Interprétation géométrique : approximation affine . . . 560
12.12 Opérations sur les dérivées . . . 561
12.12.1 Développement limité au premier ordre . . . 562
12.13 Dérivation à une variable . . . 563
12.13.1 Définition . . . 563
12.13.2 Quelques formules à connaître . . . 563
12.14 Dérivation et croissance . . . 563
12.14.1 Théorèmes de Rolle et des accroissements finis . . . 565
12.14.2 Règle de l'Hospital . . . 567
12.14.3 À propos de primitives . . . 569
12.15 Fonctions de plusieurs variables . . . 570
12.15.1 Graphes de fonctions à plusieurs variables . . . 573
12.15.2 Courbes de niveau . . . 574
12.16 Dérivée directionnelle . . . 578
12.16.1 Dérivée partielle et directionnelles . . . 578
12.16.1.1 Quelques propriétés et notations . . . 579
12.16.2 Gradient : direction de plus grande pente . . . 583
12.16.3 Gradient : orthogonal au plan tangent . . . 583
12.17 Dérivée directionnelle de fonctions composées . . . 584
12.18 Formes différentielles . . . 586
12.18.1 Décomposition dans la base duale . . . 586
12.18.2 L'isomorphisme musical . . . 587
12.18.3 Différentielles de fonctions composées . . . 587
12.18.4 Exemple de composée : les coordonnées polaires . . . 588
12.19 Différentielle . . . 589
12.19.1 Exemples introductifs . . . 590
12.19.2 Définition de la différentielle . . . 591
12.19.3 Base duale . . . 592
12.19.4 Pas la différentielle extérieure . . . 592
12.19.5 Continuité, dérivabilité et différentiabilité . . . 593
12.19.6 Calcul de valeurs approchées . . . 595
12.19.7 Prouver qu'un fonction n'est pas différentiable . . . 598
12.19.7.1 Continuité . . . 598
12.19.7.2 Linéarité . . . 598
12.19.7.3 Cohérence des dérivées partielles et directionnelle . . . 601
12.19.7.4 Un candidat dans la définition (marche toujours) . . . 602
12.19.8 Règles de calcul . . . 602
12.19.9 Linéarité . . . 603
12.19.10 Produit . . . 603
12.19.11 Différentielle de fonction composée . . . 605
12.19.12 Différentielle et dérivées partielles . . . 608
12.19.13 Plan tangent . . . 609
12.19.14 Calcul de différentielles . . . 609
12.19.15 Notes idéologiques quant au concept de plan tangent . . . 610
12.19.16 Gradient et recherche du plan tangent . . . 610

12.20 Calcul différentiel dans un espace vectoriel normé . . . 612
12.20.1 Différentielle . . . 612
12.20.2 (non ?) Différentiabilité des applications linéaires . . . 613
12.20.3 Dérivation en chaine et formule de Leibnitz . . . 613
12.20.4 Différentielle partielle . . . 617
12.20.5 Formule des accroissements finis . . . 618
12.20.6 L'inverse, sa différentielle . . . 621
12.20.7 Projection orthogonale . . . 623
12.21 Jacobienne . . . 624
12.21.1 Rappels et définitions . . . 624
12.22 Fonctions de classe $\mathcal{C}^1$ . . . 625
12.23 Théorèmes des accroissements finis . . . 626
12.24 Fonctions Lipschitziennes . . . 627
12.25 Différentielles d'ordre supérieur . . . 629
12.25.1 Identification des espaces d'applications multilinéaires . . . 629
12.25.2 Fonctions différentiables plusieurs fois . . . 629
12.25.3 Différentielle seconde, fonction de classe C^2 . . . 630
12.25.4 Ordre supérieur . . . 633
12.26 Développement asymptotique, théorème de Taylor . . . 634
12.26.1 Fonctions «petit o» . . . 635
12.26.2 Autres formulations . . . 636
12.26.3 Formule et reste . . . 637
12.26.4 Reste intégral . . . 637
12.27 Développement limité autour de zéro . . . 637
12.27.1 Généralités . . . 638
12.27.2 Formule de Taylor-Young . . . 639
12.27.3 Règles de calcul . . . 641
12.27.3.1 Linéarité des développements limités . . . 641
12.27.3.2 Développement limité d'un quotient . . . 642
12.27.3.3 Développement limité d'une fonction composée . . . 643
12.28 Développement au voisinage de $x_0 \neq 0$ . . . 644
12.29 Application au calcul de limites . . . 645
12.30 Développement au voisinage de l'infini . . . 646
12.31 Convergence uniforme . . . 646
12.31.1 Critère de Cauchy uniforme . . . 646
12.31.2 Complétude avec la norme uniforme . . . 647
12.31.3 Permuter avec les dérivées partielles . . . 649
12.32 Fonctions réelles de deux variables réelles . . . 650
12.32.1 Limites de fonctions à deux variables . . . 650
12.32.2 Dérivées partielles . . . 651
12.32.3 Différentielle et accroissement . . . 652
12.33 Les fonctions à valeurs vectorielles . . . 652
12.34 Fonctions vectorielles de plusieurs variables . . . 653
12.35 Champs de vecteurs . . . 653
12.35.1 Matrice jacobienne . . . 654
12.36 Divergence, rotationnel et l'opérateur nabla . . . 654
12.37 Interprétation de la divergence . . . 657
12.38 Quelques formules de Leibnitz . . . 659
12.39 Action de groupe et connexité . . . 659
12.40 Espaces de matrices . . . 661
12.40.1 Dilatations et transvections . . . 662
12.40.2 Connexité de certains groupes . . . 667
12.40.3 Densité . . . 669

12.40.4 Racine carré d'une matrice hermitienne positive . . . 671
12.40.5 Racine carré d'une matrice symétrique positive . . . 672
12.40.6 Décomposition polaires : cas réel . . . 673
12.40.7 Enveloppe convexe . . . 675
12.40.8 Décomposition de Bruhat . . . 678
12.41 Sous-groupes du groupe linéaire . . . 680

13 Tribus, théorie de la mesure (Vol 2) 685
13.1 Tribus . . . 685
13.1.1 Généralités . . . 685
13.1.1.1 Tribu induite . . . 686
13.1.1.2 Tribu borélienne . . . 687
13.1.1.3 Les boréliens de $\mathbb{R}$ . . . 687
13.1.2 Tribu de Baire . . . 687
13.2 Théorie de la mesure . . . 689
13.2.1 Mesure et algèbre de parties . . . 690
13.2.2 Mesure sur un espace mesurable . . . 692
13.2.3 Mesure extérieure . . . 698
13.2.4 Espace mesuré complet . . . 700
13.2.5 Prolongement . . . 708
13.3 Applications mesurables . . . 710
13.3.1 Propriétés . . . 710
13.3.2 D'une tribu à l'autre . . . 710
13.3.3 Mesure image . . . 713
13.3.4 Régularité d'une mesure . . . 714
13.3.5 Théorème de récurrence . . . 718
13.4 Mesurabilité des fonctions à valeurs réelles . . . 719
13.4.1 Quelques mots à propos de $\overline{\mathbb{R}}$ . . . 719
13.4.2 Limite supérieure et inférieure . . . 719
13.4.3 Fonctions à valeurs réelles sur un espace mesurable . . . 720
13.4.4 Fonction étagée . . . 726
13.4.5 Fonctions réelle à variables réelles . . . 729
13.5 Intégrale par rapport à une mesure . . . 729
13.5.1 Quelque propriétés . . . 731
13.5.2 Permuter limite et intégrale . . . 734
13.5.2.1 Convergence uniforme . . . 734
13.5.2.2 Convergence monotone . . . 735
13.5.2.3 Convergence dominée de Lebesgue . . . 738
13.5.3 Produit d'une mesure par une fonction (mesure à densité) . . . 739
13.5.4 Mesure et topologie . . . 740
13.6 Propriétés . . . 741
13.7 Mesure à densité . . . 742
13.7.1 Théorème de Radon-Nikodym . . . 742
13.7.2 Mesure complexe . . . 744
13.7.3 Théorème d'approximation . . . 745
13.8 Tribu produit, mesure produit . . . 745
13.8.1 Produit d'espaces mesurables . . . 745
13.8.2 Le cas des boréliens . . . 746
13.8.3 Produit de mesures . . . 747
13.9 Mesure de Lebesgue sur $\mathbb{R}$ . . . 752
13.9.1 Mesure et tribu de Lebesgue . . . 756
13.9.2 Propriétés de la mesure de Lebesgue . . . 757
13.9.3 Ensemble de Cantor . . . 762
13.9.4 Ensemble de Vitali (non mesurable) . . . 764

13.10 Tribu et mesure de Lebesgue sur $\mathbb{R}^d$ . . . 765
13.10.1 Ensembles négligeables . . . 766
13.10.2 Parties et fonctions mesurables . . . 767
13.10.3 Propriétés d'unicité . . . 767
13.10.4 Régularité . . . 769
13.11 Propriétés de l'intégrale de Lebesgue . . . 769
13.11.1 Théorème de la moyenne . . . 769
13.11.2 Primitives et intégrales . . . 770
13.11.3 Exemples et applications . . . 772
13.11.4 Intégrales impropres . . . 773
13.12 Changement de variables dans une intégrale multiple . . . 775
13.12.1 Des lemmes . . . 775
13.12.2 Déterminant et mesure de Lebesgue . . . 776
13.12.3 Le théorème et sa démonstration . . . 777
13.12.4 Exemples . . . 782
13.13 Théorème de Fubini-Tonelli et de Fubini . . . 783

14 Intégration (Vol 2) 793
14.1 Intégrale sur une variété . . . 793
14.1.1 Mesure sur une carte . . . 793
14.1.1.1 Exemple : la mesure de la sphère . . . 794
14.1.2 Intégrale sur une carte . . . 794
14.1.3 Exemples . . . 795
14.1.4 Orientation . . . 795
14.1.5 Formes différentielles . . . 797
14.1.6 Intégrale d'une fonction sur une sous-variété . . . 798
14.2 Intégrales curvilignes . . . 799
14.2.1 Chemins de classe C^1 . . . 799
14.2.2 Intégrer une fonction . . . 799
14.2.3 Intégrer un champ de vecteurs . . . 801
14.2.4 Intégrer une forme différentielle sur un chemin . . . 801
14.2.5 Intégration d'une forme différentielle sur un chemin . . . 802
14.2.6 Interprétation physique : travail . . . 803
14.2.7 Intégrer un champs de vecteurs sur un bord en $2D$ . . . 804
14.2.8 Intégrer une forme différentielle sur un bord en $2D$ . . . 804
14.2.9 Intégrer une forme différentielle sur un bord en $3D$ . . . 804
14.2.10 Intégrer d'un champ de vecteurs sur un bord en $3D$ . . . 805
14.2.11 Dérivées croisées et forme différentielle exacte . . . 805
14.3 Surfaces paramétrées . . . 806
14.3.1 Graphe d'une fonction . . . 807
14.3.2 Intégrale sur une partie de $\mathbb{R}^m$ . . . 808
14.4 Intégrales de surface . . . 809
14.4.1 Intégrale d'un champ de vecteurs . . . 809
14.5 Intégrales de surface . . . 810
14.5.1 Aire d'une surface paramétrée . . . 810
14.5.2 Intégrale d'une fonction sur une surface . . . 811
14.5.3 Aire d'une surface de révolution . . . 812
14.5.4 Intégrale d'une 2-forme . . . 813
14.6 Flux d'un champ de vecteurs à travers une surface . . . 814
14.7 Divergence, Green, Stokes . . . 817
14.7.1 Théorème de la divergence . . . 817
14.7.2 Formule de Green . . . 818
14.7.3 Formule de Stokes . . . 820
14.7.3.1 Quelle est la bonne orientation ? . . . 820

14.8 Résumé des intégrales vues . . . 821
14.8.1 L'intégrale d'une fonction sur les réels . . . 821
14.8.2 Intégrale d'une fonction sur un chemin . . . 821
14.8.3 Intégrale d'une fonction sur une surface . . . 821
14.8.4 Intégrale d'une fonction sur un volume . . . 822
14.8.5 Conclusion pour les fonctions . . . 823
14.8.6 Circulation d'un champ de vecteurs . . . 823
14.8.7 Flux d'un champ de vecteurs . . . 823
14.8.8 Conclusion pour les champs de vecteurs . . . 823
14.8.9 Attention pour les surfaces fermées ! . . . 824
14.9 Intégrales convergeant uniformément . . . 825
14.9.1 Définition et propriété . . . 825
14.9.2 Critères de convergence uniforme . . . 826
14.10 Fonctions définies par une intégrale . . . 827
14.10.1 Continuité sous l'intégrale . . . 827
14.10.2 Le coup du compact . . . 829
14.10.3 Dérivabilité sous l'intégrale . . . 829
14.10.4 Absolue continuité . . . 830
14.10.5 Différentiabilité sous l'intégrale . . . 832
14.11 Formes différentielles exactes et fermées . . . 835
14.12 Théorème d'Abel angulaire . . . 837
14.12.1 Passage à la limite sous le signe intégral . . . 841
14.12.2 Intégrale en dimension un . . . 841
14.12.3 Intégrales convergentes . . . 842
14.12.4 La méthode de Rothstein-Trager . . . 842
14.13 Rappel sur les intégrales usuelles . . . 847
14.14 Intégrales le long de chemins . . . 848
14.14.1 Circulation d'un champ de vecteur . . . 848
14.15 Circulation d'un champ conservatif . . . 850
14.16 Intégration de fonction à deux variables . . . 851
14.16.1 Intégration sur un domaine rectangulaire . . . 851
14.16.2 Intégration sur un domaine non rectangulaire . . . 852
14.16.3 Changement de variables . . . 855
14.16.3.1 Le cas des coordonnées polaires . . . 855
14.16.3.2 Les coordonnées cylindriques . . . 856
14.16.3.3 Coordonnées sphériques . . . 857
14.16.3.4 Un autre système utile . . . 857
14.17 Les intégrales triples . . . 858
14.17.1 Volume . . . 860
14.18 Un petit peu plus formel . . . 861
14.18.1 Intégration sur un domaine non rectangulaire . . . 861
14.18.2 Changement de variables . . . 862
14.18.2.1 Coordonnées polaires . . . 863
14.18.2.2 Coordonnées sphériques . . . 863
14.19 Aire et primitive . . . 863
14.19.1 Longueur d'arc de courbe . . . 864
14.19.2 Aire de révolution . . . 864
14.20 L'aire en dessous d'une courbe . . . 864
14.21 Propriétés des intégrales . . . 866
14.22 Techniques d'intégration . . . 867
14.22.1 Intégration par parties . . . 868
14.22.2 Changement de variables – pour trouver des primitives . . . 869
14.22.3 Changement de variables – pour calculer des intégrales . . . 870

14.22.4 Intégrations des fractions rationnelles réduites . . . 872
14.22.5 Quelques formules à connaître . . . 873
14.23 Trucs et astuces de calcul d'intégrales . . . 873
14.23.1 Quelques intégrales «usuelles» . . . 874
14.23.2 Reformer un carré au dénominateur . . . 876
14.23.3 Décomposer en fractions simples . . . 877
14.24 Constructions plus naïves de l'intégrale dans le cas réel . . . 877
14.24.1 Mesure de Lebesgue, version rapide . . . 877
14.24.2 Pavés et subdivisions . . . 878
14.24.3 Intégrale d'une fonction en escalier . . . 881
14.24.4 Intégrales partielles . . . 881
14.24.5 Réduction d'une intégrale multiple . . . 882
14.24.6 Propriétés de l'intégrale . . . 883
14.24.7 Intégrales multiples, cas général . . . 884
14.24.8 Réduction d'une intégrale multiple . . . 885
14.24.9 Intégrales sur des parties de $\mathbb{R}^2$ . . . 886
14.24.10 Intégrales sur des parties de $\mathbb{R}^3$ . . . 889
14.24.11 Fonctions et ensembles non bornés . . . 892

15 Suites et séries de fonctions (Vol 2) 893
15.1 Différentielle et dérivée complexe . . . 893
15.2 Série de fonctions . . . 896
15.2.1 Intégration de séries de fonctions . . . 899
15.2.2 Différentiabilité . . . 900
15.3 Séries entières . . . 901
15.3.1 Disque de convergence . . . 901
15.3.2 Propriétés de la somme . . . 904
15.3.3 Dérivation . . . 908
15.3.4 Intégration . . . 910
15.4 Dénombrement des solutions d'une équation diophantienne . . . 911
15.5 Exponentielle et logarithme . . . 912
15.6 Vitesses de x^α, de l'exponentielle et du logarithme . . . 919
15.7 Nombres de Bell . . . 922
15.8 Trigonométrie hyperbolique . . . 924
15.9 Séries entières de matrices . . . 925
15.9.1 Différentiabilité . . . 925
15.10 Étude d'asymptote . . . 927
15.11 Développement en série . . . 928
15.11.1 Série génératrice d'une suite . . . 928
15.11.2 Développement en série et Taylor . . . 929
15.11.3 Resommer une série . . . 931
15.11.3.1 Les sommes du type $\sum_n P(n)x^n$ . . . 931
15.11.3.2 Les sommes du type $\sum_n x^n/P(n)$ . . . 932
15.11.3.3 Sage, primitives et logarithme complexe . . . 934
15.11.3.4 Nombres de Bell . . . 936
15.12 Séries entières de matrices . . . 936
15.12.1 Rayon de convergence . . . 936
15.12.2 Convergence et rayon spectral . . . 936
15.12.3 Exponentielle et logarithme de matrice . . . 938
15.13 Lemme de Borel . . . 941
15.13.1 Fonctions plateaux . . . 941
15.13.2 Le lemme de Borel . . . 942

16 Trigonométrie, isométries (Vol 3) 945
16.1 Isométries de l'espace euclidien . . . 945
16.1.1 Structure du groupe Isom($\mathbb{R}^n$) . . . 945
16.1.2 Classification des isométries de $\mathbb{R}$ . . . 949
16.1.3 Segment, plan médiateur et équidistance . . . 950
16.1.4 Isométries du tétraèdre régulier . . . 950
16.2 Trigonométrie . . . 951
16.2.1 Définitions, périodicité et quelque valeurs remarquables . . . 951
16.2.2 Exemples . . . 957
16.2.3 Très modeste approximation de π . . . 959
16.2.4 Cercle trigonométrique . . . 959
16.2.5 Les fonctions tangente et arc tangente . . . 960
16.2.6 La fonction arc sinus . . . 962
16.2.7 La fonction arc cosinus . . . 964
16.2.8 Une meilleure approximation de π . . . 965
16.2.9 Forme trigonométrique des nombres complexes . . . 966
16.2.10 Angle entre deux vecteurs . . . 967
16.2.11 Aire du parallélogramme . . . 968
16.3 Classification des isométries dans $\mathbb{R}^2$ . . . 969
16.3.1 Réflexions . . . 969
16.3.2 Rotations . . . 971
16.3.3 Rotations vectorielles . . . 975
16.3.4 Angle . . . 976
16.3.5 Classification . . . 978
16.4 Isométries dans $\mathbb{R}^n$ . . . 981
16.4.1 Groupes finis d'isomorphismes . . . 985
16.4.2 Théorème de Sylvester . . . 987
16.4.3 Groupe diédral . . . 988
16.4.3.1 Définition et générateurs : vue géométrique . . . 988
16.4.3.2 Générateurs : vue abstraite . . . 990
16.4.3.3 Classes de conjugaison . . . 993
16.4.3.4 Le compte pour n pair . . . 994
16.4.3.5 Le compte pour n impair . . . 994
16.4.4 Applications : du dénombrement . . . 994
16.4.4.1 Le jeu de la roulette . . . 994
16.4.4.2 L'affaire du collier . . . 995
16.5 Angles et rotations dans $\mathbb{R}^2$ . . . 995
16.5.1 Angles orientés, rotations . . . 995
16.5.2 Angles et nombres complexes . . . 1000
16.5.3 Sous-groupes finis de SO(2) . . . 1001
16.6 Un peu de structure de O(n) . . . 1001
16.6.1 Valeurs propres dans O(n) . . . 1001
16.6.2 Sous-groupes finis de SO(3) . . . 1004
16.7 Coordonnées polaires . . . 1013
16.7.1 Les coordonnées polaires . . . 1013
16.7.1.1 Transformation inverse : théorie . . . 1013
16.7.1.2 Transformation inverse : pratique . . . 1014
16.7.2 Coordonnées polaires comme changement de variables pour intégrales . . . 1015
16.7.3 Changement de variables . . . 1017
16.7.3.1 Coordonnées polaires . . . 1017
16.7.3.2 Coordonnées sphériques . . . 1018
16.7.4 Coordonnées polaires . . . 1019
16.7.5 Coordonnées cylindriques . . . 1019

16.7.6 Coordonnées sphériques . . . 1020
16.7.7 Coordonnées cylindriques et sphériques . . . 1020
16.7.8 Récapitulatif des changements de variables . . . 1021
16.7.8.1 Coordonnées polaires . . . 1021
16.7.8.2 Coordonnées cylindriques . . . 1021
16.7.8.3 Coordonnées sphériques . . . 1021
16.7.9 une limite . . . 1022

17 Représentations et caractères (Vol 3) 1023
17.1 Représentations et caractères . . . 1023
17.1.1 Crochet de dualité et transformée de Fourier . . . 1025
17.1.2 Groupes non abéliens . . . 1026
17.1.3 Représentations linéaires des groupes finis . . . 1026
17.1.4 Module . . . 1027
17.1.5 Structure hermitienne . . . 1029
17.1.6 Caractères . . . 1029
17.2 Équivalence de représentations et caractères . . . 1030
17.2.1 Représentation régulière . . . 1032
17.2.2 Caractères et représentations : suite et fin . . . 1033
17.3 Représentation produit tensoriel . . . 1036
17.4 Exemple sur le groupe symétrique . . . 1036
17.5 Table des caractères du groupe symétrique S_4 . . . 1037
17.5.1 Calculs à partir de rien ou presque . . . 1037
17.5.2 Représentation de S_4 via les isométries du tétraèdre . . . 1039
17.5.3 À propos de la représentation ρ_u . . . 1040
17.6 Table de caractères du groupe diédral . . . 1041
17.6.1 Représentations de dimension un . . . 1041
17.6.2 Représentations de dimension deux . . . 1042
17.6.3 Le compte pour n pair . . . 1044
17.6.4 Le compte pour n impair . . . 1044

18 Suite de l'analyse (Vol 3) 1045
18.1 Densité des polynômes . . . 1045
18.1.1 Théorème de Stone-Weierstrass . . . 1045
18.2 Primitive de fonction continue . . . 1048
18.2.1 Théorème taubérien de Hardi-Littlewood . . . 1049
18.2.2 Théorème de Müntz . . . 1053
18.3 Théorèmes de point fixe . . . 1056
18.3.1 Points fixes attractifs et répulsifs . . . 1056
18.3.2 Picard . . . 1056
18.3.3 Brouwer . . . 1059
18.3.4 Théorème de Schauder . . . 1061
18.3.5 Théorème de Markov-Kakutani et mesure de Haar . . . 1062
18.4 Théorèmes de point fixes et équations différentielles . . . 1064
18.4.1 Théorème de Cauchy-Lipschitz . . . 1064
18.4.2 Théorème de Cauchy-Arzella . . . 1071
18.5 Théorèmes d'inversion locale et de la fonction implicite . . . 1072
18.5.1 Mise en situation . . . 1072
18.5.2 Théorème d'inversion locale . . . 1073
18.5.3 Théorème de la fonction implicite . . . 1076
18.5.4 Exemple . . . 1078
18.6 Décomposition polaire (régularité) . . . 1079
18.7 Théorème de Von Neumann . . . 1081
18.8 Recherche d'extrema . . . 1084

18.8.1 Extrema à une variable . . . 1084
18.8.2 Extrema libre . . . 1085
18.8.3 Extrema et Hessienne . . . 1086
18.8.4 Un peu de recettes de cuisine . . . 1088
18.8.5 Extrema liés . . . 1088
18.9 Fonctions convexes . . . 1090
18.9.1 Inégalité des pentes . . . 1091
18.9.2 Convexité et régularité . . . 1092
18.9.3 Dérivées d'une fonction convexe . . . 1092
18.9.4 Graphe d'une fonction convexe . . . 1094
18.9.5 Convexité et hessienne . . . 1098
18.9.6 Quelque inégalités . . . 1102
18.9.6.1 Inégalité de Jensen . . . 1102
18.9.6.2 Inégalité arithmético-géométrique . . . 1102
18.9.6.3 Inégalité de Kantorovitch . . . 1103
18.10 Algorithme du gradient à pas optimal . . . 1104
18.11 Ellipsoïde de John-Loewer . . . 1108
18.12 Formes quadratiques, signature, et lemme de Morse . . . 1111
18.12.1 Lemme de Morse . . . 1115
18.13 Variétés . . . 1117
18.13.1 Introduction . . . 1117
18.13.2 Définition et propriétés . . . 1117
18.13.3 Espace tangent . . . 1118
18.14 Prolongement de fonctions et complétion d'espaces métriques . . . 1119
18.15 Un petit extra . . . 1125

19 Arc paramétré (Vol 3) 1127
19.1 Définitions . . . 1127
19.2 Longueur d'arc . . . 1127
19.3 Abscisse curviligne . . . 1130
19.3.1 Formule intégrale de la longueur . . . 1131
19.4 Suite du chapitre . . . 1135
19.5 Courbes paramétrés . . . 1136
19.5.1 Définitions et exemples . . . 1136
19.6 Élément de longueur . . . 1138
19.6.1 Élément de longueur : cartésiennes . . . 1138
19.6.2 Élément de longueur : polaires (1) . . . 1138
19.6.3 Élément de longueur : polaires (2) . . . 1138
19.6.4 Approximation de la longueur par des cordes . . . 1140
19.7 Arc géométrique . . . 1142
19.7.1 Abscisse curviligne et paramétrisation normale . . . 1143
19.7.2 Tangente à une courbe paramétrée . . . 1148
19.8 Repère de Frenet . . . 1150
19.8.1 Torsion . . . 1152
19.9 Hors des coordonnées normales . . . 1153
19.10 Tracer des courbes paramétriques dans $\mathbb{R}^2$ . . . 1156
19.11 Courbes planes . . . 1157
19.11.1 Angle . . . 1157
19.11.2 Courbure signée . . . 1158
19.11.3 Degré, indice et homotopie . . . 1161
19.12 Courbes fermées planes . . . 1167
19.12.1 Cercle circonscrit . . . 1167
19.12.2 Description locale . . . 1169
19.12.3 Enveloppe convexe . . . 1170

19.12.4 Courbure et convexité . . . 1173
19.12.5 Théorème des quatre sommets . . . 1174
19.12.6 Le théorème de Jordan . . . 1176

20 Géométrie hyperbolique (Vol 3) 1177
20.1 Inversion . . . 1177
20.1.1 Cercles perpendiculaires . . . 1177
20.1.2 Inversion . . . 1178

21 Espaces projectifs (Vol 3) 1183
21.1 Sous espaces projectifs . . . 1183
21.2 Espace projectifs comme «complétés» d'espaces affines . . . 1185
21.3 Théorème de Pappus . . . 1188
21.4 Homographies . . . 1189
21.4.1 Homographies . . . 1189
21.4.2 Le groupe projectif . . . 1190
21.4.3 Repères projectifs . . . 1191
21.4.4 Identifications $P(\mathbb{K}^2)$ vers $\mathbb{K} \cup \{\infty\}$ . . . 1195
21.4.5 Birapport . . . 1196
21.5 Coordonnées homogènes . . . 1202
21.5.1 Dualité . . . 1202
21.5.2 Polynômes . . . 1203
21.6 La sphère de Riemann $P_1(\mathbb{C})$ . . . 1205
21.6.1 Éléments de géométrie dans $P_1(\mathbb{C})$ . . . 1205
21.6.1.1 Équation complexe d'une droite . . . 1206
21.6.1.2 Équation complexe d'un cercle . . . 1206
21.6.1.3 Cercle-droite . . . 1207
21.6.1.4 Rotation-homothétie . . . 1208
21.6.1.5 Application linéaire . . . 1208
21.6.1.6 Inversion . . . 1208
21.6.2 Homographies . . . 1209
21.6.3 Birapport . . . 1214
21.6.4 Division harmonique . . . 1217
21.6.5 Groupe circulaire . . . 1221
21.6.6 Action du groupe modulaire . . . 1224

22 Analyse vectorielle (Vol 3) 1229
22.1 Le théorème de Green . . . 1229
22.2 Théorème de la divergence dans le plan . . . 1233
22.2.1 La convention de sens de parcours . . . 1233
22.2.2 Théorème de la divergence . . . 1234
22.3 Théorème de Stokes . . . 1234
22.4 Théorème de Gauss . . . 1236
22.5 Coordonnées curvilignes . . . 1238
22.5.1 Base locale . . . 1238
22.5.2 Importance de l'orthogonalité . . . 1238
22.5.3 Coordonnées polaires . . . 1240
22.5.4 Coordonnées cylindriques . . . 1240
22.5.5 Coordonnées sphériques . . . 1241
22.5.6 Gradient en coordonnées curvilignes . . . 1241
22.5.6.1 Coordonnées sphériques . . . 1242
22.5.7 Divergence en coordonnées curvilignes . . . 1242
22.5.7.1 Coordonnées cylindriques . . . 1244
22.5.7.2 Coordonnées sphériques . . . 1244

22.5.8 Laplacien en coordonnées curvilignes orthogonales . . . 1244
22.5.9 Rotationnel en coordonnées curvilignes orthogonales . . . 1245
22.5.9.1 Coordonnées cylindriques . . . 1246
22.5.9.2 Coordonnées sphériques . . . 1246
22.6 Les formules . . . 1246
22.6.1 Coordonnées polaires . . . 1246
22.6.2 Coordonnées cylindriques . . . 1246
22.6.3 Coordonnées sphériques . . . 1247

23 Espaces de Hilbert (Vol 3) 1249
23.1 Espaces de Hilbert . . . 1249
23.1.1 Sous-espace vectoriel fermé ? ? ? . . . 1250
23.2 Théorème de la projection . . . 1250
23.3 Systèmes orthogonaux et bases . . . 1253
23.3.1 Orthogonal d'une partie . . . 1254
23.3.2 Dual, théorème de représentation de Riesz . . . 1255
23.3.3 Séparabilité . . . 1256
23.3.4 Base hilbertienne . . . 1258
23.3.5 Décomposition dans une base hilbertienne . . . 1263
23.3.6 Digression sur les normes opérateurs . . . 1267
23.3.7 Applications linéaires et continuité . . . 1268
23.4 Théorème de Kochen-Specker . . . 1269
23.5 Théorème de Lax-Milgram . . . 1270

24 Analyse fonctionnelle (Vol 3) 1275
24.1 Théorème d'Ascoli . . . 1275
24.2 Théorème de Banach-Steinhaus . . . 1276
24.3 Espaces L^p . . . 1278
24.3.1 Généralités . . . 1278
24.3.2 L'espace L^∞ . . . 1281
24.3.3 Quelque identifications . . . 1281
24.3.4 Inégalité de Jensen, Hölder et de Minkowski . . . 1282
24.3.5 Ni inclusions ni inégalités . . . 1284
24.3.6 Complétude . . . 1286
24.3.7 Densité des fonctions infiniment dérivables à support compact . . . 1290
24.3.8 Approximation . . . 1293
24.4 Convolution . . . 1294
24.4.1 Approximation de l'unité . . . 1295
24.4.2 Densité des polynôme trigonométrique . . . 1297
24.5 L'espace L^2 . . . 1299
24.5.1 Coefficients et série de Fourier . . . 1300
24.5.2 Dualité et théorème de représentation de Riesz . . . 1302
24.6 Théorèmes de Hahn-Banach . . . 1310
24.7 Théorème de Tietze . . . 1312
24.8 Espace de Schwartz . . . 1315
24.8.1 Topologie . . . 1316
24.8.2 Produit de convolution . . . 1319

25 Analyse complexe (Vol 3) 1321
25.1 Fonctions holomorphes . . . 1321
25.1.1 Équations de Cauchy-Riemann . . . 1321
25.1.2 Intégrales sur des chemins fermés . . . 1322
25.1.3 Lacets, indice et homotopie . . . 1325
25.1.4 Théorème de Cauchy et analycité . . . 1326

25.1.5 Théorème de Brouwer en dimension 2 . . . 1329
25.1.6 Principe des zéros isolés . . . 1330
25.1.7 Prolongement de fonctions holomorphes . . . 1331
25.1.8 Théorème de Runge . . . 1331
25.2 Intégrales de fonctions holomorphes . . . 1333
25.2.1 Mesure de Radon . . . 1338
25.3 Conditions équivalentes à l'holomorphie . . . 1339
25.4 Singularités, pôles et méromorphe . . . 1339
25.5 Fonctions d'Euler . . . 1340
25.5.1 Euler et factorielle . . . 1343
25.6 Partition d'un entier en parts fixées . . . 1344
25.7 Exponentielle et logarithme complexe . . . 1347
25.7.1 Définition et propriétés de l'exponentielle . . . 1347
25.7.2 Intégrale de Fresnel . . . 1349
25.7.3 Logarithme complexe . . . 1351
25.7.3.1 La fonction argument . . . 1351
25.7.3.2 Une définition possible du logarithme . . . 1354
25.7.3.3 Pas plus de continuité . . . 1356
25.7.3.4 Pas d'unicité : autres déterminations de l'argument . . . 1357
25.7.3.5 Pas d'unicité : développement en série . . . 1358
25.7.3.6 Pas d'unicité : laquelle choisir ? . . . 1358
25.7.3.7 Logarithme comme primitive . . . 1359
25.8 Théorème de Weierstrass . . . 1360
25.9 Théorème de Montel . . . 1361
25.10Espaces de Bergman . . . 1363

26 Série de Fourier (Vol 3) 1367
26.1 Densité des polynômes trigonométriques . . . 1367
26.1.1 Convergence pour les fonctions continues (via Weierstrass) . . . 1367
26.1.2 Convergence pour les fonctions continues (via Fejér) . . . 1367
26.1.3 Densité dans L^p . . . 1370
26.1.4 Suite équirépartie, critère de Weyl . . . 1371
26.2 Fonctions de Dirichlet . . . 1373
26.3 Coefficients et série de Fourier . . . 1374
26.3.1 Le contre-exemple que nous attendions tous . . . 1377
26.3.2 Inégalité isopérimétrique . . . 1379
26.3.3 À propos des coefficients . . . 1380

27 Transformation de Fourier (Vol 4) 1383
27.1 Transformée de Fourier dans $L^1(\mathbb{R}^d)$ . . . 1383
27.1.1 Formule sommatoire de Poisson . . . 1386
27.2 Transformée de Fourier dans l'espace de Schwartz . . . 1388
27.2.1 Quelque transformées de Fourier . . . 1391
27.3 Suite régularisante . . . 1393
27.3.1 Formule d'inversion . . . 1395
27.4 Transformée de Fourier sur $L^2(\mathbb{R}^d)$ . . . 1397
27.4.1 Extension de $L^1 \cap L^2$ vers L^2 . . . 1397

28 Distributions (Vol 4) 1401
28.1 Topologie . . . 1401
28.1.1 Dérivée partielle au sens faible . . . 1401
28.2 Distributions . . . 1404
28.2.1 Multiplication d'une distribution par une fonction . . . 1406
28.2.2 Dérivée de distribution . . . 1406

28.2.3 Ordre et support d'une distribution . . . 1407
28.3 Distributions tempérées . . . 1410
28.3.1 Topologie . . . 1411
28.3.2 Distributions associées à des fonctions . . . 1411
28.3.3 Composition avec une fonction . . . 1412
28.3.4 Transformée de Fourier d'une distribution tempérée . . . 1412
28.3.5 Convolution d'une distribution par une fonction . . . 1412
28.3.6 Approximation de la distribution de Dirac . . . 1413
28.3.7 Peigne de Dirac . . . 1416
28.4 L'espace $C^\infty(\mathbb{R}, \mathscr{D}'(\mathbb{R}^d))$ . . . 1417
28.4.1 Dérivation . . . 1419
28.5 L'espace $C^\infty(\mathbb{R}, \mathscr{S}'(\mathbb{R}^d))$ . . . 1420
28.5.1 Propriétés générales . . . 1420
28.5.2 Dérivation . . . 1422
28.6 Une équation de distribution . . . 1424

29 Espaces de Sobolev, équations elliptiques (Vol 4) 1427
29.1 Espaces de Sobolev . . . 1427
29.1.1 Sur un intervalle de $\mathbb{R}$ . . . 1427
29.1.2 Sur un ouvert de $\mathbb{R}^n$ . . . 1433
29.1.2.1 Définition . . . 1433
29.1.3 Espace de Sobolev fractionnaire . . . 1434
29.2 Trace . . . 1437
29.3 Théorème de plongement . . . 1440

30 Équations différentielles ordinaires (Vol 4) 1445
30.1 Équation homogène, solution particulière . . . 1446
30.2 Que faire avec $f(z)dz = g(t)dt$? . . . 1446
30.3 Équations linéaires du premier ordre . . . 1448
30.3.1 Pourquoi la variation des constantes fonctionne toujours ? . . . 1449
30.4 Équations à variables séparées . . . 1450
30.4.1 La méthode rapide . . . 1451
30.4.2 La méthode plus propre . . . 1451
30.4.3 Les théorèmes . . . 1451
30.5 Équations linéaires d'ordre supérieur . . . 1453
30.5.1 Équations et systèmes linéaire à coefficients constants . . . 1453
30.5.2 Si les coefficients ne sont pas constants ? . . . 1454
30.6 Système d'équations linéaires . . . 1454
30.6.1 La magie de l'exponentielle... . . . 1454
30.6.2 ...mais la difficulté . . . 1455
30.6.3 La recette . . . 1455
30.6.4 Système d'équations linéaires avec matrice constante . . . 1456
30.6.5 Système d'équations linéaires avec matrice non constante . . . 1456
30.7 Réduction de l'ordre . . . 1456
30.8 Autour de Cauchy-Lipschitz . . . 1458
30.8.1 Fuite des compacts et explosion en temps fini . . . 1458
30.8.2 Écart entre deux conditions initiales . . . 1460
30.8.3 Flot d'un champ de vecteurs . . . 1461
30.8.4 Stabilité de Lyapunov . . . 1475
30.8.5 Système proie et prédateurs : Lokta-Voltera . . . 1479
30.9 Équation du second ordre . . . 1482
30.9.1 Wronskien . . . 1482
30.9.2 Avec second membre . . . 1482
30.9.3 Équation $y'' + q(t)y = 0$ . . . 1483

30.9.4 Équation de Hill 1484
30.10 Différents types d'équations différentielles 1487
30.10.1 Équation homogène 1487
30.10.2 Équation de Bernoulli 1488
30.10.3 Équation de Riccati 1488
30.10.4 Équation différentielle exacte 1488
30.10.4.1 Résolution lorsque tout va bien 1488
30.10.4.2 Facteur intégrant (quand tout ne va pas bien) 1489
30.11 Distributions pour les équations différentielles 1489
30.11.1 Équation de Schrödinger 1490
30.12 Équations différentielles du premier ordre 1493
30.13 Équations différentielles du premier ordre à variables séparables 1496
30.14 Équations différentielles linéaires du premier ordre 1498
30.14.1 Méthode de variation de la constante 1500
30.15 Équations différentielles linéaires du second ordre 1501
30.15.1 Résolution des équations différentielles linéaires du second ordre homogènes à coefficients constants 1501
30.15.2 Équations différentielles linéaires du second ordre à coefficients constants, non homogènes 1503
30.16 Fonction de Green 1504

31 Équations aux dérivées partielles (Vol 4) 1507
31.1 Symbole principal, équation des caractéristiques 1507
31.2 Méthode des caractéristiques pour l'ordre 1 1507
31.2.1 Un exemple complet un peu minimal 1508
31.2.2 Un théorème d'existence et d'unicité 1510
31.3 Méthode des caractéristique pour l'ordre 2 1513
31.3.1 Principe général 1513
31.3.2 Exemple : l'équation d'onde 1514
31.4 Classification des équations du second ordre 1516
31.4.1 Problème au limite 1517
31.5 Principe du maximum 1518
31.6 Quelque exemples 1522
31.6.1 Un changement de variables 1522

32 Numérique (Vol 4) 1525
32.1 Introduction 1525
32.2 Représentations numériques 1525
32.2.1 Entier relatif en complément à deux (binaire) 1525
32.2.2 Représentation en virgule flottante 1527
32.2.3 Simple précision, IEEE-754 1527
32.3 Problèmes pour écrire des nombres 1530
32.3.1 Troncature : la base 1530
32.3.2 Troncature : le drift 1531
32.3.3 Quelque bonnes règles 1532
32.3.4 Erreur de "cancellation" 1532
32.3.5 Calcul d'une dérivée 1534
32.3.6 Erreur d'absorption 1535
32.4 Conditionnement et stabilité 1535
32.4.1 Comment choisir et penser le K ? 1538
32.5 Un peu de points fixes 1539
32.5.1 Choix de la fonction à point fixe 1539
32.5.2 Convergence quadratique 1541
32.5.3 Convergence 1542

32.6 Méthode de Newton . . . 1543
32.6.1 «Justification» par la formule par Taylor . . . 1544
32.6.2 «Justification» par points fixes . . . 1544
32.6.3 Convergence de la méthode de Newton . . . 1545
32.6.4 Formalisation de l'algorithme . . . 1547
32.6.5 Caractéristiques . . . 1548
32.6.6 Exemple de la racine carré . . . 1548
32.6.7 Si multiplicité . . . 1549
32.6.8 Et la dérivée ? . . . 1549
32.6.9 Méthode de Newton : le cas général . . . 1549
32.7 Estimation de l'ordre de convergence . . . 1551
32.8 Autres méthodes . . . 1552
32.8.1 Méthode de Schröder . . . 1552
32.8.2 Halley . . . 1552
32.9 Méthode des sécantes variables . . . 1553
32.9.1 Aitken . . . 1553
32.10Équations algébrique . . . 1554
32.10.1 Résoudre un système linéaire . . . 1554
32.10.2 Caractéristiques . . . 1555
32.10.3 Définitions . . . 1555
32.11Équations non linéaire . . . 1556
32.11.1 Méthode de bisection . . . 1557
32.12Efficacité . . . 1559
32.13Exercices . . . 1560
32.14Approximations de fonctions . . . 1563
32.14.1 Critère d'interpolation . . . 1563
32.14.2 Base de Newton . . . 1564
32.14.3 Méthode des minima quadratiques . . . 1566
32.14.4 Notre espace de Hilbert . . . 1567
32.14.5 Droite de régression . . . 1568
32.15Conditionnement d'une matrice . . . 1569
32.15.1 Perturbation du vecteur . . . 1571
32.15.2 Perturbation de la matrice . . . 1572
32.16Système linéaires (généralités) . . . 1573
32.16.1 Les méthodes directes . . . 1574
32.16.2 Méthodes itératives . . . 1575
32.17Système linéaires (méthodes directes) . . . 1575
32.17.1 Inversion de matrice triangulaire . . . 1575
32.17.2 Transformation gaussienne . . . 1575
32.17.3 Méthode de Gauss pour résoudre des systèmes d'équations linéaires . . . 1577
32.17.4 Méthode de Gauss sans pivot (décomposition LU) . . . 1578
32.17.5 Matrice de permutation élémentaires . . . 1581
32.18Méthode de Gauss avec pivot partiel (décomposition PLU) . . . 1582
32.18.1 L'idée . . . 1582
32.18.2 Le théorème . . . 1583
32.18.3 D'un point de vue algorithmique . . . 1587
32.18.4 Exemples . . . 1589
32.19Résolution de systèmes linéaires (suite) . . . 1592
32.19.1 Déterminant . . . 1592
32.19.2 Plusieurs termes indépendants . . . 1592
32.19.3 Cholesky . . . 1593
32.20Système linéaire (méthodes itératives) . . . 1596
32.20.1 La méthode générale . . . 1597

32.20.2 Jacobi . . . 1597
32.20.3 Gauss-Seidel . . . 1597
32.20.4 Autres . . . 1597
32.21 Indices connectés, matrice irréductible . . . 1597
32.22 Localisation des valeurs propres . . . 1599
32.22.1 Matrices à diagonale dominante . . . 1602
32.22.2 M-matrice . . . 1604

33 Méthodes des différences finies (Vol 4) 1609
33.1 Problèmes de dimension un . . . 1609
33.1.1 Un schéma à cinq points . . . 1610
33.1.1.1 Poser le système . . . 1610
33.1.1.2 Propriétés du système . . . 1612
33.1.2 Exemple . . . 1614
33.2 Problèmes de dimension deux . . . 1615
33.2.1 Discrétisation en croix . . . 1615
33.2.2 Discrétisation en carré . . . 1616
33.2.3 Résolution de la discrétisation en croix . . . 1617
33.3 Consistance, convergence . . . 1619
33.3.1 Définitions, mise en place . . . 1619
33.3.2 Exemple . . . 1621
33.3.3 Consistance, stabilité et convergence . . . 1622
33.3.4 Exemple : schéma à cinq points, laplacien en croix . . . 1623
33.4 Autres laplaciens . . . 1624
33.4.1 Travail avec le laplacien à 9 points . . . 1627

34 Variables aléatoires et théorie des probabilités (Vol 4) 1629
34.1 Espace de probabilité . . . 1629
34.2 Variables aléatoires . . . 1630
34.2.1 Indépendance . . . 1630
34.2.2 Lois conjointes et indépendance . . . 1633
34.2.3 Somme et produit de variables aléatoires indépendantes . . . 1634
34.2.4 Espérance . . . 1636
34.2.5 Variance . . . 1637
34.2.6 Covariance . . . 1638
34.2.7 Probabilité conditionnelle : événements . . . 1639
34.2.8 Espérance conditionnelle . . . 1640
34.2.9 Probabilité conditionnelle : tribu . . . 1646
34.2.10 Variables de Rademacher indépendantes . . . 1648
34.2.11 Un petit paradoxe . . . 1650
34.2.11.1 «Bonjour, je suis l'aînée» . . . 1650
34.2.11.2 «Bonjour» . . . 1651
34.2.11.3 Le parent qui répond aux questions . . . 1653
34.2.11.4 À propos des simulations . . . 1655
34.2.12 Inégalité de Jensen . . . 1655
34.2.13 Fonction de répartition . . . 1656
34.2.14 Fonction caractéristique . . . 1656
34.2.15 Fonction génératrice des moments, transformée de Laplace . . . 1657
34.2.16 Loi d'une variable aléatoire . . . 1658
34.2.17 Changement de variables . . . 1659
34.3 Convergence . . . 1660
34.4 Loi des grands nombres, théorème central limite . . . 1665
34.4.1 Loi des grands nombres . . . 1665
34.4.2 Théorème central limite . . . 1667

34.4.3 Marche aléatoire . . . 1670
34.5 Les lois usuelles . . . 1671
34.5.1 Loi de Bernoulli . . . 1671
34.5.2 Loi binomiale . . . 1672
34.5.3 Loi multinomiale . . . 1673
34.5.4 Loi géométrique . . . 1673
34.5.5 Loi de Poisson . . . 1674
34.5.6 Loi exponentielle . . . 1674
34.5.7 Approximation de la binomiale par une Poisson . . . 1677
34.5.8 Loi de Poisson et loi exponentielle . . . 1678
34.5.9 Loi normale . . . 1679
34.5.10 Vecteurs gaussiens . . . 1681
34.5.11 Variable aléatoire de Rademacher . . . 1685
34.5.12 Loi de Student . . . 1687
34.5.13 Indépendance, covariance et variance de somme . . . 1688
34.6 Estimation des grands écarts . . . 1688
34.7 Simulations de réalisations de variables aléatoires . . . 1692
34.7.1 Générateur uniforme . . . 1692
34.7.1.1 Première méthode . . . 1692
34.7.1.2 Seconde méthode . . . 1692
34.7.2 Simulation par inversion . . . 1692
34.7.2.1 Loi exponentielle . . . 1693
34.7.3 Algorithme de Box-Muller . . . 1693
34.7.4 Méthode du rejet . . . 1694
34.7.5 Simuler une loi géométrique à l'ordinateur . . . 1696
34.7.6 Simuler une loi exponentielle à l'ordinateur . . . 1696
34.7.7 Simuler une loi de Poisson à l'ordinateur . . . 1696
34.8 Sage . . . 1697
34.8.1 Loi exponentielle . . . 1697
34.8.2 Inverser des lois . . . 1697
34.9 Monte-Carlo . . . 1697
34.9.1 Intervalle de confiance . . . 1698
34.9.1.1 Principe . . . 1698
34.9.1.2 Échantillonnage préférentiel . . . 1699
34.9.1.3 Méthode de la variable de contrôle . . . 1700
34.9.1.4 Variables antithétiques . . . 1700
34.10 Résultats qui se démontrent avec des variables aléatoires . . . 1700
34.10.1 Nombres normaux . . . 1700
34.10.2 Théorème de Bernstein . . . 1702

35 Statistiques (Vol 4) 1707
35.1 Notations et hypothèses . . . 1707
35.2 Modèle statistique . . . 1707
35.3 Modèles d'échantillonnages . . . 1710
35.4 Estimation ponctuelle . . . 1713
35.5 Statistiques et estimateurs . . . 1715
35.5.1 Qualité des estimateurs . . . 1715
35.5.2 Méthode des moments . . . 1716
35.5.3 Méthode de substitution . . . 1718
35.5.4 Méthode du maximum de vraisemblance . . . 1718
35.5.5 Estimation d'une fonction de répartition . . . 1720
35.5.6 Espérance et variance d'un estimateur . . . 1723
35.6 Estimation par intervalle de confiance . . . 1723
35.6.1 Région de confiance . . . 1726

35.6.2 Fonction pivotale . . . 1726
35.6.3 Sondage de proportion . . . 1729
35.7 Estimer une densité lorsqu'on ne sait rien . . . 1730
35.7.1 Distance entre des mesures . . . 1731
35.7.2 Estimateur par fenêtres glissantes . . . 1731
35.8 Test d'hypothèses, prise de décision . . . 1734
35.8.1 Exemple : qualité des pièces d'usine . . . 1734
35.8.2 Exemple : la résistance d'un fil . . . 1734
35.8.3 Vocabulaire et théorie . . . 1735
35.8.4 Risque de première et seconde espèce . . . 1736
35.8.5 Modèle paramétrique de loi gaussienne . . . 1737
35.9 Tests paramétriques . . . 1738
35.10Tests d'adéquation . . . 1739

36 Chaînes de Markov à temps discret (Vol 4) 1745
36.1 Généralités . . . 1745
36.2 Chaînes de Markov sur un ensemble fini . . . 1746
36.3 Marche aléatoire sur $\mathbb{Z}$ . . . 1748
36.3.1 Chaînes de Markov homogènes . . . 1750
36.3.2 Graphe de transition . . . 1751
36.3.3 Chaîne de Markov définie par récurrence . . . 1751
36.3.3.1 Le cas général . . . 1751
36.3.3.2 Exemple : la file de réparation de machines à laver . . . 1753
36.4 Classification des états . . . 1753
36.4.1 Chaînes irréductibles . . . 1756
36.4.2 Nombre de visites . . . 1757
36.5 Mesure invariante . . . 1762
36.6 Convergence vers l'équilibre . . . 1764
36.7 Processus de Galton-Watson . . . 1766

37 Martingales (Vol 4) 1771
37.1 Convergence de martingales . . . 1771
37.2 Temps d'arrêt et martingale terminée . . . 1774
37.3 Décomposition de martingales . . . 1776
37.4 Problème de la ruine du joueur . . . 1778
37.4.1 Le cas où la pièce est truquée . . . 1779
37.4.1.1 Introduction d'une martingale . . . 1779
37.4.1.2 Finitude du temps d'arrêt . . . 1780
37.4.1.3 Temps moyen de jeu . . . 1780
37.4.1.4 Probabilité de victoire du joueur . . . 1781
37.4.2 Le cas où la pièce est non truquée . . . 1782
37.4.2.1 Probabilité de gagner . . . 1782
37.4.2.2 Temps moyen de jeu . . . 1784
37.4.3 Un petit complément . . . 1784

38 Processus de Poisson (Vol 4) 1785
38.1 Processus de Poisson . . . 1785
38.2 Quelques trucs sur la simulation . . . 1788
38.2.1 Le théorème central limite pour Markov . . . 1789
38.2.2 Feuille 5 . . . 1789
38.2.3 Feuille 6 . . . 1789
38.2.4 Feuille 7 . . . 1790
38.2.5 Simuler des lois conditionnelles . . . 1790

39 Utilisation dans les autres sciences (Vol 4) 1791
39.1 Démystification du MRUA . . . 1791
39.1.1 Preuve de la formule . . . 1791
39.1.2 Interprétation graphique . . . 1792
39.2 Relativité en mécanique newtonienne . . . 1792
39.2.1 Relativité du mouvement . . . 1792
39.2.2 Bob et Alice . . . 1792
39.3 Invariance de la vitesse de la lumière . . . 1793
39.3.1 Champ de gravitation et électrique . . . 1793
39.3.1.1 Finitude de la vitesse de propagation de la force électrique . . . 1793
39.3.1.2 Pourquoi pas la gravitation ? . . . 1793
39.3.2 Support du champ : pas d'éther . . . 1794
39.3.3 Le problème . . . 1794
39.4 Conséquences . . . 1794
39.4.1 Ligne d'univers . . . 1794
39.4.2 Transformations de Lorentz . . . 1795
39.4.3 Conditions d'existence . . . 1798
39.4.4 La notion d'intervalle . . . 1799
39.4.4.1 En mécanique newtonienne . . . 1799
39.4.4.2 En mécanique relativiste . . . 1799
39.4.5 Le cône de lumière d'un point . . . 1799
39.4.6 Contraction des longueurs . . . 1800
39.4.7 Dilatation des intervalles de temps . . . 1800
39.4.8 Invariance de l'intervalle . . . 1801
39.4.8.1 Rappel de trigonométrie hyperbolique . . . 1802
39.4.8.2 Les transformations de Lorentz (bis) . . . 1804
39.4.9 Vitesse limite . . . 1806
39.5 Applications . . . 1806
39.5.1 Le GPS . . . 1806
39.5.2 Les ondes électromagnétiques . . . 1806
39.6 Mécanique relativiste . . . 1806
39.6.1 Des problèmes, toujours des problèmes . . . 1807
39.6.2 Loi d'addition des vitesses . . . 1807
39.6.3 L'action d'une force . . . 1808
39.6.4 Équivalence entre la masse et l'énergie . . . 1809
39.7 Principe de correspondance . . . 1809

40 Exemples avec Sage (Vol 4) 1811

41 Épilogue : la constante de Weiner (Vol 4) 1831

42 GNU Free Documentation License (Vol 4) 1833

Bibliographie 1841

Index

λ-système, 695
p-Sylow, 319
p-groupe, 319
Éther, 1794
Événement, 1799
élément
    inversible
        dans un anneau, 142
élémentaire
    polynôme symétrique, 200
équation
    aux variations, 1467
    de Riccati, 1488
    des classes, 132
    des orbites, 132
    différentielle
        étude qualitative, 1486
        Hill, 1484
        homogène, 1487
        système, 1486
    diophantienne, 120, 158, 161
    Fredholm, 1059
    orbite-stabilisateur, 131
équation différentielle
    linéaire du premier ordre, 1498
    linéaire du premier ordre, homogène, 1499
    linéaire du second ordre, 1501
    linéaire du second ordre, homogène, 1501
    premier ordre, 1493
    second ordre, 1494
    variables séparables, 1496
équation homogène associée, 1499
équilibre
    point point une équation différentielle, 1475
équivalence
    de norme, 451
étagée
    fonction, 726
état
    apériodique, 1765
    récurrent, 1754
    récurrent positif, 1754
    transitoire, 1754
étranger
    dans leur ensemble, 166

abélianisé, 110
Abel
    angulaire, 838
    convergence radiale, 906
abscisse
    curviligne, 1143
absolument continue, 830
absorbant, 1753
accélération d'un chemin, 1127
accroissement, 652
accumulation
    dans $\mathbb{R}$, 293
    dans espace vectoriel normé, 379
action, 129
    adjointe, 130
    de groupe
        Wedderburn, 209
    domaine fondamental, 132
    fidèle, 129
    libre, 133
    transitive, 133
action de groupe, 350
    sur des matrices, 1115
adhérence, 291, 375
affine
    application, 494
    espace, 491
    sous-espace, 496
affine (application), 309
aire, 888
algébrique
    extension, 184
    nombre, 191
    par rapport à une extension de corps, 196
algébriquement
    indépendant, 242
algèbre, 148
    de parties, 690
    engendrée, 148
    polynômes, 165
algorithme, 1555
    consistant, 1555

convergeant, 1555
facteurs invariants, 163
fortement consistant, 1555
stable, 1555
alignement
dans un espace projectif, 1184
alterné
groupe, 334
polynôme, 200
alternée
forme linéaire, 343
analytique
au sens complexe, 1326
angle
d'une courbe, 1160
entre deux droites, 976
entre vecteurs, 967
orienté de vecteurs, 997
Anneau
$\mathbb{Z}/n\mathbb{Z}$
polynôme cyclotomique, 205
anneau, 93
$\mathbb{Z}/n\mathbb{Z}$, 209, 213, 329, 464
à division, 171
de séries formelles, 922
utilisation, 1344
euclidien
facteurs invariants, 163
factoriel, 153
intègre, 148
noetherien, 156
principal, 155, 175, 387
utilisation, 177
quotient par un idéal, 144
apériodique
état d'une chaîne de Markov, 1765
chaîne de Markov, 1765
application
définie positive, 357
de classe C^k, 629
différentiable, 591, 593, 626, 629, 1074, 1115
extrema lié, 1089
en escalier, 880
linéaire
théorème de Banach-Steinhaus, 1276
mesurable, 694
multilinéaire, 484
ouverte, 482
semi-définie positive, 357
tangente, 597
application réciproque, 294
approximation
de fonctions
par des polynômes, 1703
de l'unité, 1295
par polynômes, 1051
polynômiale, 1331
arc
géométriques, 1142
paramétré, 1127
arc cosinus, 964
arc sinus, 962
arc tangente, 961
archimédien, 94
associée
subdivision, 880
associés
éléments d'un anneau, 150
asymptotiquement pivotale, 1726
attractif
point fixe, 1056
automorphisme
d'espace vectoriel, 305
axiome
du choix, 89

Bézout
anneau principal, 155
calcul effectif, 117
nombres entiers, 114
polynômes, 183
Baire
espace, 298
théorème, 290, 299
tribu, 688
Banach
espace, 1249
barycentre
cas affine, 497
cas vectoriel, 985
enveloppe convexe, 501
base, 301
canonique de $\mathbb{R}^m$, 304
d'un module, 147
de Newton, 1564
de topologie, 247
dénombrable, 259
espace métrique, 259
duale, 586
espace préhilbertien, 1258
hilbertienne, 1258
utilisation, 1379
locale, 1238
Bergman (espace), 1363
Bernoulli, 1671
somme, 1689
Berry-Esséen (borne), 1670

Bessel
inégalité, 1257
biais
d'estimateur, 1716
bien
conditionné, 1536
enchaîné, 288
bijection, 294
bilinéaire, 484
binormale, 1151
birégulier
point sur une courbe, 1140
birapport, 1196
birapport dans $\mathbb{C} \cup \{\infty\}$, 1214
Bolzano-Weierstrass
espaces métriques, 283
bon
ordre, 90
borélienne
fonction, 710
tribu, 687
boréliens, 687
bord, 292
borné, 248
partie de V, 371
temps d'arrêt, 1774
bornée, 272
différentielle, 625
partie de $\mathbb{R}^m$, 535
suite, 266
boule
avec semi-normes, 295
fermée, 291, 371
ouverte, 247, 291, 371
Bruhat (décomposition), 678
Burnisde
formule, 133

Cône de lumière, 1799
canonique
base, 304
décomposition, 91
espace affine, 491
Cantor
ensemble, 762
caractéristique
d'un anneau, 145
d'une équation différentielle, 1507
polynôme, 389
sous-groupe, 107
caractère, 1029
abélien, 1023
de S_4, 1037
groupe diédral, 1041
irréductible, 1030
cardioïde, 1140
carré
dans un corps fini, 213
catégorie
ensemble de première, 290
Cauchy
critère
uniforme, 647
déterminant, 352
formule, 1326
produit, 904
théorème, 319
Cauchy-Riemann, 1321
Cauchy-Schwarz, 357, 1250
Cayley
théorème, 319
cellule d'un pavage, 879
centrale (application), 1030
centralisateur, 107, 141
centre
d'un anneau, 141
d'un groupe, 107
d'une rotation, 972
cercle
circonscrit à une courbe, 79, 1167
dans la sphère de Riemann, 1206
cercle-droite, 1207
cercles
perpendiculaires, 1177
Cesaro
moyenn, 476
chaîne, 288
de Markov, 1745
apériodique, 1765
convergence, 1766
finie, 1746
homogène, 1745
irréductible, 1751
récurrente positive, 1760
régulière, 1747
champ
conservatif, 850
de vecteurs, 801
champ dérivant d'un potentiel, 850
changement de variable, 1142
Chasles, 491
Chemin
classe C^2, 799
chemin, 799, 1127
dans $\mathbb{R}^p$, 1127
circulation, 848
clôture algébrique, 191

classe
conjugaison
dans S_4, 333
d'association, 150
de conjugaison, 109
classe C^1, 612
classe d'association, 150
codimension, 305
coefficient
de Fourier, 1367
coefficients
de Fourier, 1300
coefficients binomiaux, 143
coercion, 1270
coercive, 1104
colinéarité, 1183
combinaison
convexe, 497
combinatoire, 994
commutant, 431
commutateur
dans un groupe, 109
compacité, 283, 285, 289, 1335, 1360
sous-groupes du groupe linéaire, 680
théorème de Dini, 649
utilisation, 1109
théorème de Montel, 1361
compact, 246, 373, 537
arc paramétré, 1127
Bolzano-Weierstrass dans $\mathbb{R}^n$, 276
boule unité, 276
et fonction continue, 276, 284
fermé et borné, 275
implique fermé, 262
intervalle $[a, b]$, 274
le coup du, 829
localement, 247
opérateur, 1275
produit dénombrable, 290
produit fini, 289
quasi, 247
relatif, 1275
relativement, 247
séquentiellement, 247
suite exhaustive, 287
complément
à deux, 1526
complémentaire, 90
complété
d'un espace métrique, 1123
complétion
projective, 1186
complétude, 1121, 1123, 1286
de $\mathbb{R}$, 467
espaces L^p, 1287
complète
famille de projecteurs, 147
complet
$\mathbb{R}$, 257
corps, 94
espace mesuré, 700
espace topologique, 252
métrique, 252
composante, 652
composition
suite de, 125
concave, 1090
log-concave, 1097
condition initiale, 1494
conditionnement
absolu, 1536
d'une matrice inversible, 1570
relatif asymptotique, 1555
conjugués
éléments d'une extension, 185
connectés
indices d'une matrice, 1597, 1598
connexe
par arc, 278
connexité, 289
définition, 264
et intervalles, 273
fonction holomorphe, 1330
indice d'une courbe, 1325
le groupe $\mathrm{GL}^+(n, \mathbb{R})$, 669
par arc
fonction différentiable, 627
points d'accumulation, 285
prolongement analytique, 1125
signature d'une forme quadratique, 1112
théorème de Runge, 1331
théorème des valeurs intermédiaires, 517
utilisation
Brouwer, 1329
Conservative, 803
consistance
estimateur, 1715
ordre, 1620
constructible
angle, 237
point, 233
réel, 233
construction
des réels, 101
contenu, 167
continue, 514

fonction
en un point, 261
fonction entre espaces métriques, 279
fonction entre espaces topologiques, 261
fonction réelle, 514
forme différentielle, 586
sur espace métrique, 519
uniformément, 535
continuité, 541
fonction définie par une intégrale, 827
séquentielle, 280
sur un intervalle, 514
contraction, 1056
convergeant
schéma discret, 1620
convergence
absolue, 469
commutative, 409
dans un espace vectoriel normé, 467
de martingales, 1775
de suite, 246
en loi, 1660
en probabilité, 1660
normale, 469
ordre, 1551
presque sûrement, 1660
quadratique, 1541
rapidité, 920, 1386, 1387, 1550, 1689
série, 469
suite
dans un corps, 94
suite dans $\mathbb{R}^m$, 293
suite numérique, 266, 1051, 1371
Abel angulaire, 838
uniforme, 646
intégrale, 825
série de fonctions, 470
théorème de Dini, 649
convergent
estimateur, 1715
convexe
courbe plane, 1169
fonction sur $\mathbb{R}^n$, 1098
convexité
barycentre, 500
enveloppe de $O(n)$, 676
fonction, 1090
inégalité de Jensen, 1102
locale, 1311
méthode de Newton, 1546
stricte, 1090
utilisation, 1109
convolution, 1635, 1768
coordonnées
barycentriques, 505
cartésiennes
dans un espace affine, 504
curvilignes, 1238
cylindrique, 1020
dans un espace affine, 492
homogène, 1202
sphériques, 1020
corps, 93
archimédien, 94
complet, 94
de décomposition, 191
de rupture, 189
polynôme cyclotomique, 205
des fractions, 152
des fractions rationnelles
utilisation, 1344
extension, 193, 201
fini, 213, 217, 220
Wedderburn, 209
formellement réel, 171
ordonné, 94
premier, 173
cosinus, 951
angle entre deux vecteurs, 967
hyperbolique, 924
courbe, 1127
étude métrique, 1379
de Jordan, 1176, 1379
efficacité, 1736
fermée, 1157
simple, 1157
courbe de niveau, 571, 574
courbure, 1151
signée, 1158
totale, 1159
covariance, 1638
critère
Abel, 902
Abel pour intégrales, 826
Cauchy
uniforme, 647
de Cauchy, 467
série alternée, 474
Weierstrass, 826
série de fonctions, 898
critique
Galton-Watson, 1768
point, 1085
point d'un arc, 1140
région, 1735
valeur, 1736

cyclique
endomorphisme, 385
groupe, 107
matrice, 385
cycloïde
coordonnées normales, 1145
longueur, 1139

décalage, 1527
décimale
décomposition, 479
décomposition
Bruhat, 678
canonique, 91
corps, 191
Dunford, 416
application, 418
exponentielle de matrice, 417
Jordan
et exponentielle de matrice, 418
polaire, 674
primaire, 415
sous-espaces caractéristiques, 415
spectrale, 415
dénombrable, 92
à l'infini, 247
dénombrement, 994
partitions de $\{1, \dots, n\}$, 922
dérivé
groupe, 109
dérivée
au sens de distributions, 1427
dans Sobolev $H^1(I)$, 1427
directionnelle, 578, 582
distributionnelle, 1406
faible, 1401
fonction à valeurs dans E', 298
partielle, 578, 580, 651
seconde, 563
dérivabilité
fonction définie par une intégrale, 829
lemme de Borel, 942
dérivable, 557, 563
au sens complexe, 893
fonction, 1445
dérivation
au sens des distribution
Sobolev, 1431
déterminant, 343
Cauchy, 1054
d'un endomorphisme, 348
d'une famille de vecteurs, 345
de Cauchy, 352
et inversibilité, 348
forme linéaire alternée, 343
Gram, 352, 1054
interprétation géométrique, 776
résultant, 354, 842
utilisation, 1109
Vandermonde, 349
détermination
logarithme, 1357
principale, 1357
développable
en série entière, 929
développement
asymptotique, 646
limité
en zéro, 638
fonction holomorphe, 1321
premier ordre, 562
Taylor, 1115
degré
application $S^1 \to S^1$, 1161
d'un polynôme, 165
d'une représentation, 1023
extension de corps, 185
dense
nulle part, 290
densité, 1123
conjointe, 1634
d'une variables aléatoire, 1630
dans un espace de fonction
critère de Weyl, 1371
de $\mathbb{Q}$ dans $\mathbb{R}$, 258
utilisation, 1096
de $\mathrm{GL}(n, \mathbb{R})$ dans $\mathbb{M}(n, \mathbb{R})$, 669
de $\mathscr{D}(\mathbb{R}^n)$ dans $L^1(\mathbb{R}^n)$, 1318
de $C_c^\infty(\mathbb{R}^d)$ dans $L^p(\mathbb{R}^d)$, 1291
de $L^2\big([0,1]\big)$ dans $L^p\big([0,1]\big)$, 1292
de $S^+(n, \mathbb{R})$ dans $S^{++}(n, \mathbb{R})$, 673
des fonctions étagées dans L^p, 1291
des polynômes
dans $C_c^0[0,1]$, 1703
matrices diagonalisables dans $\mathbb{M}(n, \mathbb{C})$, 669
mesure, 78, 739
points extrémaux dans $\mathcal{L}$, 675
prolongement, 1121
densité d'une mesure, 739
diédral, 995
diagonale
dominante, 1602
fortement dominante, 1602
strictement dominante, 1602
diagonalisable, 397
et polynôme minimum scindé, 397
exponentielle, 418

diagonalisation
cas complexe, 404
cas réel, 406
endomorphisme auto-adjoint, 446
simultanée, 399
diamètre, 535
difféomorphisme, 616, 862, 1017
de classe C^k, 629
différence
centrée, 1609
divisée, 1565
progressive, 1609
régressive, 1609
différentiabilité, 627
différentiable, 597
dans un Banach, 612
deux fois, 629
sur un ouvert, 603
différentielle, 591, 592, 612
de $u \mapsto u^{-1}$, 621
partielle, 617
totale, 652
dilatation, 662
dilatation (matrice), 162
dimension, 304
n-formes multilinéaires alternées, 343
définition, 304
sous espace affine, 496
utilisation, 501
direction, 579, 1148
sous-espace affine, 496
Dirichlet
noyau, 1368
théorème, 1368
théorème (sur les nombres premiers), 209
disque de convergence, 901, 903
distance, 247
associée à une norme, 251
compatible, 249
entre deux mesures de probabilités, 1731
invariante, 249
point et ensemble, 282
distingué, 107
distribution, 1404
équation de Schrödinger, 1490
de Dirac, 1410
produit par une fonction, 1406
tempérée, 1410
divergence, 655
diviseur
dans un anneau, 141
de zéro, 148
de zéro à droite, 148
polynôme, 166
division
euclidienne, 113, 166
harmonique, 1217
domaine, 511
fondamental d'une action, 132
dominé
modèle statistique, 1714
dominée
convergence (Lebesgue), 738
mesure, 743
droite
dans la sphère de Riemann, 1206
projective, 1183
dual
algébrique, 312
d'un espace de Hilbert, 1255
de $\mathbb{M}(n, \mathbb{K})$, 316
de $L^p(\Omega)$, 1309
de L^p, 1305
topologique, 312
Dunford
décomposition, 416
dyadique, 765

écart-type, 1637
échantillon, 1708, 1710
effectif
empirique, 1739
efficacité
courbe, 1736
d'une méthode itérative, 1559
élément
de surface, 811
de torsion, 136
élément de surface, 793
ellipsoïde, 445
elliptique
équation aux dérivées partielles, 1516
endomorphisme, 305
auto-adjoint, 446
cyclique, 385
décomposition
polaire, 674
diagonalisable, 437, 671, 1486
Dunford, 416
diagonalisation, 406
nilpotent
Dunford, 416
préservant une forme quadratique, 682
sous-espace stable, 416, 1486
engendré, 144
λ-système, 695
corps et anneau, 186

idéal dans un anneau, 144
sous-espace affine, 496
sous-groupe, 108
tribu, 686
par une variable aléatoire, 1631
ensemble
de Cantor, 762
différence symétrique, 91
infini, 89
entrelacement, 1030
enveloppe
convexe, 500
équation
différentielle
linéaire, 1448
ordinaire d'ordre 1, 1445
variables séparées, 1450
générale de degré n, 243
équicontinuité, 1275
équi-intégrable, 1775
équivalence
arcs paramétrés, 1142
chemin, 1134
classe de fonctions, 1278
de représentations, 1030
de suites, 468
homotopie, 1325
norme, 451
relation, 91
suite de composition, 127
erreur, 1566, 1596
assignation, 1530
de consistance, 1620
discrétisation, 1620
quadratique, 1569
troncature, 1530
erreur relative, 1530
escalier, 726
espérance, 1636
conditionnelle, 1640, 1641
événement, 1642
variable aléatoire, 1646
Espace
de Sobolev, 1433
espace
L^2
Sobolev, 1431
L^p, 1279
Banach, 1249
de Baire, 298
de Bergman, 1363
de fonctions
L^p, 1287
Sobolev H^1, 1431
de Hilbert
espace de Sobolev H^1, 1431
de probabilité, 1629
de Schwartz, 1315, 1410
de Sobolev, 1427, 1434
euclidien, 361
métrique, 247
base de topologie, 259
mesuré, 692
complété, 703
mesurable, 685
projectif, 1183
propre, 380
tangent, 1118
topologique, 264
métrisable, 279
vectoriel
dimension, 343
espace vectoriel
topologique, 249
estimateur, 1715
biais, 1716
consistant, 1715
convergent, 1715
de fonction de répartition, 1721
maximum de vraisemblance, 1718
estimation
des grands écarts, 1689
étrangers
polynômes, 166
Euclide
algorithme étendu, 116
lemme, 118
euclidien
anneau, 156
espace, 358
Euler
indicatrice, 139
événement, 1629
exact
intervalle de confiance, 1724
excès
intervalle de confiance, 1724
exhaustive (suite de compacts), 287
exponentielle, 916
complexe, 1347
convergence, 476
de matrice, 418, 486, 938
utilisation, 1082
existence, 913
rapide, 466
unicité, 913

exposant, 326, 437
d'un groupe, 108
extension
corps de base, 420
de corps, 184, 201
algébrique, 184, 189
finie, 222
simple, 186
utilisation, 238
isométrie, 1122
extrémal
point dans un convexe, 675
extrémité
d'un intervalle, 273
extrapolation, 1564
extrema, 1086
lié, 1089
local
relatif, 1088
volume d'un ellipsoïde, 1109
extremum, 1115

facteur
intégrant, 1489
factoriel
anneau, 153
factorisation
de polynôme, 168, 187
faisceau de droites, 1203
famille
sommable, 410
Fatou, 737
Fejér
noyau, 1368
fermé, 245, 292, 373
dans un compact, 262
fermeture, 375
séquentielle, 281
fidèle (action), 129
filtration, 1771
fine
subdivision, 1128
fixateur, 129
flot, 1462, 1507
flux
d'un champ de vecteur, 818
flux d'un champ de vecteurs, 815
fonction, 511
Γ d'Euler, 1341
étagée, 1290
à décroissance rapide, 1315
borélienne, 710
caractéristique, 882
d'une variable aléatoire, 1656
continue
égales, 282
par morceaux, 729
convexe, 1090, 1096
croissante, 511
décroissante, 511
définie par une intégrale, 827, 832, 942, 1360
Γ d'Euler, 1341
utilisation, 1676
de classe $\mathcal{C}^1$, 625
de Dirichlet, 1373
de Möbius, 231
de répartition, 1656
en escalier intégrable, 881
génératrice, 1657
holomorphe, 893, 1360
théorème de Montel, 1361
image, 511
méromorphe
Γ d'Euler, 1341
monotone, 512
par morceaux, 729
valeurs vectorielles, 652
fonctionnelle
énergie, 1272
fondamental
domaine d'une action, 132
forme
bilinéaire, 357
non dégénérée, 443
différentielle, 586
exacte, 835
fermée, 835
linéaire
différentielle, 1089
quadratique, 441, 988, 1112, 1114, 1115
groupe orthogonal, 682
matrice associée, 441
forme canonique
fonction simple, 726
forme linéaire
alternée, 343
formule
Bayes, 1639
Burnside, 133
d'expulsion (produit vectoriel), 368
de Cauchy, 1326
Hadamard, 903
inversion Möbius, 231
probabilité totales, 1639
sommatoire de Poisson, 1386
Stirling, 468
Taylor

reste intégral, 637
utilisation, 1550
Fourier, 1386
série
utilisation, 1379
transformée
groupe abélien fini, 1025
fréquence
empirique, 1739, 1757
fraction
rationnelle
intégration, 842
fraction rationnelle, 872
fractions (corps), 152
Fredholm
équation, 1059
Frenet
formules, 1152
Fresnel
intégrale, 1349
Frobénius
réduction, 428
Frobenius
morphisme, 146
frontière, 292, 378
Fubini
théorème
dans $\mathbb{R}^n$, 791

générateur, 107
génératrice
partie d'un module, 147
géométrie
avec des groupes, 988, 1224
avec nombres complexes, 988, 1224
géométrique
avec des nombres complexes, 1379
Galton-Watson
sous-critique, 1768
sur-critique, 1768
Gauss
lemme
polynômes, 184
somme de, 215
Gershgorin
disque, 1599
Grönwall (lemme), 1445, 1446
gradient, 594, 603, 624
Gram (déterminant), 352
Gram-Schmidt, 361
graphe, 511, 560
de transition (chaîne de Markov), 1751
fonction, 570
fonction de deux variables, 573
groupe
p-groupe, 319
$\mathrm{GL}(n,\mathbb{R})$, 1114
action, 1224
utilisation, 682
agissant sur un ensemble
diédral, 988
alterné, 334
circulaire, 1221
dérivé, 109
de $\mathrm{GL}(n,\mathbb{K})$, 667
de $\mathrm{SL}(n,\mathbb{K})$, 667
du groupe alterné, 337
du groupe symétrique, 335
de Galois, 242
de permutation, 1041
caractères de S_4, 1037
de permutations, 994
de torsion, 136
des isométries
espace métrique, 248
des symétries, 986
diédral, 988, 995
générateurs (preuve), 988
générateurs (utilisation), 1041
en géométrie, 988
et géométrie, 343, 994, 1224
isométries du cube, 462
fini, 209, 213, 321, 329, 464, 994
alterné, 336
diédral, 988
Wedderburn, 209
linéaire, 678
décomposition polaire, 674
enveloppe convexe de $\Omega(n)$, 676
hyperplan, 317
sous-groupes compacts, 680
modulaire, 1224
orthogonal, 363
d'une forme quadratique, 682
partie génératrice, 336, 464, 1224
permutation, 343, 350, 464, 678
diédral, 988
projectif, 1190
quotient, 125
résoluble, 128
spécial orthogonal, 364
symétrique, 134
action sur un triangle, 1027
groupe dérivé
de $\mathrm{GL}(n,\mathbb{C})$, 661

Hadamard
formule, 903

Hardy-Littlewood (théorème), 1051
Hausdorff, 246
Heine (théorème), 536
hermitien
produit scalaire, 362
hessienne, 631
Hilbert, 1249
holomorphe, 893
homéomorphisme, 260
homogène
chaîne de Markov, 1745
homographie, 1189, 1224
sur $\mathbb{C} \cup \{\infty\}$, 1210
homotopie, 1163
hyperbolique
équation aux dérivées partielles, 1516
hyperplan, 981
de $\mathbb{M}(n, \mathbb{K})$, 317
sépare
au sens strict, 1311
séparer
au sens large, 1311
hypothèse
alternative, 1735
composite, 1735
multiple, 1735
nulle, 1735
simple, 1735

idéal
bilatère, 94
dans un anneau, 93
maximal, 153
maximum, 197
principal
à droite, 153
à gauche, 153
identifiable, 1714
identité
polarisation, 441
image, 511
inégalité
arithmético-géométrique, 1102
Bessel, 1257
Cauchy-Schwarz, 357, 1250
de Khintchine, 1686
de la moyenne, 626
des pentes, 1091
Hölder, 1283
utilisation, 1637, 1703
isopérimétrique, 1379
Jensen, 1102
espérance conditionnelle, 1655
pour une somme, 1102
version intégrale, 1282
Kantorovitch, 1103
Markov, 1665
Minkowski, 1284
triangulaire, 247, 250
produit scalaire, 358
incompressible
champ de vecteur, 657
indécomposable
module, 147
indépendance, 1633
événements, 1630
utilisation, 1701, 1778
affine, 504
algébrique, 242
projective, 1191
sous tribus, 1630
variables aléatoires, 1631
indicatrice d'Euler, 139
indice, 123
d'une courbe dans $\mathbb{C}$, 1325
de rotation, 1162
inductif, 90
induite
topologie, 375
tribu, 686
inférence statistique, 1707
infimum, 269
injection, 294
intégrable, 730
fonction non en escalier, 887
fonction positive, 892
intégrale
calcul, 1371
convergente, 774, 842
d'une fonction sur une carte, 794
d'une fonction sur une variété, 799
d'une forme différentielle, 802
fonction en escalier, 881
Fresnel, 1349
impropre, 773, 774
sur un chemin, 799
intégration
fraction rationnelle, 842
intérieur, 291
d'un ensemble, 372
point, 372
intègre
anneau, 148
interpolation, 1564
Intervalle, 1799
intervalle, 90, 511
fermé, 511

longueur, 753
ouvert, 511
intervalle de confiance
asymptotique, 1728
invariant
de similitude, 428
invariante
mesure
pour une chaîne de Markov, 1762
inverse généralisé, 1693
inversion, 1178
dans le groupe symétrique, 332
inversion dans $\mathbb{C} \cup \{\infty\}$, 1208
involution, 399
irréductible
chaîne de Markov, 1751
dans un anneau, 152
module, 147
polynôme, 180
représentation, 1028
irrationalité
$\sqrt{2}$, 97
isobarycentre, 498
isolé
élément de $\mathbb{R}$, 293
point dans un espace vectoriel normé, 379
isométrie
d'espaces métriques, 248
de forme quadratique, 945
de l'espace euclidien $\mathbb{R}^2$, 988
espace euclidien
isométries du cube, 462
forme bilinéaire, 358
groupe, 248
isomorphisme
$(\mathbb{Z}/p\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z}$, 213
d'anneaux, 93
d'espaces topologiques, 260
de corps, 93
espace affine, 495
espace vectoriel normé, 459
espaces vectoriels, 305
isotrope
cône, 444
totalement, 444
isotrope (vecteur), 444

jacobien, 624, 654
Jordan
chemin, 818
courbe, 1379
réduction, 430
Jordan-Hölder, 125

Kronecker, 304

lacet, 1325
Lagrange
multiplicateur, 1089
polynôme, 316
théorème, 123
lagrangien, 1089
Laplace
transformée, 1657
laplacien, 1244
Legendre
symbole, 213
Leibnitz, 559, 562
applications entre espaces vectoriels normés, 617
lemme
Borel, 942
d'Euclide, 118
de Borel-Cantelli, 1663
de Gauss
contenu de polynôme, 221
pour des entiers, 118
de Morse, 1115
de Schreider, 127
de Slutsky, 1662
de transfert, 711
de Zorn, 90
des noyaux, 381
Fatou, 737
Gauss
dans un anneau principal, 155
polynômes, 184
Grönwall, 1445, 1446
Hadamard, 834
regroupement, 1633
Schur complexe, 401
Schur réel, 405
Levi-Civita, 1239
libre
action, 133
partie, 301
partie d'un module, 147
Ligne d'univers, 1795
limite, 413
d'ensembles, 694
d'une fonction, 260
de fonctions holomorphes, 1360
de suite
espace topologique, 246
fonction, 260, 539
fonction de plusieurs variables, 540
inférieure, 719, 720
inversion, 922, 1051, 1279, 1360

permutation
utilisation, 1658
suite, 467
suite dans $\mathbb{R}^m$, 293
suite numérique, 266
supérieure, 719, 720
unicité, 260
linéaire
application, 305
Lipschitz, 1056
localement, 628
Lipschitzienne, 627
localement
intégrable, 773
log-concave, 1097
logarithme
complexe, 1354
dans $\mathbb{C}$, 1351
de matrice, 939
sur les réels positifs, 917
loi
χ^2, 1687
binomiale
comportement asymptotique, 1689
conjointe, 1633
d'une variable aléatoire, 1658
de Poisson, 1674
des grands nombres
forte, 1666
pour les chaînes de Markov, 1764
processus de Poisson, 1786
utilisation, 1689, 1701
marginale, 1633
normale
vecteur gaussien, 1681
parente, 1707, 1708
parente d'un échantillon, 1710
réciprocité quadratique, 217
sans mémoire, 1676
Student, 1687
Lokta-Voltera, 1479
longueur
élément de, 1138
arc géométrique, 1143
d'un arc paramétré compact, 1128
d'un intervalle, 753
d'une arrête, 878
longueur d'arc, 1130

M-matrice, 1604
méromorphe, 1340
Méthode
de Newton, 1545
méthode
des chemins, 548
Newton, 1550
cas convexe, 1546
métrique
discrète, 248
métrisable, 254
maigre, 687
maigre (ensemble), 290
majorant, 268, 269
essentiel, 1281
Markov
inégalité, 1665, 1713
martingale, 1771
bornée dans $L^2(\Omega)$, 1772
matrice, 678, 1224
équivalence, 307
dans le groupe linéaire, 682
associée à une forme quadratique, 441
compagnon, 426
creuse, 1574
cyclique, 385
de dilatation, 162
de permutation, 162
de similitude, 893
de Sylvester, 352
de transition, 1745
de transvection, 162
dense, 1574
hermitienne
racine carré, 671
jacobienne, 603, 624
normale, 404
orthogonale, 363
permutation
élémentaire, 1581
réductible, 1598
racine carré, 671
semblable, 671
semblables, 988, 1114
stochastique, 1745
symétrique, 1114
réelle, 1112
trigonalisable, 400
matrices
similitude, 308
maximal
idéal, 153
maximale
partie orthonormale, 1259
maximum, 268, 271
global, 1084
local, 1084
mesurable

application, 694
au sens de m^*, 699
ensemble, 692
fonction, 710
Lebesgue, 877
mesure
σ-finie, 692
absolument continue, 743
angle entre vecteurs, 998
complexe, 744
dans une carte, 793
de Borel, 716
de comptage, 693, 790
de Haar, 1062, 1063
de Lebesgue, 756
de Radon, 716, 1338
extérieure, 689
externe, 877
finie, 692
sur algèbre de partie, 690
image, 713
positive, 692
probabilité, 1629
produit, 750
régulière, 716
extérieure, 716
intérieure, 716
sur algèbre de partie, 690
minimum, 271
ensemble ordonné, 89
minorant, 269
modèle
échantillonnage, 1708, 1710
paramétrique, 1708
statistique, 1707
modulaire (groupe), 1224
module
de continuité, 1120
indécomposable, 147
irréductible, 147
simple, 147
sur un anneau, 146
moment, 1636
fonction génératrice, 1657
monôme, 165
monogène, 107
extension de corps, 186
monotonie, 1483
morceau
fonction continue ou monotone, 729
morphisme
d'algèbres, 148
d'anneaux, 93
de corps, 93
espace vectoriel normé, 459
Frobenius, 146
moyenne
de Cesaro, 476
empirique, 1638
empirique d'un échantillon, 1711
quadratique, 1637
multiplicateur
de Lagrange, 1089
multiplicité
racine d'un polynôme, 168
racine de $f(x) = 0$, 1544
valeur propre
algébrique, 390
géométrique, 390

négatif, 103
négligeable
partie d'un espace mesuré, 700
nabla, 611
Newton
méthode, 1550
nilpotent, 142
niveau de confiance, 1724
nombre
complexe
norme 1, 209
dénormalisé, 1528
de Fermat, 238
de Néper, 914
normal, 1701
normalisée, 1528
premier, 113, 175, 209, 213, 329, 464, 920
dans leur ensemble, 113
deux nombres entre eux, 113
théorème des deux carrés, 177
tours d'une courbe plane, 1161
nombre premier
polynôme cyclotomique, 205
non dénombrable, 92
normé
espace vectoriel, 250
normal
arc paramétré, 1143
endomorphisme, 404
nombre, 1701
sous-groupe, 107
normal extérieur
vecteur, 817
normale
loi réduite, 1680
principale, 1151
normalisateur, 107

norme, 250
  équivalence, 451
  d'algèbre, 455
  d'application linéaire, 453
  d'une application linéaire, 453
  euclidienne, 251
    dans $\mathbb{R}^m$, 449
  supremum, 251
  vecteur, 359
noyau
  d'une forme bilinéaire, 444
  Dirichlet, 1368
  Fejér, 1368
nulle part dense, 290

observation, 1629
opérateur
  adjoint, 362
  autoadjoint, 363
  définit positif, 406
  hermitien, 363
opérateurs
  compatibles, 1269
opposés
  chemins, 1134
orbite
  d'un point sous une action, 129
ordre, 89
  élément, 108
  bon ordre, 90
  d'un groupe, 108
  d'un polynôme, 167
  dans un corps, 94
  de convergence d'un schéma, 1620
  distribution, 1407
  partiel, 89
  sur un anneau factoriel, 153
  total, 89
orientable
  variété, 796
orientation, 795
origine
  abscisse curviligne, 1143
  repère affine, 504
orthogonal, 448, 1254
  coordonnées curviligne, 1238
  famille de projecteurs, 147
  matrice, 363
  opérateur, 363
  sous-espace, 312
  vecteur, 359
orthonormé, 449
  système, 1256
oscillation
  d'une fonction, 530
  d'une fonction en un point, 530
osculateur (cercle), 1156
ouvert, 245, 272, 292, 373
  dans $\mathbb{R}^n$, 291

pôle, 1005
parabolique
  équation aux dérivées partielles, 1516
parallèle
  sous-espaces affines, 496
paramétrages
  admissible, 1142
paramétrisation, 1134
  normale, 1143
Parseval, 1263
partie
  génératrice, 301
  régulière, 638
  totale, 1256
partie génératrice, 107
partition
  d'un entier en parts fixées, 1344
  dénombrable mesurable, 722
  de l'unité, 798
pavé, 877, 878
pavable, 878
Pearson
  theoreme, 1740
peigne de Dirac, 1416
permutation, 134
  matrice, 162
permuter
  dérivée et intégrale
    $\mathbb{R}^n$, 832
    dans $\mathbb{R}$, 829
    dans $\mathbb{R}$ avec les bornes, 831
  dérivée et limite, 649
  différentielle et intégrale
    $\mathbb{R}^n$, 833
  intégrale
    et série, 790
  limite et intégrale, 827
    convergence dominée, 738
    convergence monotone, 735
    espace mesuré, 827, 828
  série entière et dérivation, 908
  série entière et intégration, 910
  somme et intégrale, 736, 899
petit théorème de Fermat, 173
PGCD
  dans un anneau intègre, 150
  polynômes, 183
pgcd, 113, 142

calcul effectif, 116
pivotale, 1726
plan
projectif, 1183
tangent, 612, 652
Plancherel, 1263
plongement, 1121
Poincaré (demi-plan), 1224
point
d'équilibre
stable, 1475
pondéré, 497
point critique
définition, 1116
point fixe, 1768
attractif, 1056
Brouwer, 1059
Picard, 1057
Schauder, 1061
Poisson
formule sommatoire, 1386
processus, 1785
polarisation (identité), 441
polynôme
à plusieurs indéterminées, 220, 355
alterné, 200
annulateur, 381, 384
caractéristique, 389, 403
contenu, 167
cyclotomique, 203
irréductibilité, 205
propriétés, 205
d'endomorphisme, 671
décomposition de Dunford, 416
de Bernstein, 1703
irréductible, 180
séparable, 192
sur $\mathbb{F}_q$, 232
Lagrange, 316
minimal, 188, 383
d'un élément d'une extension, 184
ponctuel, 384
racines, 201
séparable, 192
scindé, 181
semi-symétrique, 200
symétrique, 200, 201, 220, 350, 355
élémentaire, 200, 202
trigonométrique, 1298
portée
mesure, 743
positif, 103
potentiel, 806, 850
PPCM
dans un anneau intègre, 150
ppcm, 113
précision
simple, 1527, 1528
préhilbertien, 1249
premier
corps, 173
deux éléments d'un anneau principal, 154
deux polynômes entre eux, 167
idéal, 154
sous corps, 173
premier temps d'atteinte, 1753
premier type
région solide, 890
presque
nulle, 136
partout, 694
surjective, 1312
primitif
élément d'un corps, 213
élément d'une extension de corps, 186
polynôme, 167
au sens du pgcd, 167
racine, 225
triplet pythagoricien, 159
primitive, 557, 848
de fonction continue, 1049
et intégrale, 770, 774
fonction, 569
principal
anneau, 154
idéal, 153
principe
prolongement analytique, 1125
zéros isolés, 1330
Principe de correspondance, 1810
probabilité
conditionnelle, 1639
problème
aux limites d'évolution, 1517
aux limites stationnaires, 1517
bien posé, 1517
limite de Dirichlet, 1517
limite de Von Neumann, 1517
problème de Cauchy, 1494
processus
adapté à une filtration, 1771
arrêté, 1776
croissant prévisible, 1776
Galton-Watson, 1766
Poisson, 1785
sans mémoire, 1676

produit
d'espaces vectoriels normés, 480
d'une mesure par une fonction, 78, 739
de Cauchy, 904
de convolution, 1294
et Fourier, 1383
distribution et fonction, 1406
espaces mesurés, 752
espaces topologiques, 246
mixte, 367, 369
scalaire
en général, 357
hermitien, 362
sur $\mathbb{M}(n, \mathbb{R})$, 461
sur $\mathbb{R}^n$, 447
semi-direct, 134
tensoriel
de représentations, 1036
vectoriel, 367, 368
produit remarquable, 1333
projecteur
dans un module, 147
projectif
complétion, 1186
droite, 1183, 1184
espace, 1183
groupe, 1190
hyperplan, 1184
plan, 1183
repère, 1191
sous-espace, 1183
projection
orthogonale, 1252
prolongement
analytique, 1125
utilisation, 1362
de fonctions, 1121
lemme de Borel, 942
méromorphe de la fonction Γ, 1341
par continuité, 526, 529
dans $H^1(I)$, 1430
par densité, 1121
théorème de Hahn, 708
propriété d'intersection non vide, 261
puissance
d'un point, 1177
d'un test, 1736
d'une inversion, 1208

quasi-compact, 247
quaternion, 196
quotient, 113, 166
dans une suite de composition, 125
de groupe, 125
de groupes, 328

réciproque
continuité, 294
dérivabilité, 561
récurrent
état, 1754
nul, 1754
point d'un système dynamique, 718
positif, 1754
réduction
d'endomorphisme, 416
Frobénius, 428
Jordan, 430
réel, 101
réflexif, 1309
réflexion, 949
dans $\mathbb{R}^2$, 969
glissée, 978
par rapport à un hyperplan, 982
région
critique, 1735
de confiance exact, 1726
de rejet, 1735
régularité
d'une mesure, 716
extérieure de la mesure de Lebesgue, 759
intérieure de la mesure de Lebesgue, 761
régulière
surface, 807
régulier
arc, 1140
chemin, 818
point d'un arc, 1140
polyèdre, 950
régulier à droite, 141
répulsif
point fixe, 1056
résidu
méthode itérative, 1596
résolvante, 1456
résultant, 353
utilisation, 355, 842
règle
du produit nul, 148
Règle de Leibnitz, 562
racine
carré
de matrice hermitienne, 671
carré de matrice
hermitienne positive, 671
d'un polynôme, 167
de l'unité, 136, 205, 209, 437, 988, 1224
primitive, 138

utilisation, 205
de polynôme, 188
multiple, 1544
primitive, 225
simple, 1544
racine carré, 518
raffinement, 1128
subdivision d'un pavé, 879
rang, 306, 343, 988
classe d'équivalence, 308
diagonalisation, 406
différentielle, 1089
utilisation, 1255
rare, 687
rayon
de convergence, 901
de courbure, 1151
de torsion, 1152
spectral, 417, 455
recouvrement, 537
rectangle
produit de tribus, 745
rectifiable, 1128
arc géométrique, 1143
rejet
région dans une prise de décision, 1735
relèvement, 1161
relations
coefficient-racines, 202
de Chasles, 491, 866
relativement
compact, 247, 1275
repère
affine, 504
cartésien
espace affine, 491
de Frenet, 1151
projectif, 1191
Représentation
virgule flottante normalisée, 1527
représentation
de groupe fini
caractères de S_4, 1037
fidèle, 1023
groupe diédral, 1041
irréductible, 1028
produit tensoriel, 1036
régulière gauche, 1032
virgule fixe, 1527
reste, 113, 166
d'un développement limité, 638
risque
première espèce, 1736
quadratique, 1715
seconde espèce, 1736
rotation
en dimension 2, 971
rotation-homothétie, 1208
rupture
corps, 189

séparé, 246
espace topologique, 246
séparable, 1256
élément d'une extension, 194
espace topologique, 246
extension de corps, 194
polynôme irréductible, 192
polynôme non constant, 192
sépare
les points, 1046
série
dans un espace vectoriel normé, 469
de Fourier, 1374, 1386
utilisation, 1379
de Laurent, 1340
de puissance, 901
divergence, 469
donnant $(1-A)^{-1}$, 471
entière, 901, 922, 1386
Abel angulaire, 838
fonctions holomorphes, 1326
processus de Markov, 1768
utilisation, 1344, 1658
fonctions, 1051, 1386
génératrice d'une suite, 929
utilisation, 1344
géométrique, 475
harmonique, 475
nombres, 1051
numérique, 920, 922
Riemann, 475
Taylor, 930
schéma
consistant, 1620
schéma numérique, 1620
Schrödinger, 1490
Schur (théorème), 1030
section, 580
de graphe, 571
propriété des, 747
segment
dans $\mathbb{R}^p$, 273
dans un espace affine, 498
semblables
matrices, 393
semi-définie positive, 406

semi-norme, 295
semi-simple
endomorphisme, 386
semi-symétrique
polynôme, 200
signature
d'une permutation, 332
similitude, 893
simple
extension de corps, 186
fonction, 726
groupe, 107
module, 147
singularité, 1339
effaçable, 1339
pôle, 1339
sinus, 951
hyperbolique, 924
sinus cardinal, 875
solfège, 137
solution
générale, 1493
particulière, 1493
somme
inférieure, 884
partielle, 469
supérieure, 884
somme directe (de représentations), 1027
somme partielles
Abel angulaire, 838
sommet, 1175
sous anneau, 144
sous arc, 1127
sous-additivité
sur algèbre de parties, 690
sous-espace
affine engendré par une partie, 496
caractéristique, 414
sous-groupe
caractéristique, 107
distingué, 329
dans le groupe alterné, 336
engendré, 108
normal, 125, 328
sous-martingale, 1771
sous-module, 147
sous-suite, 535
spectre
d'un endomorphisme, 380
matrice hermitienne, 403
matrice symétrique réelle, 406
sphère, 371
de Riemann, 1205
stabilisateur, 129
stabilité
d'un point d'équilibre, 1475
Lyapunov, 1475
stable, 1535
schéma numérique, 1623
stathme
sur $\mathbb{Z}[i]$, 174
stathme euclidien, 156
stationnaire
chaîne de Markov, 1762
statistique, 1715
statistiques
descriptives, 1707
strictement
convexe
sur $\mathbb{R}^n$, 1098
structure
complexe, 1158
structure d'anneau canonique, 141
Student, 1687, 1728
subdivision, 879
associée à une fonction, 880
d'un intervalle, 1128
subordonnée
norme, 453
suite, 136
équirépartie, 1371
critère de Weyl, 1371
arithméticogéométrique, 476
définie par itération, 1550
de Cauchy, 252
dans un corps, 94
de composition, 125
de fonctions, 1279
théorème de Montel, 1361
de fonctions intégrables, 827, 1360
de Jordan-Hölder, 125
exacte, 134
régularisante, 1393
support, 881
d'une permutation, 332
distribution, 1407
famille d'éléments, 136
supremum, 268, 269
d'une suite d'ensembles, 685
sur-martingale, 1771
surface paramétrée, 806
surjection, 294
Sylow
p-Sylow, 319
Sylvester (matrice), 352
symétrique

polynôme, 200
symbole
de Legendre, 213
symbole principal, 1507
système
fondamental, 1453
orthonormé, 1256
trigonométrique, 1257, 1297

tétraèdre, 950
tangent
vecteur unitaire, 1151
tangente, 960, 1148
tangente à un chemin, 1118
Tangente hyperbolique, 1804
tangente hyperbolique, 925
taubérien, 1049
taux d'accroissement, 1091
Taylor, 634
série entière, 929
temps d'arrêt, 1774
temps de retour, 1753
terminée
martingale, 1775
test, 1735
bilatéral, 1736
unilatéral, 1736
théorème
élément primitif, 195, 222, 223
accroissements finis, 619
dérivée directionnelle, 582
dans $\mathbb{R}$, 566
forme générale, 626
Ascoli, 1275
Bézout
polynômes, 183
utilisation, 353
Baire, 290
Banach-Steinhaus, 1276
avec semi-normes, 1277
base incomplète, 303
Beppo-Levi, 735
Bolzano-Weierstrass, 283
Bolzano-Weierstrass dans $\mathbb{R}^n$, 276
Borel-Cantelli, 1629
Borel-Lebesgue, 275
Brouwer, 1060
dimension 2, 1329
Carathéodory, 501
Cauchy
groupe, 124
Cauchy-Arzela, 1071
Cauchy-Lipschitz, 1064
Cayley-Hamilton, 392, 670
central limite, 1667
processus de Poisson, 1787
Chevalley-Warning, 220
chinois, 179
anneau des polynômes, 180
anneau principal, 154
Cochran, 1712
Cochrane, 1713
convergence
dominée de Lebesgue, 738
monotone, 735
d'Alembert-Gauss, 180
décomposition des noyaux
et exponentielle de matrice, 383
de Baire, 299
de Jordan, 1176
de représentation de Riesz, 1255
des deux carrés, 177
version faible, 175
Dini, 649
Dirichlet, 1368
forme faible, 209
Doob, 1775
du rang, 306
extension d'isométrie, 1122
extrema
lié, 1089
Fejér, 1369
fonction implicite dans $\mathbb{R}^n$, 1077
fonction implicite dans Banach, 1076
fondamental du calcul intégral, 770
Fubini
dans $\mathbb{R}^n$, 791
espace mesuré, 786
version compacte dans $\mathbb{R}^2$, 852
Fubini-Tonelli, 783
fuite des compacts, 1459
Gauss
polynômes, 183
Gauss-Wantzel, 238
Glivenko-Cantelli, 1721
Hahn-Banach, 1310
Hardy-Littlewood, 1051
Heine, 536
incidence, 1184
inversion locale, 1074
utilisation, 1089, 1115
isomorphisme
second, 111
troisième, 112
isomorphisme de Banach, 1275
Jordan, 1379
Kronecker, 355

Lagrange, 123
Lie-Kolchin, 439
Lokta-Voltera, 1479
Markov-Takutani, 1063
Montel, 1361
Pappus
affine, 1188
projectif, 1188
Pearson, 1740
petit de Fermat, 173
Picard, 1057
point fixe
Brouwer, 1329
projection
cas vectoriel, 1252
partie fermée convexe, 1250
prolongement de Hahn, 708
prolongement de Riemann, 1340
Radon-Nikodym, 743
complexe, 744
représentation de Riesz, 1309
Rolle, 565
Rothstein-Trager, 842
Runge, 1331
Schauder, 1061
Schur, 1030
spectral, 415
autoadjoint, 446
matrice symétrique, 406
matrices normales, 404
stabilité de Lyapunov, 1475
Stone-Weierstrass, 1046, 1048
Sylvester, 988
taubérien, 1049
taubérien faible, 840
transfert, 1659
Tykhonov, 288
dénombrable, 290
fini, 289
valeurs intermédiaires, 517
Von Neumann, 1082
Wedderburn, 209
Weierstrass, 284
topologie, 245, 375
$*$-faible, 297, 1405
p-adique, 488
et semi-normes, 295
faible, 455
forte, 454
induite, 263
métrique, 248, 272
produit, 246
sur $\mathscr{D}(\Omega)$, 1402
sur $\mathscr{D}(K)$, 1402
sur $C^\infty(\Omega)$, 1402
sur dual topologique, 297
usuelle sur $\mathbb{R}^n$, 375
topologique
somme directe, 1254
torsion, 1152
d'un groupe, 136
totale, 1256
trace, 1437
dual de $\mathbb{M}(n,\mathbb{K})$, 316
endomorphisme, 394
matrice, 394
produit scalaire sur $\mathbb{M}(n,\mathbb{R})$, 461
unicité pour la propriété de trace, 317
transcendant, 196
par rapport à une extension de corps, 196
transformée
de Cauchy, 1338
de Fourier, 1383, 1656
continuité, 1384
groupe abélien fini, 1025
Fourier
distribution tempérée, 1412
Laplace, 1657, 1658
transformation
Fourier, 1386
gaussienne, 1575
transient
état, 1754
transition
probabilité, 1745
transitive, 133
transitoire
état, 1754
transposée, 313
transvection, 662
transvection (matrice), 162
transversale, 132
tribu, 685
borélienne, 687
de Baire, 688
de Lebesgue, 756
engendrée, 686
par un événement, 1631
par une application, 711
par une variable aléatoire, 1631
induite, 686
produit, 745
trigonalisation
et polynôme caractéristique, 400
simultanée, 439
triplet

pythagoricien, 159
type
fini
en algèbre, 196
espace vectoriel, 302

unicité
des mesures, 697
unipotent, 142
unitaire
normale principale, 1151

valeur
principale (distribution), 1410
propre, 380
singulière, 420
valeur absolue
p-adique, 488
valeur principale, 1351
valeur propre
d'une forme quadratique, 443
valuation
p-adique, 488
d'un polynôme, 166
Vandermonde (déterminant), 349
variété, 1089
variété
orientée, 796
variable
de décision, 1736
variable aléatoire, 1630
absolument continue, 1630
Bernoulli
marche aléatoire, 1748
utilisation, 1703
binomiale
utilisation, 1778
centrée, 1636
de Bernoulli
utilisation, 1778
de Rademacher, 1685
intégrable, 1636
suite de variables aléatoire de Bernoulli, 1766
variance, 1637
empirique, 1637, 1711
empirique corrigée, 1711
vecteur gaussien, 1681
variation des constantes, 1449, 1454
vecteur
cyclique, 385
gaussien, 1681
propre, 380
unitaire normal, 1151
unitaire tangent, 1149
Vitali (ensemble), 764
vitesse d'un chemin, 1127
voisinage, 272, 375
volume
d'une région solide, 891
région bornée dans $\mathbb{R}^3$, 888
vraisemblance, 1718

Weiner
constante, 1831
Wronskien, 1482

Liste des notations

$N \lhd G$ Le sous-groupe N est normal dans G, page 107

Algèbre

$\mathcal{C}^1(U, \mathbb{R}^n)$ Les applications une fois continument dérivables, page 625

$\mathcal{L}(E, F)$ Ensemble des applications linéaires de E dans F, page 305

$\mathbb{K}(A)$ corps contenant $\mathbb{K}$ et A, page 186

$\mathbb{K}[A]$ anneau contenant $\mathbb{K}$ et A, page 186

$\mathbb{N}_0$ les naturels non nuls : $\mathbb{N}_0 = \mathbb{N}\backslash\{0\}$, page 92

$\mathcal{M}_{n\times m}$ l'ensemble des matrices $n \times m$, page 309

∇f gradient de la fonction f, page 603

$\mathtt{proj}_V$ projection de $V \times W$ sur V, page 482

$\mathrm{Span}(A)$ l'ensemble des combinaisons linéaires finies d'éléments de A, page 301

$df_a(u)$ Application de la différentielle de f sur le vecteur u, page 591

$f^{(n)}$ La n-ième dérivée de la fonction f, page 634

$o(x)$ fonction tendant rapidement vers zéro, page 635

(p) idéal engendré par p, page 144

$[\mathbb{L} : \mathbb{K}]$ degré d'une extension de corps, page 185

$\mathbb{F}_p$ lorsque p est premier, page 173

$\mathbb{F}_{p^n}$ corps fini à p^n éléments, page 211

$\mathbb{Q}$ corps des fractions sur $\mathbb{Z}$, page 152

$\mathrm{Frac}(\mathbb{A})$ Le corps des fractions de l'anneau $\mathbb{A}$, page 152

$\mathrm{Fun}(X, Y)$ les applications de X vers Y, page 141

$\mathrm{S}(E)$ Les opérateurs auto-adjoints de E, page 446

U_n Le groupe des racines n^{e}de l'unité., page 136

$\mathrm{res}(P, Q)$ résultat des polynômes P et Q, page 353

$\sqrt{A}$ racine d'une matrice hermitienne positive, page 671

$\theta_\alpha(P)$ la multiplicité de α par rapport à P, page 168

$A[X]$ tous les polynômes de degré fini à coefficients dans A, page 165

$A_n[X]$ les polynômes à coefficients dans A et de degré inférieur à n, page 165

$C(P)$ matrice compagnon, page 427

$D \mid P$ D divise P, page 166

$E_\lambda(u)$ Espace propre de u, page 380

$mat_{\mathcal{B}}(q)$ matrice de q dans la base $\mathcal{B}$, page 1108

$U(A)$ ensemble des inversibles, page 142

Ensembles de matrices

$\mathrm{Aut}(E)$ automorphisme de l'espace vectoriel E, page 305

$\mathrm{End}(E)$ les endomorphismes de E, page 305

$L(E,F)$ applications linéaires bornées (continues), page 459

$\mathrm{O}(n,\mathbb{R})$ le groupe des matrices orthogonales, page 364

$\Omega(E)$ formes quadratiques non dégénérées, page 1111

$Q(E)$ formes quadratiques réelles sur E, page 441

$Q^+(E)$ formes quadratiques positives, page 1111

$Q^{++}(E)$ formes quadratiques strictement définies positives, page 1111

$S^+(n,\mathbb{R})$ matrices symétriques définies positives, page 406

$S^{++}(n,\mathbb{R})$ matrices symétriques strictement définies positives, page 406

$S_n^{p,q}(\mathbb{R})$ matrices symétriques réelles de signature (p,q), page 1111

$\beta(s)$ Vecteur unitaire de la binormale, page 1151

$\gamma \sim g$ Équivalence d'arcs paramétrés, page 1142

$\nu(s)$ Vecteur unitaire de la normale principale, page 1151

$c(s)$ rayon de courbure, page 1151

$t(s)$ Torsion, page 1152

Géométrie

$(x_0 : \ldots : x_n)$ coordonnées homogènes dans un espace projectif, page 1202

$\mathcal{C}\mathrm{onv}(A)$ enveloppe convexe, page 500

$\mathbb{C}[G]$ combinaisons d'éléments de G à coefficients dans $\mathbb{C}$, page 1027

$\mathrm{PGL}(E)$ groupe projectif, page 1190

B^o orthogonal dans le dual, page 312

$P(E)$ l'espace projectif de E, page 1183

$P_1(\mathbb{C})$ sphère de Riemann, page 1205

Chaînes de Markov

$\pi(x)$ lié au temps de retour, page 1757

Probabilités et statistique

$\sigma(X)$ La tribu engendrée par la variable aléatoire X, page 1631

$a \wedge b$ $\min(a,b)$, page 1775

K_X matrice de covariance d'un vecteur gaussien, page 1681

$m(\mathcal{A})$ Ensemble des fonctions $\mathcal{A}$-mesurables, page 710

Théorie des groupes

$(G/H)_g$ classes à gauche, page 130

$[a]_p$ ensemble des $a + kp$, page 121

$[G, G]$ groupe dérivé, page 109

$[g, h]$ commutateur dans un groupe, page 109

$\mathrm{Aff}(\mathbb{R}^n)$ Le groupe des applications affines bijectives de $\mathbb{R}^n$., page 508

gr groupe engendré, page 107

$\hat{G}$ groupe des caractères de G, page 1023

σ_x réflexion par rapport à x, page 949

A_n groupe alterné, page 334

$D(G)$ groupe dérivé, page 109

D_n groupe diédral, page 988

G^{ab} groupe abélianisé de G, page 110

$N \times_\phi H$ produit semi-direct, page 134

S_n le groupe symétrique, page 134

Topologie et théorie des ensembles

$\complement A$ Le complémentaire de l'ensemble A, page 90

$\mathrm{Diam}(A)$ Diamètre de la partie A, page 535

∂A La frontière de l'ensemble A, page 378

$A \Delta B$ différence symétrique, page 91

A^c complémentaire de A, page 90

Analyse

$\mathrm{Isom}(X)$ Le groupe des isométries de X, page 248

$\mu \ll \nu$ La mesure μ est absolument continue par rapport à la mesure ν., page 743

$C^\infty\big(\mathbb{R}, \mathscr{S}'(\mathbb{R}^d)\big)$ Fonctions à valeurs dans les distributions., page 1489

$C^\infty(I, \mathscr{D}'(\mathbb{R}^d))$ fonctions à valeurs dans les distributions, page 1417

$(S, \hat{\mathcal{F}}, \hat{\mu})$ complété de l'espace mesuré $(S, \hat{\mathcal{F}}, \hat{\mu})$, page 703

$\mathcal{L}(E, F)$ Les applications linéaires de E vers F, page 613

$\mathcal{L}^{(n)}(V, W)$ L'espace des applications n-linéaires $V^n \to W$, page 629

$\arg(z)$ La valeur principale de l'argument de $z \in \mathbb{C}$, page 1351

L Les applications linéaires continues de E vers F, page 613

$\mathbb{D}_b$ l'ensemble de écritures décimales en base b, page 476

$\mathbb{R}$ l'ensemble des réels, page 101

$\mathbb{R}^+$ les réels positifs ou nuls, page 103

exp exponentielle, page 916

$\mathscr{H}'$ dual, page 1255

$\liminf a_n$ limite inférieure, page 720

$\limsup a_n$ limite supérieure, page 720

$\mathcal{L}^p$ espace de Lebesgue, sans les classes, page 1278

$\mu \perp \nu$ mesures perpendiculaires, page 743

μ^* La mesure extérieure associée à la mesure μ, page 698

∂_z,$\partial_{\bar{z}}$ dérivées partielles d'une fonction complexe, page 1321

$\texttt{proj}_K(x)$ projection orthogonale de x sur y, page 1252

$\sigma(\mathcal{A})$ tribu engendrée par $\mathcal{D}$, page 686

$\mathscr{D}(\Omega)$ Les fonctions C^∞ à support compact sur Ω, page 1404

$\mathscr{S}'(\mathbb{R}^d)$ espace des distributions tempérées, page 1410

$A^2(\Omega)$ espace de Bergman, page 1363

$A^\perp$ orthogonal d'une partie., page 1254

$C_c(I)$ fonctions continues à support compact dans I, page 1293

$f \sim g$ fonctions ayant des limites équivalentes, page 548

$H^1(\Omega)$ espace de Sobolev sur Ω, page 1433

$H^1(I)$ espace de Sobolev, page 1427

$H^m(M)$ espace de Sobolev, page 1434

$L^1_{loc}(I)$ fonctions intégrables sur les compacts de I, page 1428

L^p espace de Lebesgue avec les classes, page 1279

$M_i\varphi$ La fonction $x \mapsto x_i\varphi(x)$, page 1315

$S_n f$ somme partielle de série de Fourier, page 1301

Chapitre 11

Espaces affines

Ce chapitre provient principalement de [22].

Définition 11.1.
Soit E, un espace vectoriel. Un ***espace affine modelé sur*** *E est un ensemble $\mathcal{E}$ sur lequel le groupe*[1] *$(E,+)$ agit à droite transitivement et librement.*

Étant donné que E est un groupe commutatif, l'action peut être vue indifféremment à gauche ou à droite. Si $M \in \mathcal{E}$ et si $x \in E$ nous notons $M + x$ au lieu de $x \cdot M$ le résultat de l'action de x sur M.

11.2.
Lorsque nous écrivons «$M + x$», le symbole plus n'est pas une loi de composition interne de $\mathcal{E}$, mais une action.

Soient $N, M \in \mathcal{E}$. Par liberté et transitivité de l'action, il existe un unique $x \in E$ tel que $M + x = N$. Ce vecteur x sera noté $\overrightarrow{MN}$.

Proposition 11.3.
Si $A, B, C \in \mathcal{E}$ nous avons les égalités suivantes dans E :

(1) $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$ *(relations de Chasles),*

(2) $\overrightarrow{AA} = 0$,

(3) $\overrightarrow{BA} = -\overrightarrow{AB}$.

11.4.
Si E est un espace vectoriel, le groupe $(E,+)$ agit sur E par l'action $t_y(x) = y + x$. Utilisant cette action nous construisons l'**espace affine canonique** de E. En particulier nous notons $\mathcal{E}_n(\mathbb{K})$ l'espace affine canonique de $\mathbb{K}^n$ vu comme espace vectoriel sur $\mathbb{K}$.

— En tant qu'ensembles, $\mathcal{E}_n(\mathbb{K}) = \mathbb{K}^n$.

— Sur cet espace en particulier, si $M, N \in \mathcal{E}_n(\mathbb{K})$, nous avons $\overrightarrow{MN} = N - M$ où à droite, la différence est la différence vectorielle dans $\mathbb{K}^n$.

Ces deux points se généralisent immédiatement à un espace vectoriel E au lieu de $\mathbb{K}^n$.

11.1 Repères affines

Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $\mathcal{E}$ un espace affine construit sur E.

Définition 11.5.
Un multiplet $(A, e_1, \ldots, e_n)$ où A est un point de $\mathcal{E}$ et $\{e_i\}$ est une base de E est un ***repère cartésien*** *de $\mathcal{E}$.*

1. Voir exemple 3.88.

Un tel repère donne une bijection

$$\begin{aligned} \phi\colon \mathbb{K}^n &\to \mathcal{E} \\ (x_1,\dots,x_n) &\mapsto A+\sum_i x_i e_i. \end{aligned} \tag{11.1}$$

Ces nombres x_i sont les **coordonnées** du point $A+\sum_i x_i e_i$ dans le repère (A,e_i).

Soient (A,e_i) et (A',e'_i) deux repères cartésiens pour l'espace affine $\mathcal{E}$. Soit (a_{ij}) la matrice de changement de base entre $\{e_i\}$ et $\{e'_i\}$ dans E. Nous voudrions trouver les x_i en termes des x'_i.

Pour cela nous considérons un point M dans $\mathcal{E}$ et nous l'écrivons dans les deux bases. Cela fournit l'égalité

$$A+\sum_i x_i e_i = A'+\sum_i x'_i e'_i. \tag{11.2}$$

Nous considérons les coordonnées (a_i) de A' dans le repère (A,e_i), c'est à dire

$$A' = A+\sum_i a_i e_i. \tag{11.3}$$

En substituant $e'_i=\sum_k a_{jk}e_k$ et (11.3) dans (11.2) nous trouvons

$$\sum_k x_k e_k = \sum_k a_k e_k + \sum_{jk} a_{jk}x'_j e_k, \tag{11.4}$$

et par conséquent

$$x_k = a_k + \sum_j a_{jk}x'_j. \tag{11.5}$$

11.2 Classification affine des conique

Soit une conique $f(x,y)=0$ avec

$$f(x,y)=ax^2+2bxy+cy^2+2dx+2ey+f \tag{11.6}$$

dans le repère $R=(A,e_i)$. La signature de la quadratique

$$q(x,y)=ax^2+2bx+cy^2 \tag{11.7}$$

ne dépend pas de la base choisie et un changement de variables

$$\begin{cases} \tilde{x}=\alpha x+\beta y & \text{(11.8a)}\\ \tilde{y}=\gamma x+\delta y & \text{(11.8b)} \end{cases}$$

peut nous amener dans trois cas :

$$q(x,y)=\begin{cases} \tilde{x}^2+\tilde{y}^2 & \text{genre ellipse}\\ \tilde{x}^2-\tilde{y}^2 & \text{genre hyperbole}\\ \tilde{x}^2 & \text{genre parabole.} \end{cases} \tag{11.9}$$

Dans le troisième cas, la matrice de q est de rang 1.

Nous cherchons maintenant à savoir si un point $I=(x_0,y_0)$ est un centre de symétrie de $f(x,y)=0$. Pour cela nous choisissons le repère centré en I, c'est à dire que nous posons

$$\begin{cases} x=x_0+\tilde{x} & \text{(11.10a)}\\ y=y_0+\tilde{y}. & \text{(11.10b)} \end{cases}$$

Un peu de calcul montre qu'alors la conique s'écrit

$$f(x_0,y_0)+q(\tilde{x},\tilde{y})+(2ax_0+2by_0+2d)\tilde{x}+(2bx_0+2cy_0+2e)\tilde{y}=0. \tag{11.11}$$

Le point I sera un centre de symétrique si les termes linéaires en $\tilde{x}$ et $\tilde{y}$ s'annulent, c'est à dire si

$$\begin{cases} ax_0 + by_0 + d = 0 & \text{(11.12a)} \\ bx_0 + cy_0 + e = 0. & \text{(11.12b)} \end{cases}$$

Nous supposons que $(d, e) \neq (0, 0)$, sinon la conique de départ serait déjà centrée. Le déterminant du système (11.12) est

$$\delta = ac - b^2. \tag{11.13}$$

Si ce dernier est différent de zéro, le système possède une unique solution et la conique aura alors un unique centre de symétrie.

Si le déterminant du système est nul, il y a soit pas de centre de symétrie, soit une infinité. Dans le premier cas nous sommes en présence d'une parabole, et dans le second cas de deux droites parallèles.

Exemple 11.6
Soit

$$f(x, y) = x^2 + 2xy - y^2 - 6x + 2y - 1 = 0 \tag{11.14}$$

donnée dans le repère affine $R = (A, \{e_i\})$. Nous commençons par étudier la signature de $q(x, y) = x^2 + 2xy - y^2$ dont la matrice symétrique est

$$Q = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{11.15}$$

Son polynôme caractéristique est $\lambda^2 - 2$ sont les racines sont $\pm\sqrt{2}$. La signature est donc $(1, -1)$ et nous sommes en présence d'une conique de genre hyperbole. Nous cherchons le centre en posant $x = \tilde{x} + x_0$, $y = \tilde{y} + y_0$. Le système à résoudre est

$$\begin{cases} x_0 + y_0 - 3 = 0 & \text{(11.16a)} \\ x_0 - y_0 + 1 = 0, & \text{(11.16b)} \end{cases}$$

dont l'unique solution est $(x_0, y_0) = (1, 2)$. Nous considérons le repère centré en (x_0, y_0), c'est à dire le repère

$$R' = (I, \{e_i\}) \tag{11.17}$$

avec $I = A + x_0 e_1 + y_0 e_2$ où A est l'origine du repère dans lequel l'équation (11.14) était donnée.

Par construction dans ce repère nous avons la conique

$$f(x_0, y_0) + q(\tilde{x}, \tilde{y}) = 0, \tag{11.18}$$

c'est à dire

$$\tilde{x}^2 + 2\tilde{x}\tilde{y} - \tilde{y}^2 = 0. \tag{11.19}$$

Maintenant la nous avons une quadrique centrée nous voulons la mettre sous une forme plus canonique :

$$\left(\frac{1}{\sqrt{2}}(\tilde{x} + \tilde{y})\right)^2 - \tilde{y}^2 - 1 = 0. \tag{11.20}$$

Nous posons donc

$$\begin{cases} X = \dfrac{1}{\sqrt{2}}(\tilde{x} + \tilde{y}) & \text{(11.21a)} \\ Y = \tilde{y}, & \text{(11.21b)} \end{cases}$$

et nous trouvons l'hyperbole

$$X^2 - Y^2 - 1 = 0. \tag{11.22}$$

Cela revient à faire le changement de base

$$\begin{cases} e'_1 = \sqrt{2}e_1 & \text{(11.23a)} \\ e'_2 = -e_1 + e_2. & \text{(11.23b)} \end{cases}$$

Pour rappel, les vecteurs de bases se transforment avec la matrice inverse des coefficients. Étant donné que

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix}, \tag{11.24}$$

nous avons

$$\begin{pmatrix} e'_1 \\ e'_2 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}. \tag{11.25}$$

C'est de là que provient le changement (11.23). △

11.3 Applications affines

Définition 11.7.
*Soient $\mathcal{E}$ et $\mathcal{E}'$ deux espaces affines sur les espaces vectoriels E et E' (sur le même corps $\mathbb{K}$). Une application $f\colon \mathcal{E} \to \mathcal{E}'$ est dite **affine** si pour tout $M \in \mathcal{E}$, il existe une application linéaire $u_M\colon E \to E'$ telle que*

$$f(M + x) = f(M) + u(x) \tag{11.26}$$

pour tout $x \in E$.

Lemme 11.8.
Si f est affine, l'application linéaire u_M ne dépend pas de M et peut donc être notée u_f sans préciser le point.

Démonstration. Posons

$$f(M + x) = f(M) + u_M(x) \tag{11.27a}$$
$$f(N + y) = f(N) + u_N(y). \tag{11.27b}$$

Définissons $a \in E$ par $N = M + a$; nous avons d'une part que

$$f(N + y) = f(M + y + a) = f(M) + u_M(y + a), \tag{11.28}$$

et d'autre part

$$f(N + y) = f(M + a) + u_N(y) = f(M) + u_M(a) + u_N(y). \tag{11.29}$$

Par conséquent $u_M(y + a) = u_M(a) + u_N(y)$. Par linéarité $u_N = u_M$. □

Ce lemme est important car il permet de démonter qu'une application est affine en prouvant la linéarité des u_M séparément sans devoir prouver qu'elles sont égales.

Lemme 11.9 ([1])**.**
Soit $M \in \mathcal{E}$ et deux points $A, B \in \mathcal{E}$ donnés par $A = M + x_a$, $B = M + x_b$. Soit encore une application affine f sur $\mathcal{E}$. Alors

$$\overrightarrow{AB} = u_f(x_b - x_a). \tag{11.30}$$

Démonstration. En appliquant f à $A = M + x_a$ et $B = M + x_b$,

$$f(A) = f(M) + u_f(x_a) \tag{11.31a}$$
$$f(B) = f(M) + u_f(x_b). \tag{11.31b}$$

Donc $f(B) = f(A) - u_f(x_a) + u_f(x_b)$ ou encore

$$f(B) = f(A) + u_f(x_b - x_a). \tag{11.32}$$

□

Remarque 11.10.
La condition (11.26) pour tout $M \in \mathcal{E}$ est équivalente à demander

$$f \circ t_x = t_{u(x)} \circ f \tag{11.33}$$

pour tout $x \in E$.

Proposition 11.11.
Soit f une application affine.

(1) Il existe une unique application linéaire u_f telle que $f(M + x) = f(M) + u_f(x)$ pour tout $M \in \mathcal{E}$ et tout $x \in E$.

(2) L'application u_f est injective si et seulement si f est injective.

(3) L'application u_f est surjective si et seulement si f est surjective.

Si de plus les espaces $\mathcal{E}$ et $\mathcal{E}'$ ont même dimension finie, alors f est injective si et seulement si f est surjective.

11.12.
En posant $M = 0$ dans la condition (11.26), si f est une application affine il existe une application linéaire α et un vecteur v tel que $f = \tau_v \circ \alpha$.

Proposition 11.13.
Si $f\colon \mathcal{E} \to \mathcal{E}'$ et $g\colon \mathcal{E} \to \mathcal{E}''$ sont des applications affines, alors $g \circ f\colon \mathcal{E} \to \mathcal{E}''$ est affine et $u_{g\circ f} = u_g \circ u_f$.

Démonstration. Si $M \in \mathcal{E}$ et $x \in E$ nous avons

$$\begin{aligned}(g \circ f)(M + x) &= g\big(f(M) + u_f(x)\big)\\ &= f\big(f(M)\big) + u_g\big(u_f(x)\big)\\ &= (g \circ f)(M) + (u_g \circ u_f)(x).\end{aligned} \tag{11.34}$$

□

Théorème 11.14.
Soient $\mathcal{E}$ et $\mathcal{E}'$ deux espaces affines de dimensions finies p et q sur $\mathbb{K}$. Soient les repères cartésiens $R = (O, \{e_i\})$ et $R' = (O', \{e'_i\})$. Une application $f\colon \mathcal{E} \to \mathcal{E}'$ est affine si et seulement s'il existe une matrice $a \in \mathrm{M}_{p,q}(\mathbb{K})$ et $b \in \mathbb{K}^q$ tels que

$$f(x) = b + ax. \tag{11.35}$$

Remarque 11.15.
L'équation (11.35) est écrite en utilisant un abus de notation entre le vecteur $x \in \mathbb{K}^p$ et le point de $\mathcal{E}$ qui est représenté par x dans le repère $(A, \{e_i\})$.

11.4 Isomorphismes

Définition 11.16.
*Un **isomorphisme** entre les espaces affines $\mathcal{E}$ $\mathcal{E}'$ est une application affine $f\colon \mathcal{E} \to \mathcal{E}'$ inversible dont l'inverse est affine.*

Proposition 11.17.
Une application affine bijective est un isomorphisme. Si f est un isomorphisme d'espaces affines, alors $u_{f^{-1}} = (u_f)^{-1}$.

Proposition 11.18.
Un espace affine de dimension finie n sur un corps $\mathbb{K}$ est isomorphe à l'espace affine canonique $\mathcal{E}_n(\mathbb{K})$.

Démonstration. Si nous considérons le repère $R = (A, \{e_i\})$ de l'espace affine $\mathcal{E}$ alors l'application

$$\begin{aligned} \varphi\colon \mathbb{K}^n &\to \mathcal{E} \\ (x_1, \ldots, x_n) &\mapsto A + \sum_i x_i e_i \end{aligned} \tag{11.36}$$

est un isomorphisme. □

11.5 Sous espaces affines

Définition 11.19.
Soit $\mathcal{E}$ un espace affine sur l'espace vectoriel E. Un ***sous-espace affine*** *de $\mathcal{E}$ est une orbite de l'action d'un sous-espace vectoriel de E.*

Si $\mathcal{F}$ est un sous ensemble de $\mathcal{E}$, il sera un sous-espace affine de $\mathcal{E}$ si et seulement si l'ensemble

$$F = \{AB \text{ tel que } A, B \in \mathcal{F}\} \tag{11.37}$$

est un sous-espace vectoriel de E. Dans ce cas nous disons que F est la **direction** de $\mathcal{F}$. Si $A \in \mathcal{F}$, alors l'orbite de A sous F est $\mathcal{F}$. La **dimension** de $\mathcal{F}$ est la dimension de sa direction.

Si $\mathcal{F}$ et $\mathcal{G}$ sont des sous-espaces affines de $\mathcal{E}$ de directions F et G, nous disons que $\mathcal{F}$ est **parallèle** à $\mathcal{G}$ si $F \subset G$.

Proposition 11.20.
Soit $\mathcal{F}$ un sous-espace affine de dimension k dans l'espace affine $\mathcal{E}$ de dimension n. Alors il existe une application affine $f\colon \mathcal{E} \to \mathbb{K}^{n-k}$ telle que $\mathcal{F} = f^{-1}(0)$.

Démonstration. Soit F la direction de $\mathcal{F}$ et $A \in \mathcal{F}$. Nous considérons une base $\{e_i\}$ adaptée à F au sens $\{e_1, \ldots, e_k\}$ est une base de F. Nous considérons maintenant le repère cartésien $(A, \{e_i\})$ avec $A \in \mathcal{F}$ et nous construisons l'application affine

$$\begin{aligned} f\colon \mathcal{E} &\to \mathbb{K}^{n-k} \\ A + \sum_{i=1}^{n} x_i e_i &\mapsto \begin{pmatrix} x_{k+1} \\ \vdots \\ x_n \end{pmatrix}. \end{aligned} \tag{11.38}$$

Par construction nous avons $f(M) = 0$ si et seulement si $M \in \mathcal{F}$. □

Proposition 11.21 ([22])**.**
Soit σ une partie de l'espace affine $\mathcal{E}$.

(1) L'intersection de tous les sous-espaces affines contenant σ est un sous-espace affine, noté $\mathcal{F}$.

(2) Si $A \in \sigma$, alors la direction de $\mathcal{F}$ est le sous-espace vectoriel

$$F = \operatorname{Span}\{\overrightarrow{AM} \text{ tel que } M \in \sigma\}. \tag{11.39}$$

Le sous-espace affine donné par la proposition 11.21 est le sous-espace affine **engendré** par la partie σ, et il est noté $\operatorname{eae}(\sigma)$.

Proposition 11.22.
Soit $\mathcal{E}$ un espace affine de dimension n sur $\mathbb{K}$, soit $f\colon \mathcal{E} \to \mathbb{K}^r$ une fonction affine. Pour tout $a = (a_1, \ldots, a_r) \in \mathbb{K}^r$, l'ensemble $f^{-1}(a)$ est un sous-espace affine de dimension $\dim \ker(u_f)$.

Démonstration. Nous considérons le repère $(A, \{e_i\})$ de $\mathcal{E}$. Étant donné que f est affine nous avons

$$f\big(A + \sum_i x_i e_i\big) = f(A) + u_f\big(\sum_i x_i e_i\big). \tag{11.40}$$

Nous avons donc $f\big(A+\sum_i x_i e_i\big)=a$ lorsque

$$u_f(\sum_i x_i e_i)=a-f(A). \tag{11.41}$$

Nous avons donc

$$f^{-1}(a)=A+(u_f)^{-1}\big(a-f(A)\big), \tag{11.42}$$

dont la dimension est le rang de $(u_f)^{-1}=u_{f^{-1}}$ (proposition 11.17). Le rang de $(u_f)^{-1}$ est le dimension du noyau de u_f. □

Proposition 11.23.
Soit A un ensemble convexe dans un espace vectoriel et $v_1,\dots,v_n$ des éléments de A. Alors toute combinaison

$$a_1v_1+\cdots+a_nv_n \tag{11.43}$$

telle que $a_1+\cdots+a_n=1$ et $a_i\in[0,1]$ appartient à A.

Démonstration. Nous prouvons la proposition pour $n=3$. Nous devons trouver des nombres $t_1,t_2\in[0,1]$ tels que

$$t_2\big(t_1v_1+(1-t_1)v_2\big)+(1-t_2)v_3=av_1+bv_2+cv_3. \tag{11.44}$$

La réponse est immédiatement donnée par

$$t_2a=1-c \tag{11.45a}$$
$$t_1=a/t_2. \tag{11.45b}$$

Étant donné que $c\in[0,1]$ nous avons $t_2\in[0,1]$. En ce qui concerne t_1 nous avons

$$t_1=\frac{a}{t_2}\leqslant\frac{1-c}{1-c}=1. \tag{11.46}$$

□

11.6 Barycentre

Soit $\mathcal{E}$ un espace affine sur le $\mathbb{K}$-espace vectoriel E. Un couple (A,λ) avec $A\in\mathcal{E}$ et $\lambda\in\mathbb{K}$ est un **point pondéré**.

Lemme-définition 11.24 ([134])**.**
Soit une famille de points pondérés $\{(A_i,\lambda_i)\}_{i=1\dots r}$. Si $\sum_i\lambda_i\neq0$, alors il existe un unique $G\in\mathcal{E}$ tel que

$$\sum_{i=1}^{r}\lambda_i\overrightarrow{GA_i}=0. \tag{11.47}$$

*Le point G donné par le lemme 11.24 est le **barycentre** des points pondérés (A_i,λ_i).*

Notons que l'on peut toujours supposer que $\sum_i\lambda_i=1$ parce que le barycentre ne change pas lorsque tous les λ_i sont multipliés par un même nombre.

Définition 11.25 (Combinaison convexe)**.**
*Des nombres $\lambda_1,\dots,\lambda_n$ vérifiant $\sum_i\lambda_i=1$ forment une **combinaison convexe**.*

Le théorème suivant donné quelques caractérisations équivalentes du barycentre.

Théorème 11.26 ([134])**.**
Soient $\{(A_i,\lambda_i)\}_{i=1,\dots,r}$ une famille de points pondérés. Les conditions suivantes sur le point $G\in\mathcal{E}$ sont équivalentes.

(1) Le point G est barycentre de la famille.
(2) Pour tout $\alpha \in \mathbb{R}^*$, $\sum_i (\alpha\lambda_i)\overrightarrow{GA_i} = 0$.
(3) Il existe $A \in \mathcal{E}$ tel que $\big(\sum_i \lambda_i\big)\overrightarrow{AG} = \sum_i \lambda_i \overrightarrow{AA_i}$.
(4) Pour tout $B \in \mathcal{E}$, nous avons $\big(\sum_i \lambda_i\big)\overrightarrow{BG} = \sum_i \lambda_i \overrightarrow{BA_i}$.

Définition 11.27.
*Si $A, B \in \mathcal{E}$, le **segment** $[AB]$ est l'ensemble des barycentres de A et B pondérés par des poids positifs (ouvert ou fermé suivant que l'on accepte que l'un ou l'autre des poids soit nul).*

Lorsque tous les λ_i sont égaux, nous parlons d'**isobarycentre**. Autrement dit, l'isobarycentre des points A_i est le barycentre des points pondérés $(A_i, 1)$.

11.6.1 Sous-espaces affines

Proposition 11.28.
Une partie $\mathcal{F}$ des $\mathcal{E}$ est un sous-espace affine si et seulement si elle est stable par barycentrisation.

Démonstration. Soit $\mathcal{F}$ une sous-espace affine de direction F et $A_1, \dots, A_n$ des points de $\mathcal{F}$. Nous devons voir que le barycentre des points A_i pondérés de n'importe quelles masses appartient à $\mathcal{F}$. Pour ce faire nous faisons appel à la caractérisation (4) du théorème 11.26 : pour tout $B \in \mathcal{F}$,

$$\overrightarrow{BG} = \sum_i \lambda_i \overrightarrow{BA_i}. \tag{11.48}$$

Vu que B et A_i sont dans $\mathcal{F}$, nous avons $\overrightarrow{BA_i} \in F$ et donc $\overrightarrow{BG} \in F$. Mais comme $B \in \mathcal{F}$, le point G est à son tour dans $\mathcal{F}$.

Réciproquement, nous supposons que $\mathcal{F}$ est stable par barycentrisme. Nous voudrions montrer que l'ensemble

$$F = \{\overrightarrow{AB} \text{ tel que } A, B \in \mathcal{F}\} \tag{11.49}$$

est un sous-espace vectoriel. Soit $A \in \mathcal{F}$. Nous commençons par prouver que les vecteurs de la forme $\overrightarrow{AX}$ ($X \in \mathcal{F}$) forment un espace vectoriel. Considérons $\overrightarrow{AX} + \overrightarrow{AY}$ qui est un élément de E ; il existe donc $V \in \mathcal{E}$ tel que

$$\overrightarrow{AV} = \overrightarrow{AX} + \overrightarrow{AY}. \tag{11.50}$$

Par les relations de Chasles,

$$\overrightarrow{AV} = \overrightarrow{AV} + \overrightarrow{VX} + \overrightarrow{AV} + \overrightarrow{VY}, \tag{11.51}$$

donc

$$0 = \overrightarrow{VX} - \overrightarrow{VA} + \overrightarrow{VY}, \tag{11.52}$$

ce qui prouve que V est un barycentre de X, A, Y, et donc que $V \in \mathcal{F}$. De la même manière si $W \in \mathcal{E}$ est défini par $\overrightarrow{AW} = \mu\overrightarrow{AX}$, alors

$$\overrightarrow{AW} = \mu\overrightarrow{AX} = \mu(\overrightarrow{AW} + \overrightarrow{WX}), \tag{11.53}$$

ce qui signifie que

$$(1 - \mu)\overrightarrow{AW} + \mu\overrightarrow{XW} = 0 \tag{11.54}$$

et que W est un barycentre.

Afin de montrer que (11.49) est bien un espace vectoriel, nous devons considérer $A, B, X, Y \in \mathcal{F}$ et prouver que $\overrightarrow{AX} + \overrightarrow{BY} \in F$. Nous avons

$$\begin{aligned}
\overrightarrow{AX} + \overrightarrow{BY} &= \overrightarrow{AX} + \overrightarrow{BA} + \overrightarrow{AY} && && \text{(11.55a)}\\
&= \overrightarrow{AV} + \overrightarrow{BA} && V \text{ est celui donné plus haut} && \text{(11.55b)}\\
&= \overrightarrow{AV} - \overrightarrow{AB} && && \text{(11.55c)}\\
&= \overrightarrow{AV} + \overrightarrow{AW} && W \text{ est donné par } \mu = -1. && \text{(11.55d)}\\
&= \overrightarrow{AV'}. && && \text{(11.55e)}
\end{aligned}$$

□

Proposition 11.29 ([134])**.**
Soient $A_0, \dots, A_r$ des points de $\mathcal{E}$. L'ensemble des barycentres de ces points (avec des masses de somme 1) est le sous-espace affine engendré par les A_i que nous nommons $\mathcal{F}$.

Démonstration. Soit G le barycentre associé aux poids λ_i. Nous avons

$$G = A_0 + \overrightarrow{A_0G} = A_0 + \sum_{i=1}^{r} \lambda_i \overrightarrow{A_0A_i}. \tag{11.56}$$

Notons que les vecteurs $\overrightarrow{A_0A_i}$ sont dans la direction du sous-espace affine engendré par les A_i par (11.39). Donc G est bien dans $\mathcal{F}$.

Inversement si X est dans $\mathcal{F}$, on a

$$X = A_0 + \sum_i \lambda_i \overrightarrow{A_0A_i} \tag{11.57}$$

parce que $\sum_i \lambda_i \overrightarrow{A_0A_i}$ est un élément général de la direction de $\mathcal{F}$. Du coup

$$\overrightarrow{A_0X} = \sum_i \lambda_i \overrightarrow{A_0A_i}, \tag{11.58}$$

et en utilisant la relation de Chasles sur chacun des $\overrightarrow{A_0A_i}$,

$$\overrightarrow{A_0X} = \sum_i \lambda_i \big(\overrightarrow{A_0X} + \overrightarrow{XA_i}\big). \tag{11.59}$$

De là nous concluons que

$$\big(1 - \sum_i \lambda_i\big)\overrightarrow{A_0X} + \sum_i \lambda_i \overrightarrow{A_iX} = 0, \tag{11.60}$$

ce qui signifie précisément que X est un barycentre des A_i. □

Proposition 11.30.
Soient $r+1$ point $A_0, \dots, A_r$ dans $\mathcal{E}$. Le sous-espace affine engendré par les A_i est au plus de dimension r.

Démonstration. La direction de l'espace engendré $\mathrm{Aff}\{A_i\}$ est l'espace

$$\mathrm{Span}\{\overrightarrow{A_0A_i}_{i=1,\dots,r}\} \tag{11.61}$$

qui est engendré par r vecteurs et donc est au plus de dimension r. □

En deux mots, la proposition suivante signifie que le barycentre des barycentres est le barycentre.

Proposition 11.31 (Associativité des barycentres)**.**
Soit $I = \{0, 1, \dots, n\}$ et une partition $I = J_0 \cup \dots \cup J_r$. Soient des points $a_0, \dots, a_n \in \mathcal{E}$ et $\lambda_0, \dots, \lambda_m$ des nombres tels que $\sum_i \lambda_i \neq 0$. Nous supposons que $\mu_k = \sum_{i \in J_k} \lambda_i \neq 0$ pour tout k, et enfin nous nommons b_k le barycentre de la famille $\{(a_i, \lambda_i), i \in J_k\}$.

Alors le barycentre de la famille $\{(b_k, \mu_k)\}_{k=1,\dots,n}$ est le barycentre de la famille $\{(a_i, \lambda_i)\}_{i \in I}$.

Démonstration. Nous nommons b le barycentre des b_k pondérés par les μ_k, donc par définition

$$0 = \sum_{k=0}^{r} \mu_k \overrightarrow{bb_k} \tag{11.62a}$$
$$= \sum_{k} \sum_{i\in J_k} \lambda_i \overrightarrow{bb_k} \tag{11.62b}$$
$$= \sum_{k=0}^{r} \sum_{i\in J_k} \lambda_i (\overrightarrow{ba_i} + \overrightarrow{a_ib_k}) \tag{11.62c}$$
$$= \sum_{k=0}^{r} \sum_{i\in J_k} \lambda_i \overrightarrow{ba_i} + \sum_{k=0}^{r} \underbrace{\sum_{i\in J_k} \lambda_i \overrightarrow{a_ib_k}}_{=0} \tag{11.62d}$$
$$= \sum_{i\in I} \lambda_i \overrightarrow{ba_i}. \tag{11.62e}$$

Donc b est bien barycentre des a_i avec les poids λ_i. □

11.6.2 Enveloppe convexe

Définition 11.32.
*Soit A une partie d'un espace vectoriel E. L'**enveloppe convexe** de A, notée $\mathcal{C}\mathrm{onv}(A)$ est l'intersection de tous les convexes contenant A.*

L'enveloppe convexe est un convexe. En effet soit C un convexe contenant A et $x, y \in \mathcal{C}\mathrm{onv}(A)$; alors x et y sont dans C et par conséquent le segment $[x, y]$ est inclus à C. Ce segment étant inclus à tout convexe contenant A, il est inclus à $\mathcal{C}\mathrm{onv}(A)$.

Proposition 11.33.
Soit C un convexe dans l'espace affine $\mathcal{E}$ et une famille de points pondérés $\{(a_i, \lambda_i)\}_{i=1,\dots,r}$ dont tous les poids sont positifs (et non tous nuls). Alors le barycentre est aussi dans C.

En d'autre termes, un convexe est stable par barycentrage à poids positifs[2]*.*

Démonstration. Nous prouvons par récurrence. D'abord pour $r = 2$. Le barycentre des points pondérés (a_1, λ_1), (a_2, λ_2) est le point b tel que

$$\lambda_1 \overrightarrow{ba_1} + \lambda_2 \overrightarrow{ba_2} = 0. \tag{11.63}$$

Par définition, ce qui est noté $\overrightarrow{ab}$ n'est rien d'autre que $b - a$; en déballant (11.63), nous trouvons

$$\lambda_1(a_1 - b) + \lambda_2(a_2 - b) = 0 \tag{11.64}$$

et donc

$$b = \frac{\lambda_1}{\lambda_1 + \lambda_2} a_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} a_2, \tag{11.65}$$

qui est bien un point du segment $[a_1, a_2]$ parce que c'est une combinaison à coefficients positifs de somme 1.

Nous passons maintenant à la vraie récurrence avec un ensemble de points pondérés

$$A_r = \{(a_1, \lambda_1), \dots, (a_r, \lambda_r)\} \tag{11.66}$$

de masse totale non nulle ; et en vous laissant deviner ce que va désigner A_{r-1}. Si une des masses est nulle (disons λ_r), alors le barycentre de A_r est le même que celui de A_{r-1} et l'hypothèse de récurrence nous enseigne que ledit barycentre est dans C. Nous supposons donc que $\lambda_i \neq 0$ pour tout i. Dans ce cas le théorème d'associativité des barycentres 11.31 dit que le barycentre de A_r est le barycentre entre le barycentre de A_{r-1} et (a_r, λ_r), qui sont deux points de C par hypothèse de récurrence. □

2. Sauf si on prend tous les poids nuls ; mais contre ce genre d'idées, on ne peut rien faire.

Si E est un espace vectoriel et si $x_i \in E$ et $\lambda_i \in \mathbb{R}$, alors le barycentre des couples (x_i, λ_i) est le point g tel que $\sum_i \lambda_i \overrightarrow{gx_i}$, c'est à dire $\sum_i \lambda_i(x_i - g) = 0$ ou encore

$$\sum_i \lambda_i x_i = \sum_i \lambda_i g. \tag{11.67}$$

Donc quitte à diviser tous les λ_i par la somme, nous pouvons supposer que la somme des poids est 1. C'est pourquoi lorsque nous parlerons de barycentre dans un espace vectoriel sans contexte affin, nous allons toujours supposer $\sum_i \lambda_i = 0$ et avoir le barycentre

$$g = \sum_i \lambda_i x_i. \tag{11.68}$$

Proposition 11.34.
Soit E, un espace vectoriel et $A \subset E$. L'enveloppe convexe $\mathcal{C}\text{onv}(A)$ est l'ensemble des barycentres de familles finies de points affublés de masses positives.

Démonstration. Nous notons $\mathcal{B}$ l'ensemble des dits barycentres. Par la proposition 11.33, ces barycentres sont dans l'enveloppe convexe et donc $\mathcal{B} \subset \mathcal{C}\text{onv}(A)$. A contrario, si nous prouvons que $\mathcal{B}$ était convexe, alors nous aurions $\mathcal{C}\text{onv}(A) \subset \mathcal{B}$ parce que l'enveloppe convexe est l'intersection des convexes contenant A.

Soient $a, b \in \mathcal{B}$, c'est à dire que l'on a $a_0, \dots, a_n$ et $b_0, \dots, b_m$ dans A ainsi que les nombres strictement positifs $\lambda_0, \dots, \lambda_n$ et $\mu_0, \dots, \mu_m$ tels que

$$a = \sum_i \lambda_i a_i \qquad \sum_{i=1}^{n} \lambda_i = 1 \tag{11.69a}$$

$$b = \sum_j \mu_j b_j \qquad \sum_{j=1}^{n} \mu_j = 1 \tag{11.69b}$$

Un point du segment $[a, b]$ est de la forme $p = ta + (1-t)b$ avec $t \in [0, 1]$. En développant,

$$p = \sum_{i=0}^{n} (t\lambda_i)a_i + \sum_{j=0}^{m} (1-t)\mu_j b_j. \tag{11.70}$$

Cela est le barycentre de la famille $\{(a_i, \lambda_i), (b_j, \mu_j)\}$, parce que la somme des coefficients est bien 1 :

$$\sum_i (t\lambda_i) + \sum_j (1-t)\mu_j = t + (1-t) = 1. \tag{11.71}$$

□

Théorème 11.35 (Carathéodory[2])**.**
Dans un espace affine de dimension n, l'enveloppe convexe[3] *de A est l'ensemble des barycentres à coefficients positifs ou nuls de familles de $n+1$ points.*

Démonstration. Soit $x \in \mathcal{C}\text{onv}(A)$; on sait par la proposition 11.34 que x est barycentre de points de A avec des coefficients positifs :

$$x = \sum_{k=1}^{p} \lambda_k x_k \tag{11.72}$$

avec $\sum_k \lambda_k = 1$. Nous supposons que $p > n+1$ (sinon le théorème est réglé), et nous allons faire une récurrence à l'envers en montrant qu'on peut aussi écrire x sous forme d'un barycentre de strictement moins de p points.

3. Définition 11.32.

Étant donné que $p-1>n$, la famille $\{x+i-x_1\}_{i=2,\dots,p}$ est liée et il existe donc $\alpha_1,\dots,\alpha_p\in\mathbb{R}$ tels que $\sum_{i=2}^{p}\alpha_i(x_i-x_1)=0$, c'est à dire telle que

$$\sum_{i=2}^{p}\alpha_i x_i=\sum_{i=2}^{p}\alpha_i x_1. \tag{11.73}$$

Nous posons $\alpha_1=-\sum_{i=2}^{p}\alpha_1$. Remarquons qu'alors $\sum_{i=1}^{p}\alpha_i x_i=0$ parce que

$$\sum_{i=1}^{p}\alpha_i x_i=\alpha_1 x_1+\sum_{i=2}^{p}\alpha_i x_i=\alpha_1 x_1+\sum_{i=2}^{p}\alpha_i x_1=\sum_{i=1}^{p}\alpha_i x_1=0. \tag{11.74}$$

Par conséquent ça ne coûte rien de récrire (11.72) sous la forme

$$x=\sum_{i=1}^{p}(\lambda_i+t\alpha_i)x_i. \tag{11.75}$$

Les α_i ne sont pas tous nuls, mais leur somme est nulle, donc il y en a au moins un négatif. Nous notons

$$\tau=\min\{-\frac{\lambda_i}{\alpha_i}\text{ tel que }\alpha_i<0\}, \tag{11.76}$$

et J l'ensemble de i pour lesquels ce minimum est atteint. Nous considérons aussi le nombres $\mu_i=\lambda_i+\tau\alpha_i$. Plusieurs remarques.

(1) Si $j\in J$, alors $\mu_j=0$

(2) Si $\alpha_i>0$ alors $\mu_i\geqslant 0$, mais si $\alpha_i<0$ alors

$$\lambda_i+\tau\alpha_i\geqslant\lambda_i+(-\frac{\lambda_i}{\alpha_i})\alpha_i=0m \tag{11.77}$$

donc $\mu_i\geqslant 0$ quand même.

(3) $\sum_{i=1}^{p}\mu_i=1$, toujours parce que $\sum_{i=1}^{p}\alpha_i=0$.

Avec tout ça, nous avons

$$\sum_{i\notin J}\mu_i x_i=\sum_{i=1}^{p}\mu_i x_i=x. \tag{11.78}$$

Et voila, nous avons écrit x comme un barycentre à coefficients positifs de moins de p éléments parce que J n'est pas vide. □

Corollaire 11.36.
Dans un espace affine de dimension finie, l'enveloppe convexe d'un compact est compacte.

Démonstration. Soit A une partie compacte de l'espace vectoriel E, et $\mathcal{C}\text{onv}(A)$ son enveloppe convexe. Nous allons montrer que toute suite dans $\mathcal{C}\text{onv}(A)$ admet une sous-suite convergente en écrivant un point de $\mathcal{C}\text{onv}(A)$ comme le théorème de Carathéodory 11.35 nous le suggère. Pour cela nous considérons le simplexe

$$\Lambda=\left\{\lambda\in\mathbb{R}^{n+1}\text{ tel que }\sum_{k=1}^{n+1}\lambda_k=1\text{ et }\lambda_k\geqslant 0\forall k\right\}. \tag{11.79}$$

Montrons en passant que Λ est compact. Si $\lambda_k\in\Lambda$ est une suite, alors chacun des λ_k est un $(n+1)$-uple de nombres dans $[0,1]$:

$$k\mapsto(\lambda_k)_i \tag{11.80}$$

est une suite qui possède une sous-suite convergente. En passant $n+1$ fois à une sous-suite, nous tombons sur une suite convergente vers $\lambda\in\Lambda$, grâce à la convergence composante par composante. De plus pour chaque k nous avons $\sum_{i=1}^{n+1}(\lambda_k)_i=1$, et en passant à la limite, la somme étant une application continue, $\sum_i\lambda_i=1$.

Considérons l'application

$$\begin{aligned} f\colon \Lambda \times A^{n+1} &\to \mathcal{C}\text{onv}(A) \\ (\lambda, x) &\mapsto \sum_{k=1}^{n+1} \lambda_k x_k. \end{aligned} \tag{11.81}$$

C'est une application continue parce qu'elle est bilinéaire en dimension finie ; son image est contenue dans $\mathcal{C}\text{onv}(A)$ par la proposition 11.33, et elle est surjective par le théorème de Carathéodory 11.35. Bref, $\mathcal{C}\text{onv}(A) = f(\Lambda \times A^{n+1})$ est donc l'image d'un compact par une application continue ; elle est donc compacte par le théorème 6.77. □

Notons que sans le théorème de Carathéodory, peut être que le nombre de points utiles pour décomposer les différents a_k n'était pas borné ; dans ce cas nous aurions du prendre une infinité de sous-suites et rien n'aurait été sûr.

11.6.3 Applications affines et barycentre

Proposition 11.37 ([135])**.**
Une application $f\colon \mathcal{E} \to \mathcal{E}'$ entre deux espaces affines est affine si et seulement si pour tout système $\{(A_i, \lambda_i)\}_{i=1,\dots,k}$ de barycentre G et de poids total non nul, le point $f(G)$ est barycentre du système $\{\big(f(A_i), \lambda_i\big)\}$.

Démonstration. En deux parties.

Si f est affine Par définition d'un barycentre,

$$\sum_i \lambda_i \overrightarrow{GA_i} = 0. \tag{11.82}$$

Nous considérons un point arbitraire $O \in \mathcal{E}$ et nous écrivons $A_i = O + x_i$, $G = O + x_g$. Ensuite nous utilisons le lemme 11.9 pour le calcul suivant :

$$\begin{aligned} \sum_i \lambda_i \overrightarrow{g(G)f(A_i)} &= \sum_i \lambda_i u_f(x_i - x_g) && \text{(11.83a)} \\ &= u_f\big(\sum_i \lambda_i (x_i - x_g)\big) && \text{(11.83b)} \\ &= u_f\big(\sum_i \lambda_i \overrightarrow{GA_i}\big) && \text{(11.83c)} \\ &= u_f(0) = 0. && \text{(11.83d)} \end{aligned}$$

Donc $f(G)$ est bien le barycentre du nouveau système.

Si f conserve les barycentres Nous définissons u par $f(O + x) = f(O) + u(x)$. A priori, ce u dépend de O et n'est pas linéaire.

u est linéaire Soient $M, N \in \mathcal{E}$ et les éléments $x_m, x_n \in E$ tels que $\overrightarrow{OM} = x_m$ et $\overrightarrow{ON} = x_n$. Nous définissons enfin P par

$$\overrightarrow{OP} = \alpha \overrightarrow{OM} + \beta \overrightarrow{ON}, \tag{11.84}$$

et $P = O + x_p$. En décomposant $\overrightarrow{MO}$ et $\overrightarrow{NO}$ par les relations de Chasles de la proposition 11.3(1) nous avons

$$(\alpha + \beta - 1)\overrightarrow{PO} - \alpha \overrightarrow{PM} - \beta \overrightarrow{PN} \tag{11.85}$$

et donc P est barycentre du système

$$\big\{(\alpha + \beta - 1, O), (\alpha, M), (\beta, N)\big\}. \tag{11.86}$$

Le point $f(P)$ sera barycentre du système

$$\Big\{\big(\alpha + \beta - A, f(O)\big), \big(\alpha, f(M)\big), \big(\beta, f(N)\big)\Big\}. \tag{11.87}$$

Cela signifie que

$$(\alpha+\beta-1)\overrightarrow{f(P)f(O)}-\alpha\overrightarrow{f(P)f(M)}-\beta\overrightarrow{f(P)f(N)}=0. \tag{11.88}$$

En y substituant $\overrightarrow{f(P)f(O)}=u(-x_p)$, $\overrightarrow{f(P)f(M)}=u(x_m-x_p)$ et $\overrightarrow{f(P)f(N)}=u(x_mm-x_p)$ ainsi que $x_p=\alpha x_m+\beta x_b$ nous trouvons

$$u(\alpha x_m+\beta x_n)=\alpha u(x_m)+\beta u(x_n). \tag{11.89}$$

Donc u est linéaire.

u **ne dépend pas du point** O Il n'est pas besoin de démonter cela parce que la définition 11.7 ne le demande pas. Note : c'est le lemme 11.8 qui dit que c'est par ailleurs vrai.

□

11.7 Repères, coordonnées cartésiennes et barycentriques

Définition 11.38.
*On dit que les points $A_0,\dots,A_r\in\mathcal{E}$ sont **affinement indépendants** si le sous-espace affine engendré est de dimension r.*

Proposition 11.39 ([134])**.**
Pour $r+1$ points $A_0,\dots,A_r$ dans $\mathcal{E}$, les propriétés suivantes sont équivalentes.

1. *Les A_i sont affinement indépendants.*
2. *Pour tout $i=0,\dots,r$, le point A_i n'est pas dans $\mathrm{Aff}\{A_0,\dots,\hat{A}_i,\dots,A_r\}$.*
3. *Les points $A_0,\dots,A_{r-1}$ sont affinement indépendants et $A_r\notin\mathrm{Aff}\{A_0,\dots,A_{r+1}\}$.*
4. *Il existe i tel que les vecteurs $\overrightarrow{A_kA_i}$ $(k\in i)$ sont linéairement indépendants.*
5. *Pour tout $i\in\{1,\dots,r\}$, les vecteurs $\overrightarrow{A_kA_i}$ $(k\neq i)$ sont linéairement indépendants.*

Notons à propos de la condition (3) que l'existence d'un i tel que $A_i\notin\mathrm{Aff}\{A_0,\dots,\hat{A}_i,\dots,A_r\}$ n'implique pas l'indépendance des $r+1$ points. En effet dans $\mathbb{R}^2$ nous considérons les 4 points $A_0=(0,0)$, $A_1=(1,0)$, $A_2=(2,0)$ et $A_3=(0,1)$. Évidemment le point A_3 n'est pas dans l'espace engendré par les trois autres ; il n'empêche que ces points ne sont pas affinement indépendants parce que la direction est de dimension 2 au lieu de 3.

Définition 11.40.
*Soit $\mathcal{E}$ un espace affine de dimension n et $\mathcal{F}$ un sous-espace affine de dimension k. Un **repère affine** de $\mathcal{F}$ est la donnée de $k+1$ points affinement indépendants de $\mathcal{F}$.*

Si $\{A_0,\dots,A_n\}$ est un repère affine, le point A_0 est l'**origine**. C'est un choix complètement arbitraire ; et c'est bien cet arbitraire qui nous amènera à considérer les coordonnées barycentriques au lieu des coordonnées cartésiennes.

Soit $M\in\mathcal{E}$; par définition nous avons

$$M=A_0+\overrightarrow{A_0M}. \tag{11.90}$$

Mais nous savons que les vecteurs $\overrightarrow{A_0A_i}$ forment une base de E, nous avons donc des nombres λ_i tels que

$$\overrightarrow{A_0M}=\sum_{i=1}^{n}\lambda_i\overrightarrow{A_0A_i}. \tag{11.91}$$

Les nombres λ_i ainsi construits sont les **coordonnées cartésiennes** du point M dans le repère $\{A_0,\dots,A_n\}$ d'origine A_0.

À partir de ces coordonnées, le point $M\in\mathcal{E}$ se retrouve par la formule

$$M=A_0+\sum_{i=1}^{n}\lambda_i\overrightarrow{A_0A_i}. \tag{11.92}$$

Les coordonnées barycentriques sont données par la proposition suivante.

Proposition 11.41 ([134])**.**
Soient $A_0, \dots, A_r$ des points affinement indépendants dans $\mathcal{E}$ et $\mathcal{F} = \text{Aff}\{A_0, \dots, A_r\}$. Tout point $M \in \mathcal{F}$ s'écrit de façon unique comme barycentre des A_i affectés de poids λ_i tels que $\sum_{i=0}^{r} \lambda_i = 1$.

Les nombres λ_i ainsi définis sont les **coordonnées barycentriques** de M dans le repère $(A_0, \dots, A_r)$ de $\mathcal{F}$.

Démonstration. Nous avons vu plus haut (définition 11.40) que l'affine indépendance des points A_i assurait que $(A_0, \dots, A_r)$ était un repère de $\mathcal{F}$.

En ce qui concerne l'existence de l'écriture de M comme barycentre, nous savons que les sous-espace affine sont exactement les ensembles de barycentres (proposition 11.29), c'est à dire que si on a des points dans un sous-espace affine, alors les barycentres de ces points est encore dans le sous-espace affine.

L'unicité est comme suit. Si M est barycentre des A_i avec poids λ_i, nous écrivons la caractérisation (4) du théorème 11.26 avec $B = A_0$:

$$\overrightarrow{A_0M} = \sum_{i=1}^{r} \lambda_i \overrightarrow{A_0A_i} \tag{11.93}$$

où la somme à droite s'étend a priori de 0 à r, mais comme $\overrightarrow{A_0A_0} = 0$, nous l'avons limitée à 1. Si M s'écrit comme barycentre de deux façons différentes, nous aurions

$$\overrightarrow{A_0M} = \sum_{i=1}^{r} \lambda_i \overrightarrow{A_0A_i} = \sum_{i=1}^{r} \mu_i \overrightarrow{A_0A_i} \tag{11.94}$$

avec $\sum_i \lambda_i = \sum_i \mu_i = 1$. Étant donné que les points $A_0, \dots, A_r$ forment un repère, les vecteurs $\overrightarrow{A_0A_i}$ sont linéairement indépendants (point (5) de la proposition 11.39) et donc $\lambda_i = \mu_i$ pour $i = 1, \dots, r$. La condition de somme des points égale à 1 impose alors immédiatement $\lambda_0 = \mu_0$. □

11.42.
Soit $\mathbb{R}^2$ et les points non alignés A, B, C. Les coordonnées barycentriques (α, β, γ) dans ce système correspondent à l'unique $X \in \mathbb{R}^2$ tel que

$$\alpha\overrightarrow{XA} + \beta\overrightarrow{XB} + \gamma\overrightarrow{XC} = 0. \tag{11.95}$$

Exemple 11.43
Soient les points $A = (3, 1)$, $B = (-1, 2)$ et $C = (0, -1)$ dans $\mathbb{R}^2$. Nous allons montrer qu'il forment un repère affine de $\mathbb{R}^2$. L'espace engendré par ces trois points est l'espace des

$$A + \alpha\overrightarrow{AB} + \beta\overrightarrow{AC}, \tag{11.96}$$

et la direction correspondante est l'espace vectoriel donné par $\alpha\overrightarrow{AB} + \beta\overrightarrow{AC}$ qui est de dimension deux. Donc l'espace affine engendré par A, B et C est de dimension 2. △

Exemple 11.44
Dans le repère (A, B, C), quel est le point de coordonnées barycentriques $(\frac{1}{6}, \frac{1}{3}, \frac{1}{2})$? D'abord nous vérifions que

$$\frac{1}{6} + \frac{1}{3} + \frac{1}{2} = 1. \tag{11.97}$$

Ensuite nous cherchons $X \in \mathbb{R}^2$ tel que

$$\frac{1}{6}\overrightarrow{AX} + \frac{1}{3}\overrightarrow{BX} + \frac{1}{2}\overrightarrow{CX} = 0, \tag{11.98}$$

c'est à dire

$$\frac{1}{6}\begin{pmatrix} x-3 \\ y-1 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} x+1 \\ y-2 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} x \\ y+1 \end{pmatrix} = 0. \tag{11.99}$$

Nous trouvons immédiatement $x = 1/6$ et $y = 1/3$. Le point cherché est donc le point $\begin{pmatrix}1/6\\1/3\end{pmatrix}$. △

Lemme 11.45.
Une application affine $f\colon \mathcal{E} \to \mathcal{E}$ qui préserve les points d'une base affine de $\mathcal{E}$ est l'identité.

Démonstration. Une base affine de $\mathcal{E}$ consiste en $n+1$ points $\{A_0, \dots, A_n\}$ qui engendrent un espace de dimension n. Tous point de $\mathcal{E}$ peut être écrit sous la forme $\sum_i \lambda_i A_i$ avec $\sum_i \lambda_i = 1$ (proposition 11.41). Mézenplus, l'application affine f peut être écrite comme $f = \tau_c \circ \alpha$ (ça c'est vu en 11.12), et $v = \sum_i \lambda_i v$, donc

$$f\big(\lambda_i A_i\big) = \sum_i \alpha(A_i) + v = \sum_i \lambda_i \big(\alpha(A_i) + v\big) = \sum_i \lambda_i f(A_i) = \sum_i \lambda_i A_i. \tag{11.100}$$

Donc tout point de $\mathcal{E}$ est fixé par f est f est l'identité. □

11.7.1 Équation de droite

Soit $\mathcal{E}$ un espace affine de dimension trois muni d'un repère barycentrique. Une droite est donnée par trois nombres : $D = D(a,b,c)$ est l'ensemble des points dont les coordonnées barycentriques (normalisées) (x,y,z) vérifient $ax + by + cz = 0$. C'est un espace de dimension un parce qu'il y a aussi la condition $x + y + z = 1$.

La droite $D(1,1,1,)$ n'existe pas parce que ce serait $x + y + z = 0$, qui est incompatible avec $x + y + z = 1$.

Les droites $D(a,b,c)$ et $D(a',b',c')$ s'intersectent selon les solutions du système

$$\begin{cases} x + y + z = 1 & \text{(11.101a)}\\ ax + by + cz = 0 & \text{(11.101b)}\\ a'x + b'y + c'z = 0 & \text{(11.101c)} \end{cases}$$

Donc deux droites affines ont un unique point d'intersection si et seulement si

$$d = \begin{vmatrix} 1 & 1 & 1\\ a & b & c\\ a' & b' & c' \end{vmatrix} \neq 0. \tag{11.102}$$

Elles seront parallèles ou confondues si et seulement si $d = 0$.

11.7.2 Associativité, coordonnées barycentriques dans un triangle

Lemme 11.46 ([136])**.**
Soient trois points non alignés A, B, C ainsi que des nombres α, β, γ tels que $\alpha + \beta \neq 0$ et $\alpha + \beta + \gamma \neq 0$.

Soit H le barycentre du système $\{(A,\alpha),(B,\beta)\}$ et G le barycentre de $\{(A,\alpha),(B,\beta),(C,\gamma)\}$.

Alors G est barycentre de $\{(H,\alpha+\beta),(C,\gamma)\}$.

Démonstration. Vues les définition de H et G nous avons

$$\alpha\overrightarrow{HA} + \beta\overrightarrow{HB} = 0 \tag{11.103a}$$
$$\alpha\overrightarrow{GA} + \beta\overrightarrow{GB} + \gamma\overrightarrow{GC} = 0. \tag{11.103b}$$

En utilisant les relations de Chasles nous introduisons H dans la seconde relation :

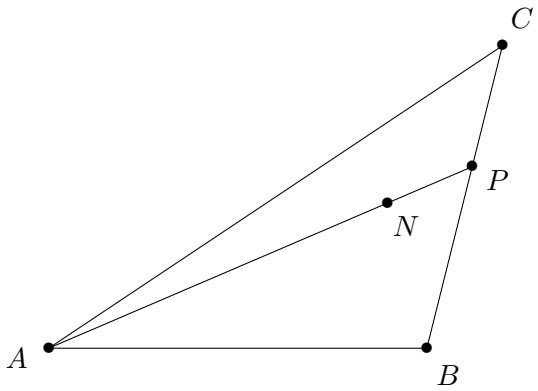
$$\alpha(\overrightarrow{GH} + \overrightarrow{HA}) + \beta(\overrightarrow{GH} + \overrightarrow{HB}) + \gamma\overrightarrow{GC} = 0 \tag{11.104a}$$
$$(\alpha+\beta)\overrightarrow{GH} + \underbrace{\alpha\overrightarrow{HA} + \beta\overrightarrow{HB}}_{=0} + \gamma\overrightarrow{GC} = 0 \tag{11.104b}$$
$$(\alpha+\beta)\overrightarrow{GH} + \gamma\overrightarrow{GC} = 0. \tag{11.104c}$$

□

Les coordonnées barycentriques dans un triangle (et plus généralement en fait) permettent de faire des projections.

Proposition 11.47.
Soient trois points non alignés A, B, C ainsi qu'un point N de coordonnées barycentriques (α, β, γ) dans le système (A, B, C). Si P est l'intersection $(AN) \cap (BC)$ alors les coordonnées de P sont $(0, \beta, \gamma)$.

Démonstration. Un dessin de la situation :



Dire que les coordonnées de N sont (α, β, γ) signifie que

$$\alpha\overrightarrow{NA} + \beta\overrightarrow{NB} + \gamma\overrightarrow{NC} = 0. \tag{11.105}$$

Nous voudrions montrer que le point P est bien le point de coordonnées $(0, \beta, \gamma)$. Soit donc le point P tel que

$$\beta\overrightarrow{PB} + \gamma\overrightarrow{PC} = 0 \tag{11.106}$$

et montrons que ce point est l'intersection $(BC) \cap (NA)$.

D'abord la relation (11.106) nous dit immédiatement que P est sur la droite (BC). Ensuite, en utilisant les relations de Chasles pour introduire N :

$$\beta(\overrightarrow{PN} + \overrightarrow{NB}) + \gamma(\overrightarrow{PN} + \overrightarrow{NC}) = 0. \tag{11.107}$$

Nous remplaçons $\beta\overrightarrow{NB} + \gamma\overrightarrow{NC}$ par $-\alpha\overrightarrow{NA}$ pour obtenir :

$$(\beta + \gamma)\overrightarrow{PN} - \alpha\overrightarrow{NA} = 0. \tag{11.108}$$

Cela montre que les vecteurs $\overrightarrow{PN}$ et $\overrightarrow{NA}$ sont colinéaires, et donc que P, N et A sont alignés. □

11.8 Structure du groupe des applications affines

Soit $v \in \mathbb{R}^n$; nous notons $\tau_v \colon \mathbb{R}^n \to \mathbb{R}^n$ la translation donnée par $\tau_v(x) = x + v$. Le groupe de toutes les translation de $\mathbb{R}^n$ est noté $T(n)$ et est isomorphe au groupe abélien $(\mathbb{R}^n, +)$.

Nous avons déjà discuté de la structure d'un espace vectoriel (en particulier $\mathbb{R}^n$) comme espace affine en 11.4.

Lemme 11.48.
Une application $f \colon E \to E$ est affine si et seulement si il existe $v \in E$ et une application linéaire α sur E telle que $f = \tau_v \circ \alpha$.

Démonstration. Nous supposons d'abord que f est affine. Alors il existe une application linéaire u_f sur E telle que

$$f(M + x) = f(M) + u_f(x) = (\tau_{f(M)} + u_f)(x) \tag{11.109}$$

pour tout x et M. De plus l'application u_f ne dépend ni de M ni de x (c'est la proposition 11.11(1)). En posant $M = 0$ nous avons :

$$f(x) = (\tau_{f(0)} \circ u_f)(x). \tag{11.110}$$

Dans l'autre sens nous supposons avoir $v \in E$ et α linéaire sur E telles que

$$f(M) = (\tau_v \circ \alpha)(M). \tag{11.111}$$

Notons qu'il y a un abus de notation entre α qui est linéaire sur l'espace *vectoriel* E et l'application α qui est une application sur l'espace *affine* E. Cet abus est légitime parce que les deux espaces sont identiques en tant qu'ensembles. Ce qui est vraiment abuser par contre, c'est de se poser ce genre de questions.

Nous avons :

$$\begin{aligned} f(M + x) = \tau_v\big(\alpha(M + x)\big) = \alpha(M + x) + v = \alpha(M) + v + \alpha(x) \\ = (\tau_v \circ \alpha)(M) + \alpha(x) = f(M) + \alpha(x). \end{aligned} \tag{11.112}$$

Donc la fonction f vérifie la définition 11.7. □

Corollaire 11.49.
Une application affine qui conserve l'origine est linéaire.

Démonstration. Conserver l'origine demande de poser $v = 0$ dans l'expression du lemme 11.48. □

Proposition 11.50.
Soit une application affine $f : \mathbb{R}^n \to \mathbb{R}^n$. L'ensemble des points fixes

$$\mathrm{Fix}(f) = \{x \in \mathbb{R}^n \text{ tel que } f(x) = x\} \tag{11.113}$$

est soit vide soit un sous-espaces affine de $\mathbb{R}^n$.

Démonstration. Soit $f = \tau_v \circ \alpha$; nous avons $x \in \mathrm{Fix}(f)$ si et seulement si

$$x = \tau_v\big(\alpha(x)\big) = \alpha(x) + v, \tag{11.114}$$

autrement dit, en considérant l'application linéaire $\beta = \mathrm{Id} - \alpha$, si et seulement si $\beta(x) = v$. Nous écrivons $\mathrm{Fix}(f) = \beta^{-1}(v)$. Supposons que ce soit non vide et considérons $x_0 \in \beta^{-1}(v)$. Nous avons

$$\begin{aligned} \beta^{-1}(v) &= \{x \in \mathbb{R}^n \text{ tel que } \beta(x) = \beta(x_0)\} && \text{(11.115a)} \\ &= \{x \text{ tel que } \beta(x - x_0) = 0\} && \text{(11.115b)} \\ &= \{x \text{ tel que } x - x_0 \in \ker(\beta)\} && \text{(11.115c)} \\ &= \ker(\beta) + x_0 && \text{(11.115d)} \\ &= \tau_{x_0}\big(\ker(\beta)\big). && \text{(11.115e)} \end{aligned}$$

Mais comme $\ker(\beta)$ est un sous-espace vectoriel, $\beta^{-1}(v)$ est le translaté d'un sous-espace vectoriel, c'est à dire un sous-espace affine. □

Définition 11.51.
Le groupe $\mathrm{Aff}(\mathbb{R}^n)$ est le groupe des applications affines bijectives de $\mathbb{R}^n$.

Remarque 11.52.
Le fait que ce soit un groupe est rapidement vu par le calcul suivant :

$$(\tau_v \circ \alpha) \circ (\tau_w \circ \beta)(x) = (\alpha\beta)(x) + \alpha(w) + v = \tau_{\alpha(w)+v} \circ (\alpha\beta). \tag{11.116}$$

Proposition 11.53.
L'ensemble $\mathrm{Aff}(\mathbb{R}^n)$ *est isomorphe au produit semi-direct*[4]

$$\mathrm{Aff}(\mathbb{R}^n) \simeq T(n) \times_{\boldsymbol{Ad}} \mathrm{GL}(n,\mathbb{R}) \tag{11.117}$$

où $\boldsymbol{Ad}$ *est l'action adjointe, c'est à dire*

$$\begin{aligned} \boldsymbol{Ad}\colon \mathrm{GL}(n,\mathbb{R}) &\to \mathrm{Aut}\big(T(n)\big) \\ \tau_v &\mapsto \alpha \circ \tau_v \circ \alpha^{-1}. \end{aligned} \tag{11.118}$$

Démonstration. L'application que nous allons montrer être un isomorphisme est ψ qui à $f = \tau_v \circ \alpha$ fait correspondre le couple $(\tau_v, \alpha) \in T(n) \times \mathrm{GL}(n,\mathbb{R})$.

Égalité d'ensembles Il faut que $\mathrm{Aff}(\mathbb{R}^n)$ soit en bijection avec $T(n) \times \mathrm{GL}(n,\mathbb{R})$. En effet si $f \in \mathrm{Aff}(\mathbb{R}^n)$, la décomposition $f = \tau_v \circ \alpha$ est unique. D'abord en appliquant à 0, $f(0) = \tau_v\big(\alpha(v)\big) = v$. Donc v est fixé par la valeur de $f(0)$. Ensuite $\alpha = f \circ \tau_v^{-1}$, donc α fixé.

L'action adjointe fonctionne Il faut vérifier que $\alpha \circ \tau_v \circ \alpha^{-1}$ est bien dans $T(n)$. Pour cela, en agissant sur $x \in \mathbb{R}^n$ nous trouvons

$$\alpha\tau_v\alpha^{-1}(x) = \alpha\big(\alpha^{-1}(x) + v\big) = x + \alpha(v) = \tau_{\alpha(v)}(x). \tag{11.119}$$

Le fait que $\mathbf{Ad}(\alpha)$ soit un automorphisme est toujours correct.

Morphisme Il faut vérifier que l'application ψ est un morphisme de groupe. D'abord la loi de groupe sur $\mathrm{Aff}(\mathbb{R}^n)$ est donnée par

$$(\tau_v \circ \alpha) \circ (\tau_w \circ \beta) = \tau_{v+\alpha(w)} \circ (\alpha \circ \beta). \tag{11.120}$$

Ensuite le loi de groupe de le produit semi-direct est donnée par

$$(\tau_v, \alpha) \cdot (\tau_w, \beta) = \big(\tau_v\, \mathbf{Ad}(\alpha)\tau_w, \alpha\beta\big) = \big(\tau_v\tau_{\alpha(w)}, \alpha\beta\big) = \big(\tau_{\alpha(w)+v}, \alpha\beta\big). \tag{11.121}$$

Nous avons donc bien

$$\psi\big((\tau_v, \beta) \cdot (\tau_w, \beta)\big) = \psi(\tau_v, \beta) \circ \psi(\tau_w, \beta). \tag{11.122}$$

□

11.9 Isométries

Définition 11.54 (Isométrie d'espace affine)**.**
Si $\mathcal{E}$ *est un espace affine muni d'une distance* d*, une isométrie de* $\mathcal{E}$ *est une application* $f\colon \mathcal{E} \to \mathcal{E}$ *préservant* d*.*

Notons que toutes les applications affines ne sont pas des isométries : par exemple les homothéties.

Si $\mathcal{E}$ est modelé sur un espace euclidien $(E, \|.\|)$ alors la formule

$$d(A,B) = \|\overrightarrow{AB}\| \tag{11.123}$$

définit une distance sur $\mathcal{E}$.

Étant donné ce qui est dit en 11.2, la formule a un sens parce qu'à A et B donnés dans $\mathcal{E}$, il est associé un unique vecteur $\overrightarrow{AB} \in E$.

Nous parlons des isométries affines du tétraèdre régulier dans la proposition 16.17.

4. Définition 3.93.

Chapitre 12

Analyse réelle

12.1 Intervalles

Définition 12.1 (Intervalle)**.**
Une partie I de $\mathbb{R}$ est un ***intervalle*** *si pour tout $a, b \in I$ nous avons $t \in I$ dès que $a \leqslant t \leqslant b$.*

Un intervalle est ***ouvert*** *s'il est de la forme $]a, b[$ avec éventuellement $a = -\infty$ ou $b = +\infty$.*
Un intervalle est ***fermé*** *s'il est de la forme $[a, b]$ ou $]-\infty, b]$ ou $[a, +\infty[$ avec $a, b \in \mathbb{R}$.*

Remarque 12.2.
L'ensemble $\mathbb{R}$ ne contient pas $-\infty$ et $-\infty$. L'intervalle $[-\infty, 5]$ par exemple, n'est pas une partie de $\mathbb{R}$.

Exemple 12.3

(1) Les ensembles $]3, 7[$ et $]-\infty, \pi[$ sont des intervalles ouverts.
(2) Les ensembles $[10, 15]$ et $[-1, +\infty[$ sont des intervalles fermés.
(3) L'ensemble $]-4, -2[\cup]2, 9[$ n'est pas un intervalle (il y a un «trou» entre -2 et 2).
(4) L'ensemble $\mathbb{R}$ lui-même est un intervalle ; par convention, il est à la fois ouvert et fermé.

Un intervalle peut n'être ni ouvert ni fermé ; par exemple $]4, 8]$. Cet intervalle est «ouvert en 4 et fermé en 8» . △

Définition 12.4 (Fonction, domaine, image, graphe)**.**
Soient X et Y deux ensembles. Une ***fonction*** *f définie sur X et à valeurs dans Y est une correspondence qui associe à chaque élément x dans X* **au plus** *un élément y dans Y. On écrit $y = f(x)$.*

— *La partie de X qui contient tous les x sur lesquels f peut opérer est dite* ***domaine*** *de f. Le domaine de f est indiqué par* Domaine f.
— *L'élément de $y \in Y$ associé par f à un élément $x \in$ Domaine f (c'est à dire $f(x) = y$) est appellé* ***image*** *de x par f. L'****image*** *de la fonction f est la partie de Y qui contient les images de tous les éléments de* Domaine f. *L'image de f est indiquée par $\Im f$.*
— *Le* ***graphe*** *de f est l'ensemble de toutes les couples $(x, f(x))$ pour $x \in$ Domaine f. Le graphe de f est une partie de l'ensemble noté $X \times Y$ et il est indiqué par* Graph f. *Dans ce cours $X = \mathbb{R}$ et $Y = \mathbb{R}$, donc le graphe de f est contenu dans le plan cartésien.*

Définition 12.5 (Fonction croissante, décroissante et monotone)**.**
Soit une fonction $f : \mathbb{R} \to \mathbb{R}$ et un intervalle $I \subset \mathbb{R}$.

(1) Le fonction f est ***croissante*** *sur I si pour tout $x < y$ dans I nous avons $f(x) \leqslant f(y)$. Elle est* strictement *croissante si $f(x) < f(y)$ dès que $x < y$.*
(2) Le fonction f est ***décroissante*** *sur I si pour tout $x < y$ dans I nous avons $f(x) \geqslant f(y)$. Elle est* strictement *décroissante si $f(x) > f(y)$ dès que $x < y$.*

*(3) La fonction f est dite **monotone** sur I si elle est soit croissante soit décroissante sur I.*

Exemple 12.6
La fonction $x \mapsto x^2$ est décroissante sur l'intervalle $]-\infty, 0]$ et croissante sur l'intervalle $[0, \infty[$. Elle n'est par contre ni croissante ni décroissante sur l'intervalle $[-4, 3]$. △

12.2 Application réciproque

12.2.1 Définitions

Les définitions d'injection, surjection, bijection et d'application réciproque sont les définitions 6.186 et 6.187.

Exemple 12.7

(1) La fonction $x \mapsto x^2$ n'est pas une bijection de $\mathbb{R}$ vers $\mathbb{R}$ parce qu'il n'existe aucun x tel que $x^2 = -1$.

(2) La fonction
$$\begin{aligned} f\colon [0, +\infty[&\to [0, +\infty[\\ x &\mapsto x^2 \end{aligned} \tag{12.1}$$
est une bijection. Notez que c'est la même fonction que celle de l'exemple précédent. Seul l'intervalle sur laquelle nous nous plaçons a changé.

(3) La fonction
$$\begin{aligned} \sin\colon \mathbb{R} &\to [-1, 1] \\ x &\mapsto \sin(x) \end{aligned} \tag{12.2}$$
n'est pas une bijection parce qu'il existe plus de un x tel que $\sin(x) = 1$.

En conclusion : il est très important de préciser les domaines des fonctions considérées. △

Remarque 12.8.
Dire que la fonction $f\colon I \to J$ est bijective, c'est dire que l'équation $f(x) = y$ d'inconnue x peut être résolue de façon univoque pour tout $y \in J$.

Remarque 12.9.
Toute fonction strictement monotone sur un intervalle I est injective.

Proposition 12.10.
Une fonction monotone et surjective d'un intervalle I sur un autre intervalle J est continue sur I.

Exemple 12.11
La fonction
$$\begin{aligned} f\colon [2, 3] &\to [4, 9] \\ x &\mapsto x^2 \end{aligned} \tag{12.3}$$
est une bijection. Sa réciproque est la fonction
$$\begin{aligned} f^{-1}\colon [4, 9] &\to [2, 3] \\ x &\mapsto \sqrt{x}. \end{aligned} \tag{12.4}$$

△

Exemple 12.12
Trouvons la fonction réciproque de la fonction affine $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto 3x - 2$. Si $y \in \mathbb{R}$ le nombre $f^{-1}(y)$ est la valeur de x pour laquelle $f(x) = y$. Il s'agit donc de résoudre

$$3x - 2 = y \tag{12.5}$$

par rapport à x. La solution est $x = \frac{y+2}{3}$ et donc nous écrivons

$$f^{-1}(y) = \frac{y+2}{3}. \tag{12.6}$$

Notons que dans les calculs, il est plus simple d'écrire «y» que «x» la variable de la fonction réciproque. Il est néanmoins (très) recommandé de nommer «x» la variable dans la réponse finale. Dans notre cas nous concluons donc

$$f^{-1}(x) = \frac{x+2}{3}. \tag{12.7}$$

△

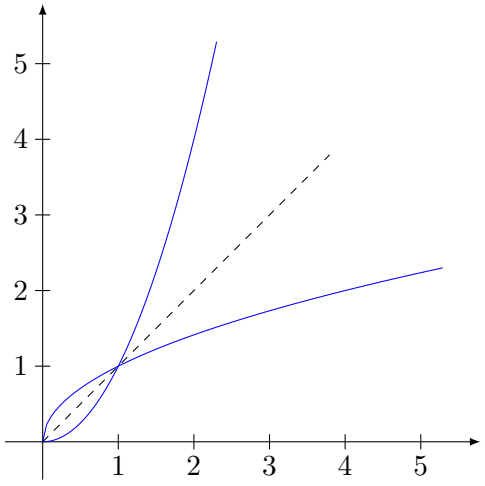
12.2.2 Graphe de la fonction réciproque

Par définition le graphe de la fonction f est l'ensemble des points de la forme (x, y) vérifiant $y = f(x)$. Affin de déterminer le graphe de la bijection réciproque nous pouvons faire le raisonnement suivant.

Le point (x_0, y_0) est sur le graphe de f
⇔
La relation $f(x_0) = y_0$ est vérifiée
⇔
La relation $x_0 = f^{-1}(y_0)$ est vérifiée
⇔
Le point (y_0, x_0) est sur le graphe de f^{-1}.

À retenir 12.13
Dans un repère orthonormal, le graphe de le bijection réciproque est obtenu à parti du graphe de f en effectuant une symétrie par rapport à la droite d'équation $y = x$.

Le dessin suivant montre le cas de la courbe de la fonction carré comparé à celle de la racine carré.



12.3 Fonctions continues

Nous allons considérer trois approches différentes de la continuité. La première sera de définir la continuité de fonctions de $\mathbb{R}$ vers $\mathbb{R}$ au moyen du critère usuel. Ensuite, nous définiront la continuité des applications entre n'importes quels espaces métriques, et nous montrerons que les deux définitions sont équivalentes dans le cas des fonctions sur $\mathbb{R}$ à valeurs réelles.

Enfin, un peu plus tard nous verrons que la continuité peut également être vue en termes de limites. Encore une fois nous verrons que dans le cas de fonctions de $\mathbb{R}$ vers $\mathbb{R}$ cette troisième approche est équivalentes aux deux premières.

La définition de fonction continue est la définition 6.69.

Nous allons donc dire qu'une fonction est continue quand plus x s'approche de a en suivant la courbe, plus $f(x)$ s'approche de $f(a)$. Voici la définition précise.

Définition 12.14.
*Nous disons que la fonction $x \mapsto f(x)$ est **continue en** a si*

$$\forall \epsilon > 0, \exists \delta \text{ tel que } \big(|x-a| \leqslant \delta\big) \Rightarrow |f(x)-f(a)| \leqslant \epsilon. \tag{12.8}$$

Définition 12.15.
*Soit une fonction $f\colon D \to \mathbb{R}$ et un point a dans D. Nous disons que f est **continue** lorsque f possède une limite en a et $\lim_{x\to a} f(x) = f(a)$.*

En remplaçant ℓ par $f(a)$ dans la définition de la limite, nous exprimons la continuité de f en a par la façon suivante. Pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que $\forall x \in D$,

$$|x-a| < \delta \Rightarrow \big|f(x)-f(a)\big| < \varepsilon. \tag{12.9}$$

Nous allons maintenant étudier quelques conséquences de cette définition.

(1) D'abord on voit que la continuité n'a été définie qu'en un point. On peut dire que la fonction f est continue *en tel point donné*, mais nous n'avons pas dit ce qu'est une fonction continue *dans son ensemble.*

(2) Si I est un intervalle de $\mathbb{R}$, on dit que f est **continue sur l'intervalle** I si elle est continue en chaque point de I.

(3) Comme la définition de f continue en a fait intervenir $f(x)$ pour tous les x pas trop loin de a, il faut au moins déjà que f soit définie sur ces x. En d'autres termes, dire que f est continue en a demande que f existe sur un intervalle autour de a.

Ceci couplé à la définition précédente laisse penser qu'il est surtout intéressant d'étudier les fonctions qui sont continues sur un intervalle.

(4) L'intuition comme quoi une fonction continue doit pouvoir être tracée sans lever la main correspond aux fonctions continues sur des intervalles. Au moins sur l'intervalle où elle est continue, elle est traçable en un morceau.

Nous allons démontrer maintenant une série de petits résultats qui permettent de simplifier la démonstration de la continuité de fonctions.

Théorème 12.16.
Si la fonction f est continue au point a, alors la fonction λf est également continue en a.

Démonstration. Soit $\epsilon > 0$. Nous avons besoin d'un $\delta > 0$ tel que pour chaque x à moins de δ de a, la fonction λf soit à moins de ϵ de $(\lambda f)(a) = \lambda f(a)$. Étant donné que la fonction f est continue en a, on sait déjà qu'il existe un δ_1 (nous notons δ_1 afin de ne pas confondre ce nombre dont on est sûr de l'existence avec le δ que nous sommes en train de chercher) tel que

$$(|x-a| \leqslant \delta_1) \Rightarrow |f(x)-f(a)| \leqslant \epsilon_1.$$

Hélas, ce δ_1 n'est pas celui qu'il faut faut parce que nous travaillons avec λf au lieu de f, ce qui fait qu'au lieu d'avoir $|f(x) - f(a)|$, nous avons $|\lambda f(x) - \lambda f(a)| = |\lambda| \cdot |f(x) - f(a)|$. Ce que δ_1 fait avec (λf), c'est

$$(|x - a| \leqslant \delta_1) \Rightarrow |(\lambda f)(x) - (\lambda f)(a)| \leqslant |\lambda|\epsilon_1.$$

Ce que nous apprend la continuité de f, c'est que pour chaque choix de ϵ_1, on a un δ_1 qui fait cette implication. Comme cela est vrai pour chaque choix de ϵ_1, essayons avec $\epsilon_1 = \epsilon/|\lambda|$ pour voir ce que ça donne. Nous avons donc un δ_1 qui fait

$$(|x - a| \leqslant \delta_1) \Rightarrow |(\lambda f)(x) - (\lambda f)(a)| \leqslant |\lambda|\epsilon_1 = \epsilon.$$

Ce δ_1 est celui qu'on cherchait. □

Théorème 12.17.
Si f et g sont deux fonctions continues en a, alors la fonction $f + g$ est également continue en a.

Démonstration. La continuité des fonctions f et g au point a fait en sorte que pour tout choix de ϵ_1 et ϵ_2, il existe δ_1 et δ_2 tels que

$$(|x - a| \leqslant \delta_1) \Rightarrow |f(x) - f(a)| \leqslant \epsilon_1.$$

et

$$(|x - a| \leqslant \delta_2) \Rightarrow |g(x) - g(a)| \leqslant \epsilon_2.$$

La quantité que nous souhaitons analyser est $|f(x) + g(x) - f(a) - g(a)|$. Tout le jeu de la démonstration de la continuité est de triturer cette expression pour en tirer quelque chose en termes de ϵ_1 et ϵ_2. Si nous supposons avoir pris $|x - a|$ plus petit en même temps que δ_1 et que δ_2, nous avons

$$|f(x) + g(x) - f(a) - g(a)| \leqslant |f(x) - g(x)| + |g(x) - g(a)| \leqslant \epsilon_1 + \epsilon_2$$

en utilisant la formule générale $|a+b| \leqslant |a|+|b|$. Maintenant si on choisit ϵ_1 et ϵ_2 tels que $\epsilon_1+\epsilon_2 < \epsilon$, et les δ_1, δ_2 correspondants, on a que

$$|f(x) + g(x) - f(a) - g(a)| \leqslant \epsilon,$$

pourvu que $|x - a|$ soit plus petit que δ_1 et δ_2. Le bon δ a prendre est donc le minimum de δ_1 et δ_2 qui eux-même sont donnés par un choix de ϵ_1 et ϵ_2 tels que $\epsilon_1 + \epsilon_2 \leqslant \epsilon$. □

Pour résumer ces deux théorèmes, on dit que si f et g sont continues en a, alors la fonction $\alpha f + \beta g$ est également continue en a pour tout α, $\beta \in \mathbb{R}$.

Parmi les propriétés immédiates de la continuité d'une fonction, nous avons ceci qui est souvent bien utile.

Corollaire 12.18.
Si la fonction f est continue en a et si $f(a) > 0$, alors f est positive sur un intervalle autour de a.

Démonstration. Prenons $\epsilon < f(a)$ et voyons[1] ce que la continuité de f en a nous offre : il existe un δ tel que

$$(|x - a| \leqslant \delta) \Rightarrow |f(x) - f(a)| \leqslant \epsilon < f(a).$$

Nous en retenons que sur un intervalle (de largeur δ), nous avons $|f(x) - f(a)| \leqslant f(a)$. Par hypothèse, $f(a) > 0$, donc si $f(x) < 0$, alors la différence $f(x) - f(a)$ donne un nombre encore plus négatif que $-f(a)$, c'est à dire que $|f(x) - f(a)| > f(a)$, ce qui est contraire à ce que nous venons de démontrer. D'où la conclusion que $f(x) > 0$. □

1. ici, nous insistons sur le fait que nous prenons ϵ *strictement* plus petit que $f(a)$.

12.3.1 La fonction la moins continue du monde

Parmi les exemples un peu sales de fonctions non continues, il y a celle-ci :

$$\chi_{\mathbb{Q}}(x) = \begin{cases} 1 \text{ si } x \in \mathbb{Q} \\ 0 \text{ sinon.} \end{cases}$$

Par exemple, $\chi_{\mathbb{Q}}(0) = 1$, et[2] $\chi_{\mathbb{Q}}(\pi) = \chi_{\mathbb{Q}}(\sqrt{2}) = 0$. Malgré que $\chi_{\mathbb{Q}}(0) = 1$, il n'existe *aucun* voisinage de 1 sur lequel la fonction reste proche de 1, parce que tout voisinage va contenir au moins un irrationnel. À chaque millimètre, cette fonction fait une infinité de bonds !

Cette fonction n'est donc continue nulle part.

À partir de là, nous pouvons construire la fonction suivante qui n'est continue qu'en un point :

$$f(x) = x\chi_{\mathbb{Q}}(x) = \begin{cases} x \text{ si } x \in \mathbb{Q} \\ 0 \text{ sinon.} \end{cases}$$

Cette fonction est continue en zéro. En effet, prenons $\delta > 0$; il nous faut un ϵ tel que $|x| \leqslant \epsilon$ implique $f(x) \leqslant \delta$ parce que $f(0) = 0$. Bon ben prendre simplement $\epsilon = \delta$ nous contente. Cette fonction est donc très facilement continue en zéro.

Et pourtant, dès que l'on s'écarte un tant soit peu de zéro, elle fait des bons une infinité de fois par millionième de millimètre ! Cette fonction est donc la plus discontinue du monde en tous les points saut un (zéro) où elle est une fonction continue !

12.3.2 Approche topologique

Nous avons vu que sur tout ensemble métrique, nous pouvons définir ce qu'est un ouvert : c'est un ensemble qui contient une boule ouverte autour de chacun de ses points. Quand on est dans un ensemble ouvert, on peut toujours un peu se déplacer sans sortir de l'ensemble.

Le théorème suivant est une très importante caractérisation des fonctions continues (de $\mathbb{R}$ dans $\mathbb{R}$) en termes de topologie, c'est à dire en termes d'ouverts.

Théorème 12.19.
Si I est un intervalle ouvert contenu dans $\operatorname{dom} f$*, alors f est continue sur I si et seulement si pour tout ouvert $\mathcal{O}$ dans $\mathbb{R}$, l'image inverse $f|_I^{-1}(\mathcal{O})$ est ouvert.*

Par abus de langage, nous exprimons souvent cette condition par « une fonction est continue si et seulement si l'image inverse de tout ouvert est un ouvert ».

Démonstration. Dans un premier temps, nous allons transformer le critère de continuité en termes de boules ouvertes, et ensuite, nous passeront à la démonstration proprement dite. Le critère de continuité de f au point x dit que

$$\forall \delta > 0, \exists\, \epsilon > 0 \text{ tel que } \big(|x - a| < \epsilon\big) \Rightarrow |f(x) - f(a)| < \delta. \tag{12.10}$$

Cette condition peut être exprimée sous la forme suivante :

$$\forall \delta > 0, \exists \epsilon \text{ tel que } a \in B(x, \epsilon) \Rightarrow f(a) \in B\big(f(x), \delta\big),$$

ou encore

$$\forall \delta > 0, \exists \epsilon \text{ tel que } f\big(B(x, \epsilon)\big) \subset B\big(f(x), \delta\big). \tag{12.11}$$

Jusque ici, nous n'avons fait que du jeu de notations. Nous avons exprimé en termes de topologie des inégalités analytiques. Si tu veux, tu peux retenir cette condition (12.11) comme définition d'une fonction continue en x. Si tu choisit de vivre comme ça, tu dois être capable de retrouver (12.10) à partir de (12.11).

2. Pour prouver que $\sqrt{2}$ n'est pas rationnel, c'est pas trop compliqué, mais pour prouver que π ne l'est pas non plus, il faudra encore manger de la soupe.

Passons maintenant à la démonstration proprement dite du théorème.

D'abord, supposons que f est continue sur I, et prenons $\mathcal{O}$, un ouvert quelconque. Le but est de prouver que $f|_I^{-1}(\mathcal{O})$ est ouvert. Pour cela, nous prenons un point $x_0 \in f|_I^{-1}(\mathcal{O})$ et nous allons trouver un ouvert autour ce ce point contenu dans $f|_I^{-1}(\mathcal{O})$. Nous écrivons $y_0 = f(x_0)$. Évidement, $y_0 \in \mathcal{O}$, donc on a une boule autour de y_0 qui est contenue dans $\mathcal{O}$, soit donc $\delta > 0$ tel que

$$B(y_0, \delta) \subset \mathcal{O}.$$

Par hypothèse, f est continue en x_0, et nous pouvons donc y appliquer le critère (12.11). Il existe donc $\epsilon > 0$ tel que

$$f\big(B(x_0, \epsilon)\big) \subset B\big(f(x_0), \delta\big) \subset \mathcal{O}.$$

Cela prouve que $B(x_0, \epsilon) \subset f|_I^{-1}(\mathcal{O})$.

Dans l'autre sens, maintenant. Nous prenons $x_0 \in I$ et nous voulons prouver que f est continue en x_0, c'est à dire que pour tout δ nous cherchons un ϵ tel que $f\big(B(x_0, \epsilon)\big) \subset B\big(f(x_0), \delta\big)$. Oui, mais $B\big(f(x_0), \delta\big)$ est ouverte, donc par hypothèse, $f|_I^{-1}\Big(B\big(f(x_0), \delta\big)\Big)$ est ouvert, inclus à I et contient x_0. Donc il existe un ϵ tel que

$$B(x_0, \epsilon) \subset f|_I^{-1}\Big(B\big(f(x_0), \delta\big)\Big),$$

et donc tel que

$$f\big(B(x_0, \epsilon)\big) \subset B\big(f(x_0), \delta\big),$$

ce qu'il fallait prouver. □

Lemme 12.20.
L'image d'un ensemble connexe par une fonction continue est connexe.

Démonstration. Nous allons encore faire la contraposée. Soit A une partie de $\mathbb{R}$ telle que $f(A)$ ne soit pas connexe. Nous allons prouver que A elle-même n'est pas connexe. Dire que $f(A)$ n'est pas connexe, c'est dire qu'il existe $\mathcal{O}_1$ et $\mathcal{O}_2$, deux ouverts disjoints qui recouvrent $f(A)$. Je prétends que $f^{-1}(\mathcal{O}_1)$ et $f^{-1}(\mathcal{O}_2)$ sont ouverts, disjoints et qu'ils recouvrent A.

- Ces deux ensembles sont ouverts parce qu'ils sont images inverses d'ouverts par une fonction continue (théorème 12.19).
- Si $x \in f^{-1}(\mathcal{O}_1) \cap f^{-1}(\mathcal{O}_2)$, alors $f(x) \in \mathcal{O}_1 \cap \mathcal{O}_2$, ce qui contredirait le fait que $\mathcal{O}_1$ et $\mathcal{O}_2$ sont disjoints. Il n'y a donc pas d'éléments dans l'intersection de $f^{-1}(\mathcal{O}_1)$ et de $f^{-1}(\mathcal{O}_2)$.
- Si $f^{-1}(\mathcal{O}_1)$ et $f^{-1}(\mathcal{O}_2)$ ne recouvrent pas A, il existe un x dans A qui n'est dans aucun des deux. Dans ce cas, $f(x)$ est dans $f(A)$, mais n'est ni dans $\mathcal{O}_1$, ni dans $\mathcal{O}_2$, ce qui contredirait le fait que ces deux derniers recouvrent $f(A)$.

Nous déduisons que A n'est pas connexe. Et donc le lemme. □

Théorème 12.21 (Théorème des valeurs intermédiaires)**.**
Soit f, une fonction continue sur $[a, b]$, et supposons que $f(a) < f(b)$. Alors pour tout y tel que $f(a) \leqslant y \leqslant f(b)$, il existe un x entre a et b tel que $f(x) = y$.

Démonstration. Nous savons que $[a, b]$ est connexe parce que c'est un intervalle (proposition 6.119). Donc $f\big([a, b]\big)$ est connexe (lemme 12.20) et donc est un intervalle (à nouveau la proposition 6.119). Étant donné que $f\big([a, b]\big)$ est un intervalle, il contient toutes les valeurs intermédiaires entre n'importe quels deux de ses éléments. En particulier toutes les valeurs intermédiaires entre $f(a)$ et $f(b)$. □

Corollaire 12.22.
L'image d'un intervalle par une fonction continue est un intervalle.

Démonstration. Soit I un intervalle et $\alpha < \beta \in f(I)$ et $\gamma \in]\alpha, \beta[$. Nous considérons $a, b \in I$ tels que $\alpha = f(a)$ et $\beta = f(b)$. Par le théorème des valeurs intermédiaires12.21, il existe $t \in]a, b[$ tel que $f(t) = \gamma$. Par conséquent $\gamma \in f(I)$. □

Corollaire-définition 12.23 (Existence de la racine carré)**.**
Si $x \geqslant 0$ alors il existe un unique $y \geqslant 0$ tel que $y^2 = x$. Ce nombre est noté $\sqrt{x}$ et est nommé ***racine carré*** *de x.*

Démonstration. La fonction $f\colon t \mapsto t^2$ est continue et strictement croissante. Nous avons $f(0) = 0$ et [3] $f(x+1) > x$. Donc le théorème des valeurs intermédiaires 12.21 nous assure qu'il existe un unique $y \in [0, x+1]$ tel que $f(y) = x$. □

Nous avons déjà vu dans la proposition 2.32 que $\sqrt{2}$ était irrationnel. En fait le théorème suivant va nous montrer que le nombre $\sqrt{n}$ est soit entier, soit irrationnel.

Théorème 12.24.
Soit $n \in \mathbb{N}$. Le nombre $\sqrt{n}$ est rationnel si et seulement si n est un carré parfait.

Démonstration. Supposons que $\sqrt{n}$ soit rationnel. Le théorème 3.46 nous donne $p, q \in \mathbb{N}$ premiers entre eux tels que $\sqrt{n} = p/q$. La proposition 3.47 nous enseigne de plus que p^2 et q^2 sont premiers entre eux. Nous avons

$$p^2 = nq^2. \tag{12.12}$$

Le nombre q est alors un diviseur commun de q^2 et de p. Donc $q = 1$ et $n = p^2$ est un carré parfait. □

12.3.3 Continuité de la racine carré, invitation à la topologie induite

Pourquoi nous intéresser particulièrement à cette fonction ? Parce qu'elle a une sale condition d'existence : son domaine de définition n'est pas ouvert. Or dans tous les théorèmes de continuité d'approche topologique que nous avons vus, nous avons donné des contions *pour tout ouvert*. Nous nous attendons donc a avoir des difficultés avec la continuité de $\sqrt{x}$ en zéro.

Prenons I, n'importe quel intervalle ouvert dans $\mathbb{R}^+$, et voyons que la fonction

$$\begin{aligned} f\colon \mathbb{R}^+ &\to \mathbb{R}^+ \\ x &\mapsto \sqrt{x} \end{aligned} \tag{12.13}$$

est continue sur I. Remarque déjà que si I est un ouvert dans $\mathbb{R}^+$, il ne peut pas contenir zéro. Avant de nous lancer dans notre propos, nous prouvons un lemme qui fera tout le travail [4].

Lemme 12.25.
Soit $\mathcal{O}$, un ouvert dans $\mathbb{R}^+$. Alors $\mathcal{O}^2 = \{x^2 \text{ tel que } x \in \mathcal{O}\}$ est également ouvert .

Démonstration. Un élément de $\mathcal{O}^2$ s'écrit sous la forme x^2 pour un certain $x \in \mathcal{O}$. Le but est de trouver un ouvert autour de x^2 qui soit contenu dans $\mathcal{O}^2$. Étant donné que $\mathcal{O}$ est ouvert, on a une boule centrée en x contenue dans $\mathcal{O}$. Nous appelons δ le rayon de cette boule :

$$B(x, \delta) \subset \mathcal{O}.$$

Étant donné que cet ensemble est connexe, nous savons par le lemme 12.20 que $B(x,\delta)^2$ est également connexe (parce que la fonction $x \mapsto x^2$ est continue). Son plus grand élément est $(x+\delta)^2 = x^2 + \delta^2 + 2x\delta > x^2 + \delta^2$, et son plus petit élément est $(x-\delta)^2 = x^2 + \delta^2 - 2x\delta$.

Ce qui serait pas mal, c'est que ces deux bornes entourent x^2 ; de cette façon elles définiraient un ouvert autour de x^2 qui soit dans $\mathcal{O}^2$. Hélas, c'est pas gagné que $x^2 + \delta^2 - 2x\delta$ soit plus petit que x^2.

Heureusement, en fait c'est vrai parce que d'une part, du fait que $\mathcal{O} \subset \mathbb{R}^+$, on a $x > 0$, et d'autre part, pour que $\mathcal{O}$ soit positif, il faut que $\delta < x$. Donc on a évidement que $\delta < 2x$, et donc que

$$x^2 + \delta^2 - 2x\delta = x^2 + \delta \underbrace{(\delta - 2x)}_{<0} < x^2.$$

3. Faites deux cas suivant $x \geqslant 1$ ou non si vous le voulez, moi je prends $x + 1$.
4. C'est toujours ingrat d'être un lemme : on fait tout le travail et c'est toujours le théorème qui est nommé.

Donc nous avons fini : l'ensemble

$$B(x,\delta)^2 =]x^2+\delta^2-2x\delta, x^2+\delta^2+2x\delta[\subset \mathcal{O}^2$$

est un intervalle qui contient x^2, et donc qui contient une boule ouverte centrée en x^2.

□

Maintenant nous pouvons nous attaquer à la continuité de la racine carré sur tout ouvert positif en utilisant le théorème 12.19. Soit $\mathcal{O}$ n'importe quel ouvert de $\mathbb{R}$, et prouvons que $f|_I^{-1}(\mathcal{O})$ est ouvert. Par définition,

$$f|_I^{-1}(\mathcal{O}) = \{x \in I \text{ tel que } \sqrt{x} \in \mathcal{O}\}. \tag{12.14}$$

Maintenant c'est un tout petit effort que de remarquer que $f|_I^{-1}(\mathcal{O}) = \mathcal{O}^2 \cap I$. De là, on a gagné parce que $\mathcal{O}^2$ et I sont des ouverts. Or l'intersection de deux ouverts est ouvert.

Nous n'en avons pas fini avec la fonction $\sqrt{x}$. Nous avons la continuité de la racine carré pour tous les réels strictement positifs. Il reste à pouvoir dire que la fonction est continue en zéro malgré qu'elle ne soit pas définie sur un ouvert autour de zéro.

Il est possible de dire que la racine carré est continue en 0, malgré qu'elle ne soit pas définie sur un ouvert autour de 0…en tout cas pas un ouvert au sens que tu as en tête. Nous allons rentabiliser un bon coup notre travail sur les espaces métriques.

Nous pouvons définir la notion de boule ouverte sur n'importe quel espace métrique A en disant que

$$B(x,r) = \{y \in A \text{ tel que } d(x,y) < r\}.$$

Définition 12.26.
*Soit $f\colon A \to B$, une application entre deux espaces métriques. Nous disons que f est **continue** au point $a \in A$ si $\forall \delta > 0$, $\exists \epsilon > 0$ tel que*

$$f\big(B(a,\epsilon)\big) \subset B\big(f(a),\delta\big). \tag{12.15}$$

Tu reconnais évidement la condition (12.11). Nous l'avons juste recopiée. Tu remarqueras cependant que cette définition généralise immensément la continuité que l'on avait travaillé à propos des fonctions de $\mathbb{R}$ vers $\mathbb{R}$. Maintenant tu peux prendre n'importe quel espace métrique et c'est bon.

Nous n'allons pas faire un tour complet des conséquences et exemples de cette définition. Au lieu de cela, nous allons juste montrer en quoi cette définition règle le problème de la continuité de la racine carré en zéro.

La fonction que nous regardons est

$$\begin{aligned} f\colon \mathbb{R}^+ &\to \mathbb{R}^+ \\ x &\mapsto \sqrt{x}. \end{aligned} \tag{12.16}$$

Mais cette fois, nous ne la voyons pas comme étant une fonction dont le domaine est une partie de $\mathbb{R}$, mais comme fonction dont le domaine est $\mathbb{R}^+$ vu comme un espace métrique en soi. Quelles sont les boules ouvertes dans $\mathbb{R}^+$ autour de zéro ? Réponse : la boule ouverte de rayon r autour de zéro dans $\mathbb{R}^+$ est :

$$B(0,r)_{\mathbb{R}^+} = \{x \in \mathbb{R}^+ \text{ tel que } d(x,0) < r\} = [0,r[.$$

Cet intervalle est un ouvert. Aussi incroyable que cela puisse paraître !

Testons la continuité de la racine carré en zéro dans ce contexte. Il s'agit de prendre $A = \mathbb{R}^+$, $B = \mathbb{R}^+$ et $a = 0$ dans la définition 12.26. Nous avons que $B(\sqrt{0},\delta) = B(0,\delta) = [0,\delta[$ pour la topologie de $\mathbb{R}^+$.

Il s'agit maintenant de trouver un ϵ tel que $f\big(B(0,\epsilon)\big) \subset [0,\delta[$. Par définition, nous avons que

$$f\big(B(0,\epsilon)\big) = [0,\sqrt{\epsilon}[,$$

le problème revient dont à trouver ϵ tel que $\sqrt{\epsilon} \leqslant \delta$. Prendre $\epsilon < \delta^2$ fait l'affaire.

Donc voila. Au sens de la topologie propre à $\mathbb{R}^+$, nous pouvons dire que la fonction racine carré est partout continue.

12.3.4 Limites en des nombres

Si tu regardes la fonction $f(x) = 5x + 3$, tu ne serais pas étonnée si je te disais par exemple que

$$\lim_{x\to 10} f(x) = 53 \qquad \text{et } \lim_{x\to 0} f(x) = 3. \tag{12.17}$$

En effet, plus x est proche de 10, plus $f(x)$ est proche de 53 et plus x est proche de 0, plus $f(x)$ est proche de 3. Pas grand chose de neuf sous le Soleil.

Oui, mais l'intérêt d'introduire le concept de limite dans le cas de l'infini était qu'on ne peut pas bêtement calculer $f(\infty)$. Il fallait donc une astuce pour parler du comportement de f quand on s'approche de l'infini.

Nous posons la définition suivante.

Définition 12.27.
*Lorsque $a \in \mathbb{R}$, on dit que la fonction f **tend vers l'infini quand** x **tend vers** a si*

$$\forall M \in \mathbb{R}, \exists \delta \text{ tel que } (|x - a| \leqslant \delta) \Rightarrow f(x) \geqslant M \text{ quand } x \in \operatorname{dom} f.$$

Cela signifie que l'on demande que dès que x est assez proche de a (c'est à dire dès que $|x - a| \leqslant \delta$), alors $f(x)$ est plus grand que M, et que l'on peut trouver un δ qui fait ça pour n'importe quel M. Une autre façon de le dire est que pour toute hauteur M, on peut trouver un intervalle de largeur δ autour de a [5] tel que sur cet intervalle, la fonction f est toujours plus grande que M.

Montrons sur un dessin pourquoi je disais que la fonction $x \to 1/x$ n'est pas de ce type.

Le problème est qu'il n'existe par exemple aucun intervalle autour de 0 sur lequel f serait toujours plus grande que 10. En effet n'importe quel intervalle autour de 0 contient au moins un nombre négatif. Or quand x est négatif, f n'est certainement pas plus grande que 10. Nous y reviendrons.

Pour l'instant, montrons que la fonction $f(x) = 1/x^2$ est une fonction qui vérifie la définition 12.27. Avant de prendre n'importe quel M, prenons par exemple 100. Nous avons besoin d'un intervalle autour de zéro sur lequel f est toujours plus grande que 100. C'est vite vu que $f(0.1) = f(-0.1) = 100$, donc l'intervalle $[-\frac{1}{10}, \frac{1}{10}]$ est le bon. Partout dans cet intervalle, f est plus grande que 100. Partout ? Ben non : en $x = 0$, la fonction n'est même pas définie, donc c'est un peu dur de dire qu'elle est plus grande que 100. C'est pour cela que nous avons ajouté la condition « quand $x \in \operatorname{dom} f$ » dans la définition de la limite.

Prenons maintenant un $M \in \mathbb{R}$ arbitraire, et trouvons un intervalle autour de 0 sur lequel f est toujours plus grande que M. La réponse est évidement l'intervalle de largeur $1/\sqrt{M}$, c'est à dire

$$\left[-\frac{1}{\sqrt{M}}, \frac{1}{\sqrt{M}}\right].$$

12.3.5 Limites quand tout va bien

D'abord définissons ce qu'on entend par la limite d'une fonction en un point quand il n'y a aucun infini en jeu.

Définition 12.28.
*On dit que la fonction f **tend vers** b **quand** x **tend vers** a si*

$$\forall \epsilon > 0, \exists \delta \text{ tel que } (|x - a| \leqslant \delta) \Rightarrow |f(x) - b| \leqslant \epsilon \text{ quand } x \in \operatorname{dom} f.$$

Dans ce cas, nous notons

$$\lim_{x\to a} f(x) = b. \tag{12.18}$$

5. C'est à dire un intervalle de la forme $[a - \delta, a + \delta]$.

Commençons par un exemple très simple : prouvons que $\lim_{x\to 0} x = 0$. C'est donc $a = b = 0$ dans la définition. Prenons $\epsilon > 0$, et trouvons un intervalle autour de zéro tel que partout dans l'intervalle, $x \leqslant \epsilon$. Bon ben c'est clair que $\delta = \epsilon$ fonctionne.

Plus compliqué maintenant, mais toujours sans surprises.

Proposition 12.29.

$$\lim_{x\to 0} x^2 = 0.$$

Démonstration. Soit $\epsilon > 0$. On veut un intervalle de largeur δ autour de zéro tel que x^2 soit plus petit que ϵ sur cet intervalle. Cette fois-ci, le δ qui fonctionne est $\delta = \sqrt{\epsilon}$. En effet un élément de l'intervalle $[-\delta, \delta]$ est un r de valeur absolue plus petite ou égale à δ :

$$|r| \leqslant \delta = \sqrt{\epsilon}.$$

En prenant le carré de cette inégalité on a :

$$r^2 \leqslant \epsilon,$$

ce qu'il fallait prouver. □

Calculer et prouver des valeurs de limites, mêmes très simples, devient vite de l'arrachage de cheveux à essayer de trouver le bon δ en fonction de ϵ si on n'a pas quelques théorèmes généraux. Nous allons donc maintenant en prouver quelque-uns.

Théorème 12.30.
Si

$$\lim_{x\to a} f(x) = b, \tag{12.19}$$

alors

$$\lim_{x\to a} (\lambda f)(x) = \lambda b \tag{12.20}$$

pour n'importe quel $\lambda \in \mathbb{R}$.

Démonstration. Soit $\epsilon > 0$. Afin de prouver la propriété (12.20), il faut trouver un δ tel que pour tout x dans $[a - \delta, a + \delta]$, on ait $|(\lambda f)(x) - \lambda b| \leqslant \epsilon$. Cette dernière inégalité est équivalente à $|\lambda||f(x) - b| \leqslant \epsilon$. Nous devons donc trouver un δ tel que

$$|f(x) - b| \leqslant \frac{\epsilon}{|\lambda|}. \tag{12.21}$$

soit vraie pour tout x dans $[a - \delta, a + \delta]$. Mais l'hypothèse (12.19) dit précisément qu'il existe un δ tel que pour tout x dans $[a - \delta, a + \delta]$ on ait cette inégalité. □

Théorème 12.31.
Si

$$\lim_{x\to a} f(x) = b_1 \tag{12.22a}$$

$$\lim_{x\to a} g(x) = b_2, \tag{12.22b}$$

alors

$$\lim_{x\to a} (f + g)(x) = b_1 + b_2. \tag{12.23}$$

Démonstration. Soit $\epsilon > 0$. Par hypothèse, il existe δ_1 tel que

$$|f(x) - b_1| \leqslant \frac{\epsilon}{2} \tag{12.24}$$

dès que $|x-a| \leqslant \delta_1$. Il existe aussi δ_2 tel que

$$|g(x) - b_2| \leqslant \frac{\epsilon}{2}. \tag{12.25}$$

dès que $|x-a| \leqslant \delta_2$. Tu notes l'astuce de prendre $\epsilon/2$ dans la définition de limite pour f et g. Maintenant, ce qu'on voudrait c'est un δ tel que l'on ait $|(f+g)(x)-(b_1+b_2)| \leqslant \epsilon$ dès que $|x-a| \leqslant \delta$. Moi je dit que $\delta = \min\{\delta_1, \delta_2\}$ fonctionne. En effet, en utilisant l'inégalité $|a+b| \leqslant |a| + |b|$, nous trouvons :

$$|(f+g)(x) - (b_1+b_2)| = |(f(x)-b_1) + (g(x)-b_2)| \leqslant |f(x)-b_1| + |g(x)-b_2|. \tag{12.26}$$

Comme on suppose que $|x-a| \leqslant \delta$, on a évidement $|x-a| \leqslant \delta_1$, et donc l'équation (12.24) tient. Mais si $|x-a| \leqslant \delta$, on a aussi $|x-a| \leqslant \delta_2$, et donc l'équation (12.24) tient également. Chacun des deux termes de (12.26) est donc plus petits que $\epsilon/2$, et donc le tout est plus petit que ϵ, ce qu'il fallait montrer.

□

Une formule qui résume ces deux théorèmes est que

$$\lim_{x\to a}\big[\alpha f(x) + \beta g(x)\big] = \alpha \lim_{x\to a} f(x) + \beta \lim_{x\to a} g(x). \tag{12.27}$$

Lemme 12.32.
Si $\lim_{x\to a} f(x) = b$ avec a, $b \in \mathbb{R}$, alors il existe un $\delta > 0$ et un $M > 0$ tels que

$$(|x-a| \leqslant \delta) \Rightarrow |f(x)| \leqslant M.$$

Ce que signifie ce lemme, c'est que quand la fonction f admet une limite finie en un point, alors il est possible de majorer la fonction sur un intervalle autour du point.

Démonstration. Cela va être démontré par l'absurde. Supposons qu'il n'existe pas de δ ni de M qui vérifient la condition. Dans ce cas, pour tout δ et pour tout M, il existe un x tel que $|x-a| \leqslant \delta$ et $|f(x)| > M$. Cela est valable pour tout M, donc prenons par exemple $b + 1000$. Donc

$$\forall \delta > 0, \exists x \text{ tel que } |x-a| \leqslant \delta \text{ et } |f(x)| > b + 1000. \tag{12.28}$$

Cela signifie qu'aucun δ ne peut convenir dans la définition de $\lim_{x\to a} f(x) = b$, ce qui contredit les hypothèses. □

Dans le même ordre d'idée, on peut prouver que si la limite de la fonction en un point est positive, alors elle est positive autour ce ce point. Plus précisément, nous avons la

Proposition 12.33.
Si f est une fonction telle que $\lim_{x\to a} f(x) > 0$, alors il existe un voisinage de a sur lequel f est positive.

Démonstration. Supposons que $\lim_{x\to a} f(x) = y_0$. Par la définition de la limite fait que si pour tout x dans un voisinage autour de a, on ait $|f(x)-a| < \epsilon$. Cela est valable pour tout ϵ, pourvu que le voisinage soit assez petit. Si je choisit un voisinage pour lequel $|f(x)-a| < \frac{y_0}{2}$, alors sur ce voisinage, f est positive. □

Théorème 12.34.
Si

$$\lim_{x\to a} f(x) = b_1 \qquad \text{et} \qquad \lim_{x\to a} g(x) = b_2, \tag{12.29}$$

alors

$$\lim_{x\to a}(fg)(x) = b_1 b_2. \tag{12.30}$$

Démonstration. Soit $\epsilon > 0$, et tentons de trouver un δ tel que $|f(x)g(x) - b_1b_2| \leqslant \epsilon$ dès que $|x - a| \leqslant \delta$. Nous avons

$$\begin{aligned} |f(x)g(x) - b_1b_2| &= |f(x)g(x) - b_1b_2 + f(x)b_2 - f(x)b_2| \\ &= \big|f(x)\big(g(x) - b_2\big) + b_2\big(f(x) - b_1\big)\big| \\ &\leqslant \big|f(x)\big(g(x) - b_2\big)\big| + \big|b_2\big(f(x) - b_1\big)\big| \\ &= |f(x)||g(x) - b_2| + |b_2||f(x) - b_1|. \end{aligned} \tag{12.31}$$

À la première ligne se trouve la subtilité de la démonstration : on ajoute et on enlève[6] $f(x)b_2$. Maintenant nous savons par le lemme 12.32 que pour un certain δ_1, la quantité $|f(x)|$ peut être majoré par un certain M dès que $|x-a| \leqslant \delta_1$. Prenons donc un tel δ_1 et supposons que $|x-a| \leqslant \delta_1$. Nous savons aussi que pour n'importe quel choix de ϵ_2 et ϵ_3, il existe des nombres δ_2 et δ_3 tels que $|f(x) - b_1| \leqslant \epsilon_2$ et $|g(x) - b_1| \leqslant \epsilon_3$ dès que $|x - a| \leqslant \delta_2$ et $|x - a| \leqslant \delta_3$. Dans ces conditions, la dernière expression (12.31) se réduit à

$$|f(x)g(x) - b_1b_2| \leqslant M\epsilon_2 + |b_2|\epsilon_3. \tag{12.32}$$

Pour terminer la preuve, il suffit de choisir ϵ_2 et ϵ_3 tels que $M\epsilon_2 + |b_2|\epsilon_3 \leqslant \epsilon$, et puis prendre $\delta = \min\{\delta_1, \delta_2, \delta_3\}$.

Remetons les choses dans l'ordre. L'on se donne ϵ au départ. La première chose est de trouver un δ_1 qui permet de majorer $|f(x)|$ par M selon le lemme 12.32, et puis choisissons ϵ_2 et ϵ_3 tels que $M\epsilon_2 + |b_2|\epsilon_3 \leqslant \epsilon$. Ensuite nous prenons, en vertu des hypothèses de limites pour f et g, les nombres δ_2 et δ_3 tels que $|f(x) - b_1| \leqslant \epsilon_2$ et $|g(x) - b_2| \leqslant \epsilon_3$ dès que $|x - a| \leqslant \delta_2$ et $|x - a| \leqslant \delta_3$.

Si avec tous ça on prend $\delta = \min\{\delta_1, \delta_2, \delta_3\}$, alors la majoration et les deux inégalités sont valables en même temps et au final

$$|f(x)g(x) - b_1b_2| \leqslant M\epsilon_2 + b_2\epsilon_3 \leqslant \epsilon,$$

ce qu'il fallait prouver.

□

À l'aide de ces petits résultats, nous pouvons déjà calculer pas mal de limites. Nous pouvons déjà par exemple calculer les limites de tous les polynomes en tous les nombrs réels. En effet, nous savons la limite de la fonction $f(x) = x$. la fonction $x \mapsto x^2$ n'est rien d'autre que le produit de f par elle-même. Donc

$$\lim_{x\to a} x^2 = \big(\lim_{x\to a} x\big) \cdot \big(\lim_{x\to a} x\big) = a^2.$$

De la même façon, nous trouvons facilement que

$$\lim_{x\to a} x^n = a^n. \tag{12.33}$$

Théorème 12.35 (Limite et continuité)**.**
La fonction f est continue au point a si et seulement si $\lim_{x\to a} f(x) = f(a)$.

Démonstration. Nous commençons par supposer que f est continue en a, et nous prouvons que $\lim_{x\to a} f(x) = a$. Soit $\epsilon > 0$; ce qu'il nous faut c'est un δ tel que $|x-a| \leqslant \delta$ implique $|f(x)-f(a)| \leqslant \epsilon$. La définition 12.14 de la continuité donne l'existence d'un δ comme il nous faut.

Dans l'autre sens, c'est à dire prouver que f est continue au point a sous l'hypothèse que $\lim_{x\to a} f(x) = f(a)$, la preuve se fait de la même façon. □

Nous en déduisons que si nous voulons gagner quelque chose à parler de limites, il faut prendre des fonctions non continues. Prenons une fonction qui fait un saut. Pour se fixer les idées, prenons celle-ci :

$$f(x) = \begin{cases} 2x & \text{si } x \in]\infty, 2[\\ x/2 & \text{si } x \in [2, \infty[\end{cases} \tag{12.34}$$

6. Comme exercice, tu peux essayer de refaire la démonstration en ajoutant et enlevant $g(x)b_1$ à la place.

Essayons de trouver la limite de cette fonction lorsque x tend vers 2. Étant donné que f n'est pas continue en 2, nous savons déjà que $\lim_{x\to 2} f(x) \neq f(2)$. Donc ce n'est pas 1. Cette limite ne peut pas valoir 4 non plus parce que si je prends n'importe quel ϵ, la valeur de $f(2+\epsilon)$ est très proche de 2, et donc ne peut pas s'approcher de 4. En fait, tu peux facilement vérifier que *aucun nombre ne vérifie la condition de limite pour f en* 2. Nous disons que la limite n'existe pas.

Pour résumer, les limites qui ne font pas intervenir l'infini ne servent à rien parce que

— si la fonction est continue, la limite est simplement la valeur de la fonction par le théorème 12.35,

— si la fonction fait un saut, alors la limite n'existe pas (nous n'avons pas prouvé cela en général, mais avoue que l'exemple est convainquant).

Nous avons même la proposition suivante :

Proposition 12.36.
Si f existe en a (c'est à dire si $a \in \text{dom}(f)$) et si $\lim_{x\to a} f(x) = b$, alors $f(a) = b$.

Démonstration. Du fait que $\lim_{x\to a} f(x) = b$, il découle que pour tout ϵ, il existe un δ tel que $|x-a| \leqslant \delta$ implique $|f(x)-b| \leqslant \epsilon$. Il est évident que pour tout δ, $|a-x| \leqslant \delta$, donc nous avons que

$$|f(a) - b| \leqslant \epsilon$$

pour tout ϵ. Cela implique que $f(a) = b$. □

Notons toutefois que l'inverse de cette proposition n'est pas vraie : la fonction (12.34) donne justement une fonction qui prend la valeur 1 en 2 sans que la limite en 2 soit 1. Quoi qu'il en soit, cette proposition achève de nous convaincre de l'inutilité d'étudier d'étudier les limites sans infinis : dès qu'on a une limite, à tous les coups c'est la valeur de la fonction ... heu ... en es-tu bien sûr ?

Proposition 12.37 ([1])**.**
Soit $f\colon \mathbb{R}^2 \to \mathbb{R}$ une application continue dont la variable y varie dans un compact I de $\mathbb{R}$. Alors la fonction

$$\begin{aligned} d\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \sup_{y\in I} f(x,y) \end{aligned} \tag{12.35}$$

est continue.

Démonstration. Soit x_0 fixé et prouvons que d est continue en x_0. Nous notons y_0 la valeur de y qui réalise le maximum (par le théorème 6.126 et le fait que les fonctions projection soient continues, lemme 9.96). Soit aussi $\epsilon > 0$ tellement fixé que même avec un tourne vis hydraulique, il ne bougerait pas. Nous considérons δ tel que si $\|(x,y)-(x_0,y_0)\| \leqslant \delta$ alors $\|f(x,y)-f(x_0,y_0)\| < \epsilon$.

Si $|x-x_0| < \delta$ alors pour y assez proche de y_0 nous avons $\|(x,y)-(x_0,y_0)\| \leqslant \delta$, et donc $\|f(x,y)-f(x_0,y_0)\| \leqslant \epsilon$. Cela montre qu'il existe δ tel que $|x-x_0| \leqslant \delta$ implique $d(x) \geqslant d(x_0) - \epsilon$.

Nous devons encore trouver un δ tel que si $|x-x_0| \leqslant \delta$ alors $d(x) \leqslant d(x_0) + \epsilon$. Supposons que non. Alors pour tout δ il existe un x tel que $|x-x_0| \leqslant \delta$ et $d(x) > d(x_0) + \epsilon$. Cela nous donne une suite $x_i \to x_0$.

Pour chaque x_i nous notons y_i la valeur de y qui réalise le supremum correspondant. La suite (y_i) étant contenue dans un compact nous supposons prendre une sous-suite de (x_i) telle que la suite (y_i) converge. Nous nommons a la limite (et non y_0 parce que nous ne savons pas si $y_i \to y_0$). Pour chaque i nous avons

$$f(x_i, y_i) > \sup_{y\in I} f(x_0, y) + \epsilon. \tag{12.36}$$

En prenant la limite et en utilisant la continuité de f,

$$f(x_0, a) > \sup_{y\in I} f(x_0, y) + \epsilon, \tag{12.37}$$

ce qui est impossible. □

12.3.6 Discussion avec mon ordinateur

Voici un extrait de ce peut donner Sage. Nous lui donnons la fonction

$$f(x) = \frac{x+4}{3x^2+10x-8}. \tag{12.38}$$

Cette fonction est faite exprès pour que le dénominateur s'annule en -4. En fait $3x^2+10x-8 = (x+4)(3x-2)$, et la fraction peut se simplifier en

$$f(x) = \frac{1}{3x-2}. \tag{12.39}$$

Et avec cela nous écririons $f(-4) = -\frac{1}{14}$. Voyons comment cela passe dans Sage.

```
----------------------------------------------------------------------
| Sage Version 5.2, Release Date: 2012-07-25                         |
| Type "notebook()" for the browser-based notebook interface.        |
| Type "help()" for help.                                            |
----------------------------------------------------------------------
sage: f(x)=(x+4)/(3*x**2+10*x-8)
sage: f(-4)
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
ValueError: power::eval(): division by zero
```

Il produit donc une erreur de division par zéro. Cela n'est pas étonnant. Pourtant si on lui demande, il est capable de simplifier. En effet :

```
sage: f.simplify_full()
x |--> 1/(3*x - 2)
sage: f.simplify_full()(-4)
-1/14
```

12.3.7 Limites et prolongement

La proposition 12.36 a une terrible limitation : il faut que la fonction existe au point considéré. Or en regardant bien la définition 12.28, nous remarquons que $\lim_{x\to a} f(x)$ peut très bien exister sans que $f(a)$ n'existe.

Reprenons l'exemple de la fonction (12.38) que mon ordinateur refusait de calculer en zéro :

$$f(x) = \frac{x+4}{3x^2+10x-8} = \frac{x+4}{(x+4)\left(x-\frac{2}{3}\right)}. \tag{12.40}$$

Cette fonction a une condition d'existence en $x=-4$. Et pourtant, tant que $x \neq 4$, cela a un sens de simplifier les $(x+4)$ et d'écrire

$$f(x) = \frac{1}{x-\frac{2}{3}} = \frac{3}{3x-2}.$$

Étant donné que pour toute valeur de x différente de -4, la fonction f s'exprime de cette façon, nous avons que

$$\lim_{x\to-4} f(x) = \lim_{x\to-4}\left(\frac{3}{3x-2}\right).$$

Oui, mais la fonction[7] $g(x) = 3/(3x-2)$ est continue en -4 et donc sa limite vaut sa valeur. Nous en déduisons que

$$\lim_{x\to-4} f(x) = -\frac{3}{14}.$$

7. Cette fonction g n'est pas f parce que g a en plus l'avantage d'être définie en -4.

Que dire maintenant de la fonction ainsi définie ?

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \neq -4 \\ -3/14 & \text{si } x = -4. \end{cases} \tag{12.41}$$

Cette fonction est continue en -4 parce qu'elle y est égale à sa limite. Les étapes suivies pour obtenir ce résultat sont :

— Repérer un point où la fonction n'existe pas,
— calculer la limite de la fonction en ce point, et en particulier vérifier que cette limite existe, ce qui n'est pas toujours le cas,
— définir une nouvelle fonction qui vaut partout la même chose que la fonction originale, sauf au point considéré où l'on met la valeur de la limite.

C'est ce qu'on appelle **prolonger la fonction par continuité** parce que la fonction résultante est continue. La prolongation de f par continuité est donc en général définie par

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } f(x) \\ \lim_{y\to x} f(y) & \text{si } f(x) \end{cases} \tag{12.42}$$

Dans le cas que nous regardions,

$$f(x) = \frac{x+4}{3x^2+10x-8},$$

le prolongement par continuité est donné par

$$\tilde{f} = \frac{3}{3x-2}. \tag{12.43}$$

Remarque que cette fonction n'est toujours pas définie en $x = 2/3$.

12.3.8 Théorèmes des valeurs intermédiaires

Théorème 12.38 (Théorème des valeurs intermédiaires)**.**
Si f est continue sur un intervalle $I = [a, b]$ avec $f(a) \neq f(b)$ alors pour tout t entre $f(a)$ et $f(b)$, il existe $x \in I$ tel que $f(x) = t$.

Nous considérons la question suivante : étant donné une fonction f définie sur $I\backslash\{x_0\}$, est-il possible de définir f en x_0 de telles façon à ce qu'elle soit continue ?

Exemple 12.39
La fonction

$$\begin{aligned} f\colon \mathbb{R}\backslash\{0\} &\to \mathbb{R} \\ x &\mapsto \frac{1}{x} \end{aligned} \tag{12.44}$$

n'est pas définie pour $x = 0$ et il n'y a pas moyen de définir $f(0)$ de telle sorte que f soit continue parce que $\lim_{x\to 0} \frac{1}{x}$ n'existe pas. △

12.3.9 Théorème de la bijection

Proposition 12.40.
Soit une bijection $f\colon I \to J$ et $f^{-1}\colon J \to I$ sa réciproque. Alors pour tout $x_0 \in I$ nous avons

$$f^{-1}\big(f(x_0)\big) = x_0 \tag{12.45}$$

et pour tout $y_0 \in J$ nous avons

$$f\big(f^{-1}(y_0)\big) = y_0. \tag{12.46}$$

Démonstration. Nous prouvons la relation (12.45) et nous laissons (12.46) comme exercice au lecteur.

Soit $x_0 \in I$, et posons $y_0 = f(x_0)$. La définition de l'application réciproque est que pour $y \in J$, $f^{-1}(y)$ est l'unique élément x de I tel que $f(x) = y$. Donc $f^{-1}(y_0)$ est l'unique élément de I dont l'image est y_0. C'est donc x_0 et nous avons $f^{-1}(y_0) = x_0$, c'est à dire

$$f^{-1}\big(f(x_0)\big) = x_0. \tag{12.47}$$

□

Théorème 12.41 (Théorème de la bijection)**.**
Soit I un intervalle et f une fonction continue et strictement monotone de I dans $\mathbb{R}$. Nous avons alors :

(1) $f(I)$ est un intervalle de $\mathbb{R}$;

(2) La fonction $f\colon I \to f(I)$ est bijective

(3) La fonction $f^{-1}\colon f(I) \to I$ est strictement monotone de même sens que f ;

(4) La fonction $f\colon I \to f(I)$ est un homéomorphisme, c'est-à-dire que $f^{-1}\colon f(I) \to I$ est continue.

Démonstration. Prouvons les choses point par point.

(1) Supposons pour fixer les idées que f est monotone croissante[8].
Soient $a < b$ dans $f(I)$. Par définition il existe $x_1, x_2 \in I$ tels que $a = f(x_1)$ et $b = f(x_2)$. La fonction f est continue sur l'intervalle $[x_1, x_2]$ et vérifie $f(x_1) < f(x_2)$. Donc le théorème des valeurs intermédiaires 12.38 nous dit que pour tout t dans $[f(x_2), f(x_2)]$, il existe un $x_0 \in [x_1, x_2]$ tel que $f(x_0) = t$. Cela montre que toutes les valeurs intermédiaires entre a et b sont atteintes par f et donc que $f(I)$ est un intervalle.

(2) Nous prouvons maintenant que f est bijective en prouvant séparément qu'elle est surjective et injective.

f est surjective Une fonction est toujours surjective depuis un intervalle I vers l'ensemble $\Im f$.

f est injective Soit $x \neq y$ dans I ; pour fixer les idées nous supposons que $x < y$. La stricte monotonie de f implique que $f(x) < f(y)$ ou que $f(x) > f(y)$. Dans tous les cas $f(x) \neq f(y)$.

La fonction f est donc bijective.

(3) Comme d'accoutumée nous supposons que f est croissante. Soient $y_1 < y_2$ dans $f(I)$; nous devons prouver que $f^{-1}(y_1) \leqslant f^{-1}(y_2)$. Pour cela nous considérons les nombres $x_1, x_2 \in I$ tels que $f(x_1) = y_1$ et $f(x_2) = y_2$. Nous allons en prouver la contraposée en supposant que $f^{-1}(y_1) > f^{-1}(y_2)$. En appliquant f (qui est croissante) à cette dernière inégalité il vient

$$f\big(f^{-1}(y_1)\big) \geqslant f\big(f^{-1}(y_2)\big), \tag{12.48}$$

ce qui signifie

$$y_1 \geqslant y_2 \tag{12.49}$$

par l'équation (12.46).

(4) La fonction $f^{-1}\colon f(I) \to I$ est une fonction monotone et surjective, donc continue par la proposition 12.10.

□

8. Traitez en tant que exercice le cas où f est décroissante.

Proposition 12.42 ([137]).
Soit $f\colon I \to J = f(I)$ une fonction bijective, continue et dérivable[9]*. Soit $x_0 \in I$ et $y_0 = f(x_0)$. Si $f'(x_0) \neq 0$ alors la fonction réciproque f^{-1} est dérivable en y_0 et sa dérivée est donnée par*

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}. \tag{12.50}$$

Démonstration. Prouvons que f^{-1} est dérivable au point $b = f(a) \in J$. Étant donné que f est dérivable en a, nous avons

$$f'(a) = \lim_{x\to a} \frac{f(x) - f(a)}{x - a}. \tag{12.51}$$

Par ailleurs, étant donnée la continuité de f^{-1} donnée par la proposition 12.41(4), nous avons

$$\lim_{\epsilon\to 0} f^{-1}(b + \epsilon) = f^{-1}(b) = a. \tag{12.52}$$

Nous pouvons donc remplacer dans (12.51) tous les x par $f^{-1}(b + \epsilon)$ et prendre la limite $\epsilon \to 0$ au lieu de $x \to a$:

$$\begin{aligned} f'(a) &= \lim_{\epsilon\to 0} \frac{f\big(f^{-1}(b + \epsilon)\big) - f(a)}{f^{-1}(b + \epsilon) - a} \\ &= \lim_{\epsilon\to 0} \frac{b + \epsilon - f(a)}{f^{-1}(b + \epsilon) - f^{-1}(b)} \\ &= \lim_{\epsilon\to 0} \frac{\epsilon}{f^{-1}(b + \epsilon) - f^{-1}(b)} \\ &= \frac{1}{\lim_{\epsilon\to 0} \frac{f^{-1}(b+\epsilon)-f^{-1}(b)}{\epsilon}} \\ &= \frac{1}{(f^{-1})'(b)}. \end{aligned} \tag{12.53}$$

Nous avons utilisé le fait que $f(a) = b$ et $a = f^{-1}(b)$. □

12.43.
Très souvent on préfère retenir la formule

$$(f^{-1})'(y_0) = \frac{1}{f'\left((f^{-1})(y_0)\right)} \tag{12.54}$$

Elle est très simple à retrouver : il suffit d'écrire

$$f^{-1}\big(f(x)\big) = x \tag{12.55}$$

puis de dériver les deux côtés par rapport à x en utilisant la règle de dérivation des fonctions composées :

$$(f^{-1})'\big(f(x)\big)f'(x) = 1. \tag{12.56}$$

Exemple 12.44(Difféomorphisme entre $\mathbb{R}$ et un ouvert borné)
Nous cherchons à construire une application dérivable et d'inverse dérivable entre $\mathbb{R}$ (en entier) et un ouvert borné de $\mathbb{R}$. Il serait tentant de prendre l'application arc tangente

$$\begin{aligned} \arctan\colon \mathbb{R} &\to \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\\ x &\mapsto \arctan(x) \end{aligned}, \tag{12.57}$$

mais elle ne sera définie que dans le théorème 16.38.

Nous posons

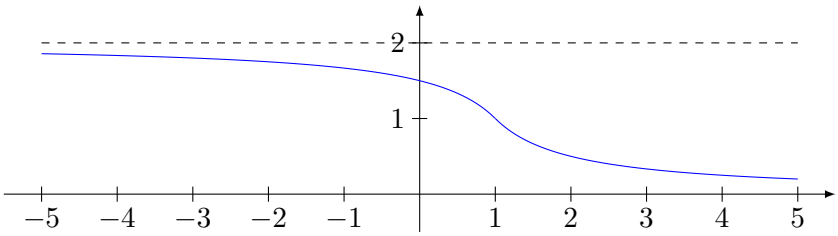
$$f(x) = \begin{cases} 2 + \frac{1}{x-2} & \text{si } x \leqslant 1 \\ \frac{1}{x} & \text{si } x > 1. \end{cases} \tag{12.58}$$

9. Pour rappel, une fonction dérivable est toujours continue ; l'hypothèse de continuité n'est pas nécessaire

Cela est continue en $x = 1$: il suffit de calculer les deux valeurs. En ce qui concerne la dérivabilité en $x = 1$, nous devons faire

$$\lim_{\epsilon\to 0} \frac{f(1+\epsilon) - f(1)}{\epsilon}. \tag{12.59}$$

La limite à gauche est égale à la dérivée de $x \mapsto 2 + \frac{1}{x-2}$ en $x = 1$ et la limite à droite est égale à la dérivée de $x \mapsto 1/x$ en $x = 1$. Dans les deux cas nous trouvons -1.



Nous voyons vite que cette fonction est strictement décroissante ; et un calcul de limite nous dit qu'il s'agit d'une bijection dérivable

$$f\colon \mathbb{R} \to]0, 2[. \tag{12.60}$$

La proposition 12.42 s'applique et la bijection réciproque est également dérivable (donc continue aussi). △

Problèmes et choses à faire

Si vous connaissez un autre exemple, plus simple, de difféomorphisme $f\colon \mathbb{R} \to]a, b[$, faites-le moi savoir. Ne pas utiliser d'exponentielle (vous pensiez à bricoler quelque chose à partir de la primitive de $x \mapsto e^{-x^2}$?) ni de fonctions trigonométriques.

Exemple 12.45
Nous aimerions donner le logarithme comme exemple, mais l'exponentielle ne sera définie que dans longtemps à partir des séries entières. Allez voir l'exemple 15.58 pour le logarithme comme inverse de l'exponentielle. △

12.3.10 Prolongement par continuité

Définition 12.46 (Prolongement par continuité)**.**
Soit $f\colon I\backslash\{x_0\} \to \mathbb{R}$ telle que $\lim_{x\to x_0} f(x) = \ell$. La fonction

$$\begin{aligned} &\tilde{f}\colon I \to \mathbb{R} \\ &\tilde{f}(x) = \begin{cases} f(x) & si\ x \neq x_0 \\ \ell & si\ x = x_0 \end{cases} \end{aligned} \tag{12.61}$$

*est une fonction continue sur I et est appelée le **prolongement par continuité** de f en x_0.*

Exemple 12.47
La fonction $f(x) = x\ln(|x|)$ n'est pas définie en $x = 0$. Cependant

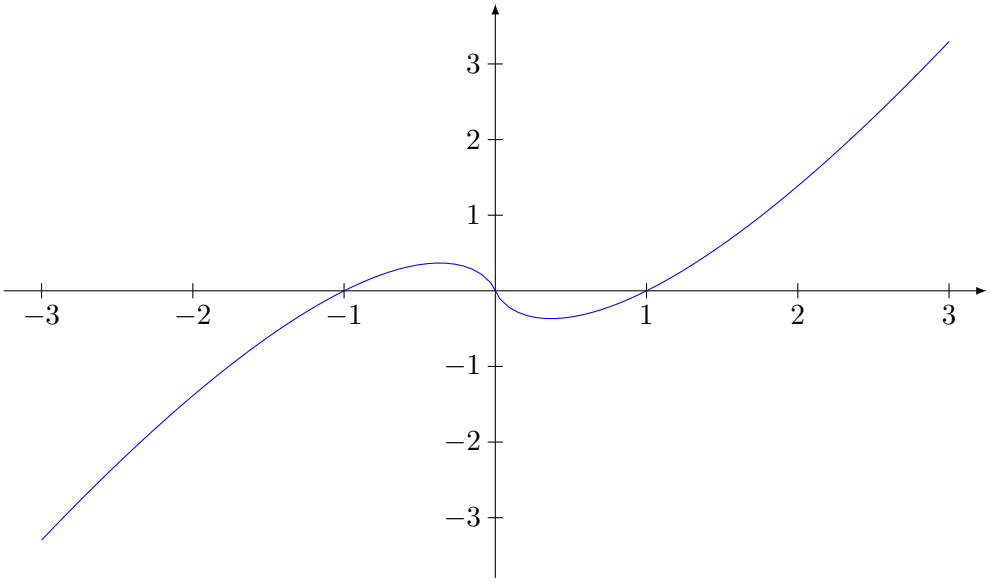
$$\lim_{x\to 0} x\ln(|x|) = 0. \tag{12.62}$$

Nous pouvons donc définir la fonction

$$\begin{aligned} &\tilde{f}\colon \mathbb{R} \to \mathbb{R} \\ &x \mapsto \begin{cases} x\ln(|x|) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0. \end{cases} \end{aligned} \tag{12.63}$$

Contrairement à la fonction initiale f, cette fonction $\tilde{f}$ est définie et continue en 0.

Notez que sur le graphe de la fonction $\tilde{f}$, la courbe est bien régulière en $x = 0$.



△

Exemple 12.48
La fonction

$$\begin{aligned} f\colon \mathbb{R}\backslash\{-3,2\} &\to \mathbb{R} \\ x &\mapsto \frac{x^2+2x-3}{(x+3)(x-2)} \end{aligned} \tag{12.64}$$

admet pour limite $\lim_{x\to -3} f(x) = \frac{4}{5}$. Son prolongement par continuité en $x=-3$ est donné par

$$\tilde{f}(x) = \frac{x-1}{x-2}. \tag{12.65}$$

Notons que les fonctions f et $\tilde{f}$ ne sont pas identiques : l'une est définie pour $x=-3$ et l'autre pas. Lorsqu'on fait le calcul

$$\frac{x^2+2x-3}{(x+3)(x-2)} = \frac{(x-1)(x+3)}{(x+3)(x-2)} = \frac{x-1}{x-2}, \tag{12.66}$$

la simplification n'est pas du tout un acte anodin. Le dernier signe «=» est discutable parce que les deux dernières expressions ne sont pas égales pour tout x ; elles ne sont égales «que» pour les x pour lesquels les deux expressions existent. △

Les fonctions trigonométriques donneront quelque exemples intéressants de prolongements par continuité. Voir l'exemple 16.32.

12.4 Espace des fonctions continues

Définition 12.49.
*Soit I, un intervalle de $\mathbb{R}$. L'**oscillation** sur I est le nombre*

$$\omega_f(I) = \sup_{x\in I} f(x) - \inf_{x\in I} f(x). \tag{12.67}$$

Pour chaque x fixé, la fonction

$$x \mapsto \omega_f\big(B(x,\delta)\big) \tag{12.68}$$

est une fonction positive, croissante et a donc une limite (pour $\delta \to 0$). Nous notons $\omega_f(x)$ cette limite qui est l'**oscillation** de f en ce point. Une propriété immédiate est que f est continue en x_0 si et seulement si $\omega_f(x_0) = 0$.

Lemme 12.50.
L'ensemble des points de discontinuité d'une fonction $f\colon \mathbb{R} \to \mathbb{R}$ est une réunion dénombrable de fermés.

Démonstration. Soit D l'ensemble des points de discontinuité de f. Nous avons

$$D = \bigcup_{n=1}^{\infty} \{x \text{ tel que } \omega_f(x) \geqslant \frac{1}{n}\}. \tag{12.69}$$

Il nous suffit donc de montrer que pour tout ϵ, l'ensemble

$$\{x \text{ tel que } \omega_f(x) < \epsilon\} \tag{12.70}$$

est ouvert. Soit en effet x_0 dans cet ensemble. Il existe δ tel que $\omega_f\big(B(x_0,\delta)\big) < \epsilon$. Si $x \in B(x_0,\delta)$, alors si on choisit δ' tel que $B(x,\delta') \subset B(x_0,\delta)$, nous avons $\omega_f\big(B(x,\delta')\big) < \epsilon$, ce qui justifie que $\omega_f(x) < \epsilon$ et donc que x est également dans l'ensemble considéré. □

Théorème 12.51.
L'ensemble des points de discontinuité d'une limite simple de fonctions continues est de première catégorie.

Démonstration. Soit (f_n) une suite de fonctions convergent simplement vers f. Nous devons écrire l'ensemble des points de discontinuité de f comme une union dénombrable d'ensembles tels que sur tout intervalle I, aucun de ces ensembles n'est dense. Nous savons déjà par le lemme 12.50 que l'ensemble des points de discontinuité de f est donné par

$$D = \bigcup_{n=1}^{\infty} \{x \text{ tel que } \omega_f(x) \geqslant \frac{1}{n}\}. \tag{12.71}$$

Nous essayons donc de prouver que pour tout ϵ, l'ensemble

$$F = \{x \text{ tel que } \omega_f(x) \geqslant \epsilon\} \tag{12.72}$$

est nulle part dense. Soit

$$E_n = \bigcap_{i,j>n} \{x \text{ tel que } |f_i(x) - f_j(x)| < \epsilon\}. \tag{12.73}$$

Nous montrons que cet ensemble est fermée en étudiant le complémentaire. Soit $x \notin E_n$; alors il existe un couple (i,j) tel que

$$|f_i(x) - f_j(x)| > \epsilon. \tag{12.74}$$

Par continuité, cette inégalité reste valide dans un voisinage de x. Donc il existe un voisinage de x contenu dans $\complement E_n$ et E_n est donc fermé.

De plus nous avons $E_n \subset E_{n+1}$ et $\bigcup_n E_n = \mathbb{R}$. Ce dernier point est dû au fait que pour tout x, il existe N tel que $i,j > N$ implique $|f_i(x) - f_j(x)| \leqslant \epsilon$. Cela est l'expression du fait que la suite $\big(f_n(x)\big)_{n\in\mathbb{N}}$ est de Cauchy.

Soit I, un intervalle fermé de $\mathbb{R}$. Nous voulons trouver un intervalle $J \subset I$ sur lequel f est continue. Nous écrivons I sous la forme

$$I = \bigcup_{n=1}^{\infty} (E_n \cap I). \tag{12.75}$$

Tous les ensembles $J_n = E_n \cap I$ ne peuvent être nulle part dense en même temps (à cause du théorème de Baire 6.166). Il existe donc un n tel que J_n contienne un ouvert J. Le but est de montrer que f est continue sur J. Pour ce faire, nous n'allons pas simplement majorer $|f(x)-f(x_0)|$ par ϵ lorsque $|x - x_0|$ est petit. Ce que nous allons faire est majorer l'oscillation de f sur $B(x_0,\delta)$ lorsque δ est petit. Pour cela nous prenons x_0 et x dans J et nous écrivons

$$|f(x) - f(x_0)| \leqslant |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)|. \tag{12.76}$$

À ce niveau nous rappelons que n est fixé par le choix de J, dans lequel ϵ est déjà inclus. Nous choisissons évidemment $|x - x_0| \leqslant \delta$ de telle sorte que le second terme soit plus petit que ϵ en vertu de la continuité de f_n. Pour le premier terme, pour tout $i, j \geqslant n$ nous avons

$$|f_i(x) - f_j(x)| < \epsilon. \tag{12.77}$$

Si nous posons $j = n$ et $i \to \infty$, en tenant compte du fait que $f_i \to f$ simplement,

$$|f(x) - f_n(x)| \leqslant \epsilon. \tag{12.78}$$

Nous avons donc obtenu $|f(x) - f_n(x_0)| \leqslant 2\epsilon$. Cela signifie que dans un voisinage de rayon δ autour de x_0, les valeurs extrêmes prises par $f(x)$ sont $f_n(x_0) \pm 4\epsilon$. Nous avons donc prouvé que pour tout ϵ, il existe δ tel que

$$\omega_f\big([x_0 - \delta, x_0 + \delta]\big) \leqslant 4\epsilon. \tag{12.79}$$

De là nous concluons que

$$\lim_{\delta \to 0} \omega_f\big([x_0 - \delta, x_0 + \delta]\big) = 0, \tag{12.80}$$

ce qui signifie que f est continue en x_0. □

Exemple 12.52
Une fonction discontinue sur $\mathbb{Q}$ et continue ailleurs. La fonction

$$f(x) = \begin{cases} 0 & \text{si } x \notin \mathbb{Q} \\ \frac{1}{q} & \text{si } x = p/q \end{cases} \tag{12.81}$$

où par «$x = p/q$» nous entendons que p/q est la fraction irréductible.

Cette fonction est discontinue sur $\mathbb{Q}$ parce que si $q \in \mathbb{Q}$ alors $f(q) \neq 0$ alors que dans tous voisinage de q il existe un irrationnel sur qui la fonction vaudra zéro.

Montrons que f est continue sur les irrationnels. Si $x_0 \notin \mathbb{Q}$ alors $f(x_0) = 0$. Mais si on prend un voisinage suffisamment petit de x_0, nous pouvons nous arranger pour que tous les rationnels aient un dénominateur arbitrairement grand. En effet si nous nous fixons un premier rayon $r_0 > 0$ alors il existe un nombre fini de fractions de la forme 1, $\frac{k}{2}$, $\frac{k}{3}$,..., $\frac{k}{N}$ dans $B(x_0, r_0)$. Il suffit maintenant de choisir $0 < r \leqslant r_0$ tel que ces fractions soient toutes hors de $B(x_0, r)$. Dans cette boule nous avons $f < \frac{1}{N}$. Du coup f est continue en x_0. △

Définition 12.53 (Point périodique)**.**
Soit $f\colon I \to I$ une application d'un ensemble I dans lui-même. Si $x \in I$ vérifie $f^n(x) = x$ et $f^k(x) \neq x$ pour $k = 1, \dots, n-1$ alors on dit que x est un point n-périodique.

Lemme 12.54.
Soit I un segment[10] *de $\mathbb{R}$ et une fonction continue $f\colon I \to I$. Si K est un segment fermé avec $K \subset f(I)$ alors il existe un segment fermé $L \subset I$ tel que $K = f(L)$.*

Démonstration. Mentionnons immédiatement que f est continue sur I qui est compact[11]. Par conséquent tous les nombres dont nous allons parler sont finis parce que f est bornée par le théorème 6.126.

Soit $K = [\alpha, \beta]$. Si $\alpha = \beta$ alors le segment $L = \{a\}$ convient. Nous supposons donc que $\alpha \neq \beta$ et nous considérons $a, b \in I$ tels que $\alpha = f(a)$ et $\beta = f(b)$. Vu que $a \neq b$ nous supposons $a < b$ (le cas $a > b$ se traite de façon similaire).

Nous posons

$$A = \{x \in [a, b] \text{ tel que } f(x) = \alpha\}. \tag{12.82}$$

10. définition 6.118. Un segment est un intervalle fermé borné.
11. Par le lemme 6.121.

C'est un ensemble borné par a et b. De plus il est fermé ; ce dernier point n'est pas tout à fait évident parce que f n'est pas définit sur $\mathbb{R}$ mais sur I qui est fermé, le corollaire 12.18 n'est donc pas immédiatement utilisable. Prouvons donc que $Z = \{x \in \mathbb{R} \text{ tel que } f(x) = \alpha\}$ est fermé. Si x_0 est hors de Z alors soit x_0 est dans I soit il est hors de I. Dans ce second cas, le complémentaire de I étant ouvert, on a un voisinage de x_0 hors de I et par conséquent hors de Z. Si au contraire $x_0 \in I$ alors il y a (encore) deux cas : soit $x_0 \in \text{Int}(I)$ soit x_0 est sur le bord de I. Dans le premier cas, le théorème des valeurs intermédiaires [12] fonctionne. Pour le second cas, nous supposons $x_0 = \max(I)$ (le cas $x_0 = \min(I)$ est similaire). Le théorème des valeurs intermédiaires dit que sur $[x_0 - \epsilon, x_0]$, $f \neq \alpha$ et en même temps, sur $]x_0, x_0 + \epsilon]$, nous sommes en dehors du domaine. Au final $\{f(x) = \alpha\}$ est fermé et A est alors fermé en tant que intersection de deux fermés.

L'ensemble A étant non vide ($a \in A$), il possède donc un maximum que nous nommons u :

$$u = \max(A). \tag{12.83}$$

Nous posons aussi

$$B = \{x \in [u, b] \text{ tel que } f(x) = \beta\} \tag{12.84}$$

qui est encore fermé, borné et non vide. Nous pouvons donc définir

$$v = \min(B). \tag{12.85}$$

Nous prouvons maintenant que $f\big([u, v]\big) = [\alpha, \beta]$. D'abord $f\big([u, v]\big)$ est un intervalle compact [13] contenant $f(u) = \alpha$ et $f(v) = \beta$. Par conséquent $[\alpha, \beta] \subset f\big([u, v]\big)$. Pour l'inclusion inverse supposons $t \in [u, v]$ tel que $f(t) > \beta$. Vu que $f(a) = \alpha$ et $\alpha < \beta$ le théorème des valeurs intermédiaires il existe $t_0 \in [a, t]$ tel que $f(t_0) = \beta$. Cela donne $t_0 < v$ et donc contredit la minimalité de v dans B. Nous en déduisons que $f\big([u, v]\big)$ ne contient aucun élément plus grand que β. Même jeu pour montrer que ça ne contient aucun élément plus petit que α.

En définitive, le segment $L = [u, v]$ fonctionne. □

Lorsque $I_2 \subset f(I_1)$ nous notons $I_1 \to I_2$ ou, si une ambiguïté est à craindre, $I_1 \xrightarrow{f} I_2$. Cette flèche se lit «recouvre».

Lemme 12.55 ([117])**.**
Soient les segments $I_0, \dots, I_{n-1}$ tels que nous ayons le cycle

$$I_0 \to I_1 \to \dots \to I_{n-1} \to I_0. \tag{12.86}$$

Alors f^n admet un point fixe $x_0 \in I_0$ tel que $f^k(x_0) \in I_k$ pour tout $k = 0, \dots, n-1$.

Démonstration. Nous prouvons les cas $n = 1$ et $n = 2$ séparément.

- $\underline{n = 1}$ Nous avons $I_0 \to I_0$, c'est à dire que $I_0 \subset f(I_0)$. Si $I_0 = [a, b]$ alors nous posons $a = f(\alpha)$ et $b = f(\beta)$ pour certains $\alpha, \beta \in I_0$. Nous posons ensuite $g(x) = f(x) - x$.
  Dans un premier temps, $g(\alpha) = a - \alpha \leqslant 0$ parce que $a =\in (I_0)$ et $\alpha \in I_0$. Pour la même raison, $g(\beta) = b - \beta \geqslant 0$. Le théorème des valeurs intermédiaires donne alors $t_0 \in [\alpha, \beta] \subset I_0$ tel que $g(t_0) = 0$. Nous avons donc $f(t_0) = t_0$.
- $\underline{n = 2}$ Nous avons $I_0 \to I_1 \to I_0$. Vu que $I_1 \subset f(I_0)$, le lemme 12.54 donne un segment $J_1 \subset I_0$ tel que $f(J_1) = I_1$. Mézalors

  $$J_1 \subset I_0 \subset f(I_1) = f^2(J_1). \tag{12.87}$$

  Nous avons donc $J_1 \xrightarrow{f^2} J_1$ et par le cas $n = 1$ traité plus haut, la fonction f^2 a un point fixe x_0 dans J_1. De plus

  $$f(x_0) \in f(J_1) = I_1, \tag{12.88}$$

  le point x_0 est donc bien celui que nous cherchions.

12. Théorème 12.21.
13. Corollaire 12.22 et théorème 6.77.

Cas général. Nous avons

$$I_0 \to I_1 \to \ldots \to I_{n-1} \to I_0. \tag{12.89}$$

Vu que $I_1 \subset f(I_0)$, il existe $J_1 \subset I_0$ tel que $f(J_1) = I_1$. Mais

$$I_2 \subset f(I_1) = f^2(J_1), \tag{12.90}$$

donc il existe $J_2 \subset J_1$ tel que $I_2 = f^2(J_2)$. En procédant encore longtemps ainsi nous construisons les ensembles $J_1, \ldots, J_{n-1}$ tels que

$$J_{n-1} \subset J_{n-2} \subset \ldots \subset J_1 \subset J_0 \tag{12.91}$$

tels que $I_k = f^k(J_k)$ pour tout $k = 1, \ldots, n-1$. La dernière de ces inclusions est $I_{n-1} = f^{n-1}(J_{n-1})$, mais $I_{n-1} \to I_0$, c'est à dire que

$$I_0 \subset f(I_{n-1}) = f^n(J_{n-1}), \tag{12.92}$$

et il existe $J_n \subset J_{n-1}$ tel que $I_0 \subset f^n(J_n)$. Mais comme $J_n \subset J_0$ nous avons en particulier $J_n \subset f^n(J_n)$.
Cela donne un point fixe $x_0 \in J_n$ pour f^n. Par construction nous avons $J_n \subset J_{n-1} \subset \ldots \subset J_1 \subset J_0$ et donc $x_0 \in J_k$ pour tout k. En particulier

$$f^k(x_0) \in f^k(J_k) = I_k \tag{12.93}$$

pour tout k.

□

Théorème 12.56 (Théorème de Sarkowski[117])**.**
Soit I, un segment de $\mathbb{R}$ et une application continue $f : I \to I$. Si f admet un point 3-périodique, alors f admet des points n-périodiques pour tout $n \geqslant 1$.

Démonstration. Soit $a \in I$ un point 3-périodique pour f et notons $b = f(a)$, $c = f(b)$. Les points b et c sont également des points 3-périodiques. Quitte à renommer, nous pouvons supposer que a est le plus petit des trois. Il reste deux possibilités : $a < b < c$ et $a < c < b$. Nous traitons d'abord le premier cas.

Supposons $a < b < c$. Nous posons $I_0 = [a, b]$ et $I_1 = [b, c]$. Nous avons immédiatement $I_1 \subset f(I_0)$ et comme $f(b) = c$ et $f(c) = a$, $f(I_1)$ recouvre $[a, c]$ et donc recouvre en même temps I_1 et I_2. Nous avons donc $I_0 \to I_1$, $I_1 \to I_0$ et $I_1 \to I_1$.

Un point 1-périodique Nous avons $I_1 \to I_1$ qui prouve que f a un point fixe dans I_1. Voila un point 1-périodique.

Un point 2-périodique Nous avons $I_0 \to I_1 \to I_0$. Par conséquent, le lemme 12.55 dit que f^2 a un point fixe $x_0 \in I_0$ tel que $f(x_0) \in I_1$. Montrons que $f(x_0) \neq f(x_0)$. Pour avoir $x_0 = f(x_0)$, il faudrait $x_0 \in I_0 \cap I_1 = \{b\}$. Mais b est un point 3-périodique, donc ne vérifiant certainement pas $f^2(b) = b$. Nous en déduisons que $f(x_0) \neq x_0$ et donc que x_0 est 2-périodique.

Un point 3-périodique On en a par hypothèse.

Un point n-périodique pour $n \geqslant 4$ Nous avons le cyle

$$I_0 \to \underbrace{I_1 \to I_1 \to \ldots \to I_1}_{\text{n-1}\,fois} \to I_0. \tag{12.94}$$

Le lemme donne alors un point fixe $x \in I_0$ pour f^n tel que $f^k(x) \in I_1$ pour $k = 1, \ldots, n-1$. Est-ce possible que $x = b$? Non parce que $f^2(b) = a \in I_0$ alors que $f^2(x) \in I_1$. Mais $I_0 \cap I_1 = \{b\}$.
Par conséquent la relation $f^k(x) \in I_1$ exclu d'avoir $f^k(x) = x$, et le point x est bien n-périodique.

Passons au cas $a < c < b$. Alors nous posons $I_0 = [a, c]$ et $I_1 = [c, b]$. Encore une fois $f(I_0)$ contient a et b, donc $I_0 \to I_0$ et $I_0 \to I_1$. Mais en même temps $f(I_1)$ contient a et c, donc $I_1 \to I_0$.

Nous pouvons donc refaire comme dans le premier cas, en inversant les rôles de I_0 et I_1. En particulier nous pouvons considérer le cycle

$$I_1 \to I_0 \to I_0 \to \ldots \to I_0 \to I_1. \tag{12.95}$$

□

12.5 Uniforme continuité

Définition 12.57.
Une partie $A \subset \mathbb{R}^m$ est dite ***bornée*** *s'il existe un $M > 0$ tel que $A \subset B(0, M)$. Le* ***diamètre*** *de la partie A est le nombre*

$$\text{Diam}(A) = \sup_{x,y \in A} \|x - y\| \in [0, \infty]. \tag{12.96}$$

Lorsque A est borné, il existe un M tel que $\|x\| \leqslant M$ pour tout $x \in A$.

Lemme 12.58.
Si A est une partie non vide de $\mathbb{R}^m$, alors $\text{Diam}(A) = \text{Diam}(\bar{A})$.

Nous n'allons pas donner de démonstrations de ce lemme.

Si (x_n) est une suite et I est un sous-ensemble infini de $\mathbb{N}$, nous désignons par x_I la suite des éléments x_n tels que $n \in I$. Par exemple la suite $x_{\mathbb{N}}$ est la suite elle-même, la suite $x_{2\mathbb{N}}$ est la suite obtenue en ne prenant que les éléments d'indice pair.

Les suites x_I ainsi construites sont dites des **sous-suites** de la suite (x_n).

Pour une fonction $f\colon D \subset \mathbb{R}^m \to \mathbb{R}$, la continuité au point a signifie que pour tout $\varepsilon > 0$,

$$\exists \delta > 0 \text{ tel que } 0 < \|x - a\| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon. \tag{12.97}$$

Le δ qu'il faut choisir dépend évidement de ε, mais il dépend en général aussi du point a où l'on veut tester la continuité. C'est à dire que, étant donné un $\varepsilon > 0$, nous pouvons trouver un δ qui fonctionne pour certains points, mais qui ne fonctionne pas pour d'autres points.

Il peut cependant également arriver qu'un même δ fonctionne pour tous les points du domaine. Dans ce cas, nous disons que la fonction est uniformément continue sur le domaine.

Définition 12.59.
Une fonction $f\colon D \subset \mathbb{R}^m \to \mathbb{R}$ est dite ***uniformément continue*** *sur D si*

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ tel que } \forall x, y \in D, \|x - y\| \leqslant \delta \Rightarrow |f(x) - f(a)| < \varepsilon. \tag{12.98}$$

Il est intéressant de voir ce que signifie le fait de *ne pas* être uniformément continue sur un domaine D. Il s'agit essentiellement de retourner tous les quantificateurs de la condition (12.98) :

$$\exists \varepsilon > 0 \text{ tel que } \forall \delta > 0, \exists x, y \in D \text{ tel que } \|x - y\| < \delta \text{ et } \big|f(x) - f(y)\big| > \varepsilon. \tag{12.99}$$

Dans cette condition, les points x et y peuvent être fonction du δ. L'important est que pour tout δ, on puisse trouver deux points δ-proches dont les images par f ne soient pas ε-proches.

Exemple 12.60
Prenons la fonction $f(x) = \frac{1}{x}$, et demandons nous pour quel δ nous sommes sûr d'avoir

$$|f(a + \delta) - f(a)| = \left| \frac{1}{a + \delta} - \frac{1}{a} \right| < \varepsilon. \tag{12.100}$$

Pour simplifier, nous supposons que $a > 0$. Nous calculons

$$\begin{aligned} \frac{1}{a} - \frac{1}{a+\delta} &< \varepsilon \\ \frac{\delta}{a(a+\delta)} &< \varepsilon \\ \delta &< \varepsilon a^2 + \varepsilon a\delta \\ \delta(1-\varepsilon a) &< \varepsilon a^2 \\ \delta &< \frac{\varepsilon a^2}{1-\varepsilon a}. \end{aligned} \tag{12.101}$$

Notons que, à ε fixé, plus a est petit, plus il faut choisir δ petit. La fonction $x \mapsto \frac{1}{x}$ n'est donc pas uniformément continue. Cela correspond au fait que, proche de zéro, la fonction monte très vite. Une fonction uniformément continue sera une fonction qui ne montera jamais très vite. △

Proposition 12.61.
Quelques propriétés des fonctions uniformément continues.

(1) Toute application uniformément continue est continue ;

(2) la composée de deux fonctions uniformément continues est uniformément continue ;

(3) tout application lipschitzienne est uniformément continues.

Une fonction peut être uniformément continue sur un domaine et pas sur un autre. Le théorème suivant donne une importante indication à ce sujet.

Théorème 12.62 (Heine)**.**
Une fonction continue sur un compact (fermé et borné) est uniformément continue.

La démonstration qui suit est valable pour une fonction $f\colon \mathbb{R}^n \to \mathbb{R}^m$ et utilise le fait que le produit cartésien de compacts est compact. Dans le cas de fonctions sur $\mathbb{R}$, nous pouvons modifier la démonstration pour ne pas utiliser ce résultat ; voir plus bas.

Démonstration. Nous allons prouver ce théorème par l'absurde. Nous commençons par écrire la condition (12.99) qui exprime que f n'est pas uniformément continue sur le compact K :

$$\exists \varepsilon > 0 \text{ tel que } \forall \delta > 0,\ \exists x, y \in K \text{ tels que } \|x-y\| < \delta \text{ et } \big|f(x) - f(y)\big| > \varepsilon. \tag{12.102}$$

En particulier (en prenant $\delta = \frac{1}{n}$ pour tout n), pour chaque n nous pouvons trouver x_n et y_n dans K qui vérifient simultanément les deux conditions suivantes :

$$\begin{cases} \|x_n - y_n\| < \dfrac{1}{n} & \text{(12.103a)} \\ \big|f(x_n) - f(y_n)\big| > \varepsilon. & \text{(12.103b)} \end{cases}$$

Nous insistons que c'est le même ε pour chaque n. L'ensemble K étant compact, l'ensemble $K \times K$ est compact (théorème 6.162) et nous pouvons trouver une sous-suite convergente *du couple* (x_n, y_n) dans $K \times K$. Quitte à passer à ces sous-suites, nous nous supposons que (x_n, y_n) converge dans $K \times K$ et en particulier que les suites (x_n) et (y_n) sont convergentes. Étant donné que pour chaque n elles vérifient $\|x_n - y_n\| < \frac{1}{n}$, les limites sont égales :

$$\lim x_n = \lim y_n = x. \tag{12.104}$$

L'ensemble K étant fermé, la limite x est dans K. Par continuité de f, nous avons finalement

$$\lim f(x_n) = \lim f(y_n) = f(x), \tag{12.105}$$

mais alors

$$\lim_{n\to\infty} \big|f(x_n) - f(y_n)\big| = 0, \tag{12.106}$$

ce qui est en contradiction avec le choix (12.103b).

Tout ceci prouve que $f(K)$ est bornée supérieurement et que f atteint son supremum (qui est donc un maximum). Le fait que $f(K)$ soit borné inférieurement se prouve en considérant la fonction $-f$ au lieu de f.

□

Remarque 12.63.
Nous pouvons ne pas utiliser le fait que le produit de compacts est compact. Cela est particulièrement commode lorsqu'on considère des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ parce que dans ce cadre nous ne pouvons pas supposer connue la notion de produit d'espace topologiques.

Pour choisir les sous-suites (x_n) et (y_n), il suffit de prendre une sous-suite convergente de (x_n) et d'invoquer le fait que $\|x_n - y_n\| \leqslant \frac{1}{n}$. Les suites (x_n) et (y_n) étant adjacentes, la convergence de (x_n) implique la convergence de (y_n) vers la même limite.

Il est donc un peu superflus de parler de la convergence du couple (x_n, y_n).

12.6 Compacité

Soit E, un sous ensemble de $\mathbb{R}$. Nous pouvons considérer les ouverts suivants :

$$\mathcal{O}_x = B(x, 1) \tag{12.107}$$

pour chaque $x \in E$. Évidement,

$$E \subseteq \bigcup_{x \in E} \mathcal{O}_x. \tag{12.108}$$

Cette union est très souvent énorme, et même infinie. Elle contient de nombreuses redondances. Si par exemple $E = [-10, 10]$, l'élément $3 \in E$ est contenu dans $\mathcal{O}_{3.5}$, $\mathcal{O}_{2.7}$ et bien d'autres. Pire : même si on enlève par exemple $\mathcal{O}_2$ de la liste des ouverts, l'union de ce qui reste continue à être tout E. La question est : *est-ce qu'on peut en enlever suffisamment pour qu'il n'en reste qu'un nombre fini ?*

Définition 12.64.
Soit E, un sous ensemble de $\mathbb{R}$. Une collection d'ouverts $\mathcal{O}_i$ est un ***recouvrement*** *de E si $E \subseteq \bigcup_i \mathcal{O}_i$. Un sous ensemble E de $\mathbb{R}$ tel que de tout recouvrement par des ouverts, on peut extraire un sous-recouvrement fini est dit* ***compact****.*

Proposition 12.65.
Les ensembles compacts sont fermés et bornés.

Démonstration. Prouvons d'abord qu'un ensemble compact est borné. Pour cela, supposons que K est un compact non borné vers le haut [14]. Donc il existe une suite infinie de nombres strictement croissante $x_1 < x_2 < \dots$ tels que $x_i \in K$. Prenons n'importe quel recouvrement ouvert de la partie de K plus petite ou égale à x_1, et complétons ce recouvrement par les ouverts $\mathcal{O}_i =]x_{i-1}, x_i[$. Le tout forme bien un recouvrement de K par des ouverts.

Il n'y a cependant pas moyen d'en tirer un sous recouvrement fini parce que si on ne prends qu'un nombre fini parmi les $\mathcal{O}_i$, on en aura fatalement un maximum, disons $\mathcal{O}_k$. Dans ce cas, les points x_{k+1}, x_{k+1},... ne seront pas dans le choix fini d'ouverts.

Cela prouve que K doit être borné.

Pour prouver que K est fermé, nous allons prouver que le complémentaire est ouvert. Et pour cela, nous allons prouver que si le complémentaire n'est pas ouvert, alors nous pouvons construire un recouvrement de K dont on ne peut pas extraire de sous recouvrement fini.

Si $\mathbb{R}\backslash K$ n'est pas ouvert, il possède un point, disons x, tel que tout voisinage de x intersecte K. Soit $B(x, \epsilon_1)$, un de ces voisinages, et prenons $k_1 \in K \cap B(x, \epsilon_1)$. Ensuite, nous prenons ϵ_2 tel que k_1 n'est pas dans $B(x, \epsilon_1)$, et nous choisissons $k_2 \in K \cap B(x, \epsilon_2)$. De cette manière, nous construisons une suite de $k_i \in K$ tous différents et de plus en plus proches de x. Prenons un recouvrement

14. Nous laissons à titre d'exercice le cas où K est borné par le haut et pas par le bas.

quelconque par des ouverts de la partie de K qui n'est pas dans $B(x, \epsilon_1)$. Les nombres k_i ne sont pas dans ce recouvrement.

Nous ajoutons à ce recouvrement les ensembles $\mathcal{O} =]k_i, k_{i+1}[$. Le tout forme un recouvrement (infini) par des ouverts dont il n'y a pas moyen de tirer un sous recouvrement fini, pour exactement la même raison que la première fois. □

Le résultat suivant le théorème de Borel-Lebesgue, et la démonstration vient de wikipédia.

Théorème 12.66 (borel-Lebesgue)**.**
Les intervalles de la forme $[a, b]$ sont compacts.

Démonstration. Soit Ω, un recouvrement du segment $[a, b]$ par des ouverts, c'est à dire que

$$[a, b] \subseteq \bigcup_{\mathcal{O} \in \Omega} \mathcal{O}. \tag{12.109}$$

Nous notons par M le sous-ensemble de $[a, b]$ des points m tels que l'intervalle $[a, m]$ peut être recouvert par un sous-ensemble fini de Ω. C'est à dire que M est le sous ensemble de $[a, b]$ sur lequel le théorème est vrai. Le but est maintenant de prouver que $M = [a, b]$.

- M **est non vide** En effet, $a \in M$ parce que il existe un ouvert $\mathcal{O} \in \Omega$ tel que $a \in \mathcal{O}$. Donc $\mathcal{O}$ tout seul recouvre l'intervalle $[a, a]$.
- M **est un intervalle** Soient m_1, $m_2 \in M$. Le but est de montrer que si $m' \in [m_1, m_2]$, alors $m' \in M$. Il y a un sous recouvrement fini de l'intervalle $[a, m_2]$ (par définition de $m_2 \in M$). Ce sous recouvrement fini recouvre évidement aussi $[a, m']$ parce que $[a, m'] \subseteq [a, m_2]$, donc $m' \in M$.
- M **est une ensemble ouvert** Soit $m \in M$. Le but est de prouver qu'il y a un ouvert autour de m qui est contenu dans M. Mettons que Ω' soit un sous recouvrement fini qui contienne l'intervalle $[a, m]$. Dans ce cas, on a un ouvert $\mathcal{O} \in \Omega'$ tel que $m \in \mathcal{O}$. Tous les points de $\mathcal{O}$ sont dans M, vu qu'ils sont tous recouverts par Ω'. Donc $\mathcal{O}$ est un voisinage de m contenu dans M.
- M **est un ensemble fermé** M est un intervalle qui commence en a, en contenant a, et qui finit on ne sait pas encore où. Il est donc soit de la forme $[a, m]$, soit de la forme $[a, m[$. Nous allons montrer que M est de la première forme en démontrant que M contient son supremum s. Ce supremum est un élément de $[a, b]$, et donc il est contenu dans un des ouverts de Ω. Disons $s \in \mathcal{O}_s$. Soit c, un élément de $\mathcal{O}_s$ strictement plus petit que c ; étant donné que s est supremum de M, cet élément c est dans M, et donc on a un sous recouvrement fini Ω' qui recouvre $[a, c]$. Maintenant, le sous recouvrement constitué de Ω' et de $\mathcal{O}_s$ est fini et recouvre $[a, s]$.

Nous pouvons maintenant conclure : le seul intervalle non vide de $[a, b]$ qui soit à la fois ouvert et fermé est $[a, b]$ lui-même, ce qui prouve que $M = [a, b]$, et donc que $[a, b]$ est compact. □

Par le théorème des valeurs intermédiaires, l'image d'un intervalle par une fonction continue est un intervalle, et nous avons l'importante propriété suivante des fonctions continues sur un compact.

Le théorème suivant est un cas particulier du théorème 6.126.

Théorème 12.67.
Si f est une fonction continue sur l'intervalle compact $[a, b]$. Alors f est bornée sur $[a, b]$ et elle atteint ses bornes.

Démonstration. Étant donné que $[a, b]$ est un intervalle compact, son image est également un intervalle compact, et donc est de la forme $[m, M]$. Ceci découle du théorème 6.77 et le corollaire 12.22. Le maximum de f sur $[a, b]$ est la borne M qui est bien dans l'image (parce que $[m, M]$ est fermé). Idem pour le minimum m. □

12.7 Limite et continuité

12.7.1 Définition

Proposition 12.68 (Caractérisation de la limite)**.**
Soit une fonction $f\colon D \subset \mathbb{R} \to \mathbb{R}$ et a un point d'accumulation de D. On dit que f admet une ***limite*** *en a s'il existe un réel ℓ tel que*

$$\forall \varepsilon > 0,\ \exists \delta > 0 \text{ tel que } \forall x \in D,\ 0 < |x-a| < \delta \Rightarrow |f(x) - \ell| < \varepsilon. \tag{12.110}$$

Démonstration. Il s'agit seulement de recopier la définition 6.69. □

Si aucun nombre ℓ ne vérifie la condition de la définition, alors on dit que la fonction n'admet pas de limite en a. Lorsque f possède la limite ℓ en a, nous notons

$$\lim_{x \to a} f(x) = \ell. \tag{12.111}$$

Proposition 12.69.
Soit une fonction $f\colon D \to \mathbb{R}$. Si a est un point d'accumulation de D et s'il existe une limite de f en a, alors il en existe une seule.

De façon équivalente, il ne peut pas exister deux nombres $\ell \neq \ell'$ vérifiant tout les deux la condition (12.110).

Démonstration. Soient ℓ et ℓ' deux limites de f au point a. Par définition, pour tout ε nous avons des nombres δ et δ' tels que

$$\begin{aligned} |x-a| < \delta &\Rightarrow \big|f(x) - \ell\big| < \varepsilon \\ |x-a| < \delta' &\Rightarrow \big|f(x) - \ell'\big| < \varepsilon \end{aligned} \tag{12.112}$$

Pour fixer les idées, supposons que $\delta < \delta'$ (le cas $\delta \geqslant \delta'$ se traite de la même manière).

Étant donné que a est un point d'accumulation du domaine D de f, il existe un $x \in D$ tel que $|x-a| < \delta$. Évidemment, nous avons aussi $|x-a| < \delta'$. Les conditions (12.112) signifient alors que ce x vérifie en même temps

$$|f(x) - \ell| < \varepsilon, \tag{12.113}$$

et

$$|f(x) - \ell'| < \varepsilon. \tag{12.114}$$

Afin de prouver que $\ell = \ell'$, nous allons maintenant calculer $|\ell - \ell'|$ et montrer que cette distance est plus petite que tout nombre. Nous avons (voir remarque 12.70)

$$|\ell - \ell'| = |\ell - f(x) + f(x) - \ell'| \leqslant |\ell - f(x)| + |f(x) - \ell'| < \varepsilon + \varepsilon. \tag{12.115}$$

En résumé, pour tout $\varepsilon > 0$ nous avons

$$|\ell - \ell'| < 2\varepsilon, \tag{12.116}$$

et donc $|\ell - \ell'| = 0$, ce qui signifie que $\ell = \ell'$. □

Remarque 12.70.
Les inégalités (12.115) utilisent deux techniques très classiques en analyse qu'il convient d'avoir bien compris. La première est de faire

$$|A - B| = |A - C + C - B|. \tag{12.117}$$

Il s'agit d'ajouter $-C + C$ dans la norme. Évidemment, cela ne change rien.

La seconde technique est l'inégalité

$$|A + B| \leqslant |A| + |B|. \tag{12.118}$$

Exemple 12.71
Considérons la fonction $f(x) = 2x$, et calculons la limite $\lim_{x\to 3} f(x)$. Vu que $f(3) = 6$, nous nous attendons à avoir $\ell = 6$. C'est ce que nous allons prouver maintenant. Pour chaque $\varepsilon > 0$ nous devons trouver un $\delta > 0$ tel que $|x - 3| < \delta$ implique $|f(x) - 6| < \varepsilon$. En remplaçant $f(x)$ par sa valeur en fonction de x et avec quelques manipulations nous trouvons :

$$\begin{aligned} |f(x) - 6| &< \varepsilon \\ |2x - 6| &< \varepsilon \\ 2|x - 3| &< \varepsilon \\ |x - 3| &< \frac{\varepsilon}{2} \end{aligned} \tag{12.119}$$

Donc dès que $|x - 3| < \frac{\varepsilon}{2}$, nous avons $|f(x) - 6| < \varepsilon$. Nous posons donc $\delta = \frac{\varepsilon}{2}$.

Plus généralement, nous avons $\lim_{x\to a} f(x) = 2a$, et cela se prouve en étudiant $|f(x) - 2a|$ exactement de la même manière. △

12.7.2 Propriétés de base

Proposition 12.72.
La limite est une opération linéaire, c'est à dire que si f et g sont des fonctions qui admettent des limites en a et si λ est un nombre réel,

(1) $\lim_{x\to a}(\lambda f)(x) = \lambda \lim_{x\to a} f(x)$,

(2) $\lim_{x\to a}(f + g)(x) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x)$.

En combinant les deux propriétés de la proposition 12.72, nous pouvons écrire

$$\lim_{x\to a}(\lambda f + \mu g)(x) = \lambda \lim_{x\to a} f(x) + \mu \lim_{x\to a} g(x). \tag{12.120}$$

pour toutes fonctions f et g admettant une limite en a et pour tout réels λ et μ.

En plus d'être linéaire, la limite possède les deux propriétés suivantes.

Proposition 12.73.
Si f et g sont deux fonctions qui admettent une limite en a, alors

$$\lim_{x\to a}(fg)(x) = \lim_{x\to a} f(x) \cdot \lim_{x\to a} g(x). \tag{12.121}$$

Si de plus $\lim_{x\to a} g(x) \neq 0$, *alors*

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)}. \tag{12.122}$$

12.7.3 Limites de fonctions

Définition 12.74.
Soit $f\colon D \subset \mathbb{R}^m \to \mathbb{R}$ une fonction et a un point d'accumulation de D. On dit que f possède une ***limite*** *s'il existe un élément $\ell \in \mathbb{R}$ tel que*

$$\forall \varepsilon > 0,\ \exists \delta > 0 \text{ tel que } 0 < \|x - a\| < \delta \Rightarrow |f(x) - \ell| < \varepsilon. \tag{12.123}$$

Pour une fonction $f\colon D \subset \mathbb{R}^m \to \mathbb{R}^n$, la définition est la même, sauf que nous remplaçons la valeur absolue par la norme dans $\mathbb{R}^n$. Nous disons donc que ℓ est la limite de f lorsque x tend vers a, et nous notons $\lim_{x\to a} f(x) = \ell$ lorsque pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que

$$0 < \|x - a\|_{\mathbb{R}^m} < \delta \Rightarrow \|f(x) - \ell\|_{\mathbb{R}^n} < \varepsilon. \tag{12.124}$$

Remarque 12.75.
Dans l'équation (12.124), nous avons explicitement écrit les normes $\|.\|_{\mathbb{R}^m}$ et $\|.\|_{\mathbb{R}^n}$. Dans la suite nous allons le plus souvent noter $\|.\|$ sans plus de précision. Il est important de faire l'exercice de bien comprendre à chaque fois de quelle norme nous parlons.

Remarque 12.76.
Il est important de remarquer à quel point les définitions 12.74, et les caractérisatons 9.188, 12.68 sont analogues. En réalité, la définition fondamentale est la définition de la limite dans les espaces vectoriels normés ; les deux autres sont des cas particuliers, adaptés à $\mathbb{R}$ et $\mathbb{R}^m$. Il en sera de même pour les définitions de fonctions continues : il y aura une définition pour la continuité de fonctions entre espaces vectoriels normés, et ensuite une définition pour les fonctions de $\mathbb{R}^m$ dans $\mathbb{R}^n$ qui en sera un cas particulier.

Tentons de comprendre ce que signifie qu'un nombre ℓ *ne soit pas* la limite de f lorsque $x \to a$. Il s'agit d'inverser la condition (12.123). Le nombre ℓ n'est pas une limite de f pour $x \to a$ lorsque

$$\exists \varepsilon > 0 \text{ tel que } \forall \delta > 0, \exists x \text{ tel que } 0 < \|x - a\| < \delta \text{ et } \|f(x) - \ell\| > \varepsilon, \tag{12.125}$$

c'est à dire qu'il existe un certain seuil ε tel qu'on a beau s'approcher aussi proche qu'on veut de a (distance δ), on trouvera toujours un x tel que $f(x)$ n'est pas ε-proche de ℓ.

Lemme 12.77 (Unicité de la limite)**.**
Si ℓ et ℓ' sont deux limites de $f(x)$ lorsque x tend vers a, alors $\ell = \ell'$.

Démonstration. Soit $\varepsilon > 0$. Nous considérons δ tel que $\|f(x)-\ell\| < \varepsilon$ pour tout x tel que $\|x-a\| < \delta$. De la même manière, nous prenons δ' tel que $\|x - a\| < \delta'$ implique $\|f(x) - \ell'\| < \varepsilon$. Pour les x tels que $\|x - a\|$ est plus petit que δ et δ' en même temps, nous avons

$$\|\ell - \ell'\| = \|\ell - f(x) + f(x) - \ell'\| \leqslant \|\ell - f(x)\| + \|f(x) - \ell'\| < 2\varepsilon, \tag{12.126}$$

et donc $\|\ell - \ell'\| = 0$ parce que c'est plus petit que 2ε pour tout ε. □

Le concept de limite appelle immédiatement celui de continuité.

Définition 12.78.
*Soit $f\colon D \subset \mathbb{R}^m \to \mathbb{R}^n$ et $a \in D$. On dit que f est **continue** en a lorsque la limite $\lim_{x\to a} f(x)$ existe et est égale à $f(a)$.*

On dit que f est continue sur une partie $A \subset D$ si elle est continue en tous les points de a.

La continuité peut évidement être récrite avec une formule du même type que celle de la limite.

Proposition 12.79.
La fonction $f\colon D \subset \mathbb{R}^m \to \mathbb{R}^n$ est continue en $a \in D$ si et seulement si

$$\forall \varepsilon, \exists \delta > 0 \text{ tel que } x \in D \cap B(a, \delta) \Rightarrow \|f(x) - f(a)\| < \varepsilon. \tag{12.127}$$

Quasiment toutes les propriétés des limites ont un équivalent concernant la continuité.

Proposition 12.80.
Soit $f\colon D \subset \mathbb{R}^m \to \mathbb{R}^n$. Nous avons

$$\lim_{x\to a} f(x) = \ell \tag{12.128}$$

si et seulement si

$$\lim_{x\to a} f_i(x) = \ell_i \tag{12.129}$$

pour tout $i \in \{1, \dots, n\}$ où $f_i(x)$ dénote la i-ème composante de $f(x)$ et ℓ_i la i-ème composante de $\ell \in \mathbb{R}^n$.

Cette proposition revient à dire que la convergence d'une fonction est équivalente à la convergence de chacune de ses composantes.

Démonstration. L'élément clef de la preuve est le fait que pour tout vecteur $u \in \mathbb{R}^p$, nous ayons l'inégalité

$$|u_i| \leqslant \sqrt{\sum_{k=1}^{p} |u_k|^2} = \|u\|. \tag{12.130}$$

La norme (dans $\mathbb{R}^p$) d'un vecteur est plus grande ou égale à la valeur absolue de chacune de ses composantes.

Supposons que nous ayons une fonction dont chacune des composantes a une limite en a : $\lim_{x\to a} f_i(x) = \ell_i$. Montrons que dans ce cas la fonction f tend vers ℓ. Si nous considérons $\varepsilon > 0$, par définition de la limite de chacune des fonctions f_i, il existent des δ_i tels que

$$\|x - a\|_{\mathbb{R}^m} < \delta_i \Rightarrow |f_i(x) - \ell_i| < \varepsilon. \tag{12.131}$$

Notez que la norme à gauche est une norme dans $\mathbb{R}^m$ et que celle à droite est une simple valeur absolue dans $\mathbb{R}$. Considérons $\delta = \min\{\delta_i\}_{i=1,\dots n}$. Si $\|x - a\| < \delta$, alors

$$\|f(x) - \ell\| = \sqrt{\sum_{i=1}^{n} |f_i(x) - \ell_i|^2} < \sqrt{\sum_{i=1}^{n} \varepsilon^2} = \sqrt{n\varepsilon^2} = \sqrt{n}\varepsilon. \tag{12.132}$$

Nous voyons qu'en choisissant les δ_i tels que $|f_i(x) - \ell_i| < \varepsilon$, nous trouvons $\|f(x) - \ell\| < \sqrt{n}\varepsilon$. Afin d'obtenir $\|f(x) - \ell\| < \varepsilon$, nous choisissons donc les δ_i de telle manière a avoir $|f_i(x) - \ell_i| < \varepsilon/\sqrt{n}$.

Nous avons donc prouvé que la limite composante par composante impliquait la limite de la fonction. Nous devons encore prouver le sens inverse.

Supposons donc que $\lim_{x\to a} f(x) = \ell$, et prouvons que nous ayons $\lim_{x\to a} f_i(x) = \ell_i$ pour chaque i. Soit $\varepsilon > 0$ et $\delta > 0$ tel que $\|x - a\| < \delta$ implique $\|f(x) - \ell\| < \varepsilon$. Avec ces choix, nous avons

$$|f_i(x) - \ell_i| \leqslant \|f(x) - \ell\| < \varepsilon \tag{12.133}$$

où nous avons utilisé la majoration (12.130) avec $f(x) - \ell$ en guise de u. □

De même, pour la continuité nous avons la proposition suivante :

Proposition 12.81.
Soit une fonction $f\colon D \subset \mathbb{R}^m \to \mathbb{R}^n$ et $a \in D$. La fonction f est continue en a si et seulement si chacune de ses composantes l'est, c'est à dire si et seulement si chacune des fonctions $f_i\colon D \to \mathbb{R}$ est continue en a.

Essayez de prouver cette proposition directement par la définition de la continuité, en suivant pas à pas la démonstration de la proposition 12.80.

Proposition 12.82.
Soit $f\colon \mathbb{R}^m \to \mathbb{R}$ et a, un point du domaine de f telle que $f(a) > 0$. Alors il existe un rayon r tel que $f(x) > 0$ pour tout x dans $B(a, r)$.

Cette proposition signifie que si la fonction est strictement positive en un point, alors elle restera strictement positive en tous les points «pas trop loin».

Démonstration. Prenons $\varepsilon = f(a)/2$ dans la définition de la continuité. Il existe donc un rayon δ tel que pour tout x dans $B(a, \delta)$,

$$|f(x) - f(a)| \leqslant \frac{f(a)}{2}, \tag{12.134}$$

en d'autres termes, $f(x) \in B\big(f(a), \frac{f(a)}{2}\big)$. Évidement aucun nombre négatif ne fait partie de cette dernière boule lorsque $f(a)$ est strictement positif. □

Corollaire 12.83.
Si $f\colon \mathbb{R}^m \to \mathbb{R}$ est une fonction continue, alors l'ensemble

$$A = \{x \in \mathbb{R}^m \text{ tels que } f(x) \neq 0\} \tag{12.135}$$

est ouvert.

Démonstration. Soit $x \in A$. Si $x > 0$ (le cas $x < 0$ est laissé en exercice), alors il existe une boule autour de x sur laquelle f reste strictement positive (proposition 12.82). Cette boule est donc contenue dans A. Étant donné qu'autour de chaque point de A nous pouvons trouver une boule contenue dans A, ce dernier est ouvert. □

Exemple 12.84
Soit $GL_n(\mathbb{R})$ l'ensemble des matrices $n \times n$ inversibles. Nous allons montrer que $GL_n(\mathbb{R})$ est un ouvert de $\mathbb{R}^{n^2}$. L'identification entre les vecteurs et les matrices consiste simplement à «déplier» la matrice pour en faire un vecteur. Par exemple, en dimension deux,

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \in \mathbb{R}^4. \tag{12.136}$$

En dimension 3,

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{pmatrix} \in \mathbb{R}^9. \tag{12.137}$$

Une matrice est inversible si et seulement si son déterminant est non nul. Or le déterminant est un polynôme en les composantes de la matrice. En dimension deux, nous avons

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc, \tag{12.138}$$

mais en écriture «dépliée», nous pouvons aussi bien écrire

$$\det \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = ad - bc. \tag{12.139}$$

En dimension 3, le déterminant est donc un polynôme des 9 variables qui apparaissent dans le vecteur «déplié». En général, dans $\mathbb{R}^{n^2}$, nous considérons donc le polynôme $\det\colon \mathbb{R}^{n^2} \to \mathbb{R}$ qui à un vecteur $X \in \mathbb{R}^{n^2}$ fait correspondre le déterminant de la matrice obtenue en «repliant» le vecteur X.

Donc dans $\mathbb{R}^{n^2}$, l'ensemble des matrices inversibles est donné par l'ensemble des vecteurs sur lesquels le polynôme det ne s'annule pas, c'est à dire

$$\{X \in \mathbb{R}^{n^2} \text{ tels que } \det(X) \neq 0\}. \tag{12.140}$$

Mais le déterminant est un polynôme, et donc une fonction continue. Cet ensemble est par conséquence ouvert par le corollaire 12.83. △

La proposition suivante montre que la limite peut «passer à travers» les fonctions continues.

Proposition 12.85 (limite de fonction composée)**.**
Soit $f\colon \mathbb{R}^n \to \mathbb{R}^q$ et $g\colon \mathbb{R}^m \to \mathbb{R}^n$ telles que

$$\lim_{x\to a} g(x) = p \tag{12.141a}$$
$$\lim_{y\to p} f(y) = q \tag{12.141b}$$

Alors nous avons $\lim_{x\to a}(f\circ g)(x) = q$.

Démonstration. Comme presque toute preuve à propos de limite ou de continuité, nous commençons par choisir $\varepsilon > 0$. Nous devons montrer qu'il existe un δ tel que $\|x-a\| \leqslant \delta$ implique $\|f\big(g(x)\big) - q\| \leqslant \varepsilon$.

La limite (12.141b) impose l'existence d'un $\tilde{\delta}$ tel que $\|y-p\| \leqslant \tilde{\delta}$ implique $\|f(y)-q\| \leqslant \varepsilon$, tandis que la limite (12.141a) donne un δ tel que $\|x-a\| \leqslant \delta$ implique $\|g(x)-p\| \leqslant \tilde{\delta}$ (nous avons pris $\tilde{\delta}$ en guise de ε dans la définition de la limite pour g).

Avec ces choix, si $\|x-a\| \leqslant \delta$, alors $\|g(x)-p\| \leqslant \tilde{\delta}$, et par conséquent,

$$\|f\big(g(x)\big) - q\| \leqslant \varepsilon, \tag{12.142}$$

ce que nous voulions. □

De façon pragmatique, la proposition 12.85 nous fournit une formule pour les limites de fonctions composée :

$$\lim_{x\to a}(f\circ g)(x) = \lim_{y\to \lim_{x\to a} g(x)} f(y) \tag{12.143}$$

lorsque f est continue.

Remarque 12.86.
La formule (12.143) ne peut pas être utilisée à l'envers. Il existe des cas où $\lim_{x\to a}(g\circ f)(x) = q$, et $\lim_{x\to a} f(x) = p$ sans pour autant avoir $\lim_{y\to q} g(y) = q$. Par exemple

$$g(x) = \begin{cases} 2 & \text{si } x \geqslant 0 \\ 0 & \text{si } x < 0 \end{cases} \tag{12.144a}$$
$$f(x) = |x|. \tag{12.144b}$$

Nous avons $(g\circ f)(x) = 2$ pour tout x, ainsi que $\lim_{x\to 0} f(x) = 0$, mais la limite $\lim_{y\to 0} g(y)$ n'existe pas.

Théorème 12.87 (Caractérisation de la limite par les suites)**.**
Une fonction $f\colon D \subset \mathbb{R}^m \to \mathbb{R}^n$ admet une limite ℓ en un point d'accumulation a de D si et seulement si pour toute suite (x_n) dans $D\backslash\{a\}$ convergente vers a, la suite $\big(f(x_n)\big)$ dans $\mathbb{R}^n$ converge vers ℓ.

Démonstration. Supposons d'abord que la fonction ait une limite ℓ lorsque $x \to a$, et considérons une suite (x_n) dans $D\backslash\{a\}$ convergente vers a. Nous devons montrer que la suite $y_n = f(x_n)$ converge vers ℓ, c'est à dire que si nous choisissons $\varepsilon > 0$ nous devons montrer qu'il existe un N tel que $n > N$ implique $\|y_n - \ell\| = \|f(x_n) - \ell\| < \varepsilon$.

Nous avons deux hypothèses. La première est la convergence de la fonction et la seconde est la convergence de la suite (x_n). L'hypothèse de convergence de la fonction nous dit que (le ε a déjà été choisit dans le paragraphe précédent)

$$\exists \delta \text{ tel que } 0 < \|x-a\| < \delta \Rightarrow \|f(x) - \ell\| < \varepsilon. \tag{12.145}$$

Une fois choisit ce δ qui «va avec» le ε qui a été choisit précédemment, la définition de la convergence de la suite nous enseigne que

$$\exists N \text{ tel que } n > N \Rightarrow \|x_n - a\| < \delta. \tag{12.146}$$

Récapitulons ce que nous avons fait. Nous avons choisi un ε, et puis nous avons construit un N. Lorsque $n > N$, nous avons $\|x_n - a\| < \delta$. Mais alors, par construction de ce δ, nous avons $\|f(x_n) - \ell\| < \varepsilon$. Au final, $n > N$ implique bien $\|y_n - \ell\| < \varepsilon$, ce qu'il nous fallait.

Nous supposons maintenant que la fonction f *ne* converge *pas* vers ℓ, et nous allons construire une suite d'éléments x_n qui converge vers a sans que $(y_n) = f(x_n)$ ne converge vers ℓ. La fonction f vérifie la condition (12.125). Nous prenons donc un ε tel que $\forall\delta$, il existe un x qui vérifie *en même temps* les deux conditions

$$\begin{cases} 0 < \|x - a\| < \delta & \text{(12.147a)} \\ \|f(x) - \ell\| > \varepsilon. & \text{(12.147b)} \end{cases}$$

Un tel x existe pour tout choix de δ. Choisissons un n arbitraire et $\delta = \frac{1}{n}$. Nous nommons x_n le x correspondant à ce choix de n. La suite (x_n) ainsi construite converge vers a parce que

$$\|x_n - a\| < \delta_n = \frac{1}{n}, \tag{12.148}$$

donc dès que n est grand, $\|x_n - a\|$ est petit. Mais la suite $y_n = f(x_n)$ ne converge pas vers ℓ parce que

$$\|f(x_n) - \ell\| > \varepsilon \tag{12.149}$$

pour tout n. La suite y_n ne s'approche donc jamais à moins d'une distance ε de ℓ. □

Nous avons déjà vu par le corollaire 6.128 qu'une suite croissante et bornée était convergente. Il en va de même pour les fonctions.

Proposition 12.88 ([1])**.**
Si la fonction réelle $f\colon I = [a, b[\to \mathbb{R}$ est croissante et bornée, alors la limite

$$\lim_{x \to b} f(x) \tag{12.150}$$

existe et est finie.

Démonstration. Commençons par prouver que si (x_n) est une suite dans I convergent vers b, alors $f(x_n)$ est une suite convergente. Dans un second temps nous allons prouver que si (x_n) et (x'_n) sont deux suites convergent vers b, alors les suites convergentes $f(x_n)$ et $f(x'_n)$ convergent vers la même limite. Alors le critère critère séquentiel de la limite d'une fonction conclura (proposition 6.138).

Nous pouvons extraire de x_n une sous-suite croissante $(x_{\alpha(n)})$. Alors la suite $f\big(x_{\alpha(n)}\big)$ est une suite croissante et majorée, donc convergente par le corollaire 6.128 [15]. Nommons ℓ la limite et montrons qu'elle est aussi limite de f sur la suite originale.

Pour tout $\epsilon > 0$, il existe K tel que si $n > K$ alors $\big|f\big(x_{\alpha(n)}\big) - \ell\big| < \epsilon$. Soit K' tel que pour tout $n > K'$ nous ayons $x_n > x_{\alpha(K')}$. Cela est possible parce que la suite est bornée par b et converge vers b : il suffit de prendre K' de telle sorte que $|x_n - b| \leqslant |x_{\alpha(n)} - b|$. Si $n > K'$ alors $x_n > x_{\alpha(K)}$ et

$$f(x_n) \geqslant f(x_{\alpha(n)}) \geqslant \ell - \epsilon; \tag{12.151}$$

en résumé si $n > K$ alors $|f(x_n) - \ell| < \epsilon$. Cela prouve que $f(x_n) \to \ell$.

Soit maintenant une autre suite (x'_n) convergent également vers b. Comme nous venons de le voir la suite $f(x'_n)$ est convergente et nous nommons ℓ' la limite. Si nous considérons (x''_n) la suite «alternée» $(x_1, x'_1, x_2, x'_2, \cdots)$ alors nous avons encore une suite convergente vers b et donc $f(x''_n) \to \ell'$.

Mais étant donné que $f(x_n)$ et $f(x'_n)$ sont des sous-suites, elles doivent converger vers la même valeur. Donc $\ell = \ell' = \ell''$. □

15. En gros nous sommes en train de dire que toute la théorie des fonctions convexes est un vulgaire corollaire de Bolzano-Weierstrass.

12.7.4 Règles simples de calcul

Les opérations simples passent à la limite, sauf la division pour laquelle il faut faire attention au dénominateur.

Proposition 12.89.
Soient f et g deux fonctions telles que $\lim_{x\to a} f(x) = \alpha$ et $\lim_{x\to a} g(x) = \beta$. Alors

(1) $\lim_{x\to a} f(x) + g(x) = \alpha + \beta$,

(2) $\lim_{x\to a} f(x)g(x) = \alpha\beta$,

(3) s'il existe un voisinage de a sur lequel g ne s'annule pas, alors $\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\alpha}{\beta}$.

Un résultat pratique pour calculer des limites est la

Proposition 12.90.
Quand la limite existe, nous avons

$$\lim_{x\to a} f(x) = \lim_{\epsilon\to 0} f(a+\epsilon),$$

ce qui correspond à un «changement de variables» dans la limite.

Démonstration. Si $A = \lim_{x\to a} f(x)$, par définition,

$$\forall \epsilon' > 0,\ \exists \delta \text{ tel que } |x-a| \leqslant \delta \Rightarrow |f(x) - A| \leqslant \epsilon'. \tag{12.152}$$

La seule subtilité de la démonstration est de remarquer que si $|x-a| \leqslant \delta$, alors x peut être écrit sous la forme $x = a+\epsilon$ pour un certain $|\epsilon| \leqslant \delta$. En remplaçant x par $a+\epsilon$ dans la condition 12.152, nous trouvons

$$\forall \epsilon' > 0,\ \exists \delta \text{ tel que } |\epsilon| \leqslant \delta \Rightarrow |f(x+\epsilon) - A| \leqslant \epsilon', \tag{12.153}$$

ce qui signifie exactement que $\lim_{\epsilon\to 0} f(x+\epsilon) = A$. □

Il y a une petite différence de point de vue entre $\lim_{x\to a} f(x)$ et $\lim_{\epsilon\to 0} f(a+\epsilon)$. Dans le premier cas, on considère $f(x)$, et on regarde ce qu'il se passe quand x se rapproche de a, tandis que dans le second, on considère $f(a)$, et on regarde ce qu'il se passe quand on s'éloigne un tout petit peu de a. Dans un cas, on s'approche très près de a, et dans l'autre on s'en éloigne un tout petit peu. Le contenu de la proposition 12.90 est de dire que ces deux points de vue sont équivalents.

12.8 Limites à plusieurs variables

Prenons une fonction $f\colon \mathbb{R}^n \to \mathbb{R}$. Nous disons que

$$\lim_{x\to x_0} f(x) = l \in \mathbb{R} \tag{12.154}$$

lorsque $\forall \epsilon > 0$, $\exists \delta$ tel que $\|x - x_0\| \leqslant \delta$ implique $|f(x) - l| \leqslant \epsilon$.

Remarquez qu'ici, $x \in \mathbb{R}^n$, et sachez distinguer $\|.\|$, la norme dans $\mathbb{R}^n$ de $|.|$ qui est la valeur absolue dans $\mathbb{R}$. Une autre façon d'exprimer cette définition est que l'ensemble des valeurs atteintes par f dans une boule de rayon δ autour de x_0 n'est pas très loin de l. Nous définissons donc

$$E_\delta = \{f(x) \text{ tel que } x \in B(x_0, \delta)\}. \tag{12.155}$$

Notez que si f n'est pas définie en x_0, il n'y a pas de valeurs correspondantes au centre de la boule dans E_δ. Ceci est évidement la situation générique lorsqu'il y a une indétermination à lever dans le calcul de la limite. Nous avons alors que

$$\lim_{x\to x_0} f(x) = l \tag{12.156}$$

lorsque $\forall \epsilon > 0$, $\exists \delta$ tel que

$$\sup\{|v - l| \text{ tel que } v \in E_\delta\} \leqslant \epsilon. \tag{12.157}$$

Une façon classique de montrer qu'une limite n'existe pas, est de prouver que, pour tout δ, l'ensemble E_δ contient deux valeurs constantes. Si par exemple $0 \in E_\delta$ et $1 \in E_\delta$ pour tout δ, alors aucune valeur de l (même pas $l = \pm\infty$) ne peut satisfaire à la condition (12.157) pour toute valeur de ϵ.

Nous laissons à la sagacité de l'étudiant le soin d'adapter tout ceci pour le cas $\lim_{x\to x_0} f(x) = \pm\infty$.

La proposition suivante semble évidente, mais nous sera tellement utile qu'il est préférable de l'expliciter :

Proposition 12.91.
Soit $f : D \to \mathbb{R}$ une fonction dont le domaine s'écrit comme une réunion finie

$$D = \bigcup_{i=1}^{k} A_i$$

où k est un entier. Soit $a \in \operatorname{Adh} D$ tel que $a \in \operatorname{Adh} A_i$ pour tout $i \leqslant k$, et soit $b \in \mathbb{R}$. Alors, la limite

$$\lim_{x\to a} f(x)$$

existe et vaut b si et seulement si chacune des limites

$$\lim_{\substack{x\to a \\ x\in A_i}} f(x)$$

existe et vaut b.

Démonstration. On sait déjà que si la limite de $f : D \to \mathbb{R}$ existe, alors toute restriction à A_i admet la même limite. Il suffit donc de prouver la réciproque.

Par hypothèse, pour tout $i = 1 \dots k$, nous savons que

$$\forall \epsilon > 0\, \exists \delta_i > 0 \text{ tel que } (x \in A_i) \text{ et } (\|x - a\| < \delta_i) \Rightarrow \|f(x) - b\| < \epsilon$$

Si ϵ est fixé, posons $\delta = \min_i\{\delta_i\}$. Nous savons alors que

(1) pour tout $x \in D$, il existe i tel que $x \in A_i$, et

(2) si x vérifie $\|x - a\| < \delta$, alors pour tout i, $\|x - a\| < \delta_i$ par définition de δ.

On en déduit que si $x \in D$ et $\|x - a\| < \delta$, alors il existe i tel que $x \in A_i$ et $\|x - a\| < \delta_i$, ce qui implique $|f(x) - b| < \epsilon$ et prouve la continuité. □

Exemple 12.92

(1) Pour qu'une fonction $f : \mathbb{R} \to \mathbb{R}$ admette une limite en $a \in \mathbb{R}$, il faut et il suffit qu'elle y admette une limite à droite et une limite à gauche qui soient égales.

(2) Une suite (x_k) admet une limite si et seulement si les sous suites (x_{2k}) et (x_{2k+1}) convergent vers la même limite. Ceci n'est pas une application directe de la proposition, mais la teneur est la même.

△

12.8.1 Règle de l'étau

Une première façon de calculer la limite d'une fonction est de la «coincer» entre deux fonctions dont nous connaissons la limite. Le théorème, que nous acceptons sans démonstration, est le suivant :

Théorème 12.93 (Règle de l'étau)**.**
Soit $\mathcal{O}$, un ouvert de $\mathbb{R}^m$ contenant le point a. Soient f, g et h, trois fonctions définies sur $\mathcal{O}$ (éventuellement pas en a lui-même). Supposons que pour tout $x \in \mathcal{O}$ (à part éventuellement a), nous ayons les inégalités

$$g(x) \leqslant f(x) \leqslant h(x). \tag{12.158}$$

Supposons de plus que

$$\lim_{x\to a} g(x) = \lim_{x\to a} h(x) = \ell. \tag{12.159}$$

Alors la limite $\lim_{x\to a} f(x)$ existe et vaut ℓ.

Nous insistons sur le fait que les deux fonctions entre lesquelles nous coinçons f doivent tendre vers *la même* valeur.

Cette méthode est très pratique lorsqu'on a des fonctions trigonométriques qui se factorisent parce qu'elles sont toujours majorables par 1.

Exemple 12.94
Prouvons que la fonction $f(x) = x\sin(x)$ tend vers zéro lorsque x tend vers 0. D'abord, nous coinçons la fonction entre deux fonctions connues :

$$0 \leqslant |x\sin(x)| = |x||\sin(x)| \leqslant |x|. \tag{12.160}$$

Donc $|x\sin(x)|$ est coincé entre $g(x) = 0$ et $h(x) = |x|$. Ces deux fonctions tendent vers 0 lorsque $x \to 0$, et donc $f(x)$ tend vers zéro. △

Exemple 12.95
Prouver la continuité en $(0,0)$ de la fonction

$$f(x,y) = \begin{cases} \frac{x|y|}{\sqrt{x^2+y^2}} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{sinon.} \end{cases} \tag{12.161}$$

Considérons une suite $(x_n, y_n) \in \mathbb{R}^2$ qui tend vers $(0,0)$. Étant donné que $\frac{|y|}{\sqrt{x^2+y^2}} < 1$ pour tout x et y, nous avons

$$0 \leqslant |f(x_n,y_n)| = \left|\frac{x_n|y_n|}{\sqrt{x_n^2+y_n^2}}\right| \leqslant |x_n| \to 0. \tag{12.162}$$

Donc nous avons

$$\lim_{(x,y)\to(0,0)} f(x,y) = 0 = f(0,0), \tag{12.163}$$

ce qui prouve que la fonction est continue en $(0,0)$ par la proposition 6.138. Nous avons utilisé la règle de l'étau (théorème 12.93). △

Nous notons $f \sim g$ pour $x \to a$ lorsque $\lim_{x\to a} \frac{f(x)}{g(x)} = 1$. Cela signifie que f et g tendent vers la même limite, à la même vitesse. Par exemple nous avons $\ln(1-x) \sim -x$ pour $x \to 0$ parce que

$$\lim_{x\to 0} -\frac{\ln(1-x)}{x} = \lim_{x\to 0} -\frac{\frac{-1}{1-x}}{1} = \lim_{x\to 0} \frac{1}{1-x} = 1 \tag{12.164}$$

où nous avons utilisé la règle de l'Hospital.

12.8.2 Méthode des chemins

Lorsque la limite n'existe pas, il y a une façon en général assez simple de le savoir, c'est la **méthode des chemin**.

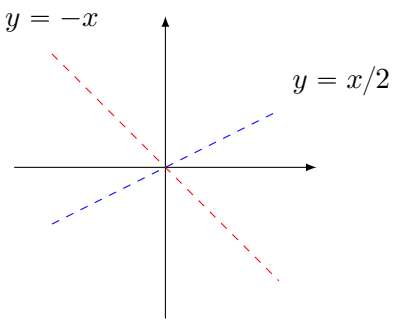


FIGURE 12.1 – Sur toute la droite $y = -x$, la fonction vaut $-1/2$, tandis que sur toute la droite $y = x/2$, elle vaut $\frac{2}{5}$. Il est donc impossible que la fonction ait une limite en $(0, 0)$, parce que dans toute boule autour de zéro, il y aura toujours un point de chacune de ces deux droites.

Exemple 12.96
Considérons la fonction

$$f(x, y) = \frac{xy}{x^2 + y^2}, \tag{12.165}$$

et remarquons que, quelle que soit la valeur de y, cette fonction est nulle lorsque $x = 0$. De la même manière, nous voyons que si $x = y$, alors la fonction vaut [16] $\frac{1}{2}$.

Il est impossible que la fonction ait une limite en $(0, 0)$ parce qu'on ne peut pas trouver un ℓ dont on s'approche à la fois en suivant la ligne $x = 0$ et la ligne $x = y$.

Deux autres chemins avec encore deux autres valeurs sont dessinés sur la figure 12.1.

△

Nous pouvons formaliser cet exemple en utilisant le théorème 12.87. Considérons les deux suites $x_n = (0, \frac{1}{n})$ et $y_n = (\frac{1}{n}, \frac{1}{n})$. Ce sont deux suites dans $\mathbb{R}^2$ qui tendent vers $(0, 0)$. Si la fonction f convergeait vers ℓ, alors nous aurions au moins

$$\lim f(x_n) = \ell \tag{12.166a}$$
$$\lim f(y_n) = \ell, \tag{12.166b}$$

mais nous savons que pour tout n, $f(x_n) = f(0, \frac{1}{n}) = 0$ et $f(y_n) = f(\frac{1}{n}, \frac{1}{n}) = \frac{1}{2}$. Il n'y a donc aucun nombre ℓ qui vérifie les deux équations (12.166) parce que $\lim f(x_n) = 0$ et $\lim f(y_n) = \frac{1}{2}$.

Tout ceci est formalisé et généralisé dans la proposition suivante.

Proposition 12.97.
Soit $f : D \subset \mathbb{R}^m \to \mathbb{R}^n$ et a un point d'adhérence de D. Alors nous avons

$$\lim_{x \to a} f(x) = \ell \tag{12.167}$$

si et seulement si pour toute fonction $\gamma : \mathbb{R} \to \mathbb{R}^m$ telle que $\lim_{t \to 0} \gamma(t) = a$, nous avons

$$\lim_{t \to 0} (f \circ \gamma)(t) = \ell. \tag{12.168}$$

Corollaire 12.98.
Soient $f : D \subset \mathbb{R}^m \to \mathbb{R}^n$ et a un point d'accumulation de D. Si nous avons deux fonctions $\gamma_1, \gamma_2 : \mathbb{R} \to \mathbb{R}^m$ telles que

$$\lim_{t \to 0} \gamma_1(t) = \lim_{t \to 0} \gamma_2(t) = a \tag{12.169}$$

tandis que

$$\lim_{t \to 0} (f \circ \gamma_1)(t) \neq \lim_{t \to 0} (f \circ \gamma_2)(t), \tag{12.170}$$

ou bien que l'une des deux limites n'existe pas, alors la limite de $f(x)$ lorsque $x \to a$ n'existe pas.

16. En fait ce que nous sommes en train de faire est de poser $\theta = \pi/2$ et $\theta = \pi/4$ dans (12.178).

Corollaire 12.99.
Soient $f\colon D \subset \mathbb{R}^m \to \mathbb{R}^n$ et a un point d'accumulation de D. Si il existe une fonction $\gamma\colon \mathbb{R} \to \mathbb{R}^m$ avec $\gamma(0) = a$ telle que la limite $\lim_{t\to 0}(f \circ \gamma)(t)$ n'existe pas, alors la limite $\lim_{x\to a} f(x)$ n'existe pas.

En ce qui concerne le calcul de limites, la méthode des chemins peut être utilisé de trois façons :

(1) Dès que l'on trouve une fonction $\gamma\colon \mathbb{R} \to \mathbb{R}^m$ telle que $\lim_{t\to 0}(f \circ \gamma)(t) = \ell$, alors nous savons que *si la limite* $\lim_{x\to a} f(x)$ *existe*, alors cette limite vaut ℓ.
(2) Dès que l'on a trouvé deux fonctions γ_i qui tendent vers a, mais dont les limites de $\lim_{t\to 0}(f \circ \gamma_i)(t)$ sont différentes, alors la limite $\lim_{x\to a} f(x)$ n'existe pas.
(3) Dès qu'on trouve une chemin le long duquel il n'y a pas de limite, alors la limite n'existe pas (corollaire 12.99).

La méthode des chemins ne permet donc pas de de calculer une limite quand elle existe. Elle permet uniquement de la «deviner», ou bien de prouver que la limite n'existe pas.

Exemple 12.100
Soit à calculer

$$\lim_{(x,y)\to(0,0)} \frac{x-y}{x+y}. \tag{12.171}$$

Si nous prenons le chemin $\gamma_1(t) = (t, t)$, nous avons bien $\lim_{t\to 0} \gamma_1(t) = (0, 0)$, et nous avons

$$\lim_{t\to 0}(f \circ \gamma_1)(t) = \lim_{t\to 0} \frac{t-t}{t+t} = 0. \tag{12.172}$$

Donc si la limite (12.171) existait, elle vaudrait obligatoirement 0. Mais si nous considérons $\gamma_2(t) = (0, t)$, nous avons

$$(f \circ \gamma_2)(t) = \frac{-t}{t} = -1, \tag{12.173}$$

donc si la limite existe, elle doit obligatoirement valoir -1. Ne pouvant être égale à 0 et à -1 en même temps, la limite (12.171) n'existe pas. △

12.8.3 Méthode des coordonnées polaires

La proposition suivante exprime la définition de la limite en d'autres termes, et va être pratique dans le calcul de certaines limites.

Proposition 12.101.
Soit $f\colon D \subset \mathbb{R}^m \to \mathbb{R}^n$, a un point d'accumulation de D et $\ell \in \mathbb{R}^n$. Nous définissons

$$E_r = \{f(x) \text{ tel que } x \in B(a, r) \cap D\}, \tag{12.174}$$

et

$$s_r = \sup\{\|v - \ell\| \text{ tel que } v \in E_r\}. \tag{12.175}$$

Alors nous avons $\lim_{x\to a} f(x) = \ell$ si et seulement si $\lim_{r\to 0} s_r = 0$.

Dans cette proposition, E_r représente l'ensemble des valeurs atteintes par f dans un rayon r autour de a. Le nombre s_r sélectionne, parmi toutes ces valeurs, celle qui est la plus éloignée de ℓ et donne la distance. En d'autres termes, s_r est la distance maximale entre $f(x)$ et ℓ lorsque x est à une distance au maximum r de a.

Lorsque nous avons affaire à une fonction $f\colon \mathbb{R}^2 \to \mathbb{R}$, cette proposition nous permet de calculer facilement les limites en passant aux coordonnées polaires.

Exemple 12.102
Reprenons la fonction de l'exemple 12.96 :

$$f(x, y) = \frac{xy}{x^2 + y^2}. \tag{12.176}$$

Son domaine est $\mathbb{R}^2\backslash\{(0,0)\}$. Nous voulons calculer $\lim_{(x,y)\to(0,0)} f(x,y)$. Écrivons la définition de E_r :

$$E_r = \{f(x,y) \text{ tel que } (x,y) \in B\big((0,0),r\big)\}. \tag{12.177}$$

Les points de la boule sont, en coordonnées polaires, les points de la forme (ρ,θ) avec $\rho < r$. La chose intéressante est que $f(\rho,\theta)$ est relativement simple (plus simple que la fonction départ). En effet en remplaçant tous les x par $\rho\cos(\theta)$ et tous les y par $\rho\sin(\theta)$, et en utilisant le fait que $\cos^2(\theta)+\sin^2(\theta)=1$, nous trouvons

$$f(\rho,\theta) = \frac{\rho^2\cos(\theta)\sin(\theta)}{\rho^2} = \cos(\theta)\sin(\theta). \tag{12.178}$$

Cela signifie que

$$E_r = \{\cos(\theta)\sin(\theta) \text{ tel que } \theta \in [0,2\pi[\}. \tag{12.179}$$

Prenons ℓ quelconque. Le nombre s_r est le supremum des

$$\|\ell - \cos(\theta)\sin(\theta)\| \tag{12.180}$$

lorsque θ parcours $[0,2\pi]$. Nous ne sommes pas obligés calculer la valeur exacte de s_r. Ce qui compte ici est que s_r ne vaut certainement pas zéro, et ne dépend pas de r. Donc il est impossible d'avoir $\lim_{r\to 0} s_r = 0$, et la fonction donnée n'a pas de limite en $(0,0)$. △

Nous pouvons retenir cette règle pour calculer les limites lorsque $(x,y)\to(0,0)$ de fonctions $f\colon\mathbb{R}^2\to\mathbb{R}$:

(1) passer en coordonnées polaires, c'est à dire remplacer x par $\rho\cos(\theta)$ et y par $\rho\sin(\theta)$;

(2) nous obtenons une fonction g de ρ et θ. Si la limite $\lim_{r\to 0} g(r,\theta)$ n'existe pas ou dépend de θ, alors la fonction n'a pas de limite. Si on peut majorer g par une fonction ne dépendant pas de θ, et que cette fonction a une limite lorsque $r\to 0$, alors cette limite est la limite de la fonction.

La vraie difficulté de la technique des coordonnées polaire est de trouver le supremum de E_r, ou tout au moins de montrer qu'il est borné par une fonction qui a une limite qui ne dépend pas de θ. Une de situations classiques dans laquelle c'est facile est lorsque la fonction se présente comme une fonction de r multiplié par une fonction de θ.

Exemple 12.103
Soit à calculer la limite

$$\lim_{(x,y)\to(0,0)} xy\left(\frac{x^2-y^2}{x^2+y^2}\right). \tag{12.181}$$

Le passage aux coordonnées polaires donne

$$f(r,\theta) = r^2\sin\theta\cos\theta(\cos^2\theta-\sin^2\theta). \tag{12.182}$$

Déterminer le supremum de cela est relativement difficile. Mais nous savons que de toutes façons, la quantité $\sin\theta\cos\theta(\cos^2\theta-\sin^2\theta)$ est bornée par 1. Donc

$$\|f(r,\theta)\| \leqslant r^2. \tag{12.183}$$

Maintenant la règle de l'étau montre que $\lim_{(x,y)\to(0,0)} f(x,y)$ est zéro.

La situation vraiment gênante serait celle avec une fonction de θ qui risque de s'annuler dans un dénominateur. △

Exemple 12.104
Soit à calculer

$$\lim_{(x,y)\to(0,0)} \frac{x^2+y^2}{x-y}. \tag{12.184}$$

Le passage en polaires donne

$$f(r,\theta) = \frac{r^2}{r\big(\cos(\theta) - \sin(\theta)\big)} = \frac{r}{\cos(\theta) - \sin(\theta)}. \tag{12.185}$$

Certes *pour chaque* θ nous avons $\lim_{r\to 0} f(r,\theta) = 0$, mais il ne faut pas en déduire trop vite que la limite $\lim_{(x,y)\to(0,0)} f(x,y)$ vaut zéro parce que prendre la limite $r \to 0$ avec θ fixé revient à prendre la limite le long de la droite d'angle θ.

Il n'est pas possible de majorer $f(r,\theta)$ par une fonction ne dépendant pas de θ parce que cette fonction tend vers l'infini lorsque $\theta \to \pi/4$. Est-ce que cela veut dire que la limite n'existe pas ? Cela veut en tout cas dire que la méthode des coordonnées polaires ne parvient pas à résoudre l'exercice. Pour conclure, il faudra encore un peu travailler.

Nous pouvons essayer de calculer le long d'un chemin plus général $(r(t),\theta(t))$. Choisissons $r(t) = t$ puis cherchons $\theta(t)$ de telle sorte à avoir

$$\cos\theta(t) - \sin\theta(t) = t^2. \tag{12.186}$$

Le mieux serait de résoudre cette équation pour trouver $\theta(t)$. Mais en réalité il n'est pas nécessaire de résoudre : montrer qu'il existe une solution suffit. Nous pouvons supposer que $t^2 < 1$. Pour $\theta = \pi/4$ nous avons $\cos(\theta) - \sin(\theta) = 0$ et pour $\theta = 0$ nous avons $\cos(\theta) - \sin(\theta) = 1$. Le théorème des valeurs intermédiaires nous enseigne alors qu'il existe une valeur de θ qui résout l'équation (12.186).

Pour être rigoureux, nous devons aussi montrer que la fonction $\theta(t)$ est continue. Pour cela il faudrait utiliser le théorème de la fonction implicite. Nous verrons dans l'exemple 12.106 comment s'en sortir sans théorème de la fonction implicite, au prix de plus de calculs. △

Exemple 12.105
Considérons encore la fonction

$$f(x,y) = \frac{x^2 + y^2}{x - y}. \tag{12.187}$$

Une mauvaise idée pour prouver que la limite n'existe pas pour $(x,y) \to (0,0)$ est de considérer le chemin (t,t). En effet, la fonction n'existe pas sur ce chemin. Or la méthode des chemins parle uniquement de chemins contenus dans le domaine de la fonction. △

Exemple 12.106
Revenons encore et toujours sur la fonction

$$(x^2 + y^2)/(x - y). \tag{12.188}$$

Nous prouvons que la limite n'existe pas en trouvant des chemins le long desquels les limites sont différentes. Si nous essayons le chemin (t,kt) avec k constant, nous trouvons

$$f(t,kt) = \frac{t(1+k^2)}{1-k}. \tag{12.189}$$

La limite $t \to 0$ est hélas toujours 0. Nous ne pouvons donc pas conclure.

Nous allons maintenant utiliser la même technique que celle utilisée en coordonnées polaires. Vous noterez que dans ce cas, travailler en cartésiennes donne lieu à des calculs plus longs. L'astuce consiste à prendre k non constant et à chercher par exemple $k(t)$ de façon à avoir

$$\frac{1 + k(t)^2}{1 - k(t)} = \frac{1}{t}. \tag{12.190}$$

Avec une telle fonction, la fonction $t \mapsto f(t,tk(t))$ serait la constante 1. L'équation à résoudre pour k est

$$tk^2 + k + (t-1) = 0, \tag{12.191}$$

et les solutions sont

$$k(t) = \frac{-1 \pm \sqrt{1 - 4t(t-1)}}{2t}. \tag{12.192}$$

Nous proposons donc les chemins

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} t \\ \frac{-1 \pm \sqrt{1-4t(t-1)}}{2} \end{pmatrix} \tag{12.193}$$

Nous devons vérifier deux points. D'abord que ce chemin est bien défini, et ensuite que $tk(t)$ tend bien vers zéro lorsque $t \to 0$ (sinon $(t, k(t)t)$) n'est pas un chemin passant par $(0,0)$. Lorsque t est petit, ce qui se trouve sous la racine est proche de 1 et ne pose pas de problèmes. Ensuite,

$$\lim_{t \to 0} tk(t) = \frac{-1 \pm 1}{2}. \tag{12.194}$$

En choisissant le signe $+$, nous trouvons un chemin qui nous convient.

Ce que nous avons prouvé est que

$$f\left(t, \frac{-1 + \sqrt{1 - 4t(t-1)}}{2}\right) = 1 \tag{12.195}$$

pour tout t. Le long de ce chemin, la limite de f est donc 1. Cette limite est différente des limites obtenues le long de chemins avec k constant. La limite $\lim_{(x,y)\to(0,0)} f(x,y)$ n'existe donc pas. △

Exemple 12.107

Considérons la fonction (figure 12.2)

$$f(x,y) = \begin{cases} \sqrt{x^2+y^2} \sin \frac{1}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0), \end{cases} \tag{12.196}$$

et cherchons la limite $(x,y) \to (0,0)$. Le passage en coordonnées polaires donne

$$f(\rho, \theta) = \rho \sin \frac{1}{\rho}. \tag{12.197}$$

Pour calculer la limite de cela lorsque $\rho \to 0$, nous remarquons que

$$0 \leqslant |\rho \sin \frac{1}{\rho}| \leqslant \rho \tag{12.198}$$

parce que $\sin(\frac{1}{\rho}) \leqslant 1$ quel que soit ρ. Or évidement $\lim_{\rho \to 0} \rho = 0$, donc la limite de la fonction (12.197) est zéro et ne dépend pas de θ. Nous en concluons que $\lim_{(x,y)\to(0,0)} f(x,y) = 0$. △

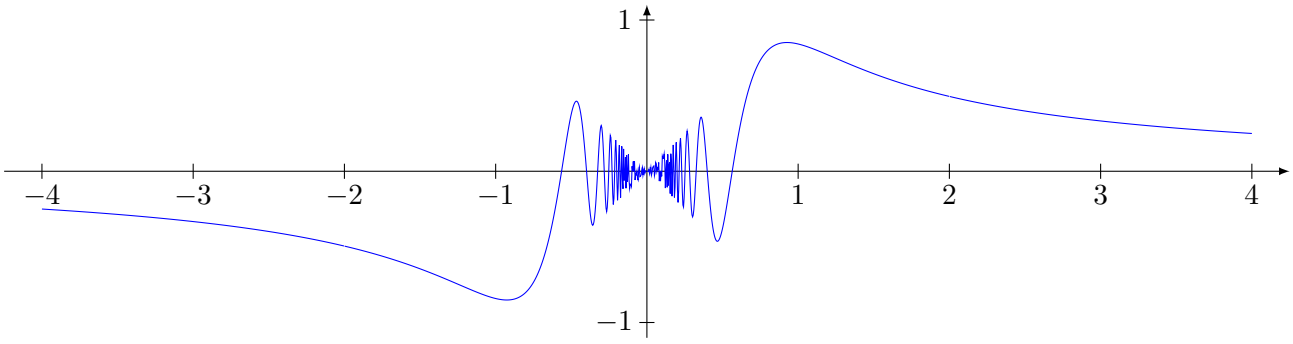


FIGURE 12.2 – La fonction de l'exemple 12.107.

12.8.4 Méthode du développement asymptotique

Nous savons que nous pouvons développer certaines fonctions en série grâce au développement de Taylor (théorème 12.263). Lorsque nous avons une limite à calculer, nous pouvons remplacer certaines parties de la fonction à traiter par la formule (12.677b). Cela est très utile pour comparer des fonctions trigonométrique à des polynômes.

Exemple 12.108
La limite $\lim_{x\to 0} \frac{\sin(x)}{x} = 1$ est bien connue. Une manière de la prouver des d'écrire

$$\sin(x) = x + h(x) \tag{12.199}$$

avec $h \in o(x)$, c'est à dire $\lim_{x\to 0} h(x)/x = 0$. Alors nous avons

$$\lim_{x\to 0} \frac{\sin(x)}{x} = \lim_{x\to 0} \frac{x + h(x)}{x} = \lim_{x\to 0} \frac{x}{x} + \lim_{x\to 0} \frac{h(x)}{x} = 1. \tag{12.200}$$

△

L'utilisation de la proposition 12.85 permet d'utiliser cette technique dans le cadre de limites à plusieurs variables. Reprenons l'exemple 12.108 un tout petit peu modifié :

Exemple 12.109
Soit à calculer $\lim_{(x,y)\to(0,0)} f(x,y)$ où

$$f(x,y) = \frac{\sin(xy)}{xy}. \tag{12.201}$$

La première chose à faire est de voir f comme la composée de fonctions $f = f_1 \circ f_2$ avec

$$\begin{aligned} f_1 \colon \mathbb{R} &\to \mathbb{R} \\ t &\mapsto \frac{\sin(t)}{t} \end{aligned} \tag{12.202}$$

et

$$\begin{aligned} f_2 \colon \mathbb{R}^2 &\to \mathbb{R} \\ (x,y) &\mapsto xy. \end{aligned} \tag{12.203}$$

Étant donné que $\lim_{(x,y)\to(0,0)} f_2(x,y) = 0$, nous avons $\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{t\to 0} f_1(t) = 1$.
△

12.9 Dérivée : exemples introductifs

12.9.1 La vitesse

Lorsqu'un mobile se déplace à une vitesse variable, nous obtenons la *vitesse instantanée* en calculant une vitesse moyenne sur des intervalles de plus en plus petits. Si le mobile a un mouvement donné par $x(t)$, la vitesse moyenne entre $t = 2$ et $t = 5$ sera

$$v_{\text{moy}}(2 \to 5) = \frac{x(5) - x(2)}{5 - 2}.$$

Plus généralement, la vitesse moyenne entre 2 et $2 + \Delta t$ est donnée par

$$v_{\text{moy}}(2 \to 2 + \Delta t) = \frac{x(2 + \Delta t) - x(2)}{\Delta t}.$$

Cela est une fonction de Δt. Oui, mais je te rappelle qu'on a dans l'idée de calculer une vitesse instantanée, c'est à dire de voir ce que vaut la vitesse moyenne sur un intervalle très très très très

petit. La notion de limite semble toute indiquée pour décrire mathématiquement l'idée physique de vitesse instantanée.

Nous allons dire que la vitesse instantanée d'un mobile est la limite quand Δt tends vers zéro de sa vitesse moyenne sur l'intervalle de temps Δt, ou en formule :

$$v(t_0) = \lim_{\Delta t \to 0} \frac{x(t_0) - x(t_0 + \Delta t)}{\Delta t}. \tag{12.204}$$

12.9.2 La tangente à une courbe

Passons maintenant à tout autre chose, mais toujours dans l'utilisation de la notion de limite pour résoudre des problèmes intéressants. Comment trouver l'équation de la tangente à la courbe $y = f(x)$ au point $(x_0, f(x_0))$?

Essayons de trouver la tangente au point P donné de la courbe donnée à la figure 12.3.

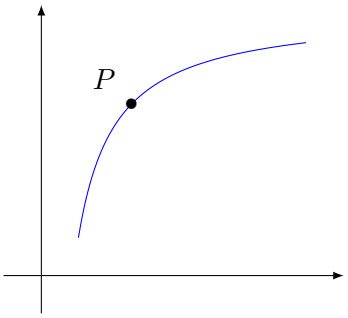


FIGURE 12.3 – Comment trouver la tangente à la courbe au point P ?

La tangente est la droite qui touche la courbe en un seul point sans la traverser. Afin de la construire, nous allons dessiner des droites qui touchent la courbe en P et un autre point Q, et nous allons voir ce qu'il se passe quand Q est très proche de P. Cela donnera une droite qui, certes, touchera la courbe en deux points, mais en deux point *tellement proche que c'est comme si c'étaient les mêmes*. Tu sens que la notion de limite va encore venir.

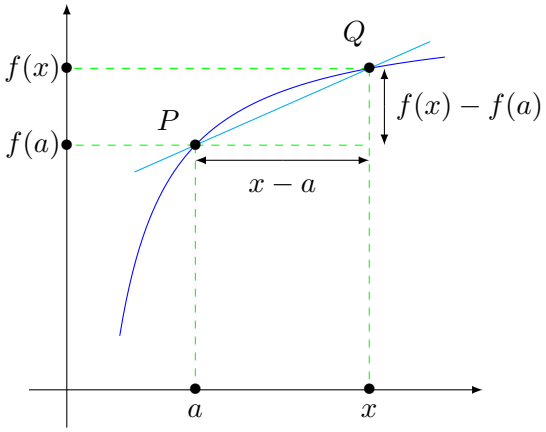


FIGURE 12.4 – Traçons d'abord une corde entre le point P et un point Q un peu plus loin.

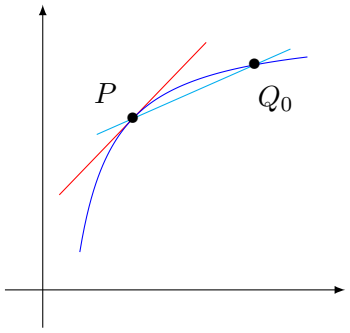
Nous avons placé le point, sur la figure 12.4, le point P en a et le point Q un peu plus loin x. En d'autres termes leurs coordonnées sont

$$P = \big(a, f(a)\big) \qquad\qquad Q = \big(x, f(x)\big). \tag{12.205}$$

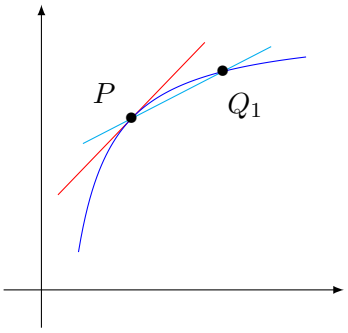
Comme tu devrais le savoir sans même regarder la figure 12.4, le coefficient directeur de la droite qui passe par ces deux points est donné par

$$\frac{f(x) - f(a)}{x - a}, \tag{12.206}$$

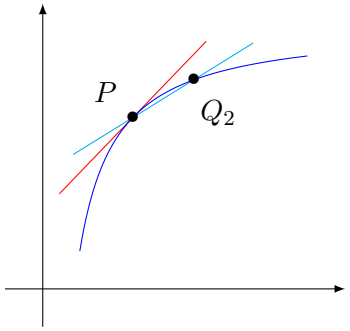
et bang ! Encore le même rapport que celui qu'on avait trouvé à l'équation (12.204) en parlant de vitesses. Si tu regardes la figure 12.5, tu verras que réellement en faisant tendre x vers a on obtient la tangente.



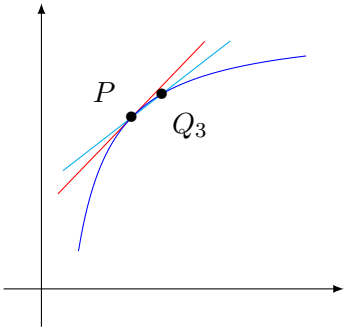
(a) Pas très bon ...



(b) ... de mieux en mieux ...



(c) ... de mieux en mieux ...



(d) ... presque parfait

FIGURE 12.5 – Recherche de la tangente par approximations successives.

12.9.3 L'aire en dessous d'une courbe

Encore un exemple. Nous voudrions bien pouvoir calculer l'aire en dessous d'une courbe. Nous notons $S_f(x)$ l'aire en dessous de la fonction f entre l'abscisse 0 et x, c'est à dire l'aire bleue de la figure 12.6.

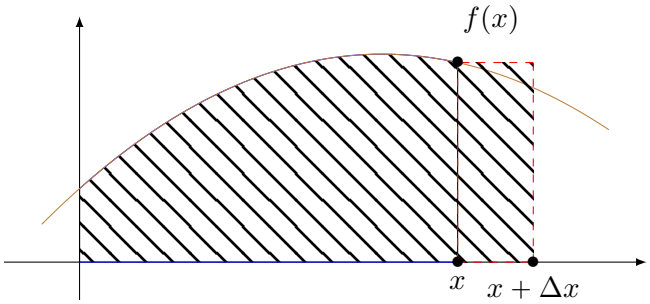


FIGURE 12.6 – L'aire en dessous d'une courbe. Le rectangle rouge d'aire $f(x)\Delta x$ approxime l'augmentation de l'aire lorsqu'on passe de x à $x + \Delta x$.

Si la fonction f est continue et que Δx est assez petit, la fonction ne varie pas beaucoup entre x et $x + \Delta x$. L'augmentation de surface entre x et $x + \Delta x$ peut donc être approximée par le rectangle de surface $f(x)\Delta x$. Ce que nous avons donc, c'est que quand Δx est très petit,

$$S_f(x + \Delta x) - S_f(x) = f(x)\Delta x, \tag{12.207}$$

c'est à dire

$$f(x) = \lim_{\Delta x \to 0} \frac{S_f(x + \Delta x) - S_f(x)}{\Delta x}. \tag{12.208}$$

Donc, la fonction f est la dérivée de la fonction qui représente l'aire en dessous de f. Calculer des surfaces revient donc au travail inverse de calculer des dérivées.

Nous avons déjà vu que calculer la dérivée d'une fonction n'est pas très compliqué. Aussi étonnant que cela puisse paraître, il se fait que le processus inverse est très compliqué : il est en général extrêmement difficile (et même souvent impossible) de trouver une fonction dont la dérivée est une fonction donnée.

Une fonction dont la dérivée est la fonction f s'appelle une **primitive** de f, et la fonction qui donne l'aire en dessous de la fonction f entre l'abscisse 0 et x est notée

$$S_f(x) = \int_0^x f(t)dt. \tag{12.209}$$

Nous pouvons nous demander si, pour une fonction f donnée, il existe une ou plusieurs primitives, c'est à dire s'il existe une ou plusieurs fonctions F telles que $F' = f$. La réponse viendra...

12.10 Continuité et dérivabilité

On considère dans la suite une fonction $f : A \to \mathbb{R}$, où $a \in A \subset \mathbb{R}$; cependant, les notions de continuité et de dérivabilité se généralisent immédiatement au cas de fonctions à valeurs vectorielles ; la notion de continuité se généralise au cas des fonctions à plusieurs variables (la notion de dérivabilité est remplacée par celle de différentiabilité dans ce cadre).

Définition 12.110.
*La fonction f est **dérivable** en a si $a \in \operatorname{int} A$ et si*

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{f(x) - f(a)}{x - a}$$

existe. On note alors cette quantité $f'(a)$, c'est le nombre dérivé de f en a. La fonction dérivée *de f est*

$$f' : A' \to \mathbb{R} : a \mapsto f'(a)$$

définie sur l'ensemble noté A' des points a où f est dérivable.

Exemple 12.111
Montrons que la fonction $f : \mathbb{R} \to \mathbb{R} : x \mapsto x$ est continue et dérivable. Exceptionnellement (bien qu'on sache que la dérivabilité implique la continuité), montrons ces deux assertions séparément.

Continuité Pour prouver la continuité au point $a \in \mathbb{R}$ nous devons montrer que

$$\lim_{x \to a} x = a \tag{12.210}$$

c'est-à-dire

$$\forall \epsilon > 0, \exists \delta > 0 : \forall x \in \mathbb{R}\, |x - a| < \delta \Rightarrow |x - a| < \epsilon \tag{12.211}$$

ce qui est clair en prenant $\delta = \epsilon$.

Dérivabilité Soit $a \in \mathbb{R}$. Calculons la limite du quotient différentiel

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{x - a}{x - a} = \lim_{\substack{x \to a \\ x \neq a}} 1 = 1 \tag{12.212}$$

ce qui prouve que f est dérivable et que sa dérivée vaut 1 en tout point a de $\mathbb{R}$.

On a donc montré que la fonction $x \mapsto x$ est continue, dérivable, et que sa dérivée vaut 1 en tout point a de son domaine.

△

Proposition 12.112.
Une fonction dérivable sur un intervalle y est continue.

Démonstration. Soit I un intervalle sur lequel la fonction f est dérivable, et soit $x_0 \in I$. Nous allons prouver la continuité de f en x_0. Le fait que la limite

$$f'(x_0) = \lim_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h} \tag{12.213}$$

existe implique a fortiori que

$$\lim_{h\to 0} f(x_0+h) - f(x_0) = 0. \tag{12.214}$$

Cela signifie la continuité de f en vertu du critère 12.35. □

12.11 Dérivation de fonctions d'une variable réelle

12.11.1 Exemples

12.11.1.1 La fonction $f(x) = x$

Commençons par la fonction $f(x) = x$. Dans ce cas nous avons

$$\frac{f(x) - f(a)}{x - a} = \frac{x - a}{x - a} = 1. \tag{12.215}$$

La dérivée est donc 1.

Proposition 12.113.
La dérivé de la fonction $x \mapsto x$ vaut 1, en notations compactes : $(x)' = 1$.

Démonstration. D'après la définition de la dérivée, si $f(x) = x$, nous avons

$$f(x) = \lim_{\epsilon\to 0} \frac{(x+\epsilon) - x}{\epsilon} = \lim_{\epsilon\to 0} \frac{\epsilon}{\epsilon} = 1, \tag{12.216}$$

et c'est déjà fini. □

12.11.1.2 La fonction $f(x) = x^2$

Prenons ensuite $f(x) = x^2$. En utilisant le produit remarquable $(x^2 - a^2) = (x-a)(x+a)$ nous trouvons

$$\frac{f(x) - f(a)}{x - a} = x + a. \tag{12.217}$$

Lorsque $x \to a$, cela devient $2a$. Nous avons par conséquent

$$f'(x) = 2x. \tag{12.218}$$

Lemme 12.114.
Si $f(x) = x^2$, alors $f'(x) = 2x$.

Démonstration. Utilisons la définition, et remplaçons f par sa valeur :

$$\begin{aligned}
f'(x) &= \lim_{\epsilon\to 0} \frac{f(x+\epsilon) - f(x)}{\epsilon} && \text{(12.219a)}\\
&= \lim_{\epsilon\to 0} \frac{(x+\epsilon)^2 - x^2}{\epsilon} && \text{(12.219b)}\\
&= \lim_{\epsilon\to 0} \frac{x^2 + 2x\epsilon + \epsilon^2 - x^2}{\epsilon} && \text{(12.219c)}\\
&= \lim_{\epsilon\to 0} \frac{\epsilon(2x+\epsilon)}{\epsilon} && \text{(12.219d)}\\
&= \lim_{\epsilon\to 0} (2x+\epsilon) && \text{(12.219e)}\\
&= 2x, && \text{(12.219f)}
\end{aligned}$$

ce qu'il fallait prouver. □

12.11.1.3 La fonction $f(x) = \sqrt{x}$

Considérons maintenant la fonction $f(x) = \sqrt{x}$. Nous avons

$$\begin{aligned}\frac{f(x)-f(a)}{x-a} &= \frac{\sqrt{x}-\sqrt{a}}{x-a}\\ &= \frac{(\sqrt{x}-\sqrt{a})(\sqrt{x}+\sqrt{x})}{(x-a)(\sqrt{x}+\sqrt{x})}\\ &= \frac{1}{\sqrt{x}+\sqrt{x}}.\end{aligned} \tag{12.220}$$

Lorsque $x \to 0$, nous obtenons

$$f'(a) = \frac{1}{2\sqrt{a}}. \tag{12.221}$$

Notons que la dérivée de $f(x) = \sqrt{x}$ n'existe pas en $x = 0$. En effet elle serait donnée par le quotient

$$f'(0) = \lim_{x\to 0} \frac{\sqrt{x}-\sqrt{0}}{x} = \lim_{x\to 0} \frac{\sqrt{x}}{x} = \lim_{x\to 0} \frac{1}{\sqrt{x}}. \tag{12.222}$$

Mais si x devient très petit, la dernière fraction tend vers l'infini.

12.11.2 Calcul de la dérivée

Proposition 12.115.
Soit $f, g\colon I \subset \mathbb{R} \to \mathbb{R}$ deux fonctions dérivables. Alors nous avons les propriétés suivante

(1) la fonction $h = f + g$ est dérivable et $h'(x) = f'(x) + g'(x)$.

(2) la fonction $h = fg$ est dérivable et

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x). \tag{12.223}$$

*Cette formule est appelée **règle de Leibnitz**.*

(3) la fonction $h = \frac{f}{g}$ est dérivable en tout point x tel que $g(x) \neq 0$ et

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}. \tag{12.224}$$

(4) la fonction $h = f \circ g$ est dérivable et

$$(f \circ g)'(x) = f'\big(g(x)\big)g'(x). \tag{12.225}$$

Proposition 12.116.
Si $f(x) = x^n$ avec $n \in \mathbb{N}$, alors $f'(x) = nx^{n-1}$.

Démonstration. Nous avons déjà vu que la proposition était vraie avec $n = 1$ et $n = 2$. Supposons qu'elles soit vraie avec $n = k$, et prouvons qu'elle est vraie pour $n = k + 1$. Nous avons

$$x^{k+1} = xx^k. \tag{12.226}$$

En utilisant la règle de Leibnitz et l'hypothèse de récurrence,

$$\begin{aligned}\big(x^{k+1}\big)' &= (x)'x^k + x\big(x^k\big)'\\ &= x^k + x\big(kx^{k-1}\big)\\ &= x^k + kx^k\\ &= (k+1)x^k,\end{aligned} \tag{12.227}$$

ce qu'il fallait démontrer. □

12.11.3 Interprétation géométrique de la dérivée : tangente

Considérons le **graphe** de la fonction f sur I, c'est à dire l'ensemble

$$\big\{\big(x, f(x)\big) \text{ tel que } x \in I\big\}. \tag{12.228}$$

Le nombre

$$\frac{f(x) - f(a)}{x - a} \tag{12.229}$$

est la pente de la droite qui joint les points $\big(x, f(x)\big)$ et $\big(a, f(a)\big)$, voir la figure 12.7.

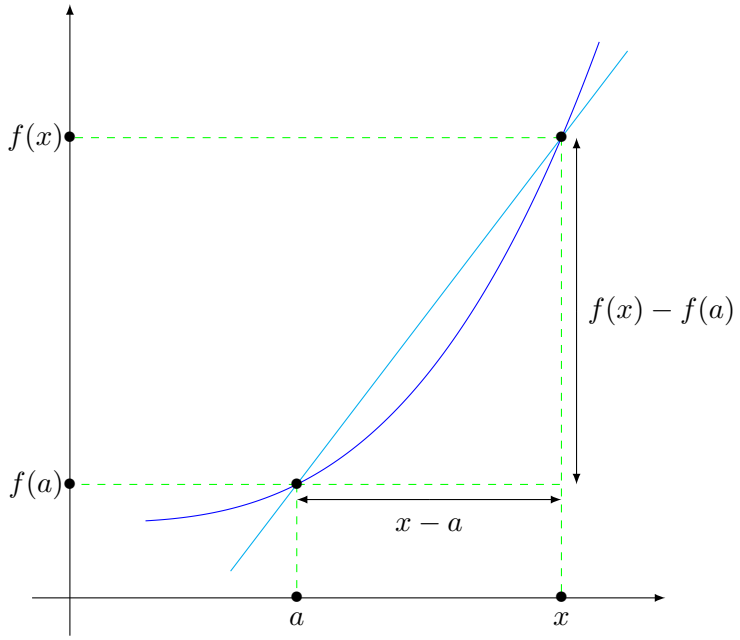


FIGURE 12.7 – Le coefficient directeur de la corde entre a et x.

Étant donné que $f'(a)$ est le coefficient directeur de la tangente au point $\big(a, f(a)\big)$, l'équation de la tangente est

$$y - f(a) = f'(a)(x - a). \tag{12.230}$$

12.11.4 Interprétation géométrique de la dérivée : approximation affine

Le fait que la fonction f soit dérivable au point $a \in I$ signifie que

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \ell \tag{12.231}$$

pour un certain nombre ℓ. Cela peut être récrit sous la forme

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} - \ell = 0, \tag{12.232}$$

ou encore

$$\lim_{x \to a} \frac{f(x) - f(a) - \ell(x - a)}{x - a} = 0. \tag{12.233}$$

Introduisons la fonction

$$\alpha(t) = \frac{f(a + t) - f(a) - t\ell}{t}. \tag{12.234}$$

Cette fonction est faite exprès pour que

$$\alpha(x - a) = \frac{f(x) - f(a) - \ell(x - a)}{x - a}; \tag{12.235}$$

par conséquent $\lim_{x\to a}\alpha(x-a)=0$. Nous récrivons l'équation (12.235) sous la forme

$$f(x)-f(a)-\ell(x-a)=(x-a)\alpha(x-a). \tag{12.236}$$

Le second membre tend vers zéro lorsque x tend vers a avec une «vitesse au carré» : c'est le produit de deux facteurs tous deux tendant vers zéro. Si x n'est pas très loin de a, il n'est donc pas une mauvaise approximation de dire

$$f(x)-f(a)-\ell(x-a)\simeq 0, \tag{12.237}$$

c'est à dire

$$f(x)\simeq f(a)+f'(a)(x-a). \tag{12.238}$$

Nous avons retrouvé l'équation (12.230). La manipulation que nous venons de faire revient donc à dire que la fonction f, au voisinage de a, est bien approximée par sa tangente.

L'équation (12.238) peut être aussi écrite sous la forme

$$f(x+\Delta x)\simeq f(x)+f'(x)\Delta x \tag{12.239}$$

qui est une approximation d'autant meilleure que Δx est petit.

Proposition 12.117 ([138])**.**
Soit f une fonction dérivable et strictement monotone de l'intervalle I sur l'intervalle J (f est alors une bijection de I vers J). Si ne s'annule par sur alors

(1) la fonction f est une bijection de I vers J,

(2) la fonction f^{-1} est dérivable sur J,

(3) et nous avons la formule

$$(f^{-1})'=\frac{1}{f'\circ f^{-1}}. \tag{12.240}$$

12.12 Opérations sur les dérivées

Pour continuer, nous allons en faire une un peu plus abstraite.

Proposition 12.118.
La dérivation est une opération linéaire, c'est à dire que

(1) $(\lambda f)'=\lambda f'$ pour tout réel λ où, pour rappel, la fonction (λf) est définie par $(\lambda f)(x)=\lambda\cdot f(x)$,

(2) $(f+g)'=f'+g'$.

Démonstration. Ces deux propriétés découlent des propriétés correspondantes de la limite. Nous allons faire la première, et laisser la seconde à titre d'exercice. Écrivons la définition de la dérivée avec (λf) au lieu de f, et calculons un petit peu :

$$\begin{aligned}(\lambda f)'(x)&=\lim_{\epsilon\to 0}\frac{(\lambda f)(x+\epsilon)-(\lambda f)(x)}{\epsilon}\\&=\lim_{\epsilon\to 0}\frac{\lambda\big(f(x+\epsilon)\big)-\lambda f(x)}{\epsilon}\\&=\lim_{\epsilon\to 0}\lambda\frac{f(x+\epsilon)-f(x)}{\epsilon}\\&=\lambda\lim_{\epsilon\to 0}\frac{f(x+\epsilon)-f(x)}{\epsilon}\\&=\lambda f'(x).\end{aligned} \tag{12.241}$$

□

Proposition 12.119.
La dérivée d'un produit obéit à la ***règle de Leibnitz*** *:*

$$(fg)'(x) = f'(x)g(x) + f(g)g'(x). \tag{12.242}$$

Cette règle est souvent écrite sous la forme compacte $(fg)' = f'g + g'f$.

Démonstration. La définition de la dérivée dit que

$$(fg)'(x) = \lim_{\epsilon\to 0} \frac{f(x+\epsilon)g(x+\epsilon) - f(x)g(x)}{\epsilon}. \tag{12.243}$$

La subtilité est d'ajouter au numérateur la quantité $-f(x)g(x+\epsilon)+f(x)g(x+\epsilon)$, ce qui est permit parce que cette quantité est nulle [17]. Le numérateur de (12.243) devient donc

$$\begin{aligned} f(x+\epsilon)g(x+\epsilon) - f(x)g(x+\epsilon) + f(x)g(x+\epsilon) - f(x)g(x) \\ = g(x+\epsilon)\big(f(x+\epsilon) - f(x)\big) + f(x)\big(g(x+\epsilon) - g(x)\big), \end{aligned} \tag{12.244}$$

où nous avons effectué deux mises en évidence. Étant donné que nous avons deux termes, nous pouvons couper la limite en deux :

$$\begin{aligned} (fg)'(x) &= \lim_{\epsilon\to 0} g(x+\epsilon)\frac{f(x+\epsilon)-f(x)}{\epsilon} \qquad + \lim_{\epsilon\to 0} f(x)\frac{g(x+\epsilon)-g(x)}{\epsilon} \\ &= \lim_{\epsilon\to 0} g(x+\epsilon) \lim_{\epsilon\to 0} \frac{f(x+\epsilon)-f(x)}{\epsilon} \quad + f(x) \lim_{\epsilon\to 0} \frac{g(x+\epsilon)-g(x)}{\epsilon}, \end{aligned} \tag{12.245}$$

où nous avons utilisé le théorème 12.34 pour scinder la première limite en deux, ainsi que la propriété (12.20) pour sortir le $f(x)$ de la limite dans le second terme. Maintenant, dans le premier terme, nous avons évidement [18] $\lim_{\epsilon\to 0} g(x+\epsilon) = g(x)$. Les limites qui restent sont les définitions classiques des dérivées de f et g au point x :

$$(fg)'(x) = g(x)f'(x) - f(x)g'(x), \tag{12.246}$$

ce qu'il fallait démontrer. □

12.12.1 Développement limité au premier ordre

Si une fonction est dérivable en a alors elle peut être approximée «au premier ordre» par une formule simple.

Proposition 12.120 (Développement limité au premier ordre)**.**
Si f est dérivable en a alors nous avons la formule

$$f(a+h) = f(a) + hf'(a) + \alpha(h) \tag{12.247}$$

pour une fonction α telle que

$$\lim_{h\to 0} \frac{\alpha(h)}{h} = 0. \tag{12.248}$$

Ce résultat sera généralisé pour des dérivées d'ordre supérieures avec les séries de Taylor, théorème 12.263.

Démonstration. La fonction f étant dérivable en a nous avons l'existence de la limite suivante :

$$f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}, \tag{12.249}$$

17. Le coup d'ajouter et enlever la même chose a déjà été fait durant la démonstration du théorème 12.34. C'est une technique assez courante en analyse.

18. Pas tout à fait évidemment : selon le théorème 12.35, *limite et continuité*, il faut que g soit continue.

ce qui revient à dire qu'en définissant la fonction β par

$$f'(a) = \frac{f(a+h) - f(a)}{h} + \beta(h) \tag{12.250}$$

alors $\beta(h) \to 0$ lorsque $h \to 0$. En multipliant par h et en nommant $\alpha(h) = h\beta(h)$ nous trouvons le résultat :

$$f(a+h) = f(a) + hf'(a) + \alpha(h) \tag{12.251}$$

avec

$$\lim_{h\to 0} \frac{\alpha(h)}{h} = \lim_{h\to 0} \beta(h) = 0. \tag{12.252}$$

□

12.13 Dérivation à une variable

12.13.1 Définition

Définition 12.121 (Fonction dérivable)**.**
*Nous disons qu'une fonction f est **dérivable** au point $x_0 \in I$ si la limite*

$$\lim_{\epsilon\to 0} \frac{f(x_0+\epsilon) - f(x_0)}{\epsilon} \tag{12.253}$$

existe.

Si f est une fonction dérivable, rien n'empêche la fonction dérivée f' d'être elle-même dérivable. Dans ce cas nous notons f'' ou $f^{(2)}$ la dérivée de la fonction f'. Cette fonction f'' est la **dérivée seconde** de f. Elle peut encore être dérivable ; dans ce cas nous notons $f^{(3)}$ sa dérivée, et ainsi de suite. Nous définissons $f^{(n)} = (f^{(n-1)})'$ la dérivée n^{e}de f. Nous posons évidemment $f^{(0)} = f$.

Théorème 12.122.
Toute fonction f dérivable au point x_0 est continue au point x_0.

Remarque 12.123.
La réciproque du théorème précédent n'est pas vraie : il existent bien des fonctions qui sont continues à un point x_0 mais qui ne sont pas dérivables en x_0. La fonction valeur absolue, $x \mapsto |x|$, par exemple est continue sur tout $\mathbb{R}$ mais elle n'est pas dérivable en 0.

12.13.2 Quelques formules à connaître

À retenir 12.124

$$(\alpha f(x) + \beta g(x))' = \alpha f'(x) + \beta g'(x). \tag{12.254a}$$

$$(f(x)g(x))' = f'(x)g(x) + f(x)g'(x). \tag{12.254b}$$

$$(f(u(x)))' = f'(u(x))u'(x). \tag{12.254c}$$

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}. \tag{12.254d}$$

12.14 Dérivation et croissance

Supposons une fonction dont la dérivée est positive. Étant donné que la courbe est « collée » à ses tangentes, tant que les tangentes montent, la fonction monte. Or, une tangente qui monte correspond à une dérivée positive, parce que la dérivée est le coefficient directeur de la tangente.

Ce résultat très intuitif peut être prouvé rigoureusement. C'est la tache à laquelle nous allons nous atteler maintenant.

Proposition 12.125.
Si f et f' sont des fonctions continues sur l'intervalle $[a,b]$ et si f' est strictement positive sur $[a,b]$, alors f est croissante sur $[a,b]$.

De la même manière, si f' est strictement négative sur $[a,b]$, alors f est décroissante sur $[a,b]$.

Démonstration. Nous n'allons prouver que la première partie. La seconde partie se prouve en considérant $-f$ et en invoquant alors la première [19]. Prenons x_1 et x_2 dans $[a,b]$ tels que $x_1 < x_2$. Par hypothèse, pour tout x dans $[x_1, x_2]$, nous avons

$$f'(x) = \lim_{\epsilon \to 0} \frac{f(x+\epsilon) - f(x)}{\epsilon} > 0. \tag{12.255}$$

Maintenant, la proposition 12.33 dit que quand une limite est positive, alors la fonction dans la limite est positive sur un voisinage. En appliquant cette proposition à la fonction

$$r(\epsilon) = \frac{f(x+\epsilon) - f(x)}{\epsilon}, \tag{12.256}$$

dont la limite en zéro est positive, nous trouvons que $r(\epsilon) > 0$ pour tout ϵ pas trop éloigné de zéro. En particulier, il existe un $\delta > 0$ tel que $\epsilon < \delta$ implique $r(\epsilon) > 0$; pour un tel ϵ, nous avons donc

$$r(\epsilon) = \frac{f(x+\epsilon) - f(x)}{\epsilon} > 0. \tag{12.257}$$

Étant donné que $\epsilon > 0$, nous avons que $f(x+\epsilon) - f(x) > 0$, c'est à dire que f est strictement croissante entre x et $x + \delta$.

Jusqu'ici, nous avons prouvé que la fonction f était strictement croissante dans un voisinage autour de chaque point de $[a,b]$. Cela n'est cependant pas encore tout à fait suffisant pour conclure. Ce que nous voudrions faire, c'est de dire, c'est prendre un voisinage $]a, m_1[$ autour de a sur lequel f est croissante. Donc, $f(m_1) > f(a)$. Ensuite, on prend un voisinage $]m_1, m_2[$ de m_1 sur lequel f est croissante. De ce fait, $f(m_2) > f(m_1) > f(a)$. Et ainsi de suite, nous voulons construire des m_3, m_4,... jusqu'à arriver en b. Hélas, rien ne dit que ce processus va fonctionner. Il faut trouver une subtilité. Le problème est que les voisinages sur lesquels la fonction est croissante sont peut-être de plus en plus petit, de telle sorte à ce qu'il faille une infinité d'étapes avant d'arriver à bon port (en b).

Heureusement, nous pouvons drastiquement réduire le nombre d'étapes en nous souvenant du théorème de Borel-Lebesgue (numéro 12.66). Nous notons par $\mathcal{O}_x$, un ouvert autour de x tel que f soit strictement croissante sur $\mathcal{O}_x$. Un tel voisinage existe. Cela fait une infinité d'ouverts tels que

$$[a,b] \subseteq \bigcup_{x \in [a,b]} \mathcal{O}_x. \tag{12.258}$$

Ce que le théorème dit, c'est qu'on peut en choisir un nombre fini qui recouvre encore $[a,b]$. Soient $\{\mathcal{O}_{x_1}, \dots, \mathcal{O}_{x_n}\}$, les heureux élus, que nous supposons prit dans l'ordre : $x_1 < x_2 < \dots < x_n$. Nous avons

$$[a,b] \subseteq \bigcup_{i=1}^{n} \mathcal{O}_i. \tag{12.259}$$

Quitte à les rajouter à la collection, nous supposons que $x_1 = a$ et que $x_n = b$. Maintenant nous allons choisir encore un sous ensemble de cette collection d'ouverts. On pose $\mathcal{A}_1 = \mathcal{O}_{x_1}$. Nous savons que $\mathcal{A}_1$ intersecte au moins un des autres $\mathcal{O}_{x_i}$. Cette affirmation vient du fait que $[a,b]$ est connexe (proposition 6.119), et que si $\mathcal{O}_{x_1}$ n'intersectait personne, alors

$$\mathcal{O}_{x_1} \quad \text{et} \quad \bigcup_{i=2}^{n} \mathcal{O}_{x_i} \tag{12.260}$$

19. Méditer cela.

forment une partition de $[a,b]$ en deux ouverts disjoints, ce qui n'est pas possible parce que $[a,b]$ est connexe. Nous nommons $\mathcal{A}_2$, un des ouverts $\mathcal{O}_{x_i}$ qui intersecte $\mathcal{A}_1$. Disons que c'est $\mathcal{O}_k$. Notons que $\mathcal{A}_1 \cup \mathcal{A}_2$ est un intervalle sur lequel f est strictement croissante. En effet, si y_{12} est dans l'intersection, $f(a) < f(y_{12})$ parce que f est strictement croissante sur $\mathcal{A}_1$, et pour tout $x > y_{12}$ dans $\mathcal{A}_2$, $f(x) > f(y_{12})$ parce que f est strictement croissante dans $\mathcal{A}_2$.

Maintenant, nous éliminons de la liste des $\mathcal{O}_{x_i}$ tous ceux qui sont inclus à $\mathcal{A}_1 \cup \mathcal{A}_2$. Dans ce qu'il reste, il y en a automatiquement un qui intersecte $\mathcal{A}_1 \cup \mathcal{A}_2$, pour la même raison de connexité que celle invoquée plus haut. Nous appelons cet ouvert $\mathcal{A}_3$, et pour la même raison qu'avant, f est strictement croissante sur $\mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{A}_3$.

En recommençant suffisamment de fois, nous finissons par devoir prendre un des $\mathcal{O}_{x_i}$ qui contient b, parce qu'au moins un des $\mathcal{O}_{x_i}$ contient b. À ce moment, nous avons finit la démonstration. □

Il est intéressant de noter que ce théorème concerne la croissance d'une fonction sous l'hypothèse que la dérivée est positive. Il nous a fallu très peu de temps, en utilisant la positivité de la dérivée, pour conclure qu'autour de tout point, la fonction était strictement croissante. À partir de là, c'était pour ainsi dire gagné. Mais il a fallu un réel travail de topologie très fine [20] pour conclure. Étonnant qu'une telle quantité de topologie soit nécessaire pour démontrer un résultat essentiellement analytique dont l'hypothèse est qu'une limite est positive, n'est-ce pas ?

Une petite facile, maintenant.

Proposition 12.126.
Si f est croissante sur un intervalle, alors $f' \geqslant 0$ à l'intérieur cet intervalle, et si f est décroissante sur l'intervalle, alors $f' \leqslant 0$ à l'intérieur de l'intervalle.

Note qu'ici, nous demandons juste la croissance de f, et non sa *stricte* croissance.

Démonstration. Soit f, une fonction croissante sur l'intervalle I, et x un point intérieur de I. La dérivée de f en x vaut

$$f'(x) = \lim_{\epsilon \to 0} \frac{f(x+\epsilon) - f(x)}{\epsilon}, \tag{12.261}$$

mais, comme f est croissante sur I, nous avons toujours que $f(x+\epsilon) - f(x) \geqslant 0$ quand $\epsilon > 0$, et $f(x+\epsilon) - f(x) \leqslant 0$ quand $\epsilon < 0$, donc cette limite est une limite de nombre positifs ou nuls, qui est donc positive ou nulle. Cela prouve que $f'(x) \geqslant 0$. □

12.14.1 Théorèmes de Rolle et des accroissements finis

Théorème 12.127 (Théorème de Rolle[139])**.**
Soit f, une fonction continue sur $[a,b]$ et dérivable sur $]a,b[$. Si $f(a) = f(b)$, alors il existe un point $c \in]a,b[$ tel que $f'(c) = 0$.

Démonstration. Étant donné que $[a,b]$ est un intervalle compact, l'image de $[a,b]$ par f est un intervalle compact, soit $[m,M]$ (théorème 6.77). Si $m = M$, alors le théorème est évident : c'est que la fonction est constante, et la dérivée est par conséquent nulle. Supposons que $M > f(a)$ (il se peut que $M = f(a)$, mais alors si f n'est pas constante, il faut avoir $m < f(a)$ et le reste de la preuve peut être adaptée).

Comme M est dans l'image de $[a,b]$ par f, il existe $c \in]a,b[$ tel que $f(c) = M$. Considérons maintenant la fonction

$$\tau(x) = \frac{f(c+x) - f(c)}{x}. \tag{12.262}$$

Par définition, $\lim_{x \to 0} \tau(x) = f'(c)$. Par hypothèse, si $u < c$,

$$\tau(u-c) = \frac{f(u) - f(c)}{u - c} > 0 \tag{12.263}$$

20. et je te rappelle que nous avons utilisé la proposition 6.119, qui elle même était déjà un très gros boulot !

parce que $u-c<0$ et $f(u)-f(c)<0$. Par conséquent, $\lim_{x\to 0}\tau(x)\geqslant 0$. Nous avons aussi, pour $v>c$,

$$\tau(v-c)=\frac{f(v)-f(c)}{v-c}<0 \tag{12.264}$$

parce que $v-c>0$ et $f(v)-f(c)<0$. Par conséquent, $\lim_{x\to 0}\tau(x)\leqslant 0$. Mettant les deux ensemble, nous avons $f'(c)=\lim_{x\to 0}\tau(x)=0$, et c est le point que nous cherchions. □

Voici une généralisation du théorème de Rolle, dans le cas où nous n'aurions pas deux points sur lesquels la fonction est identique, mais deux points en lesquelles la limite de la fonction est identique. Typiquement, lorsque les points en question sont $\pm\infty$.

Théorème 12.128 (Généralisation de Rolle[139])**.**
Soient $-\infty\leqslant a<b\leqslant+\infty$. Soit une fonction dérivable $f\colon]a,b[\to\mathbb{R}$ telle que

$$\lim_{x\to a}f(x)=\lim_{x\to b}f(x)=\ell \tag{12.265}$$

avec $\ell\in\bar{\mathbb{R}}$. Alors il existe $x\in]a,b[$ tel que $f'(x)=0$.

Démonstration. Soit un difféomorphisme strictement croissant $\varphi\colon\mathbb{R}\to]\alpha,\beta[$. Pour cela vous pouvez bricoler à partir de l'exemple 12.44. Mais n'utilisez pas la fonction arc tangente, parce qu'elle n'est définie qu'au théorème 16.38.

Nous posons $a'=\varphi(a)$, $b'=\varphi(b)$ et

$$g=\varphi\circ f\circ\varphi^{-1}\colon]a',b'[\to]\alpha,\beta[. \tag{12.266}$$

Cela est une fonction dérivable et continue sur $[a',b']$ en posant $g(a')=g(b')=\varphi(\ell)$.

Donc il existe $c'\in]a',b'[$ tel que $g'(c')=0$. En posant $c=\varphi^{-1}(c')$ nous avons $c\in]a,b[$ et, en utilisant de nombreuses fois la règle de dérivation des fonctions composées 12.115(4),

$$\begin{aligned}
f'(c)&=f'\big(\varphi^{-1}(c')\big) && \text{(12.267a)}\\
&=(\varphi^{-1})'\Big((g\circ\varphi)\big(\varphi^{-1}(c')\big)\Big)(g\circ\varphi)'\big(\varphi^{-1}(c')\big) && \text{(12.267b)}\\
&=(\varphi^{-1})'\big(g(c')\big)\underbrace{g'(c')}_{=0}\varphi'\big(\varphi^{-1}(c')\big) && \text{(12.267c)}\\
&=0. && \text{(12.267d)}
\end{aligned}$$

□

Le théorème suivant est le théorème des **accroissements finis**.

Théorème 12.129 (Accroissements finis)**.**
Soit f, une fonction continue sur $[a,b]$ et dérivable sur $]a,b[$.

(1) Il existe au moins un réel $c\in]a,b[$ tel que

$$f(b)-f(a)=(b-a)f'(c). \tag{12.268}$$

Autrement dit, la tangente en c est parallèle à la corde entre a et b.

(2) Nous avons la majoration

$$\big|f(b)-f(a)\big|\leqslant\sup_{x\in[a,b]}|f'(x)||b-a|. \tag{12.269}$$

Démonstration. Considérons la fonction

$$\tau(x)=f(x)-\big(\frac{f(b)-f(a)}{b-a}x+f(a)-a\frac{f(b)-f(a)}{b-a}\big), \tag{12.270}$$

c'est à dire la fonction qui donne la distance entre f et le segment de droite qui lie $(a, f(a))$ à $(b, f(b))$. Par construction, $\tau(a) - \tau(b) = 0$, donc le théorème de Rolle s'applique à τ pour laquelle il existe donc un $c \in]a, b[$ tel que $\tau'(c) = 0$.

En utilisant les règles de dérivation, nous trouvons que la dérivée de τ vaut

$$\tau'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}, \tag{12.271}$$

donc dire que $\tau'(c) = 0$ revient à dire que $f(b) - f(a) = (b - a)f'(c)$, ce qu'il fallait démontrer.

La majoration est une conséquence immédiate, parce que le supremum de $|f'(x)|$ est forcément plus grand que $|f'(c)|$. □

Une généralisation pour une fonction sur un intervalle $]a, b[$ où a et b peuvent être infinis.

Théorème 12.130 (Généralisation des accroissements finis)**.**
Soient $-\infty \leqslant a < b \leqslant +\infty$ et f, g des fonctions continues sur $[a, b]$ et dérivables sur $]a, b[$.

Alors il existe $c \in]a, b[$ tel que

$$\big(f(b) - f(a)\big)g'(c) = \big(g(b) - g(a)\big)f'(c). \tag{12.272}$$

Quelque précisions sur l'énoncé avant la preuve. Si a ou b est infini, l'intervalle de continuité s'ouvre du ou des côtés qu'il faut ; dans ce cas la continuité est matérialisé par le fait que nous notons $f(a)$, $f(b)$, $g(a)$ et $g(b)$ les limites correspondantes et que nous supposons ces limites finies.

Démonstration. Nous posons

$$h(t) = \big(g(b) - g(a)\big)f(t) - \big(f(b) - f(a)\big)g(t). \tag{12.273}$$

Nous avons $\lim_{t\to a} h(t) = \lim_{t\to a} h(t)$, de telle sorte que le théorème de Rolle généralisé 12.128 s'applique et il existe $c \in]a, b[$ tel que $h'(c) =$. Pour ce c nous avons

$$h'(c) = \big(f(b) - f(a)\big)g'(c) - \big(g(b) - g(a)\big)f'(c). \tag{12.274}$$

□

12.14.2 Règle de l'Hospital

Proposition 12.131 (Règle de l'Hospital pour $\frac{0}{0}$[140])**.**
Soient des fonctions f, g dérivables sur $]a, b[$ et dont la limite en a est nulle. Si g' ne s'annule pas sur $]a, b[$ et si

$$\lim_{x\to a^+} \frac{f'(x)}{g'(x)} = \ell \tag{12.275}$$

alors

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \ell. \tag{12.276}$$

Ici $\ell \in \bar{\mathbb{R}}$, et les hypothèses garantissent l'existence de la limite (12.276)*.*

Démonstration. Soit $x \in]a, b[$. Les fonctions f et g sont dérivables sur $]a, x[$ et continues sur $[a, x]$, de telle sorte que le théorème 12.130 s'applique et nous avons $c_x \in]a, x[$ tel que

$$\big(f(x) - f(a)\big)g'(c_x) = \big(g(x) - g(a)\big)f'(c_x). \tag{12.277}$$

Nous nous souvenons de ce que signifient les notations dans le théorème : les notations $f(a)$, $f(x)$, $g(a)$ et $g(x)$ désignent en réalité les limites. Donc dans (12.277), nous avons $f(a) = g(a) = 0$.

D'autre part nous avons $g(x) \neq g(a)$, sinon le théorème de Rolle 12.128 annulerait g' quelque part dans $]a, x[$. Nous pouvons donc récrire (12.277) sous la forme

$$\frac{f(x)}{g(x)} = \frac{f'(c_x)}{g'(c_x)}. \tag{12.278}$$

Mais $\lim_{x\to a^+} c_x = a$ parce que $c_x \in]a, x[$. Donc la limite du membre de droite de (12.278) lorsque $x \to a^+$ existe et vaut ℓ. La même limite à gauche doit alors exister et valoir la même valeur. □

Proposition 12.132.
Soit f et g deux fonctions

(1) dérivables sur $]a,b[$,

(2) dont les limites en a sont toutes deux ∞,

(3) $g' \neq 0$ sur $]a,b[$.

(4)

$$\lim_{x\to a^+} \frac{f'(x)}{g'(x)} = \ell \in \bar{\mathbb{R}}. \tag{12.279}$$

Alors

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \ell. \tag{12.280}$$

Cette dernière égalité signifie «la limite existe et vaut ℓ».

Démonstration. Soit un intervalle $]x,y[$ strictement inclus à $]a,b[$ avec $x,y \in \mathbb{R}$. Par le théorème des accroissements finis généralisés 12.130 il existe $x \in]x,y[$ tel que

$$\frac{f(x)-f(y)}{g(x)-g(y)} = \frac{f'(c)}{g'(c)}. \tag{12.281}$$

Notons que le dénominateur à gauche n'est pas nul à cause de Rolle et de l'hypothèse que g' ne s'annule pas sur $[x,y]$. Nous isolons $f(x)$:

$$f(x) = \frac{f'(c)}{g'(c)}\Big(g(x)-g(y)\Big) + f(y). \tag{12.282}$$

Avant de diviser par $g(x)$ nous devons prendre quelque précautions. Soit V, un voisinage de ℓ[21]. Vu la limite (12.279), il existe $y \in]a,b[$ tel que

$$\frac{f'(t)}{g'(t)} \in V \tag{12.283}$$

pour tout $t \in]a,y[$. Nous utilisons ici avec subtilité le fait que ces intervalles sont une base de la topologie autour de ∞. Maintenant $f(y)$ et $g(y)$ sont fixés et sont des nombres réels. Vu que $\lim_{x\to a} g(x) = 0$ nous pouvons choisir $r < y$ tel que nous ayons simultanément

(1) $g(x) \neq 0$ sur $]a,r[$,

(2)

$$|\frac{g(y)}{g(x)}| \leqslant \epsilon \tag{12.284}$$

et

$$|\frac{f(y)}{g(x)}| \leqslant \epsilon \tag{12.285}$$

pour tout $x \in]a,r[$.

Nous sommes maintenant armés de y et r satisfaisant tout cela et nous pouvons traiter avec la formule (12.282) en ne la considérant que pour $x \in]a,r[$. Soit $x \in]a,r[$; il existe $c_x \in]a,x[$ tel que

$$\frac{f(x)}{g(x)} = \frac{f'(c_x)}{g'(c_x)}\left(1-\frac{g(y)}{g(x)}\right) + \frac{f(y)}{g(x)}. \tag{12.286}$$

Nous avons :

(1) $\lim_{x\to a^+} c_x = a$,

(2)

$$\lim_{x\to a^+} \frac{f'(c_x)}{g'(c_x)} = \lim_{x\to a^+} \frac{f'(x)}{g'(x)} = \ell, \tag{12.287}$$

21. Vous savez ce que signifie un «voisinage de ∞» ? Allez voir la sous-section 6.10.6.

(3)
$$\lim_{x\to a^+}\left(1-\frac{g(y)}{g(x)}\right)=1, \tag{12.288}$$

(4) $\lim_{x\to a^+}\frac{f(y)}{g(x)}=0$.

Donc chaque partie du membre de droite de (12.286) a une limite bien déterminée pour $x\to a^+$. Les règles de calcul s'appliquent, la limite du membre de gauche existe et vaut

$$\lim_{x\to a^+}\frac{g(x)}{g(x)}=\ell\times 1+0=\ell. \tag{12.289}$$

□

12.14.3 À propos de primitives

Corollaire 12.133.
Soit f une fonction dérivable sur $[a,b]$ telle que $f'(x)=0$ pour tout $x\in[a,b]$. Alors f est constante sur $[a,b]$.

Démonstration. Si f n'était pas constante sur $[a,b]$, il existerait un $x_1\in]a,b[$ tel que $f(a)\neq f(x_1)$, et dans ce cas, il existerait, par le théorème des accroissements finis 12.129, un $c\in]a,x_1[$ tel que

$$f'(c)=\frac{f(x_1)-f(a)}{x_1-a}\neq 0, \tag{12.290}$$

ce qui contredirait les hypothèses. □

Corollaire 12.134.
Soient f et g, deux fonctions dérivables sur $[a,b]$ telles que

$$f'(x)=g'(x) \tag{12.291}$$

pour tout $x\in[a,b]$. Alors existe un réel C tel que $f(x)=g(x)+C$ pour tout $x\in[a,b]$.

Démonstration. Considérons la fonction $h(x)=f(x)-g(x)$, dont la dérivée est, par hypothèse, nulle. L'annulation de la dérivée entraine par le corollaire 12.134 que h est constante. Si $h(x)=C$, alors $f(x)=g(x)+C$, ce qu'il fallait prouver. □

Définition 12.135.
Soit I un intervalle ouvert de $\mathbb{R}$ et une fonction $f\colon I\to\mathbb{R}$. La fonction $F\colon I\to\mathbb{R}$ est une ***primitive*** *de f si F est dérivable sur I et si $F'(x)=f(x)$ pour tout x dans I.*

Exprimé en termes des primitives, le corollaire 12.134 signifie que

Corollaire 12.136.
Si F et G sont deux primitives de la même fonction f sur un intervalle, alors il existe une constante C pour laquelle $F(x)=G(x)+C$.

Cela signifie qu'il n'y a, en réalité, pas des milliards de primitives différentes à une fonction. Il y en a essentiellement une seule, et puis les autres, ce sont juste les mêmes, mais décalées d'une constante.

Remarque 12.137.
L'hypothèse de se limiter à un intervalle est importante parce que si on considère la fonction sur deux intervalles disjoints, nous pouvons choisir la constante indépendamment dans l'un et dans l'autre. Par exemple la fonction

$$F(x)=\begin{cases}\ln(x)+1 & \text{si } x>0\\ \ln(x)-7 & \text{si } x<0\end{cases} \tag{12.292}$$

est une primitive de $\frac{1}{x}$ sur l'ensemble $\mathbb{R}\backslash\{0\}$.

Certains ne s'en privent pas. Le logiciel Sage par exemple fait ceci :

```
sage: f(x)=1/x
sage: F=f.integrate(x)
sage: A=F(x)-F(-x)
sage: A.full_simplify()
I*pi
```

En réalité lorsque $x > 0$, Sage définit $\ln(-x) = \ln(x) + i\pi$. Cela a une certaine logique parce que $\ln(-1) = i\pi$ (du fait que $e^{i\pi} = -1$), mais si on ne le sait pas, ça peut étonner.

12.138.
Il existe plusieurs primitives à une fonction donnée. En physique, la constante arbitraire est souvent fixée par une condition initiale, comme nous le verrons dans la section 39.1.

12.15 Fonctions de plusieurs variables

La physique, et les sciences en général, regorge de fonctions à plusieurs variables.

Accélération centripète [22] Si une masse m tourne sur un cercle, elle subira une accélération dirigée vers l'intérieur égale à

$$F(v, r) = \frac{mv^2}{r} \tag{12.293}$$

où r est le rayon du cercle et v est la vitesse.

Pression dans un gaz Si on a n moles de gaz dans un volume V a une température T, alors la pression sera donnée par la fonction de trois variables

$$p = \frac{nRT}{V} \tag{12.294}$$

où R est la constante des gaz parfaits.

En mathématique, on peut inventer de nombreuses fonctions de plusieurs variables. La fonction

$$f(x, y) = x^2 + xy\cos(x^2 + y^3) \tag{12.295}$$

est définie sur $\mathbb{R}^2$. La fonction

$$f(x, y, z) = \frac{x + y - 2z}{1 - x^2 - y^2 - z^2} \tag{12.296}$$

est définie sur $\mathbb{R}^3$ moins la sphère unité $\{x^2 + y^2 + z^2 = 1\}$.

La plus grande partie de ce cours est consacrée à l'étude des fonction de plusieurs variables. Nous allons maintenant donner quelques indication sur comment «dessiner» une telle fonction. Vous connaissez déjà la définition de graphe pour une fonction f d'une seule variable à valeurs dans $\mathbb{R}$: c'est l'ensemble des point du plan de la forme $(x, f(x))$. Vous voyez que cet ensemble n'est pas vraiment un gros morceau de $\mathbb{R}^2$ parce que son intérieur est vide : il y a une seule valeur de f qui correspond au point x, donc une boule de $\mathbb{R}^2$ centrée en $(x, f(x))$ de n'importe quel rayon contient toujours des points qui ne font pas partie du graphe de f.

Nous voulons donner une définition assez générale pour le graphe d'une fonction

Définition 12.139.
*Soit f une fonction de $\mathbb{R}^m$ dans $\mathbb{R}^n$. Le **graphe** de f est la partie de $\mathbb{R}^m \times \mathbb{R}^n$ de la forme*

$$\operatorname{Graph} f = \{(x, y) \in \mathbb{R}^m \times \mathbb{R}^n \,|\, y = f(x)\}. \tag{12.297}$$

22. Appelez la «centrifuge» si vous voulez ; ça ne me fait ni chaud ni froid.

Cette définition se spécialise de la façon suivante dans les cas communs. Soit f une fonction de $\mathbb{R}^m$ dans $\mathbb{R}$. Le graphe de f est la partie de $\mathbb{R}^m \times \mathbb{R}$ donné par

$$\text{Graph } f = \{(x, y) \in \mathbb{R}^m \times \mathbb{R} \,|\, y = f(x)\}. \tag{12.298}$$

Et pour les fonctions $\mathbb{R}^2 \to \mathbb{R}$:

$$\text{Graph } f = \{(x, y, z) \in \mathbb{R}^2 \text{ tel que } z = f(x, y)\}. \tag{12.299}$$

C'est cette définition qu'il faut garder à l'esprit lorsqu'on travaille sur des dessins en trois dimensions.

Si f est une fonction de deux variables indépendantes x et y à valeurs dans $\mathbb{R}$, alors un point dans le graphe de f est un point $(x, y, z) \in \mathbb{R}^3$ tel que

$$z = f(x, y), \tag{12.300}$$

ou encore, un point de la forme

$$\big(x, y, f(x, y)\big). \tag{12.301}$$

Nous avons parfois besoin de donner des représentation graphiques d'une fonction. Nous pouvons, par exemple, penser à la fonction que associe à un point de la Terre son altitude. Lorsqu'on part pour une promenade en montagne on a envie de connaitre le graphe de cette fonction qui correspond en fait à la surface de la montagne. Bien sur nous ne voulons pas amener avec nous un modèle en 3D de la montagne donc il nous faut une méthode efficace pour projeter le graphe de f sur le plan x-y tout en gardant les informations fondamentales. Pour cela nous avons besoin de deux définitions (à ne pas confondre !)

Définition 12.140.
*Soit f une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ et soit c dans $\mathbb{R}$. La **z-section** de* Graph f *à la hauteur c est donné par*

$$S_c^z = \{(x, y, c) \in \mathbb{R}^3 \,|\, f(x, y) = c\}.$$

Définition 12.141.
*Soit f une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ et soit c dans $\mathbb{R}$. La **courbe de niveau** de f à la hauteur c est l'ensemble*

$$N_c = \{(x1, \ldots, x_n) \in \mathbb{R}^n \,|\, f(x1, \ldots, x_n) = c\}. \tag{12.302}$$

On peut représenter la fonction f d'une façon très précise en traçant quelques unes de ses courbes de niveau. Dans la suite on pourra considérer aussi les x-sections et les y-sections du graphe d'une fonction de deux variables. La x-section de Graph f à la hauteur a est

$$S_a^x = \{(a, y, z) \in \mathbb{R}^3 \,|\, f(a, y) = z\}.$$

Comme vous avez peut être déjà compris, S_a^x est le graphe de la fonction de y qu'on obtient de f en fixant $x = a$. Cette fonction est appelée x-section de f pour $x = a$.

Certaines surfaces dans $\mathbb{R}^3$ sont le graphe d'une fonction.

Exemple 12.142
Quelques graphes importants.

Un plan non vertical Tout plan dans $\mathbb{R}^3$ peut être décrit par une équation de la forme

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = r,$$

où, (x_0, y_0, z_0) est vecteur dans $\mathbb{R}^3$, et a, b, c et r sont des nombres réels. Si $c \neq 0$ alors le plan n'est pas vertical et on peut dire que il est le graphe de la fonction

$$P(x, y) = \frac{r + cz_0 - a(x - x_0) - b(y - y_0)}{c},$$

quitte à choisir des nouvelles constantes s, t, q,

$$P(x, y) = sx + ty + q.$$

Un paraboloïde elliptique Pour tous α et β dans $\mathbb{R}$ les graphes des fonctions

$$PE_1(x,y) = \frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2}$$

ou de la fonction

$$PE_2(x,y) = -\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2}$$

sont des paraboloïdes elliptiques. Le premier est contenu dans le demi-espace $z \geqslant 0$, l'autre dans $z \leqslant 0$. Le nom de cette surface vient de la forme de ses sections. En fait toutes sections S_c^z sont des ellipses, alors que les section S_a^x et S_b^y sont des paraboles.

Un paraboloïde hyperbolique (selle) Pour tous α et β dans $\mathbb{R}$ les graphes des fonctions

$$PH_1(x,y) = \frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2}$$

ou de la fonction

$$PH_2(x,y) = -\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2}$$

sont des paraboloïdes hyperboliques. Remarquez que les sections S_c^z de ce graphe sont des hyperboles, alors que les section S_a^x et S_b^y sont des paraboles.

Une demi-sphère La fonction $S^+ = \sqrt{R^2 - x^2 - y^2}$ a pour graphe la demi-sphère supérieure centrée en l'origine et de rayon R. Le dernier de ces exemples nous signale une chose très importante : une sphère entière n'est pas le graphe d'une fonction de x et y. Par contre, une demi-sphère est bien le graphe de la fonction $f(x,y) = \sqrt{1 - x^2 - y^2}$.

L'équation que nous utilisons pour d'écrire une sphère de rayon R centrée en l'origine est

$$x^2 + y^2 + z^2 = R^2$$

Donc, à chaque point (x,y) dans le disque $x^2 + y^2 \leqslant R^2$ (notez que ce disque est contenu dans la section S_0^z), on peut associer deux valeurs de z : $z_1 = \sqrt{R^2 - x^2 - y^2}$ et $z_2 = -\sqrt{R^2 - x^2 - y^2}$. Par définition, une fonction n'associe qu'un seul valeur à chaque point de son domaine, d'où l'impossibilité de décrire cette sphère comme le graphe d'une fonction de x et y.

△

Considérons la fonction $Sp : \mathbb{R}^3 \to \mathbb{R}$ qui associe à (x,y,z) la valeur $x^2 + y^2 + z^2$. La sphère de rayon R centrée en l'origine est l'ensemble de niveau N_{R^2} de Sp. L'ensemble de niveau N_0 de Sp est l'origine, et tous les ensemble de niveau de hauteur négative sont vides. La même chose est vraie pour les ellipsoïdes centrées en l'origine avec les axes x, y et z comme axes principaux et comme longueurs de demi-axes a, b et c. Voici la fonction dont il sont les ensemble de niveau

$$El(x,y,z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}.$$

Exemple 12.143

Des ensembles de niveau importants.

Tout graphe Le graphe de toute fonction f de $\mathbb{R}^2$ dans $\mathbb{R}$ peut être considéré comme l'ensemble de niveau zéro de la fonction $F(x,y,z) = z - f(x,y)$.

Hyperboloïdes Les hyperboloïdes, comme les ellipsoïdes, sont une famille d'ensemble de niveau. En particulier, nous considérons des hyperboloïdes dont l'axe de symétrie est l'axe des z et qui sont symétriques par rapport un plan x-y. Une fois que les paramètres a, b et c sont fixés la fonction que nous intéresse est

$$Hyp(x,y,z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2}.$$

Les ensembles de niveau N_d pour $d > 0$ sont connexes, on les appelle *hyperboloïdes à une feuille*. L'ensemble de niveau N_0 est *cône (elliptique)*, le deux moitiés du cône se touchent en l'origine. Enfin, les ensembles de niveau N_d pour $d < 0$ ne sont pas connexes et pour cette raison on les appelle *hyperboloïdes à deux feuilles*.

△

12.15.1 Graphes de fonctions à plusieurs variables

Le **graphe** d'une fonction de deux variables $f\colon D \subset \mathbb{R}^2 \to \mathbb{R}$ est l'ensemble

$$\Big\{\big(x, y, f(x,y)\big) \text{ tel que } (x,y) \in D\Big\} \subset \mathbb{R}^3. \tag{12.303}$$

Ce graphe est une surface dans $\mathbb{R}^3$.

Exemple 12.144

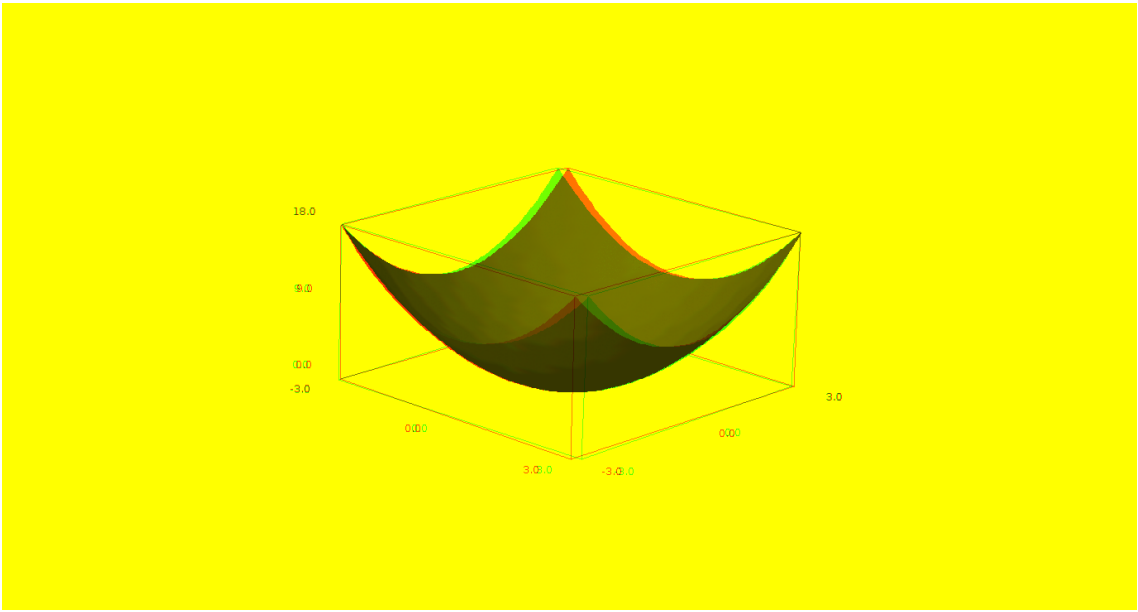
Tracer le graphe de la fonction

$$(x,y) \mapsto x^2 + y^2. \tag{12.304}$$

Le plus simple est de demander à Sage de nous fournir une représentation 3D

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y)=x**2+y**2
sage: plot3d(f,(x,-3,3),(y,-3,3))
```

Voici ce que cela donne [23] : (à regarder avec des lunettes bleues et rouges) :



À part que l'ordinateur l'a dit, est-ce qu'on peut comprendre pourquoi le graphe de la fonction $x^2 + y^2$ ressemble à un bol ? En coordonnées cylindriques, le graphe s'écrit

$$z = r^2. \tag{12.305}$$

23. En vrai, ce que Sage donne est un objet qu'on peut même faire bouger.

Donc il se fait que plus on s'éloigne du point $(0,0)$ dans le plan XY, plus le graphe va monter. Et il monte à quelle vitesse ? Il monte à la vitesse r^2. Il s'agit donc de dessiner la fonction $z = r^2$ dans le plan et de la «faire tourner».

△

12.15.2 Courbes de niveau

Une technique utile pour se faire une idée de la forme d'une fonction en trois dimensions est de tracer les **courbes de niveau**. La courbe de niveau de hauteur h est la courbe dans le plan donnée par l'équation

$$f(x,y) = h. \tag{12.306}$$

Exemple 12.145

Dessinons par exemple les courbes de niveau de la fonction

$$f(x,y) = x + y + 2. \tag{12.307}$$

La courbe de niveau h est donnée par l'équation $x + y + 2 = h$, c'est à dire

$$y(x) = -x + h - 2. \tag{12.308}$$

Par conséquent la courbe de niveau de hauteur 0 est $y = -x - 2$, celle de hauteur 5 est $y = -x + 3$, etc.

Nous pouvons également nous aider de Sage pour ce faire :

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y)=x+y+2
sage: var('h')
h
sage: niveau(h,x)=solve(f(x,y)==h,y)[0].rhs()
sage: g1(x)=niveau(1,x)
sage: g1
x |--> -x - 1
```

Ici la fonction `g1` est la courbe de niveau 1.

Si on veut faire tracer une courbe de niveau, Sage peut le faire :

```
sage: implicit_plot(f(x,y)==1,(x,-3,3),(y,-4,4))
```

Cela tracera la courbe de niveau $h = 1$ dans la partie du plan $x \in [-3,3]$ et $y \in [-4,4,]$.

△

Il est bien entendu possible de créer automatiquement 50 courbes de niveau et de demander de les tracer toutes sur le même graphe.

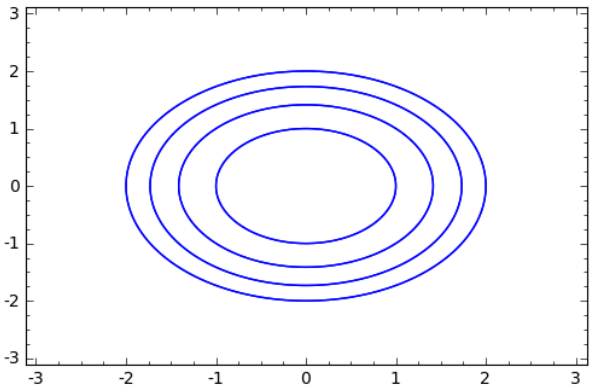
```
1 #! /usr/bin/sage -python
2 # -*- coding: utf8 -*-
3
```

```
from sage.all import *

var('x,y')
f=x**2+y**2
G=Graphics()
a=3
for i in range(0,5):
  G=G+implicit_plot(f==i,(x,-a,a),(y,-a,a))
show(G)
```

tex/frido/courbeNiveau.py

Le résultat est :



Notez que les courbes sont censées être des cercles : les axes X et Y n'ont pas la même échelle.

Exemple 12.146
Un exemple plus riche en enseignements est celui de la fonction

$$f(x,y)=x^2-y^2. \tag{12.309}$$

La courbe de niveau h est donnée par l'équation $x^2-y^2=h$.

Commençons par $h=0$. Dans ce cas nous avons $(x+y)(x-y)=0$ et par conséquent les courbes de niveau de hauteur zéro sont les deux droites $x+y=0$ et $x-y=0$.

Voyons ensuite la courbe de niveau $h=1$. Cela est l'équation $x^2-y^2=1$, c'est à dire

$$y(x)=\pm\sqrt{x^2-1}. \tag{12.310}$$

C'est une fonction qui n'est définie que pour $|x|\geqslant 1$. Avec $x=1$ nous avons $y=1$. Ensuite, lorsque x grandit, y grandit également, mais la courbe ne peut pas croiser la courbe de niveau $h=0$. Donc, suivant les notations de la figure 12.8, la courbe de niveau «part» de P et doit monter sans croiser les diagonales.

En ce qui concerne la courbe de niveau $h=-1$, elle correspond à la courbe $y=\pm\sqrt{1+x^2}$ qui est définie pour tous les $x\in\mathbb{R}$. Le même raisonnement que précédemment nous amène à la figure 12.9.

△

Une autre façon de voir les courbe de niveau est de dire que la courbe de niveau de hauteur h est la projection dans le plan XY de la section du graphe de f par le plan $z=h$.

On peut également définir le graphe de fonctions de trois (ou plus) variables. Le graphe de la fonction $f\colon D\subset\mathbb{R}^3\to\mathbb{R}$ est l'ensemble

$$\big\{\big(x,y,z,f(x,y,z)\big)\text{ tel que }(x,y,z)\in D\big\}\subset\mathbb{R}^4. \tag{12.311}$$

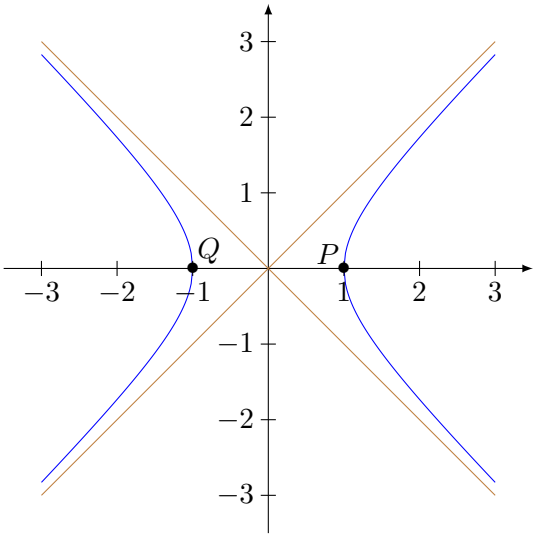


FIGURE 12.8 – La courbe de niveau $h = 1$ de $x^2 - y^2$. Notez qu'elle est en deux morceaux.

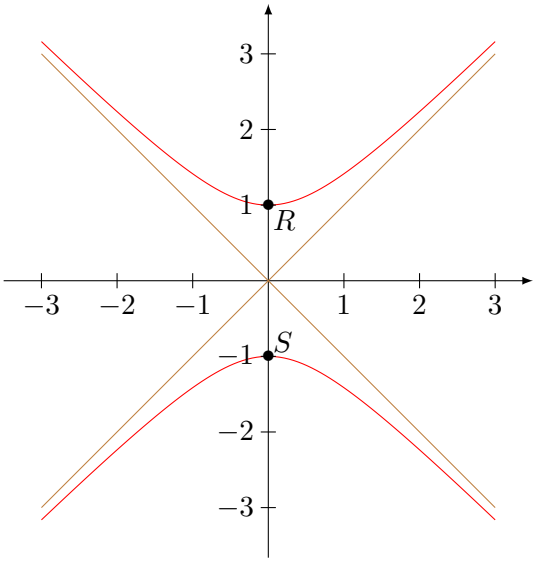


FIGURE 12.9 – La courbe de niveau $x^2 - y^2 = -1$.

De tels graphes ne peuvent pas être représentés sur une feuille de papier. Il est toutefois possible de définir les ensembles de niveaux :

$$E_h = \big\{(x, y, z) \in D \text{ tel que } f(x, y, z) = h\big\}. \tag{12.312}$$

Ce sont des surfaces dans $\mathbb{R}^3$ que l'on peut dessiner.

Exemple 12.147
Les surfaces de niveau de la fonction $f(x, y, z) = x^2 + y^2 + z^2$ sont des sphères. Il n'y a pas de surfaces de niveau pour les «hauteurs» négatives. △

Exemple 12.148
Considérons la fonction $f(x, y, z) = x^2 + y^2 - z^2$. En coordonnées cylindrique, cette fonction s'écrit

$$f(r, \theta, z) = r^2 - z^2. \tag{12.313}$$

La surface de niveau 0 est donnée par l'équation $r = |z|$. Cela fait un cercle à chaque hauteur, dont le rayon grandit linéairement avec la hauteur ; le tout est donc un cône. C'est d'ailleurs le

cône obtenu par rotation de la courbe de niveau $h = 0$ que nous avions obtenue pour la fonction $x^2 - y^2$.

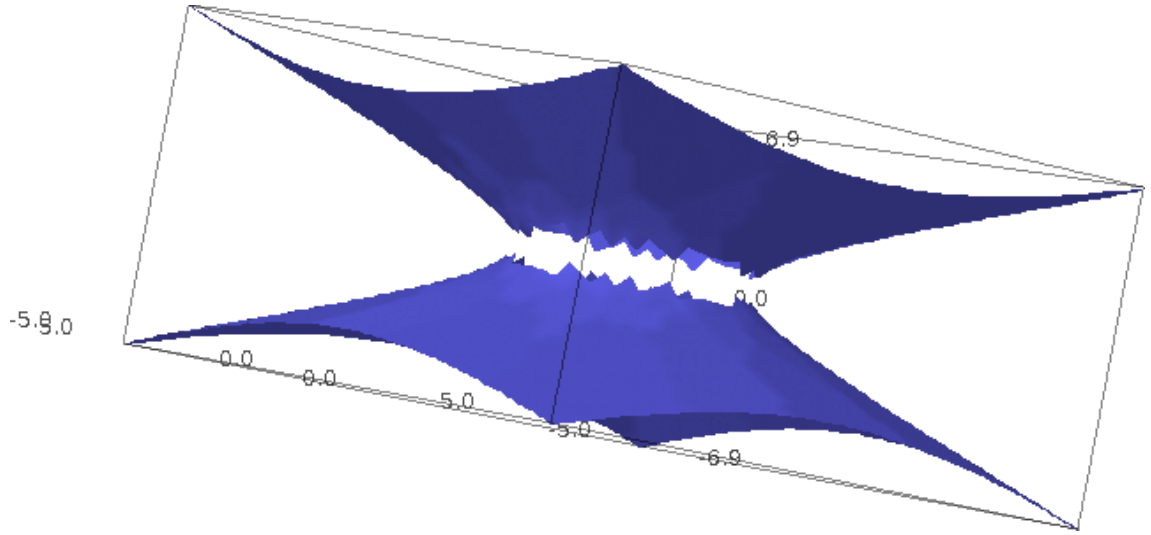
En ce qui concerne les ensembles de niveau positifs, ils sont donnés par

$$z = \pm\sqrt{x^2 + y^2 - h}. \tag{12.314}$$

Notez qu'ils ne sont pas définis pour $r \geqslant h$. Cela pose un petit problème quand on veut le tracer à l'ordinateur :

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: var('x,y')
(x, y)
sage: f(x,y)=sqrt(x**2+y**2-3)
sage: F=plot3d(f(x,y),(x,-5,5),(y,-5,5))
sage: G=plot3d(-f(x,y),(x,-5,5),(y,-5,5))
sage: F+G
```

Le résultat est [24] :

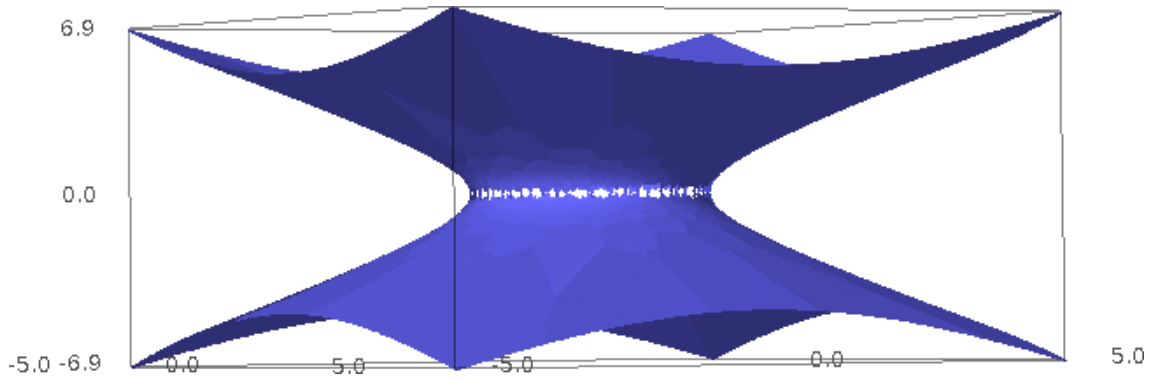


On voit qu'il y a un grand trou au centre correspondant aux z proches de zéro. Or d'après l'équation, il n'en est rien : en $z = 0$ il y a bel et bien tout un cercle. Afin d'obtenir une meilleur image, il faut demander de tracer avec un maillage plus fin :

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: var('x,y')
(x, y)
sage: f(x,y)=sqrt(x**2+y**2-3)
sage: F=plot3d(f(x,y),(x,-5,5),(y,-5,5),plot_points=300)
sage: G=plot3d(-f(x,y),(x,-5,5),(y,-5,5),plot_points=300)
sage: F+G
```

Le temps de calcul est un peu plus long, mais le résultat est meilleur :

24. Encore une fois : ça donne mieux à l'écran, et vous pouvez le faire bouger ; je vous encourage à le faire !



△

12.16 Dérivée directionnelle

Nous sommes capables de dériver une fonction de deux variables $f(x,y)$ par rapport à x et par rapport à y. C'est à dire que nous sommes capables de donner la variation de la fonction lorsqu'on bouge le long des axes horizontal et vertical. Il est évidemment souhaitable de parler de la variation de la fonction lorsqu'on se déplace le long d'autre droites.

Soit donc $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ un vecteur unitaire (c'est à dire $u_1^2 + u_2^2 = 1$), et considérons la fonction de une variable

$$\begin{aligned} \varphi\colon \mathbb{R} &\to \mathbb{R} \\ t &\mapsto f(a + tu_1, b + tu_2). \end{aligned} \tag{12.315}$$

La fonction φ n'est rien d'autre que la fonction f vue le long de la droite de direction donnée par le vecteur u. Nous pouvons aussi l'écrire $\varphi(t) = f(p + tu)$.

Soit $f\colon \mathbb{R}^2 \to \mathbb{R}$ une fonction de deux variables et soit $(a,b) \in \mathbb{R}^2$. La façon la plus naturelle de définir une dérivée à deux variables est de considérer les **dérivées partielles** définies par

$$\begin{aligned} \frac{\partial f}{\partial x}(a,b) &= \lim_{x\to a} \frac{f(x,b) - f(a,b)}{x - a} \\ \frac{\partial f}{\partial y}(a,b) &= \lim_{y\to b} \frac{f(a,y) - f(a,b)}{y - b}. \end{aligned} \tag{12.316}$$

Ces nombres représentent la façon dont le nombre $f(x,y)$ varie lorsque soit seul x varie soit seul y varie. Les dérivées partielles se calculent de la même façon que les dérivées normales. Pour calculer $\partial_x f$, on fait «comme si» y était une constante, et pour calculer $\partial_y f$, on fait comme si x était une constante.

12.16.1 Dérivée partielle et directionnelles

Soit une fonction $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$. Si $n \neq 1$, la notion de *dérivée* de la fonction f n'a plus de sens puisqu'on ne peut plus parler de pente de *la* tangente au graphe de f en un point. On introduit alors quelques notions qui feront, en dimension quelconque, le même travail que la dérivée en dimension un : les dérivées directionnelles et la différentielle. Nous allons voir qu'en dimension un, la différentielle coïncide avec la dérivée.

Définition 12.149.
*Soit f une application de $U \subset \mathbb{R}^m$ dans $\mathbb{R}$, a un point dans U et v un vecteur de $\mathbb{R}^m$. On dit que f admet une **dérivée suivant le vecteur v au point** a si la fonction $t \mapsto f(a + tv)$ admet une dérivée en $t = 0$. La dérivée de f suivant le vecteur v au point a est alors cette dérivée, et f est dite dérivable suivant v en a,*

$$\partial_v f(a) = \lim_{\substack{t\to 0 \\ t\neq 0}} \frac{f(a + tv) - f(a)}{t}.$$

Définition 12.150.
*La fonction $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ de composantes $(f_1, \dots, f_n)$, est dite **dérivable suivant** v **au point** a si toute ses composante f_i, $i = 1, \dots, n$ sont dérivables suivant v au point a. Dans ce cas, nous écrivons*

$$\partial_v f(a) = (\partial_v f_1(a), \dots, \partial_v f_n(a))^T . \tag{12.317}$$

On parle aussi souvent de dérivé **dans la direction** du vecteur v. Une **direction** dans $\mathbb{R}^m$ est un vecteur de norme 1. Tant que u est un élément non nul de $\mathbb{R}^m$, nous pouvons parler de la direction de u.

Géométriquement, il s'agit du taux de variation instantané de f en a dans la direction du vecteur u, c'est-à-dire de la pente de la tangente dans la direction du vecteur u au graphe de f au point $(a, f(a))$.

Remarque 12.151.
On peut reformuler la définition en écrivant $x = a + u$, on obtient :

$$\lim_{\substack{u \to 0 \\ u \neq 0}} \frac{f(a+u) - f(a) - T(u)}{\|u\|} = 0. \tag{12.318}$$

Remarque 12.152.
Pourquoi avons-nous posé la condition $\|u\| = 1$? Le but de la dérivée directionnelle dans la direction u est de savoir à quelle vitesse la fonction monte lorsque l'on se déplace en suivant la direction u. Cette information n'aura un caractère « objectif » que si l'on avance à une vitesse donnée. En effet, si on se déplace deux fois plus vite, la fonction montera deux fois plus vite. Par convention, nous demandons donc d'avancer à vitesse 1.

Cas particulier où $n = 2$:

$a = (a_1, a_2)$, $u = (u_1, u_2)$ et

$$\frac{\partial f}{\partial u}(a_1, a_2) = \lim_{t \to 0} \frac{f(a_1 + tu_1, a_2 + tu_2) - f(a_1, a_2)}{t}$$

Un cas particulier des dérivées directionnelles est la dérivée partielle. Si nous considérons la base canonique e_i de $\mathbb{R}^n$, nous notons

$$\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial e_i}. \tag{12.319}$$

Dans le cas d'une fonction à deux variables, nous avons donc les deux dérivées partielles

$$\frac{\partial f}{\partial x}(a) \quad \text{et} \quad \frac{\partial f}{\partial y}(a) \tag{12.320}$$

qui correspondent aux dérivées directionnelles dans les directions des axes. Ces deux nombres représentent de combien la fonction f monte lorsqu'on part de a en se déplaçant dans le sens des axes X et Y.

12.16.1.1 Quelques propriétés et notations

(1) $\forall \alpha \in \mathbb{R}$, si $v = \alpha\, u$, alors $\frac{\partial f}{\partial v}(a) = \alpha\, \frac{\partial f}{\partial u}(a)$.

(2) Si on prend $u = e_j$ le $j$ème vecteur de la base canonique de $\mathbb{R}^n$, alors

$$\frac{\partial f}{\partial e_j}(a) = \frac{\partial f}{\partial x_j}(a)$$

c'est-à-dire que la dérivée de f au point a dans la direction e_j est la dérivée partielle de f par rapport à sa $j$ème variable.

(3) Une fonction peut être dérivable dans certaines directions mais pas dans d'autres (rappelez vous que si la limite à droite est différente de la limite à gauche, la limite n'existe pas).

(4) Même si une fonction est dérivable en un point dans toutes les directions, on n'est pas sûr qu'elle soit continue en ce point. La dérivabilité directionnelle n'est donc pas une notion suffisante pour assurer la continuité. C'est pourquoi on introduit le concept de *différentiabilité.*

Proposition 12.153.
Soit u un vecteur de norme 1 dans $\mathbb{R}^m$ et soit $v = \lambda u$, avec λ dans $\mathbb{R}$. La fonction f est dérivable suivant v au point a si et seulement si f est dérivable suivant u au point a, en outre

$$\partial_v f(a) = \lambda \partial_u f(a).$$

Démonstration.

$$\begin{aligned} \partial_v f(a) &= \lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(a+tv)-f(a)}{t} = \lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(a+t\lambda u)-f(a)}{t} = \\ &= \lambda \lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(a+t\lambda u)-f(a)}{\lambda t} = \lambda \partial_u f(a). \end{aligned} \tag{12.321}$$

□

Définition 12.154.
Soit f une application de $U \subset \mathbb{R}^m$ dans $\mathbb{R}$. On appelle ***dérivées partielles de f au point*** *a les dérivées de f suivant les vecteurs de base $e_1, \ldots, e_m$ au point a, si elles existent.*

Si $m = 2, 3$ on peut utiliser la notation f_x, ∂_x ou ∂_1 pour la dérivée partielle suivant e_1, f_y, ∂_y ou ∂_2 pour la dérivée partielle suivant e_2 et f_z, ∂_z ou ∂_3 pour la dérivée partielle suivant e_3. En général, nous écrivons ∂_i pour noter la la dérivée partielle suivant e_i.

Exemple 12.155
Les dérivées partielles de la fonction $f(x, y) = xy^3 + \sin y$ au point $(0, \pi)$ sont

$$\partial_x f(0,\pi) = \frac{\partial f}{\partial x}(0,\pi) = f_x(0,\pi) = \lim_{\substack{t\to 0\\ t\neq 0}} \frac{(t\pi^3 + \sin\pi) - (\sin\pi)}{t} = \pi^3,$$

$$\partial_y f(0,\pi) = \frac{\partial f}{\partial y}(0,\pi) = f_y(0,\pi) = \lim_{\substack{t\to 0\\ t\neq 0}} \frac{0(\pi+t)^3 + \sin(t+\pi) - 0\,\cdot\,\pi^3}{t} = \cos\pi = -1,$$

△ La fonction d'une

seule variable qu'on obtient à partir de f en fixant les $p-1$ variables $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_p$ et qui associe à x_i la valeur $f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_p)$, est appelée x_i-ème **section** de f en $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_p$. L'$i$-ème dérivée partielle de f au point $a = (x_1, \ldots, x_m)$ est la dérivée de l'i-ème section de f au point x_i. En pratique, pour calculer les dérivées partielles d'une fonction on fait une dérivation par rapport à la variable choisie en considérant les autres variables comme des constantes.

Exemple 12.156
Considérons la fonction $f(x, y) = 2xy^2$. Lorsque nous calculons $\partial_x f(x, y)$, nous faisons comme si y était constant. Nous avons donc $\partial_x f(x, y) = 2y^2$. Par contre lors du calcul de $\partial_y f(x, y)$, nous prenons x comme une constante. La dérivée de y^2 par rapport à y est évidement $2y$, et par conséquent, $\partial_y f(x, y) = 4xy$. △

Exemple 12.157
La fonction $f(x, y) = x^y$ est dérivable au point $(1, 2)$ et on a

$$\partial_x f(1,2) = (yx^{y-1})_{(x,y)=(1,2)} = 2,$$

$$\partial_y f(1,2) = \partial_y \left(e^{y \ln x} \right)_{(x,y)=(1,2)} = \left(\ln x e^{y \ln x} \right)_{(x,y)=(1,2)} = \ln \left(1 - e^{2 \ln(1)} \right) = 0.$$

△

Définition 12.158.
*Soit f une application de $U \subset \mathbb{R}^m$ dans $\mathbb{R}$ et u un vecteur de $\mathbb{R}^m$. La fonction f est **dérivable sur U suivant le vecteur** u, si f est dérivable suivant le vecteur u en tout point de U.*

Pour les fonctions d'une seule variable la dérivabilité en un point a implique la continuité en a. Cela n'est pas vrai pour les fonctions de plusieurs variables : il existe des fonction f qui sont dérivables suivant tout vecteur au point a sans pour autant être continue en a.

Exemple 12.159
Considérons la fonction $f : \mathbb{R}^2 \to \mathbb{R}$

$$f(x,y) = \begin{cases} \frac{x^2 y}{x^4+y^2} & \text{si } (x,y) \neq (0,0), \\ 0 & \text{sinon.} \end{cases} \tag{12.322}$$

Pour voir que f n'est pas continue en $(0,0)$ il suffit de calculer la limite de f restreinte à la parabole $y = x^2$

$$\lim_{x \to 0} f(x, x^2) = \frac{1}{2} \neq 0.$$

Pourtant la fonction f est dérivable en $(0,0)$ dans toutes les directions. En effet, soit $v = (v_1, v_2)$. Si $v_2 \neq 0$, alors

$$\partial_v f(a) = \lim_{\substack{t \to 0 \\ t \neq 0}} \frac{t^3 v_1^2 v^2}{t^5 v_1^4 + t^3 v_2^2} = \frac{v_1^2}{v_2},$$

tandis que si $v_2 = 0$, alors la valeur de $f(tv_1, 0)$ est 0 pour tout t et v_1, donc la dérivée partielle de f par rapport à x en l'origine existe et est nulle. △

Exemple 12.160
Pour une fonction réelle à variable réelle, la dérivabilité entraine la continuité. Il n'en va pas de même pour les fonctions à plusieurs variables, comme le montre l'exemple suivant :

$$f(x,y) = \begin{cases} 0 & \text{si } x = 0 \\ \frac{y}{x}\sqrt{x^2+y^2} & \text{sinon.} \end{cases} \tag{12.323}$$

Nous avons tout de suite

$$\frac{\partial f}{\partial y}(0,0) = 0. \tag{12.324}$$

De plus si $u_x \neq 0$ nous avons

$$\frac{\partial f}{\partial u}(0,0) = \frac{u_y}{u_x}\|u\|. \tag{12.325}$$

Donc toutes les dérivées directionnelles de f en $(0,0)$ existent alors que la fonction n'y est manifestement pas continue. En effet sous forme polaire,

$$f(r,\theta) = \frac{r \sin(\theta)}{\cos(\theta)}, \tag{12.326}$$

et quelle que soit la valeur de r, en prenant θ suffisamment proche de $\pi/2$, la fraction peut être arbitrairement grande.

Nous verrons par la proposition 12.176 que la différentiabilité d'une fonction implique sa continuité. △

Théorème 12.161 (Accroissement finis pour les dérivées suivant un vecteur)**.**
Soit U un ouvert dans $\mathbb{R}^m$ et soit $f : U \to \mathbb{R}^n$ une fonction. Soient a et b deux points distincts dans U, tels que le segment[25] *$[a, b]$ soit contenu dans U. Soit u le vecteur*

$$u = \frac{b-a}{\|b-a\|_m}.$$

Si $\partial_u f(x)$ existe pour tout x dans $[a, b]$ on a

$$\|f(b) - f(a)\|_n \leqslant \sup_{x\in[a,b]} \|\partial_u f(x)\|_n \|b-a\|_m.$$

Démonstration. Nous considérons la fonction $g(t) = f\big((1-t)a - tb\big)$. Elle décrit la droite entre a et b parce que $g(0) = a$ et $g(1) = b$. En ce qui concerne la dérivée,

$$\begin{aligned} g'(t) &= \lim_{h\to 0} \frac{g(t+h) - g(t)}{h} \\ &= \lim_{h\to 0} \frac{f\big((1-t-h)a - (t+h)b\big)}{h} \\ &= \lim_{h\to 0} \frac{f\big(a + (t+h)(b-a)\big) - f\big(a + t(b-a)\big)}{h} \\ &= \frac{\partial f}{\partial u}\big(a + t(b-a)\big)\|b-a\|. \end{aligned} \tag{12.327}$$

Le dernier facteur $\|b-a\|$ apparaît pour la normalisation du vecteur u. En effet dans la limite, il apparaît $h(b-a)$, ce qui donnerait la dérivée le long de $b-a$, tandis que u vaut $(b-a)/\|b-a\|$.

Par le théorème des accroissements finis pour g, il existe $t_0 \in]0, 1[$ tel que

$$g(1) = g(0) + g'(t_0)(1-0). \tag{12.328}$$

Donc

$$\|g(1) - g(0)\| \leqslant \sup_{t_0} \|g'(t_0)\| = \sum_{t_0\in]0,1[} \left\| \frac{\partial f}{\partial u}(a + t_0(b-a)) \right\| \|b-a\|. \tag{12.329}$$

Mais lorsque t_0 parcours $]0, 1[$, le point $a + t_0(b-a)$ parcours le segment $]a, b[$, d'où le résultat. □

Corollaire 12.162.
Dans les mêmes hypothèses, si $n = 1$, alors il existe $\bar{x}$ dans $]a, b[$ tel que

$$f(b) - f(a) = \partial_u f(\bar{x})\|b-a\|_m.$$

Définition 12.163.
Le nombre

$$\lim_{t\to 0} \frac{f\big(a + tu_1, b + tu_2\big) - f(a, b)}{t} \tag{12.330}$$

*est la **dérivée directionnelle** de f dans la direction de u au point (a, b). Il sera noté*

$$\frac{\partial f}{\partial u}(a, b), \tag{12.331}$$

ou plus simplement $\partial_u f(a, b)$.

Lorsque f est différentiable, la dérivée directionnelle est donnée par

$$\frac{\partial f}{\partial u}(p) = \nabla f(p) \cdot u. \tag{12.332}$$

25. Définition 6.118.

12.16.2 Gradient : direction de plus grande pente

Étant donné que u est de norme 1, l'inégalité de Cauchy-Schwartz donne

$$\left|\nabla f(a,b) \cdot \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\right| \leqslant \|\nabla f(a,b)\|. \tag{12.333}$$

Donc

$$-\|\nabla f(p)\| \leqslant \nabla f(p) \cdot u \leqslant \|\nabla f(p)\|. \tag{12.334}$$

La norme de la dérivée directionnelle (qui est la valeur absolue du nombre au centre) est donc «coincée» entre $-\|\nabla f(p)\|$ et $\|\nabla f(p)\|$. Prenons par exemple

$$u = \frac{\nabla f(p)}{\|\nabla f(p)\|}. \tag{12.335}$$

Dans ce cas, nous avons exactement

$$\nabla f(p) \cdot u = \|\nabla f(p)\|, \tag{12.336}$$

qui est la valeur maximale que la dérivée directionnelle peut prendre.

La direction du gradient est donc la direction suivant laquelle la dérivée directionnelle est la plus grande. Pour la même raison, la dérivée directionnelle est la plus petite dans le sens opposé au gradient.

En termes bien clairs : lorsqu'on veut aller le plus vite possible au ski, on prend la direction du gradient de la piste de ski. C'est dans cette direction que ça descend le plus vite. Dans quelle direction vont les débutants ? Ils vont perpendiculairement à la pente (ce qui ennuie tout le monde, mais c'est un autre problème). Les débutants vont donc dans la direction perpendiculaire au gradient. Prenons donc $u \perp \nabla f(p)$ et calculons la dérivée directionnelle de f dans la direction u en utilisant la formule 12.332 :

$$\frac{\partial f}{\partial u}(p) = \nabla f(p) \cdot u = 0 \tag{12.337}$$

parce que nous avons choisi $u \perp \nabla f(p)$. Nous voyons donc que les débutants en ski ont eu la bonne intuition que la direction dans laquelle la piste ne descend pas, c'est la direction perpendiculaire au gradient.

C'est aussi pour cela que l'on a tendance à faire du zig-zag à vélo lorsqu'on monte une pente très forte et qu'on est fatigué. C'est toujours pour cela que les routes de montagne font de longs lacets. La montée est moins rude en suivant une direction proche d'être perpendiculaire au gradient !

Théorème 12.164.
Le gradient des fonction suit à peu près les mêmes règles que les dérivées. Soient f et g deux fonctions différentiables. Nous avons entre autres

(1) $\nabla(f+g) = \nabla f + \nabla g$;

(2) $\nabla(fg)(a,b) = g(a,b)\nabla f(a,b) + f(a,b)\nabla g(a,b)$;

(3) Dès que $g(a,b) \neq 0$, *nous avons*

$$\nabla\frac{f}{g} = \frac{g(a,b)\nabla f(a,b) - f(a,b)\nabla g(a,b)}{g(a,b)^2}. \tag{12.338}$$

12.16.3 Gradient : orthogonal au plan tangent

Vu que le gradient d'une fonction est la direction de plus grande pente et que le plan tangent est le plan de plus petite pente, quoi de plus naturel que de penser que le gradient est orthogonal au plan tangent ?

Lemme 12.165.
Soit $\phi\colon \mathbb{R}^n \to \mathbb{R}$ une fonction de classe C^1 et la partie

$$\Gamma = \{x \in \mathbb{R}^n \text{ tel que } \phi(x) = C\} \tag{12.339}$$

pour une certaine constante C.

Soit $x_0 \in \Gamma$. Le gradient de ϕ en x_0 est orthogonal au plan tangent à Γ en x_0.

Démonstration. Un vecteur tangent à Γ en x_0 est de la forme $\gamma'(0)$ où $\gamma\colon \mathbb{R} \to \Gamma$ vérifie $\gamma(0) = x_0$. Vu que ϕ est constante sur Γ nous avons

$$\frac{d}{ds}\Big[\phi\big(\gamma(s)\big)\Big]_{s=0} = 0, \tag{12.340}$$

ce qui donne

$$\sum_i \frac{\partial \phi}{\partial x_i}\big(\gamma(0)\big)\gamma_i'(0) = 0, \tag{12.341}$$

ce qui signifie exactement $\langle (\nabla\phi)(x_0), \gamma'(0)\rangle = 0$. Le vecteur $(\nabla\phi)(x_0)$ est donc perpendiculaire à tout vecteur tangent de Γ en x_0. □

12.17 Dérivée directionnelle de fonctions composées

Nous savons déjà comment dériver les fonctions composées de $\mathbb{R}$ dans $\mathbb{R}$. Si nous avons deux fonctions $f\colon \mathbb{R} \to \mathbb{R}$ et $u\colon \mathbb{R} \to \mathbb{R}$, nous formons la composée $\varphi = f \circ u\colon \mathbb{R} \to \mathbb{R}$ dont la dérivée vaut

$$\varphi'(a) = f'\big(u(a)\big)u'(a). \tag{12.342}$$

Considérons maintenant le cas un peu plus compliqué des fonctions $f\colon \mathbb{R} \to \mathbb{R}$ et $u\colon \mathbb{R}^2 \to \mathbb{R}$, et de la composée

$$\begin{aligned} &\varphi\colon \mathbb{R}^2 \to \mathbb{R} \\ &\varphi(x, y) = f\big(u(x, y)\big). \end{aligned} \tag{12.343}$$

Afin de calculer la dérivée partielle de φ par rapport à x, nous admettons que pour tout a, b et t, il existe $c \in [a, a + t]$ tel que

$$u(a + t, b) = u(a, b) + t\frac{\partial u}{\partial x}(c, b). \tag{12.344}$$

Cela est une généralisation immédiate du théorème 12.129. Nous devons calculer

$$\frac{\partial \varphi}{\partial x}(a, b) = \lim_{t\to 0} \frac{\varphi(a + t, b) - \varphi(a, b)}{t} = \lim_{t\to 0} \frac{f\big(u(a + t, b)\big) - g\big(u(a, b)\big)}{t}. \tag{12.345}$$

Étant donné l'hypothèse que nous avons faite sur u, nous avons

$$f\big(u(a + t, b)\big) = f\big(u(a, b) + t\frac{\partial u}{\partial x}(c, b)\big). \tag{12.346}$$

En utilisant le théorème des accroissements finis pour f, nous avons un point d entre $u(a, b)$ et $u(a, b) + t\frac{\partial u}{\partial x}(c, b)$ tel que

$$f\big(u(a, b) + t\frac{\partial u}{\partial x}(c, b)\big) = f\big(u(a, b)\big) + t\frac{\partial u}{\partial x}(c, b)f'(d). \tag{12.347}$$

Le numérateur de (12.345) devient donc

$$t\frac{\partial u}{\partial x}(c, b)f'(d). \tag{12.348}$$

Certes les points c et d sont inconnus, mais nous savons que c est entre a et $a+t$ ainsi que d se situe entre $u(a,b)$ et $u(a,b)+t\frac{\partial u}{\partial x}(c,b)$. Lorsque nous prenons la limite $t\to 0$, nous avons donc $\lim_{t\to 0} c=a$ et $\lim_{t\to 0} d=u(a,b)$. Nous avons alors

$$\lim_{t\to 0}\frac{t\frac{\partial u}{\partial x}(c,b)f'(d)}{t}=\frac{\partial u}{\partial x}(a,b)f'\big(u(a,b)\big). \tag{12.349}$$

La formule que nous avons obtenue (de façon pas très rigoureuse) est

$$\frac{\partial}{\partial x}f\big(u(x,y)\big)=\frac{\partial u}{\partial x}(x,y)f'\big(u(x,y)\big). \tag{12.350}$$

Prenons maintenant un cas un peu plus compliqué où nous voudrions savoir les dérivées partielles de la fonction φ donnée par

$$\varphi(x,y,z)=f\big(u(x,y),v(x,y,z)\big) \tag{12.351}$$

où $f\colon\mathbb{R}^2\to\mathbb{R}$, $u\colon\mathbb{R}^2\to\mathbb{R}$ et $v\colon\mathbb{R}^3\to\mathbb{R}$.

Commençons par la dérivée partielles par rapport à z. Étant donné que φ ne dépend de z que via la seconde entrée de f, il est normal que seule la dérivée partielle de f par rapport à sa seconde entrée arrive dans la formule :

$$\frac{\partial\varphi}{\partial z}(x,y,z)=\frac{\partial f}{\partial v}\big(u(x,y),v(x,y,z)\big)\frac{\partial v}{\partial z}(x,y,z). \tag{12.352}$$

La dérivée partielle par rapport à y demande de tenir compte en même temps de la façon dont f varie avec sa première entrée et la façon dont elle varie avec sa seconde entrée ; cela nous fait deux termes :

$$\frac{\partial\varphi}{\partial y}(x,y,z)=\frac{\partial f}{\partial u}\big(u(x,y),v(x,y,z)\big)\frac{\partial u}{\partial y}(x,y)+\frac{\partial f}{\partial v}\big(u(x,y),v(x,y,z)\big)\frac{\partial v}{\partial y}(x,y,z). \tag{12.353}$$

Cette formule a une interprétation simple. Lançons un caillou du sommet d'une falaise. Son mouvement est une chute libre avec une vitesse initiale horizontale :

$$\begin{cases} x(t)=v_0t & \text{(12.354a)}\\ y(t)=h_0-\dfrac{gt^2}{2} & \text{(12.354b)}\end{cases}$$

où v_0 est la vitesse initiale horizontale et h_0 est la hauteur de la falaise. Si nous sommes intéressés à la distance entre le caillou et le bas de la falaise (point $(0,0)$), le théorème de Pythagore nous dit que

$$d(t)=\sqrt{x^2(t),y^2(t)}. \tag{12.355}$$

Pour trouver la variation de la distance par rapport au temps il faut savoir de combien la distance varie lorsque x varie et multiplier par la variation de x par rapport à t, et puis faire la même chose avec y.

Théorème 12.166.
Soit $g\colon\mathbb{R}^m\to\mathbb{R}^n$ une fonction différentiable en a, et $f\colon\mathbb{R}^n\to\mathbb{R}^p$ une fonction différentiable en $g(a)$. Si nous définissons $\varphi(x)=(f\circ g)(x)$, alors pour tout $i=1,\dots,m$, nous avons

$$\frac{\partial\varphi}{\partial x_i}(a)=\sum_{k=1}^{n}\frac{\partial f}{\partial y_k}\big(g(a)\big)\frac{\partial g}{\partial x_i} \tag{12.356}$$

où $\frac{\partial f}{\partial y_k}$ dénote la dérivée partielle de f par rapport à sa k-ième variable.

Donnons un exemple d'utilisation de cette formule. Si

$$\begin{aligned} g&\colon\mathbb{R}^2\to\mathbb{R}^3\\ f&\colon\mathbb{R}^3\to\mathbb{R},\end{aligned} \tag{12.357}$$

nous avons $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$. Les dérivées partielles de φ sont données par les formules

$$\frac{\partial\varphi}{\partial x}(x,y) = \frac{\partial f}{\partial x_1}\big(g(x,y)\big)\frac{\partial g_1}{\partial x}(x,y) + \frac{\partial f}{\partial x_2}\big(g(x,y)\big)\frac{\partial g_2}{\partial y}(x,y) + \frac{\partial f}{\partial x_3}\big(g(x,y)\big)\frac{\partial g_3}{\partial x}(x,y) \tag{12.358}$$

et

$$\frac{\partial\varphi}{\partial y}(x,y) = \frac{\partial f}{\partial x_1}\big(g(x,y)\big)\frac{\partial g_1}{\partial y}(x,y) + \frac{\partial f}{\partial x_2}\big(g(x,y)\big)\frac{\partial g_2}{\partial y}(x,y) + \frac{\partial f}{\partial x_3}\big(g(x,y)\big)\frac{\partial g_3}{\partial y}(x,y) \tag{12.359}$$

Notez que les dérivées de φ et des composantes de g sont calculées en (x,y), tandis que celles de f sont calculées en $g(x,y)$.

12.18 Formes différentielles

Nous parlerons de formes différentielles exactes et fermées dans la section 14.77.

12.18.1 Décomposition dans la base duale

Définition 12.167.
Soit U, un ouvert dans $\mathbb{R}^n$. Une 1-***forme différentielle*** *ω sur U est une application*

$$\begin{aligned}\omega\colon U &\to (\mathbb{R}^n)^*\\ x &\mapsto \omega_x.\end{aligned} \tag{12.360}$$

Remarque 12.168.
L'ensemble des 1-formes différentielles forment un espace vectoriel avec les définitions

$$\begin{aligned}(\lambda\omega)_x(v) &= \lambda\omega_x(v)\\ (\omega+\mu)_x(v) &= \omega_x(v) + \mu_x(v).\end{aligned} \tag{12.361}$$

Nous connaissons la **base duale** de $(\mathbb{R}^n)^*$, ce sont les formes e_i^* définies par $e_i^*(e_j) = \delta_{ij}$. Nous allons noter ces formes par dx_i :

$$\begin{aligned}e_1^* = dx_1\colon v &\mapsto v_1\\ &\vdots\\ e_n^* = dx_n\colon v &\mapsto v_n\end{aligned} \tag{12.362}$$

Toute forme différentielle s'écrit

$$\omega_x = \sum_{i=0}^{n} a_i(x)dx_i \tag{12.363}$$

où $a_1,\ldots,a_n$ sont les composantes de ω dans la base usuelle, et sont des fonctions à valeurs réelles.

Lemme 12.169.
Une 1*-forme différentielle est* ***continue*** *si les fonctions a_i sont continues. La forme sera C^k quand les a_i seront C^k.*

Pour un vecteur $v = (v_1,\ldots,v_n)$ on a donc par définition de dx_i

$$\omega_x(v) = \sum_{i=0}^{n} a_i(x)v_i. \tag{12.364}$$

Ces fonctions a_i peuvent être trouvées en appliquant ω aux éléments de la base canonique de $\mathbb{R}^n$:

$$a_j(x) = \omega_x(e_j) \tag{12.365}$$

parce que $\omega_x(e_j) = \sum_i a_i(x)dx_i(e_i) = \sum_i a_i(x)\delta_{ij} = a_j(x)$.

12.18.2 L'isomorphisme musical

Si G est un champ de vecteur sur $\mathbb{R}^n$, et si $x \in \mathbb{R}^n$, nous pouvons définissons

$$\begin{aligned} G_x^\flat \colon \mathbb{R}^n &\to \mathbb{R} \\ v &\mapsto \langle G(x), v \rangle \end{aligned} \tag{12.366}$$

Pour chaque x, l'application $G_x^\flat$ est une forme sur $\mathbb{R}^n$, c'est à dire une application linéaire de $\mathbb{R}^n$ vers $\mathbb{R}$. Nous écrivons que

$$G_x^\flat \in \big(\mathbb{R}^n\big)^*. \tag{12.367}$$

Nous pouvons ainsi déterminer le développement de $G^\flat$ dans la base des dx_i en faisant le calcul

$$G_x^\flat(e_i) = \langle G(x), e_i \rangle = G_i(x), \tag{12.368}$$

donc les composantes de $G^\flat$ dans la base dx_i sont exactement les composantes de G dans la base e_i :

$$G_x^\flat = G_1(x)dx_1 + \cdots + G_n(x)dx_n. \tag{12.369}$$

La construction inverse existe également. Si ω est une 1-forme différentielle, nous pouvons définir le champ de vecteur $\omega^\sharp$ par la formule (implicite)

$$\omega_x(v) = \langle \omega^\sharp(x), v \rangle \tag{12.370}$$

pour tout $v \in \mathbb{R}^n$. Par définition, $(\omega^\sharp)^\flat = \omega$.

Lemme 12.170.
En composantes nous avons :

$$\omega^\sharp(x) = \big(a_1(x), \dots, a_n(x)\big). \tag{12.371}$$

Si G est un champ de vecteurs, alors $(G^\flat)^\sharp = G$.

Si $F \colon \mathbb{R}^2 \to \mathbb{R}$ est une fonction C^2, sa différentielle est la forme

$$dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy. \tag{12.372}$$

Si nous nommons f et g les fonctions $\partial_x F$ et $\partial_y F$, nous avons donc

$$Df = fdx + gdy, \tag{12.373}$$

qui vérifie

$$\partial_y f = \partial_x g, \tag{12.374}$$

parce que $\frac{\partial f}{\partial y} = \frac{\partial^2 F}{\partial x \partial y} = \frac{\partial^2 F}{\partial y \partial x} = \frac{\partial g}{\partial x}$. Ce que nous avons donc prouvé, c'est que

Lemme 12.171.
Si $fdx + gdy$ est la différentielle d'une fonction de classe C^2, alors $\partial_y f = \partial_x g$.

12.18.3 Différentielles de fonctions composées

Cette façon de voir la différentielle nous permet de jeter un nouveau regard sur la formule de différentiation des fonctions composées. Soient

$$\begin{aligned} f &\colon \mathbb{R}^p \to \mathbb{R}^n \\ g &\colon \mathbb{R}^n \to \mathbb{R}, \end{aligned} \tag{12.375}$$

et $h \colon \mathbb{R}^p \to \mathbb{R}$ définie par

$$h(u) = h\big(f(u)\big) = (g \circ f)(u). \tag{12.376}$$

Nous allons noter x les coordonnées de $\mathbb{R}^p$, a un point de $\mathbb{R}^p$ et u, un vecteur de $\mathbb{R}^p$ accroché au point a. Pour $\mathbb{R}^n$, les notations seront que les coordonnées sont y, b est un point de $\mathbb{R}^n$ et v est un vecteur «accroché» au point b.

Nous avons

$$dg_b(v) = \sum_{i=1}^{n} \frac{\partial g}{\partial y_i}(b) dy_i(v). \tag{12.377}$$

Ici $dy_i(v)$ signifie la $i$ème composante de v. C'est simplement v_i. Cette formule étant valable pour tout point $b \in \mathbb{R}^n$ et pour tout vecteur v, nous pouvons l'écrire en particulier pour

$$\begin{cases} b = f(a) & \text{(12.378a)} \\ v = df_a(u). & \text{(12.378b)} \end{cases}$$

Cela donne

$$dg_{f(a)}\big(df_a(u)\big) = \sum_{i=1}^{n} \frac{\partial g}{\partial y_i}\big(f(a)\big) dy_i\big(df_a(u)\big). \tag{12.379}$$

Mais

$$df_a(u) = \sum_{j=1}^{p} \frac{\partial f}{\partial x_j}(a) dx_j(u), \tag{12.380}$$

donc la $i$ème composante de ce vecteur est

$$\big(df_a(u)\big)_i = \sum_{j=1}^{p} \frac{\partial f_i}{\partial x_j}(a) dx_j(u). \tag{12.381}$$

En remplaçant $dy_i\big(df_a(u)\big)$ par cela dans l'expression (12.379), nous trouvons

$$dg_{f(a)}\big(df_a(u)\big) = \sum_{i=1}^{n} \frac{\partial g}{\partial y_i}\big(f(a)\big) \sum_{j=1}^{p} \frac{\partial f_i}{\partial x_j}(a) dx_j(u). \tag{12.382}$$

Nous pouvons vérifier que cela est la différentielle de $g \circ f$ au point a appliquée au vecteur u. En effet

$$d(g \circ f)_a(u) = \sum_{j=1}^{p} \frac{\partial (g \circ f)}{\partial x_j}(a) dx_j(u), \tag{12.383}$$

tandis que, par la dérivation de fonctions composées,

$$\frac{\partial (g \circ f)}{\partial x_j}(a) = \sum_{i=1}^{n} \frac{\partial g}{\partial y_i}\big(f(a)\big) \frac{\partial f_i}{\partial x_j}(a). \tag{12.384}$$

Au final, ce que nous avons prouvé est que

$$d(g \circ f)_a(u) = dg_{f(a)}\big(df_a(u)\big). \tag{12.385}$$

12.18.4 Exemple de composée : les coordonnées polaires

Le changement de coordonnées pour les polaires est la fonction

$$f\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}. \tag{12.386}$$

Considérons une fonction g sur $\mathbb{R}^2$, et définissons la fonction $\tilde{g}$ par

$$\tilde{g}(r, \theta) = g(r\cos\theta, r\sin\theta). \tag{12.387}$$

La formule (12.384) permet de trouver les dérivées partielles de g par rapport à r et θ en termes de celles par rapport à x et y de g.

Pour faire le lien avec les notations du point précédent, nous avons

$$\begin{aligned} f_1(r,\theta) &= r\cos(\theta)\\ f_2(r,\theta) &= r\sin(\theta)\\ (x_1,x_2) &\to (r,\theta)\\ (y_1,y_2) &\to (x,y). \end{aligned} \tag{12.388}$$

Nous avons donc

$$\begin{aligned} \frac{\partial\tilde{g}}{\partial r}(r,\theta) &= \sum_{i=1}^{2}\frac{\partial g}{\partial x_i}\big(f(r,\theta)\big)\frac{\partial f_i}{\partial r}(r,\theta)\\ &= \frac{\partial g}{\partial x}(r\cos\theta,r\sin\theta)\frac{\partial\big(r\cos\theta\big)}{\partial r}(r,\theta)\\ &\quad+\frac{\partial g}{\partial y}(r\cos\theta,r\sin\theta)\frac{\partial\big(r\sin\theta\big)}{\partial r}(r,\theta)\\ &= \cos\theta\frac{\partial g}{\partial x}(r\cos\theta,r\sin\theta)+\sin\theta\frac{\partial g}{\partial y}(r\cos\theta,r\sin\theta). \end{aligned} \tag{12.389}$$

Prenons par exemple $g(x,y)=\frac{1}{x^2+y^2}$. Étant donné que

$$\frac{\partial g}{\partial x}=\frac{-2x}{(x^2+y^2)^2}, \tag{12.390}$$

nous avons

$$\frac{\partial g}{\partial x}(r\cos\theta,r\sin\theta)=\frac{-2\cos\theta}{r^3}. \tag{12.391}$$

En utilisant la formule,

$$\frac{\partial\tilde{g}}{\partial r}(r,\theta)=\cos(\theta)\left(\frac{-2\cos\theta}{r^3}\right)+\sin(\theta)\left(\frac{-2\sin\theta}{r^3}\right)=-\frac{2}{r^3}. \tag{12.392}$$

Nous pouvons vérifier directement que cela est correct. En effet

$$\tilde{g}(r,\theta)=g(r\cos\theta,r\sin\theta)=\frac{1}{r^2}, \tag{12.393}$$

dont la dérivée par rapport à r vaut $-2/r^3$.

En ce qui concerne la dérivée par rapport à θ, nous avons

$$\begin{aligned} \frac{\partial\tilde{g}}{\partial\theta} &= \frac{\partial g}{\partial x}(r\cos\theta,r\sin\theta)\frac{\partial\big(r\cos(\theta)\big)}{\partial\theta}+\frac{\partial g}{\partial y}(r\cos\theta,r\sin\theta)\frac{\partial\big(r\sin(\theta)\big)}{\partial\theta}\\ &= \left(\frac{-2\cos\theta}{r^3}\right)(-r\sin\theta)+\left(\frac{-2\sin\theta}{r^3}\right)(r\cos\theta)\\ &= 0. \end{aligned} \tag{12.394}$$

En résumé et avec quelques abus de notations :

$$\begin{aligned} \frac{\partial\tilde{g}}{\partial r} &= \cos(\theta)\frac{\partial g}{\partial x}+\sin(\theta)\frac{\partial g}{\partial y}\\ \frac{\partial\tilde{g}}{\partial\theta} &= -r\sin(\theta)\frac{\partial g}{\partial x}+r\cos(\theta)\frac{\partial g}{\partial y} \end{aligned} \tag{12.395}$$

12.19 Différentielle

Note : pour savoir des choses sur la différentielle de $f\colon E\to F$ avec E et F de dimension infinie, il faut aller voir la section 12.20. Ici nous ne parlerons que de dimension finie.

12.19.1 Exemples introductifs

La notion de dérivée est associée à la recherche de la droite tangente à une courbe. Reprenons rapidement le cheminement. La dérivée de $f\colon \mathbb{R}\to\mathbb{R}$ au point a est un nombre $f'(a)$, qui définit donc une application linéaire dont le coefficients angulaire est $f'(a)$, et que nous notons df_a :

$$\begin{aligned} df_a\colon \mathbb{R}&\to\mathbb{R}\\ u&\mapsto f'(a)u. \end{aligned} \tag{12.396}$$

La droite donnée par l'équation

$$y(a+u)=f'(a)u \tag{12.397}$$

est parallèle à la tangente en a. Pour trouver la tangente, il suffit de la décaler de la hauteur qu'il faut. L'équation de la droite tangente au graphe de f au point $\big(a,f(a)\big)$ devient

$$y(x)=f(a)+f'(a)(x-a)=f(a)+df_a(x-a). \tag{12.398}$$

Nous nous proposons de généraliser cette formule au cas de la recherche du plan tangent à une surface.

Exemple 12.172
Considérons $f(x,y)=x^2y+y^2e^x$. Les dérivées partielles sont

$$\begin{aligned} \frac{\partial f}{\partial x}&=2xy+y^2e^x\\ \frac{\partial f}{\partial y}&=x^2+2ye^x. \end{aligned} \tag{12.399}$$

△

Cet exemple était l'exemple facile où tout se passe bien.

Exemple 12.173
Les choses sont moins simples lorsqu'on considère la fonction suivante :

$$f(x,y)=\begin{cases} \frac{xy}{x^2+y^2} & \text{si } (x,y)\neq(0,0)\\ 0 & \text{si } (x,y)=(0,0). \end{cases} \tag{12.400}$$

On voit que pour tout x et tout y, nous avons $f(x,0)=f(0,y)=0$. Donc cette fonction est nulle sur les axes horizontaux et verticaux. Nous avons en particulier

$$\begin{aligned} \frac{\partial f}{\partial x}(0,0)&=0\\ \frac{\partial f}{\partial y}(0,0)&=0. \end{aligned} \tag{12.401}$$

Donc ces dérivées partielles existe.

Il n'est par contre pas question de dire que cette fonction «va bien» autour du point $(0,0)$. En effet si nous regardons sa valeur sur la droite diagonale $y=x$, nous avons

$$f(x,x)=\frac{x^2}{2x^2}=\frac{1}{2}. \tag{12.402}$$

Par conséquent si nous suivons la fonction le long de la droite $y=x$, la hauteur vaut $\frac{1}{2}$ en permanence, sauf juste en $(0,0)$ où la fonction fait un grand plongeon !

```
sage: var('x,y')
(x, y)
sage: f(x,y)=(x*y)/(x**2+y**2)
sage: plot3d(f,(x,-2,2),y(-2,2))
```

D'ailleurs elle fait un plongeon le long de toutes les droites (sauf verticale et horizontale). En effet si nous regardons la fonction le long de la droite $y = mx$, nous avons

$$f(x, mx) = \frac{mx^2}{x^2 + m^2x^2} = \frac{m}{1 + m^2}. \tag{12.403}$$

La fonction est donc *constante* sur chacune de ces droites. Il n'est donc pas question de dire que cette fonction est «dérivable» en $(0, 0)$, vu qu'elle fait des grands sauts dans presque toutes les directions. △

Nous devons donc trouver mieux que les dérivées partielles pour étudier le comportement des fonctions un peu problématiques.

12.19.2 Définition de la différentielle

Nous nous souvenons de l'équation (12.236) qui nous dit que pour une fonction d'une variable la dérivabilité signifiait qu'il existait un nombre ℓ et une fonction α tels que

$$f(x) = f(a) + \ell(x - a) + (x - a)\alpha(x - a) \tag{12.404}$$

et $\lim_{t\to 0} \alpha(t) = 0$.

En nous inspirant de cela, nous posons la définition suivante.

Proposition-définition 12.174 ([1])**.**
Soit U un ouvert dans $\mathbb{R}^m$, $a \in U$ et une application $f : U \to \mathbb{R}^n$. Si il existe une une application linéaire $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ satisfaisant

$$\lim_{\substack{h\to 0\\ h\in\mathbb{R}^m}} \frac{f(a + h) - f(a) - T(h)}{\|h\|_m} = 0, \tag{12.405}$$

alors il en existe une seule.

*Dans ce cas nous disons que f est **différentiable au point** a et l'application T ainsi définie est appelée **différentielle** de f au point a, et on la note df_a.*

Démonstration. Soient deux applications linéaires T_1, T_2 satisfaisant la condition (12.405). Nous avons

$$\frac{\|T_1(h) - T_2(h)\|}{\|h\|} \leqslant \frac{\|T_1(h) - f(a + h) + f(a)\|}{\|h\|} + \frac{\|f(a + h) - f(a) - T_2(h)\|}{\|h\|} \to 0. \tag{12.406}$$

Nous avons donc

$$\lim_{h\to 0} \frac{\|(T_1 - T_2)(h)\|}{\|h\|} \tag{12.407}$$

En particulier pour un vecteur de base e_i nous avons $\|e_i\| = 1$ et la limite [26]

$$0 = \lim_{t\to 0} \frac{\|(T_1 - T_2)(te_i)\|}{\|te_i\|} = \|(T_1 - T_2)(e_i)\|. \tag{12.408}$$

Les applications linéaires T_1 et T_2 sont donc égales sur une base. Elles sont donc égales. □

Note : df_a est *en soi* une application $df_a : \mathbb{R}^m \to \mathbb{R}^n$. Nous notons $df_a(u)$ la valeur de df_a sur le vecteur $u \in \mathbb{R}^m$.

Si f est différentiable en tout point de l'ouvert U, alors nous disons que f est différentiable sur U et l'application

$$\begin{aligned} df : U &\to \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n) \\ x &\mapsto df_x \end{aligned} \tag{12.409}$$

26. Parce qu'une limite dans $\mathbb{R}^m$ donne en particulier la limite suivant tous les chemins par la proposition 12.97.

est la **différentielle** de f sur U. Cet objet est une forme différentielle au sens de la définition 12.167.

Les propositions 12.179 et 12.180 vont montrer qu'en étudiant bien les dérivées partielles, nous pouvons conclure à la différentiabilité d'une fonction. Attention cependant, nous verrons dans l'exemple 12.187 que l'existence des dérivées directionnelles partielles ne permettait pas de conclure à la différentiabilité.

12.19.3 Base duale

Nous avons déjà parlé en (12.362) de la base $\{dx_i\}_{i=1,\dots,n}$ des formes différentielles sur $\mathbb{R}^n$.

Proposition 12.175.
La forme de base dx_i est la différentielle de la fonction de projection

$$\begin{aligned} \texttt{proj}_i\colon \mathbb{R}^n &\to \mathbb{R} \\ v &\mapsto v_i. \end{aligned} \tag{12.410}$$

Autrement dit nous avons

$$d(\texttt{proj}_i)_a = dx_i \tag{12.411}$$

pour tout i et pour tout a.

Démonstration. Le quotient

$$\frac{\texttt{proj}_i(a+h) - \texttt{proj}_i(a) - dx_i(h)}{\|h\|} \tag{12.412}$$

est toujours nul. La limite est a fortiori nulle. □

Nous avons donc $(d\,\texttt{proj}_i)_a = dx_i$ pour tout a. Notons que les fonctions dx_i et $\texttt{proj}_i$ sont les mêmes. Cela justifie la notation «dx_i» pour les formes différentielles de base, parce que ce sont les différentielles des fonctions «coordonnées» que nous pouvons noter x_i.

Étant donné une fonction f, il est légitime de nous demander comment (si elle existe) la différentielle se décompose en chaque point dans la base duale. C'est à dire fixer les fonctions a_i en termes des dérivées de f pour avoir

$$df_a = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) dx_i. \tag{12.413}$$

C'est ce que nous allons faire dans le corollaire 12.177.

12.19.4 Pas la différentielle extérieure

Il existe une notion de différentielle extérieure, mais ce n'est pas celle-là que nous allons utiliser la majorité du temps. En particulier si E et F sont des espaces vectoriels normés, lorsque $f\colon E \to F$ est une fonction, df est une application

$$df\colon E \to \mathcal{L}(E,F) \tag{12.414}$$

et la différentielle seconde est la différentielle de cette application-là. Chose faisable parce que $\mathcal{L}(E,F)$ est un espace vectoriel on ne peut plus respectable.

Soit $D \subset \mathbb{R}^n$. Par définition de la différentielle extérieure d'une 1-forme, nous avons une formule de Leibnitz

$$d(f\omega) = df \wedge \omega + f d\omega. \tag{12.415}$$

En particulier,

$$d(fdx) = df \wedge dx + f \underbrace{d(dx)}_{=0} = \frac{\partial f}{\partial x} dx \underbrace{dx \wedge dx}_{=0} + \frac{\partial f}{\partial y} dy \wedge dx. \tag{12.416}$$

Attention : la différentielle extérieure n'est pas la différentielle usuelle. Certes dans le cas d'une 0-forme (c'est à dire d'une fonction), les deux notions coïncident, mais ça ne va pas plus loin. La différentielle extérieure vérifie $d^2\omega = 0$ pour tout ω, y compris pour les fonctions : si $\omega = df$ alors $d\omega = 0$.

Nous mentionnerons la différentielle extérieure dans le cas de

(1) Théorème de Stockes 14.48.

12.19.5 Continuité, dérivabilité et différentiabilité

Le théorème suivant reprend les principales propriétés d'une fonction différentiable. Il est à ne pas confondre avec la proposition 12.204 qui dira que si les dérivées partielles *sont continues* sur un voisinage de a, alors f est différentiables en a.

Proposition 12.176.
Si f est différentiable au point $a \in \mathbb{R}^n$ alors

(1) elle est continue en a,

(2) elle admet une dérivée dans toutes les directions de $\mathbb{R}^m$,

(3) toute les dérivées directionnelles $\partial_u f(a)$ existent et nous avons l'égalité

$$\begin{aligned} df_a \colon \mathbb{R}^n &\to \mathbb{R}^m \\ u &\mapsto df_a(u) = \frac{\partial f}{\partial u}(a) = \sum_i \frac{\partial f}{\partial x_i}(a)u_i, \end{aligned} \tag{12.417}$$

si les u^i sont les composantes de u dans la base canonique de $\mathbb{R}^n$.

La dernière égalité sera de temps en temps utilisée sous la forme

$$df_a(u) = \frac{d}{dt}\Big[f(a+tu)\Big]_{t=0}. \tag{12.418}$$

Démonstration. La limite

$$\lim_{h\to 0_m} \frac{\|f(a+h) - f(a) - T(h)\|_n}{\|h\|_m} = 0,$$

implique que

$$\lim_{h\to 0_m} \|f(a+h) - f(a) - T(h)\|_n = 0.$$

Comme T est dans $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$, on a $\lim_{h\to 0} T(h) = 0$, d'où la continuité de f au point a.

Si u est un vecteur non nul, la différentiabilité de f au point a implique

$$\lim_{t\to 0} \frac{\|f(a+tu) - f(a) - T(tu)\|_n}{\|tu\|_m} = 0,$$

par la linéarité de T et par l'égalité $\|tu\|_m = |t|\|u\|_m$ on obtient

$$\lim_{t\to 0} \frac{f(a+tu) - f(a)}{|t|} = T(u).$$

Donc f est dérivable suivant le vecteur u et $\partial_u f(a) = T(u) = df_a(u)$. □

Corollaire 12.177.
Si f est différentiable, alors la forme différentielle df_a se décompose en

$$da_a f = \sum_i (\partial_i f)(a) dx_i. \tag{12.419}$$

Démonstration. Vue la définition des formes dx_i nous pouvons remplacer u_i par $dx_i(u)$ dans l'égalité (12.417) et écrire

$$df_a(u) = \sum_i (\partial_i f)(a) dx_i(u) \tag{12.420}$$

et donc écrire l'égalité demandée. □

Le lemme suivant regroupe quelques égalités avec lesquelles nous allons souvent travailler. Il explique comment sont liés les dérivées directionnelles, les dérivées partielles et la différentielle.

Lemme 12.178.
Si $f\colon \mathbb{R}^m \to \mathbb{R}^n$ est une fonction différentiable, alors

$$df_a(u) = \frac{\partial f}{\partial u}(a) = \frac{d}{dt}\Big[f(a+tu)\Big]_{t=0} = \sum_{i=1}^{m} u_i \frac{\partial f}{\partial x_i}(a) = \nabla f(a) \cdot u \tag{12.421}$$

pour tout vecteur $u \in \mathbb{R}^m$

Démonstration. La première égalité est la proposition 12.176, et la seconde est seulement la définition de la dérivée directionnelle avec des notations un peu plus snob. En particulier nous avons

$$df_a(e_i) = \frac{\partial f}{\partial x_i}(a). \tag{12.422}$$

Pour le reste c'est la linéarité de la différentielle qui joue : le vecteur u peut être écrit de façon unique comme combinaison linéaire des vecteurs de base

$$u = \sum_{i=1}^{m} u_i e_i, \qquad u_i \in \mathbb{R},\ \forall i \in \{1, \dots, m\}.$$

Alors, la linéarité de df_a nous donne

$$df_a(u) = df_a\left(\sum_{i=1}^{m} u_i e_i\right) = \sum_{i=1}^{m} u_i \left(df_a e_i\right) = \sum_{i=1}^{m} u_i \frac{\partial f}{\partial x_i}(a). \tag{12.423}$$

Le lien avec le gradient est la définition du produit scalaire (9.263). □

Proposition 12.179.
Soit f une fonction de x et y et un point $(a,b) \in \mathbb{R}^2$. Si les nombres $\partial_x f(a,b)$ et $\partial_y f(a,b)$ existent et s'il existe une fonction $\alpha\colon \mathbb{R} \to \mathbb{R}$ telle que

$$\begin{aligned} f(x,y) = f(a,b) &+ \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b) \\ &+ \|(x,y)-(a,b)\|\alpha\Big(\|(x,y)-(a,b)\|\Big) \end{aligned} \tag{12.424}$$

et

$$\lim_{t\to 0} \alpha(t) = 0, \tag{12.425}$$

alors f est différentiable en (a,b).

Dans cet énoncé nous avons écrit $d\big((x,y),(a,b)\big)$ la distance entre (x,y) et (a,b), c'est à dire le nombre $\sqrt{(x-a)^2+(y-b)^2}$. Afin d'écrire l'équation (12.424) sous forme plus compacte, nous introduisons le vecteur

$$\nabla f(a,b) = \begin{pmatrix} \frac{\partial f}{\partial x}(a,b) \\ \frac{\partial f}{\partial y}(a,b). \end{pmatrix} \tag{12.426}$$

L'équation (12.424) devient alors

$$f(X) = f(P) + \nabla f(a,b) \cdot (X-P) + \|X-P\|\alpha\big(\|X-P\|\big). \tag{12.427}$$

Le vecteur $\nabla f(a,b)$ est appelé le **gradient** de f au point (a,b).

Proposition 12.180.
Soit f une fonction de deux variables admettant des dérivées partielles $\partial_x f(x,y)$ et $\partial_y f(x,y)$ qui sont elles-mêmes des fonctions continues de x et y. Alors la fonction f est différentiable partout.

Proposition 12.181.
Si f est différentiable en (a,b) alors pour tout vecteur u, la fonction

$$\begin{aligned}\varphi\colon \mathbb{R} &\to \mathbb{R}\\ t &\mapsto f(a+tu_1,b+tu_2)\end{aligned} \tag{12.428}$$

est dérivable en 0 et on a

$$\varphi'(0) = \nabla f(p)\cdot u \tag{12.429}$$

où nous avons noté $p=(a,b)$.

Démonstration. Récrivons la formule (12.427) sous la forme

$$f(x) = f(p) + \nabla f(p)\cdot(x-p) + \|x-p\|\alpha(\|x-p\|). \tag{12.430}$$

Cela étant vrai pour tout x, nous l'écrivons en particulier pour $x = p+tu$ où t est un réel et u est le vecteur unitaire choisi. Nous avons donc

$$f(p+tu) = f(p) + t\nabla f(p)\cdot u + \|tu\|\alpha(\|tu\|). \tag{12.431}$$

En utilisant le fait que u est unitaire, $\|tu\| = |t|\|u\| = |t|$. La dérivée de φ en 0 est alors donnée par

$$\lim_{t\to 0}\frac{f(p+tu)-f(p)}{t} = \lim_{t\to 0}\nabla f(p)\cdot u + \alpha(|t|). \tag{12.432}$$

Lorsque nous prenons la limite, le membre de gauche devient $\varphi'(0)$ tandis que dans le membre de droite, le second terme disparaît. Nous avons finalement

$$\varphi'(0) = \nabla f(p)\cdot u \tag{12.433}$$

□

12.19.6 Calcul de valeurs approchées

Si nous remplaçons les accroissements $x-a$ et $y-b$ par h et k, le critère de différentiabilité s'écrit

$$\begin{aligned}f(a+h,b+k) = f(a,b) + \frac{\partial f}{\partial x}(a,b)h + \frac{\partial f}{\partial y}(a,b)k\\ + \sqrt{h^2+k^2}\alpha\big(\sqrt{h^2+k^2}\big).\end{aligned} \tag{12.434}$$

Le dernier terme du membre de droite tend vers zéro à une vitesse double lorsque h et k tendent vers zéro : d'une part parce que $\sqrt{h^2+k^2}$ tend vers zéro et d'autre part parce que $\alpha\big(\sqrt{h^2+k^2}\big)$ tend vers zéro. Nous avons donc la «bonne» approximation

$$f(x,y) \simeq f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b). \tag{12.435}$$

lorsque (x,y) n'est pas trop loin de (a,b). Cette expression est évidemment une généralisation immédiate de l'équation (12.239). Elle exprime que l'on peut obtenir des information sur la valeur d'une fonction en (x,y) si on peut calculer la fonction et ses dérivées en un point (a,b) non loin de (x,y).

Cette formule peut aussi être vue sous la forme suivante, plus pratique dans certains calculs :

$$f(a+\Delta x, b+\Delta y) \simeq f(a,b) + \Delta x\frac{\partial f}{\partial x}(a,b) + \Delta y\frac{\partial f}{\partial y}(a,b). \tag{12.436}$$

Exemple 12.182
Prenons la fonction $f(x,y)=\cos(x)\sin(y)$ et calculons une approximation de

$$f\big(\frac{\pi}{3}+0.01,\frac{\pi}{2}+0.03\big). \tag{12.437}$$

D'abord les dérivées partielles sont

$$\begin{aligned}\frac{\partial f}{\partial x}(x,y)&=-\sin(x)\sin(y)\\ \frac{\partial f}{\partial y}(x,y)&=\cos(x)\cos(y).\end{aligned} \tag{12.438}$$

Nous allons utiliser l'approximation

$$f\big(\frac{\pi}{3}+0.01,\frac{\pi}{2}+0.03\big)\simeq f\big(\frac{\pi}{3},\frac{\pi}{2}\big)+0.01\frac{\partial f}{\partial x}\big(\frac{\pi}{3},\frac{\pi}{2}\big)+0.03\frac{\partial f}{\partial y}\big(\frac{\pi}{3},\frac{\pi}{2}\big). \tag{12.439}$$

Nous avons

$$\begin{aligned}\frac{\partial f}{\partial x}\big(\frac{\pi}{3},\frac{\pi}{2}\big)&=-\sin\frac{\pi}{3}\sin\frac{\pi}{2}=-\frac{\sqrt{3}}{2}\\ \frac{\partial f}{\partial y}\big(\frac{\pi}{3},\frac{\pi}{2}\big)&=\cos\frac{\pi}{3}\cos\frac{\pi}{2}=0.\end{aligned} \tag{12.440}$$

Par conséquent

$$f\big(\frac{\pi}{3}+0.01,\frac{\pi}{2}+0.03\big)\simeq\frac{1}{2}-0.01\frac{\sqrt{3}}{2}=\frac{1}{2}-\frac{\sqrt{3}}{200}. \tag{12.441}$$

```
sage: var('x,y')
(x, y)
sage: f(x,y)=cos(x)*sin(y)
sage: a=f(pi/3+0.01,pi/2+0.03)
sage: numerical_approx(a)
0.491093815387986
sage: b=1/2-sqrt(3)/200
sage: numerical_approx(b)
0.491339745962156
sage: numerical_approx(a-b)
-0.000245930574169814
```

Cela fait une erreur de l'ordre du dix millième.

△

Remarque 12.183.
Les esprits les plus critiques diront que cette vérification pas Sage n'en est pas une parce que Sage a certainement utilisé un algorithme d'approximation qui se base sur la même idée que ce que nous venons de faire, et que par conséquent le fait qu'il obtienne le même résultat que nous est un peu tautologique.

Ils n'auront pas tord. Cependant, le code source de Sage est disponible publiquement [27] ; vous pouvez aller le lire et vérifier qu'il y a effectivement une *preuve* que le résultat fourni par Sage possède une bonne dizaine de décimales correctes.

Cette disponibilité publique du code source est une des nombreuses différence fondamentale entre Sage et votre calculatrice [28]. Dois-je vous rappeler qu'un des principes fondamentaux de l'éthique scientifique est que les résultats et les méthodes utilisées doivent être absolument ouverts à la vérification et à la critique de tous ?

27. Voir `http://www.sagemath.org`
28. et les autres logiciels de type fenêtre, pomme ou feuille d'érable.

$$df_p(u) = \nabla f(p) \cdot u. \tag{12.442}$$

Définition 12.184.
*Soit un point $a \in \operatorname{int} A$. La fonction f est **différentiable** au point a s'il existe une application linéaire $df_a \colon \mathbb{R}^n \to \mathbb{R}^m$ telle que*

$$\lim_{x \to a} \frac{f(x) - f(a) - df_a(x-a)}{\|x-a\|} = 0. \tag{12.443}$$

Si f est différentiable en a, l'application df_a est appelée la différentielle de f en a. Voyons comment cette application linéaire agit sur les vecteurs de $\mathbb{R}^n$.

La notion de dérivée partielle (ou de dérivée suivant un vecteur) pour une fonction de plusieurs variables n'est pas une généralisation de la notion de dérivée en une variable d'espace. En fait, du point de vue géométrique, la dérivée de la fonction $g : \mathbb{R} \to \mathbb{R}$ au point a est la pente de la ligne droite tangente au graphe de g au point $(a, g(a))$. Cette ligne, d'équation $r(x) = g'(a)x + g(a)$, est la meilleure approximation affine du graphe de g au point a, comme à la figure 12.10.

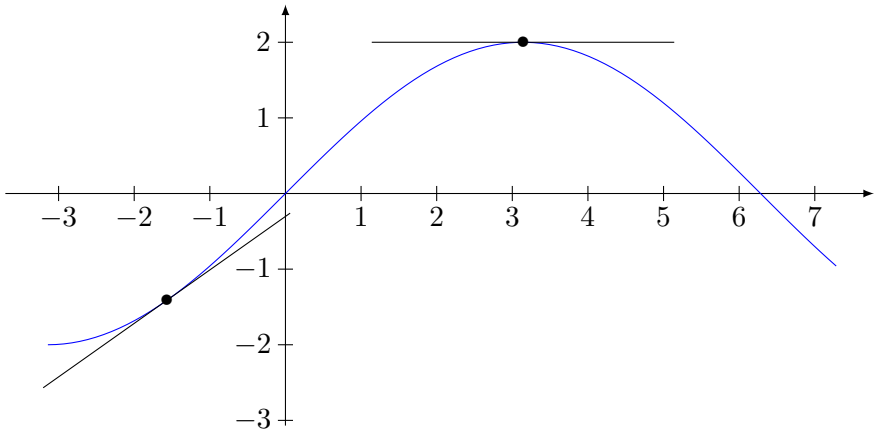


FIGURE 12.10 – Tangentes au graphe d'une fonction d'une variable

Le graphe d'une fonction f de $\mathbb{R}^2$ dans $\mathbb{R}$ est une surface de deux paramètres dans $\mathbb{R}^3$. Si l'approximation affine d'une telle surface au point $(x, y, f(x, y))$ existe, alors elle est un plan tangent. En dimension plus haute, le graphe de la fonction $f : \mathbb{R}^m \to \mathbb{R}$ est une surface de m paramètres dans $\mathbb{R}^{m+1}$ et son approximation affine (si elle existe) est un hyperplan de $\mathbb{R}^m$.

Nous allons voir que si f prend ses valeurs dans $\mathbb{R}^n$ l'approximation affine de f au point a est l'élément de $f(a) + \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ qui ressemble le plus à f au voisinage de a. Plus précisément, on utilise les définitions suivantes.

Définition 12.185.
*Soient f et g deux applications d'un ouvert U de $\mathbb{R}^m$ dans $\mathbb{R}^n$. On dit que g est **tangente** à f au point $a \in U$ si $f(a) = g(a)$ et*

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{\|f(x) - g(x)\|_n}{\|x-a\|_m} = 0.$$

La relation de tangence est une relation d'équivalence. Nous sommes particulièrement intéressés par le cas où f admet une application affine tangente au point a.

En ce qui concerne l'interprétation géométrique, si nous regardons la figure 12.11, et d'ailleurs aussi en voyant la définition 12.405, la fonction est différentiable et la différentielle est T s'il existe une fonction α telle que

$$f(a+u) - f(a) - T(u) = \alpha(u) \tag{12.444}$$

où la fonction α satisfait

$$\lim_{u \to 0} \frac{\|\alpha(u)\|}{\|u\|} = 0 \tag{12.445}$$

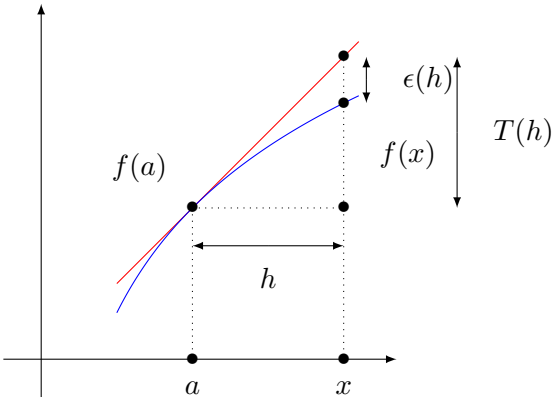


FIGURE 12.11 – Interprétation géométrique de la différentielle.

C'est cela qui fait écrire $f(a+u) - f(a) - df_a(u) = o(\|u\|)$ à ceux qui n'ont pas peur de la notation o.

La différentielle df_a est donc la partie linéaire de l'application affine qui approxime au mieux la fonction f autour du point a. La notion de différentielle est la vraie généralisation du concept de dérivée pour fonctions de plusieurs variables, en outre elle nous permet d'expliciter la relation qui associe au vecteur u la dérivée $\partial_u f(a)$, pour f et a fixés.

Remarque 12.186.
Si on remplace les normes $\|\cdot\|_m$ et $\|\cdot\|_n$ par d'autres normes, l'existence et la valeur de la différentielle de f au point a ne sont pas remises en cause. En effet, soient $\|\cdot\|_M$ une norme sur $\mathbb{R}^m$ et $\|\cdot\|_N$ une norme sur $\mathbb{R}^n$. Par le théorème 10.3, ces normes sont équivalentes à $\|.\|_m$ et $\|.\|_m$ respectivement ; il existe donc des constantes k, K, l, $L > 0$ telles que pour tout vecteur u de $\mathbb{R}^m$ et tout vecteur v de $\mathbb{R}^n$

$$k\|u\|_M \leqslant \|u\|_m \leqslant K\|u\|_M,$$
$$l\|v\|_N \leqslant \|v\|_n \leqslant L\|v\|_N.$$

Les éléments de $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ sont les mêmes et on a

$$\begin{aligned}\frac{l}{K}\frac{\|f(a+h)-f(a)-T(h)\|_N}{\|h\|_M} \leqslant \frac{\|f(a+h)-f(a)-T(h)\|_n}{\|h\|_m} \leqslant \\ \leqslant \frac{L}{k}\frac{\|f(a+h)-f(a)-T(h)\|_N}{\|h\|_M}.\end{aligned} \tag{12.446}$$

Il est donc possible, pour démontrer la différentiabilité ou pour calculer la différentielle, d'utiliser le critère (12.405) avec une norme au choix. Parfois c'est utile.

12.19.7 Prouver qu'un fonction n'est pas différentiable

Chacun des point du théorème 12.176 est en soi un critère pour montrer qu'une fonction n'est pas différentiable en un point.

12.19.7.1 Continuité

Le premier critère à vérifier est donc la continuité. Si une fonction n'est pas continue en un point, alors elle n'y sera pas différentiable. Pour rappel, la continuité en a se teste en vérifiant si $\lim_{x\to a} f(x) = f(a)$.

12.19.7.2 Linéarité

Un second test est la linéarité de la dérivée directionnelle par rapport à la direction : l'application $u \mapsto \frac{\partial f}{\partial u}(a)$ doit être linéaire, sinon df_a n'existe pas.

Exemple 12.187
Examinons la fonction

$$\begin{aligned} f\colon \mathbb{R}^2 &\to \mathbb{R} \\ (x,y) &\mapsto \begin{cases} \frac{xy^2}{x^2+y^4} & \text{si } (x,y)\neq(0,0) \\ 0 & \text{sinon.} \end{cases} \end{aligned} \tag{12.447}$$

Prenons $u=(u_1,u_2)$ et calculons la dérivée de f dans la direction de u au point $(0,0)$:

$$\begin{aligned} \frac{\partial f}{\partial u}(0,0) &= \lim_{t\to 0}\frac{f(tu_1,tu_2)-f(0,0)}{t} \\ &= \lim_{t\to 0}\frac{1}{t}\left(\frac{tu_1t^2u_2}{t^2u_1^2+t^4u_2^4}\right) \\ &= \lim_{t\to 0}\left(\frac{u_1u_2^2}{u_1^2+t^2u_2^4}\right) \\ &= \begin{cases} \frac{u_2^2}{u_1} & \text{si } u_1\neq 0 \\ 0 & \text{si } u_1=0. \end{cases} \end{aligned} \tag{12.448}$$

Cette application n'est pas linéaire par rapport à u. En effet, notons

$$\begin{aligned} A\colon \mathbb{R}^n &\to \mathbb{R} \\ u &\mapsto \frac{\partial f}{\partial u}(0,0), \end{aligned} \tag{12.449}$$

et vérifions que pour tout u et v dans $\mathbb{R}^n$ et $\lambda\in\mathbb{R}$, nous ayons $A(\lambda u)=\lambda A(u)$ et $A(u+v)=A(u)+A(v)$. Le premier fonctionne parce que

$$A(\lambda u)=A(\lambda u_1,\lambda u_2)=\frac{\lambda^2u_2^2}{\lambda u_1}=\lambda\frac{u_2^2}{u_1}=\lambda A(u). \tag{12.450}$$

Mais nous avons par exemple

$$A\big((0,1)+(2,3)\big)=A(2,4)=\frac{16}{2}=8, \tag{12.451}$$

tandis que

$$A(0,1)+A(2,3)=0+\frac{9}{2}\neq 8. \tag{12.452}$$

La fonction f n'est donc pas différentiable en $(0,0)$, parce que la candidate différentielle, $df_{(0,0)}(u)=\frac{\partial f}{\partial u}(0,0)$, n'est même pas linéaire.

△

Voici une autre façon de traiter la fonction de l'exemple 12.187.

Exemple 12.188
La figure 12.12 représente le domaine d'une fonction $f\colon\mathbb{R}^2\to\mathbb{R}$, et sur chacune des parties, elle est définie différemment.

L'expression de f est ici

$$f(x,y)=\begin{cases} xy & \text{si } x<0 \text{ et } y>0 \\ x-y & \text{si } x\geqslant 0 \text{ et } y\geqslant 0 \\ x^2y & \text{si } x>0 \text{ et } y<0 \\ x+y & \text{sinon.} \end{cases} \tag{12.453}$$

On note que les deux axes forment une zone à problèmes. La zone hors des axes est un ouvert sur lequel f est différentiable car composée de polynômes. Analysons chacun des points de la forme (a,b) dans la zone à problèmes (c'est-à-dire si $ab=0$).

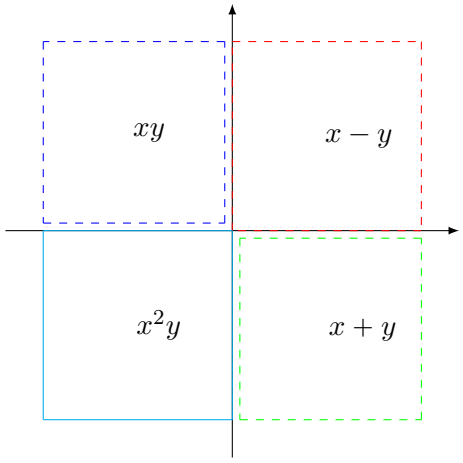


FIGURE 12.12 – La fonction de l'exemple 12.188.

Si $a=0$ **et** $b>0$ Un tel point $(0,b)$ est sur l'axe verticale, dans la moitié supérieure. Pour calculer la limite de f en ce point, on peut restreindre notre étude au demi-plan ouvert $y>0$, ce qui revient à comparer la limite

$$\lim_{\substack{(x,y)\to(0,b)\\ y>0\\ x\geqslant 0}} f(x,y) = \lim_{\substack{(x,y)\to(0,b)\\ y>0\\ x\geqslant 0}} x-y = 0-b = -b$$

avec la limite

$$\lim_{\substack{(x,y)\to(0,b)\\ y>0\\ x<0}} f(x,y) = \lim_{\substack{(x,y)\to(0,b)\\ y>0\\ x<0}} xy = 0b = 0$$

qui sont différentes puisque b est supposé non-nul.

Conclusion : f n'est pas continue en un point du type $(0,b)$ avec $b>0$.

Si $a=0$ **et** $b<0$ Un tel point $(0,b)$ est sur l'axe verticale, dans la moitié inférieure. Pour calculer la limite de f en ce point, on peut restreindre notre étude au demi-plan ouvert $y<0$, ce qui revient à comparer la limite

$$\lim_{\substack{(x,y)\to(0,b)\\ y<0\\ x\geqslant 0}} f(x,y) = \lim_{\substack{(x,y)\to(0,b)\\ y<0\\ x\geqslant 0}} x^2y = 0^2b = 0$$

avec la limite

$$\lim_{\substack{(x,y)\to(0,b)\\ y<0\\ x<0}} f(x,y) = \lim_{\substack{(x,y)\to(0,b)\\ y<0\\ x<0}} x+y = 0+b = b$$

qui sont différentes puisque b est supposé non-nul.

Conclusion : f n'est pas continue en un point du type $(0,b)$ avec $b<0$.

Si $a>0$ **et** $b=0$ Un tel point $(a,0)$ est sur l'axe horizontal, dans la moitié droite. Pour calculer la limite de f en ce point, on peut restreindre notre étude au demi-plan ouvert $x>0$, ce qui revient à comparer la limite

$$\lim_{\substack{(x,y)\to(a,0)\\ x>0\\ y\geqslant 0}} f(x,y) = \lim_{\substack{(x,y)\to(a,0)\\ x>0\\ y\geqslant 0}} x-y = a-0 = a$$

avec la limite

$$\lim_{\substack{(x,y)\to(a,0)\\ x>0\\ y<0}} f(x,y) = \lim_{\substack{(x,y)\to(a,0)\\ x>0\\ y<0}} x^2y = a^20 = 0$$

qui sont différentes puisque a est supposé non-nul.

Conclusion : f n'est pas continue en un point du type $(a,0)$ avec $a>0$.

Si $a<0$ et $b=0$ Un tel point $(a,0)$ est sur l'axe horizontal, dans la moitié gauche. Pour calculer la limite de f en ce point, on peut restreindre notre étude au demi-plan ouvert $x<0$, ce qui revient à comparer la limite

$$\lim_{\substack{(x,y)\to(a,0)\\x<0\\y>0}} f(x,y) = \lim_{\substack{(x,y)\to(a,0)\\x<0\\y>0}} xy = a0 = 0$$

avec la limite

$$\lim_{\substack{(x,y)\to(a,0)\\x<0\\y\leqslant 0}} f(x,y) = \lim_{\substack{(x,y)\to(a,0)\\x<0\\y\leqslant 0}} x+y = a+0 = a$$

qui sont différentes puisque a est supposé non-nul.

Conclusion : f n'est pas continue en un point du type $(a,0)$ avec $a<0$.

Si $a=0$ et $b=0$ Le cas du point $(0,0)$ est particulier, puisque il est adhérent aux quatre composantes du domaine où la fonction est définie différemment. Pour étudier la continuité, il faut donc étudier quatre limites. Ces limites ont déjà été étudiées ci-dessus et valent toutes 0, ce qui prouve la continuité de f en $(0,0)$.

En ce qui concerne la différentiabilité, on sait qu'il est nécessaire que toutes les dérivées directionnelles existent. Calculons la dérivée dans la direction $(0,1)$ (au point $(0,0)$) :

$$\lim_{\substack{t\to 0\\t\neq 0}} \frac{f((0,0)+t(0,1))-f(0,0)}{t} = \lim_{\substack{t\to 0\\t\neq 0}} \frac{f(0,t)}{t} = \dots$$

qu'on sépare en deux cas, car $f(0,t)$ possède une formule différente si $t<0$ ou si $t\geqslant 0$:

$$\lim_{\substack{t\to 0\\t\neq 0}} \frac{f(0,t)}{t} = \begin{cases} \lim_{\substack{t\to 0\\t<0}} \frac{f(0,t)}{t} = \lim_{\substack{t\to 0\\t<0}} \frac{0+t}{t} = 1 \\ \lim_{\substack{t\to 0\\t\geqslant 0}} \frac{f(0,t)}{t} = \lim_{\substack{t\to 0\\t\geqslant 0}} \frac{0-t}{t} = -1 \end{cases}$$

ce qui prouve que la limite n'existe pas, donc que la dérivée directionnelle n'existe pas, et finalement que la fonction n'est pas différentiable.

Conclusion : La fonction donnée est continue hors des axes et au point $(0,0)$, mais discontinue partout ailleurs sur les axes. Elle est différentiable hors des axes, mais ne l'est pas sur les axes.

△

12.19.7.3 Cohérence des dérivées partielles et directionnelle

Dans la pratique, nous pouvons calculer $\partial_u f(a)$ pour une direction u générale, et puis en déduire $\partial_x f$ et $\partial_y f$ comme cas particuliers en posant $u=(1,0)$ et $u=(0,1)$. Une chose incroyable, mais pourtant possible est qu'il peut arriver que

$$\frac{\partial f}{\partial u}(a) \neq \sum_i \frac{\partial f}{\partial x_i}(a) u^i. \tag{12.454}$$

Ceci se produit lorsque f n'est pas différentiable en a. En voici un exemple.

12.19.7.4 Un candidat dans la définition (marche toujours)

Lorsqu'une fonction est donné, un candidat différentielle au point (a_1, a_2) est souvent assez simple à trouver en un point :

$$T(u_1, u_2) = \frac{\partial f}{\partial x}(a_1, a_2)u_1 + \frac{\partial f}{\partial y}(a_1, a_2)u_2. \tag{12.455}$$

L'application T est la candidate différentielle en ce sens que si la différentielle existe, alors elle est égale à T. Ensuite, il faut vérifier si

$$\lim_{(x,y)\to(a_1,a_2)} \frac{f(x,y) - f(a_1,a_2) - T\big((x,y)-(a_1,a_2)\big)}{\|(x,y)-(a_1,a_2)\|} = 0 \tag{12.456}$$

ou non. Si oui, alors la différentielle existe et $df_{(a,b)}(u) = T(u)$, sinon [29], la différentielle n'existe pas.

Attention : dans la ZAP, les dérivées partielles $\partial_x f$ et $\partial_y f$ ne peuvent en général pas être calculées en utilisant les règles de calcul (c'est bien pour ça que la ZAP est une zone à problèmes). Il faut d'office utiliser la définition

$$\frac{\partial f}{\partial x}(a_1, a_2) = \lim_{t\to 0} \frac{f(a_1 + t, a_2) - f(a_1, a_2)}{t}, \tag{12.457}$$

et la définition correspondante pour $\partial_y f$.

Conclusion

Soient $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, et $a \in int\, A$. Si f est différentiable en a,

$$(df_a(e_j))_i = d(f_i)_a(e_j) = \frac{\partial f_i}{\partial x_j}(a) = [Jac(f)_{|a}]_{ij}$$

et la matrice de l'application linéaire df_a est la matrice jacobienne $m \times n$ de f en a notée $Jac(f)_{|a}$.

12.19.8 Règles de calcul

Proposition 12.189 (Règles de calculs)**.**
Soient f et g des fonctions différentiables en $g(a)$ et a respectivement, alors la composée $f \circ g$ est différentiable en a et

$$d(f \circ g)_a = df_{g(a)} \circ dg_a$$

et de plus les jacobiennes correspondantes vérifient

$$J_{f\circ g}(a) = J_f\big(g(a)\big)J_g(a)$$

où le membre de droite est le produit (non-commutatif!) des deux matrices.

Corollaire 12.190 (Chain rule)**.**
Si $f : \mathbb{R}^p \to \mathbb{R}$ et $g : \mathbb{R} \to \mathbb{R}^p$, alors

$$(f \circ g)'(t) = \sum_{i=1}^{p} \frac{\partial f}{\partial x_i}(g(t))g_i'(t).$$

Remarque 12.191.
(1) Si $p = 1$, on retrouve la règle usuelle de dérivation de fonctions composées.
(2) Si g est à plusieurs variables, cette règle permet de déterminer les dérivées partielles de $f \circ g$, puisqu'une dérivée partielle peut être vue comme dérivée usuelle par rapport à une seule variable (voir remarque page 609).

29. y compris si la limite (12.456) n'existe même pas.

(3) Si f est à valeurs vectorielles, cette formule permet de retrouver la jacobienne de $f \circ g$ puisqu'il suffit de traiter chaque composante de f séparément.

Définition 12.192.
*Une fonction $f : \mathbb{R}^m \to \mathbb{R}^n$ est dite **différentiable sur l'ouvert** $U \subset \mathbb{R}^m$, si f est différentiable en tout point de U. Dans ce cas, la différentielle de f est l'application*

$$\begin{aligned} df : U \subset \mathbb{R}^m &\to \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n) \\ x &\mapsto df(x). \end{aligned} \tag{12.458}$$

Remarque 12.193.
Tout élément T de $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ est différentiable en tout point de $\mathbb{R}^m$ et coïncide avec sa différentielle. En effet, pour tout a et h dans $\mathbb{R}^m$ on a

$$\frac{\|T(a+h) - T(a) - T(h)\|_n}{\|h\|_m} = 0.$$

La proposition 12.176 nous donne une recette très pratique pour calculer la différentielle d'une fonction de $\mathbb{R}^m$ dans $\mathbb{R}^n$.

Définition 12.194.
*Soit f une fonction différentiable de $\mathbb{R}^m$ dans $\mathbb{R}$. On appelle **gradient** de f la fonction $\nabla f : \mathbb{R}^m \to \mathbb{R}^m$ de composantes*

$$\partial_1 f, \ldots, \partial_m f.$$

*Soit f une fonction de $\mathbb{R}^m$ dans $\mathbb{R}^n$, $f(a) = (f_1(a), \ldots, f_n(a))^T$. On appelle **matrice jacobienne** de f la fonction $J(f) : \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^n$ définie par*

$$a \mapsto \begin{pmatrix} \partial_1 f_1(a) & \ldots & \partial_m f_1(a) \\ \vdots & \ddots & \vdots \\ \partial_1 f_n(a) & \ldots & \partial_m f_n(a) \end{pmatrix} \tag{12.459}$$

12.19.9 Linéarité

La proposition suivante signifie que différentiation est une opération linéaire sur l'ensemble des fonctions différentiables.

Proposition 12.195.
Soient f et g deux fonction de $U \subset \mathbb{R}^m$ dans $\mathbb{R}^n$ différentiables au point $a \in U$, et soit λ dans $\mathbb{R}$. Alors les fonctions $f + g$ et λf sont différentiables au point a et on a

$$\begin{aligned} d(f+g)(a) &= df(a) + dg(a), \\ d(\lambda f)(a) &= \lambda df(a), \end{aligned} \tag{12.460}$$

Démonstration.

$$\begin{aligned} &\lim_{h \to 0_m} \frac{\|(f(a+h) + g(a+h)) - (f(a) + g(a)) - df(a).h - dg(a).h\|_n}{\|h\|_m} \leqslant \\ &\lim_{h \to 0_m} \frac{\|f(a+h) - f(a) - df(a).h\|_n}{\|h\|_m} + \lim_{h \to 0_m} \frac{\|g(a+h) - g(a) - dg(a).h\|_n}{\|h\|_m} = 0. \end{aligned} \tag{12.461}$$

De même on démontre la propriété $d(\lambda f)(a) = \lambda df(a)$. □

12.19.10 Produit

Soient f et g deux fonctions de $\mathbb{R}^m$ dans $\mathbb{R}^n$. Nous notons $f \cdot g$ la fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ donnée par le produit scalaire point par point, c'est à dire

$$(f \cdot g)(x) = f(x) \cdot g(x) \tag{12.462}$$

pour tout $x \in \mathbb{R}^m$. Le point dans le membre de droite est le produit scalaire dans $\mathbb{R}^n$. Le cas particulier $n = 1$ revient au produit usuel de fonctions :

$$(fg)(x) = f(x)g(x). \tag{12.463}$$

Lemme 12.196.
Si f et g sont des fonctions différentiables sur $\mathbb{R}^m$ à valeurs dans $\mathbb{R}$, alors la fonction produit fg est également différentiable et

$$d(fg)(a) = df(a)g(a) + f(a)dg(a) \tag{12.464}$$

au sens où pour chaque u dans $\mathbb{R}^m$,

$$d(fg)(a).u = g(a)df(a).u + f(a)dg(a).u. \tag{12.465}$$

Remarquons qu'ici, $f(a)$ et $g(a)$ sont des réels, donc nous pouvons écrire $f(a)dg(a)$ aussi bien que $dg(a)f(a)$ sans ambigüités.

Démonstration. Ce que nous devons faire pour vérifier la formule 12.464, c'est de vérifier le critère (12.405) en remplaçant f par fg et $T(h)$ par $g(a)df(a).h + f(a)dg(a).h$.

Ce que nous avons au numérateur est

$$\begin{aligned} \clubsuit &= (fg)(a+h) - (fg)(a) - g(a)df(a).h - f(a)dg(a).h \\ &= f(a+h)g(a+h) - f(a)g(a) - g(a)df(a).h - f(a)dg(a).h. \end{aligned} \tag{12.466}$$

Maintenant, nous allons faire apparaître $\big(f(a+h)-f(a)-df(a)\big)g(a+h)$ en ajoutant et soustrayant ce qu'il faut pour conserver $\clubsuit$:

$$\begin{aligned} \clubsuit = {} & \big(f(a+h) - f(a) - df(a).h\big)g(a+h) \\ & + f(a)g(a+h) + g(a+h)df(a).h \\ & - f(a)g(a) - g(a)df(a).h - f(a)dg(a).h. \end{aligned} \tag{12.467}$$

Nous mettons maintenant $f(a)$ et $fd(a).h$ en évidence là où c'est possible :

$$\begin{aligned} \clubsuit = {} & \big(f(a+h) - f(a) - df(a).h\big)g(a+h) \\ & + f(a)\big(g(a+h) - g(a) - dg(a).h\big) \\ & + \big(g(a+h) - g(a)\big)df(a).h. \end{aligned} \tag{12.468}$$

Nous devons maintenant considérer la limite

$$\lim_{h\to 0} \frac{\|\clubsuit\|}{\|h\|}. \tag{12.469}$$

Étant donné que f et g sont différentiables, les deux premiers termes sont nuls :

$$\begin{aligned} \lim_{h\to 0} \frac{\big(f(a+h) - f(a) - df(a).h\big)}{\|h\|} g(a+h) = 0 \\ \lim_{h\to 0} f(a) \frac{\big(g(a+h) - g(a) - dg(a).h\big)}{\|h\|} = 0. \end{aligned} \tag{12.470}$$

En ce qui concerne le troisième terme, en utilisant la norme d'une application linéaire, nous avons

$$\lim_{h\to 0} \frac{\|df(a).h\|}{\|h\|} \leqslant \sup_{h\in\mathbb{R}^m} \frac{\|df(a).h\|}{\|h\|} = \|df(a)\|, \tag{12.471}$$

et par conséquent

$$\begin{aligned} 0 &\leqslant \lim_{h\to 0} \|g(a+h) - g(a)\| \frac{\|df(a).h\|\|h\|}{\|h\|} \\ &\leqslant \lim_{h\to 0} \|g(a+h) - g(a)\|\|df(a)\| = 0 \end{aligned} \tag{12.472}$$

parce que g est continue (la limite du premier facteur est nulle tandis que la norme de $df(a)$ est un nombre constant). Nous avons donc bien prouvé que la formule (12.464) est la différentielle de fg au point a. □

Ce résultat se généralise pour des fonctions f et g de $\mathbb{R}^m$ dans $\mathbb{R}^n$.

Proposition 12.197.
Soient f et g deux fonction de $U \subset \mathbb{R}^m$ dans $\mathbb{R}^n$ différentiables au point $a \in U$. Alors la fonction $f \cdot g$ est différentiable au point a et on a

$$g(f \cdot g)(a) = g(a) \cdot df(a) + f(a) \cdot dg(a) \tag{12.473}$$

au sens où

$$d(f \cdot g)_a(u) = g(a) \cdot \big(df_a(u)\big) + f(a) \cdot \big(dg_a(u)\big) \tag{12.474}$$

pour tout $u \in \mathbb{R}^m$.

Note : il faut être bien attentif en lisant la formule (12.474). Les points à l'intérieur des grandes parenthèses marquent l'application des différentielles sur u. Le contenu de ces parenthèses sont donc des éléments de $\mathbb{R}^n$. Les points devant les parenthèses dénotent le produit scalaire dans $\mathbb{R}^n$ ($f(a)$ et $dg_a(u)$ sont des éléments de $\mathbb{R}^n$).

Démonstration. La preuve du cas $n = 1$ est déjà faite ; c'est la formule (12.464). Pour le cas général $n \geqslant 2$, nous passons au composantes en nous rappelant que

$$(f \cdot g)(a) = \sum_{i=1}^{n} f_i(a)g_i(a) = \sum_{i=1}^{n}(f_ig_i)(a). \tag{12.475}$$

En utilisant la linéarité de la différentiation, nous nous réduisons donc au cas des produits f_ig_i qui sont des fonctions de $\mathbb{R}^m$ dans $\mathbb{R}$:

$$\begin{aligned} d(f \cdot g)(a) &= d\left(\sum_{i=1}^{n} f_ig_i\right)(a) \\ &= \sum_{i=1}^{n} \big(df_i(a)g_i(a) + f_i(a)dg_i(a)\big) \\ &= g(a) \cdot df(a) + f(a) \cdot dg(a). \end{aligned} \tag{12.476}$$

Ceci termine la preuve. □

12.19.11 Différentielle de fonction composée

La plus importante entre les règles de différentiation est la règle de différentiation d'une fonction composée (*chain rule* dans les livres anglais et américains). Cette règle généralise la règle de dérivation pour fonctions de $\mathbb{R}$ dans $\mathbb{R}$. Il est utile d'introduire d'abord une formulation équivalente de la définition de différentielle

Lemme 12.198.
Soit U un ouvert de $\mathbb{R}^m$. La fonction $f : U \to \mathbb{R}^n$ est différentiable au point a dans U, si et seulement s'il existe une fonction $\sigma_f : U \times U \to \mathbb{R}^n$ telle que

$$\sigma_f(a, a) = \lim_{x \to a} \sigma_f(a, x) = 0 \tag{12.477a}$$

$$f(x) = f(a) + T(x - a) + \sigma_f(a, x)\|x - a\|_m, \tag{12.477b}$$

pour une certaine application linéaire $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$.

Démonstration. Si les conditions (12.477) sont satisfaites alors T est la différentielle de f en a. En effet, dans ce cas nous avons

$$f(a + h) = f(a) + T(h) + \sigma_f(a, a + h)\|h\|, \tag{12.478}$$

et la condition (12.405) devient

$$\lim_{h\to 0} \frac{\|\sigma_f(a,a+h)\|\|h\|}{\|h\|} = \lim_{h\to 0} \|\sigma_f(a,a+h)\| = 0 \tag{12.479}$$

Si f est différentiable au point a il suffit de prendre $T = df(a)$ et

$$\sigma_f(a,x) = \frac{f(x) - f(a) - df(a).(x-a)}{\|x-a\|_m}.$$

□

Remarque 12.199.
La fonction $\sigma_f(a,x)\|x-a\|_m$ est ce qui avait été appelle $\epsilon(h)$ sur la figure 12.11.

Proposition 12.200.
Soient U un ouvert de $\mathbb{R}^m$ et V un ouvert de $\mathbb{R}^n$. Soient $f : U \to V$ et $g : V \to \mathbb{R}^p$ deux fonctions différentiables respectivement au point a dans U et $b = f(a)$ dans V. Alors la fonction composée $g \circ f : U \to \mathbb{R}^p$ est différentiable au point a et

$$d(g\circ f)_a = dg_{f(a)} \circ df_a. \tag{12.480}$$

Note : la formule (12.480) est à comprendre de la façon suivante. Si $u \in \mathbb{R}^m$, alors

$$d(g\circ f)_a(u) = \underbrace{dg_{f(a)}}_{\in \mathcal{L}(\mathbb{R}^n,\mathbb{R}^p)} \Big(\underbrace{df_a(u)}_{\in \mathbb{R}^n} \Big) \in \mathbb{R}^p. \tag{12.481}$$

Démonstration. En tenant compte du lemme 12.198 on peut écrire

$$f(a+h) - f(a) = df_a(h) + \sigma_f(a,a+h)\|h\|_m, \qquad \forall h \in U-a, \tag{12.482a}$$
$$g(b+k) - g(b) = dg_b(k) + \sigma_g(b,b+k)\|k\|_n, \qquad \forall k \in V-b. \tag{12.482b}$$

On sait que $f(a) = b$ et que $f(a+h)$ est un élément de V et $f(a+h) = f(a) + k$ pour $k = df(a).h + \sigma_f(a,a+h)\|h\|_m$. Par substitution dans la deuxième équation on obtient

$$\begin{aligned}
&g\big(f(a+h)\big) - g\big(f(a)\big) \\
&\quad = dg_{f(a)}\Big(df_a(h) + \sigma_f(a,a+h)\|h\|_m\Big) \\
&\qquad + \sigma_g\left(f(a), f(a+h)\right)\left\|df_a(h) + \sigma_f(a,a+h)\|h\|_m\right\|_n \\
&\quad = g\circ f(a+h) - g\circ f(a) \\
&\quad = dg_{f(a)} \circ df_a(h) \\
&\qquad + \|h\|_m\Big[dg_{f(a)}\sigma_f(a,a+h) \\
&\qquad + \sigma_g\left(f(a), f(a+h)\right)\left\|df_a\frac{h}{\|h\|_m} + \sigma_f(a,a+h)\right\|_n\Big],
\end{aligned} \tag{12.483}$$

donc

$$(g\circ f)(a+h) - (g\circ f)(a) = dg_f(a)\circ df_a(h) + S(a,a+h)\|h\|_m \tag{12.484}$$

où S représente le contenu du dernier grand crochet. Il ne reste plus qu'à prouver que $S(a,a+h)$ est $o(\|h\|_m)$. En tenant compte du fait que $\sigma_f(a,a+h)$ et $\sigma_g\left(f(a),f(a+h)\right)$ sont $o(\|h\|_m)$,

$$\begin{aligned}
\lim_{h\to 0_m} \frac{S(a,a+h)}{\|h\|_m} &= \lim_{h\to 0_m} \frac{dg_{f(a)}\sigma_f(a,a+h)}{\|h\|_m} + \\
+ \lim_{h\to 0_m} &\frac{\sigma_g\left(f(a),f(a+h)\right)\left\|df_a\frac{h}{\|h\|_m} + \sigma_f(a,a+h)\right\|_n}{\|h\|_m} = 0.
\end{aligned} \tag{12.485}$$

□

En appliquant la proposition précédente point par point, nous obtenons le résultat suivant.

Proposition 12.201.
Soient U un ouvert de $\mathbb{R}^m$ et V un ouvert de $\mathbb{R}^n$. Soient $f : U \to V$ et $g : V \to \mathbb{R}^p$ deux fonctions différentiables respectivement sur U et sur V. Alors la fonction composée $g \circ f : U \to \mathbb{R}^p$ est différentiable sur U.

La matrice jacobienne de $g \circ f$ au point a est le produit matriciel des matrices jacobiennes de f et de f. Plus précisément, nous avons

$$J_{g\circ f}(a) = J_g\big(f(a)\big)J_f(a). \tag{12.486}$$

Remarquez que nous considérons la matrice jacobienne de g au point $f(a)$.

Dans la cas particulier où $m = 1$ et f est une fonction d'un intervalle I dans $\mathbb{R}^n$, dérivable au point a, on a que la fonction composée $g \circ f$ est dérivable au point a si g est différentiable et alors

$$(g \circ f)'(a) = dg\,(f(a))\,.f'(a).$$

En fait, pour les fonction d'une seule variable la dérivabilité coïncide avec la différentiabilité.

Nous avons aussi une formule importante pour la différentielle des formes bilinéaires.

Lemme 12.202.
Toute application bilinéaire

$$\begin{aligned} B\colon \mathbb{R}^m \times \mathbb{R}^n &\to \mathbb{R}^p \\ B(a_1, a_2) &= a_1 \star a_2 \end{aligned} \tag{12.487}$$

est différentiable en tout point (a_1, a_2) de $\mathbb{R}^m \times \mathbb{R}^n$, et on a

$$dB(a_1, a_2).(h_1, h_2) = h_1 \star a_2 + a_1 \star h_2.$$

Démonstration.

$$\begin{aligned} &\frac{\|B(a_1 + h_1, a_2 + h_2) - B(a_1, a_2) - (h_1 \star a_2 + a_1 \star h_2)\|_p}{\|(h_1, h_2)\|_{\mathbb{R}^m\times\mathbb{R}^n}} = \\ &= \frac{\|(a_1 + h_1) \star (a_2 + h_2) - a_1 \star a_2 - (h_1 \star a_2 + a_1 \star h_2)\|_p}{\|(h_1, h_2)\|_{\mathbb{R}^m\times\mathbb{R}^n}} = \spadesuit \end{aligned} \tag{12.488}$$

on rajoute et on enlève la quantité $(a_1 + h_1) \star a_2$ dans le numérateur, et on obtient

$$\begin{aligned} \spadesuit &= \frac{\|(a_1 + h_1) \star h_2 + h_1 \star a_2 - (h_1 \star a_2 + a_1 \star h_2)\|_p}{\|(h_1, h_2)\|_{\mathbb{R}^m\times\mathbb{R}^n}} = \\ &= \frac{\|h_1 \star h_2\|_p}{\|(h_1, h_2)\|_{\mathbb{R}^m\times\mathbb{R}^n}} \leqslant C\frac{\|h_1\|_m\|h_2\|_n}{\|(h_1, h_2)\|_{\mathbb{R}^m\times\mathbb{R}^n}} \leqslant \\ &\leqslant C\frac{\|(h_1, h_2)\|^2_{\mathbb{R}^m\times\mathbb{R}^n}}{\|(h_1, h_2)\|_{\mathbb{R}^m\times\mathbb{R}^n}} = C\|(h_1, h_2)\|_{\mathbb{R}^m\times\mathbb{R}^n}. \end{aligned} \tag{12.489}$$

Si on prend la limite de cette expression pour $(h_1, h_2) \to (0_m, 0_n)$ on obtient 0, donc la preuve est complète. À noter, que dans l'avant-dernier passage on a utilisé la continuité des applications linéaires $\texttt{proj}_m : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m$ et $\texttt{proj}_n : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ qui à chaque point (a_1, a_2) de $\mathbb{R}^m \times \mathbb{R}^n$ associent a_1 et a_2 respectivement. □

Proposition 12.203.
Soit V et W deux espaces vectoriels et $\varphi\colon V \to W$ un isomorphisme. Soit $f\colon \mathbb{C} \to V$ une application telle que $\varphi \circ f\colon \mathbb{C} \to W$ soit différentiable.

Alors f est différentiable et $df = \varphi^{-1} \circ d(\varphi \circ f)$.

Démonstration. Si T est la différentielle de $\varphi \circ f$ au point z nous avons

$$\lim_{\substack{h\to 0\\ h\in\mathbb{C}}} \frac{(\varphi\circ f)(z+h)-(\varphi\circ f)(z)+T(h)}{h}=0. \tag{12.490}$$

En appliquant φ aux deux membres, et en permutant avec la limite (parce que φ est continue),

$$\varphi \lim_{h\to 0} \frac{f(z+h)-f(z)+\varphi^{-1}T(h)}{h}=0, \tag{12.491}$$

ce qui signifie que f est différentiable et que $df=\varphi^{-1}\circ T=\varphi^{-1}\circ d(\varphi\circ f)$. □

12.19.12 Différentielle et dérivées partielles

Proposition 12.204.
Soit U un ouvert dans $\mathbb{R}^m$ et a un point dans U. Soit f une application de U dans $\mathbb{R}^n$. Si toute les dérivée partielles de f existent sur U et sont continues au point a alors f est différentiable au point a.

Démonstration. On se limite au cas $m=2$. Pour rendre les calculs plus simples on utilise ici la norme $\|\cdot\|_\infty$ dans l'espace $\mathbb{R}^2$, mais comme on a vu plus en haut, cela ne peut pas avoir des conséquences sur la différentiabilité de f. Si la différentielle de f au point a existe alors elle est définie par la formule

$$df_a(v)=\frac{\partial f}{\partial x}(a)v_1+\frac{\partial f}{\partial y}(a)v_2$$

pour tout v dans $\mathbb{R}^m$.

On commence par prouver le résultat en supposant que les dérivées partielles de f au point a sont nulles. La différentiabilité de f signifie que pour toute constante $\varepsilon>0$ il y a une constante $\delta>0$ telle que si $\|v\|_\infty\leqslant\delta$ alors

$$\frac{\|f(a_1+v_1,a_2+v_2)-f(a_1,a_2)\|_n}{\|v\|_\infty}\leqslant\varepsilon.$$

On écrit alors

$$\begin{aligned}&\|f(a_1+v_1,a_2+v_2)-f(a_1,a_2)\|_n=\\&=\|f(a_1+v_1,a_2+v_2)-f(a_1+v_1,a_2)+f(a_1+v_1,a_2)-f(a_1,a_2)\|_n\leqslant\\&\leqslant\|f(a_1+v_1,a_2+v_2)-f(a_1+v_1,a_2)\|_n+\|f(a_1+v_1,a_2)-f(a_1,a_2)\|_n.\end{aligned} \tag{12.492}$$

Comme la dérivée partielle $\partial_x f$ est nulle au point a on sait que pour toute constante $\varepsilon>0$ il y a une constante $\delta_1>0$ telle que si $|v_1|\leqslant\delta_1$ alors

$$\|f(a_1+v_1,a_2)-f(a_1,a_2)\|_n\leqslant\varepsilon|v_1|.$$

Pour l'autre terme on a, par la proposition 12.161,

$$\begin{aligned}&\|f(a_1+v_1,a_2+v_2)-f(a_1+v_1,a_2)\|_n\leqslant\\&\leqslant\sup\{\|\partial_y f(x)\|_n\,|\,x\in S\}|v_2|.\end{aligned} \tag{12.493}$$

où S est le segment d'extrémités (a_1+v_1,a_2) et (a_1+v_1,a_2+v_2). Comme la dérivée partielle $\partial_y f$ est continue et nulle au point a on sait que pour toute constante $\varepsilon>0$ il existe une constante $\delta_2>0$ telle que si $\|(u_1,u_2)\|_\infty\leqslant\delta_2$ alors $\|\partial_y f(a_1+u_1,a_2+u_2)\|_n\leqslant\varepsilon$. Si on choisit $\delta=\min\{\delta_1,\,\delta_2\}$ le segment S est contenu dans la boule de rayon δ centrée au point a et on obtient

$$\|f(a_1+v_1,a_2+v_2)-f(a_1,a_2)\|_n\leqslant\varepsilon|v_1|+\varepsilon|v_2|\leqslant 2\varepsilon\|v\|_\infty.$$

Cela prouve que f est différentiable en (a_1,a_2) et que la différentielle est nulle :

$$df_{(a_1,a_2)}=0. \tag{12.494}$$

Dans le cas général, où les dérivées partielles de f au point a ne sont pas spécialement nulles, on peut considérer la fonction [30]

$$g(x,y) = f(x,y) - \partial_1 f(a)x - \partial_2 f(a)y, \tag{12.495}$$

qui a dérivées partielles nulles au point a. La fonction g est donc différentiables. La fonction f est maintenant la somme de g et de la fonction linéaire et continue $(x,y) \mapsto \partial_1 f(a)x - \partial_2 f(a)y$. On verra dans la prochaine section que la somme de deux fonctions différentiables est une fonction différentiable. Par conséquent, la fonction f est différentiable. □

Remarque 12.205.
En dimension infinie, il n'est pas vrai que l'existence et la continuité de toutes les dérivées partielles en un point implique la différentiabilité en ce point. Pour donner un exemple, nous allons continuer l'exemple 10.20 avec la fonction 23.127 sur un espace de Hilbert.

En dimension infinie nous aurons le théorème 12.231 qui donnera quelque chose de moins fort.

Étant donné que pour tout vecteur u dans $\mathbb{R}^m$ on a $\partial_u f(a) = \nabla f(a) \cdot u$, le gradient de f nous donne la direction dans laquelle la croissance de f est maximale. Soit C une colline et soit f la fonction que a chaque point (x,y) de la Terre associe son altitude. Si nous voulons monter la colline le plus vite possible nous n'avons qu'a suivre la direction ∇f à chaque point. Elle est la projection sur le plan x-y de la direction de pente maximale. Au contraire, la direction $-\nabla f$ est la direction de croissance minimale.

La matrice jacobienne calculé au point a est la matrice associée canoniquement à l'application linéaire $df_a : \mathbb{R}^m \to \mathbb{R}^n$.

12.19.13 Plan tangent

On a dit au début de cette section que si f est une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ alors le graphe de f est une surface à deux paramètres et que l'application affine tangente au graphe de f au point $(a, f(a))$ est un plan. Maintenant on sait que ce plan est celui d'équation

$$T_a(x,y) = f(a_1,a_2) + \frac{\partial f}{\partial x}(a_1,a_2)(x-a_1) + \frac{\partial f}{\partial y}(a_1,a_2)(y-a_2). \tag{12.496}$$

Le plan tangent au graphe de f au point a est le graphe de cette fonction T_a.

Remarque 12.206.
Il existe cependant des fonctions différentiables dont les dérivées partielles ne sont pas continues. La construction d'un tel exemple est cependant délicate, et nous le ferons pas ici. Retenez cependant que si dans un exercice vous obtenez que les dérivées partielles ne sont pas continues, vous ne pouvez pas immédiatement en conclure que la fonction ne sera pas différentiable.

12.19.14 Calcul de différentielles

Remarque 12.207.
En pratique, ayant une formule pour la fonction f, on dérive –grâce aux règles usuelles de dérivation– par rapport à la variable x_i en considérant que les autres (x_j avec $j \neq i$) sont des constantes.

Exemple 12.208
Pour $f(x,y) = xy + x^2$, les dérivées partielles s'écrivent

$$\frac{\partial f}{\partial x} = y + 2x \quad \text{et} \quad \frac{\partial f}{\partial y} = x$$

△

30. Vous verrez dans la discussion à propos de la fonction (23.127) pourquoi cette fonction ne fonctionne pas dans le cas de la dimension infinie.

Des *règles de calcul* sont d'application. En particulier, quand ces opérations existent, les sommes, différences, produits, quotients et compositions d'applications différentiables sont différentiables.

Toute application linéaire est différentiable, et sa différentielle en tout point est égale à l'application elle-même. En particulier, les *projections canoniques*, c'est-à-dire les applications du type $(x, y, z) \mapsto y$, sont linéaires donc différentiables.

Exemple 12.209
Les cas suivants sont faciles :

(1) En combinant les projections canoniques avec les règles de calculs, on obtient que toute fonction polynômiale à n variables est différentiable comme application de $\mathbb{R}^n$ dans $\mathbb{R}$.

(2) Toute fonction rationnelle, du type $f(x) \stackrel{\text{def}}{=} \frac{P(x)}{Q(x)}$ où P et Q sont des polynômes, est différentiable en tout point a tel que $Q(a) \neq 0$.

(3) Pour une fonction d'une variable $f : D \subset \mathbb{R} \to \mathbb{R}$, le caractère différentiable et le caractère dérivable coïncident. De plus, on a

$$df_a(u) = f'(a)u.$$

△

12.19.15 Notes idéologiques quant au concept de plan tangent

Notons G, le graphe d'une fonction f, c'est à dire

$$G = \{(x, y, z) \in \mathbb{R}^3 \text{ tel que } z = f(x, y)\}. \tag{12.497}$$

Première affirmation : si $\gamma : \mathbb{R} \to G$ est une courbe telle que $\gamma(0) = \big(a, f(a)\big)$, alors $\gamma'(0) \in \mathbb{R}^n$ est dans le plan tangent à G au point $\big(a, f(a)\big)$.

Plus fort : tous les éléments du plan tangent sont de cette forme.

Le plan tangent à G en un point $x \in G$ est donc constitué des vecteurs vitesse de tous les chemins qui passent par x.

Prenons maintenant S, une courbe de niveau de G, c'est à dire

$$S = \{(x, y) \in \mathbb{R}^2 \text{ tel que } f(x, y) = C\}. \tag{12.498}$$

Si nous prenons un chemin dans G qui est, de plus, contraint à S, c'est à dire tel que $\gamma(t) \in S$, alors $\gamma'(0)$ sera tangent à G (ça, on le savait déjà), mais en plus, $\gamma'(0)$ sera tangent à S, ce qui est logique.

La morale est que si vous prenez un chemin qui se ballade dans n'importe quoi, alors la dérivée du chemin sera un vecteur tangent à ce n'importe quoi.

En outre, si $\gamma(t) \in S$ et $\gamma(0) = a$, alors

$$\langle \nabla f(a), \gamma'(0) \rangle = 0, \tag{12.499}$$

c'est à dire que le vecteur tangent à la courbe de niveau est perpendiculaire au gradient. Cela est intuitivement logique parce que la tangente à la courbe de niveau correspond à la direction de *moins* grande pente.

12.19.16 Gradient et recherche du plan tangent

Nous avons maintenant en main les concepts utiles pour trouver l'équation du plan tangent à une surface.

De la même manière que la tangente à une courbe était la droite de coefficient directeur donné par la dérivée, maintenant, le plan tangent à une surface est le plan dont les vecteurs directeurs sont les dérivées partielles :

La généralisation de l'équation (12.398) est

$$T_a(x) = f(a) + \sum_i \frac{\partial f}{\partial x_i}(a)(x-a)^i \tag{12.500}$$

Nous introduisons aussi souvent l'opérateur différentiel abstrait **nabla**, noté ∇ et qui est donné par le vecteur

$$\nabla = \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right). \tag{12.501}$$

Les égalités suivantes sont juste des notations, sommes toutes logiques, liées à ∇ :

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right), \tag{12.502}$$

et

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right). \tag{12.503}$$

Ce dernier est un élément de $\mathbb{R}^n$: chaque entrée est un nombre réel.

Définition 12.210.
Le vecteur gradient de f au point a est le vecteur donné par la formule (12.503).

La notation ∇ permet d'écrire la différentielle sous forme un peu plus compacte. En effet, la formule (12.417) peut être notée

$$df_a(u) = \langle \nabla f(a), u \rangle. \tag{12.504}$$

En utilisant ce produit scalaire, l'équation (12.500) peut se récrire

$$T_a(x) = f(a) + \sum_i \frac{\partial f}{\partial x_i}(a)(x-a)^i = f(a) + \langle \nabla f(a), x-a \rangle. \tag{12.505}$$

Afin d'éviter les confusions, il est parfois souhaitable de bien mettre les parenthèses et noter $(\nabla f)(a)$ au lieu de $\nabla f(a)$.

Proposition 12.211.
$\nabla f(a) \perp S_a$

$$z = f(a) + \sum_i \frac{\partial f}{\partial f}(a)(x-a)^i. \tag{12.506}$$

Cas particulier où $n = 2$:

Le plan T_a avec $a = (a_1, a_2)$ a pour équation dans $\mathbb{R}^3$:

$$z = f(a_1, a_2) + \frac{\partial f}{\partial x}(a_1, a_2)\,(x - a_1) + \frac{\partial f}{\partial y}(a_1, a_2)\,(y - a_2). \tag{12.507}$$

Définition 12.212.
Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction différentiable en un point a. Le plan tangent *au graphe de f en $(a, f(a))$ est l'ensemble des points*

$$\begin{aligned} T_a f &= \{(x, z) \in \mathbb{R}^n \times \mathbb{R} \text{ tel que } z = f(a) + df_a(x-a)\} \\ &= \{(x, z) \in \mathbb{R}^n \times \mathbb{R} \text{ tel que } z = f(a) + \langle \nabla f(a), x-a \rangle\} \end{aligned}$$

Nous avons vu que, de la même façon qu'en deux dimensions nous avions l'approximation (12.238) d'une fonction par sa tangente, en trois dimensions nous avons l'approximation suivante d'une fonction de deux variables :

$$f(x, y) \simeq f(a, b) + \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b) \tag{12.508}$$

lorsque (x, y) n'est pas trop loin de (a, b). Cela signifie que le graphe de f ressemble au graphe de la fonction $T_{(a,b)}$ donnée par

$$T_{(a,b)}(x, y) = f(a, b) + \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(x - a). \tag{12.509}$$

En notations compactes :

$$T_p(x) = f(p) + \nabla f(p) \cdot (x - p). \tag{12.510}$$

Le graphe de la fonction T_p sera le **plan tangent** au graphe de f au point p. L'équation du plan tangent sera donc

$$z - f(p) = \nabla f(p) \cdot (x - p). \tag{12.511}$$

Remarque 12.213.
Lorsque nous utilisons la notation vectorielle, la lettre «x» désigne le vecteur (x, y). Il faut être attentif. Dans un cas x est un vecteur dans l'autre c'est une composante d'un vecteur.

12.20 Calcul différentiel dans un espace vectoriel normé

Nous développons dans cette section le concept de différentielle de fonction de et vers des espaces vectoriels normés au lieu de $\mathbb{R}^n$.

12.20.1 Différentielle

Définition 12.214.
*Soit une application $f : E \to F$ entre deux espaces de Banach. Nous disons que f est **différentiable** en $a \in E$ s'il existe une application linéaire continue*[31] *$T : E \to F$ telle que*

$$\lim_{h \to 0} \frac{f(a + h) - f(a) - T(h)}{\|h\|} = 0. \tag{12.512}$$

L'application $a \mapsto T$ est la **différentielle** de f au point a et est notée df_a. L'application différentielle

$$\begin{aligned} df : E &\to \mathcal{L}(E, F) \\ a &\mapsto df_a \end{aligned} \tag{12.513}$$

est également très importante.

Définition 12.215.
*Une application $f : E \to F$ est de **classe** C^1 lorsque l'application différentielle $df : E \to \mathcal{L}(E, F)$ est continue. Voir aussi les définitions 12.251 pour les applications de classe C^k.*

Remarque 12.216.
L'application norme étant continue, le critère du théorème 6.149 est en réalité assez général. Par exemple à partir d'une application différentiable[32] $f : X \to Y$ nous pouvons considérer la fonction réelle

$$a \mapsto \|df_a\| \tag{12.514}$$

où la norme est la norme opérateur[33]. Si f est de classe C^1 alors cette application est continue et donc bornée sur un compact K de X.

31. Nous demandons bien que le candidat différentielle soit continue ; en dimension infinie ce n'est pas le cas de toutes les fonctions linéaires, comme le montre l'exemple 10.20.
32. Définition 12.214.
33. Définition 10.5.

12.20.2 (non ?) Différentiabilité des applications linéaires

Si E et F sont deux espaces vectoriels nous notons $\mathcal{L}(E,F)$ l'ensemble des applications linéaires de E vers F et $L(E,F)$ l'ensemble des applications linéaires continues de E vers F. Ces espaces seront bien entendu, sauf mention du contraire, toujours munis de la norme opérateur de la définition 10.5.

Lemme 12.217.
Soit une application linéaire f.

(1) Si f est continue, alors elle est différentiable et $df_a(u) = f(u)$ pour tout a et u.

(2) Si f n'est pas continue, alors elle n'est pas différentiable.

Démonstration. La linéarité de f donne :

$$f(a+h) - f(a) - f(h) = 0, \tag{12.515}$$

et donc prendre $T = f$ dans la définition 12.214 fait fonctionner la limite. De plus T est alors continue par hypothèse ; elle est donc bien la différentielle de f.

Supposons que f ne soit pas continue, prenons une application linéaire continue T, et calculons

$$\frac{f(a+h) - f(a) - T(h)}{\|h\|} = \frac{(f-T)(h)}{\|h\|} = (f-T)(e_h) \tag{12.516}$$

où e_h est le vecteur unitaire dans la direction de h. Vu que f n'est pas continue et que T l'est, l'application $f - T$ n'est pas continue. Elle n'est pas pas bornée par la proposition 10.18. Il existe alors un vecteur h tel que $\|(f-T)(e_h)\| > 1$ (et même plus grand que ce qu'on veut).

Donc la limite de (12.516) pour $h \to 0$ ne peut pas être nulle. □

Lemme 12.218.
Une application linéaire continue est de classe C^∞.

Démonstration. Soit $a \in E$. Étant donné que f est linéaire et continue, elle est différentiable et

$$\begin{aligned} df \colon E &\to L(E,F) \\ a &\mapsto f \end{aligned} \tag{12.517}$$

est une fonction constante et en particulier continue ; nous avons donc $f \in C^1$. Pour la différentielle seconde nous avons $d(df)_a = 0$ parce que $df(a+h) - df(a) = f - f = 0$. Toutes les différentielles suivantes sont nulles. □

12.20.3 Dérivation en chaine et formule de Leibnitz

Proposition 12.219.
Soient $f_i \colon U \to F_i$, des fonctions de classe C^r où U est ouvert dans l'espace vectoriel normé E et les F_i sont des espaces vectoriels normés. Alors l'application

$$\begin{aligned} f = f_1 \times \cdots \times f_n \colon U &\to F_1 \times \cdots \times F_n \\ x &\mapsto \big(f_1(x), \dots, f_n(x)\big) \end{aligned} \tag{12.518}$$

est de classe C^r et

$$d^r f = d^r f_1 \times \dots d^r f_n. \tag{12.519}$$

Démonstration. Soit $x \in U$ et $h \in E$. La différentiabilité des fonctions f_i donne

$$f_i(x+h) = f_i(x) + (df_i)_x(h) + \alpha_i(h) \tag{12.520}$$

avec $\lim_{h\to 0} \alpha_i(h)/\|h\| = 0$. Par conséquent

$$f(x+h) = \big(\dots, f_i(x) + (df_i)_x(h) + \alpha_i(h), \dots\big) = \big(\dots, f_i(x), \dots\big) + \big(\dots, (df_i)_x(h), \dots\big) + \big(\dots, \alpha_i(h), \dots\big). \tag{12.521}$$

Mais la définition 10.73 de la norme dans un espace produit donne

$$\lim_{h\to 0}\frac{\|\big(\alpha_1(h),\dots,\alpha_n(h)\big)\|}{\|h\|}=0, \tag{12.522}$$

ce qui nous permet de noter $\alpha(h)=\big(\alpha_1(h),\dots,\alpha_n(h)\big)$ et avoir $\lim_{h\to 0}\alpha(h)/\|h\|=0$. Avec tout ça nous avons bien

$$f(x+h)=f(x)+\big((df_1)_x(h)+\cdots+(df_n)_x(h)\big)+\alpha(h), \tag{12.523}$$

ce qui signifie que f est différentiable et

$$df_x=\big(df_1,\dots,df_n\big). \tag{12.524}$$

□

Théorème 12.220 (Différentielle de fonctions composées[141])**.**
Soient E, F et G des espaces vectoriels normés, U ouvert dans E et V ouvert dans F. Soient des applications de classe C^r ($r\geqslant 1$)

$$f\colon U\to V \tag{12.525a}$$
$$g\colon V\to G. \tag{12.525b}$$

Alors l'application $g\circ f\colon V\to G$ est de classe C^r et

$$d(g\circ f)_x=dg_{f(x)}\circ df_x. \tag{12.526}$$

Démonstration. Nous nous fixons $x\in U$. La fonction f est différentiable en $x\in U$ et g en $f(x)$, donc nous pouvons écrire

$$f(x+h)=f(x)+df_x(h)+\alpha(h) \tag{12.527}$$

et

$$g\big(f(x)+u\big)=g\big(f(x)\big)+dg_{f(x)}(u)+\beta(u) \tag{12.528}$$

où la fonction α a la propriété que

$$\lim_{h\to 0}\frac{\|\alpha(h)\|}{\|h\|}=0; \tag{12.529}$$

et la même chose pour β. La fonction composée en $x+h$ s'écrit donc

$$(g\circ f)(x+h)=g\big(f(x)+df_x(h)+\alpha(h)\big)=g\big(f(x)\big)+dg_{f(x)}\big(df_x(h)+\alpha(h)\big)+\beta\big(df_x(h)+\alpha(h)\big). \tag{12.530}$$

Nous montrons que tous les «petits» termes de cette formule peuvent être groupés. D'abord si h est proche de 0, nous avons

$$\frac{\|df_x(h)+\alpha(h)\|}{\|h\|}\leqslant\frac{\|df_x\|\|h\|}{\|h\|}+\frac{\|\alpha(h)\|}{\|h\|}. \tag{12.531}$$

Si h est petit, le second terme est arbitrairement petit, donc en prenant n'importe que $M>\|df_x\|$ nous avons

$$\frac{\|df_x(h)+\alpha(h)\|}{\|h\|}\leqslant M. \tag{12.532}$$

Par ailleurs, nous avons

$$\frac{\|\beta\big(df_x(h)+\alpha(h)\big)\|}{\|h\|}=\frac{\|\beta\big(df_x(h)+\alpha(h)\big)\|}{\|df_x(h)+\alpha(h)\|}\frac{\|df_x(h)+\alpha(h)\|}{\|h\|}\leqslant M\frac{\|\beta\big(df_x(h)+\alpha(h)\big)\|}{\|df_x(h)+\alpha(h)\|}. \tag{12.533}$$

Vu que la fraction est du type $\frac{\beta(f(h))}{f(h)}$ avec $\lim_{h\to 0}f(h)=0$, la fraction tend vers zéro lorsque $h\to 0$. En posant

$$\gamma_1(h)=\beta\big(df_x(h)+\alpha(h)\big) \tag{12.534}$$

nous avons $\lim_{h\to 0}\gamma_1(h)/\|h\|=0$.

L'autre candidat à être un petit terme dans (12.530) est traité en utilisant le lemme 10.14 :

$$\|dg_{f(x)}\big(\alpha(h)\big)\|\leqslant\|dg_{f(x)}\|\|\alpha(h)\|. \tag{12.535}$$

Donc

$$\frac{\|dg_{f(x)}\big(\alpha(h)\big)\|}{\|h\|}\leqslant\|dg_{f(x)}\|\frac{\|\alpha(h)\|}{\|h\|}, \tag{12.536}$$

ce qui nous permet de poser

$$\gamma_2(h)=dg_{f(x)}\big(\alpha(h)\big) \tag{12.537}$$

avec γ_2 qui a la même propriété que γ_1. Avec tout cela, en posant $\gamma=\gamma_1+\gamma_2$ nous récrivons

$$(g\circ f)(x+h)=g\big(f(x)\big)+dg_{f(x)}\big(df_x(h)\big)+\gamma(h) \tag{12.538}$$

avec $\lim_{h\to 0}\frac{\gamma(h)}{\|h\|}=0$. Tout cela pour dire que

$$\lim_{h\to 0}\frac{(g\circ f)(x+h)-(g\circ f)(x)-\big(dg_{f(x)}\circ df_x\big)(h)}{\|h\|}=0, \tag{12.539}$$

ce qui signifie que

$$d(g\circ f)_x=dg_{f(x)}\circ df_x. \tag{12.540}$$

Nous avons donc montré que si f et g sont différentiables, alors $g\circ f$ est différentiable avec différentielle donnée par (12.526).

Nous passons à la régularité. Nous supposons maintenant que f et g sont de classe C^r et nous considérons l'application

$$\begin{aligned}\varphi\colon L(F,G)\times L(E,F)&\to L(E,G)\\(A,B)&\mapsto A\circ B.\end{aligned} \tag{12.541}$$

Montrons que l'application φ est continue en montrant qu'elle est bornée [34]. Pour cela nous écrivons la norme opérateur

$$\|\varphi\|=\sup_{\|(A,B)\|=1}\|\varphi(A,B)\|=\sup_{\|(A,B)\|=1}\|A\circ B\|\leqslant\sup_{\|(A,B)\|=1}\|A\|\|B\|\leqslant 1. \tag{12.542}$$

Pour ce calcul nous avons utilisé le fait que la norme opérateur soit une norme algébrique (lemme 10.14) ainsi que la définition 10.73 de la norme sur un espace produit pour la dernière majoration. L'application φ est donc continue et donc C^∞ par le lemme 12.218. Nous considérons également l'application

$$\begin{aligned}\psi\colon U&\to L(F,G)\times L(E,F)\\x&\mapsto\big(dg_{f(x)},df_x\big).\end{aligned} \tag{12.543}$$

Vu que f et g sont C^1, l'application ψ est continue. Ces deux applications φ et ψ sont choisies pour avoir

$$(\varphi\circ\psi)(x)=\varphi\big(dg_{f(x)},df_x\big)=dg_{f(x)}\circ df_x, \tag{12.544}$$

c'est à dire $\varphi\circ\psi=d(g\circ f)$. Les applications φ et ψ étant continues, l'application $d(g\circ f)$ est continue, ce qui prouve que $g\circ f$ est C^1.

Si f et g sont C^r alors $dg\in C^{r-1}$ et $dg\circ f\in C^{r-1}$ où il ne faut pas se tromper : $dg\colon F\to L(F,G)$ et $f\colon U\to F$; la composée est $dg\circ f\colon x\mapsto dg_{f(x)}\in L(F,G)$.

Pour la récurrence nous supposons que $f,g\in C^{r-1}$ implique $g\circ f\in C^{r-1}$ pour un certain $r\geqslant 2$ (parce que nous venons de prouver cela avec $r=1$ et $r=2$). Soient $f,g\in C^r$ et montrons que $g\circ f\in C^r$. Par la proposition 12.219 nous avons

$$\psi=dg\circ f\times df\in C^{r-1}, \tag{12.545}$$

et donc $d(g\circ f)=\varphi\circ\psi\in C^{r-1}$, ce qui signifie que $g\circ f\in C^r$. □

34. Proposition 10.18.

On fixe maintenant une définition largement utilisée dans la suite.

Définition 12.221.
*Soient U et V, deux ouverts d'un espace vectoriel normé. Une application f de U dans V est un **difféomorphisme** si elle est bijective, différentiable et dont l'inverse $f^{-1} : V \to U$ est aussi différentiable.*

Remarque 12.222.
Il n'est pas possible d'avoir une application inversible d'un ouvert de $\mathbb{R}^m$ vers un ouvert de $\mathbb{R}^n$ si $m \neq n$. Il n'y a donc pas de notion de difféomorphismes entre ouverts de dimensions différentes.

Lemme 12.223.
Si $f : U \to V$ est une difféomorphisme[35] *alors pour tout $a \in U$, l'application df_a est inversible et*

$$(df_a)^{-1} = (df^{-1})_{f(a)}. \tag{12.546}$$

Démonstration. Il suffit d'apercevoir qu'en vertu de la règle de différentiation en chaîne 12.526,

$$(df_a)(df^{-1})_{f(a)} = f(f \circ f^{-1})_{f(a)} = \mathrm{Id}\,. \tag{12.547}$$

□

Lemme 12.224 (Leibnitz pour les formes bilinéaires[141])**.**
Si $B : E \times F \to G$ est bilinéaire et continue, elle est C^∞ et

$$dB_{(x,y)}(u, v) = B(x, v) + B(u, y). \tag{12.548}$$

Démonstration. D'abord le membre de droite de (12.548) est une application linéaire et continue, donc c'est un bon candidat à être différentielle. Nous allons prouver que ça l'est, ce qui prouvera la différentiabilité de B. Avec ce candidat, le numérateur de la définition (12.512) s'écrit dans notre cas

$$B\big((x, y) + (u, v)\big) - B(x, y) - B(x, v) - B(u, y) = B(u, v). \tag{12.549}$$

Il reste à voir que

$$\lim_{(u,v)\to(0,0)} \frac{B(u, v)}{\|(u, v)\|} = 0 \tag{12.550}$$

Par l'équation (10.189) nous avons

$$\frac{\|B(u, v)\|}{\|(u, v)\|} \leqslant \frac{\|B\|\|u\|\|v\|}{\|u\|} = \|B\|\|v\| \tag{12.551}$$

parce que $\|(u, v)\| \geqslant \|u\|$. À partir de là il est maintenant clair que

$$\lim_{(u,v)\to(0,0)} \frac{\|B(u, v)\|}{\|(u, v)\|} = 0, \tag{12.552}$$

ce qu'il fallait. □

Proposition 12.225 (Règle de Leibnitz[141])**.**
Soient E, F_1, F_2 des espaces vectoriels normés, U ouvert dans E et des applications de classe C^r $(r \geqslant 1)$

$$f_1 : U \to F_1 \tag{12.553a}$$
$$f_2 : U \to F_2 \tag{12.553b}$$
$$\tag{12.553c}$$

35. Définition 12.221

et $B \in L(F_1 \times F_2, G)$. Alors l'application

$$\begin{aligned} \varphi\colon U &\to G \\ x &\mapsto B\big(f_1(x), f_2(x)\big) \end{aligned} \tag{12.554}$$

est de classe C^r et

$$d\varphi_x(u) = \varphi\big((df_1)_x(u), f_2(x)\big) + \varphi\big(f_1(x), (df_2)_x(u)\big). \tag{12.555}$$

Démonstration. Par hypothèse B est continue (c'est la définition de l'espace L), et donc C^∞ par le lemme 12.224. Par ailleurs la fonction $f_1 \times f_2$ est de classe C^r parce que f_1 et f_2 le sont et parce que la proposition 12.219 le dit. L'application composée $B \circ (f_1 \times f_2)$ est donc également de classe C^r par le théorème 12.220.

Il ne nous reste donc qu'à prouver la formule 12.555. En utilisant la différentielle du produit cartésien [36] nous avons

$$f\big(B \circ (f_1 \times f_2)\big)_x(h) = dB_{(f_1 \times f_2)(x)}\big((df_1)_x(h), (df_2)_x(h)\big). \tag{12.556}$$

Nous développons cela en utilisant le lemme 12.224 :

$$d\big(B \circ (f_1 \times f_2)\big)_x(h) = dB_{\left(f_1(x), f_2(x)\right)}\big((df_1)_x(h), (df_2)_x(h)\big) \tag{12.557a}$$

$$= B\big(f_1(x), (df_2)_x(h)\big) + B\big((df_1)_x(h), f_2(x)\big), \tag{12.557b}$$

comme souhaité. □

12.20.4 Différentielle partielle

Définition 12.226 (Différentielle partielle)**.**
*Soient E, F et G des espaces vectoriels normés et une fonction $f\colon E \times F \to G$. Nous définissons sa **différentielle partielle** sur l'espace E par*

$$\begin{aligned} d_1 f_{(x_0,y_0)}\colon E &\to G \\ u &\mapsto \frac{d}{dt}\Big[f(x_0 + tu, y_0\Big]_{t=0}. \end{aligned} \tag{12.558}$$

La différentielle d_2 se définit de la même façon.

Proposition 12.227 ([141])**.**
Soient E_1, E_2 et F des espaces vectoriels normés, soit un ouvert $U \subset E_1 \times E_2$ et une fonction $f\colon U \to F$.

(1) Si f est différentiable alors les différentielles partielles existent et

$$d_1 f_{(x_0,y_0)}(u) = df_{(x_0,y_0)}(u, 0) \tag{12.559a}$$

$$d_2 f_{(x_0,y_0)}(v) = df_{(x_0,y_0)}(0, v) \tag{12.559b}$$

où $u \in E_1$ et $v \in E_2$.

(2) Si f est différentiable alors

$$df_{(x_0,y_0)}(u, v) = d_1 f_{(x,y_0)}(u) + d_2 f_{(x_0,y_0)}(v). \tag{12.560}$$

Démonstration. Nous posons $\alpha = (x_0, y_0) \in U$ et

$$\begin{aligned} j_\alpha^{(1)}\colon E_1 &\to E_1 \times E_2 \\ x &\mapsto (x, y_0). \end{aligned} \tag{12.561}$$

36. Proposition 12.219.

C'est une fonction de classe C^∞ et

$$(dj_\alpha^{(1)})_{x_0}(u) = \frac{d}{dt}\Big[j_\alpha^{(1)}(x_0+tu)\Big]_{t=0} = \frac{d}{dt}\Big[(x_0+tu,y_0)\Big]_{t=0} = (u,0). \tag{12.562}$$

D'autre part

$$\begin{aligned}(d_1f)_\alpha(u) &= \frac{d}{dt}\Big[f(x_0+tu,y_0)\Big]_{t=0} && (12.563a)\\ &= \frac{d}{dt}\Big[(f\circ j_\alpha^{(1)})(x_0+tu)\Big]_{t=0} && (12.563b)\\ &= \big(d(f\circ j_\alpha^{(1)})\big)_{x_0}(u). && (12.563c)\end{aligned}$$

À ce moment nous utilisons la règle des différentielles composées 12.220 pour dire que

$$(d_1f)_\alpha(u) = df_{j_\alpha^{(1)}(x_0)}\circ(dj_\alpha^{(1)})_{x_0}(u) = df_\alpha(u,0). \tag{12.564}$$

Voila qui prouve déjà le point (1).

Pour la suite nous considérons les fonctions

$$\begin{aligned}P_1(x,y)&=x, \qquad & J_1(u)&=(u,0),\\ P_2(x,y)&=y, \qquad & J_2(v)&=(0,v)\end{aligned} \tag{12.565}$$

et nous avons l'égalité évidente

$$J_1\circ P_1 + J_2\circ P_2 = \mathbb{1} \tag{12.566}$$

sur $E_1\times E_2$. En appliquant df_α à cette dernière égalité, en appliquant à (u,v) et en utilisant la linéarité de df_α nous trouvons

$$\begin{aligned}df_\alpha(u,v) &= df_\alpha\big((J_1\circ P_1)(u,v)\big) + df_\alpha\big((J_2\circ P_2)(u,v)\big) && (12.567a)\\ &= df_\alpha(u,0)+df_\alpha(0,v) && (12.567b)\\ &= (d_1f)_\alpha(u)+(d_2f)_\alpha(v) && (12.567c)\end{aligned}$$

où nous avons utilisé le point (1) pour la dernière égalité. □

12.20.5 Formule des accroissements finis

Proposition 12.228.
Soient $a<b$ dans $\mathbb{R}$ et deux fonctions

$$\begin{aligned}f&\colon [a,b]\to E && (12.568a)\\ g&\colon [a,b]\to\mathbb{R} && (12.568b)\end{aligned}$$

continues sur $[a,b]$ et dérivables sur $]a,b[$. Si pour tout $t\in]a,b[$ nous avons $\|f'(t)\|\leqslant g'(t)$ alors

$$\|f(b)-f(a)\|\leqslant g(b)-g(a). \tag{12.569}$$

Démonstration. Soit $\epsilon>0$ et la fonction

$$\begin{aligned}\varphi_\epsilon\colon [a,b]&\to\mathbb{R}\\ t&\mapsto \|f(t)-f(a)\|-g(t)-\epsilon t.\end{aligned} \tag{12.570}$$

Cela est une fonction continue réelle à variable réelle. En particulier pour tout $u\in]a,b[$ la fonction φ_ϵ est continue sur le compact $[u,b]$ et donc y atteint son minimum en un certain point $c\in[u,b]$; c'est le bon vieux théorème de Weierstrass 6.149. Nous commençons par montrer que pour tout u, ledit minimum ne peut être que b. Pour cela nous allons montrer que si $t\in[u,b[$, alors $\varphi_\epsilon(s)<\varphi_\epsilon(t)$ pour un certain $s>t$. Par continuité si s est proche de t nous avons

$$\left\|\frac{f(s)-f(t)}{s-t}\right\| - \frac{\epsilon}{2} < \|f'(t)\| < g'(t)+\frac{\epsilon}{2} = \frac{g(s)-g(t)}{s-t}+\frac{\epsilon}{2}. \tag{12.571}$$

Ces inégalités proviennent de la limite

$$\lim_{s\to t}\frac{f(s)-f(t)}{s-t}=f'(t),\tag{12.572}$$

donc si s et t sont proches,

$$\left\|\frac{f(s)-f(t)}{s-t}-f'(t)\right\|\tag{12.573}$$

est petit. Si $s>t$ nous pouvons oublier des valeurs absolues et transformer l'inégalité en

$$\|f(s)-f(t)\|<g(s)-g(t)+\epsilon(s-t).\tag{12.574}$$

Utilisant cela et l'inégalité triangulaire,

$$\begin{aligned}\varphi_\epsilon(s)&\leqslant\|f(s)-f(t)\|+\|f(t)-f(a)\|-g(s)-\epsilon s&(12.575a)\\&\leqslant g(s)-g(t)+\epsilon s-\epsilon t+\|f(t)-f(a)\|-g(s)-\epsilon s&(12.575b)\\&=\varphi_\epsilon(t).&(12.575c)\end{aligned}$$

Donc nous avons bien $\varphi_\epsilon(s)<\varphi_\epsilon(t)$ avec l'inégalité stricte. Par conséquent pour tout $u\in]a,b[$ nous avons $\varphi_\epsilon(b)<\varphi_\epsilon(u)$ et en prenant la limite $u\to a$ nous avons

$$\varphi_\epsilon(b)\leqslant\varphi_\epsilon(a).\tag{12.576}$$

Cette inégalité donne immédiatement

$$\|f(b)-f(a)\|\leqslant g(b)-g(a)+\epsilon(b-a)\tag{12.577}$$

pour tout $\epsilon>0$ et donc

$$\|f(b)-f(a)\|\leqslant g(b)-g(a).\tag{12.578}$$

□

Théorème 12.229 (Théorème des accroissements finis)**.**
Soient E et F des espaces vectoriels normés, U ouvert dans E et une application différentiable $f\colon U\to F$. Pour tout segment $[a,b]\subset U$ nous avons

$$\|f(b)-f(a)\|\leqslant\left(\sup_{x\in[a,b]}\|df_x\|\right)\|b-a\|.\tag{12.579}$$

Démonstration. Nous prenons les applications

$$\begin{aligned}k\colon[0,1]&\to E\\t&\mapsto f\big((1-t)a+tb\big)\end{aligned}\tag{12.580}$$

et

$$\begin{aligned}g\colon[0,1]&\to\mathbb{R}\\t&\mapsto t\sup_{x\in[a,b]}\|df_x\|\|b-a\|.\end{aligned}\tag{12.581}$$

Pour tout t nous avons $g'(t)=M\|b-a\|$ où il n'est besoin de dire ce qu'est M. D'un autre côté nous avons aussi

$$\begin{aligned}k'(t)&=\lim_{\epsilon\to0}\frac{f\big((1-t-\epsilon)a+(t+\epsilon)b\big)-f\big((1-t)a+tb\big)}{\epsilon}\\&=\frac{d}{d\epsilon}\Big[f\big((1-t)a+tb+\epsilon(b-a)\big)\Big]_{\epsilon=0}\\&=df_{(1-t)a+tb}(b-a)\end{aligned}\tag{12.582}$$

où nous avons utilisé l'hypothèse de différentiabilité de f sur $[a,b]$ et donc en $(1-t)a+tb$. Nous avons donc

$$\|k'(t)\| \leqslant \|b-a\| \|df_{(1-t)a+tb}\| \leqslant M\|b-a\| = g'(t) \tag{12.583}$$

La proposition 12.228 est donc utilisable et

$$\|k(1)-k(0)\| = g(1)-g(0), \tag{12.584}$$

c'est à dire

$$\|f(b)-f(a)\| = M\|b-a\| \tag{12.585}$$

comme il se doit. □

Proposition 12.230.
Soient E et F des espaces vectoriels normés, U ouvert dans E et une application $f\colon U \to F$. Soient $a,b \in U$ tels que $[a,b] \subset U$. Nous posons $u = (b-a)/\|b-a\|$ et nous supposons que pour tout $x \in [a,b]$, la dérivée directionnelle

$$\frac{\partial f}{\partial u}(x) = \frac{d}{dt}\Big[f(x+tu)\Big]_{t=0} \tag{12.586}$$

existe. Nous supposons de plus que $\frac{\partial f}{\partial u}(x)$ est continue en $x=a$. Alors

$$\|f(b)-f(a)\| \leqslant \left(\sup_{x\in[a,b]} \|\frac{\partial f}{\partial u}(x)\|\right) \|b-a\|. \tag{12.587}$$

Démonstration. Nous posons évidemment

$$M = \sup_{x\in[a,b]} \|\frac{\partial f}{\partial u}(x)\| \tag{12.588}$$

et nous considérons les fonctions

$$k(t) = f\big((1-t)a+tb\big) \tag{12.589}$$

et

$$g(t) = tM\|b-a\|. \tag{12.590}$$

Pour alléger les notations nous posons $x = (1-t)a+tb$ et nous calculons avec un petit changement de variables dans la limite :

$$k'(t) = \frac{d}{d\epsilon}\Big[f\big(x+\epsilon(b-a)\big)\Big]_{\epsilon=0} = \|b-a\|\frac{d}{d\epsilon}\Big[f\big(x+\frac{\epsilon}{\|b-a\|}(b-a)\big)\Big]_{\epsilon=0} = \|b-a\|\frac{\partial f}{\partial u}(x), \tag{12.591}$$

et donc encore une fois nous avons

$$\|k'(t)\| \leqslant g'(t), \tag{12.592}$$

ce qui donne

$$\|k(1)-k(0)\| = g(1)-g(0), \tag{12.593}$$

c'est à dire

$$\|f(b)-f(a)\| \leqslant \sup_{x\in[a,b]} \|\frac{\partial f}{\partial u}(x)\|\|b-a\|. \tag{12.594}$$

□

Théorème 12.231.
Soient E,V deux espaces vectoriels normés, une application $f\colon E \to V$, un point $a \in E$ tel que pour tout $u \in E$, la dérivée

$$\frac{d}{dt}\Big[f(x+tu)\Big]_{t=0} \tag{12.595}$$

existe pour tout $x \in B(a,r)$ et est continue (par rapport à x) en $x = a$. Nous supposons de plus que[37]

$$\frac{\partial f}{\partial u}(a) = 0 \tag{12.596}$$

pour tout $u \in E$. Alors f est différentiable en a et

$$df_a = 0 \tag{12.597}$$

Démonstration. Soit $\epsilon > 0$. Pourvu que $\|h\|$ soit assez petit pour que $a+h \in B(a,r)$, la proposition 12.230 nous donne

$$\|f(a+h) - f(a)\| \leqslant \sup_{x\in[a,a+h]} \|\frac{\partial f}{\partial u}(x)\||h| \tag{12.598}$$

où $u = h/\|h\|$. Par continuité de $\partial_u f(x)$ en $x = a$ et par le fait que cela vaut 0 en $x = a$, il existe un $\delta > 0$ tel que si $\|h\| < \delta$ alors

$$\|\frac{\partial f}{\partial u}(a+h)\| \leqslant \epsilon. \tag{12.599}$$

Pour de tels h nous avons

$$\|f(a+h) - f(a)\| \leqslant \epsilon\|h\|, \tag{12.600}$$

ce qui prouve que l'application linéaire $T(u) = 0$ convient parfaitement pour faire fonctionner la définition 12.214. □

12.20.6 L'inverse, sa différentielle

Si E est un espace de Banach, nous sommes intéressé à l'espace $\mathrm{GL}(E)$ des endomorphismes inversibles de E sur E. Cet ensemble est métrique par la formule usuelle

$$\|T\| = \sup_{\|x\|=1} \|T(x)\|_E. \tag{12.601}$$

Théorème 12.232 (Inverse dans $\mathrm{GL}(E)$[142, 141])**.**
Soient E et F des espaces vectoriels normés.

(1) L'ensemble $\mathrm{GL}(E)$ est ouvert dans $\mathrm{End}(E)$.

(2) L'application inverse

$$\begin{aligned} i\colon \mathrm{GL}(E,F) &\to \mathrm{GL}(F,E) \\ u &\mapsto u^{-1} \end{aligned} \tag{12.602}$$

est de classe C^∞ et

$$di_{u_0}(h) = -u_0^{-1} \circ h \circ u_0^{-1} \tag{12.603}$$

pour tout $h \in \mathrm{End}(E)$

Démonstration. Nous supposons que $\mathrm{GL}(E,F)$ n'est pas vide, sinon ce n'est pas du jeu.

Cas de dimension finie Si la dimension de E et F est finie, elles doivent être égales, sinon il n'y a pas de fonctions inversibles $E \to F$. L'ensemble $\mathrm{GL}(E,F)$ est donc naturellement $\mathrm{GL}(n,\mathbb{R})$. Un élément de $\mathbb{M}(n,\mathbb{R})$ est dans $\mathrm{GL}(n,\mathbb{R})$ si et seulement si son déterminant est non nul. Le déterminant étant une fonction continue (polynomiale) en les entrées de la matrice, l'ensemble $\mathrm{GL}(n,\mathbb{R})$ est ouvert dans $\mathbb{M}(n,\mathbb{R})$.

Même idée pour la régularité de la fonction $i\colon \mathrm{GL}(n,\mathbb{R}) \to \mathrm{GL}(n,\mathbb{R})$, $X \mapsto X^{-1}$. Les entrées de X^{-1} sont les cofacteurs de X divisé par $\det(X)$, et donc des polynômes en les entrées de X divisés par un polynôme qui ne s'annule pas sur $\mathrm{GL}(n,\mathbb{R})$, et donc sur un ouvert autour de X et de X^{-1}. Bref, tout est C^∞.

Le reste de la preuve parle de la dimension infinie.

37. Je ne suis pas certain que cette hypothèse soit nécessaire, voir la question (5) de la page 83.

Ouvert autour de l'identité Nous commençons par prouver que $B(\mathbb{1},1)\subset \mathrm{GL}(E)$. Pour cela il suffit de remarquer que si $\|u\|<1$ alors le lemme 10.52 nous donne un inverse de $(1+u)$ en la personne de $\sum_{k=0}^{\infty}(-u)^k$.

Ouvert en général Soit maintenant $u_0\in\mathrm{GL}(E)$. Si $\|u\|<\frac{1}{\|u_0^{-1}\|}$ alors $\|u_0^{-1}u\|<1$, ce qui signifie que

$$\mathbb{1}+u_0^{-1}u \tag{12.604}$$

est inversible. Mais $u_0+u=u_0(\mathbb{1}+u_0^{-1}u)$, donc $u_0+u\in\mathrm{GL}(E)$ ce qui signifie que

$$B\left(u_0,\frac{1}{\|u_0^{-1}\|}\right)\subset\mathrm{GL}(E). \tag{12.605}$$

Différentielle en l'identité Nous commençons par prouver que $di_{\mathbb{1}}(u)=-u$. Pour cela nous posons

$$\alpha(h)=\sum_{k=2}^{\infty}(-1)^k h^k \tag{12.606}$$

et nous calculons

$$di_{\mathbb{1}}(u)=\frac{d}{dt}\Big[i(\mathbb{1}+tu)\Big]_{t=0}=\frac{d}{dt}\Big[\mathbb{1}-tu+\alpha(tu)\Big]_{t=0}. \tag{12.607}$$

Il suffit de prouver que $\frac{d}{dt}\Big[\alpha(tu)\Big]_{t=0}=0$ pour conclure que $di_{\mathbb{1}}(u)=-u$. Pour cela, nous remarquons que $\alpha(0)=0$ et donc que

$$\frac{d}{dt}\Big[\alpha(tu)\Big]_{t=0}=\lim_{t\to0}\frac{\alpha(tu)-\alpha(0)}{t} \tag{12.608a}$$

$$=\lim_{t\to0}\sum_{k=2}^{\infty}(-1)^k\frac{(tu)^k}{t} \tag{12.608b}$$

$$=-\lim_{t\to0}u\sum_{k=1}^{\infty}(-1)^k t^k u^k. \tag{12.608c}$$

La norme de ce qui est dans la limite est majorée par

$$\|u\|\sum_{k=1}^{\infty}\|tu\|^k=\|u\|\left(\frac{1}{1-\|tu\|}-1\right), \tag{12.609}$$

et cela tend vers zéro lorsque $t\to\infty$. Nous avons utilisé la somme 10.129 de la série géométrique. Nous avons bien prouvé que $di_{\mathbb{1}}(u)=-u$.

Différentielle en général Soit maintenant $u_0\in\mathrm{GL}(E)$ et $h\in\mathrm{End}(E)$ tel que $u_0+h\in\mathrm{GL}(E)$; par le premier point, il suffit de prendre $\|h\|$ suffisamment petit. Vu que $u_0+h=u_0(\mathbb{1}+u_0^{-1}h)$ nous avons

$$(u_0+h)^{-1}=(\mathbb{1}+u_0^{-1}h)^{-1}u_0^{-1}. \tag{12.610}$$

Nous pouvons donc calculer

$$(u_0+h)^{-1}=\big(\mathbb{1}-u_0^{-1}h+\alpha(u_0^{-1}h)\big)u_0^{-1}=u_0^{-1}-u_0^{-1}hu_0^{-1}+\alpha(u_0^{-1}h)u_0^{-1}, \tag{12.611}$$

et ensuite

$$di_{u_0}(h)=\frac{d}{dt}\Big[i(u_0+th)\Big]_{t=0}=\frac{d}{dt}\Big[u_0^{-1}-tu_0^{-1}hu_0^{-1}+\alpha(tu_0^{-1}h)u_0^{-1}\Big]_{t=0}, \tag{12.612}$$

mais nous avons déjà vu que

$$\frac{d}{dt}\Big[\alpha(th)\Big]_{t=0}=0, \tag{12.613}$$

donc

$$di_{u_0}(h)=-u_0^{-1}hu_0^{-1} \tag{12.614}$$

Cela donne la différentielle de l'application inverse.

Continuité de l'inverse L'application i est continue parce que différentiable.

L'inverse est C^∞ Nous allons écrire la fonction inverse comme une composée. Soient les applications

$$\begin{aligned} B\colon L(F,E)\times L(F,E) &\to L\big(L(E,F),L(F,E)\big)\\ B(\psi_1,\psi_2)(A) &= -\psi_1\circ A\circ\psi_2 \end{aligned} \tag{12.615}$$

et

$$\begin{aligned} \Delta\colon L(F,E) &\to L(F,E)\times L(F,E)\\ \varphi &\mapsto (\varphi,\varphi) \end{aligned} \tag{12.616}$$

Nous avons alors

$$di = B\circ\Delta\circ i. \tag{12.617}$$

L'application Δ est de classe C^∞. Nous devons voir que B l'est aussi. Pour le voir nous commençons par prouver qu'elle est bornée :

$$\begin{aligned} \|B\| &= \sup_{\|\psi_1\|,\|\psi_2\|=1} \|B(\psi_1,\psi_2)\|_{\mathcal{L}\left(L(E,F),L(F,E)\right)}\\ &= \sup_{\|\psi_1\|,\|\psi_2\|=1}\sup_{\|A\|=1} \|\psi_1\circ A\circ\psi_2\|_{L(F,E)}\\ &\leqslant \sup_{\|\psi_1\|,\|\psi_2\|=1}\sup_{\|A\|=1} \|\psi_1\|\|A\|\|\psi_2\|\\ &\leqslant 1. \end{aligned} \tag{12.618}$$

Donc B est bien bornée et par conséquent continue. Une application bilinéaire continue est C^∞ par le lemme 12.224. La décomposition $di = B\circ\Delta\circ i$ nous donne donc que $i\in C^\infty$ dès que i est continue, ce que nous avions déjà montré.

□

12.20.7 Projection orthogonale

Le théorème suivant n'est pas indispensablissime parce qu'il est le même que le théorème de la projection sur les espaces de Hilbert [38]. Cependant la partie existence est plus simple en se limitant au cas de dimension finie.

Théorème 12.233 (Théorème de la projection)**.**
Soit E un espace vectoriel réel ou complexe de dimension finie, $x\in E$, et C un sous ensemble fermé convexe de E.

(1) Les deux conditions suivantes sur $y\in E$ sont équivalentes :
  (a) $\|x-y\| = \inf\{\|x-z\| \text{ tel que } z\in C\}$,
  (b) pour tout $z\in C$, $\mathrm{Re}\langle x-y, z-y\rangle \leqslant 0$.

(2) Il existe un unique $y\in E$, noté $y = \texttt{proj}_C(x)$ vérifiant ces conditions.

Démonstration. Nous commençons par prouver l'existence et l'unicité d'un élément dans C vérifiant la première condition. Ensuite nous verrons l'équivalence.

Existence Soit $z_0\in C$ et $r = \|x-z_0\|$. La boule fermée $\overline{B(x,r)}$ est compacte [39] et intersecte C. Vu que C est fermé, l'ensemble $C' = C\cap\overline{B(x,r)}$ est compacte. Tous les points qui minimisent la distance entre x et C sont dans C' ; la fonction

$$\begin{aligned} C' &\to \mathbb{R}\\ z &\mapsto d(x,z) \end{aligned} \tag{12.619}$$

est continue sur un compact et donc a un minimum qu'elle atteint [40]. Un point P réalisant ce minimum prouve l'existence d'un point vérifiant la première condition.

38. Théorème 23.6
39. C'est ceci qui ne marche plus en dimension infinie.
40. Théorème 6.126.

Unicité Soient y_1 et y_2, deux éléments de C minimisant la distance avec x, et soit d ce minimum. Nous avons par l'identité du parallélogramme (9.101) que

$$\|y_1 - y_2\|^2 = -4\left\|\frac{y_1 + y_2 - x}{2}\right\|^2 + 2\|y_1 - x\|^2 + 2\|y_2 - x\|^2 \leqslant -4d + 2d + 2d = 0. \quad (12.620)$$

Par conséquent $y_1 = y_2$.

(1)a⇒ (1)b Soit $z \in C$ et $t \in]0,1[$; nous notons $P = \texttt{proj}_C x$. Par convexité le point $z = ty + (1-t)P$ est dans C, et par conséquent,

$$\|x - P\|^2 \leqslant \|x - tz - (1-t)P\|^2 = \|(x-P) - t(z-P)\|^2. \quad (12.621)$$

Nous sommes dans un cas $\|a\|^2 \leqslant |a-b|^2$, qui implique $2\,\mathrm{Re}\langle a, b\rangle \leqslant \|b\|^2$. Dans notre cas,

$$2\,\mathrm{Re}\langle x - P, t(z-P)\rangle \leqslant t^2\|z-P\|^2. \quad (12.622)$$

En divisant par t et en faisant $t \to 0$ nous trouvons l'inégalité demandée :

$$2\,\mathrm{Re}\langle x - P, z - P\rangle \leqslant 0. \quad (12.623)$$

(1)b⇒ (1)a Soit un point $P \in C$ vérifiant

$$\mathrm{Re}\langle x - P, z - P\rangle \leqslant 0 \quad (12.624)$$

pour tout $z \in C$. Alors en notant $a = x - P$ et $b = P - z$,

$$\begin{aligned}\|x - z\|^2 = \|x - P + P - z\|^2 &= \|a + b\|^2 \\ &= \|a\|^2 + \|b\|^2 + 2\,\mathrm{Re}\langle a, b\rangle \\ &= \|a\|^2 + \|b\|^2 - 2\,\mathrm{Re}\langle x - P, z - P\rangle \\ &\geqslant \|b\|^2,\end{aligned} \quad (12.625)$$

ce qu'il fallait.

□

12.21 Jacobienne

12.21.1 Rappels et définitions

Dans cette section nous considérons des fonctions $f : D \to \mathbb{R}^m$ où $D \subset \mathbb{R}^n$, et un point $a \in \mathrm{Int}\, D$ où f est différentiable.

Remarque 12.234.
La définition de continuité (resp. différentiabilité) pour une fonction à valeurs vectorielles est celle introduite précédemment, et on remarque que pour avoir la continuité (resp. différentiabilité) de f en un point, il faut et il suffit de chacune des composantes de $f = (f_1, \dots, f_m)$, vues séparément comme fonctions à n variables et à valeurs réelles, soit continue (resp. différentiable) en ce point.

Définition 12.235.
*La **jacobienne** de f en a est la matrice de l'application linéaire donnée par la différentielle. Elle a de nombreuses notations*

$$J_f(a) = \frac{\partial(f_1, \dots, f_m)}{\partial x_1, \dots, x_m} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix} \quad (12.626)$$

*Autrement dit, c'est la matrice composée de l'ensemble des dérivées partielles de f. Le **jacobien** de f au point a est le déterminant de cette matrice.*

*Si $m = 1$, cette matrice ne contient qu'une ligne ; c'est donc un vecteur appelé le **gradient** de f au point a et noté $\nabla f(a)$.*

Remarque 12.236. (1) Si la fonction est supposée différentiable, calculer la jacobienne revient à connaître la différentielle. En effet, par linéarité de la différentielle et par définition des dérivées partielles, nous avons

$$df_a(u) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

où $u = (u_1, \dots, u_n)$ et où le membre de droite est un produit matriciel

(2) Remarquons que la jacobienne peut exister en un point donné sans que la fonction soit différentiable en ce point !

12.22 Fonctions de classe $\mathcal{C}^1$

Soit f une fonction différentiable de U, ouvert de $\mathbb{R}^m$, dans $\mathbb{R}^n$. L'application différentielle de f est une application de $\mathbb{R}^m$ dans $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$

$$\begin{array}{rccc} df: & \mathbb{R}^m & \to & \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n) \\ & a & \mapsto & df_a. \end{array} \tag{12.627}$$

Nous savons que $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ est un espace vectoriel normé avec la définition 7.19. Si T est un élément dans $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ alors la norme de T est définie par

$$\|T\|_{\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)} = \sup_{x \in \mathbb{R}^m} \frac{\|T(x)\|_n}{\|x\|_m} = \sup_{\substack{x \in \mathbb{R}^m \\ \|x\|_m \leqslant 1}} \|T(x)\|_n.$$

Lorsqu'il existe un $M > 0$ tel que $\|df(a)\|_{\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)} < M$ pour tout a dans U, nous disons que la différentielle de f est **bornée** sur U.

Définition 12.237.
*La fonction f est dite **de classe** $\mathcal{C}^1$ de $U \subset \mathbb{R}^m$ dans $\mathbb{R}^n$ si son application différentielle df est continue de $\mathbb{R}^m$ dans $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$. Nous écrivons $f \in \mathcal{C}^1(U, \mathbb{R}^n)$.*

Proposition 12.238.
Une fonction $f: U \to \mathbb{R}^n$ où U est ouvert dans $\mathbb{R}^m$ est de classe C^1 si et seulement si les dérivées partielles de f existent et sont continues.

Démonstration. Supposons que les dérivées partielles de f existent et sont continues. Nous savons alors déjà par la proposition 12.204 que la fonction f est différentiable et qu'elle s'exprime sous la forme

$$df_a(h) = \sum_{i=1}^m \partial_i f(a) h_i, \qquad \forall a \in U, \forall h \in \mathbb{R}^m.$$

Pour montrer que df est continue, nous devons montrer que la quantité $\|df(x) - df(a)\|_{\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)}$ peut être rendue arbitrairement petite si $\|x - a\|_m$ est rendu petit. Nous avons

$$\begin{aligned} \|df_x - df_a\|_{\mathcal{L}} &= \sup_{\|h\|=1} \|df_x(h) - df_a(h)\| \\ &= \sup_{\|h\|_m=1} \left\| \sum_{i=1}^m (\partial_i f(x) - \partial_i f(a)) h_i \right\|_n \leqslant \\ &\leqslant \sup_{\|h\|_m=1} \sum_{i=1}^m \|(\partial_i f(x) - \partial_i f(a))\|_n |h_i| \leqslant \\ &\leqslant \sup_{\|h\|_m=1} \|h\|_\infty \sum_{i=1}^m \|(\partial_i f(x) - \partial_i f(a))\|_n \\ &\leqslant \sum_{i=1}^m \|\partial_i f(x) - \partial_i f(a)\|. \end{aligned} \tag{12.628}$$

Dans ce calcul, nous avons utilisé le fait que si $\|h\|_m \leqslant 1$, alors $\|h\|_\infty \leqslant 1$. Étant donné la continuité de $\partial_i f$, la dernière ligne peut être rendue arbitrairement petite lorsque x est proche e a.

Supposons maintenant que f soit dans $\mathcal{C}^1(U, \mathbb{R}^n)$. Alors

$$\|\partial_i f(x) - \partial_i f(a)\|_n = \|df(x).e_i - df(a).e_i\|_n \leqslant \|df(x) - df(a)\|_{\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)},$$

la continuité de df implique donc celle de $\partial_i f$ pour tout i dans $\{1, \dots, m\}$. □

Proposition 12.239.
Soient U un ouvert de $\mathbb{R}^m$ et V un ouvert de $\mathbb{R}^n$. Soient $f : U \to V$ dans $\mathcal{C}^1(U, V)$ et $g : V \to \mathbb{R}^p$ dans $\mathcal{C}^1(V, \mathbb{R}^n)$. Alors la fonction composée $g \circ f : U \to \mathbb{R}^p$ est dans $\mathcal{C}^1(U, \mathbb{R}^p)$.

Démonstration. On fixe a dans U

$$\begin{aligned}
&\big\|d(g \circ f)(x) - d(g \circ f)(a)\big\|_{\mathcal{L}(\mathbb{R}^m, \mathbb{R}^p)} \\
&\qquad = \|dg(f(x)) \circ df(x) - dg(f(a)) \circ df(a)\|_{\mathcal{L}(\mathbb{R}^m, \mathbb{R}^p)} \leqslant \\
&\qquad \leqslant \|(dg(f(x)) - dg(f(a))) \circ df(x)\|_{\mathcal{L}(\mathbb{R}^m, \mathbb{R}^p)} + \\
&\qquad\quad + \|dg(f(a)) \circ (df(x) - df(a))\|_{\mathcal{L}(\mathbb{R}^m, \mathbb{R}^p)} \leqslant \\
&\qquad \leqslant \|dg(f(x)) - dg(f(a))\|_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)} \|df(x)\|_{\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)} + \\
&\qquad\quad + \|dg(f(a))\|_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)} \|df(x) - df(a)\|_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)}.
\end{aligned} \tag{12.629}$$

On peut conclure en passant à la limite $x \to a$ parce que les fonctions f, g, df et dg sont continues, de telle sorte que

$$\begin{aligned}
\lim_{x \to a} dg\big(f(x)\big) &= dg\big(f(a)\big) \\
\lim_{x \to a} df(x) &= df(a).
\end{aligned} \tag{12.630}$$

□

Remarque 12.240.
On peut prouver le même résultat en utilisant la continuité de l'application bilinéaire

$$\begin{array}{rccc}
\circ : & \mathcal{C}^1(U, V) \times \mathcal{C}^1(V, \mathbb{R}^p) & \to & \mathcal{L}(U, \mathbb{R}^p) \\
& (T, S) & \mapsto & T \circ S.
\end{array} \tag{12.631}$$

12.23 Théorèmes des accroissements finis

Nous avons déjà démontré (lemme 12.178) que si f est différentiable au point x alors $df_x(u) = \partial_u f(x)$. Une importante conséquence est le théorème des accroissements finis

Théorème 12.241 (Accroissements finis, inégalité de la moyenne)**.**
Soit U un ouvert dans $\mathbb{R}^m$ et soit $f : U \to \mathbb{R}^n$ une fonction différentiable. Soient a et b deux point dans U, $a \neq b$, tels que le segment $[a, b]$ soit contenu dans U. Alors

$$\|f(b) - f(a)\|_n \leqslant \sup_{x \in [a,b]} \|df(x)\|_{\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)} \|b - a\|_m. \tag{12.632}$$

Démonstration. On utilise le théorème 12.161 et le fait que

$$\|\partial_u f(x)\|_n \leqslant \|df(x)\|_{\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)} \|u\|_m,$$

pour tout u dans $\mathbb{R}^m$. □

La proposition suivante est une application fondamentale du théorème des accroissements finis 12.241.

Proposition 12.242.
Soit U un ouvert connexe par arcs de $\mathbb{R}^m$ et une fonction $f\colon U \to \mathbb{R}^n$. Les conditions suivantes sont équivalentes :

(1) f est constante ;

(2) f est différentiable et $df(a) = 0$ pour tout $a \in U$;

(3) les dérivées partielles $\partial_1 f, \dots, \partial_m f$ existent et sont nulles sur U.

Démonstration. Nous allons démonter les équivalences en plusieurs étapes. D'abord (1) ⇒ (2), puis (2) ⇒ (3), ensuite (3) ⇒ (2) et enfin (2) ⇒ (1).

Commençons par montrer que la condition (1) implique la condition (2). Si $f(x)$ est constante, alors la condition (12.405) est vite vérifiée en posant $T(h) = 0$.

Afin de voir que la condition (2) implique la condition (3), remarquons d'abord que la différentiabilité de f implique que les dérivées partielles existent (proposition 12.176) et que nous avons l'égalité $df(a).u = \sum_i u_i \partial_i f(a)$ pour tout $u \in \mathbb{R}^m$ (lemme 12.178). L'annulation de $\sum_i u_i \partial_i f(a)$ pour tout u implique l'annulation des $\partial_i f(a)$ pour tout i.

Prouvons maintenant que la propriété (3) implique la propriété (2). D'abord, par la proposition 12.204, l'existence et la continuité des dérivées partielles $\partial_i f(a)$ implique la différentiabilité de f. Ensuite, la formule $df(a).u = \sum_i u_i \partial_i f(a)$ implique que $df(a) = 0$.

Il reste à montrer que (2) implique la condition (1), c'est à dire que l'annulation de la différentielle implique la constance de la fonction. C'est ici que nous allons utiliser le théorème des accroissements finis. En effet, si a et b sont des points de U, le théorème 12.241 nous dit que

$$\|f(b) - f(a)\|_n \leqslant \sup_{x\in[a,b]} \|df(x)\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)} \|b - a\|_m. \tag{12.633}$$

Mais $\|df(x)\| = 0$ pour tout $x \in U$, donc ce supremum est nul et $f(b) = f(a)$, ce qui signifie la constance de la fonction. □

12.24 Fonctions Lipschitziennes

Définition 12.243.
Soient (E, d_E) et (F, d_F) deux espaces métriques[41]*, $f\colon E \to F$ une application et un réel k strictement positif. Nous disons que f est **Lipschitzienne** de constante k sur E si pour tout $x, y \in E$,*

$$d_F\big(f(x) - f(y)\big) \leqslant k d_E(x, y). \tag{12.634}$$

Soit f une fonction k-Lipschitzienne. Si $y \in \overline{B(x,\delta)}$ alors $\|x-y\| \leqslant \delta$ et donc $\big\|f(x)-f(y)\big\| \leqslant k\delta$. Cela signifie que la condition Lipschitz pour s'énoncer en termes de boules fermées par

$$f\big(\overline{B(x,\delta)}\big) \subset \overline{B\big(f(x), k\delta\big)} \tag{12.635}$$

tant que $\overline{B(x,\delta)}$ est contenue dans le domaine sur lequel f est Lipschitz.

Proposition 12.244.
Soit U un ouvert convexe de $\mathbb{R}^m$, et soit $f : U \to \mathbb{R}^n$ une fonction différentiable. La fonction f est Lipschitzienne sur U si et seulement si df est bornée sur U.

Démonstration. Le fait que l'application différentielle df soit bornée signifie qu'il existe un $M > 0$ dans $\mathbb{R}$ tel que $\|df_a\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)} \leqslant M$, pour tout a dans U. Si cela est le cas, alors le théorème 12.241 et la convexité[42] de U impliquent évidemment que f est de Lipschitz de constante plus petite ou égale à M.

41. Pour rappel, les espaces métriques sont définis par la définition 6.18 et le théorème 6.19 ; je précise que nous ne supposons pas que E soit vectoriel ; en particulier il peut être un ouvert de $\mathbb{R}^n$.

42. La convexité de U sert à assurer que la droite reliant a à b est contenue dans U ; c'est ce que nous utilisons dans la démonstration du théorème 12.241.

Inversement, si f est Lipschitz de constante k, alors pour tout a dans U et u dans $\mathbb{R}^m$ on a

$$\left\| \frac{f(a+tu)-f(a)}{t} \right\|_n \leqslant k\|u\|_m,$$

En passant à la limite pour $t \to 0$ on a

$$\|\partial_u f(a)\|_n = \|df_a(u)\|_n \leqslant k\|u\|_m,$$

donc la norme de df_a est majorée par k pour tout a dans U. □

Notez cependant qu'une fonction peut être Lipschitzienne sans être différentiable.

Proposition 12.245.
Une fonction Lipschitzienne $f\colon \mathbb{R} \to \mathbb{R}$ est continue.

Démonstration. Nous utilisons la caractérisation de la continuité donnée par le théorème 6.70. Prouvons donc la continuité en $a \in \mathbb{R}$. Pour tout x nous avons

$$\big|f(x)-f(a)\big| \leqslant k|x-a|. \tag{12.636}$$

Si $\epsilon > 0$ est donné, il suffit de prendre $\delta < \frac{\epsilon}{k}$ pour avoir

$$\big|f(x)-f(a)\big| \leqslant k\frac{\epsilon}{k} = \epsilon. \tag{12.637}$$

Donc f est continue en a. □

Définition 12.246.
Une fonction

$$\begin{aligned} f\colon \mathbb{R}^n \times \mathbb{R}^m &\to \mathbb{R}^p \\ (t,y) &\mapsto f(t,y) \end{aligned} \tag{12.638}$$

est ***localement Lipschitz*** *en y au point (t_0,y_0) s'il existe des voisinages V de t_0 et W de y_0 et un nombre $k>0$ tels que pour tout $(t,y) \in V \times W$ on ait*

$$\big\|f(t_0,y_0)-f(t,y)\big\| \leqslant k\|y-y_0\|. \tag{12.639}$$

La fonction est localement Lipschitz sur un ouvert U de $\mathbb{R}^n \times \mathbb{R}^m$ si elle est localement Lipschitz en chaque point de U.

12.247.
Autrement dit, une fonction est localement Lipschitzienne en sa deuxième variable lorsque tout point admet un voisinage sur lequel elle est Lipschitzienne.

Proposition 12.248.
Si f et g sont deux fonctions localement Lipschitz alors $f+g$ l'est.

Démonstration. Il s'agit d'un simple calcul avec une majoration standard :

$$\begin{aligned} \|(f+g)(t_0,y_0)-(f+g)(t,y)\| &\leqslant \|f(t_0,y_0)-f(t,y)\| + \|g(t_0,y_0)-g(t,y)\| && \text{(12.640a)} \\ &\leqslant k_f\|y-y_0\| + k_g\|y-y_0\| && \text{(12.640b)} \\ &= (k_f+k_g)\|y-y_0\|. && \text{(12.640c)} \end{aligned}$$

□

Lemme 12.249.
La fonction donné par

$$f(t,(x,y)) = xy \tag{12.641}$$

est localement Lipschitz en tout point.

Démonstration. Nous avons la majoration classique

$$|f\big(t,(x_0,y_0)\big)-f\big(t,(x,y)\big)|=|x_0y_0-xy|\leqslant|x_0y_0-x_0y|+|x_0y-xy|\leqslant|x_0||y_0-y|+|y||x_0-x|. \tag{12.642}$$

Vu que nous parlons de fonction *localement Lipschitzienne*, nous pouvons majorer $|y|$ et $|x_0|$ par un même nombre k dans un voisinage de (x_0,y_0). Cela donne

$$|f\big(t,(x_0,y_0)\big)-f\big(t,(x,y)\big)|\leqslant k\big(|y_0-y|+|x_0-x|\big)\leqslant\sqrt{2}k\|\begin{pmatrix}x_0-x\\y_0-y\end{pmatrix}\|. \tag{12.643}$$

Nous avons utilisé l'équivalence de norme de la proposition 10.2(1). □

12.25 Différentielles d'ordre supérieur

Définition 12.250.
Soit U un ouvert de $\mathbb{R}^m$ et $f:U\subset\mathbb{R}^m\to\mathbb{R}^n$ une fonction. La fonction f est dite ***deux fois différentiable*** *au point a dans U, si f est différentiable dans un voisinage de a, et sa différentielle df est différentiable au point a en tant que application de U dans $\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)$.*

La fonction f sera dite deux fois différentiable sur l'ensemble U si elle est deux fois différentiable en chaque point de U.

12.25.1 Identification des espaces d'applications multilinéaires

La différentielle de la différentielle de f est notée

$$d(df)(a)=d^2f(a),$$

et est une application de U dans $\mathcal{L}(\mathbb{R}^m,\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n))$. Comme on a vu dans la proposition 10.87, l'espace $\mathcal{L}(\mathbb{R}^m,\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n))$ est isométriquement isomorphe à l'espace $\mathcal{L}(\mathbb{R}^m\times\mathbb{R}^m,\mathbb{R}^n)$. On verra comment cette propriété est utilisé dans l'exemple 12.253.

Soient V et W deux espace vectoriel normés de dimension finie et $\mathcal{O}$ un ouvert autour de $x\in V$. D'une part l'espace des applications linéaires $\mathcal{L}(V,W)$ est lui-même un espace vectoriel normé de dimension finie, et on peut identifier $\mathcal{L}\big(V,\mathcal{L}^{(k)}(V,W)\big)$ avec $\mathcal{L}^{(k+1)}(V,W)$, ce qui nous permet de dire que la k^{e} différentielle est une application

$$d^kf\colon\mathcal{O}\to\mathcal{L}^{(k)}(V,W). \tag{12.644}$$

Plus précisément, l'identification se fait de la façon suivante : si $\omega\in\mathcal{L}\big(V,\mathcal{L}^{(k)}(V,W)\big)$, alors ω vu dans $\mathcal{L}^{(k+1)}(V,W)$ est définie par

$$\omega(u_1,\dots,u_{k+1})=\omega(u_1)(u_2,\dots,u_{k+1}). \tag{12.645}$$

Cela étant posé nous pouvons donner les définition.

12.25.2 Fonctions différentiables plusieurs fois

Définition 12.251.
La fonction $f\colon\mathcal{O}\subset V\to W$ est

1. *de classe C^0 si elle est continue,*
2. *de classe C^1 si $df\colon\mathcal{O}\to\mathcal{L}(V,W)$ est continue,*
3. *de classe C^k si $d^kf\colon\mathcal{O}\to\mathcal{L}^{(k)}(V,W)$ est continue,*
4. *de classe C^∞ si f est dans $\bigcap_{k=0}^{\infty}C^k(V,W)$.*

Définition 12.252.
*Un C^k-**difféomorphisme** est une application inversible de classe C^k dont l'inverse est également de classe C^k.*

Exemple 12.253
Soit $B : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^n$ une application bilinéaire. On définit $f : \mathbb{R}^m \to \mathbb{R}^n$ par $f(x) = B(x, x)$. Le lemme 12.202 nous dit que B est différentiable. Cela implique la différentiabilité de f. Pour trouver la différentielle de la fonction f, nous écrivons $f = B \circ s$ où $s\colon \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m$ est l'application $s(x) = (x, x)$. En utilisant la règle de différentiation de fonctions composées,

$$df(a) = dB\big(s(a)\big) \circ ds(a). \tag{12.646}$$

Mais $ds(a).u = (u, u)$ parce que $s(a + h) - s(a) - (h, h) = 0$. Par conséquent,

$$df(a).u = dB\big(s(a)\big)(u, u) = B(u, a) + B(a, u) \tag{12.647}$$

où nous avons utilisé la formule du lemme 12.202. La formule (12.647) peut être écrite sous la forme compacte

$$df(a) = B(\ \cdot\ , a) + B(a,\ \cdot\ ) \tag{12.648}$$

La fonction $df(a)$ ainsi écrite est linéaire par rapport à a, donc différentiable. En outre elle coïncide avec sa différentielle, comme on a vu dans le remarque 12.193, au sens que la différentielle de df au point a sera l'application que à chaque x dans $\mathbb{R}^m$ associe l'application linéaire $B(x,\ \cdot\ ) + B(\ \cdot\ , x)$. On voit bien que d^2f au point a est une application de $\mathbb{R}^m$ vers l'espace des applications linéaires $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$. On peut utiliser d'autre part l'isomorphisme des espaces $\mathcal{L}(\mathbb{R}^m, \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n))$ et $\mathcal{L}(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^n)$ et dire que, une fois que a est fixé, l'application $d^2f(a)$ est une application bilinéaire sur $\mathbb{R}^m \times \mathbb{R}^m$. On écrit alors $d^2f(a)(x, y) = B(x, y) + B(y, x)$. △

12.25.3 Différentielle seconde, fonction de classe C^2

Une condition nécessaire et suffisante pour l'existence de la différentielle seconde est la suivante

Proposition 12.254.
Soit U un ouvert de $\mathbb{R}^m$ et $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ une fonction. La fonction f est deux fois différentiable au point a si et seulement si les dérivées partielles $\partial_1 f, \dots, \partial_m f$ sont différentiables en a.

Cela veut dire, en particulier, que f est deux fois différentiable si et seulement si ses dérivées partielles secondes, $\partial_i \partial_j f$, pour toute couple d'indices i, j dans $\{1, \dots, m\}$, existent et sont continues. Pour les différentielles d'ordre supérieur on a la proposition suivante.

La différentielle seconde dans l'exemple 12.253 est symétrique, c'est à dire que $d^2f(a)(x_1, x_2) = d^2f(a)(x_2, x_1)$. En fait toute différentielle seconde est symétrique.

Théorème 12.255 (Schwarz)**.**
Soit U un ouvert de $\mathbb{R}^m$ et $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ une fonction de classe $\mathcal{C}^2$. Alors, pour toute couple i, j d'indices dans $\{1, \dots, m\}$ et pour tout point a dans U, on a

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

Démonstration. Pour simplifier l'exposition nous nous limitons ici au cas $m = 2$. Soit (h, g) un vecteur fixé dans $\mathbb{R}^2$. Pour tout $v = (x, y)$ dans $\mathbb{R}^2$ on note

$$\begin{aligned} \Delta_h f(v) &= f(v + he_1) - f(v) = f(x + h, y) - f(x, y), \\ \Delta_g f(v) &= f(v + ge_2) - f(v) = f(x, y + g) - f(x, y), \end{aligned} \tag{12.649}$$

Nous avons

$$\begin{aligned} \Delta_g \Delta_h f(v) &= \big(f(x + h, y + g) - f(x, y + g)\big) - \big(f(x + h, y) - f(x, y)\big)\,, \\ \Delta_h \Delta_g f(v) &= \big(f(x + h, y + g) - f(x + h, y)\big) - \big(f(x, y + g) - f(x, y)\big)\,, \end{aligned} \tag{12.650}$$

donc,

$$\frac{1}{g}\Delta_g\left(\frac{1}{h}\Delta_h f(v)\right) = \frac{1}{h}\Delta_h\left(\frac{1}{g}\Delta_g f(v)\right). \tag{12.651}$$

On utilise alors le théorème des accroissements finis 12.129

$$\frac{1}{h}\Delta_h f(v) = \frac{1}{h}\big(f(x+h,y) - f(x,y)\big) = \frac{1}{h}\partial_1 f(x+t_1h,y)h = \partial_1 f(x+t_1h,y), \tag{12.652}$$

pour un certain t_1 dans $]0,1[$. De même on obtient

$$\frac{1}{g}\Delta_g f(v) = \partial_2 f(x,y+t_2g),$$

pour un certain t_2 dans $]0,1[$. Alors

$$\frac{1}{g}\Delta_g\big(\partial_1 f(x+t_1h,y)\big) = \frac{1}{h}\Delta_h\big(\partial_2 f(x,y+t_2g)\big). \tag{12.653}$$

En appliquant encore une fois le théorème des accroissements finis on a

$$\partial_2\partial_1 f(x+t_1h,y+s_1g) = \partial_1\partial_2 f(x+s_2h,y+t_2g). \tag{12.654}$$

Il suffit maintenant de passer à la limite pour $(h,g)\to(0,0)$ et de se souvenir du fait que f est $\mathcal{C}^2$ seulement si ses dérivées partielles secondes sont continues pour avoir $\partial_2\partial_1 f(v) = \partial_1\partial_2 f(v)$. □

Si f est deux fois différentiable $d^2f(a)$ est l'application bilinéaire associée avec la matrice symétrique

$$H_f(a) = \begin{pmatrix} \partial_1^2 f(a) & \dots & \partial_1\partial_m f(a) \\ \vdots & \ddots & \vdots \\ \partial_1\partial_m f(a) & \dots & \partial_1^2 f(a), \end{pmatrix} \tag{12.655}$$

Cette matrice est dite la matrice **hessienne** de f.

Exemple 12.256
Montrons qu'il n'existe pas de fonctions f de classe $\mathcal{C}^2$ telles que

$$\begin{cases} \partial_x f(x,y) = 5\sin x & \text{(12.656a)} \\ \partial_y(x,y) = 6x + y. & \text{(12.656b)} \end{cases}$$

Ceci est vite fait en appliquant le théorème de Schwarz, 12.255 ; ce que nous trouvons est

$$\partial_y(\partial_x f) = 0 \neq \partial_x(\partial_y f) = 6.$$

Donc, l'existence d'une fonction f de classe $\mathcal{C}^2$ telle que $\partial_x(x,y) = 5\sin x$ et $\partial_y f(x,y) = 6x+y$ serait en contradiction avec le théorème. △

Soit une fonction de classe C^2 $f\colon V\to\mathbb{R}$ où V est un espace vectoriel de dimension $n<\infty$. Nous avons

$$f\colon V\to\mathbb{R} \tag{12.657a}$$
$$df\colon V\to\mathcal{L}(V,\mathbb{R}) \tag{12.657b}$$
$$d^2f\colon V\to\mathcal{L}\Big(V,\mathcal{L}(V,\mathbb{R})\Big), \tag{12.657c}$$

avec, en suivant les différentes formules du lemme 12.178,

$$df_a(u) = \frac{d}{dt}\Big[f(v+tu)\Big]_{t=0} \tag{12.658}$$

et

$$(d^2f)_a(u) = \frac{d}{dt}\Big[df_{v+tu}\Big]_{t=0} \tag{12.659}$$

pour tout $a, u \in V$. Notons que dans le deuxième cas, il s'agit d'une limite dans $\mathcal{L}(V, \mathbb{R})$. Si $\dim(V) = n$, alors $\dim \mathcal{L}(V, \mathbb{R}) = n$ et avec un choix de base, nous pouvons trouver une matrice $n \times n$ pour $(d^2 f)_a$.

Soit une base $\{e_i\}$ de V et la base duale $\{e_i^*\}$ de $\mathcal{L}(V, \mathbb{R})$. Nous allons chercher la matrice de $(d^2 f)_a$ pour ces bases. L'élément de matrice

$$\big[(d^2 f)_a\big]_{ij} \tag{12.660}$$

est la composante e_j^* de $(d^2 f)_a$ appliqué à e_i. Trouver cette composante e_j^* revient à appliquer l'élément $(d^2 f)_a e_i$ de $\mathcal{L}(V, \mathbb{R})$ à e_j. Le calcul est donc :

$$\begin{aligned}
\big[(d^2 f)_a\big]_{ij} &= \big((d^2 f)_a e_i\big)(e_j) && (12.661a)\\
&= \frac{d}{dt}\Big[df_{a+te_i}(e_j)\Big]_{t=0} && (12.661b)\\
&= \frac{d}{dt}\Big[\frac{d}{ds}\Big[f(a + te_i + se_j)\Big]_{s=0}\Big]_{t=0} && (12.661c)\\
&= \frac{\partial^2 f}{\partial x_i \partial x_j}(a). && (12.661d)
\end{aligned}$$

Attention : le passage à (12.661b) n'est pas une trivialité. Le fait est que si $t \mapsto A(t)$ est une application continue $\mathbb{R} \to \mathcal{L}(V, \mathbb{R})$ alors

$$\lim_{t\to 0} \big(A(t)v\big) = \big(\lim_{t\to 0} A(t)\big)v. \tag{12.662}$$

Donc la matrice de $d^2 f$ est la matrice des dérivées secondes. Il s'agit d'une matrice symétrique par le théorème de Schwarz 12.255.

12.257.
Si $a \in v$, nous pouvons aussi voir $(d^2 f)_a$ comme une forme bilinéaire sur V grâce à la proposition 10.87. Si $u, v \in V$ nous notons

$$(d^2 f)_a(u, v) = (d^2 f)_a(u)v. \tag{12.663}$$

À droite, il s'agit de la définition réelle de $d^2 f$ sans abus de notations, et à gauche, il s'agit d'une notation. Cette application bilinéaire $(d^2 f)_a \in \mathcal{L}^{(2)}(V, \mathbb{R})$ a pour matrice symétrique la matrice des dérivées secondes calculées en a.

Exemple 12.258
Voyons commet la différentielle seconde fonctionne entre deux espaces vectoriels. Soient deux espaces vectoriels de dimension finie V et W. Pour que les choses soient claires, nous avons :

$$\begin{aligned}
f &: V \to W && (12.664a)\\
df &: V \to \mathcal{L}(V, W) && (12.664b)\\
d^2 f &: V \to \mathcal{L}\Big(V, \mathcal{L}(V, W)\Big). && (12.664c)
\end{aligned}$$

Si $a \in V$, alors $(d^2 f)_a$ est une application $V \to \mathcal{L}(V, W)$. Il faut donc l'appliquer à $u \in V$ et ensuite à $v \in V$ pour obtenir un élément de W :

$$\begin{aligned}
(d^2 f)_a(u)v &= \frac{d}{dt}\Big[df_{a+tu}\Big]_{t=0} v && (12.665a)\\
&= \frac{d}{dt}\Big[df_{a+tu}(v)\Big]_{t=0} && (12.665b)\\
&= \frac{d}{dt}\Big[\frac{d}{ds}\Big[f(a + tu + sv)\Big]_{s=0}\Big]_{t=0} && (12.665c)\\
&= \frac{\partial^2 f}{\partial u \partial v}(a). && (12.665d)
\end{aligned}$$

Par conséquent nous voyons

$$\begin{aligned} d^2f\colon V &\to \mathcal{L}^{(2)}(V,W) \\ d^2f_a(u,v) &= \frac{\partial^2 f}{\partial u\partial v}(a). \end{aligned} \tag{12.666}$$

Dans le cas d'une fonction $f\colon \mathbb{R}\to\mathbb{R}$, nous avons une seule direction et par linéarité de (12.666) par rapport à u et v, nous avons

$$d^2f_a(u,v) = f''(a)uv \tag{12.667}$$

où les produits sont des produits usuels dans $\mathbb{R}$ et f'' est la dérivée seconde usuelle. △

Tout ceci est un peu résumé dans la proposition suivante.

Proposition 12.259.
Soit une fonction $f\colon \mathbb{R}^n\to\mathbb{R}$ de classe C^2. Alors en désignant par H_af sa matrice hessienne au point a nous avons

$$(d^2f)_a(u,v) = \frac{\partial^2 f}{\partial u\partial v}(a) = \langle (H_af)u, v\rangle \tag{12.668}$$

pour tout $u,v\in\mathbb{R}^n$.

Démonstration. La première égalité est l'équation (12.667) déjà faite. Pour la seconde, il faut se rappeler du lien entre dérivée partielle et dérivée directionnelle, donné en le lemme 12.178. En particulier ici nous avons

$$\frac{\partial^2 f}{\partial u\partial v} = \sum_{kl}\frac{\partial^2 f}{\partial x_k\partial x_l}(a)u_kv_l = \langle (H_af)u,v\rangle. \tag{12.669}$$

□

En particulier, la matrice hessienne H_af est symétrique et donc diagonalisable (théorème spectral 9.165). Si e_i est un vecteur propre unitaire pour la valeur propre λ_i nous avons

$$(d^2f)_a(e_i,e_i) = \langle (H_af)e_i, e_i\rangle = \lambda_i\langle e_i,e_i\rangle = \lambda. \tag{12.670}$$

Enfin pour celles qui aiment les notations matricielles de tout poil, il y a cette façon-ci d'écrire :

$$(d^2f)_a(\alpha,\beta) = \begin{pmatrix}\alpha & \beta\end{pmatrix}\begin{pmatrix}\partial_x^2 f(a) & \partial_{xy}^2 f(a)\\ \partial_{xy}^2 f(a) & \partial_y^2 f(a)\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}. \tag{12.671}$$

12.25.4 Ordre supérieur

Soit une fonction $f\colon\mathbb{R}^n\to\mathbb{R}$ différentiable l fois. L'application

$$d^lf\colon \mathbb{R}^n\to\mathcal{L}\Big(\mathbb{R}^n,\mathcal{L}\big(\mathbb{R}^n,\mathcal{L}(\dots)\big)\Big) \tag{12.672}$$

au point x appliquée à $v^{(1)}$ appliquée au point $v^{(2)}$, …, appliquée à $v^{(l)}$ est notée

$$(d^lf)_x(v^{(1)},\dots,v^{(l)})\in\mathbb{R}. \tag{12.673}$$

Proposition 12.260.
Soit une fonction $f\colon\mathbb{R}^n\to\mathbb{R}$ différentiable l fois. Avec la notation (12.673) *nous avons*

$$(d^lf)_x(v^{(1)},\dots v^{(l)}) = \sum_{k_1,\dots,k_l} v_{k_1}^{(1)}\dots v_{k_l}^{(l)}\frac{\partial^l f}{\partial x_{k_1}\dots\partial x_{k_l}}(x). \tag{12.674}$$

Démonstration. La preuve se fait par récurrence sur l, en sachant que la formule est déjà vraie pour $l = 1$ et $l = 2$. Si la formule est valable pour l, nous avons

$$(d^{l+1}f)_x(v^{(1)},\dots,v^{(l+1)}) = \frac{d}{dt}\Big[(d^l)_{x+tv^{(l+1)}}(v^{(1)},\dots,v^{(l)})\Big]_{t=0} \tag{12.675a}$$

$$= \sum_{k_1\dots k_l} v_{k_1}^{(1)}\dots v_{k_l}^{(l)}\frac{d}{dt}\Big[\frac{\partial^l f}{\partial x_1\dots\partial x_l}(x+tv^{l+1})\Big]_{t=0} \tag{12.675b}$$

$$= \sum_{k_1\dots k_l} v_{k_1}^{(1)}\dots v_{k_l}^{(l)}\sum_i \frac{\partial}{\partial x_i}\frac{\partial^l f}{\partial x_{k_1}\dots\partial x_k}(x). \tag{12.675c}$$

Cela donne le résultat attendu. □

12.261.
La formule de la proposition 12.260 nous permet d'écrire de jolies formules comme

$$(d^3f)_x(h,h,h) = \sum_{ijk} h_ih_jh_k(\partial^3_{ijk}f)(x). \tag{12.676}$$

Proposition 12.262 (Dérivées partielles et fonctions C^k)**.**
Soit U un ouvert de $\mathbb{R}^m$ et $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$. La fonction f est de classe C^k si et seulement si les dérivées partielles $\partial_1 f,\dots,\partial_m f$ existent et sont de classe C^k.

12.26 Développement asymptotique, théorème de Taylor

Théorème 12.263 (Théorème de Taylor[120, 143])**.**
Soit $I \subset$ un intervalle non vide et non réduit à un point de $\mathbb{R}$ ainsi que $a \in I$. Soit une fonction $f\colon I \to \mathbb{R}$ telle que $f^{(n)}(a)$ existe. Alors il existe une fonction α définie sur I et à valeurs dans $\mathbb{R}$ vérifiant les deux conditions suivantes :

$$f(x) = \sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k + \alpha(x)(x-a)^n, \tag{12.677a}$$

$$\lim_{t\to a}\alpha(t) = 0 \tag{12.677b}$$

pour tout $x \in I$. Ici $f^{(k)}$ dénote la k-ième dérivée de f (en particulier, $f^{(0)} = f$, $f^{(1)} = f'$).

Nous insistons sur le fait que la formule (12.677b) est une égalité, et non une approximation. Ce qui serait une approximation serait de récrire la formule dans le terme contenant α.

Le polynôme $T^a_{f,n}$ est le **polynôme de Taylor** de f au point a à l'ordre n.

Les conditions (12.677) sont souvent aussi énoncées sous la forme qu'il existe une fonction α telle que

$$\begin{cases} \lim_{t\to 0}\dfrac{\alpha(t)}{t^n} = 0 & (12.678a)\\ f(a+h) = f(a) + hf'(a) + \dfrac{h^2}{2}f''(a) + \cdots + \dfrac{h^n}{n!}f^{(n)}(a) + \alpha(h). & (12.678b)\end{cases}$$

Voici un énoncé pour les fonctions à plusieurs variables.

Théorème 12.264 ([144])**.**
Si $f\colon E \to \mathbb{R}$ est une application n fois différentiable en $a \in E$ alors il existe une fonction $\epsilon\colon \mathbb{R} \to \mathbb{R}$ telle que

$$\begin{cases} f(a+h) = f(a) + df_a(h) + \dfrac{1}{2}(d^2f)_a(h,h) + \dots + & (12.679a)\\ \qquad + \dots + \dfrac{1}{n!}(d^nf)_a(h,\dots,h) + \|h\|^n\epsilon(\|h\|) & (12.679b)\\ \lim_{t\to 0}\epsilon(t) = 0. & (12.679c)\end{cases}$$

12.26.1 Fonctions «petit o»

Nous voulons formaliser l'idée d'une fonction qui tend vers zéro « plus vite » qu'une autre. Nous disons que $f \in o\big(\varphi(x)\big)$ si

$$\lim_{x\to 0} \frac{f(x)}{\varphi(x)} = 0. \tag{12.680}$$

En particulier, nous disons que $f \in o(x)$ lorsque $\lim_{x\to 0} f(x)/x = 0$.

Remarque 12.265.
À titre personnel, l'auteur de ces lignes déconseille d'utiliser cette notation qui est un peu casse-figure pour qui ne la maîtrise pas bien.

En termes de notations, nous définissons l'ensemble $o(x)$ l'ensemble des fonctions f telles que

$$\lim_{x\to 0} \frac{f(x)}{x} = 0. \tag{12.681}$$

Plus généralement si g est une fonction telle que $\lim_{x\to 0} g(x) = 0$, nous disons $f \in o(g)$ si

$$\lim_{x\to 0} \frac{f(x)}{g(x)} = 0. \tag{12.682}$$

De façon intuitive, l'ensemble $o(g)$ est l'ensemble des fonctions qui tendent vers zéro «plus vite» que g.

Nous pouvons donner un énoncé alternatif au théorème 12.263 en définissant $h(x) = \epsilon(x+a)x^n$. Cette fonction est définie exprès pour avoir

$$h(x-a) = \epsilon(x)(x-a)^n, \tag{12.683}$$

et donc

$$\lim_{x\to 0} \frac{h(x)}{x^n} = \lim_{x\to 0} \epsilon(x-a) = \lim_{x\to a} \epsilon(x) = 0. \tag{12.684}$$

Donc $h \in o(x^n)$.

Le théorème dit donc qu'il existe une fonction $\alpha \in o(x^n)$ telle que

$$f(x) = T^a_{f,n}(x) + \alpha(x-a). \tag{12.685}$$

pour tout $x \in I$.

Exemple 12.266
L'exemple 16.31 du développement du cosinus sera traité quand les fonctions trigonométriques seront définies. △

Proposition 12.267 (Ordre deux sur $\mathbb{R}^n$[1])**.**
Soit un ouvert Ω de $\mathbb{R}^n$ et $a \in \Omega$ ainsi qu'une fonction $f\colon \Omega \to \mathbb{R}$ de classe C^2. Alors il existe une fonction $\alpha\colon \mathbb{R}^n \to \mathbb{R}$ telle que

$$\begin{cases} f(a+h) = f(a) + df_a(h) + \frac{1}{2}(d^2f)_a(h,h) + \|h\|^2\alpha(h) & \text{(12.686a)} \\ \lim_{h\to 0} \alpha(h) = 0. & \text{(12.686b)} \end{cases}$$

Ici, la notation $(d^2f)_a(h,h)$ réfère à ce qui est expliqué en 12.257.

Démonstration. Dans la suite nous considérons t et h tels que toutes les expressions suivantes aient un sens, c'est à dire que tous les trucs comme $a + th$ restent dans Ω. Pour $h \in \mathbb{R}^n$ nous nommons e_h le vecteur unitaire dans la direction de h, c'est à dire $e_h = h/\|h\|$ et nous posons

$$k_h(t) = f(a + te_h). \tag{12.687}$$

et nous lui appliquons Taylor 12.263 à l'ordre deux : il existe une fonction β_h telle que

$$k_h(x) = k_h(0) + xk'_h(0) + \frac{x^2}{2}k''_h(0) + x^2\beta_h(x). \tag{12.688}$$

avec $\lim_{x\to 0}\beta_h(x) = 0$.

En ce qui concerne les dérivées de k_h nous avons

$$k'_h(0) = df_a(e_h) \tag{12.689}$$

et

$$k''_h(0) = (d^2f)_a(e_h, e_h). \tag{12.690}$$

Il est maintenant temps d'écrire $f(a+h) = k(\|h\|)$ et de substituer les dérivées de k par les différentielles de f dans (12.688) :

$$f(a+h) = k(\|h\|) = f(a) + df_a(h) + \frac{1}{2}(d^2f)_a(h,h) + \|h^2\|\beta_h(\|h\|). \tag{12.691}$$

Il reste à voir que la fonction $\alpha\colon h \mapsto \beta_h(\|h\|)$ tend vers zéro pour $h \to 0$. En prenant la limite $h \to 0$ dans (12.691), il est manifeste que la limite du membre de gauche existe et vaut $f(a)$. Donc la limite du membre de droite doit exister et valoir également $f(a)$. Nous en déduisons que la limite de

$$df_a(h) + \frac{1}{2}(d^2f)_a(h,h) + \|h\|^2\beta_h(\|h\|) \tag{12.692}$$

existe et vaut zéro. La limite des deux premiers termes existe et vaut zéro, donc la limite du troisième existe et vaut zéro :

$$\lim_{h\to 0}\|h\|^2\beta_h(\|h\|) = 0. \tag{12.693}$$

□

Proposition 12.268.
Soit un ouvert Ω de $\mathbb{R}^n$ et une fonction $f\colon \Omega \to \mathbb{R}$ de classe C^1 et deux fois différentiable sur $]x, x+h[$. Alors il existe $\theta \in]0,1[$ tel que

$$f(x+h) = f(x) + df_x(h) + \frac{1}{2}(d^2f)_{x+\theta h}(h,h). \tag{12.694}$$

12.26.2 Autres formulations

Exemple 12.269
Une des façons les plus courantes d'utiliser les formules (12.677) est de développer $f(a+t)$ pour des petits t en posant $x = a+t$ dans la formule :

$$f(a+t) = f(a) + f'(a)t + f''(a)\frac{t^2}{2} + \epsilon(a+t)t^2 \tag{12.695}$$

avec $\lim_{t\to 0}\epsilon(a+t) = 0$. Ici, la fonction T dont on parle dans le théorème est $T^a_{f,2}(a+t) = f(a) + f'(a)t + f''(a)\frac{t^2}{2}$.

Lorsque x et y sont deux nombres «proches[43]», nous pouvons développer $f(y)$ autour de $f(x)$:

$$f(y) = f(x) + f'(x)(y-x) + f''(x)\frac{(y-x)^2}{2} + \epsilon(y-x)(y-x)^2, \tag{12.696}$$

et donc écrire

$$f(x) - f(y) = -f'(x)(y-x) - f''(x)\frac{(y-x)^2}{2} - \epsilon(y-x)(y-x)^2. \tag{12.697}$$

De cette manière nous obtenons une formule qui ne contient plus que y dans la différence $y-x$.
△

43. par exemple dans une limite $(x,y) \to (h,h)$.

12.26.3 Formule et reste

Proposition 12.270.
Soient $f\colon I \subset \mathbb{R} \to \mathbb{R}$ et $a \in \operatorname{Int}(I)$. Soit un entier $k \geqslant 1$. Si f est k fois dérivable en a, alors il existe un et un seul polynôme P de degré $\leqslant k$ tel que

$$f(x) - P(x-a) \in o\big(|x-a|^k\big) \tag{12.698}$$

lorsque $x \to a$, $x \neq a$. Ce polynôme est donné par

$$P(h) = f(a) + f'(a)h + \frac{f''(a)}{2!}h^2 + \cdots + \frac{f^{(k)}(a)}{k!}h^k. \tag{12.699}$$

Notons encore deux façons alternatives d'écrire le résultat. Si $f \in C^k$ il existe une fonction α telle que $\lim_{t\to 0}\alpha(t) = 0$ et

$$f(x) = \sum_{n=0}^{k} \frac{f^{(n)}(a)}{n!}(x-a)^n + (x-a)^n\alpha(x-a). \tag{12.700}$$

Si $f \in C^{k+1}$ alors

$$f(x) = \sum_{n=0}^{k} \frac{f^{(n)}(a)}{n!}(x-a)^n + (x-a)^{n+1}\xi(x-a) \tag{12.701}$$

où ξ est une fonction telle que $\xi(t)$ tend vers une constante lorsque $t \to 0$.

La proposition suivant donne une intéressante façon de trouver le reste d'un développement de Taylor.

Proposition 12.271.
Soient I, un intervalle dans $\mathbb{R}$ et $f\colon I \to \mathbb{R}$ une fonction de classe C^k sur I telle que $f^{(k+1)}$ existe sur I. Soient $a \in \operatorname{Int}(I)$ et $x \in I$. Alors il existe $c \in]x, a[$ tel que

$$f(x) = \sum_{n=0}^{k} \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(k+1)}(c)}{(k+1)!}(x-a)^{k+1}. \tag{12.702}$$

12.26.4 Reste intégral

Proposition 12.272 (Formule de Taylor avec reste intégral[145])**.**
Soient X et Y des espaces normés et un ouvert $\mathcal{O} \subset X$. Si $f \in C^m(\mathcal{O}, Y)$ et si $[p, x] \subset \mathcal{O}$ alors

$$\begin{aligned} f(x) = f(p) &+ \sum_{k=1}^{m-1} \frac{1}{k!}(d^k f)_p(x-p)^k \\ &+ \frac{1}{(m-1)!}\int_0^1 (1-t)^{m-1}(d^m f)_{p+t(x-p)}(x-p)^m \end{aligned} \tag{12.703}$$

où $\omega_p u^k$ signifie $\omega_p(u, \ldots, u)$ lorsque $\omega \in \Omega^k$.

Comme expliqué dans l'exemple 12.258, toute ces applications de différentielles se réduisent à des termes de la forme

$$f^{(k)}(p)(x-p)^k \tag{12.704}$$

dans le cas d'une fonction $\mathbb{R} \to \mathbb{R}$.

12.27 Développement limité autour de zéro

Dans cette sections nous supposons toujours que les fonctions sont définies sur un intervalle ouvert de $\mathbb{R}$, I, contenant 0.

12.27.1 Généralités

Définition 12.273.
*Soit $f\colon I \to 0$ une fonction définie sur un ouvert I autour de zéro. Nous disons que f admet un **développement limité** autour de 0 à l'ordre n s'il existe une fonction $\alpha\colon I \to \mathbb{R}$ telle que*

$$\begin{cases} f(x) = P_n(x) + x^n\alpha(x) & (12.705a)\\ \lim_{x\to 0}\alpha(x) = 0 & (12.705b)\end{cases}$$

*où $P(x) = a_0 + a_1x + \cdots + a_nx^n$ est une polynôme de degré n. Le polynôme P_n est appelé la **partie régulière** du développement.*

La fonction α est appelé le **reste** du développement et sera parfois noté α_f. Lorsque P est la partie régulière d'un développement limité de f nous notons parfois $f \sim P$.

Proposition 12.274 (Troncature)**.**
Si f admet un développement limité d'ordre n alors il admet également un développement limité d'ordre n' pour tout $n' < n$. Ce dernier s'obtient en tronquant le polynôme d'ordre n à l'ordre n'.

Proposition 12.275 (Unicité)**.**
Si f admet une développement limité alors ce dernier est unique : il existe un unique polynôme P_n d'ordre n et une unique fonction α vérifiant simultanément les deux conditions

$$\begin{cases} f(x) = P_n(x) + x^n\alpha(x), & (12.706a)\\ \lim_{x\to 0}\alpha(x) = 0. & (12.706b)\end{cases}$$

Exemple 12.276
En ce qui concerne les séries géométriques de raison x nous savons les formules

$$1 + x + x^2 + \cdots + x^n = \frac{1 - x^{n+1}}{1 - x} \tag{12.707}$$

et

$$1 + x + x^2 + x^3 + \cdots = \frac{1}{1 - x} \tag{12.708}$$

pour tout $x \in]-\infty, 1[$. Comparant les deux, il est naturel d'essayer de prendre $1 + x + x^2 + \cdots + x^n$ comme développement limité de la fonction $f(x) = \frac{1}{1-x}$. Pour voir si cela fonctionne, il faut vérifier si «le reste» est bien de la forme $x^n\alpha(x)$ avec $\lim_{x\to 0}\alpha(x) = 0$.

Le reste en question est donné par

$$\frac{1}{1 - x} - 1 - x - x^2 - \ldots - x^n = \frac{1}{1 - x} - \frac{1 - x^{n+1}}{1 - x} = \frac{x^{n+1}}{1 - x} = x^n\frac{x}{1 - x}. \tag{12.709}$$

En posant $\alpha(x) = \frac{x}{1-x}$ nous avons donc bien

$$f(x) = \frac{1}{1 - x} = 1 + x + x^2 + \cdots + x^n + x^n\frac{x}{1 - x} \tag{12.710}$$

et $\lim_{x\to 0}\frac{x}{1-x} = 0$. Cela est le développement limité de f à l'ordre n autour de 0. △

La formule des accroissements finis est un cas particulier de développement fini. Supposons que f soit dérivable en 0. En effet nous pouvons facilement trouver la fonction α qui convient. Sachant que $f(0) + xf'(0)$ donne l'approximation affine de f autour de 0, nous cherchons α en écrivant

$$f(x) = f(0) + xf'(0) + x\alpha(x). \tag{12.711}$$

Cela nous pousse à définir

$$\alpha(x) = \frac{f(x) - f(0)}{x} - f'(0). \tag{12.712}$$

Notons que cette fonction n'est pas définie en $x = 0$, mais cela n'a pas d'importance : seule la limite $\lim_{x\to 0} \alpha(x)$ nous intéresse. Par définition de la dérivée,

$$\lim_{x\to 0} \alpha(x) = \lim_{x\to 0} \frac{f(x) - f(0)}{x} - f'(0) = 0. \tag{12.713}$$

En conclusion si f est dérivable, son développement limité à l'ordre 1 est donné par

$$f(x) = f(0) + xf'(0) + x\alpha(x) \tag{12.714}$$

où $\alpha(x)$ est donnée par la formule (12.712).

12.27.2 Formule de Taylor-Young

Plus généralement nous avons la proposition suivante qui donne le développement limité de toute fonction dérivable n fois.

Proposition 12.277 (Formule de Taylor-Young)**.**
Soit f une fonction n fois dérivable sur un intervalle I contenant 0. Alors il existe une fonction $\alpha\colon I \to \mathbb{R}$ telle que

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + x^n\alpha(x) \tag{12.715}$$

et

$$\lim_{x\to 0} \alpha(x) = 0. \tag{12.716}$$

Cette proposition nous permet de trouver facilement des développements limités. Dans l'exemple 12.276 nous avons dû utiliser des astuces et des formules pour déterminer le développement limité de $\frac{1}{1-x}$. Au contraire la formule (12.715) nous permet de trouver le polynôme en calculant seulement de dérivées.

Exemple 12.278
Utilisation de la formule (12.715) pour déterminer le développement limité de la fonction

$$f(x) = \frac{1}{1-x}. \tag{12.717}$$

Il faut calculer les dérivées successives de f :

$$f(x) = \frac{1}{1-x} \tag{12.718a}$$

$$f'(x) = \frac{1}{(1-x)^2} \tag{12.718b}$$

$$f''(x) = \frac{2}{(1-x)^3} \tag{12.718c}$$

Avec ces résultats, nous devinons que

$$f^{(n)}(x) = \frac{n!}{(1-x)^{n+1}}. \tag{12.719}$$

Pour en être sûr nous le prouvons par récurrence. La dérivée de $\frac{n!}{(1-x)^{n+1}}$ est donnée par

$$\frac{n!(n+1)(1-x)^n}{(1-x)^{2n+2}} = \frac{(n+1)!}{(1-x)^{n+2}}. \tag{12.720}$$

Évaluées en $x = 0$, les dérivées successives de f sont $f(0) = 0$, $f'(0) = 1$, $f''(0) = 2$,…,$f^{(n)}(0) = n!$. Utilisant la formule (12.715) nous avons

$$f(x) = 1 + x + x^2 + \cdots + x^n + x^n\alpha(x), \tag{12.721}$$

conformément à ce que nous avions déjà trouvé. △

Remarque 12.279.
Pour alléger la notation et ne pas écrire $\dots + x^n\alpha(x)$ nous pouvons aussi écrire

$$f(x) \sim 1 + x + x^2 + \cdots + x^n, \tag{12.722}$$

mais il est interdit d'écrire

$$f(x) = 1 + x + x^2 + \cdots + x^n \tag{12.723}$$

en mettant un signe d'égalité entre une fonction et son développement limité [44].

Notons cependant que la proposition 12.277 ne donne pas de moyen simple de trouver la fonction α. Si la fonction f est très régulière dans l'intervalle I on a le résultat suivant.

Proposition 12.280 (Reste dans la forme de Lagrange)**.**
Si la fonction f est dérivable $n+1$ fois dans I alors il existe $\bar{x}$ dans l'intervalle $[0, x]$ tel que

$$f(x) = P_n(x) + \frac{1}{(n+1)!}f^{n+1}(\bar{x})x^{n+1}. \tag{12.724}$$

Voici quelque développements limités à savoir. Ils sont calculables en utilisant la formule de Taylor-Young (proposition 12.277).

$$e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + x^n\alpha(x) \qquad \text{ordre } n \tag{12.725a}$$

$$\cos(x) = \sum_{k=0}^{p} \frac{(-1)^k x^{2k}}{(2k)!} + x^{2p+1}\alpha(x) \qquad \text{ordre } 2p+1 \tag{12.725b}$$

$$\sin(x) = \sum_{k=0}^{p} \frac{(-1)^k x^{2k+1}}{(2k+1)!} + x^{p+2}\alpha(x) \qquad \text{ordre } 2p+1 \tag{12.725c}$$

$$\ln(1+x) = \sum_{k=0}^{n-1} (-1)^k \frac{x^{k+1}}{k+1} + x^n\alpha(x) \qquad \text{ordre } n \tag{12.725d}$$

$$(1+x)^l = \sum_{k=0}^{l} \binom{l}{k} x^k \qquad \text{exact si } l \text{ est entier.} \tag{12.725e}$$

$$(1+x)^\alpha = 1 + \sum_{k=1}^{n} \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!}x^k + x^n\alpha(x) \qquad \text{ordre } n. \tag{12.725f}$$

Il est fortement suggéré au lecteur de vérifier ces développements par lui-même.

Remarque 12.281.
Quelque remarques.

(1) Les développements de sinus et de cosinus ont un terme sur deux qui est nul. C'est pour cela qu'en ayant une polynôme de degré $2p$, nous avons le développement d'ordre $2p+1$.

(2) En ce qui concerne le développement de $\ln(1+x)$, il faut noter que la somme va jusqu'à $k = n-1$ (et non $k = n$) parce que nous voulons aller jusqu'à x^n et que nous écrivons x^{k+1}.

(3) Si l est entier, le développement de $(1+x)^l$ est exact. Dans le développement de $(1+x)^\alpha$, nous reconnaissons la formule de $\binom{k}{\alpha}$, sauf que nous ne pouvons pas l'écrire avec cette notation lorsque α n'est pas entier.

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

44. Il faut cependant être très prudents avec la notation abrégée. Elle pourrait nous faire oublier des informations importantes, voir les développements des fonction trigonométriques pour un exemple.

12.27.3 Règles de calcul

Les règles suivantes permettent de calculer les développements limités des fonctions qu'on peut écrire comme combinaison des fonctions dans le tableau des équation s(12.725) de la section précédente.

Remarque 12.282.
Il est toujours possible de calculer le développement limité d'une fonction par la formule de Taylor-Young. Les règles suivantes peuvent nous economiser de l'effort et du temps.

12.27.3.1 Linéarité des développements limités

L'opération qui consiste à prendre le développement limité d'une fonction est une opération linéaire : connaissant les développements limités de f et de g, il suffit de les sommer pour obtenir celui de $f+g$. De même, si λ est une constante, le développement limité de λf est le développement limité de f fois λ.

Proposition 12.283.
Soient λ et μ dans $\mathbb{R}$. Si f et g sont deux fonctions acceptant des développements limités d'ordre n

$$f(x) = P(x) + x^n\alpha_f(x) \tag{12.726a}$$
$$g(x) = Q(x) + x^n\beta(x) \tag{12.726b}$$

avec $\lim_{x\to 0}\alpha(x) = \lim_{x\to 0}\beta(x) = 0$, alors la fonction $\lambda f + \mu g$ admet le développement limité

$$(f+g)(x) = (\lambda P + \mu Q)(x) + (\lambda\alpha + \mu\beta)(x). \tag{12.727}$$

Remarque 12.284.
La forme explicite du reste ne nous interesse pas. Dans la pratique on écrira toujours $(f+g)(x) = (P+Q)(x) + \alpha(x)$, où on appelle α une fonction apportune telle que $\lim_{x\to 0}\alpha(x) = 0$.

Démonstration. Vu les définitions (12.726) des polynômes P, Q et des restes α et β, l'égalité (12.727) est une conséquence de la linéarité de la dérivation et de la proposition 12.277

De plus $P+Q$ est un polynôme de degré n dès que P et Q sont des polynômes de degré n, et

$$\lim_{x\to 0}(\lambda\alpha + \mu\beta)(x) = \lim_{x\to 0}\lambda\alpha(x) + \lim_{x\to 0}\mu\beta(x) = 0. \tag{12.728}$$

Par conséquent $\lambda\alpha + \mu\beta$ est la fonction de reste de $\lambda f + \mu g$. □

Exemple 12.285
Calculer le développement de la fonction

$$f(x) = 3\sqrt[3]{1+x} + e^{-2x}. \tag{12.729}$$

Le développement de $\sqrt[3]{1+x}$ est donné par la formule (12.725f) avec $\alpha = \frac{1}{3}$. Nous avons donc dans un premier temps

$$\sqrt[3]{1+x} = 1 + \frac{1}{3}x + \frac{\frac{1}{3}\left(\frac{1}{3}-1\right)}{2}x^2 + \frac{\frac{1}{3}\left(\frac{1}{3}-1\right)\left(\frac{1}{3}-2\right)}{6}x^3 + x^3\alpha(x) \tag{12.730a}$$
$$= 1 + \frac{1}{3}x - \frac{1}{9}x^2 + \frac{5}{81}x^3 + x^3\alpha(x). \tag{12.730b}$$

Nous avons alors

$$3\sqrt[3]{1+x} + e^{-2x} = 3\Big[1 + \frac{1}{3}x - \frac{1}{9}x^2 + \frac{5}{81}x^3 + x^3\alpha(x)\Big] + 1 - 2x + 2x^2 - \frac{4}{3}x^3 + x^3\beta(x) \tag{12.731a}$$
$$= 4 - x + \frac{5}{3}x^2 - \frac{31}{27}x^3 + x^3\big(\alpha(x) + \beta(x)\big). \tag{12.731b}$$

△

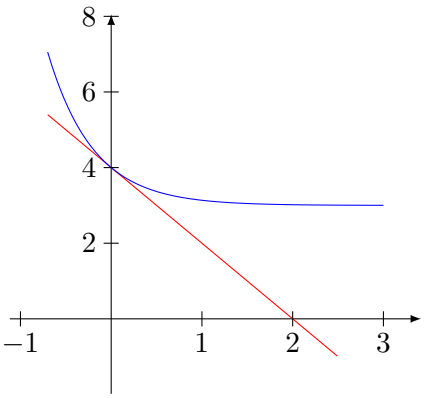
La condition $\lim_{x\to 0}\alpha(x)=0$ signifie que l'approximation qui consiste à remplacer $f(x)$ par le polynôme n'est pas une trop mauvaise approximation lorsque x est petit. Cela ne signifie rien de plus. En particulier si x est grand, l'approximation polynômiale peut-être (et est souvent) très mauvaise.

À ce propos, notez qu'un polynôme tend toujours vers $\pm\infty$ lorsque x est grand. Une approximation polynômiale d'une fonction bornée est donc toujours (très) mauvaise pour les grandes valeurs de x.

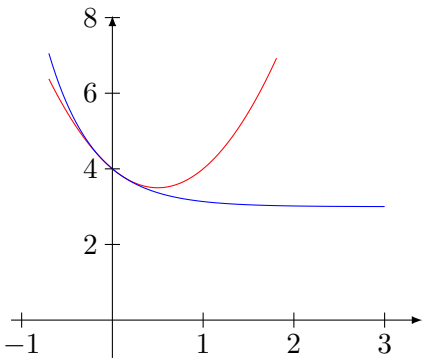
À titre d'exemple nous avons tracé sur la figure 12.13 la fonction

$$f(x)=3\sqrt[3]{x+1}+e^{-2x} \tag{12.732}$$

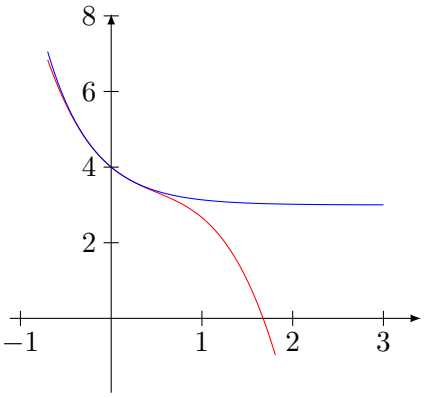
et ses développements limités d'ordre 1 à 3. Il est particulièrement visible que l'approximation est assez bonne pour la partie gauche du graphe sur laquelle la fonction est bien croissante, alors qu'elle est franchement mauvaise sur la droite où le graphe ressemble plutôt à une constante [45].



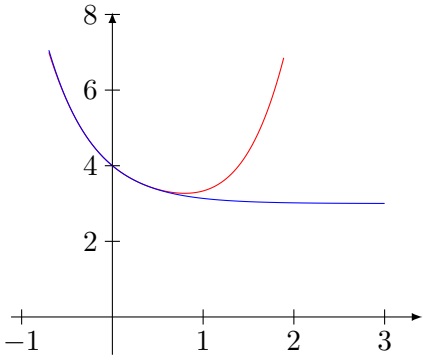
(a) Polyôme d'ordre 1



(b) Polyôme d'ordre 2



(c) Polyôme d'ordre 3



(d) Polyôme d'ordre 4

FIGURE 12.13 – Les développements limités d'ordre de plus en plus grand de la fonction de l'exemple 12.285. La fonction est en bleu et les «approximations» sont en rouge.

12.27.3.2 Développement limité d'un quotient

Proposition 12.286.
Si P_f est le polynôme du développement limité de f à l'ordre n et P_g celui de g, alors nous obtenons le développement limité de f/g à l'ordre n en effectuant la division selon les puissances croissantes de P_f par P_g.

45. Pouvez-vous cependant dire que vaut $\lim_{x\to\infty}f(x)$?

Attention : il s'agit bien de faire une division selon les puissances croissantes, et non une divisions euclidienne. La division euclidienne de A par B consiste à écrire $A = BQ + R$ avec le reste R de degré le plus *petit* possible. Ici nous voulons avoir un reste de degré le plus *grand* possible.

Exemple 12.287
Cherchons le développement limité à l'ordre 5 de $\tan(x) = \frac{\sin(x)}{\cos(x)}$. Nous utilisons les formules (12.725b) et (12.725c) pour savoir les développements de sinus et cosinus [46] :

$$\sin(x) = x - \frac{x^3}{6} + \frac{x^5}{120} + x^5\alpha_1(x) \tag{12.733a}$$

$$\cos(x) = 1 - \frac{x^2}{2} + \frac{x^4}{24} + x^5\alpha_2(x). \tag{12.733b}$$

Nous calculons alors la division des deux polynômes, en classant les puissances dans l'ordre croissant (c'est le sens inverse de ce qui est fait pour la divisions euclidienne !) :

$$\begin{array}{l|l}
x - \frac{1}{6}x^3 + \frac{1}{120}x^5 & 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 \\
-\big(x - \frac{1}{2}x^3 + \frac{1}{24}x^5\big) & \hline
\frac{1}{3}x^3 - \frac{1}{30}x^5 & x + \frac{1}{3}x^3 + \frac{2}{15}x^5 \\
-\big(\frac{1}{3}x^3 - \frac{1}{6}x^5 + \frac{1}{72}x^7\big) & \\
\frac{2}{15}x^5 - \frac{1}{72}x^7 & \\
-\big(\frac{2}{15}x^5 - \frac{1}{15}x^7 + \frac{1}{180}x^9\big) & \\
\frac{29}{360}x^7 - \frac{1}{180}x^9 &
\end{array}$$

Nous avons continué la division jusqu'à obtenir un reste de degré plus grand que 5. Le développement à l'ordre 5 de la fonction tangente autour de zéro est alors

$$\tan(x) = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + x^5\alpha(x). \tag{12.734}$$

Notons que, vu que le reste ne nous intéresse pas vraiment, nous aurions pu ne pas calculer les coefficients des termes en x^7 et x^8. La dernière soustraction était également inutile. △

12.27.3.3 Développement limité d'une fonction composée

Proposition 12.288.
Soient f et g des fonctions admettant des développements limités d'ordre n au voisinage de 0. Nous supposons que $\lim_{x\to 0} g(x) = 0$. Alors la composée $f\big(g(x)\big)$ admet un développement limité d'ordre n au voisinage de 0 qui s'obtient en substituant le développement de g à chaque «x» du développement de f, et en supprimant tous les termes de degré plus élevé que n.

Exemple 12.289
Pour trouver le développement de la fonction $f(x) = e^{-2x}$, il suffit d'écrire celui de e^t et de remplacer ensuite t par $-2x$. Le développement à l'ordre 3 de la fonction exponentielle est :

$$e^t = 1 + t + \frac{t^2}{2} + \frac{t^3}{6} + t^3\alpha(t). \tag{12.735}$$

Le développement de $f(x) = e^{-2x}$ sera donc

$$f(x) = 1 - 2x + \frac{4x^2}{2} - \frac{8x^3}{6} - 8x^3\alpha(-2x). \tag{12.736}$$

46. Encore une fois, vous êtes très fortement encouragés à calculer vous-même ces développements à partir de la formule de Taylor-Young 12.715.

Donc le polynôme de degré 3 partie régulière de g est :

$$1 - 2x + 2x^2 - \frac{4}{3}x^3, \tag{12.737}$$

et la fonction reste correspondante est :

$$\alpha_g(x) = -8\alpha(-2x). \tag{12.738}$$

△

Exemple 12.290
Nous savons les développements

$$f(x) = \ln(1+x) \sim x - \frac{x^2}{2} + \frac{x^3}{3} \tag{12.739}$$

et

$$\sin(x) \sim x - \frac{x^3}{6}. \tag{12.740}$$

Nous obtenons le développement d'ordre 3 de la fonction $x \mapsto \ln\big(1+\sin(x)\big)$ en écrivant

$$\ln\big(1+\sin(x)\big) \sim \big(x - \frac{x^3}{6}\big) - \frac{1}{2}\left(x - \frac{x^3}{6}\right)^2 + \frac{1}{3}\left(x - \frac{x^3}{6}\right)^3. \tag{12.741}$$

Il s'agit maintenant de trouver les termes qui sont de degré inférieur ou égale à 3.

D'abord

$$\left(x - \frac{x^3}{6}\right)^2 = x^2 - \frac{x^4}{3} + \frac{x^6}{36} \sim x^2 \tag{12.742}$$

Nous avons alors aussi

$$\left(x - \frac{x^3}{6}\right)^6 \sim x^2\left(x - \frac{x^3}{6}\right) \sim x^3. \tag{12.743}$$

En replaçant tout ça dans (12.741) nous trouvons

$$\ln\big(1+\sin(x)\big) \sim x - \frac{x^2}{2} + \frac{x^3}{6}. \tag{12.744}$$

△

12.28 Développement au voisinage de $x_0 \neq 0$

Il est intéressant de développer une fonction au voisinage de zéro lorsque nous nous intéressons à son comportement pour les x pas très grands. Il est toutefois souvent souhaitable de savoir le comportement d'une fonction au voisinage d'autres valeurs que zéro.

Pour développer la fonction f autour de x_0, nous considérons la fonction $h \mapsto f(x_0+h)$ que nous développons autour de zéro (pour h). L'objectif est de trouver une polynôme P et une fonction α tels que

$$\begin{cases} f(x) = P(x) + (x-x_0)^n\alpha(x) & \text{(12.745a)} \\ \lim\limits_{x\to x_0} \alpha(x) = 0. & \text{(12.745b)} \end{cases}$$

En pratique, le développement limité à l'ordre n d'une fonction autour d'un point x_0 quelconque à l'intérieur de son domaine prend la forme suivante, qui généralise la formule de Taylor-Young vue dans la proposition 12.277

Proposition 12.291 (Formule de Taylor-Young, cas général)**.**
Soit f une fonction n fois dérivable sur un intervalle I contenant x_0. Alors il existe une fonction $\alpha\colon I\to\mathbb{R}$ telle que

$$\begin{aligned} f(x) = f(x_0)+f'(x_0)(x-x_0) + \frac{f''(x_0)}{2}(x-x_0)^2+ \\ + \frac{f^{(3)}(x_0)}{3!}(x-x_0)^3 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + (x-x_0)^n\alpha(x-x_0) \end{aligned} \tag{12.746}$$

et

$$\lim_{t\to 0}\alpha(t) = 0. \tag{12.747}$$

Exemple 12.292
Développer la fonction cos autour de $x = \frac{\pi}{3}$. Nous développons autour de $h = 0$ la fonction $\cos(\frac{\pi}{3}+h)$:

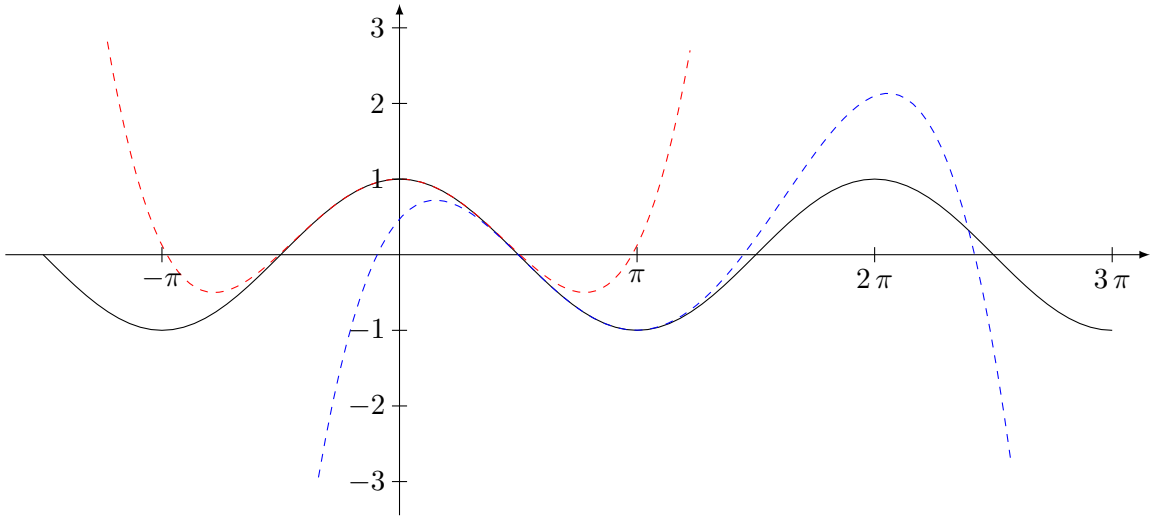
$$\cos\left(\frac{\pi}{3}+h\right) \sim \cos\left(\frac{\pi}{3}\right) + h\cos'(\frac{\pi}{3}) + \frac{h^2}{2}\cos''\left(\frac{\pi}{3}\right) = \frac{1}{2} - \frac{\sqrt{3}}{2}h - \frac{1}{4}h^2. \tag{12.748}$$

Il est aussi possible d'écrire cela en notant $x = x_0 + h$, c'est à dire en remplaçant h par $x - \frac{\pi}{3}$:

$$\cos(x) \sim \frac{1}{2} - \frac{\sqrt{3}}{2}(x-\frac{\pi}{3}) - \frac{1}{4}(x-\frac{\pi}{3})^2. \tag{12.749}$$

△

Pour donner une idée nous avons dessiné sur le graphe suivant la fonction sinus et ses développements d'ordre 4 autour de zéro et autour de $3\pi/4$.



12.29 Application au calcul de limites

Lors d'un calcul de limite, développer une partie d'une expression peut être utile.

Exemple 12.293
À calculer :

$$\lim_{x\to 0}\frac{\ln(1+x)}{x}. \tag{12.750}$$

Cela est une indétermination de type $\frac{0}{0}$. Le développement limité du numérateur nous donne une fonction $\alpha(x)$ telle que $\lim_{x\to 0}\alpha(x) = 0$ et

$$\frac{\ln(1+x)}{x} = \frac{x - \frac{x^2}{2} + x^2\alpha(x)}{x} = 1 - \frac{x}{2} + x\alpha(x). \tag{12.751}$$

Sur le membre de droite la limite est facile à calculer :

$$\lim_{x\to 0}\frac{\ln(1+x)}{x}=\lim_{x\to 0}\Big(1-\frac{x}{2}+x\alpha(x)\Big)=1. \tag{12.752}$$

△

12.30 Développement au voisinage de l'infini

Il est souvent utile de connaître le comportement d'une fonction pour les grandes valeurs de x et de déterminer ses asymptotes éventuelles. La technique que nous allons utiliser consiste à poser $x=\frac{1}{h}$ et de développer la fonction "auxiliaire" $g(h)=f(1/h)$ autour de $h=0$. La limite avec $h\to 0^+$ donnera le comportement pour $x\to\infty$ et la limite $h\to 0^-$ donnera le comportement pour $x\to-\infty$.

Dans le cas d'une développement autour de $\pm\infty$ nous ne parlons plus de développement *limité* mais de **développement asymptotique**.

Exemple 12.294
Calculer

$$\lim_{x\to\infty} e^{1/x}\sqrt{1+4x^2}-2x. \tag{12.753}$$

Nous allons effectuer un développement asymptotique de la partie «difficile» de l'expression posant d'abord $x=1/h$. Si $f(x)=e^{1/x}\sqrt{1-4x^2}$ alors

$$g(h)=\frac{1}{|h|}e^h\sqrt{h^2+4}=\frac{1}{h}\big(1+h+h\alpha(h)\big)\big(2+h\beta(h)\big). \tag{12.754}$$

La première parenthèse est le développement de e^h et la seconde celui de $\sqrt{h^2+4}$. Nous nous apprêtons à faire la limite $x\to\infty$ qui correspond à $h\to 0^+$, nous pouvons donc supposer que $h>0$ et omettre la valeur absolue. En effectuant le produit et en regroupant tous les termes contenant h^2, $\alpha(h)$ ou $\beta(h)$ dans un seul terme $h\gamma(h)$,

$$f(h)=\frac{1}{h}\big(2+2h+h\gamma(h)\big)=\frac{2}{h}+2+\gamma(h)=2x+2+\gamma(1/x) \tag{12.755}$$

où γ est une fonction vérifiant $\lim_{t\to 0}\gamma(t)=0$.

Nous sommes maintenant en mesure de calculer la limite (12.753) :

$$\lim_{x\to\infty} e^{1/x}\sqrt{1+x^2}-2x=\lim_{x\to\infty}\big(2x+2+\gamma(1/x)-2x\big)=2. \tag{12.756}$$

△

12.31 Convergence uniforme

12.31.1 Critère de Cauchy uniforme

Définition 12.295 ([120]).
*Nous disons qu'une suite de fonctions (f_n) définies sur un ensemble A **converge uniformément** vers une fonction f si*

$$\lim_{n\to\infty}\|f_n-f\|_A=0 \tag{12.757}$$

où $\|g\|_A=\sup_{x\in A}\|g(x)\|$.

Proposition 12.296 (Critère de Cauchy uniforme[146]).
Soit X un espace topologique et (Y,d) un espace topologique complet. La suite de fonction $f_n\colon X\to Y$ converge uniformément sur A si et seulement si pour tout $\epsilon>0$ il existe $N\in\mathbb{N}$ tel que si $k,l>N$ alors

$$d\big(f_k(x),f_l(x)\big)\leqslant\epsilon \tag{12.758}$$

pour tout $x\in X$.

Grosso modo, cela dit que si qu'une suite de Cauchy pour la norme uniforme est une suite uniformément convergente. Le fait que la suite converge fait partie du résultat et n'est pas une hypothèse. Ce critère sera utilisé pour montrer que $\big(C(K),\|.\|_\infty\big)$ est complet, proposition 12.299.

Démonstration. Si $f_n \xrightarrow{unif} f$ alors le critère est satisfait ; c'est dans l'autre sens que la preuve est intéressante.

Soit donc une suite de fonctions satisfaisant au critère et montrons qu'elle converge uniformément. Pour tout $x \in X$ la suite $n \mapsto f_n(x)$ est de Cauchy dans l'espace complet Y ; nous avons donc convergence ponctuelle $f_n \to f$. Nous devons prouver que cette convergence est uniforme. Soit $\epsilon > 0$ et $N \in \mathbb{N}$ tel que si $k, l > N$ alors

$$d\big(f_k(x), f_l(x)\big) \leqslant \epsilon \tag{12.759}$$

pour tout $x \in X$. Si nous nous fixons un tel k et un $x \in A$ nous considérons l'inégalité

$$d\big(f_k(x), f_l(x)\big) \leqslant \epsilon \tag{12.760}$$

qui est vraie pour tout l. En passant à la limite $l \to \infty$ (limite qui commute avec la fonction distance par définition de la topologie) nous avons

$$d\big(f_k(x), f(x)\big) \leqslant \epsilon. \tag{12.761}$$

Cette inégalité étant valable pour tout $x \in X$, cela signifie que $f_n \xrightarrow{unif} f$. □

Théorème 12.297 (Limite uniforme de fonctions continues)**.**
Soit A, un ensemble mesuré et $f_n \colon A \to \mathbb{R}^n$, une suite de fonctions continues convergeant uniformément vers f. Si les fonctions f_n sont toutes continues en $x_0 \in A$, alors f est continue en x_0.

Démonstration. Soit $\epsilon > 0$. Si $x \in A$ nous avons, pour tout n, la majoration

$$\begin{aligned}\|f(x) - f(x_0)\| &\leqslant \|f(x) - f_n(x)\| + \|f_n(x) - f_n(x_0)\| + \|f_n(x_0) - f(x_0)\| &\text{(12.762a)}\\ &\leqslant \|f_n(x) - f_n(x_0)\| + 2\|f_n - f\|_\infty. &\text{(12.762b)}\end{aligned}$$

Grâce à l'uniforme convergence, nous considérons $N \in \mathbb{N}$ tel que $\|f_n - f\| \leqslant \epsilon$ pour tout $n \geqslant N$. Pour de tels n, nous avons

$$\|f(x) - f(x_0)\| \leqslant 2\epsilon\|f_n - f\| + \|f_n(x) - f_n(x_0)\|. \tag{12.763}$$

La continuité de f_n nous fournit un $\delta > 0$ tel que $\|f_n(x_0) - f_n(x)\| < \epsilon$ dès que $\|x - x_0\| < \delta$. Pour ce δ, nous avons alors $\|f(x) - f(x_0)\| < \epsilon$. □

12.31.2 Complétude avec la norme uniforme

Proposition 12.298 (Limite uniforme de fonctions continues)**.**
Soit X un espace topologique et (Y, d) un espace métrique. Si une suite de fonctions $f_n \colon X \to Y$ continues converge uniformément, alors la limite est séquentiellement continue[47]*.*

Démonstration. Soit $a \in X$ et prouvons que f est séquentiellement continue en a. Pour cela nous considérons une suite $x_n \to a$ dans X. Nous savons que $f(x_n) \xrightarrow{Y} f(x)$. Pour tout $k \in \mathbb{N}$, tout $n \in \mathbb{N}$ et tout $x \in X$ nous avons la majoration

$$\big\|f(x_n)-f(x)\big\| \leqslant \big\|f(x_n)-f_k(x_n)\big\|+\big\|f_k(x_n)-f_k(x)\big\|+\big\|f_k(x)-f(x)\big\| \leqslant 2\|f-f_k\|_\infty+\big\|f_k(x_n)-f_k(x)\big\|. \tag{12.764}$$

Soit $\epsilon > 0$. Si nous choisissons k suffisamment grand la premier terme est plus petit que ϵ. Et par continuité de f_k, en prenant n assez grand, le dernier terme est également plus petit que ϵ. □

47. Si X est métrique, alors c'est la continuité usuelle par la proposition 6.138.

Proposition 12.299.
Soit X un espace topologique métrique (Y, d) un espace espace métrique complet. Alors les espaces

(1) $\big(C_b^0(X,Y), \|.\|_\infty\big)$ *des fonctions continues et bornées $X \to Y$,*

(2) $\big(C_0^0(X,Y), \|.\|_\infty\big)$ *des fonctions continues et s'annulant à l'infini*

(3) $\big(C_0^k(X,Y), \|.\|_\infty\big)$ *des fonctions de classe C^k et s'annulant à l'infini*

sont complets.

Démonstration. Soit (f_n) une suite de Cauchy dans $C(X,Y)$, c'est à dire que pour tout $\epsilon > 0$ il existe $N \in \mathbb{N}$ tel que si $k, l > N$ nous avons $\|f_k - f_l\|_\infty \leqslant \epsilon$. Cette suite vérifie le critère de Cauchy uniforme 12.296 et donc converge uniformément vers une fonction $f\colon X \to Y$. La continuité (ou l'aspect C^k) de la fonction f découle de la convergence uniforme et de la proposition 12.298 (c'est pour avoir l'équivalence entre la continuité séquentielle et la continuité normale que nous avons pris l'hypothèse d'espace métrique).

Si les fonctions f_k sont bornées ou s'annulent à l'infini, la convergence uniforme implique que la limite le sera également. □

Notons que si X est compact, les fonctions continues sont bornées par le théorème 6.77 et nous pouvons simplement dire que $C^0(X,Y)$ est complet, sans préciser que nous parlons des fonctions bornées.

Lemme 12.300.
Soit A compact et B complet. L'ensemble des fonctions continues de A vers B muni de la norme uniforme est complet.

Dit de façon courte : $\big(C(A,B), \|.\|_\infty\big)$ *est complet.*

Démonstration. Soit (f_k) une suite de Cauchy de fonctions dans $C(A,B)$. Pour chaque $x \in A$ nous avons

$$\|f_k(x) - f_l(x)\|_B \leqslant \|f_k - f_l\|_\infty, \tag{12.765}$$

de telle sorte que la suite $(f_k(x))$ est de Cauchy dans B et converge donc vers un élément de B. La suite de Cauchy (f_k) converge donc ponctuellement vers une fonction $f\colon A \to B$. Nous devons encore voir que cette fonction est continue ; ce sera l'uniformité de la norme qui donnera la continuité. En effet soit $x_n \to x$ une suite dans A convergent vers $x \in A$. Pour chaque $k \in \mathbb{N}$ nous avons

$$\|f(x_n) - f(x)\| \leqslant \|f(x_n) - f_k(x_n)\| + \|f_k(x_n) - f_k(x)\| + \|f_k(x) - f(x)\|. \tag{12.766}$$

En prenant k et n assez grands, cette expression peut être rendue aussi petite que l'on veut ; le premier et le troisième terme par convergence ponctuelle $f_k \to f$, le second terme par continuité de f_k. La suite $f(x_n)$ est donc convergente vers $f(x)$ et la fonction f est continue. □

Problèmes et choses à faire

Il serait sans doute bon de revoir cette preuve à la lumière du critère de Cauchy uniforme 12.296.

12.301 ([147])**.**
Le théorème de Stone-Weierstrass indique que les polynômes sont denses pour la topologie uniforme dans les fonctions continues. Donc il existe des limites uniformes de fonctions C^∞ qui ne sont même pas dérivables. Les espaces de type C^p munis de $\|.\|_\infty$ ne sont donc pas complets sans quelque hypothèses. Voir la proposition 12.299 et le thème 22.

Théorème 12.302 (Théorème de Dini[148])**.**
Soit D un espace métrique compact et une suite de fonctions $f_n \in C(D, \mathbb{R})$ telle que

(1) $f_n \to g$ *ponctuellement,*

(2) $g \in C(D, \mathbb{R})$*,*

(3) la suite (f_n) est croissante, c'est à dire que pour tout $x \in D$ et pour tout $n \geqslant 0$ nous avons $f_{n+1}(x) \geqslant f_n(x)$.

Alors la convergence est uniforme.

Démonstration. Soit $x \in D$ et $\epsilon > 0$. Il existe $N(x) \in \mathbb{N}$ tel que

$$g(x) - \epsilon \leqslant f_{N(x)} \leqslant g(x). \tag{12.767}$$

De plus g et $f_{N(x)}$ sont des fonctions continues, donc il existe $\eta(x)$ tel que si $y \in B\big(x, \eta(x)\big)$ alors

$$g(y) \in B\big(g(x), \epsilon\big) \tag{12.768a}$$
$$f_{N(x)}(y) \in B\big(f_{N(x)}(x), \epsilon\big). \tag{12.768b}$$

Si $n \geqslant N(x)$ et si $y \in B(x, \eta(x))$ alors nous avons les majorations

$$g(y) \geqslant f_n(y) \geqslant f_{N(x)}(y) \geqslant f_{N(x)}(x) - \epsilon \geqslant g(x) - 2\epsilon \geqslant g(y) - 3\epsilon. \tag{12.769}$$

Justifications :

(1) Les deux première inégalités sont la croissance de la suite.
(2) La suivante est (12.768b).
(3) Ensuite il y a le choix de $N(x)$.
(4) Et enfin il y a (12.768a).

Nous retenons que si $x \in D$ et si $n \geqslant N(x)$ alors

$$g(y) \geqslant f_n(y) \geqslant g(y) - 3\epsilon \tag{12.770}$$

pour tout $y \in B(x, \eta(x))$.

Nous utilisons maintenant la compacité de D. Pour chaque $x \in D$ nous pouvons considérer la boule ouverte $B\big(x, \eta(x)\big)$; ces boules recouvrent D. Nous en extrayons un sous-recouvrement fini, c'est à dire un ensemble fini d'éléments $x_1, \ldots, x_K$ tels que

$$D = \bigcup_{k=1}^{K} B\big(x_k, \eta(x_k)\big). \tag{12.771}$$

Si à ce moment vous ne comprenez pas pourquoi c'est une égalité au lieu d'une inclusion, il faut lire l'exemple 6.81. Considérons

$$n \geqslant N = \max\{N(x_1), \ldots, N(x_K)\}. \tag{12.772}$$

Pour tout $y \in D$ il existe $k \in \{1, \ldots, K\}$ tel que $y \in B\big(x_k, \eta(x_k)\big)$, et vu que $n \geqslant N(x_k)$ nous reprenons la majoration (12.770) :

$$g(y) \geqslant f_n(y) \geqslant g(y) - 3\epsilon. \tag{12.773}$$

Pour le n choisi nous avons ces inégalités pour tout $y \in D$, c'est à dire que nous avons $\|f_n - g\| \leqslant 3\epsilon$ et donc la convergence uniforme. □

12.31.3 Permuter avec les dérivées partielles

Théorème 12.303.
Soit $U \subset \mathbb{R}^n$ ouvert, $f_k \colon U \to \mathbb{R}$ et f_k de classe C^1. Supposons que f_k converge simplement vers f et que $\partial_i f_k$ converge uniformément sur tout compact vers une fonction g_i pour $i = 1, \ldots, n$. Alors f est de classe C^1 et $\partial_i f = g_i$. De plus, f_k converge vers f uniformément.

12.32 Fonctions réelles de deux variables réelles

Une **fonction réelle de 2 variables réelles** est une fonction $f : A \subset \mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto z = f(x, y)$.

Le **graphe de** f, noté $\operatorname{Gr} f$, est un sous-ensemble de $\mathbb{R}^3$:

$$\operatorname{Gr} f = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in A \text{ et } z = f(x, y)\}$$

Les **courbes de niveau** de la fonction f sont obtenues en posant $f(x, y) = \lambda$.

12.32.1 Limites de fonctions à deux variables

Ici nous n'allons pas entrer dans tous les détails, mais simplement mentionner les quelques techniques les plus courantes.

Théorème 12.304.
Soient deux fonctions $f : \mathbb{R}^n \to \mathbb{R}^p$ et $g : \mathbb{R}^p \to \mathbb{R}^q$. Si a est un point adhérent au domaine de $g \circ f$ et si

$$\begin{aligned} \lim_{x \to a} f(x) &= b \\ \lim_{y \to b} g(y) &= c, \end{aligned} \tag{12.774}$$

alors

$$\lim_{x \to a} (g \circ f)(x) = c. \tag{12.775}$$

Les techniques usuelles sont

(1) La règle de l'étau. Cette technique demande un peu plus d'imagination parce qu'il faut penser à un «truc» différent pour chaque exercice. En revanche, la justification est facile : il y a un théorème qui dit que ça marche.

(2) Lorsqu'on applique la règle de l'étau, penser à

$$|x| = \sqrt{x^2} \leqslant \sqrt{x^2 + y^2}. \tag{12.776}$$

Cela permet de majorer le numérateur. Attention : ce genre de majoration ne fonctionnent qu'au numérateur : agrandir le dénominateur ferait diminuer la fraction.

(3) Il n'est pas vrai que

$$|x| = \sqrt{x^2} \leqslant \sqrt{x^4} \leqslant \sqrt{x^4 + 2y^4}. \tag{12.777}$$

En effet, si x est petit, alors $x^2 > x^4$, et non le contraire.

Une technique très efficace pour les limites $(x, y) \to (0, 0)$ est le passage aux coordonnées polaires. Il s'agit de poser

$$\begin{cases} x = r\cos(\theta) & \text{(12.778a)} \\ y = r\sin(\theta) & \text{(12.778b)} \end{cases}$$

et puis de faire la limite $r \to 0$.

Si la limite obtenue **ne dépend pas de** θ, alors c'est la limite cherchée. L'exercice suivant en donne des exemples.

Exercice 1

Calculer les limites suivantes :

(1) $\lim_{(x,y)\to(0,0)} \frac{x-y}{x+y}$

(2) $\lim_{(x,y)\to(0,0)} \frac{(xy)^2}{(x+y)^2+(x-y)^2}$

(3) $\lim_{(x,y)\to(0,0)} \frac{xy^3}{x^2+y^2}$

(4) $\lim_{(x,y)\to(0,0)} \frac{x\sin(y)}{\sqrt{x^2+y^2}}$

Correction de l'exercice 1

(1) Ici la méthode des chemins pour est particulièrement éclairante. Regardons d'abord la fonction sur la droite $x = y$. Nous avons

$$f(x,y) = \frac{x-x}{2x} = 0. \tag{12.779}$$

Donc la fonction est nulle sur toute la ligne.

Si nous regardons maintenant la ligne verticale $x = 0$, nous avons

$$f(0,y) = \frac{-y}{y} = -1, \tag{12.780}$$

donc la fonction vaut -1 sur toute la ligne verticale.

(2)

(3) Regardons la technique des coordonnées polaires. Nous remplaçons x par $r\cos(\theta)$ et y par $r\sin(\theta)$:

$$f(r,\theta) = \frac{r^4\cos(\theta)\sin^3(\theta)}{r^2} = r^2\cos(\theta)\sin^3(\theta). \tag{12.781}$$

Cette fonction tend vers zéro quand $r \to 0$. Nous avons donc

$$\lim_{(x,y)\to(0,0)} f(x,y) = 0. \tag{12.782}$$

Pour cet exercice nous pouvons aussi utiliser la règle de l'étau en écrivant d'abord

$$0 \leqslant |f(x,y)| \leqslant \frac{|x||y^3|}{|x^2+y^2|}. \tag{12.783}$$

Mais on a que $|x| \leqslant \sqrt{x^2+y^2}$, $|y| \leqslant \sqrt{x^2+y^2}$ et $|x^2+y^2| = \left(\sqrt{x^2+y^2}\right)^2$, donc

$$0 \leqslant |f(x,y)| \leqslant \frac{\sqrt{x^2+y^2}\left(\sqrt{x^2+y^2}\right)^3}{\left(\sqrt{x^2+y^2}\right)^2} = \left(\sqrt{x^2+y^2}\right)^2 \to 0. \tag{12.784}$$

(4) En passant aux polaires, nous avons

$$f(r,\theta) = \frac{r\cos\theta\sin\big(r\sin\theta\big)}{r} = \cos(\theta)\sin\big(r\sin\theta\big). \tag{12.785}$$

La limite de cette dernière fonction lorsque $r \to 0$ vaut zéro.

Une autre façon de procéder consiste à multiplier et diviser par y de telle façon à faire apparaître $\sin(y)/y$ dont nous connaissons la limite :

$$f(x,y) = \frac{\sin(y)}{y} \cdot \frac{xy}{\sqrt{x^2+y^2}}. \tag{12.786}$$

La limite du premier facteur est 1, tandis que le second peut être traité de façon classique en prenant la valeur absolue et en majorant $|x|$ par $\sqrt{x^2+y^2}$.

12.32.2 Dérivées partielles

La **dérivée partielle** par rapport à x au point (x,y) est notée

$$\frac{\partial f}{\partial x}(x,y) \tag{12.787}$$

et se calcule en dérivant f par rapport à x en considérant que y est constante.

De la même manière, la dérivée partielle par rapport à y au point (x,y) est notée

$$\frac{\partial f}{\partial y}(x,y) \tag{12.788}$$

et se calcule en dérivant f par rapport à y en considérant que x est constante.

Pour les dérivées partielles secondes,

— $f''_{xx}(x,y) = (f'_x)'_x = \frac{\partial^2 f}{\partial x^2}(x,y) = \frac{\partial}{\partial x}(\frac{\partial f}{\partial x})$.
— $f''_{yy}(x,y) = (f'_y)'_y = \frac{\partial^2 f}{\partial y^2}(x,y) = \frac{\partial}{\partial y}(\frac{\partial f}{\partial y})$.
— $f''_{xy}(x,y) = (f'_x)'_y = (f'_y)'_x = f''_{yx}(x,y)$ ou $\frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial}{\partial x}(\frac{\partial f}{\partial y}) = \frac{\partial}{\partial y}(\frac{\partial f}{\partial x}) = \frac{\partial^2 f}{\partial y \partial x}(x,y)$.

12.32.3 Différentielle et accroissement

La **différentielle totale** de f au point (a,b) est donnée, quand elle existe (!), par la formule

$$df(a,b) = \frac{\partial f}{\partial x}(a,b)dx + \frac{\partial f}{\partial y}(a,b)dy. \tag{12.789}$$

De la même façon que la formule des accroissements finis disait que $f(x+a) \simeq f(x) + af'(x)$, en deux dimensions nous avons que l'**accroissement** approximatif de f au point (a,b) pour des accroissements Δx et Δy est

$$f(x+\Delta x, y+\Delta y) = f(x,y) + \Delta x \frac{\partial f}{\partial x}(x,y) + \Delta y \frac{\partial f}{\partial y}(x,y). \tag{12.790}$$

Le **plan tangent** au graphe de f au point $\big(a,b,f(a,b)\big)$ est

$$T_{(a,b)}(x,y) = f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b) \tag{12.791}$$

essayez d'écrire l'équation de la droite tangente au graphe de $f(x)$ au point $x=a$ en terme de la dérivée de f, et comparez votre résultat à cette formule.

Un des principaux théorèmes pour tester la différentiabilité d'une fonction est le suivant.

Théorème 12.305.
Soit une fonction $f\colon \mathbb{R}^m \to \mathbb{R}^p$. Si les dérivées partielles existent dans un voisinage de a et donc continues en a, alors f est différentiable en a.

Le plus souvent, nous prouvons qu'une fonction est différentiable en calculant les dérivées partielles et en montrant qu'elles sont continues.

Dérivation implicite : Soit $F(x,f(x)) = 0$ la représentation implicite d'une fonction $y = f(x)$ alors

$$y' = f'(x) = -\frac{F'_x}{F'_y}.$$

12.33 Les fonctions à valeurs vectorielles

Jusqu'à présent nous avons vu des fonctions de plusieurs variables qui prenaient leurs valeurs dans $\mathbb{R}$. Nous allons maintenant voir ce qu'il se passe lorsque les fonctions prennent leurs valeurs dans $\mathbb{R}^3$.

Une fonction d'une variable est dite **à valeurs vectorielles** lorsque

$$\begin{aligned} f\colon I \subset \mathbb{R} &\to \mathbb{R}^3 \\ f(x) &= \begin{pmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \end{pmatrix}. \end{aligned} \tag{12.792}$$

Les fonctions $f_i\colon \mathbb{R} \to \mathbb{R}$ sont les **composantes** de f. Ce que nous avons raconté à propos des dérivées passe facilement :

$$\frac{f(a+\epsilon)-f(a)}{\epsilon} = \begin{pmatrix} \frac{f_1(a+\epsilon)-f_1(a)}{\epsilon} \\ \frac{f_2(a+\epsilon)-f_2(a)}{\epsilon} \\ \frac{f_3(a+\epsilon)-f_3(a)}{\epsilon} \end{pmatrix}. \tag{12.793}$$

En particulier dès que les fonctions f_i sont dérivables, nous avons

$$f'(a) = \begin{pmatrix} f_1'(a) \\ f_2'(a) \\ f_3'(a) \end{pmatrix} \tag{12.794}$$

comme dérivée de la fonction. Cette dérivée est un vecteur.

Exemple 12.306
Si

$$f\colon x \in \mathbb{R} \mapsto \begin{pmatrix} x^2e^x \\ \cos(x^2) \\ x^3 + x \end{pmatrix}, \tag{12.795}$$

alors

$$f'(x) = \begin{pmatrix} 2xe^x + x^2e^x \\ -2x\sin(x^2) \\ 3x^2 + 1 \end{pmatrix}. \tag{12.796}$$

△

12.34 Fonctions vectorielles de plusieurs variables

Ce sont les fonctions de la forme

$$\begin{aligned} f\colon \mathbb{R}^3 &\to \mathbb{R}^3 \\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} &\mapsto \begin{pmatrix} f_1(x,y,z) \\ f_2(x,y,z) \\ f_3(x,y,z) \end{pmatrix}. \end{aligned} \tag{12.797}$$

En ce qui concerne les dérivées, tout se passe comme avant. Si les dérivées partielles des composantes f_i existent au point $a \in \mathbb{R}^3$, alors

$$\frac{\partial f}{\partial x}(a) = \begin{pmatrix} \partial_x f_1(a) \\ \partial_x f_2(a) \\ \partial_x f_3(a) \end{pmatrix}, \quad \frac{\partial f}{\partial y}(a) = \begin{pmatrix} \partial_y f_1(a) \\ \partial_y f_2(a) \\ \partial_y f_3(a) \end{pmatrix}, \quad \frac{\partial f}{\partial z}(a) = \begin{pmatrix} \partial_z f_1(a) \\ \partial_z f_2(a) \\ \partial_z f_3(a) \end{pmatrix}. \tag{12.798}$$

12.35 Champs de vecteurs

Un champ de vecteur est une fonction $f\colon \mathbb{R}^3 \to \mathbb{R}^3$. Géométriquement, il s'agit simplement de mettre un vecteur en chaque point de l'espace. Cela arrive très souvent en physique.

Exemple 12.307
Si un fluide (eau, gaz) coule dans un tube, en tout point le point a une vitesse, qui sera un vecteur généralement dirigé le long du tube. △

Exemple 12.308
La force d'attraction de la Terre sur une masse m située au point $r = (x, y, z)$ est donnée par

$$F(r) = -G\frac{Mmr}{\|r\|^3}. \tag{12.799}$$

Dans cette expression, tant r que $F(r)$ sont des vecteurs. Nous l'avons représenté sur la figure 12.14.

L'application

$$\begin{aligned} F\colon \mathbb{R}^3 &\to \mathbb{R}^3 \\ r &\mapsto F(r) \end{aligned} \tag{12.800}$$

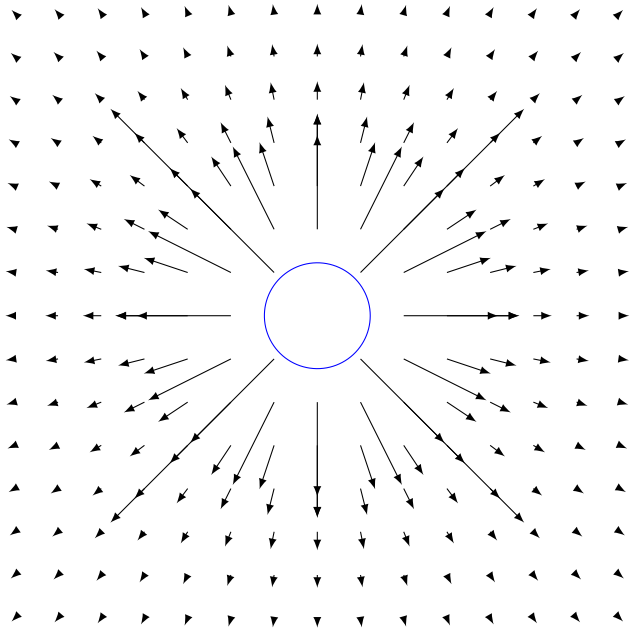


FIGURE 12.14 – Le champ de gravitation de la Terre.

est le champ gravitationnel de la Terre.

△

12.35.1 Matrice jacobienne

La **matrice jacobienne** de la fonction $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ au point $a \in \mathbb{R}^3$ est la matrice dont les colonnes sont les vecteurs $\frac{\partial f}{\partial x}(a)$, $\frac{\partial f}{\partial y}(a)$ et $\frac{\partial f}{\partial z}(a)$, c'est à dire

$$J_f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(a) & \frac{\partial f_1}{\partial y}(a) & \frac{\partial f_1}{\partial z}(a) \\ \frac{\partial f_2}{\partial x}(a) & \frac{\partial f_2}{\partial y}(a) & \frac{\partial f_2}{\partial z}(a) \\ \frac{\partial f_3}{\partial x}(a) & \frac{\partial f_3}{\partial y}(a) & \frac{\partial f_3}{\partial z}(a) \end{pmatrix}. \tag{12.801}$$

Exemple 12.309
Si

$$f(x,y,z) = \begin{pmatrix} xye^z \\ x^2 + \cos(yz) \\ xyz \end{pmatrix}, \tag{12.802}$$

alors

$$J_f(x,y,z) = \begin{pmatrix} ye^z & xe^z & xye^z \\ 2x & -z\sin(yz) & -y\sin(yz) \\ yz & xz & xy \end{pmatrix}. \tag{12.803}$$

△

12.36 Divergence, rotationnel et l'opérateur nabla

Nous avons déjà vu le gradient d'une fonction $f\colon \mathbb{R}^3 \to \mathbb{R}$

$$\nabla f(x,y,z) = \begin{pmatrix} \partial_x f(x,y,z) \\ \partial_y f(x,y,z) \\ \partial_z f(x,y,z) \end{pmatrix} \tag{12.804}$$

Afin de définir la divergence et le rotationnel, nous introduisons ∇ sous une forme un peu plus abstraite comme le «vecteur»

$$\nabla = \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix}. \tag{12.805}$$

Vue comme ça, la formule (12.804) est claire.

Si F est un champ de vecteurs, nous introduisons la **divergence** de F par

$$\nabla \cdot F = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}. \tag{12.806}$$

Cela est une fonction. Et nous introduisons le rotationnel du champ de vecteur F par

$$\begin{aligned} \nabla \times F &= \begin{vmatrix} e_x & e_y & e_z \\ \partial_x & \partial_y & \partial_z \\ F_x & F_y & F_z \end{vmatrix} \\ &= \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) e_x - \left(\frac{\partial F_z}{\partial x} - \frac{\partial F_x}{\partial z} \right) e_y + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) e_z. \end{aligned} \tag{12.807}$$

Cela est un champ de vecteur. En utilisant le symbole complètement antisymétrique ϵ_{ijk}, le rotationnel d'un champ de vecteur peut s'écrire

$$\nabla \times F = \sum_{ijk} \epsilon_{ijk} \partial_i F_j e_k. \tag{12.808}$$

Le gradient, la divergence et le rotationnel consistent à appliquer simplement à ∇ est trois produits qu'on peut effectuer sur un vecteur :

(1) Le produit d'un vecteur par un scalaire multiplie chacune des composantes :

$$\begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} f = \begin{pmatrix} \partial_x f \\ \partial_y f \\ \partial_z f \end{pmatrix}. \tag{12.809}$$

(2) Le produit scalaire d'un vecteur avec un autre vecteur donne lieu à la divergence :

$$\begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} \cdot \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}. \tag{12.810}$$

(3) Le produit vectoriel de deux vecteurs :

$$\begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} \times \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix} = \begin{vmatrix} e_x & e_y & e_z \\ \partial_x & \partial_y & \partial_z \\ F_x & F_y & F_z \end{vmatrix}. \tag{12.811}$$

Ces trois opérations joueront un rôle central en électromagnétisme dans les équations de Maxwell.

Exemple 12.310
Soit $F(x, y, z) = xe_x + xye_y + e_z$, c'est à dire

$$F(x, y, z) = \begin{pmatrix} x \\ xy \\ 1 \end{pmatrix}. \tag{12.812}$$

Son rotationnel est donné par

$$\nabla \times F = \begin{vmatrix} e_x & e_y & e_z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial y} \\ x & xy & 1 \end{vmatrix} = (0-0)e_x - (0-0)e_y + (y-0)e_z = ye_z = \begin{pmatrix} 0 \\ 0 \\ y \end{pmatrix}. \tag{12.813}$$

△

Afin d'étudier comment se comporte la composition de ces opérateurs, nous aurons besoin de ce lemme que nous n'énoncerons pas précisément.

Lemme 12.311.
Si $f\colon \mathbb{R}^3 \to \mathbb{R}$ est une fonction de classe C^2, alors on peut permuter l'ordre des dérivées :

$$\begin{aligned}\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\\ \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial z}\right) &= \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial x}\right)\\ \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial y}\right) &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial z}\right)\end{aligned} \tag{12.814}$$

La fonction

$$(x,y,z) \mapsto \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(x,y,z) \tag{12.815}$$

sera notée

$$\frac{\partial^2 f}{\partial x \partial y}. \tag{12.816}$$

Il y a deux propriétés importantes :

Théorème 12.312.
Soit $f\colon \mathbb{R}^3 \to \mathbb{R}$ une fonction de classe C^2. Alors

$$\nabla \times (\nabla f) = 0. \tag{12.817}$$

Si $F\colon \mathbb{R}^3 \to \mathbb{R}^3$ est un champ de vecteurs de classe C^2, alors

$$\nabla \cdot (\nabla \times F) = 0. \tag{12.818}$$

Démonstration. Ce sont seulement deux calculs qui manipulent les définitions. Pour le premier, la divergence de f est le champ de vecteurs

$$\nabla f = \frac{\partial f}{\partial x}e_x + \frac{\partial f}{\partial y}e_y + \frac{\partial f}{\partial z}e_z. \tag{12.819}$$

En mettant ce champ dans la définition du rotationnel,

$$\begin{aligned}\nabla \times (\nabla f) = \begin{vmatrix} e_x & e_y & e_z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{vmatrix} &= \left[\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial z}\right) - \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial y}\right)\right]e_x\\ &\quad - \left[\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial z}\right) - \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial x}\right)\right]e_y\\ &\quad + \left[\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) - \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\right]e_z.\end{aligned} \tag{12.820}$$

En utilisant le lemme 12.311, chacun des termes fait zéro.

La seconde propriété se démontre en utilisant le même type de calcul. □

Remarque 12.313.
Il n'y a pas de propriétés du même style pour la combinaison $\nabla \times (\nabla \cdot F)$ pour le rotationnel de la divergence. En effet la divergence d'un champ de vecteur est une fonction, et il n'y a pas de rotationnel pour une fonction.

12.37 Interprétation géométrique et physique de la divergence

En physique, on dit qu'un champ de vecteurs à divergence nulle est **incompressible**. Nous allons essayer de comprendre pourquoi. Lorsqu'un fluide incompressible se déplace, il faut qu'en chaque point il y autant de fluide qui rentre que de fluide qui sort. Nous allons voir sur quelques exemples que la divergence d'un champ de vecteurs est le «bilan de masse» d'un fluide qui se déplace selon le champ de vecteurs.

Si en un point la divergence est positive, cela signifie qu'il y a une perte de masse et si la divergence est négative, cela signifie qu'il y a une accumulation de masse.

Prenons par exemple un fluide qui se déplace selon le champ de vitesse montré à figure 12.15.

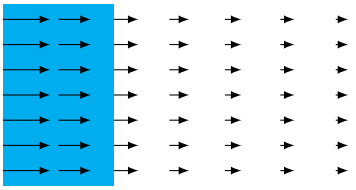


FIGURE 12.15 – Le champ de vecteurs $F(x,y) = \frac{1}{x}(1,0)$.

Étant donné que la vitesse diminue lorsque x avance, il y a une accumulation de fluide. Regardez en effet la quantité de fluide qui rentre dans le rectangle par rapport à la quantité de fluide qui en sort. Ce champ de vecteurs a pour équation :

$$F(x,y) = \frac{1}{x}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1/x\\0\end{pmatrix}. \tag{12.821}$$

Sa divergence vaut donc

$$(\nabla \cdot F)(x,y) = \frac{\partial F_x}{\partial x}(x,y) + \underbrace{\frac{\partial F_y}{\partial y}(x,y)}_{=0} = -\frac{1}{x^2}. \tag{12.822}$$

Cette divergence étant négative, il y a bien accumulation de fluide en tout point, et d'autant plus que x est petit.

Exemple 12.314

Prenons le champ de vecteurs tournant

$$F(x,y) = \frac{1}{\sqrt{x^2+y^2}}\begin{pmatrix}y\\-x\end{pmatrix} \tag{12.823}$$

représenté à la figure 12.16. Cela est un vecteur qui est constamment perpendiculaire au rayon.

Un fluide dont la vitesse serait donné par ce champ de vecteur se contente de tourner. Intuitivement il ne devrait pas y avoir de divergence parce qu'il n'y a aucune accumulation de fluide. En effet,

$$\nabla \cdot F(x,y) = \frac{-2xy}{(x^2+y^2)^2} + \frac{2xy}{(x^2+y^2)^2} = 0. \tag{12.824}$$

△

Exemple 12.315

Prenons le cas du champ de force de gravitation :

$$F(x,y,z) = \frac{1}{(x^2+y^2+z^2)^{3/2}}\begin{pmatrix}x\\y\\z\end{pmatrix}. \tag{12.825}$$

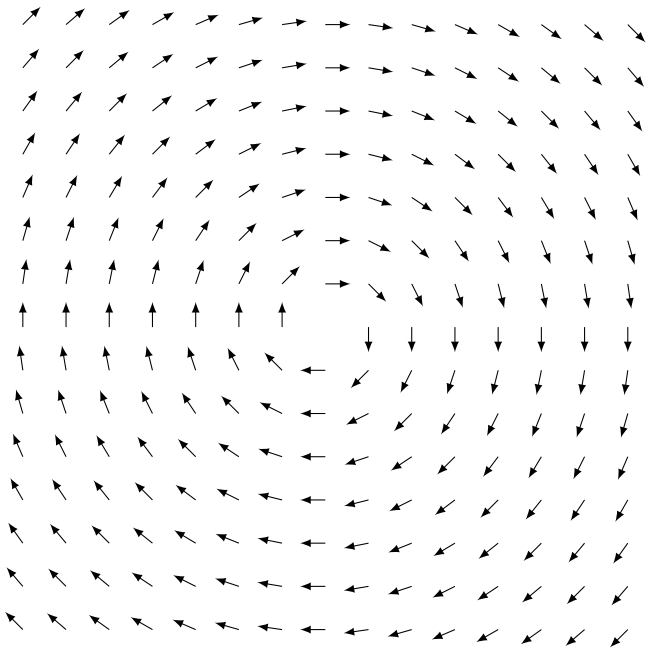


FIGURE 12.16 – Le champ de vecteurs $F(x,y) = (y,-x)$.

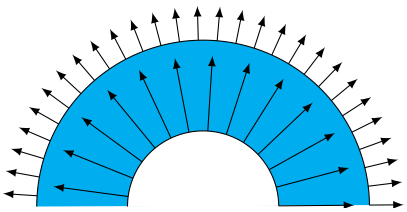


FIGURE 12.17 – Le champ de vecteur de la gravité. Nous avons tracé, sur les deux cercles la même densité de vecteurs, c'est à dire le même nombre de vecteurs par unité de surface.

Nous pouvons rapidement remarquer que $\nabla \cdot F = 0$. Est-ce que cela peut se comprendre sur le dessin de la figure 12.17.

Essayons de voir combien de fluide entre dans la zone bleue et combien en sort. D'abord, il est certain que les vecteurs qui sortent sont plus courts que ceux qui rentrent, ce qui voudrait dire qu'il y a plus de fluide qui rentre. Mais on voit également que le *nombre* de vecteurs qui sortent est plus grand parce que la seconde sphère est plus grande et qu'il y a un vecteur en chaque point de la sphère.

Intuitivement nous pouvons dire que la quantité qui rentre dans la sphère de rayon r_1 donnée par la taille des vecteurs entrants multiplié par la surface de la sphère, c'est à dire

$$4\pi r_1^2 \|F(x,y,z)\|, \tag{12.826}$$

mais $\|F(x,y,z)\| = \frac{1}{r_1^2}$, donc la quantité de fluide entrant est 4π. La quantité de fluide sortant sera la même.

Cela explique deux choses

(1) Pourquoi les forces de gravitation et électromagnétiques sont en $1/r^2$; c'est parce que nous vivons dans un monde avec trois dimensions d'espace. En étudiant très précisément le champ de gravitation, certains physiciens espèrent trouver des déviations expérimentales par rapport à la règle du $1/r^2$; cela *pourrait* être un signe que l'espace contient des dimensions supplémentaires.

(2) Pourquoi il y a un 4π comme coefficient dans beaucoup d'équations en électromagnétisme ; en particulier dans certaines anciennes unités de flux.

△

Remarque 12.316.
Nous allons voir plus loin comment s'assurer que l'équation (12.826) représente bien la «quantité de fluide» qui rentre dans la zone délimitée

12.38 Quelques formules de Leibnitz

La divergence étant une combinaison de dérivées, il n'est pas tellement étonnant que la divergence de produits donne lieux à des formules en deux termes. Si f est une fonction et si F et G sont des champs de vecteurs, nous avons (sans démonstrations) :

$$\begin{aligned}\nabla \cdot (fF) &= f\nabla \cdot F + F \cdot \nabla f\\ \nabla \cdot (F \times G) &= G \cdot \nabla \times F - F \cdot \nabla \times G.\end{aligned} \tag{12.827}$$

Nous avons aussi, pour le rotationnel,

$$\nabla \times (fF) = f\nabla \times F + \nabla f \times F. \tag{12.828}$$

12.39 Action de groupe et connexité

Sources : [149] et wikipédia.

Théorème 12.317.
Soit G un groupe topologique localement compact et dénombrable à l'infini[48] *agissant continument et transitivement sur un espace topologique localement compact E. Alors l'application*

$$\begin{aligned}\varphi\colon G/G_x &\to E\\ [g] &\mapsto g \cdot x\end{aligned} \tag{12.829}$$

est un homéomorphisme.

Lemme 12.318.
Si G et H sont des groupes topologiques tels que G/H et H sont connexes[49]*, alors G est connexe.*

Démonstration. Soit $f\colon G \to \{0,1\}$ une fonction continue. Considérons l'application

$$\begin{aligned}\tilde{f}\colon G/H &\to \{0,1\}\\ [g] &\mapsto f(g).\end{aligned} \tag{12.830}$$

D'abord nous montrons qu'elle est bien définie. En effet si $h \in H$ nous aurions $\tilde{f}([gh]) = f(gh)$, mais étant donné que H est connexe, l'ensemble gH est également connexe ; la fonction continue f est donc constante sur gH. Nous avons donc $f(gh) = f(g)$.

Étant donné que G/H est également connexe, la fonction $\tilde{f}$ doit être constante. Si g_1 et g_2 sont deux éléments du groupe, nous avons $f(g_1) = \tilde{f}([g_1]) = \tilde{f}([g_2]) = f(g_2)$. Nous en déduisons que f est constante et que G est connexe. □

Théorème 12.319.
Le groupe $\mathrm{SO}(n)$ *est connexe, le groupe* $\mathrm{O}(n)$ *a deux composantes connexes.*

48. Cela signifie qu'il est une réunion dénombrable de compacts
49. Définition 6.84.

Démonstration. La seconde assertion découle de la première parce que les matrices de déterminant 1 et celles de déterminant -1 ne peuvent pas être reliées par un chemin continu tandis que l'application

$$M \mapsto \begin{pmatrix} -1 & & \\ & 1 & \\ & & 1 \end{pmatrix} M \tag{12.831}$$

est un homéomorphisme entre les matrices de déterminant 1 et celles de déterminants -1. Montrons donc que $G = \mathrm{SO}(n)$ est connexe par arcs pour $n \geqslant 2$ en procédant par récurrence sur la dimension.

Nous acceptons le résultat pour $G = \mathrm{SO}(2)$. Notons que nous en avons besoin pour prouver que la sphère S^{n-1} est connexe.

Le groupe $\mathrm{SO}(n)$ agit, par définition, de façon transitive sur la sphère S^{n-1}. Soit $a \in S^{n-1}$, nous avons

$$G \cdot a = S^{n-1} \tag{12.832a}$$
$$G_a \simeq \mathrm{SO}(n-1) \tag{12.832b}$$

où G_a est le fixateur de a dans G. Pour montrer le second point, nous considérons $\{e_i\}$, la base canonique de $\mathbb{R}^n$ et $M \in G$ telle que $Ma = e_1$. Le fixateur de e_1 est évidemment isomorphe à $\mathrm{SO}(n-1)$ parce qu'il est constitué des matrices de la forme

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & a_{11} & \dots & a_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n-1,1} & \dots & a_{n-1,n-1} \end{pmatrix} \tag{12.833}$$

où $(a_{ij}) \in \mathrm{SO}(n-1)$. L'application

$$\begin{aligned} \alpha\colon G_{e_1} &\to G_a \\ A &\mapsto M^{-1}AM \end{aligned} \tag{12.834}$$

est un isomorphisme entre G_a et $\mathrm{SO}(n-1)$. Le théorème 12.317 nous montre alors que, en tant qu'espaces topologiques,

$$G/G_a = S^{n-1}. \tag{12.835}$$

L'hypothèse de récurrence montre que $G_a = \mathrm{SO}(n-1)$ est connexe tandis que nous savons que S^{n-1} est connexe. Le lemme 12.318 conclut que $G = \mathrm{SO}(n)$ est connexe. □

Lemme 12.320.
Une bijection continue entre un espace compact et un espace séparé est un homéomorphisme.

Proposition 12.321.
Les groupes $\mathrm{U}(n)$ *et* $\mathrm{SU}(n)$ *sont connexes.*

Démonstration. Soit $G(n)$ le groupe $\mathrm{SU}(n)$ ou $\mathrm{U}(n)$. Ce groupe opère transitivement sur la sphère complexe

$$S^{n-1}_{\mathbb{C}} = \{z \in \mathbb{C}^n \text{ tel que } \langle z, z\rangle = \sum_k |z_k|^2 = 1\}. \tag{12.836}$$

Cet ensemble est le même que S^{2n-1} parce que $|z_k| = x_k^2 + y_k^2$. Nous avons une bijection continue entre S^{n-1} et $S^{n-1}_{\mathbb{C}}$ et donc un homéomorphisme (lemme 12.320). Soit $a \in S^{n-1}_{\mathbb{C}}$, nous avons

$$G \cdot a = S^{n-1}_{\mathbb{C}} \tag{12.837a}$$
$$G_a \simeq G(n-1). \tag{12.837b}$$

La seconde ligne est un isomorphisme de groupe et un homéomorphisme. Il est donné de la façon suivante. D'abord le fixateur de e_1 dans $G(n)$ est donné par les matrices de la forme

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & a_{11} & \dots & a_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n-1,1} & \dots & a_{n-1,n-1} \end{pmatrix} \tag{12.838}$$

où $(a_{ij}) \in G(n-1)$. Par ailleurs si M est une matrice de $G(n)$ telle que $Ma = e_1$, nous avons l'homéomorphisme

$$\begin{aligned} \alpha\colon G_{e_1} &\to G_a \\ A &\mapsto M^{-1}AM. \end{aligned} \tag{12.839}$$

Encore une fois, cela est un homéomorphisme par le lemme 12.320. Par composition nous avons $G_a \simeq G(n-1)$ et un homéomorphisme

$$G(n)/G_a = S^{n-1}_{\mathbb{C}}. \tag{12.840}$$

Le groupe G_a et l'ensemble $S^{n-1}_{\mathbb{C}}$ étant connexes, le groupe $G(n)$ est connexe par le lemme 12.318. □

Lemme 12.322 ([117])**.**
Si G est un sous-groupe connexe de $\mathrm{GL}(n,\mathbb{C})$ *alors son groupe dérivé*[50] *l'est également.*

Démonstration. Soit S_m l'ensemble des produits de m commutateurs de G :

$$S_m = \{g_1, \dots, g_m \text{ où les } g_i \text{ sont des commutateurs}\}. \tag{12.841}$$

La partie S_m est l'image de G par l'application continue

$$\begin{aligned} \underbrace{G \times \dots \times G}_{2m \text{ facteurs}} &\to G \\ (g_1, h_1, g_2, h_2, \dots, g_m, h_m) &\mapsto [g_1, h_1] \dots [g_m, h_m] \end{aligned} \tag{12.842}$$

En tant qu'image d'un connexe par une application continue, S_m est connexe par la proposition 6.86. Vu que les S_m ont l'identité en commun, le groupe dérivé

$$D(G) = \bigcup_{m=1}^{\infty} S_m \tag{12.843}$$

est également connexe. □

12.40 Espaces de matrices

L'ensemble des matrices est un espace vectoriel. Nous identifions $\mathbb{M}(n,\mathbb{R})$ avec $\mathbb{R}^{n^2}$; plus précisément, nous identifions une matrice

$$A = (a_{i,j})_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant n} \tag{12.844}$$

avec le vecteur $x = (x_1, x_2, \dots, x_{n^2}) \in \mathbb{R}^{n^2}$, où $a_{i,j} = x_{(n-1)i+j}$.

50. Définition 3.19.

12.40.1 Dilatations et transvections

Soit un corps commutatif $\mathbb{K}$ et $n \geqslant 2$.

Théorème-définition 12.323 ([117]).
Soit une application linéaire $u\colon E \to E$ dont les points fixes forment un hyperplan noté H d'équation $H = \ker(f)$ avec $f \in E^*$.

(1) Les affirmations suivantes sont équivalentes :

(a) $\det(u) \neq 1$

(b) L'application u est diagonalisable et a une valeur propre qui vaut $\det(u) \neq 1$.

(c) $\text{Image}(u - \text{Id}) \nsubseteq H$.

(d) Il existe une base de E dans laquelle la matrice de u est $\text{diag}(1, \ldots, 1, \lambda)$ *avec* $\lambda \neq 1$.

(2) Les affirmation suivantes sont équivalentes :

(i) Il existe $a \in H$ tel que pour tout $x \in E$, $u(x) = x + f(x)a$.

(ii) Dans une base adaptée, la matrice de u est donnée par

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix}. \tag{12.845}$$

(3) Les conditions (1)a-(1)d sont respectées si et seulement si les conditions (2)i-(2)ii ne sont pas respectées (elles sont les négations l'une de l'autre.).

Un endomorphisme qui est soit l'identité soit respecte les conditions (1) est une ***dilatation****. Un endomorphisme qui est soit l'identité soit qui vérifie les conditions (2) est une* ***transvection*** *(dans les deux cas il faut que les points fixes forment un hyperplan).*

Notons que selon cette terminologie, l'application $x \mapsto \lambda x$ n'est pas une dilation mais un produit de dilations.

Démonstration. Nous allons prouver plein d'implications ...

(1)a implique (1)b Le théorème de la base incomplete (voir remarque 7.11) permet de considérer une base $\{e_1, \ldots, e_n\}$ de E telle que $\{e_1, \ldots, e_{n-1}\}$ soit une base de H. Dans cette base, la matrice de u est de la forme suivante (les cases non remplies sont nulles et les étoiles correspondent à des valeurs inconnues mais pas spécialement nulles) :

$$\begin{pmatrix} 1 & & & * \\ & \ddots & & \vdots \\ & & 1 & * \\ & & & \lambda \end{pmatrix} \tag{12.846}$$

Le fait que le déterminant de u ne soit pas 1 implique que $\lambda \neq 1$. Par conséquent le polynôme caractéristique

$$\chi_u(X) = (1 - X)^{n-1}(\lambda - X) \tag{12.847}$$

possède une racine $\lambda \neq 1$, et donc u possède un vecteur propre v pour cette valeur [51]. Le vecteur v est linéairement indépendant de $\{e_1, \ldots, e_{n-1}\}$ (parce que vecteur propre de valeur propre différente). Par conséquent l'ensemble $\{e_1, \ldots, e_{n-1}, v\}$ est une base par le théorème 7.14(3). Cela est une base de vecteurs propres et donc une base de diagonalisation [52].

51. Proposition 9.133.

52. Nous pourrions en dire à peine plus et prouver le point (1)d, mais cela ne servirait à rien parce que nous voulons prouver les équivalences et qu'il faudra quand même prouver que (1)c implique (1)d.

(1)b implique (1)c Nous nommons maintenant $\{e_1, \dots, e_n\}$ la base de diagonalisation. Nous avons $u(e_n) = \lambda e_n$ avec $\det(u) = \lambda \neq 1$. Nous avons

$$(u - \mathrm{Id})(e_n) = (\lambda - 1)e_n \notin H, \tag{12.848}$$

ce qui prouver que l'image de e_n par $u - \mathrm{Id}$ n'est pas dans H.

(1)c implique (1)d Reprenons une base $\{e_1, \dots, , e_n\}$ donnant la matrice (12.846). Il existe $x \in E$ tel que $u(x) - x$ n'est pas dans H, c'est à dire tel que $u\big(u(x) - x\big) \neq u(x) - x$. Nous en déduisons que

$$u^2(x) - 2u(x) + x \neq 0 \tag{12.849}$$

ou encore que

$$(X - 1)^2(u)x \neq 0. \tag{12.850}$$

C'est à dire que $(X - 1)^2$ n'est pas un polynôme annulateur de u. Or ce serait le cas si $X - 1$ était le polynôme minimal (proposition 9.110). Le polynôme caractéristique étant $(X - 1)^{n-1}(X - \lambda)$ (et étant annulateur [53]), le polynôme minimal est de la forme

$$\mu_u(X) = \begin{cases} (X - 1)(X - \lambda) & \text{si } \lambda \neq 1 \\ X - 1 & \text{si } \lambda = 1. \end{cases} \tag{12.851}$$

Dans notre cas nous venons de voir que ce n'est pas $X - 1$ et donc c'est $(X - 1)(X - \lambda)$ avec $\lambda \neq 1$.

Nous devons trouver une base de diagonalisation ... Supposons

$$u(e_n) = \sum_{k=1}^{n-1} a_k e_k + \lambda e_n, \tag{12.852}$$

dans lequel nous venons de prouver que $\lambda \neq 1$, et cherchons

$$e'_n = \sum\nolimits_{-} j = 1^n p_j e_j \tag{12.853}$$

de telle sorte à avoir $u(e'_n) = \lambda e_n$. Nous avons

$$u(e'_n) = \sum_{j=1}^{n-1} p_j u(e_j) + p_n u(e_n) \tag{12.854a}$$

$$= \sum_{j=1}^{n-1} (p_j + p_n a_j)e_j + p_n \lambda e_n. \tag{12.854b}$$

En égalisant à $\lambda \sum_{j=1}^{n} p_j e_j$, il vient

$$p_j + p_n a_j = \lambda p_j \tag{12.855}$$

pour tout $j = 1, \dots, n-1$ et la condition triviale $p_n \lambda = \lambda p_n$ pour $j = n$. Nous en déduisons que le choix

$$p_j = \frac{p_n a_j}{\lambda - 1} \tag{12.856}$$

fonctionne (parce que $\lambda \neq 1$ comme nous l'avons démontré plus haut). En bref, il suffit de poser

$$e'_n = \sum_{j=1}^{n-1} \frac{p_n a_j}{\lambda - 1} e_j + p_n e_n \tag{12.857}$$

avec p_n au choix pour avoir une base $\{e_1, \dots, e_{n-1}, e'_n\}$ de diagonalisation de u avec $\lambda \neq 1$ comme dernière valeur propre.

53. Théorème de Cayley-Hamilton 9.130.

(1)d implique (1)a Évident ... encore qu'il faut invoquer l'invariance du déterminant par changement de base.

Nous avons terminé la première série d'équivalences. Nous continuons avec la seconde.

(2)i implique (2)ii Nous prenons $e_{n-1} = a$ et nous complétons en une base de H. Pour e_n il suffit de prendre n'importe quel vecteur v tel que $f(v) \neq 0$ (qui existe parce que $f = 0$ est seulement un hyperplan), et de le normaliser.

Dans cette base, la matrice de u a la forme désirée parce que $u(e_n) = e_n + f(e_n)a = e_n + e_{n-1}$ du fait que $e_{n-1} = a$ et $f(e_n) = 1$.

(2)ii implique (2)i Soit $\{e_1, \dots, e_n\}$ cette base. En prenant $a = e_{n-1}$ et en posant $x = \sum_k x_k e_k$ nous avons

$$u(x) = \sum_{k=1}^{n-1} x_k e_k + x_n(e_{n-1} + e_n) = x + x_n e_{n-1} = x_n a. \tag{12.858}$$

Mais vu que $f(x) = \sum_i f_i x_i$, et que $f(e_i) = 0$ pour tout $i = 1, \dots, n-1$ nous avons $f(x) = f_n x_n$. Il n'y a cependant pas de raisons d'avoir $f_n = 1$. Cependant en définissant

$$e'_i = \frac{1}{f_n} e_i \tag{12.859}$$

nous avons bien $u(e'_n) = \frac{1}{f_n}(e_{n-1} + e_n) = e'_{n-1} + e'_n$. Donc dans cette base nous avons encore la matrice de u de la forme

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix}, \tag{12.860}$$

mais cette fois avec $f(e'_n) = 1$.

Nous avons terminé avec la seconde série d'équivalences. Il nous reste à prouver que la première est équivalente à la négation de la seconde.

non (1)c implique (2)i Considérons $x_0 \in E$ tel que $f(x_0) = 1$ et posons $a = u(x_0) - x_0 \in \text{Image}(u - \text{Id})$. Par la négation de (1)c nous avons $a \in H$. De plus $x_0 \notin H$ (sinon $f(x_0) = 0$) donc $u(x_0) \neq x_0$ et $a \neq 0$.

Nous montrons que ce choix de a fonctionne : $u(x) = x + f(x)a$ pour tout $x \in E$. Nous faisons cela séparément pour $x \in H$ et pour $x = x_0$.

Si $h \in H$ alors $u(h) = h$ et $f(h) = 0$ donc $h + f(h)a = h = u(h)$. Si $x = x_0$ alors $u(x_0) = a + x_0$ (cela est la définition de a) et$x_0 + f(x_0)a = x_0 + a$.

(2)ii implique non (1)a Dans une base adaptée nous avons

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix}, \tag{12.861}$$

et donc $\det(u) = 1$, ce qui contredit (1)a.

□

Remarque 12.324.
Nous notons E_{ij} la matrice qui possède uniquement 1 en position (i, j). C'est à dire que $\big(E_{ij}\big)_{kl} = \delta_{ik}\delta_{jl}$. Soit H l'hyperplan des points fixes de f. Dans une base contenant une base de H, la matrice d'une transvection a pour forme type :

$$T_{ij}(\lambda) = \mathbb{1} + \lambda E_{ij} \tag{12.862}$$

avec $i \neq j$ et $\lambda \in \mathbb{K}$, et une dilatation a pour forme type la matrice diagonale

$$D_i(\alpha) = \mathbb{1} + (\alpha - 1)E_{ii} \tag{12.863}$$

avec $\alpha \in \mathbb{K}^*$.

Bien entendu, en choisissant une base quelconque, les matrices des dilatations et des translations peuvent avoir des formes différentes.

Lemme 12.325.
Quelque manipulations de lignes et de colonnes pour les matrices.

(1) La multiplication à gauche par $T_{ij}(\lambda)$ revient à effectuer le remplacement de ligne

$$L_i \to L_i + \lambda L_j. \tag{12.864}$$

(2) La multiplication à droite par $T_{ij}(\lambda)$ revient à effectuer le remplacement de colonne

$$C_j \to C_j + \lambda C_i. \tag{12.865}$$

(3) La multiplication à gauche par $T_{ij}(1)T_{ji}(-1)T_{ij}(1)$ revient à la substitution de lignes

$$\begin{cases} L_i \to L_j & \text{(12.866a)} \\ L_j \to -L_i. & \text{(12.866b)} \end{cases}$$

Note qu'il n'est pas possible d'inverser deux lignes à l'aide de transvections sans changer un signe parce que les transvections sont de déterminant 1 alors que l'inversion de lignes change le signe du déterminant.

Démonstration. Point par point.

Pour (1) Nous devons prouver que

$$\big(T_{ij}(\lambda)A\big)_{kl} = \begin{cases} A_{kl} & \text{si } k \neq i \\ A_{il} + \lambda A_{jl} & \text{si } k = i. \end{cases} \tag{12.867}$$

Un peu de calcul matriciel avec utilisation modérée des indices donner :

$$\begin{aligned} \big(T_{ij}(\lambda)A\big)_{kl} &= \sum_s \big(T_{ij}(\lambda)\big)_{ks} A_{sl} && \text{(12.868a)} \\ &= \sum_s \delta_{ks}A_{sl} + \lambda\delta_{ik}\delta_{js}A_{sl} && \text{(12.868b)} \\ &= A_{kl} + \lambda\delta_{ik}A_{jl}. && \text{(12.868c)} \end{aligned}$$

Pour (2) C'est la même chose.

Pour (3) Si nous appliquons successivement ces trois matrices (de droite à gauche) nous effectuons les substitutions :

$$\begin{cases} L'_i = L_i + L_j \\ L'_j = L_j \end{cases} \text{ suivit de } \begin{cases} L''_i = L'_i \\ L''_j = L'_j - L'_i \end{cases} \text{ et de } \begin{cases} L'''_i = L''_i + L''_j \\ L'''_j = L''_j. \end{cases} \tag{12.869}$$

En effectuant ces substitutions,

$$L'''_i = L''_i + L''_j = L'_i + (L'_j - L'_i) = L'_j = L_j \tag{12.870}$$

et

$$L'''_j = L''_j = L'_j - L'_i = L_j - (L_i + L_j) = -L_i, \tag{12.871}$$

ce qu'il fallait.

□

Proposition 12.326 ([150]).
Soit $n \geqslant 2$ et un corps commutatif $\mathbb{K}$.

(1) Si $A \in \mathrm{GL}(n,\mathbb{K})$, il existe des transvections $U_1,\dots,U_r$, $V_1,\dots,V_s$ telles que

$$A = U_1 \dots U_r \, D_n\big(\det(A)\big) \, V_1 \dots V_s. \tag{12.872}$$

(2) L'ensemble des transvections engendre le groupe spécial linéaire $\mathrm{SL}(n,\mathbb{K})$.

(3) L'ensemble des transvections et des dilatations engendre le groupe linéaire $\mathrm{GL}(n,\mathbb{K})$.

Démonstration. Nous allons montrer que toutes les matrices de $\mathrm{SL}(n,\mathbb{K})$ peuvent être écrites comme produits de matrices de la forme (12.862). Cela montrera qu'étant donné un endomorphisme f et une base *pas spécialement liée à f*, il est possible décrire la matrice de f comme produit de transvections dont les hyperplans invariants sont «contenus» dans cette base. Cela suffit à prouver que les transvections engendrent $\mathrm{SL}(n,\mathbb{K})$ grâce au lemme 3.9.

Toutes les transvections ont un déterminant égal à 1. Donc le groupe engendré par les transvections est inclus à $\mathrm{SL}(2,\mathbb{K})$. Soit $A \in \mathrm{GL}(n,\mathbb{K})$; nous allons utiliser le pivot de Gauss pour la diagonaliser. Étant donné que A est inversible, sa première colonne n'est pas nulle. Si $A_{i1} \neq 0$ alors une multiplication à gauche par $L_{1i}\big((A_{11}-1)/A_{i1}\big)$ effectue la substitution

$$L_1 \to L_1 - \frac{A_{11}-1}{A_{i1}} L_i \tag{12.873}$$

qui met un 1 en la position $(1,1)$. Notons que si la première colonne est de la forme

$$\begin{pmatrix} s \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{12.874}$$

avec $s \neq 0$ alors il faut plutôt faire les substitutions $L_2 \to L_2 + L_1$ et ensuite $L_1 \to L_1 - \frac{1}{s}L_2$ pour obtenir le même résultat. En effectuant le pivot avec A_{11}, une suite d'opérations sur les lignes et les colonnes donnent

$$M_1 \dots M_p A N_1 \dots N_q = \begin{pmatrix} 1 & 0 \\ 0 & A_1 \end{pmatrix} \tag{12.875}$$

où $A_1 \in \mathrm{GL}(n-1,\mathbb{K})$ et $\det(A_1) = \det(A)$. En continuant de la sorte nous arrivons sur une matrice diagonale [54]

$$M_1 \dots M_{p'} A N_1 \dots N_{q'} = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \alpha \end{pmatrix} \tag{12.876}$$

avec $\alpha = \det(A)$. En d'autres termes nous avons prouvé qu'il existe des transvections $U_1,\dots,U_r$ et $V_1,\dots,V_s$ telles que

$$A = U_1 \dots U_r \, D_n\big(\det(A)\big) \, V_1 \dots V_s. \tag{12.877}$$

Cela prouve que les transvections et les translations engendrent $\mathrm{GL}(n,\mathbb{K})$. Si $A \in \mathrm{SL}(n,\mathbb{K})$ alors $D_n\big(\det(A)\big) = 1$ et l'équation (12.877) est un produit de transvections. □

Proposition 12.327.
Le groupe $\mathrm{GL}(n,\mathbb{R})$ est engendré par les endomorphismes inversibles diagonalisables.

54. Attention : les opérations sur les lignes et le colonnes ne sont pas des opérations de similitude. Il n'est pas question de prétendre ici que toutes les matrices de $\mathrm{GL}(n,\mathbb{K})$ sont diagonales, voir la définition 7.27.

Démonstration. Par la proposition 12.326, le groupe $\mathrm{GL}(n,\mathbb{R})$ est engendré par les dilatations et les transvections. Il suffit donc de montrer qu'à leur tour, ces deux types d'endomorphismes sont engendrés par les endomorphismes inversibles et diagonalisables.

Les dilatations sont diagonalisables et inversibles. Soit une transvection u, et une base $\{e_i\}_{i=1,\dots,n}$ dans laquelle u est de la forme (12.845). Nous considérons l'endomorphisme $d\colon E\to E$ défini par $d(e_k)=ke_k$. Cet endomorphisme est diagonalisable parce que son polynôme minimal, $\mu_d=\prod_{k=1}^{n}(X-k)$, est scindé à racines simples (voir le théorème 9.143).

Nous avons évidemment $u=d^{-1}\circ(d\circ u)$ où d^{-1} est diagonalisable et inversible. Voyons que $d\circ u$ est également diagonalisable en montrant que μ_d est son polynôme minimal (qui est scindé à racines simples).

Il suffit de montrer que $\mu_d(d\circ u)(e_k)=0$ pour tout k. Ainsi μ_d sera un polynôme annulateur de $d\circ u$ de degré n, et donc minimal.

Si $k\leqslant n-1$ Alors $u(e_k)=e_k$ et $(d\circ u-n)e_k=(k-n)e_k$. En tout :

$$\mu_d(d\circ u)(e_k)=(d\circ u-1)(d\circ u-2)\dots(d\circ u-n)e_k=(k-1)(k-2)\dots(k-n)e_k=0 \quad (12.878)$$

parce que dans le produit des $k-i$, il y en a forcément un de nul.

Si $k=n$ Dans un premier temps,

$$(d\circ u-n)e_n=d(e_n+e_{n-1})-ne_n=ne_n+(n-1)e_{n-1}-ne_n=(n-1)e_{n-1}. \quad (12.879)$$

Ensuite

$$\begin{aligned}\big(d\circ u-(n-1)\big)e_{n-1}&=d(e_{n-1})-(n-1)e_{n-1} && (12.880a)\\ &=d(e_{n-1})-(n-1)e_{n-1} && (12.880b)\\ &=(n-1)e_{n-1}-(n-1)e_{n-1} && (12.880c)\\ &=0 && (12.880d)\end{aligned}$$

Le polynôme μ_d est donc un polynôme scindé à n racines simples annulateur de $d\circ u$, qui est alors diagonalisable et inversible (parce que u et d le sont).

Donc sous la forme $u=d^{-1}(du)$, la transvection u est écrite comme produit de diagonalisables inversibles. □

Proposition 12.328 ([91])**.**
Soit $n\geqslant 3$ et $\mathbb{K}$ un corps de caractéristique différente de 2. Alors

(1) le groupe dérivé de $D(\mathrm{GL}(n,\mathbb{K}))$ est $\mathrm{SL}(n,\mathbb{K})$;

(2) le groupe dérivé de $\mathrm{SL}(n,\mathbb{K})$ est $\mathrm{SL}(n,\mathbb{K})$.

La preuve utilise le fait que les transvections engendrent $\mathrm{SL}(n,\mathbb{K})$ et que les transvections avec les dilatations engendrent $\mathrm{GL}(n,\mathbb{K})$. Voir la proposition 12.326.

12.40.2 Connexité de certains groupes

Lemme 12.329.
Le groupe $\mathrm{O}(n),\mathbb{R}$ n'est pas connexe.

Démonstration. La non connexité par arcs est facile parce que les éléments de déterminant 1 ne peuvent pas être reliés aux éléments de déterminant -1 par un chemin continu restant dans $\mathrm{O}(n)$ à cause du théorème des valeurs intermédiaires 12.21.

En ce qui concerne la connexité, il faut en dire un peu plus.

Les éléments de $\mathrm{O}(n,\mathbb{R})$ ont des déterminants égaux à 1 ou à -1. Ces deux parties sont des ouverts (pour la topologie induite de $\mathbb{M}(n,\mathbb{R})$). En effet soit $A\in\mathrm{SO}(n,\mathbb{R})$ (la partie contenant les déterminants 1 ; ce que l'on va dire tient pour l'autre partie). Alors, vu que le déterminant est une fonction continue sur $\mathbb{M}(n,\mathbb{R})$ il exist un voisinage $\mathcal{O}$ de A dans
$eM(n,\mathbb{R})$ dans lequel le déterminant reste entre $\frac{1}{2}$ et $\frac{3}{2}$ (c'est la définition de la continuité avec

$\epsilon = 1/2$). L'ensemble $\mathcal{O} \cap \mathrm{O}(n, \mathbb{R})$ est par définition un ouvert de $\mathrm{O}(n, \mathbb{R})$ et ne contient que des éléments de déterminant 1.

La partie $\mathrm{O}(n, \mathbb{R})$ de
$eM(n, \mathbb{R})$ est donc non-connexe selon la définition 6.84. □

Lemme 12.330.
Les groupes $\mathrm{U}(n)$ *et* $\mathrm{SU}(n)$ *sont connexes par arcs.*

Démonstration. Soit A, une matrice unitaire et Q une matrice unitaire qui diagonalise A. Étant donné que les valeurs propres arrivent par paires complexes conjuguées,

$$QAQ^{-1} = \begin{pmatrix} e^{i\theta_1} & & & & & \\ & e^{-i\theta_1} & & & & \\ & & \ddots & & & \\ & & & e^{i\theta_r} & & \\ & & & & e^{-i\theta_r} & \end{pmatrix}. \tag{12.881}$$

Le chemin $U(t)$ obtenu en remplaçant θ_i par $t\theta_i$ avec $t \in [0, 1]$ joint QAQ^{-1} à l'identité. Par conséquent $Q^{-1}U(t)Q$ joint A à l'unité. □

Proposition 12.331.
Le groupe $\mathrm{SO}(n)$ *est connexe.*

Théorème 12.332.
Les matrices normales[55] *forment un espace connexe par arc.*

Démonstration. Soit A une matrice normale, et U une matrice unitaire qui diagonalise A. Nous considérons $U(t)$, un chemin qui joint $\mathbb{1}$ à U dans $\mathrm{U}(n)$. Pour chaque t, la matrice

$$A(t) = U(t)^{-1}AU(t) \tag{12.882}$$

est normale. Nous avons donc trouvé un chemin dans les matrices normales qui joint A à une matrice diagonale. Il est à présent facile de la joindre à l'identité.

Toutes les matrices normales étant connexes à l'identité, l'ensemble des matrices normales est connexe. □

Proposition 12.333.
Le groupe $\mathrm{SL}(n, \mathbb{K})$ *est connexe par arcs.*

Démonstration. Soit $A \in \mathrm{SL}(n, \mathbb{K})$; par la proposition 12.326(2) nous pouvons écrire

$$A = \prod_{c \in X} T_c(\lambda_c) \tag{12.883}$$

où X est une partie de l'ensemble des couples (i, j) dans $\{1, \ldots, n\}$. En posant

$$\begin{aligned} \varphi\colon [0, 1] &\to \mathrm{SL}(n, \mathbb{K}) \\ t &\mapsto \prod_{c \in X} T_c(t\lambda_c) \end{aligned} \tag{12.884}$$

nous avons une application continue de A vers $\mathbb{1}$, dont pour tout t la matrice $\varphi(t)$ est inversible de déterminant1.

Donc tous les éléments de $\mathrm{SL}(n, \mathbb{K})$ peuvent être reliés à $\mathbb{1}$. Donc $\mathrm{SL}(n, \mathbb{K})$ est connexe par arcs. □

Proposition 12.334 ([151])**.**
Le groupe $\mathrm{GL}(n, \mathbb{C})$ *est connexe par arcs.*

55. Définition 9.161.

Démonstration. Soit $A \in \mathrm{GL}(n,\mathbb{C})$ et sa décomposition (12.872). Comme fait précédemment, chacune des transvections peut être reliée à $\mathbb{1}$ par un chemin continu dans $\mathrm{SL}(n,\mathbb{C})$. En ce qui concerne le facteur de translation, nous ne pouvons pas simplement prendre le chemin donné par $t \mapsto D_n\big(t\det(A)\big)$ parce que le résultat n'est pas inversible en $t = 0$.

Vu que C^* il existe une application continue $\alpha\colon [0,1] \to \mathbb{C}^*$ telle que $\alpha(0) = \det(A) \in \mathbb{C}^*$ et $\alpha(1) = 1$. Il suffit alors de prendre $D_n\big(\alpha(t)\big)$ et nous avons un chemin continu de A vers $\mathbb{1}$ restant dans $\mathrm{GL}(n,\mathbb{C})$. □

Proposition 12.335.
Le groupe $\mathrm{GL}(n,\mathbb{R})$ *a exactement deux composantes connexes par arcs.*

Démonstration. Nous notons $\mathrm{GL}^+(n,\mathbb{R})$ et $\mathrm{GL}^-(n,\mathbb{R})$ les parties de $\mathrm{GL}(n,\mathbb{R})$ formées des applications de déterminant ± 1 respectivement. Vu le théorème des valeurs intermédiaires (théorème 12.21), il n'existe pas d'applications continues dans $\mathrm{GL}(n,\mathbb{R})$ reliant $\mathrm{GL}^+(n,\mathbb{R})$ à $\mathrm{GL}^-(n,\mathbb{R})$ tout en restant dans les applications de déterminant non nul [56].

Montrons que $\mathrm{GL}^{\pm}(n,\mathbb{R})$ sont connexes par arcs. Si $A \in \mathrm{GL}^+(n,\mathbb{R})$ alors grâce à la décomposition (12.872), il existe un chemin continu de A vers $D_n\big(\det(A)\big)$. Vu que $\mathbb{R}^{\pm}$ sont connexes par arc, il est possible de relier $D_n\big(\det(A)\big)$ à $D_n(\pm 1)$ par un chemin continu. □

12.40.3 Densité

Proposition 12.336.
Les matrices diagonalisables sont denses dans $\mathbb{M}(n,\mathbb{C})$.

Démonstration. D'après le lemme de Schur 9.152, une matrice de $\mathbb{M}(n,\mathbb{C})$ est de la forme

$$A = Q\begin{pmatrix} \lambda_1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_n \end{pmatrix}Q^{-1}. \tag{12.885}$$

Les valeurs propres sont sur la diagonale. La matrice est diagonalisable si les éléments de la diagonales sont tous différents. Il suffit maintenant de considérer n suites $(\epsilon_k^{(r)})_{k\in\mathbb{N}}$ convergentes vers zéro telles que pour chaque k les nombres $\lambda_r + \epsilon_k^{(r)}$ soient tous différents. La suite de matrices

$$A_k = Q\begin{pmatrix} \lambda_1 + \epsilon_k^{(1)} & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_n + \epsilon_k^{(n)} \end{pmatrix}Q^{-1}. \tag{12.886}$$

est alors diagonalisable pour tout k et nous avons $\lim_{k\to\infty} A_k = A$. □

Proposition 12.337.
Les matrices inversibles sont denses dans l'ensemble des matrices. C'est à dire que $\mathrm{GL}(n,\mathbb{R})$ *est dense dans* $\mathbb{M}(n,\mathbb{R})$.

Démonstration. Soit $A \in \mathbb{M}(n,\mathbb{R})$; le lemme de Schur réel 9.164 nous permet d'écrire

$$A = Q\begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & \begin{pmatrix} a & b \\ c & d \end{pmatrix} & & \\ & & & & \ddots & \end{pmatrix}Q^{-1} \tag{12.887}$$

56. Si $\varphi\colon [0,1] \to \mathrm{GL}(n,\mathbb{R})$ est le chemin, la fonction à mettre dans le théorème des valeurs intermédiaires est la fonction $f\colon [0,1] \to \mathbb{R}$ $t \mapsto \det\big(\varphi(t)\big)$.

avec Q orthogonale.

Pour définir A_k nous remplaçons λ_i par $\lambda_i + \epsilon^{(i)k}$ de façon à avoir $\epsilon_k^{(i)} \to 0$ et $\lambda_i + \epsilon_k^{(i)} \neq 0$. En ce qui concerne les blocs, ceux dont le déterminant est non nul, nous n'y touchons pas, et ceux dont le déterminant est nul, nous remplaçons a par $a + \epsilon_k$.

Avec cela, QA_kA^{-1} est une suite dans $\mathrm{GL}(n, \mathbb{R})$ qui converge vers A. □

Proposition 12.338.
Si $A \in \mathbb{M}(n, \mathbb{C})$ alors

$$e^{\mathrm{Tr}(A)} = \det(e^A). \tag{12.888}$$

Démonstration. Ici, e^A est l'exponentielle soit d'endomorphisme soit de matrice définie par la proposition 10.88.

Le résultat est un simple calcul pour les matrices diagonalisable. Si A n'est pas diagonalisable, nous considérons une suite de matrices diagonalisables A_k dont la limite est A (proposition 12.336). La suite

$$a_k = e^{\mathrm{Tr}(A_k)} \tag{12.889}$$

converge vers $e^{\mathrm{Tr}(A)}$ tandis que la suite

$$b_k = \det(e^{A_k}) \tag{12.890}$$

converge vers $\det(e^A)$. Mais nous avons $a_k = b_k$ pour tout k ; les limites sont donc égales. □

Théorème 12.339 (Cayley-Hamilton[152, 153])**.**
Tout endomorphisme d'un espace vectoriel de dimension finie sur un corps commutatif quelconque annule son propre polynôme caractéristique

Une autre démonstration est donnée en le théorème 9.130.

Démonstration. La preuve est divisée en plusieurs étapes.

Endomorphisme diagonalisable Soit u un endomorphisme sur un espace vectoriel V de dimension n sur un corps $\mathbb{K}$ et χ_u sont polynôme caractéristique. Nous savons que si λ est une valeur propre de u alors $\chi_u(\lambda) = 0$ le théorème 9.125(2). En combinant avec le lemme 9.104, si x est vecteur propre pour la valeur propre λ de u nous avons

$$\chi_u(u)x = \chi_u(\lambda)x = 0. \tag{12.891}$$

Donc tant que u possède une base de vecteurs propres nous avons $\chi_u(u) = 0$.

Le cas complexe Nous nous restreignons à présent (et provisoirement) au cas $\mathbb{K} = \mathbb{C}$, ce qui nous donne $u \in \mathbb{M}(n, \mathbb{C})$. Les matrices diagonalisables sont denses dans $\mathbb{M}(n, \mathbb{C})$ par la proposition 12.336. Si $A \in \mathbb{M}(n, \mathbb{C})$ nous considérons une suite de matrices diagonalisables $A_k \xrightarrow{\mathbb{M}(n,\mathbb{C})} A$. Pour chaque k nous avons par le point précédent

$$\chi_{u_k}(u_k) = 0. \tag{12.892}$$

Chacune des composantes de $\chi_{u_k}(u_k)$ est un polynôme en les composantes de u_k, ce qui légitime le passage à la limite :

$$\chi_u(u) = 0. \tag{12.893}$$

Le théorème est établi pour toutes les matrices de $\mathbb{M}(n, \mathbb{C})$ et donc aussi pour tous les sous-corps de $\mathbb{C}$ comme $\mathbb{R}$ ou $\mathbb{Z}$.

La cas général Par définition, $\chi_u(X) = \det(u - X\mathbb{1})$; les coefficients de X sont des polynômes à coefficients entiers en les composantes de u. En substituant u à X nous obtenons une matrice dont chacune des entrées est un polynôme à coefficients entiers en les coefficients de u. Pour chaque i et j entre 1 et n il existe donc un polynôme $P_{ij} \in \mathbb{Z}(X_1, \ldots, X_{n^2})$ tel que

$$\chi_u(u)_{ij} = P(u_{11}, \ldots, u_{nn}). \tag{12.894}$$

Ces polynômes ne dépendent pas de u ni du corps sur lequel on travaille. Notre but est maintenant de prouver que $P_{ij} = 0$.

Étant donné que le cas complexe (et a fortiori entier) est déjà prouvé nous savons que pour tout $u \in \mathbb{M}(n, \mathbb{Z})$ nous avons $P(u_{11}, \ldots, u_{nn}) = 0$. La proposition 5.114 nous donne effectivement $P = 0$, en conséquence de quoi l'endomorphisme $\chi_u(u)$ est nul.

□

Exemple 12.340
Pour montrer que chaque composante $\chi_u(u)$ est bien un polynôme à coefficients entiers en les coefficients de u, voyons l'exemple 2×2 : $u = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. D'abord

$$\chi_u(X) = \det \begin{pmatrix} a - X & b \\ c & d - X \end{pmatrix} = X^2 - (a + d)X + ad - cb. \tag{12.895}$$

Le coefficient de X^2 est 1, celui de X est $-a - d$ et le terme indépendant est $ad - cb$; tout trois sont des polynômes à coefficients entiers en a, b, c, d. Après substitution de X par u,

$$\chi_u(u)_{ij} = (u^2)_{ij} - (a + d)u_{ij} + ad - cb. \tag{12.896}$$

Cela est bien un polynôme à coefficients entiers en les entrées de la matrice u. △

12.40.4 Racine carré d'une matrice hermitienne positive

Proposition 12.341.
Si $A \in \mathbb{M}(n, \mathbb{C})$ est une matrice hermitienne[57] *positive, alors il existe une unique matrice hermitienne positive R telle que $A = R^2$. De plus R est un polynôme (de $\mathbb{R}[X]$) en A.*

La matrice R ainsi définie est la **racine carré de** de A, et est notée $\sqrt{A}$. Une des applications usuelles de cette proposition est la décomposition polaire.

Démonstration. **Existence** Étant donné que A est hermitienne, elle est diagonalisable par une unitaire (proposition 9.162), et ses valeurs propres sont réelles et positives (parce que A est positive). Soit donc P une matrice unitaire telle que

$$P^*AP = \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix} \tag{12.897}$$

avec $\alpha_i > 0$. Si on pose

$$R = P \begin{pmatrix} \sqrt{\alpha_1} & & \\ & \ddots & \\ & & \sqrt{\alpha_n} \end{pmatrix} P^*, \tag{12.898}$$

alors $R^2 = A$ parce que $P^*P = \mathbb{1}$.

Hermitienne positive La matrice R est hermitienne parce que, avec un peu de notation raccourcie, $R = P^*\sqrt{\alpha}P$ et $R^* = P^*\sqrt{\alpha}P$. D'autre part, elle est positive parce que ses valeurs propres sont les $\sqrt{\alpha_i}$ qui sont positives.

Polynôme Nous montrons maintenant que la matrice R est un polynôme en A. Pour cela nous considérons un polynôme Q tel que $A(\alpha_i) = \sqrt{\alpha_i}$ pour tout i. Soit $\{e_i\}$ une base de

57. Définition 9.45.

diagonalisation de A : $Ae_i = \alpha_i e_i$. Alors c'est encore une base de diagonalisation de $Q(A)$. En effet si $Q = \sum_k a_k X^k$, alors

$$Q(A)e_i = (\sum_k a_k A^k)e_i = (\sum_k a_k \alpha_i^k)e_i = Q(\alpha_i)e_i = \sqrt{\alpha_i}e_i. \tag{12.899}$$

Les valeurs propres de $Q(A)$ sont donc $\sqrt{\alpha_i}$. Nous savons maintenant que $Q(A)$ a la même base de diagonalisation de A (et donc la même matrice unitaire P qui diagonalise), c'est à dire que

$$Q(A) = P^* \begin{pmatrix} \sqrt{\alpha_1} & & \\ & \ddots & \\ & & \sqrt{\alpha_n} \end{pmatrix} = R. \tag{12.900}$$

Donc oui, R est un polynôme en A.

Notons que ce Q n'est pas du tout unique ; il existe une infinité de polynômes qui envoient n nombres donnés sur n nombres donnés.

Unicité Soit S une matrice hermitienne positive telle que $R^2 = S^2 = A$. D'abord S commute avec A parce que

$$SA = S^3 = S^2S = AS. \tag{12.901}$$

Donc S commute aussi avec $Q(A) = R$. Étant donné que S et R commutent et sont diagonalisables, ils sont simultanément diagonalisables par le corollaire 9.144. Soient $D_R = PRP^*$ et $D_S = PSP^*$ les formes diagonales de R et S dans une base de simultanée diagonalisation. Les carrés des valeurs propres de R et S étant identiques (ce sont les valeurs propres de A) et les valeurs propres de R et S étant positives, nous déduisons que $D_R = D_S$ et donc que $R = P^*D_RP = P^*D_SP = S$.

□

12.40.5 Racine carré d'une matrice symétrique positive

Lemme 12.342 ([154])**.**
Le groupe orthogonal $\mathrm{O}(n, \mathbb{R})$ *est compact.*

Démonstration. Nous avons $O(n) = f^{-1}\big(\{\mathbb{1}_n\}\big)$ où f est l'application continue $A \mapsto A^tA$. En tant qu'image inverse d'un fermé par une application continue, le groupe $O(n)$ est fermé.

De plus il est borné parce que tous les coefficients d'une matrice orthogonale sont $\leqslant 1$, donc $\|A\|_\infty$ pour tout $A \in O(n)$. □

Proposition 12.343.
Une matrice symétrique semi (ou pas) définie positive admet une unique racine carré symétrique. Le spectre de la racine carré est la racine carré du spectre de la matrice de départ.

Démonstration. Ceci est une phrase pour que les titres se mettent bien.

Existence Soit T une matrice symétrique et Q une matrice orthogonale qui diagonalise[58] T : $QTQ^{-1} = D$ avec $D = \mathrm{diag}(\lambda_i)$ et $\lambda_i \geqslant 0$. En posant $R = Q^{-1}\sqrt{D}Q$, il est vite vérifié que $R^2 = T$ et que R est symétrique. En ce qui concerne le spectre, R a pour valeurs propres les $\sqrt{\lambda_i}$.

Unicité Soit R une matrice symétrique de T : $R^2 = T$. Du coup R et T commutent : $RT = R^3 = TR$. Par conséquent les espaces propres de T sont stables sous R. Soit E_λ l'un d'eux de dimension d, et T_F, R_F les restrictions de T et R à E_λ. L'application T_F est une homothétie et $R_F^2 = T_F = \lambda\mathbb{1}$. Mais R_F est encore une matrice symétrique définie positive, donc nous

58. Théorème 9.165.

pouvons considérer une base $\{e_1, \dots, e_d\}$ de E_λ qui diagonalise R_F avec les valeurs propres μ_i ; nous avons donc en même temps

$$R_f^2(e_i) = \mu_i^2 e_i \tag{12.902a}$$
$$T_F(e_i) = \lambda e_i, \tag{12.902b}$$

de telle sorte que $\mu_i^2 = \lambda$. Mais les valeurs propres de R_F sont positives, sont $\mu_i = \sqrt{\lambda}$ pour tout i. En conclusion R_F est univoquement déterminé par la donnée de T. Vu que cela est valable pour tous les espaces propres de T et que ces espaces propres engendrent tout E, l'opérateur R est déterminé de façon univoque par T.

□

Notons que nous n'avons démontré l'unicité qu'au sein des matrices symétriques.

12.40.6 Décomposition polaires : cas réel

Nous rappelons que $S^{++}(n, \mathbb{R})$ est l'ensemble des matrice symétriques strictement définies positives définies en 9.168.

Lemme 12.344.
La partie $S^+(n, \mathbb{R})$ est fermée dans $\mathbb{M}(n, \mathbb{R})$.

Démonstration. En effet si S_k est une suite de matrices symétriques convergeant dans $\mathbb{M}(n, \mathbb{R})$ vers la matrice A, les suites $(S_k)_{ij}$ et $(S_k)_{ji}$ des composantes ij et ji sont des suites égales, et donc leurs limites sont égales [59]. Donc la limite est symétrique.

En ce qui concerne le spectre, le théorème 9.165 nous permet de diagonaliser : $S_k = Q_k D_k Q_k^{-1}$ où les D_k sont des matrices diagonales remplies de nombres positifs ou nuls. Vu que $O(n)$ est compact [60], nous avons une sous-suite $Q_{\varphi(k)}$ convergente : $Q_{\varphi(k)} \to Q$. Pour chaque k, nous avons

$$S_{\varphi(k)} = Q_{\varphi(k)} D_{\varphi(k)} Q_{\varphi(k)}^{-1}, \tag{12.903}$$

dont la limite existe et vaut A. Vu que pour tout k, $D_{\varphi(k)} = Q_{\varphi(k)}^{-1} S_{\varphi(k)} Q_{\varphi(k)}$ et que le produit matriciel est continu, la suite $k \mapsto D_{\varphi(k)}$ est une suite convergente dans $\mathbb{M}(n, \mathbb{R})$. Nous notons D sa limite qui est encore une matrice diagonale contenant des nombres positifs ou nuls sur la diagonale.

$$A = \lim_{k\to\infty} S_{\varphi(k)} = QDQ^{-1}, \tag{12.904}$$

et donc le spectre de A est la limite de ceux des matrices $D_{\varphi(k)}$. Chacun étant positif, la limite est positive. Donc $A \in S^+(n, \mathbb{R})$. □

Lemme 12.345.
La fermeture de l'ensemble des matrice symétriques strictement définies positives est l'ensemble des matrices définies positives : $\overline{S^{++}(n, \mathbb{R})} = S^+(n, \mathbb{R})$.

Démonstration. Le lemme 12.344 nous a à peine dit que $S^+(n, \mathbb{R})$ était fermé. Nous devons prouver que pour tout élément de $S^+(n, \mathbb{R})$, il existe une suite (S_k) dans $S^{++}(n, \mathbb{R})$ convergeant vers S.

Si $S \in S^+(n, \mathbb{R})$ alors nous avons la diagonalisation

$$S = QDQ^{-1} = Q \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} Q^{-1} \tag{12.905}$$

59. Ici nous utilisons le critère de convergence composante par composante et le fait que nous ne sommes pas trop inquiétés par la norme que nous choisissons parce que toutes les normes sont équivalentes par le théorème 10.3.
60. Lemme 12.342.

où $\lambda_i \geqslant 0$ pour tout i. Nous définissons

$$D_k = \begin{pmatrix} \lambda_1 + \epsilon_k^{(1)} & & \\ & \ddots & \\ & & \lambda_n + \epsilon_k^{(n)} \end{pmatrix} \tag{12.906}$$

où ϵ_k^i est une suite convergent vers 0 telle que $\lambda_i + \epsilon_n^{(i)} > 0$ pour tout n. Typiquement si $\lambda_i > 0$ alors $\epsilon_k^{(i)} = 0$ et sinon $\epsilon_k^{(i)} = 1/k$.

Pour tout k nous avons $QD_kQ^{-1} \in S^{++}(n, \mathbb{R})$ et de plus $QD_kQ^{-1} \to QDQ = S$. □

Théorème 12.346 (Décomposition polaire de matrices symétriques définies positives[154, 155, 156])**.**
En ce qui concerne les matrices inversibles :

$$\begin{aligned} f\colon O(n, \mathbb{R}) \times S^{++}(n, \mathbb{R}) &\to \mathrm{GL}(n, \mathbb{R}) \\ (Q, S) &\mapsto SQ \end{aligned} \tag{12.907}$$

est un homéomorphisme[61]*.*

En ce qui concerne les matrices en général :

$$\begin{aligned} g\colon O(n, \mathbb{R}) \times S^{+}(n, \mathbb{R}) &\to \mathbb{M}(n, \mathbb{R}) \\ (Q, S) &\mapsto SQ \end{aligned} \tag{12.908}$$

est une surjection mais pas une injection.

De plus les mêmes conclusions tiennent si nous regardons $(Q, S) \mapsto QS$ au lieu de SQ.

Démonstration. Nous commençons par prouver les résultats concernant les matrices inversibles.

Existence et unicité Si $M = SQ$, alors $MM^t = SQQ^tS^t = S^2$, donc S doit être une racine carré symétrique de la matrice définie positive MM^t. La proposition 12.343 nous dit que ça existe et que c'est unique. Donc S est univoquement déterminé par M. Maintenant avoir $Q = MS^{-1}$ est obligatoire (unicité) et fonctionne :

$$Q^tQ = (S^{-1})^tM^tMS^{-1} = S^{-1}S^2S^{-1} = \mathbb{1}, \tag{12.909}$$

donc Q ainsi défini est orthogonale.

Notons que ceci ne fonctionne pas lorsque M n'est pas inversible parce qu'alors S n'est pas inversible.

Homéomorphisme Le fait que f soit continue n'est pas un problème : c'est un produit de matrice. Nous devons vérifier que f^{-1} est continue. Soit une suite convergente $M_k \to M$ dans $\mathrm{GL}(n, \mathbb{R})$. Si nous nommons (Q_k, S_k) la décomposition polaire de M_k et (Q, S) celle de M, nous devons prouver que $Q_k \to Q$ et $S_k \to S$. En effet dans ce cas nous aurions

$$\lim_{k\to\infty} f^{-1}(M_k) = \lim_{k\to\infty} (Q_k, S_k) = (Q, S) = f^{-1}(M). \tag{12.910}$$

Étant donné que $O(n)$ est compact (lemme 12.342), la suite (Q_k) admet une sous-suite convergente (Bolzano-Weierstrass, théorème 6.147) que nous nommons

$$Q_{\varphi(k)} \to F \in O(n). \tag{12.911}$$

Vu que la suite (M_k) converge, sa sous-suite converge vers la même limite : $M_{\varphi(k)} \to M$ et vu que pour tout k nous avons $S_k = M_kQ_k^{-1}$,

$$S_{\varphi(k)} \to G = MF^{-1}. \tag{12.912}$$

61. Cela est en réalité en difféomorphisme, voir la remarque 12.347.

Vu que chacune des matrices $S_{\varphi(k)}$ est symétrique définie positive, la limite est symétrique et semi-définie positive [62]. Donc $G \in S^+(n,\mathbb{R}) \cap \mathrm{GL}(n,\mathbb{R})$ parce que de plus M et F étant inversibles, G est inversible. En ce qui concerne la sous-suite nous avons

$$M_{\varphi(k)} = S_{\varphi(k)}Q_{\varphi(k)} \to GF = M \tag{12.913}$$

où $F \in O(n)$ et $G \in S^+(n,\mathbb{R})$. Par unicité de la décomposition polaire de M (partie déjà démontrée), nous avons $G = S$ et $F = Q$.

Nous avons prouvé que toute sous-suite convergente de Q_k a Q pour limite. Donc la suite elle-même converge [63] vers Q. Donc $Q_k \to Q$. Du coup vu que $S_k = M_k Q_k^{-1}$ est un produit de suites convergentes, S_k converge également, vers S : $S_k \to S$.

Au final l'application f^{-1} est bien continue parce que les égalités (12.910) ont bien lieu.

Nous passons maintenant à la preuve dans le cas des matrices en général.

Soit $A \in \mathbb{M}(n,\mathbb{R})$; par densité (lemme 12.337), il existe une suite (A_k) dans $\mathrm{GL}(n,\mathbb{R})$ telle que $A_k \to A$. Pour chacun des k nous appliquons la décomposition polaire déjà prouvée : $A_k = Q_k S_k$. D'abord (Q_k) est une suite dans le compact [64] $\mathrm{O}(n,\mathbb{R})$ et accepte donc une sous-suite convergente. Quitte à redéfinir la suite de départ, nous supposons pour alléger les notations que $Q_k \to Q \in \mathrm{O}(n,\mathbb{R})$. Vu que Q_k est inversible,

$$S_k = Q_k^{-1}A_k \tag{12.914}$$

Le produit matriciel étant continu nous avons $S_k \to S$ dans $\mathbb{M}(n,\mathbb{R})$. Mais $S^+(n,\mathbb{R})$ étant fermé (lemme 12.344) nous avons aussi $S \in S^+(n,\mathbb{R})$. □

Remarque 12.347.
Pour démontrer que f est différentiable, nous devons utiliser le théorème d'inversion locale 18.47 ; cela est fait dans la proposition 18.56.

Corollaire 12.348.
Toute matrice peut être écrite sous la forme Q_1DQ_2 où Q_1 et Q_2 sont orthogonale et D est diagonale.

Démonstration. Si $A \in \mathbb{M}(n,\mathbb{R})$ alors la décomposition polaire 12.346 nous donne $A = SQ$ où S est symétrique définie positive et Q est orthogonale. La matrice S peut ensuite être diagonalisée par le théorème 9.165 : $S = RDR^{-1}$ où D est diagonale et R est orthogonale. Avec ces deux décompositions en main, $A = SQ = RDR^{-1}Q$. La matrice $R^{-1}Q$ est orthogonale. □

12.40.7 Enveloppe convexe

Définition 12.349.
*Sur C est un ensemble convexe, un point $x \in C$ est un **point extrémal** si $C\backslash\{x\}$ est encore convexe.*

Théorème 12.350 ([2])**.**
Soit E un espace euclidien de dimension $n \geqslant 1$ et $\mathcal{L}(E)$ l'espace des opérateurs linéaires sur E sur lequel nous considérons la norme subordonnée [65] à celle sur E. L'ensemble des points extrémaux de la boule unité fermée de $\mathcal{L}(E)$ est le groupe orthogonal $O(n,\mathbb{R})$.

Démonstration. Nous notons $\mathcal{B}$ la boule unité fermée de $\mathcal{L}(E)$. Montrons pour commencer que les éléments de $O(n)$ sont extrémaux dans $\mathcal{B}$. D'abord si $A \in O(E)$ alors $\|A\| = 1$ parce que $\|Ax\| = \|x\|$. Supposons maintenant que A n'est pas extrémal, c'est à dire qu'il est le milieu d'un

62. Lemme 12.345
63. Proposition 6.155, pas difficile.
64. Lemme 12.342.
65. Définition 10.5.

segment joignant deux points (distincts) de la boule unité de $\mathcal{L}(E)$. Soient donc $T, U \in \mathcal{B}$ tels que $A = \frac{1}{2}(T+U)$. Pour tout $x \in E$ tel que $\|x\| = 1$ nous avons

$$1 = \|x\| = \|Ax\| = \frac{1}{2}\|Tx + Ux\| \leqslant \frac{1}{2}\big(\|Tx\| + \|Ux\|\big) \leqslant \frac{1}{2}\big(\|T\| + |U|\big) \leqslant 1 \tag{12.915}$$

Toutes les inégalités sont en réalité des égalités. En particulier nous avons

$$\|Tx + Ux\| = \|Tx\| + \|Ux\|, \tag{12.916}$$

mais alors nous sommes dans un cas d'égalité dans l'inégalité de Cauchy-Schwartz (théorème 9.25) et donc il existe $\lambda \geqslant 0$ tel que $Tx = \lambda Ux$. Mais de plus les ~~inégalité~~ égalités (12.915) nous donnent

$$\frac{1}{2}\big(\|Tx\| + \|Ux\|\big) = 1 \tag{12.917}$$

alors que nous savons que $\|Tx\|, \|Ux\| \leqslant 1$, donc $\|Tx\| = \|Ux\| = 1$. La seule possibilité est d'avoir $\lambda = 1$ et donc que $U = T$ parce que nous avons $Tx = Ux$ pour tout x de norme 1. Au final A n'est pas le milieu d'un segment dans $\mathcal{B}$.

Nous passons donc à l'inclusion inverse : nous prouvons que les points extrémaux de $\mathcal{B}$ sont dans $O(E)$. Pour cela nous prenons $U \in \mathcal{B}\backslash O(E)$ et nous allons montrer que U n'est pas un point extrémal : nous allons l'écrire comme milieu d'un segment dans $\mathcal{B}$.

Par la seconde partie du théorème de décomposition polaire 12.346, il existe $Q \in O(n, \mathbb{R})$ et $S \in S^+(n, \mathbb{R})$ tels que $U = QS$. Nous diagonalisons S à l'aide de la matrice orthogonale P :

$$S = PDP^{-1} \tag{12.918}$$

avec $D = \operatorname{diag}(\lambda_i)$. En termes de normes, nous avons

$$\|U\| = \|S\| = \|S\|. \tag{12.919}$$

En effet vu que Q est orthogonale, $\|Ux\| = \|QSx\| = \|Sx\|$ pour tout x, donc $\|U\| = \|S\|$. De plus pour tout x nous avons

$$\|Sx\| = \|PDP^{-1}x\| = \|DP^{-1}x\|. \tag{12.920}$$

Étant donné que P^{-1} est une bijection, le supremum des $\|Sx\|$ sera le même que celui des $\|Dx\|$ et donc $\|S\| = \|D\|$. Étant donné que par définition $\|U\| \leqslant 1$, nous avons aussi $\|D\| \leqslant 1$ et donc $0 \leqslant \lambda_i \leqslant 1$ (pour rappel, les valeurs propres de D sont positives ou nulles parce que S est ainsi).

Comme $U \notin O(E)$, au moins une des valeurs propres n'est pas 1, supposons que ce soit λ_1. Alors nous avons $\alpha, \beta \in [-1, 1]$ avec $-1 \leqslant \alpha < \beta \leqslant 1$ et $\lambda_1 = \frac{1}{2}(\alpha + \beta)$. Nous posons alors

$$D_1 = \operatorname{diag}(\alpha, \lambda_2, \ldots, \lambda_n) \tag{12.921a}$$
$$D_2 = \operatorname{diag}(\beta, \lambda_2, \ldots, \lambda_n). \tag{12.921b}$$

Nous avons bien $D_1 \neq D2$ et $D_1 + D_2 = D$. Par conséquent

$$U = \frac{1}{2}\big(QPD_1P^{-1} + QPD_2P^{-1}\big) \tag{12.922}$$

avec $QPD_1P^{-1} \neq QPD_2^{-1}$. La matrice U est donc le milieu d'un segment. Reste à montrer que ce segment est dans $\mathcal{B}$. Pour ce faire, prenons $x \in E$ et calculons :

$$\|QPD_iP^{-1}x\| = \|D_iP^{-1}x\| \leqslant \|P^{-1}x\| = \|x\| \tag{12.923}$$

parce que $\|D_i\| \leqslant 1$ et P^{-1} est orthogonale. Au final la norme de QPD_iP est plus petite que 1 et donc U est bien le milieu d'un segment dans $\mathcal{B}$, et donc non extrémal. □

Théorème 12.351 ([157]).
L'enveloppe convexe de $O(n)$ dans $\mathbb{M}_n(\mathbb{R})$ est la boule unité pour la norme induite de $\|.\|_2$ sur $\mathbb{R}^n$.

Démonstration. Nous notons $\mathcal{B}$ la boule unité fermée de $\mathbb{M}(n,\mathbb{R})$ et $\mathcal{C}\text{onv}\big(O(n,\mathbb{R})\big)$ l'enveloppe convexe de $O(n,\mathbb{R})$. Vu que $\mathcal{B}$ est convexe nous avons $\mathcal{C}\text{onv}\big(O(n)\big) \subset \mathcal{B}$.

Maintenant nous devons prouver l'inclusion inverse. Pour ce faire nous supposons avoir un élément $A \in \mathcal{B} \backslash \mathcal{C}\text{onv}\big(O(n)\big)$ et nous allons dériver une contradiction.

Remarquons que $O(n)$ est compact par le lemme 12.342 et que par conséquent $\mathcal{C}\text{onv}(O(n))$ est compacte par le corollaire 11.36 et donc fermée. Nous considérons un produit scalaire $(X,Y) \mapsto X \cdot Y$ sur $\mathbb{M}$. Vu que $\mathcal{C}\text{onv}\big(O(n)\big)$ est un fermé convexe nous pouvons considérer la projection [66] sur $\mathcal{C}\text{onv}(A)$ relativement au produit scalaire choisis.

Nous notons $P = \texttt{proj}_{\mathcal{C}\text{onv}\left(O(n)\right)}(A)$. En vertu du théorème de projection, nous avons

$$(A-P) \cdot (M-P) \leqslant 0 \tag{12.924}$$

pour tout $M \in \mathcal{C}\text{onv}\, O(n)$. Notons $B = A - P$ pour alléger les notations. L'équation (12.924) s'écrit

$$B \cdot M \leqslant B \cdot P. \tag{12.925}$$

D'autre par vu que $B \neq 0$ nous avons $B \cdot B > 0$, c'est à dire $B \cdot (A-P) > 0$ et donc

$$B \cdot A > B \cdot P. \tag{12.926}$$

En combinant avec (12.925),

$$B \cdot M \leqslant B \cdot P < B \cdot A. \tag{12.927}$$

Nous utilisons maintenant la décomposition polaire, théorème 12.346, pour écrire $B = QS$ avec $Q \in O(n)$ et $S \in S^+(n,\mathbb{R})$. Vu que l'inégalité (12.927) tient pour tout $M \in \mathcal{C}\text{onv}(O(n))$, elle tient en particulier pour $Q \in O(n)$. Donc

$$B \cdot Q = B \cdot A. \tag{12.928}$$

Nous nous particularisons à présent au produit scalaire $(X,Y) \mapsto \text{Tr}(X^tY)$ de la proposition 10.27. D'abord

$$B \cdot Q = \text{Tr}(B^tQ) = \text{Tr}(S^tQ^tQ) = \text{Tr}(S^t) = \text{Tr}(S), \tag{12.929}$$

et ensuite l'inégalité (12.929) devient

$$\text{Tr}(S) < B \cdot A = \text{Tr}(S^tQ^tA). \tag{12.930}$$

Nous choisissons une basse $\{e_i\}$ diagonalisant S : $Se_i = \lambda_i e_i$ vérifiant automatiquement $\lambda_i \geqslant 0$ parce que S est semi-définie positive [67]. Alors

$$\begin{aligned}
\text{Tr}(S) &< \text{Tr}(S^tQ^tA) && (12.931a)\\
&= \sum_i \langle S^tQ^tAe_i, e_i\rangle && (12.931b)\\
&= \sum_i \langle Ae_i, QSe_i\rangle && (12.931c)\\
&\leqslant \sum_i \|Ae_i\||\lambda_i| \underbrace{\|Qe_i\|}_{=1} && (12.931d)\\
&\leqslant \sum_i \lambda_i && A \in \mathcal{B} \Rightarrow \|Ae_i\| \leqslant 1 \quad (12.931e)\\
&= \text{Tr}(S). && (12.931f)
\end{aligned}$$

Il faut noter que la première inégalité est stricte, et donc nous avons une contradiction. □

66. Le théorème de projection : théorème 12.233.
67. Définition 9.167.

12.40.8 Décomposition de Bruhat

Théorème 12.352 (Décomposition de Bruhat)**.**
Soit $\mathbb{K}$ un corps ; un élément $M \in \mathrm{GL}(n,\mathbb{R})$ s'écrit sous la forme

$$M = T_1 P_\sigma T_2 \tag{12.932}$$

où T_1 et T_2 sont des matrices triangulaires supérieures inversibles et où P_σ est une matrice de permutation $\sigma \in S_n$. De plus il y a unicité de σ.

Démonstration. Afin de rendre les choses plus visuelles, nous nous permettons de donner des exemples au fur et à mesure de la preuve. Nous prenons l'exemple de la matrice

$$\begin{pmatrix} 1 & 3 & 4 \\ 2 & 5 & 6 \\ 0 & 7 & 8 \end{pmatrix}. \tag{12.933}$$

Existence Soit $M \in \mathrm{GL}(n,\mathbb{R})$; vu qu'elle est inversible, on a un indice i_1 maximum tel que $M_{i_1,1} \neq 0$. Nous changeons toutes les lignes jusque là, c'est à dire que nous faisons, pour $1 \leqslant i < i_1$,

$$L_i \to L_i - \frac{M_{i1}}{M_{i_1 1}} L_{i_1}. \tag{12.934}$$

Voir le lemme 12.325(3).

Nous avons donc obtenu une matrice dont la première colonne est nulle sauf la case numéro i_1. L'opération (12.934) revient à considérer la multiplication par la matrice de transvection

$$T_1^{(i)} = T_{ii_1}\left(-\frac{M_{i1}}{M_{i_1 1}}\right) \tag{12.935}$$

pour tout $i < i_1$. Pour rappel nous ne changeons que les lignes *au-dessus* de la i_1. Du coup les matrices $T_1^{(i)}$ sont triangulaires supérieures. Nous avons donc la nouvelle matrice $M_1 = \left(\prod_{i<i_1} T_1^{(i)}\right) M$ pour laquelle toute la première colonne est nulle sauf un élément.

Dans le cas de l'exemple, le «pivot» sera la ligne $(2,5,6)$ et la matrice se transforme à l'aide de la matrice $T_1 = T_{12}(-1/2)$:

$$\begin{pmatrix} 1 & -1/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 3 & 4 \\ 2 & 5 & 6 \\ 0 & 7 & 8 \end{pmatrix} = \begin{pmatrix} 0 & 1/2 & 1 \\ 2 & 5 & 6 \\ 0 & 7 & 8 \end{pmatrix}. \tag{12.936}$$

Maintenant nous faisons de même avec les colonnes (en renommant M la matrice obtenue à l'étape précédente) :

$$C_j \to C_j - \frac{M_{i_1 j}}{M_{i_1 1}} C_1, \tag{12.937}$$

qui revient à multiplier à droite par les matrices $T_{1j}(\frac{M_{i_1 i}}{M_{i_1 1}})$ avec $j > 1$. Encore une fois ce sont des matrices triangulaires supérieures.

Dans l'exemple, pour traiter la seconde colonne, nous multiplions (12.936) à droite par la matrice $T_{12}(-5/2)$:

$$\begin{pmatrix} 0 & 1/2 & 1 \\ 2 & 5 & 6 \\ 0 & 7 & 8 \end{pmatrix}\begin{pmatrix} 1 & -5/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1/2 & 1 \\ 2 & 0 & 6 \\ 0 & 7 & 8 \end{pmatrix}. \tag{12.938}$$

Appliquer encore la matrice $T_{13}(-6/2)$ apporte la matrice

$$\begin{pmatrix} 0 & 1/2 & 1 \\ 2 & 0 & 0 \\ 0 & 7 & 8 \end{pmatrix}. \tag{12.939}$$

Enfin nous multiplions la matrice obtenue par $\frac{1}{M_{i_1 1}}\mathbb{1}$ pour normaliser à 1 l'élément «pivot» que nous avions choisit. Dans notre exemple nous multiplions par 1/2 pour trouver

$$\begin{pmatrix} 0 & 1/4 & 1/2 \\ 1 & 0 & 0 \\ 0 & 7/2 & 4 \end{pmatrix}. \tag{12.940}$$

La matrice obtenue jusqu'ici possède une ligne et une colonne de zéros avec un 1 à leur intersection, et elle est de la forme

$$M' = T_1 M T_2 \tag{12.941}$$

où T_1 et T_2 sont triangulaires supérieures et inversibles, produits de matrices de transvection (et d'une matrice scalaire pour la normalisation).

Il reste à recommencer l'opération avec la seconde colonne (qui n'est pas toute nulle parce que le déterminant est encore non nul) puis la suivante etc. Dans notre exemple de l'équation (12.940), nous éliminerions le 1/4 et le 4 en utilisant le 7/2.

Encore une fois tout cela se fait à l'aide de matrice supérieures parce qu'à chaque étape, les colonnes précédent le pivot sont déjà nulles (saut un 1) et ne doivent donc pas être touchées.

À la fin de ce processus, ce qui reste est une matrice TMT' qui ne contient plus que un seul 1 sur chaque ligne et chaque colonne, c'est à dire une matrice de permutation : $P_\sigma = TMT'$ et donc

$$M = T_\sigma^{-1}(T')^{-1}. \tag{12.942}$$

Unicité Soient $\sigma, \sigma \in S_n'$ tels que $T_1 P_\sigma T_2 = S_1 P_\tau S_2$ avec T_i et S_i triangulaires supérieures et inversibles. En posant $T = T_2 S_2^{-1}$ et $S = T_1^{-1} S_1$, nous avons

$$P_\sigma T = S P_\tau \tag{12.943}$$

où S et T sont des matrices triangulaires supérieures et inversibles. Par les calculs de la preuve du lemme 4.98,

$$\begin{cases} (P_\sigma T)_{kl} = T_{\sigma^{-1}(k)l} & \text{(12.944a)} \\ (SP_\tau)_{kl} = S_{k\tau(l)}, & \text{(12.944b)} \end{cases}$$

et donc

$$T_{\sigma^{-1}(k)l} = S_{k\tau(l)}. \tag{12.945}$$

En écrivant cette équation avec $k = \sigma(i)$ (nous rappelons que σ est bijective),

$$T_{il} = S_{\sigma(i)\tau(l)}. \tag{12.946}$$

Nous savons que les termes diagonaux de T sont non nuls parce que T est triangulaire supérieure et inversible (donc pas de colonnes entières nulles). Nous avons donc, en prenant $i = l = k$,

$$0 \neq T_{kk} = S_{\sigma(k)\tau(k)}. \tag{12.947}$$

La matrice étant triangulaire supérieure, cela implique

$$\sigma(k) \leqslant \tau(k). \tag{12.948}$$

De la même manière en écrivant (12.945) avec $l = \tau^{-1}(i)$,

$$S_{ki} = T_{\sigma^{-1}(k)\tau^{-1}(i)} \tag{12.949}$$

et donc

$$\sigma^{-1}(k) \leqslant \tau^{-1}(k). \tag{12.950}$$

En écrivant cela avec $k = \sigma(j)$, nous avons $j \leqslant \tau^{-1}\sigma(j)$ et en appliquant enfin τ,

$$\tau(j) \leqslant \sigma(j). \tag{12.951}$$

En comparant avec (12.948), nous avons $\sigma = \tau$.

□

12.41 Sous-groupes du groupe linéaire

Lemme 12.353 ([2]).
Soit V un espace vectoriel de dimension finie muni d'une norme euclidienne $\|.\|$. Soit K un compact convexe de V et G, un sous groupe compact de $\mathrm{GL}(V)$ tel que

$$u(K) \subset K \tag{12.952}$$

pour tout $u \in G$. Alors il existe $a \in K$ tel que $u(a) = a$ pour tout $u \in G$.

Démonstration. Avant de nous lancer dans la preuve, nous avons besoin d'un petit résultat.

Un pré-résultat Nous commençons par prouver que si $v \in \mathcal{L}(V)$ vérifie $v(K) \subset K$, alors v a un point fixe dans K. Pour cela nous considérons $x_0 \in K$ et la suite

$$x_k = \frac{1}{k+1} \sum_{i=0}^{k} v^i(x_0). \tag{12.953}$$

Étant donné que K est convexe et stable par v, la suite (x_k) est contenue dans K et accepte une sous-suite convergente [68] que nous allons noter $x_{\varphi(n)}$ avec $\varphi\colon \mathbb{N} \to \mathbb{N}$ strictement croissante. Soit $a \in K$ la limite :

$$\lim_{n\to\infty} x_{\varphi(n)} = a. \tag{12.954}$$

Tant que nous y sommes nous pouvons aussi calculer $v(x_k)$:

$$\begin{aligned} v(x_k) &= v\left(\frac{1}{k+1} \sum_{i=1}^{k} v^i(x_0) \right) && (12.955a)\\ &= \frac{1}{k+1} \sum_{i=0}^{k} v^{i+1}(x_0) && (12.955b)\\ &= x_k + \frac{1}{k+1}\Big(v^{k+1}(x_0) - x_0 \Big). && (12.955c) \end{aligned}$$

La norme $\|v^{k+1}(x_0) - x_0\|$ est bornée par le diamètre de K, donc en prenant la limite $k \to \infty$ le second terme de (12.955c) tend vers zéro. En prenant ces égalités en $k = \varphi(n)$ et en prenant $n \to \infty$, nous trouvons

$$v(a) = a, \tag{12.956}$$

c'est à dire le résultat que nous voulions dans un premier temps.

Une norme sur V Nous passons maintenant à la preuve du lemme. D'abord nous remarquons que le groupe G agit sur V par $u \cdot x = u(x)$ et de plus, considérant la fonction continue

$$\begin{aligned} \alpha\colon G &\to V \\ u &\mapsto u(x), \end{aligned} \tag{12.957}$$

nous voyons que les orbites de cette action sont compactes en tant qu'image par α du compact G (théorème 6.77). Nous posons

$$\begin{aligned} \nu\colon V &\to \mathbb{R}^+ \\ x &\mapsto \max_{u\in G} \|u(x)\|. \end{aligned} \tag{12.958}$$

Cette définition a un sens parce que l'orbite $\{u(x) \text{ tel que } u \in G\}$ est compacte dans V et donc l'ensemble des normes est compact dans $\mathbb{R}$ et admet un maximum. De plus cela donne une norme sur V parce que nous vérifions les conditions de la définition 6.31 :

68. C'est Bolzano-Weierstrass, théorème 6.147.

(1) Pour tout $x, y \in V$ nous avons :

$$\nu(x+y) = \max_{u\in G} \|u(x)+u(y)\| \leqslant \max_{u\in G} (\|u(x)\| + \|u(y)\|) \leqslant \nu(x)+\nu(y). \tag{12.959}$$

(2) Si $\nu(x) = 0$, alors l'égalité $\max_{u\in G} \|u(x)\| = 0$ nous enseigne que $\|u(x)\| = 0$ pour tout $u \in G$ et donc en particulier avec $u = \mathrm{Id}$ nous trouvons $x = 0$.

(3) Pour tout $\lambda \in \mathbb{R}$ et $x \in V$,

$$\nu(\lambda x) = \max_{u\in G} \|u(\lambda x)\| = \max \|\lambda u(x)\| = \max |\lambda| \|u(x)\| = |\lambda|\nu(x). \tag{12.960}$$

De plus la fonction ν est constante sur les orbites de G.

Un point fixe Pour tout $u \in G$ nous posons

$$F_u = \{x \in K \text{ tel que } u(x) = x\}; \tag{12.961}$$

par le pré-résultat, aucun de ces ensembles n'est vide. Ils sont de plus tous fermés par continuité de u (le complémentaire est ouvert). Nous devons prouver que $\bigcap_{u\in G} F_u \neq \varnothing$ parce qu'une intersection serait un point fixe de tous les éléments de G. Supposons donc que $\bigcap_{u\in G} F_u = \varnothing$. Alors les complémentaires des F_u forment un recouvrement ouvert de K et nous pouvons en extraire un sous-recouvrement fini par compacité. Soient $\{u_i\}_{i=1,\dots,p}$ les éléments qui réalisent ce recouvrement. Alors

$$\bigcap_{i=1}^{p} F_{u_i} = \varnothing. \tag{12.962}$$

Nous considérons l'opérateur

$$v = \frac{1}{p}\sum_{i=1}^{p} u_i \in \mathcal{L}(V). \tag{12.963}$$

Vu que K est convexe et stable sous chacun des u_i, nous avons aussi $v(K) \subset K$ et donc il existe $a \in K$ tel que $v(a) = a$. Pour ce a, nous avons

$$\begin{aligned}
\nu\big(v(a)\big) &= \nu\left(\frac{1}{p}\sum_{i=1}^{p} u_i(a)\right) && \text{(12.964a)}\\
&\leqslant \frac{1}{p}\sum_{i=1}^{p} \nu\left(u_i(a)\right) && \text{(12.964b)}\\
&= \frac{1}{p}\sum_{i=1}^{p} \nu(a) && \text{(12.964c)}\\
&= \nu(a) && \text{(12.964d)}
\end{aligned}$$

où nous avons utilisé la constance de ν sur les orbites de G. Par ailleurs nous savons que $v(a) = a$, donc en réalité à gauche dans (12.964a) nous avons $\nu(a)$ et toutes les inégalités sont des égalités. Nous avons en particulier

$$\nu\left(\sum_{i=1}^{p} u_i(a)\right) = \sum_{i=1}^{p} \nu\left(u_i(a)\right). \tag{12.965}$$

Notons $u_0 \in G$ l'élément qui réalise le maximum de la définition de ν pour le vecteur $\sum_i u_i(a)$:

$$\nu\left(\sum_i u_i(a)\right) = \|u_0\left(\sum_i u_i(a)\right)\| \leqslant \sum_i \|u_0 u_i(a)\| \leqslant \sum_i \nu\big(u_i(a)\big). \tag{12.966}$$

Mais nous venons de voir (équation (12.965)) que l'expression de gauche est égale à celle de droite. Donc les inégalités sont des égalités et en particulier la première inégalité devient l'égalité

$$\| \sum_i u_0 u_i(a) \| = \sum_i \| u_0 u_i(a) \|. \tag{12.967}$$

En vertu du lemme 9.30, il existe des nombres positifs λ_i tels que

$$u_0 u_1(a) = \lambda_2 u_0 u_2(a) = \ldots = \lambda_p u_0 u_p(a). \tag{12.968}$$

Du fait que u_0 est inversible nous avons aussi

$$u_1(a) = \lambda_2 u_2(a) = \ldots = \lambda_p u_p(a). \tag{12.969}$$

Mais par constance de ν sur les orbites nous avons $\nu(u_i(a)) = \nu(u_j(a))$ pour tout i et j ; en appliquant ν à la série d'égalités (12.969), nous trouvons que tous les λ_i doivent être égaux à 1. En particulier

$$u_1(a) = u_2(a) = \ldots = u_p(a). \tag{12.970}$$

Nous récrivons maintenant l'équation $v(a) = a$ avec la définition de v :

$$a = v(a) = \frac{1}{p} \sum_{i=1}^{p} u_i(a) = u_j(a) \tag{12.971}$$

pour n'importe quel j. Donc

$$a \in \bigcap_{i=1}^{p} F_{u_i}, \tag{12.972}$$

ce qui contredit notre hypothèse de départ.

□

Proposition 12.354 ([157, 2, 158])**.**
Soit G un sous-groupe compact de $\mathrm{GL}(n, \mathbb{R})$. Alors

(1) Il existe une forme quadratique définie positive q sur $\mathbb{R}^n$ telle que $G \subset \mathrm{O}(q)$.

(2) Le groupe G est conjugué à un sous-groupe de $\mathrm{O}(n, \mathbb{R})$.

Démonstration. Nous considérons le (pas tout à fait) morphisme de groupe

$$\begin{aligned} \rho \colon G &\to \mathrm{GL}\big(\mathrm{S}(n, \mathbb{R})\big) \\ u &\mapsto \rho_u \colon s \to u^t s u, \end{aligned} \tag{12.973}$$

et tant que nous y sommes à considérer, nous considérons l'ensemble

$$H = \{M^t M \text{ tel que } M \in G\} \subset \mathrm{S}(n, \mathbb{R}). \tag{12.974}$$

Cet ensemble est constitué de matrices définies positives parce que si $\langle M^t M x, x \rangle = 0$, alors $0 = \langle Mx, Mx \rangle = \|Mx\|$, mais M étant inversible, cela implique que $x = 0$. Qui plus est cet ensemble est compact dans $\mathrm{GL}(n, \mathbb{R})$ en tant qu'image du compact G par l'application continue $M \mapsto M^t M$. L'enveloppe convexe $K = \mathcal{C}\mathrm{onv}(H)$ est alors également compacte par le théorème 11.36. Enfin nous considérons $L = \rho(G)$, qui est un sous-groupe compact de $\mathrm{GL}\big(\mathrm{S}(n, \mathbb{R})\big)$ parce que $\rho_u \rho_v = \rho_{vu} \in \rho(G)$. Nous remarquons que ρ_u étant linéaire, elle préserve les combinaisons convexes et donc pour tout $u \in G$, $\rho_u(K) \subset K$.

Bref, L est un sous-groupe compact de $\mathrm{GL}(n, \mathbb{R})$ préservant le compact K de $\mathrm{S}(n, \mathbb{R})$. Par le lemme 12.353, il existe $s \in K$ tel que $\rho_u(s) = s$ pour tout $u \in G$. Ou encore :

$$u^t s u = s \tag{12.975}$$

pour tout $u \in G$. Fort de ce s bien particulier, nous considérons la forme quadratique associée : $q(x) = x^t s x$. Cette forme est définie positive parce que s l'est. Nous avons $G \subset \mathrm{O}(q)$ parce que si $u \in G$ alors

$$q\big(ux\big) = (ux)^t sux = x^t \underbrace{u^t s u}_{=s} x = q(x). \tag{12.976}$$

Le premier point est prouvé.

La matrice s est symétrique et définie positive. Elle peut donc être diagonalisée [69] en $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ avec $\lambda_i > 0$, et ensuite transformée en la matrice $\mathbb{1}_n$ par la matrice $\mathrm{diag}(1/\sqrt{\lambda_i})$. Nous avons donc une matrice $a \in \mathrm{GL}(n, \mathbb{R})$ telle que $a^t s a = \mathbb{1}_n$. Avec ça, si $u \in G$, nous avons

$$(a^{-1}ua)^t(a^{-1}ua) = (a^{-1}ua)^t \mathbb{1}_n (a^{-1}ua) = a^t u^t (a^t)^{-1} a^t s a a^{-1} u a = a^t u^t s u a = a^t s a = \mathbb{1}, \tag{12.977}$$

ce qui prouve que $a^{-1}ua$ est dans $\mathrm{O}(n, \mathbb{R})$, et donc que $a^{-1}Ga \subset \mathrm{O}(n, \mathbb{R})$. □

69. Théorème 9.165

Chapitre 13

Tribus, théorie de la mesure

13.1 Tribus

Vous pouvez voir le thème 24 pour voir plus vite où sont les définitions associées.

13.1.1 Généralités

Définition 13.1 (Tribu, espace mesurable[159])**.**
Si Ω est un ensemble, un ensemble $\mathcal{A}$ de sous-ensembles de Ω est une ***tribu*** *si*

(1) $\Omega \in \mathcal{A}$ *;*

(2) $\complement A \in A$ *pour tout* $A \in \mathcal{A}$ *;*

(3) si $(A_i)_{i\in\mathbb{N}}$ *une suite dénombrable d'éléments de* $\mathcal{A}$*, alors* $\bigcup_{i\in\mathbb{N}} A_i \in \mathcal{A}$*.*

Le couple $(\Omega, \mathcal{A})$ *est alors un* ***espace mesurable****.*

Remarque 13.2.
Nous trouvons parfois la notation

$$\bigcup_{k\in\mathbb{N}} A_k = \sup_{k\geqslant 0} A_k. \tag{13.1}$$

Lemme 13.3.
Opérations ensemblistes sur les tribus.

(1) Une tribu est stable par intersections au plus dénombrables.

(2) Une tribu est stable par différence ensembliste.

Démonstration. Soit $(A_i)_{i\in I}$ une famille au plus dénombrable d'éléments de la tribu $\mathcal{A}$. Nous devons prouver que $\bigcap_{i\in I} A_i$ est également un élément de $\mathcal{A}$. Pour cela nous passons au complémentaire :

$$\complement\left(\bigcap_{i\in I} A_i\right) = \bigcup_{i\in I} \complement A_i. \tag{13.2}$$

La définition d'une tribu implique que le membre de droite est un élément de la tribu. Par stabilité d'une tribu par complémentaire, l'ensemble $\bigcap_{i\in I} A_i$ est également un élément de la tribu.

La seconde assertion est immédiate à partir de la première parce que $A\backslash B = A \cap \complement B$. □

Si $(\mathcal{A}_i)_{i\in I}$ est un ensemble de tribus (indexé par un ensemble I quelconque) alors

$$\mathcal{A} = \bigcap_{i\in I} \mathcal{A}_i \tag{13.3}$$

est également une tribu.

Définition 13.4.
*Soit $\mathcal{D}$ un ensemble de parties de Ω. La **tribu engendrée** par $\mathcal{D}$ est l'intersection de toutes les tribus de Ω contenant $\mathcal{D}$. C'est la plus petite tribu contenant $\mathcal{D}$. Nous la noterons le plus souvent $\sigma(\mathcal{A})$*

Note : une tribu engendrée par une application sera la définition 13.69.

Proposition 13.5 ([160])**.**
Soit S un ensemble et $\mathcal{F}$ une tribu de S. Soit une classe $\mathcal{N} \subset \mathcal{P}(S)$ telle que

(1) Si $A \in \mathcal{N}$ alors il existe $Y \in \mathcal{F} \cap \mathcal{N}$ tel que $A \subset Y$.

(2) Si $A \in \mathcal{N}$ et $B \subset A$ alors $B \in \mathcal{N}$.

(3) La classe $\mathcal{N}$ est stable par union dénombrable.

Alors la classe

$$\mathcal{T} = \{X \cup A \text{ avec } A \in \mathcal{N} \text{ et } X \in \mathcal{F}\} \tag{13.4}$$

est une tribu.

Démonstration. L'ensemble $\mathcal{T}$ est stable par union dénombrable parce que $\mathcal{F}$ et $\mathcal{N}$ le sont. De plus S et $\varnothing$ sont dans $\mathcal{F}$ et donc dans $\mathcal{T}$. Nous devons voir que $\mathcal{T}$ est stable par complémentarité.

Soit donc $A \in \mathcal{N}$ et $X \in \mathcal{F}$; nous savons que $(A \cup X)^c = A^c \cap X^c$. De plus il existe $Y \in \mathcal{F} \cap \mathcal{N}$ tel que $A \subset Y$ et nous pouvons exprimer A^c en termes de Y : $A^c = Y^c \cup (Y \backslash A)$. Donc

$$(A \cup X)^c = \big(Y^c \cup (Y \backslash A)\big) \cap X^c = \underbrace{(Y^c \cap X^c)}_{\in \mathcal{F}} \cup \underbrace{\big((Y \backslash A) \cap X^c\big)}_{\in \mathcal{N}}. \tag{13.5}$$

Le fait que la seconde partie soit dans $\mathcal{N}$ est due au fait que ce soit une partie de $Y \in \mathcal{N}$. Nous avons donc bien $(A \cup X)^c \in \mathcal{T}$. □

13.1.1.1 Tribu induite

Proposition 13.6.
Soit un espace mesurable $(S, \mathcal{F})$ et une partie $R \subset S$. L'ensemble

$$\mathcal{F}_R = \{A \cap R \text{ tel que } A \in \mathcal{F}\} \tag{13.6}$$

est une tribu.

Démonstration. D'abord R et $\varnothing$ dont dans $\mathcal{F}_R$. Si $C \in \mathcal{F}_R$ alors $C = A \cap R$ pour un certain $A \in \mathcal{F}$ et nous devons prouver que $R \cap C^c$ est dans $\mathcal{F}_R$ (le complémentaire de C dans R). Nous avons

$$R \cap C^c = R \cap (A \cap R)^c = R \cap A^c \in \mathcal{F}_R \tag{13.7}$$

parce que $A^c \in \mathcal{F}$. Enfin si $C_i \in \mathcal{F}_R$ alors $C_i = R \cap A_i$ pour des A_i dans $\mathcal{F}$. Nous avons

$$\bigcup_{i \in \mathbb{N}} C_i = \bigcup_{i \in \mathbb{N}} (R \cap A_i) = R \cap \big(\bigcup_{i \in \mathbb{N}} A_i\big), \tag{13.8}$$

mais $\bigcup_i A_i \in \mathcal{F}$ donc $\bigcup_i C_i \in \mathcal{F}_R$. □

Définition 13.7.
Soit un espace mesurable $(S, \mathcal{F})$ et une partie $R \subset S$. L'ensemble

$$\mathcal{F}_R = \{A \cap R \text{ tel que } A \in \mathcal{F}\} \tag{13.9}$$

*est la **tribu induite** de R depuis S.*

13.1.1.2 Tribu borélienne

Définition 13.8 (Tribu borélienne)**.**
La tribu des ***boréliens****, notée $\mathcal{B}or(\mathbb{R}^d)$ est la tribu engendrée par les ouverts de $\mathbb{R}^d$. Plus généralement si Y est un espace topologique, la tribu des boréliens est la tribu engendrée par les ouverts de Y.*

Proposition 13.9.
La tribu engendrée par une base dénombrable de la topologie est celle des boréliens.

Démonstration. Si une base de topologie est donnée, tout ouvert peut être écrit comme union d'élément de la base, proposition 6.17. Dans le cas d'une base dénombrable, cette union sera forcément dénombrable. Une tribu étant stable par union dénombrable, tout ouvert est dans la tribu engendrée par la base de topologie. Les autres boréliens suivent automatiquement.

Dit avec plus de lettres et moins de phrases, si $\mathcal{D}$ est une base dénombrable de la topologie de X, et si $\mathcal{O}$ est un ouvert de X, nous avons $\mathcal{O} = \bigcup_{i=1}^{\infty} A_i$ avec $A_i \in \mathcal{D}$. Vu qu'une tribu est stable par union dénombrable [1], nous avons $\mathcal{O} \in \sigma(\mathcal{D})$. En conséquence de quoi $\mathcal{B}\text{or}(X) \subset \sigma(\mathcal{D})$.

Mais comme $\mathcal{D} \subset \mathcal{B}\text{or}(X)$ l'inclusion inverse est automatique. D'où l'égalité $\mathcal{B}\text{or}(X) = \sigma(\mathcal{D})$. □

13.1.1.3 Les boréliens de $\mathbb{R}$

Nous rappelons que la topologie de $\mathbb{R}$ est celle des boules donnée par le théorème 6.19. Nous rappelons (voir la proposition 6.60 et sa preuve) que les boules ouvertes de la forme $B(q, r)$ avec $q, r \in \mathbb{Q}$ forment une base dénombrable de la topologique de $\mathbb{R}$.

Lemme 13.10.
Soit $\{q_i\}$ une énumération des rationnels. La tribu engendrée par les ouverts $\sigma_i =]q_i, \infty[$ est la tribu des boréliens.

Démonstration. Si $a < b$ dans $\mathbb{Q}$ alors $\sigma_a \backslash \sigma_b =]a, b]$. Ensuite

$$\bigcup_{n\in\mathbb{N}^*} \sigma_a \backslash \sigma_{b-\frac{1}{n}} = \bigcup_{n\in\mathbb{N}^*}]a, b - \frac{1}{n}] =]a, b[. \tag{13.10}$$

Par union dénombrable, tous les intervalles $]a, b[$ avec $a, b \in \mathbb{Q}$ sont dans la tribu engendrée par les σ_i.

Ces boules ouvertes forment une base de la topologie de $\mathbb{R}$ par la proposition 6.60 et la proposition 13.9 conclu. □

Exemple 13.11
Les singletons sont des boréliens de $\mathbb{R}$ parce que

$$\{x\} = \Big(]-\infty, x[\cup]x, +\infty[\Big)^c. \tag{13.11}$$

Vu qu'une tribu est stable par union dénombrable, l'ensemble $\mathbb{Q}$ est un borélien de $\mathbb{R}$. Et comme les tribus sont stables par différence ensembliste (13.3(2)), l'ensemble des irrationnels est un borélien de $\mathbb{R}$. △

13.1.2 Tribu de Baire

Définition 13.12.
Une partie d'un espace topologique est ***rare*** *si elle est contenue dans un fermé d'intérieur vide.*

Une partie est ***maigre*** *si elle est réunion finie ou dénombrable de parties rares.*

1. Définition 13.1(3)

Exemple 13.13
L'ensemble $\mathbb{Q}$ dans $\mathbb{R}$ est maigre mais n'est pas rare parce que $\bar{\mathbb{Q}} = \mathbb{R}$. △

Proposition 13.14 ([160]).
Soit X un espace topologique. L'ensemble de parties

$$\mathcal{B}a(X) = \{B \cup A \text{ avec } B \text{ borélien et } A \text{ maigre}\} \tag{13.12}$$

est une tribu. Elle est appelée la ***tribu de Baire*** *de l'espace X.*

Démonstration. Nous allons montrer que les boréliens et les maigres vérifient les conditions de la proposition 13.5.

(1) Si A est maigre, il s'écrit comme $A = \bigcup_{i\in\mathbb{N}} R_i$ où les R_i sont rares. Il existe donc des fermés d'intérieur vide F_i tels que $R_i \subset F_i$; en particulier $A \subset \bigcup_i F_i$. En tant que fermés, $F_i \in \mathcal{B}\text{or}(X)$; de plus chaque F_i est rare, donc $\bigcup_i F_i$ est maigre. L'ensemble A est donc bien contenu dans un ensemble maigre et borélien.

(2) Soi A maigre et $B \subset A$. Nous avons, avec les mêmes notations, $A = \bigcup_i R_i$ et $B = \bigcup_i (R_i \cap B)$. Les ensembles $R_i \cap B$ sont encore rares, donc B est une union dénombrable d'ensembles rares. L'ensemble B est donc maigre.

(3) Si les ensembles (A_i) sont maigres, alors ils sont unions dénombrables de rares : $A_i = \bigcup_k R_k^{(i)}$. Nous avons alors

$$\bigcup_i A_i = \bigcup_{(i,k)\in\mathbb{N}^2} R_k^{(i)}, \tag{13.13}$$

et donc $\bigcup_i A_i$ est encore une union dénombrable d'ensembles rares.

□

Proposition 13.15 ([160]).
Une partie B de l'espace topologique X est dans la tribu de Baire de X si et seulement s'il existe un ouvert U tel que $B\Delta U$ est maigre.

Démonstration. Nous définissons la relation d'équivalence[2] suivante sur $\mathcal{P}(X)$: nous disons que $A \sim B$ si et seulement si $A\Delta B$ est maigre.

Réflexive Nous avons $A\Delta A = \varnothing$, donc $A \sim A$.

symétrique Nous avons $A\Delta B = B\Delta A$, donc $\sim$ est symétrique.

transitive Si A, B, C sont des parties de X alors nous avons toujours

$$A\Delta C \subset (A \cup B \cup C)\backslash(A \cap B \cap C) = (A\Delta B) \cup (B\Delta C). \tag{13.14}$$

Donc si $A \sim B$ et $B \sim C$ alors $A\Delta C$ est contenu dans une union de maigres et est donc maigre.

Autres propriétés de $\sim$ De plus la relation d'équivalence $\sim$ vérifie $A \sim B$ si et seulement si $A^c \sim B^c$, par le lemme 2.14(1).

Pour compléter les propriétés de $\sim$ mentionnons encore le fait que si F est fermé alors $F \sim \text{Int}(F)$. En effet $F \cup \text{Int}(F) = F$ et $F \cap \text{Int}(F) = \text{Int}(F)$, de telle sorte que $F\Delta\,\text{Int}(F) = F\backslash\,\text{Int}(F)$. Cela est un fermé parce que son complémentaire est $F^c \cup \text{Int}(F)$ qui est une union d'ouverts. De plus $F \subset \text{Int}(F)$ est d'intérieur vide, de telle sorte qu'il est rare et donc maigre.

Pour la suite de la preuve nous posons

$$\mathcal{F} = \{A \subset X \text{ tel que il existe ouvert } U \text{ avec } U \sim A\}, \tag{13.15}$$

et nous devons prouver que $\mathcal{F} = \mathcal{B}\text{a}(X)$.

2. Définition 2.15

$\underline{\mathcal{F} \subset \mathcal{B}\mathbf{a}(X)}$ Soit $A \in \mathcal{F}$ et un ouvert U tel que $U \sim A$. Alors nous posons $M = U\Delta A$ qui est maigre. En vertu du lemme 2.14(2), nous avons

$$A = M\Delta U = (M \cup U)\backslash(M \cap U), \tag{13.16}$$

ce qui prouve que A est dans la tribu engendrée par les ouverts et les maigres, laquelle tribu est contenue dans $\mathcal{B}\mathrm{a}(X)$.

$\underline{\mathcal{B}\mathbf{a}(X) \subset \mathcal{F}}$ Nous allons montrer que $\mathcal{F}$ est une tribu contenant tous les ouverts et tous les maigres. Alors en particulier $\mathcal{F}$ contiendra $\mathcal{B}\mathrm{a}(X)$. Si U est ouvert, $U \sim U$ et donc $U \in \mathcal{F}$. Si M est maigre, alors $M \sim \varnothing$ et donc $M \in \mathcal{F}$. Il reste à prouver que $\mathcal{F}$ est une tribu.

Vide et ensemble Cela est facile : X et $\varnothing$ sont dans $\mathcal{F}$.

Comlémentaire Commençons par nous souvenir que $F \sim \mathrm{Int}(F)$ dès que F est fermé. Si $A \in \mathcal{F}$ alors il existe un ouvert U tel que $A \sim U$ et donc aussi $A^c \sim U^c$. D'autre part U^c est fermé, donc $U^c \sim \mathrm{Int}(U^c)$, donc

$$A^c \sim U^c \sim \mathrm{Int}(U^c), \tag{13.17}$$

ce qui implique que $A^c \in \mathcal{F}$.

Union dénombrable Soit $A_n \in \mathcal{F}$ et $M_n = A_n\Delta U_n$ avec M_n maigre et U_n ouvert. Nous allons prouver que

$$\bigcup_n A_n \sim \bigcup_n U_n. \tag{13.18}$$

Pour cela il faut remarque que

$$\Big(\bigcup_n A_n\Big)\Delta\Big(\bigcup_n U_n\Big) \subset \bigcup_n (A_n\Delta U_n) = \bigcup_n M_n. \tag{13.19}$$

Le terme le plus à droite est maigre, ce qui signifie que celui le plus à gauche est contenu dans un maigre et donc est maigre lui-même.

□

Proposition 13.16.
Si B est un borélien de X, alors il existe un ouvert U et un maigre M tels que

(1) $B\Delta U$ est maigre,

(2) $M\Delta U = B$,

(3) $D\Delta M$ est ouvert.

Démonstration. Vu que B est borélien, il est aussi dans la tribu de Baire et il existe par la proposition 13.15 un ouvert U tel que $M = B\Delta U$ est maigre. En prenant ce U et ce M, les trois conditions sont vérifiées parce que

$$M\Delta U = (B\Delta U)\Delta U = B \tag{13.20}$$

et

$$B\Delta M = M\Delta B = (U\Delta B)\Delta B = U. \tag{13.21}$$

Tout par le lemme 2.14(2). □

13.2 Théorie de la mesure

Définition 13.17 ([161])**.**
*Une **mesure extérieure** sur un ensemble S est une application $m^*\colon \mathcal{P}(S) \to [0,\infty]$ telle que*

(1) $m^*(\varnothing) = 0$,

(2) Si $A \subset B$ dans S alors $m^*(A) \leqslant m^*(B)$

(3) Si les A_n sont des parties de S alors

$$m^*\big(\bigcup_{n\in\mathbb{N}} A_n\big) \leqslant \sum_{n\in\mathbb{N}} m^*(A_n). \tag{13.22}$$

La différence avec une mesure est que nous ne demandons pas que (13.22) soit une égalité lorsque les A_n sont disjoints.

13.2.1 Mesure et algèbre de parties

Définition 13.18 (Algèbre de parties[161])**.**
*Soit S, un ensemble. Une classe $\mathcal{D}$ de parties de S est une **algèbre de parties** de S si*

(1) $S \in \mathcal{D}$ et $\varnothing \in \mathcal{D}$,

(2) si $A \in \mathcal{D}$ alors $A^c \in \mathcal{D}$,

(3) si $A, B \in \mathcal{D}$ alors $A \cup B \in \mathcal{D}$.

Les algèbre de parties ne sont pas des classes si sauvages que ça ; en témoigne le lemme suivant.

Lemme 13.19.
Une algèbre de partie est stable par intersection (finie) et par différence ensembliste.

Démonstration. Il suffit de remarquer que $A \cap B = \big(A^c \cup B^c\big)^c$ et que $A\backslash B = A \cap B^c$. □

Définition 13.20 (Mesure sur une algèbre de parties)**.**
*Soit S un ensemble et $\mathcal{D}$ une algèbre de parties de S. Une **mesure positive** sur $(S, \mathcal{D})$ est une application $\mu\colon \mathcal{D} \to [0, \infty]$ telle que*

(1) $\mu(\varnothing) = 0$,

(2) Si $A_n \in \mathcal{D}$ sont des ensembles deux à deux disjoints et tels que $\bigcup_n A_n \in \mathcal{D}$ alors

$$\mu\big(\bigcup_n A_n\big) = \sum_n \mu(A_n). \tag{13.23}$$

*La mesure est **finie** si $\mu(S) < \infty$ et σ-finie s'il existe une suite (S_n) dans $\mathcal{D}$ telle que $S = \bigcup_n S_n$ et $\mu(S_n) < \infty$.*

Lemme 13.21 ([161])**.**
Si $\mathcal{D}$ est une algèbre de parties de S et si μ est une mesure sur $(S, \mathcal{D})$ alors

(1) si $A, B \in \mathcal{D}$ avec $A \subset B$ alors $\mu(A) \leqslant \mu(B)$

(2) si $A_n \in \mathcal{D}$ et $\bigcup_n A_n \in \mathcal{D}$ alors

$$\mu\big(\bigcup_n A_n\big) \leqslant \sum_n \mu(A_n). \tag{13.24}$$

La propriété (2) est la σ-**sous-additivité**.

Démonstration. Si $A \subset B$ alors $B = A \cup (B\backslash A)$ avec A et $B\backslash A$ disjoints donc

$$\mu(B) = \mu(A) + \mu(B\backslash A) \geqslant \mu(A). \tag{13.25}$$

Pour la seconde, on passe par les compléments deux à deux : nous posons

$$\begin{cases} B_0 = \varnothing & \text{(13.26a)} \\ B_n = A_n\backslash \bigcup_{k<n} B_k. & \text{(13.26b)} \end{cases}$$

Ces ensembles sont deux à deux disjoints et $\bigcup_n B_n = \bigcup_n A_n \in \mathcal{D}$, donc

$$\mu\big(\bigcup_n A_n\big) = \sum_n \mu\big(A_n\backslash \bigcup_{k<n} B_k\big) \leqslant \sum_n \mu(A_n), \tag{13.27}$$

où nous avons utilisé la première partie du lemme. □

Proposition 13.22 (Mesure extérieure à partir d'une algèbre de parties[161])**.**
Soit $\mathcal{D}$ une algèbre de partie sur l'ensemble S et μ une mesure sur $(S,\mathcal{D})$. Alors l'application

$$\begin{aligned}\mu^*\colon \mathcal{P}(S) &\to [0,+\infty]\\ X &\mapsto \inf\{\sum_n \mu(A_n) \text{ tel que } A_n\in\mathcal{D}, X\subset\bigcup_n A_n\}\end{aligned}\tag{13.28}$$

est une mesure extérieure[3] *sur S et pour tout $A\in\mathcal{D}$ nous avons $\mu^*(A)=\mu(A)$.*

Démonstration. Ceci est une phrase servant à aligner correctement le `description` qui suit.

La définition est bonne Notons d'abord que la définition est bonne : l'ensemble sur lequel l'infimum est pris n'est pas vide : il suffit de prendre $A_1=S$ et $A_{n\geqslant 2}=\varnothing$.

Le vide D'abord $\mu^*(\varnothing)=0$ parce que $\varnothing\in\mathcal{D}$. Prendre $A_n=\varnothing$.

Inégalité d'inclusion Soient $X\subset Y$ dans $\mathcal{P}(S)$. Si $Y\subset A$ alors $X\subset A$, donc

$$\{\sum_n \mu(A_n) \text{ tel que } A_n\in\mathcal{D}, Y\subset\bigcup_n A_n\}\subset\{\sum_n\mu(A_n) \text{ tel que } A_n\in\mathcal{D}, X\subset\bigcup_n A_n\},\tag{13.29}$$

ce qui prouve que $\mu^*(X)\leqslant\mu^*(Y)$.

Inégalité par union dénombrable Soit $(X_n)_{n\in\mathbb{N}}$ une suite de parties de S. S'il existe n_0 tel que $\mu^*(X_{n_0})=\infty$ alors nous avons automatiquement $\sum_n\mu^*(X_n)=\infty$ et l'inégalité demandée est évidente parce que n'importe quoi est plus petit ou égal à ∞. Nous supposons donc que $\mu^*(X_n)<\infty$ pour tout n.

Soit $\epsilon>0$. Pour tout $n\geqslant 1$ il existe une suite $(B_k^{(n)})_{k\in\mathbb{N}}$ dans $\mathcal{D}$ tells que $X_n\subset\bigcup_k B_k^{(n)}$ et

$$\mu^*(X_n)+\frac{\epsilon}{2^n}\geqslant\sum_k\mu(B_k^{(n)}).\tag{13.30}$$

Étant donné que

$$\bigcup_n X_n\subset\bigcup_n\big(\bigcup_k B_k^{(n)}\big),\tag{13.31}$$

nous avons

$$\mu^*\big(\bigcup_n X_n\big)\leqslant\sum_n\sum_k\mu^*(B_k^{(n)})\leqslant\sum_n\Big(\mu^*(X_n)+\frac{\epsilon}{2^n}\Big)=\sum_n\mu^*(X_n)+\epsilon.\tag{13.32}$$

Cette inégalité étant valable pour tout ϵ, nous avons bien

$$\mu^*\big(\bigcup_n X_n\big)=\sum_n\mu^*(X_n).\tag{13.33}$$

Restriction Soit $A\in\mathcal{D}$. Nous avons automatiquement $\mu^*(A)\leqslant\mu(A)$ parce que $\mu(A)$ est dans l'ensemble dont nous prenons l'infimum (prendre $A_1=A$ et $A_{n\geqslant 2}=\varnothing$).

En ce qui concerne l'inégalité inverse nous considérons une suite A_n dans $\mathcal{D}$ telle que $A\subset\bigcup_n A_n$. Vu que $A\in\mathcal{D}$ et que $\mathcal{D}$ est une algèbre de parties nous avons $A\cap A_n\in\mathcal{D}$ et $\bigcup_n(A\cap A_n)=A\in\mathcal{D}$. Par conséquent

$$\mu(A)=\mu\big(\bigcup_n(A\cap A_n)\big)\leqslant\sum_n\mu(A\cap A_n)\leqslant\sum_n\mu(A_n).\tag{13.34}$$

Donc tous les éléments de l'ensemble sur lequel nous prenons l'infimum sont plus grands que $\mu(A)$. Nous en déduisons que $\mu^*(A)\geqslant\mu(A)$.

□

3. Définition 13.17.

13.2.2 Mesure sur un espace mesurable

Définition 13.23 (Mesure positive)**.**
Une ***mesure positive*** *sur l'espace mesurable*[4] $(\Omega, \mathcal{A})$ *est une application* $\mu\colon \mathcal{A} \to [0,\infty]$ *telle que*

(1) $\mu(A) \geqslant 0$ *pour tout* $A \in \mathcal{A}$ *;*

(2) $\mu\left(\bigcup_{i=0}^{\infty} A_i\right) = \sum_{i=0}^{\infty} \mu(A_i)$ *si les* A_i *sont des éléments de* $\mathcal{A}$ *deux à deux disjoints.*

Le triple $(\Omega, \mathcal{A}, \mu)$ *est alors un* ***espace mesuré****.*

Une mesure est σ***-finie*** *s'il existe un recouvrement dénombrable de* Ω *par des ensembles de mesure finie. Si la mesure est* σ*-finie, nous disons que l'espace* $(\Omega, \mathcal{A}, \mu)$ *est un espace mesuré* σ*-fini.*

La mesure μ *sur* Ω *est* ***finie*** *si* $\mu(\Omega) < \infty$*.*

Remarque 13.24.
La définition implique automatiquement que $\mu(\varnothing) = 0$. En effet si A est une partie mesurable, $\mu(A \cup \varnothing) = \mu(A)$, mais A et $\varnothing$ sont disjoints, donc $\mu(A \cup \varnothing) = \mu(A) + \mu(\varnothing)$.

Si $(\Omega, \mathcal{A}, \mu)$ et $(S, \mathcal{F}, \nu)$ sont deux espaces mesurés, alors nous notons

$$(\Omega, \mathcal{A}, \mu) \subset (S, \mathcal{F}, \nu) \tag{13.35}$$

lorsque $\Omega \subset S$, $\mathcal{A} \subset \mathcal{F}$ et pour tout $A \in \mathcal{A}$, $\mu(A) = \nu(A)$.

Définition 13.25 (Ensemble mesurable)**.**
Les éléments de $\mathcal{A}$ *sont les ensembles* ***mesurables*** *pour la mesure* μ*.*

Si la mesure des σ-finie, nous pouvons choisir le recouvrement croissant pour l'inclusion. En effet si $(E_n)_{n\in\mathbb{N}}$ est le recouvrement, il suffit de considérer $F_n = \bigcup_{k\leqslant n} E_k$. Ces ensembles F_n forment tout autant un recouvrement dénombrable, mais il est évidemment croissant.

Le lemme suivant complète la propriété 13.23(2) lorsque les ensembles ne sont pas disjoints.

Lemme 13.26.
Si $A \subset B$ *sont deux ensembles* μ*-mesurables de mesure finie alors*

$$\mu(B\backslash A) = \mu(B) - \mu(A) \tag{13.36}$$

et en particulier

$$\mu(B) \geqslant \mu(A). \tag{13.37}$$

Si (M_n) *est une suite d'éléments de* $\mathcal{A}$ *pas spécialement disjoints, alors*

$$\mu\Big(\bigcup_k M_k\Big) \leqslant \sum_k \mu(M_k). \tag{13.38}$$

Démonstration. Vu que les ensembles $B\backslash A$ et A sont disjoints par la propriété (2) de la définition de mesure nous avons

$$\mu\big((B\backslash A) \cup A\big) = \mu(B\backslash A) + \mu(A) \tag{13.39}$$

et donc

$$\mu(B) = \mu(B\backslash A) + \mu(A) \tag{13.40}$$

comme demandé.

Pour la seconde partie nous considérons la suite disjointe

$$\begin{cases} M'_0 = \varnothing & \text{(13.41a)} \\ M'_k = M_k \backslash M'_{k-1}. & \text{(13.41b)} \end{cases}$$

Nous avons $\bigcup_k M'_k = \bigcup_k M_k$. Le calcul suivant est alors immédiat :

$$\mu\Big(\bigcup_k M_k\Big) = \mu\Big(\bigcup_k M'_k\Big) = \sum_k \mu(M'_k) = \sum_k \mu(M_k \backslash M'_{k-1}) \leqslant \sum_k \mu(M_k). \tag{13.42}$$

□

4. Les définitions de tribus et d'espaces mesurables sont en 13.1.

Lemme 13.27 ([1])**.**
Résultats sur les unions croissantes d'ensembles mesurables dans $(S, \mathcal{A}, \mu)$*.*

(1) Si (A_k) *est une suite croissante d'ensembles* μ*-mesurables dont l'union est mesurable, alors*

$$\lim_{n\to\infty} \mu(A_k) = \mu(\bigcup_k A_k). \tag{13.43}$$

(2) Soit K_n*, une suite emboîtée d'éléments de* $\mathcal{A}$ *tels que* $K_n \to S$*. Si* $A \in \mathcal{A}$ *alors*

$$\lim_{n\to\infty} \mu(A \cap K_n) = \mu(A). \tag{13.44}$$

Démonstration. Pour prouver (1), nous faisons le coup de l'union télescopique, en posant $A_0 = \varnothing$:

$$\bigcup_{k=1}^{\infty} A_k = \bigcup_{k=1}^{\infty} (A_k \backslash A_{k-1}). \tag{13.45}$$

Les ensembles $A_k \backslash A_{k-1}$ sont deux à deux disjoints, donc la propriété (2) de la définition d'une mesure donne

$$\mu(\bigcup_{k=1}^{\infty} A_k) = \mu\left(\bigcup_{k=1}^{\infty} (A_k \backslash A_{k-1})\right) \tag{13.46a}$$

$$= \sum_{k=1}^{\infty} \mu(A_k \backslash A_{k-1}) \tag{13.46b}$$

$$= \sum_{k=1}^{\infty} \big(\mu(A_k) - \mu(A_{k-1})\big) \tag{13.46c}$$

$$= \lim_{k\to\infty} \mu(A_k) - \mu(A_0) \tag{13.46d}$$

$$= \lim_{k\to\infty} \mu(A_k). \tag{13.46e}$$

où pour obtenir 13.46c, nous avons utilisé le lemme 13.26.

Le point (2) est une application du point (1). □

Définition 13.28 (mesure de comptage)**.**
Soit $(S, \mathcal{F})$ *un ensemble mesurable. La* ***mesure de comptage*** *sur* $(S, \mathcal{F})$ *est la mesure définie par*

$$m(A) = \begin{cases} \text{Card}(A) & \text{si } A \text{ est fini} \\ +\infty & \text{sinon.} \end{cases} \tag{13.47}$$

Cette mesure est utilisée pour voir des séries comme des intégrales sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}), m)$.

Exemple 13.29
La mesure de comptage m sur $\mathbb{N}$ muni de la tribu de ses parties est σ-finie parce que $E_n = \{0, \dots, n\}$ est de mesure finie et $\bigcup_{n\in\mathbb{N}} E_n = \mathbb{N}$. △

Exemple 13.30
L'intervalle $I = [0, 1]$ muni de la tribu de toutes ses parties et de la mesure de comptage est un espace mesuré non σ-fini. △

Exemple 13.31
L'intégration à la Riemann n'est pas dans la théorie des espaces mesurés. En effet l'ensemble

$$\mathcal{A} = \{A \subset [0, 1] \text{ tel que } \mathbb{1}_A \text{ est intégrable au sens de Riemann}\} \tag{13.48}$$

n'est pas une tribu. Par exemple les singletons en font partie tandis que $[0, 1] \cap \mathbb{Q}$ n'en fait pas partie malgé que ce soit une union dénombrable de singletons. △

Définition 13.32.
*Si μ est une mesure nous disons qu'une propriété est vraie μ-**presque partout** si elle est fausse seulement sur un ensemble de mesure nulle.*

Par exemple la fonction de Dirichlet est presque partout égale à la fonction 1 (pour la mesure de Lebesgue).

Définition 13.33 (fonction mesurable)**.**
Une application entre espace mesurés

$$f\colon (\Omega, \mathcal{A}) \to (\Omega', \mathcal{A}') \tag{13.49}$$

*est **mesurable** si pour tout $B \in \mathcal{A}'$, l'ensemble $f^{-1}(B)$ est dans $\mathcal{A}$.*

Lemme 13.34.
Une union dénombrable d'ensemble de mesure nulle est de mesure nulle.

Démonstration. C'est une conséquence immédiate du point (2) de la définition d'une mesure : si les A_i sont de mesure nulle,

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leqslant \mu(A_i) = 0 \tag{13.50}$$

□

Définition 13.35.
*Si (A_n) est une suite croissante d'ensembles alors la **limite** est*

$$\lim_n A_n = \bigcup_{i=0}^{\infty} A_i. \tag{13.51}$$

Si la suite est décroissante alors la limite est

$$\lim_n A_n = \bigcap_{i=0}^{\infty} A_i. \tag{13.52}$$

Proposition 13.36 ([162])**.**
Soit μ une mesure sur Ω et (S_n) une suite croissante d'ensembles μ-mesurables de Ω. Nous notons

$$S = \lim_n S_n. \tag{13.53}$$

Alors pour tout ensemble mesurable[5] *$A \subset \Omega$ nous avons*

$$\mu(A \cap S) = \lim_{n\to\infty} \mu(A \cap S_n). \tag{13.54}$$

Note : dans la référence le résultat fonctionne pour tout ensemble A (et non seulement les mesurables) parce que la définition de la mesurabilité est un peu différente.

Démonstration. L'inégalité $\lim \mu(A \cap S_n) \leqslant \mu(A \cap S)$ est simple à prouver. En effet pour tout n nous avons $A \cap S_n \subset A \cap S$ et donc par le lemme 13.26 nous avons

$$\mu(A \cap S_n) \leqslant \mu(A \cap S). \tag{13.55}$$

En passant à la limite (qui respecte les inégalités) nous avons l'inégalité.

Nous passons à l'inégalité dans l'autre sens. D'abord si $\mu(A \cap S_n) = \infty$ pour un certain n, alors il cela vaut encore ∞ pour tous les n suivants et la limite est ∞ sans problèmes. Donc nous

5. Définition 13.25

supposons que $\mu(A \cap S_n) < \infty$ pour tout $n \in \mathbb{N}$. De plus, quitte à renommer les indices, nous pouvons supposer que $S_0 = \varnothing$.

Un élément x est dans S si et seulement s'il existe $n \geqslant 0$ tel que $x \in S_{n+1}$. En prenant le plus petit de ces n nous avons $x \neq S_n$ (éventuellement $n = 0$) et donc

$$S = \bigcup_{n=0}^{\infty} \big(S_{n+1} \backslash S_n\big). \tag{13.56}$$

Par conséquent

$$A \cap S = A \cap \bigcup_{n=0}^{\infty} (S_{n+1} \backslash S_n) = \bigcup_{n=0}^{\infty} A \cap (S_{n+1} \backslash S_n) \tag{13.57}$$

Étant donné que les ensembles $A \cap (S_{n+1} \backslash S_n)$ sont disjoints,

$$\begin{aligned}
\mu(A \cap S) &= \sum_{n=0}^{\infty} \mu\big(A \cap (S_{n+1} \backslash S_n)\big) && \text{(13.58a)}\\
&= \sum_{n=0}^{\infty} \mu\Big((A \cap S_{n+1}) \backslash (A \cap S_n)\Big) && \text{(13.58b)}\\
&= \sum_{n=0}^{\infty} \big[\mu(A \cap S_{n+1}) - \mu(A \cap S_n)\big] && \text{(13.58c)}\\
&= \lim_{n \to \infty} \mu(A \cap S_{n+1}) - \underbrace{\mu(A \cap S_0)}_{=0} && \text{(13.58d)}\\
&= \lim_{n \to \infty} \mu(A \cap S_n). && \text{(13.58e)}
\end{aligned}$$

Dans ce calcul nous avons utilisé plusieurs fois le fait que les S_n et A étaient mesurables (et la propriété de tribu qui dit que $A \cap S_n$ est également mesurable) ainsi que le lemme 13.26. Nous avons aussi utilisé la série télescopique dans $\mathbb{R}$ pour obtenir (13.58d). □

Définition 13.37 (λ-système[163])**.**
Soit E un ensemble. Un ensemble $\mathcal{D}$ de parties de E est un ***λ-système*** *lorsqu'il vérifie les conditions suivantes :*

(1) si $A, B \in \mathcal{D}$ avec $A \subset B$ alors $B \backslash A \in \mathcal{D}$,

(2) si $(A_k)_{k \geqslant 1}$ est une suite croissante d'éléments de $\mathcal{D}$ alors $\bigcup_k A_k \in \mathcal{D}$.

Note : une tribu est un λ-système.

Lemme 13.38 ([163])**.**
Une intersection quelconque de λ-systèmes dans E est un λ-système dans E.

Démonstration. Soient $\{\mathcal{D}_l\}_{l \in L}$ des λ-systèmes indicés par un ensemble L. Si $A, B \in \bigcap_{l \in L} \mathcal{D}_l$ alors $B \backslash A \in \mathcal{D}_l$ pour tout $l \in L$ et donc $A \backslash B \in \bigcap_{l \in L} \mathcal{D}_l$. De la même façon si (A_k) est une suite croissante dans $\bigcap_{l \in L} \mathcal{D}_l$ alors pour tout $l \in L$ nous avons $\bigcup_k A_k \in \mathcal{D}_l$. Donc $\bigcup_k A_k \in \bigcap_l \mathcal{D}_l$. □

Ce lemme est ce qui permet de définir le λ-système **engendré** par une classe $\mathcal{A}$ de parties de E : c'est l'intersection de tous les λ-systèmes de E contenant $\mathcal{A}$.

Lemme 13.39 ([163])**.**
Soit $\mathcal{C}$ une classe de parties de E (contenant E lui-même) qui soit stable par intersection finie. Alors le λ-système engendré par $\mathcal{C}$ coïncide avec la tribu engendrée par $\mathcal{C}$.

Démonstration. Nous notons $\mathcal{E}$ le λ-système engendré par $\mathcal{C}$ et $\mathcal{F}$ la tribu engendrée par $\mathcal{C}$. Étant donné que $\mathcal{F}$ est un λ-système nous avons $\mathcal{E} \subset \mathcal{F}$. Pour montrer l'inclusion inverse nous allons prouver que $\mathcal{E}$ est une tribu.

D'abord pour $C \in \mathcal{C}$ nous posons

$$\mathcal{G}_C = \{A \subset \mathcal{E} \text{ tel que } A \cap C \in \mathcal{E}\}. \tag{13.59}$$

et pour $F \in \mathcal{E}$,

$$\mathcal{H}_F = \{A \in \mathcal{E} \text{ tel que } A \cap F \in \mathcal{E}\}. \tag{13.60}$$

Nous allons montrer que $\mathcal{G}_C$ et $\mathcal{H}_F$ sont des λ-systèmes et que $\mathcal{G}_C = \mathcal{H}_F = \mathcal{E}$.

Nous commençons par $\mathcal{G}_C$. Si $A, B \in \mathcal{G}_C$ avec $A \subset B$ alors

$$(B\backslash A) \cap C = \underbrace{(B \cap C)}_{\in\mathcal{E}} \backslash \underbrace{(A \cap C)}_{\in\mathcal{E}}. \tag{13.61}$$

Vu que $\mathcal{E}$ est un λ-système et que $(A \cap C) \subset (B \cap C)$ nous avons bien $(B\backslash A) \cap C \in \mathcal{E}$ et donc $B\backslash A \in \mathcal{G}_C$. Soit maintenant (A_k) une suite croissante dans $\mathcal{G}_C$. Nous avons

$$\big(\bigcup_{k=1}^{\infty} A_k\big) \cap C = \bigcup_{k=1}^{\infty} (A_k \cap C) \tag{13.62}$$

qui est une union d'une suite croissante d'éléments de $\mathcal{E}$. Donc $\bigcup_{k=1}^{\infty}(A_k \cap C) \in \mathcal{E}$, ce qui signifie que $\bigcup_{k=1}^{\infty} A_k \in \mathcal{G}_C$. Cela termine la preuve du fait que $\mathcal{G}_C$ soit une λ-système.

Étant donné que $\mathcal{C}$ est stable par intersection finie, si $K \in \mathcal{C}$ nous avons $C \cap K \in \mathcal{C}$, ce qui signifie que $K \in \mathcal{G}_C$. Nous avons donc $\mathcal{C} \subset \mathcal{G}_C$. Donc $\mathcal{G}_C$ est un λ-système vérifiant $\mathcal{C} \subset \mathcal{G}_C \subset \mathcal{E}$. Mais comme $\mathcal{E}$ est le plus petit λ-système contenant $\mathcal{C}$ nous avons en fait $\mathcal{G}_C = \mathcal{E}$.

Nous montrons à présent que $\mathcal{H}_F$ est un λ-système. Si $A, B \in \mathcal{H}_F$ avec $A \subset B$ alors $(B\backslash A)\cap F = (B\cap F)\backslash(A\cap F)$. Vu que $\mathcal{E}$ est une λ-système et que $A\cap F$ et $B\cap F$ sont dans $\mathcal{E}$ avec $A\cap F \subset B\cap F$, nous avons

$$(B \cap F)\backslash(A \cap F) \in \mathcal{H}_F. \tag{13.63}$$

Soit maintenant $(A_k)_{k\geqslant 1}$ une suite croissante dans $\mathcal{H}_F$. Pour tout k nous avons $A_k \cap F \in \mathcal{E}$, ce qui donne

$$\big(\bigcap_{k=1}^{\infty} A_k\big) \cap F = \bigcap_{k=1}^{\infty} (A_k \cap F) \in \mathcal{E}. \tag{13.64}$$

Donc $\mathcal{H}_F$ est un λ-système vérifiant $\mathcal{C} \subset \mathcal{H}_F \subset \mathcal{E}$. Nous en concluons que pour tout $C \in \mathcal{C}$ et pour tout $F \in \mathcal{E}$,

$$\mathcal{G}_C = \mathcal{H}_F = \mathcal{E}. \tag{13.65}$$

Nous allons maintenant prouver que $\mathcal{E}$ est une tribu [6].

(1) Si $F \in \mathcal{E}$ alors $E \cap F = F \in \mathcal{E}$, ce qui signifie que $E \in \mathcal{H}_F = \mathcal{E}$.

(2) Si $A \in \mathcal{E}$ alors $E\backslash A \in \mathcal{E}$ parce que $\mathcal{E}$ est un λ-système et $E \in \mathcal{E}$. Donc $\complement A \in \mathcal{E}$.

(3) Montrons que $\mathcal{E}$ est stable par union finie en considérant $A, B \in \mathcal{E}$. Vu que E est également un élément de $\mathcal{E}$ nous avons

$$E\backslash(A \cup B) = (E\backslash A) \cap (E\backslash B) \in \mathcal{E}. \tag{13.66}$$

Cela prouve que $\complement(A \cup B) \in \mathcal{E}$. Par complémentarité nous avons aussi $A \cup B \in \mathcal{E}$.

Soient $A_k \in \mathcal{E}$, et nommons $B_p = A_1 \cup \ldots \cup A_p$. Les ensembles B_p forment une suite croissante d'éléments de $\mathcal{E}$. L'union est donc dans $\mathcal{E}$ et ce dernier est au final stable par union dénombrable.

Maintenant que $\mathcal{E}$ est une tribu nous avons $\mathcal{F} \subset \mathcal{E}$ parce que $\mathcal{F}$ est la plus petite tribu contenant $\mathcal{C}$. Nous en déduisons que $\mathcal{E} = \mathcal{F}$, ce qu'il fallait démontrer. □

Le théorème suivant permet de prouver que la mesure de Lebesgue est l'unique mesure possible ayant les bonnes valeurs sur les intervalles (théorème 13.171).

6. Définition 13.1.

Théorème 13.40 (Unicité des mesures[163])**.**
Soient μ et ν, deux mesures sur $(E, \mathcal{A})$ et une classe $\mathcal{E}$ de parties de E telles que

(1) La tribu engendrée par $\mathcal{E}$ soit $\mathcal{A}$.

(2) si $A, B \in \mathcal{E}$ alors $A \cap B \in \mathcal{E}$

(3) il existe une suite croissante (E_n) dans $\mathcal{E}$ telle que $E = \lim E_n$.

Alors si les mesures μ et ν coïncident sur $\mathcal{E}$, elles coïncident sur $\mathcal{A}$ en entier.

Démonstration. Soit (E_n) la suite des hypothèses ; nous considérons μ_n et ν_n, les restrictions de μ et ν à E_n, c'est à dire

$$\mu_n(A) = \mu(A \cap E_n) \tag{13.67a}$$
$$\nu_n(A) = \nu(A \cap E_n). \tag{13.67b}$$

Vu que les E_n sont dans $\mathcal{E} \subset \mathcal{A}$ ils sont mesurables au sens de μ et ν. Par la proposition 13.36, pour tout $A \in \mathcal{E}$ nous avons alors

$$\lim_{n\to\infty} \mu_n(A) = \mu(A) \tag{13.68a}$$
$$\lim_{n\to\infty} \nu_n(A) = \nu(A) \tag{13.68b}$$

Nous devons donc seulement montrer que pour tout $A \in \mathcal{A}$ et pour tout $n \in \mathbb{N}$, $\mu_n(A) = \nu_n(A)$. Pour cela nous nous fixons un n et nous considérons la classe

$$\mathcal{D} = \{A \in \mathcal{A} \text{ tel que } \mu_n(A) = \nu_n(A)\}. \tag{13.69}$$

Le but sera de prouver que $\mathcal{D} = \mathcal{A}$.

Par hypothèse $A \cap E_n \in \mathcal{E}$ et donc

$$\mu(A \cap E_n) = \nu(A \cap E_n) < \infty, \tag{13.70}$$

c'est à dire que $\mu_n = \nu_n$ sur $\mathcal{E}$. Par ailleurs, $E \cap E_n = E_n \in \mathcal{E}$, donc

$$\mu_n(E) = \nu_n(E) < \infty. \tag{13.71}$$

Par conséquent $\mu_n = \nu_n$ sur la classe $\mathcal{E}' = \mathcal{E} \cup \{E\} : \mathcal{E}' \subset \mathcal{D}$.

Montrons que $\mathcal{D}$ est un λ-système. Soient $A, B \in \mathcal{D}$ avec $A \subset B$. Alors, étant donné que les mesures μ_n et ν_n sont finies, le lemme 13.26 nous donne

$$\mu_n(B\backslash A) = \mu_n(B) - \mu_n(A) \tag{13.72a}$$
$$\nu_n(B\backslash A) = \nu_n(B) - \nu_n(A). \tag{13.72b}$$

Donc $\mu_n(B\backslash A) = \nu_n(B\backslash A)$ et $B\backslash A \in \mathcal{D}$.

Soit par ailleurs une suite croissante $(A_k)_{k\geqslant 1}$ d'éléments de $\mathcal{D}$. En posant $B_p = \bigcup_{k=1}^{p} A_k$, le lemme 13.27(1) nous donne

$$\mu_n(\bigcup_{k=1}^{\infty} A_k) = \lim_{p\to\infty} \mu_n(A_p). \tag{13.73}$$

Mais vu que pour chaque p nous avons $\mu_n(A_p) = \nu_n(A_p)$, nous avons aussi

$$\mu_n(\bigcup_{p=1}^{\infty} A_p) = \nu_n(\bigcup_{p=1}^{\infty} A_p). \tag{13.74}$$

Donc $\mathcal{D}$ est bel et bien un λ-système contenant $\mathcal{E}'$. Par le lemme 13.39, le λ-système engendré par $\mathcal{E}'$ est égal à la tribu engendrée par $\mathcal{E}'$, mais par hypothèse la tribu engendrée par $\mathcal{E}$ est $\mathcal{A}$, donc le λ-système engendré par $\mathcal{E}'$ est $\mathcal{A}$. Vu que $\mathcal{D}$ est une λ-système contenant $\mathcal{E}'$, nous avons alors $\mathcal{A} \subset \mathcal{D}$ et donc $\mathcal{A} = \mathcal{D}$, ce qu'il fallait. □

Exemple 13.41
La partie $\mathcal{E}$ des intervalles de $\mathbb{R}$ de la forme $]a,b[$ engendre les boréliens par la proposition 6.60. Par conséquent pour vérifier que deux mesures sont égales sur les boréliens de $\mathbb{R}$ il suffit de prouver qu'elles sont égales sur les intervalles ouverts. △

13.2.3 Mesure extérieure

Nous avons déjà défini la notion de mesure extérieure en la définition 13.17.

Lemme 13.42 ([161])**.**
Soit $(S,\mathcal{F},\mu)$ un espace mesuré et $X\subset S$. Alors

$$\inf_{\substack{A\in\mathcal{F}\\ X\in A}}\mu(A)=\inf\{\sum_n\mu(A_n)\ \textit{tel que}\ A_n\in\mathcal{F}, X\subset\bigcup_k A_k\}. \tag{13.75}$$

Démonstration. Pour montrer l'inégalité $\geqslant$, nous remarquons qu'il y a plus d'éléments dans l'ensemble du second membre que dans le premier. En effet si $A\in\mathcal{F}$ avec $X\subset A$ alors dans le membre de gauche nous pouvons prendre $A_1=A$ et $A_{n\geqslant 1}=\varnothing$.

Pour l'inégalité dans l'autre sens, nous montrons que tout élément de

$$\{\sum_n\mu(A_n)\text{ tel que }A_n\in\mathcal{F}, X\subset\bigcup_k A_k\} \tag{13.76}$$

est plus grand qu'un élément de

$$\{\mu(A)\text{ tel que }A\in\mathcal{F}, X\subset A\}. \tag{13.77}$$

En effet si $A_n\in\mathcal{F}$ avec $X\subset\bigcup_k A_k$ alors en posant $A=\bigcup_k A_k$ nous avons $A\in\mathcal{F}$ avec $X\subset A$ ainsi que $\mu(A)\leqslant\sum_n\mu(A_n)$. Cela prouve que l'élément $\sum_n\mu(A_n)$ de (13.76) est plus grand que l'élément $\mu(A)$ de (13.77). □

Proposition 13.43 ([161])**.**
Soit un espace mesuré $(S,\mathcal{F},\mu)$ et l'application

$$\begin{aligned}\mu^*\colon\mathcal{P}(S)&\to[0,\infty]\\ X&\mapsto\inf\{\mu(A)\ \textit{tel que}\ A\in\mathcal{F}, X\subset A\}.\end{aligned} \tag{13.78}$$

Alors μ^* est une mesure extérieure sur S et sa restriction à $\mathcal{F}$ est égale à μ.

Cela est un cas particulier de 13.22 en utilisant 13.42. Nous en donnons cependant une preuve directe, qui est presque identique à celle de 13.22, mais avec une ou deux simplifications.

Démonstration. Notons que la définition est bonne parce que l'ensemble sur lequel l'infimum est pris n'est pas vide : prendre $A=S$.

Le vide D'abord $\mu^*(\varnothing)=O$ parce que $\varnothing\in\mathcal{F}$.

Inégalité d'inclusion Soient $X\subset Y$ dans $\mathcal{P}(S)$. Si $Y\subset A$ alors $X\subset A$, donc

$$\inf\{\mu(A)\text{ tel que }A\in\mathcal{F}, X\subset A\}\leqslant\inf\{\mu(A)\text{ tel que }A\in\mathcal{F}, Y\subset A\}, \tag{13.79}$$

ce qui signifie que $\mu^*(X)\leqslant\mu^*(Y)$.

Inégalité par union dénombrable Soit $(X_n)_{n\in\mathbb{N}}$ une suite de parties de S. S'il existe n_0 tel que $\mu^*(X_{n_0})=\infty$ alors nous avons automatiquement $\sum_n\mu^*(X_n)=\infty$ et l'inégalité demandée est évidente parce que n'importe quoi est plus petit ou égal à ∞. Nous supposons donc que $\mu^*(X_n)<\infty$ pour tout n.

Soit $\epsilon>$ et par définition pour chaque n, il existe un $A_n\in\mathcal{F}$ tel que $X_n\subset A_n$ et $\mu(A_n)\leqslant\mu^*(X_n)+\frac{\epsilon}{2^n}$. Bien entendu nous avons

$$\bigcup_n X_n\subset\bigcup_n A_n\in\mathcal{F}. \tag{13.80}$$

Nous en déduisons que

$$\mu^*\big(\bigcup_n X_n\big) \leqslant \mu\big(\bigcup_n A_n\big). \tag{13.81}$$

Mais $(S, \mathcal{F}, \mu)$ étant un espace mesuré,

$$\mu\big(\bigcup_n A_n\big) \leqslant \sum_n \mu(A_n). \tag{13.82}$$

Au final nous avons les inégalités

$$\mu^*\big(\bigcup_n X_n\big) \leqslant \mu\big(\bigcup_n A_n\big) \leqslant \sum_n \mu(A_n) \leqslant \sum_n \mu^*(X_n) + \epsilon \sum_n \frac{1}{2^n} = \sum_n \mu^*(X_n) + \epsilon. \tag{13.83}$$

Cela étant vrai pour tout ϵ,

$$\mu^*\big(\bigcup_n X_n\big) \leqslant \sum_n \mu^*(X_n), \tag{13.84}$$

ce qui prouve que μ^* est une mesure extérieure.

Restriction Supposons que $X \in \mathcal{F}$. Alors si $X \subset A$ nous avons $\mu(X) \leqslant \mu(A)$; mais en même temps, $\mu(X)$ est dans l'infimum qui définit $\mu^*(X)$ donc

$$\mu^*(X) \leqslant \mu(X) \leqslant \inf\{\mu(A) \text{ tel que } A \in \mathcal{F}, X \subset A\} \leqslant \mu(X) \leqslant \mu^*(X). \tag{13.85}$$

Donc nous avons égalité de tous les éléments de cette chaîne d'inégalité.

□

Définition 13.44.
*Soit S un ensemble et m^* une mesure extérieure sur S. Une partie $A \subset X$ est m^*-**mesurable** si pour tout $X \subset S$,*

$$m^*(X) = m^*(X \cap A) + m^*(X \cap A^c). \tag{13.86}$$

Remarque 13.45.
L'inégalité

$$m^*(X) \leqslant m^*(X \cap A) + m^*(X \cap A^c) \tag{13.87}$$

étant toujours vraie, pour prouver qu'un ensemble est m^*-mesurable, il est suffisant de prouver l'inégalité inverse :

$$m^*(X) \geqslant m^*(X \cap A) + m^*(X \cap A^c) \tag{13.88}$$

La définition 13.44 est motivée par la proposition suivante.

Proposition 13.46.
Soit un espace mesuré $(S, \mathcal{F}, \mu)$ et μ^* la mesure extérieure qui va avec. Alors tous les éléments de $\mathcal{F}$ sont μ^*-mesurables.

En d'autres termes, pour tout $A \in \mathcal{F}$ et tout $X \subset S$ nous avons

$$\mu^*(X) = \mu^*(X \cap A) + \mu^*(X \cap A^c). \tag{13.89}$$

Démonstration. Vu que $X = (X \cap A) \cup (X \cap A^c)$, et que μ^* est une mesure extérieure,

$$\mu^*(X) \leqslant \mu^*(X \cap A) + \mu^*(X \cap A^c). \tag{13.90}$$

Nous devons montrer l'inégalité inverse.

Soit $B \in \mathcal{F}$ tel que $X \subset B$. D'une part nous avons $X \cap A \subset B \cap A \in \mathcal{F}$, donc

$$\mu^*(X \cap A) \leqslant \mu^*(B \cap A) = \mu(B \cap A). \tag{13.91}$$

Et d'autre part, $X \cap A^c \subset B \cap A^c \in \mathcal{F}$, donc

$$\mu^*(X \cap A^c) \leqslant \mu(B \cap A^c). \tag{13.92}$$

En remettant ensemble,

$$\mu^*(X\cap A)+\mu^*(X\cap A^c)\leqslant \mu(B\cap A)+\mu(B\cap A^c)=\mu(B). \tag{13.93}$$

La dernière égalité est le fait que $B\cap A$ et $B\cap A^c$ sont disjoints et que μ est une mesure. L'inégalité (13.93) étant vraie pour tout $B\in\mathcal{F}$ tel que $X\subset B$, elle est encore vraie pour l'infimum :

$$\mu^*(X\cap A)+\mu^*(X\cap A^c)\leqslant \inf\{\mu(B) \text{ tel que } B\in\mathcal{F}, X\subset B\}=\mu^*(X). \tag{13.94}$$

Nous avons donc prouvé que

$$\mu^*(X\cap A)+\mu^*(X\cap A^c)\leqslant \mu^*(X). \tag{13.95}$$

□

Remarque 13.47.
Notons la duplicité du vocabulaire. Les ensembles μ-mesurables sont les éléments de $\mathcal{F}$, qui sont a priori les seuls sur lesquels μ est calculable[7], alors que les μ^*-mesurables sont les parties de S qui vérifient une certaine propriété (et μ^* est calculable sur toutes les parties de S).

13.2.4 Espace mesuré complet

Définition 13.48.
*Soit un espace mesuré $(X,\mathcal{A},\mu)$. Une partie N de X est **négligeable** pour μ s'il existe $Y\in\mathcal{A}$ tel que $N\subset Y$ et $\mu(Y)=0$.*

Lemme 13.49.
L'ensemble des parties négligeables est stable par union dénombrable.

Démonstration. Si les ensembles N_i sont négligeables, alors pour chaque i nous avons $Y_i\in\mathcal{A}$ tel que $N_i\subset Y_i$ et $\mu(Y_i)=0$. Alors bien entendu $\bigcup_i N_i\subset\bigcup_i Y_i$ et en utilisant (13.38),

$$\mu\big(\bigcup_i Y_i\big)\leqslant\sum_i\mu(Y_i)=0. \tag{13.96}$$

□

Définition 13.50.
*L'espace mesuré $(X,\mathcal{F},\mu)$ est **complet** si tout ensemble μ-négligeable est dans $\mathcal{F}$.*

Notons que la proposition 13.5 s'applique si $(X,\mathcal{F},\mu)$ est un espace mesuré et $\mathcal{N}$ est l'ensemble des parties μ-négligeables. C'est ce qui permet de donner le théorème suivant, que nous redémontrons de façon indépendante de la proposition 13.5.

Théorème 13.51 (Complétion d'espace mesuré[161, 164, 165])**.**
Soit un espace mesuré $(X,\mathcal{F},\mu)$ et $\mathcal{N}$ l'ensemble des parties μ-négligeables de X.

(1) Les ensembles suivants sont égaux :

$$\mathcal{A}=\{A\subset X \text{ tel que } \exists B,C\in\mathcal{F} \text{ tel que } B\subset A\subset C, \mu(C\backslash B)=0\} \tag{13.97a}$$
$$\mathcal{B}=\{B\cup N \text{ tel que } B\in\mathcal{F}, N\in\mathcal{N}\} \tag{13.97b}$$
$$\mathcal{C}=\{A\subset X \text{ tel que } \exists B\in\mathcal{F} \text{ tel que } A\Delta B\in\mathcal{N}\}. \tag{13.97c}$$

Ici $A\Delta B$ est la différence symétrique de A et B, définition 2.13.

(2) L'ensemble $\hat{\mathcal{F}}=\mathcal{A}=\mathcal{B}=\mathcal{C}$ est une tribu.

7. «calculable» au sens où μ y vaut un nombre bien définit ; après, que ce soit facile ou pas à calculer dans la pratique, c'est une autre histoire.

(3) *La définition*

$$\begin{aligned}\mu' \colon \mathcal{B} &\to [0,\infty] \\ A \cup N &\mapsto \mu(A)\end{aligned} \tag{13.98}$$

est cohérente.

(4) *L'application μ' ainsi définie est une mesure sur $(X,\mathcal{A})$.*

(5) *L'espace $(X,\mathcal{A},\mu')$ est complet.*

(6) *La mesure μ' prolonge μ.*

(7) *La mesure μ' est minimale au sens où toute mesure complète prolongeant μ prolonge μ'.*

Démonstration. Commençons par prouver que les trois ensembles $\mathcal{A}$, $\mathcal{B}$ et $\mathcal{C}$ sont égaux.

$\underline{\mathcal{A} \subset \mathcal{B}.}$ Soit $A \in \mathcal{A}$. Alors nous avons des ensembles $B, C \in \mathcal{F}$ tels que $B \subset A \subset V$ avec $\mu(C\backslash B) = 0$. Alors nous avons aussi $A = B \cup (C\backslash B)$, ce qui prouve que $A \in \mathcal{B}$.

$\underline{\mathcal{B} \subset \mathcal{C}.}$ Soit $A \in \mathcal{B}$, c'est à dire que $A = B \cup N$ avec $B \in \mathcal{F}$ et $N \in \mathcal{N}$. Nous avons évidemment $A \cup B = A$ et donc

$$A\Delta B = (A \cup B)\backslash(A \cap B) = A\backslash(A \cap B) = (B \cup N)\backslash(A \cap B) \subset N. \tag{13.99}$$

Pour comprendre la dernière inclusion, si x appartient à $A = B \cup N$ sans être dans N alors $x \in B$ et donc $x \in A \cap B$. Par conséquent nous avons $A\Delta B \subset N$ et donc $A\Delta B \in \mathcal{N}$.

$\underline{\mathcal{C} \subset \mathcal{A}}$ Soit donc $A \in \mathcal{C}$; il existe $B \in \mathcal{F}$ tel que $A\Delta B \in \mathcal{N}$ ou encore, il existe $D \in \mathcal{F}$ tel que $A\Delta B \subset D$ avec $\mu(D) = 0$. Si nous posons $B' = B \cap D^c$ et $C' = B \cup D$ alors nous prétendons avoir

$$B' \subset A \subset C'. \tag{13.100}$$

Et nous le prouvons. En effet si $x \in B \cap D^c$ alors en remarquant que B se divise en

$$B = (B \cap A) \cup \big(B \cap (A\Delta B)\big), \tag{13.101}$$

et en nous souvenant que $B \cap (A\Delta B) \subset D$, il vient que $B \cap D^c \subset B \cap A$. Et en particulier $x \in A$. D'autre part

$$A \subset B \cup (A\Delta B) \subset B \cup D. \tag{13.102}$$

Nous avons donc bien $B' \subset A \subset C'$. Par stabilité de la tribu $\mathcal{F}$ sous les intersections et complémentaires nous avons aussi $B', C' \in \mathcal{F}$. De plus

$$C'\backslash B' = (B \cup D)\backslash(B \cap D^c) \subset D, \tag{13.103}$$

et donc

$$\mu(C'\backslash B') \leqslant \mu(D) = 0. \tag{13.104}$$

Nous avons donc prouvé que $\mathcal{A} \subset \mathcal{B} \subset \mathcal{C} \subset \mathcal{A}$, et donc que $\mathcal{A} = \mathcal{B} = \mathcal{C}$. Nous allons donc maintenant noter $\mathcal{A}$ indifféremment les trois ensembles. Nous prouvons à présent que c'est une tribu.

Tribu : le vide Pas de problèmes à $\varnothing \in \mathcal{A}$

Tribu : complémentaire Soit $A \in \mathcal{A}$. Alors il existe $B, C \in \mathcal{F}$ tels que $B \subset A \subset C$ avec $\mu(C\backslash B) = 0$. En passant au complémentaire,

$$C^c \subset A^c \subset B^c. \tag{13.105}$$

Mais $B^c\backslash C^c = C\backslash B$, donc $\mu(B^c\backslash C^c) = 0$.

Tribu : union dénombrable Soit (A_n) des éléments de $\mathcal{A}$. Pour chaque n nous avons des ensembles $B_n, C_n \in \mathcal{F}$ tels que$B_n \subset A_n \subset C_n$ avec $\mu(C_n \backslash B_n) = 0$. En ce qui concerne les unions nous avons

$$\bigcup_n B_n \subset \bigcup_n A_n \subset \bigcup_n C_n, \tag{13.106}$$

et

$$(\bigcup_n C_n) \backslash (\bigcup_n B_n) \subset \bigcup_n (C_n \backslash B_n). \tag{13.107}$$

Par conséquent, en utilisant (13.38),

$$\mu\left((\bigcup_n C_n) \backslash (\bigcup_n B_n) \right) \leqslant \mu\left(\bigcup_n (C_n \backslash B_n) \right) \leqslant \sum_n \mu(C_n \backslash B_n) = 0. \tag{13.108}$$

Cela prouve que $\bigcup_n A_n \in \mathcal{A}$, et donc que $\mathcal{A}$ est une tribu.

Définition cohérente Soient $A, A' \in \mathcal{F}$ et $N, N' \in \mathcal{N}$ tels que $A \cup N = A' \cup N'$. Nous considérons $Y, Y' \in \mathcal{F}$ tel que $N \subset Y$, $N' \subset Y'$ et $\mu(Y) = \mu(Y') = 0$. En vertu de (13.38) nous avons

$$\mu(A) \leqslant \mu(A \cup Y) \leqslant \mu(A' \cup Y \cup Y') \leqslant \mu(A') + \mu(Y) + \mu(Y') = \mu(A'). \tag{13.109}$$

En écrivant la même chose en échangeant les primes nous prouvons également $\mu(A') \leqslant \mu(A)$. Au final $\mu(A) = \mu(A')$, c'est à dire

$$\mu'(A \cup N) = \mu'(A' \cup N'). \tag{13.110}$$

La définition de μ' est donc cohérente.

μ' est une mesure Le fait que μ' soit positive et que $\mu'(\varnothing)$ soit nul ne pose pas de problèmes. Il faut voir l'union dénombrable disjointe. Si les ensembles $A_i = B_i \cup N_i$ sont disjoints, alors les B_i et le N_i sont tous disjoints deux à deux. De plus l'ensemble $\bigcup_i N_i$ est négligeable parce que nous avons déjà vu que $\mathcal{N}$ était stable par union dénombrable (13.38). Donc

$$\mu'\left(\bigcup_i B_i \cup N_i \right) = \mu'\Big((\bigcup_i B_i) \cup \underbrace{(\bigcup_i N_i)}_{\in \mathcal{N}} \Big) = \mu\big(\bigcup_i B_i\big) = \sum_u \mu(B_i) = \sum_i \mu'(B_i \cup N_i). \tag{13.111}$$

Espace complet Un ensemble μ'-négligeable est automatiquement μ-négligeable. En effet si H est μ'-négligeable, il existe $B \in \mathcal{F}$ et $N \in \mathcal{N}$ tels que $H \subset B \cup N$ avec $\mu(B) = 0$. Vu que N est μ-négligeable, il existe $Y \in \mathcal{F}$ tel que $N \subset Y$ et $\mu(Y) = 0$. Donc $H \subset B \cup N \subset B \cup Y$ avec $\mu(B \cup Y) = 0$.

Tous les ensembles μ-négligeables faisant partie de $\mathcal{B}$, tous les ensembles μ'-négligeables font partie de $\mathcal{A}$.

Prolongement La mesure μ' prolonge μ. En effet si $A \in \mathcal{F}$ alors $A = A \cup \varnothing \in \mathcal{B}$ et A est m'-mesurable. De plus $\mu'(A) = \mu'(A \cup \varnothing) = \mu(A)$.

Minimalité Soit un espace mesuré complet $(X, \mathcal{M}, \nu)$ prolongeant $(X, \mathcal{F}, \mu)$. Pour $A \in \mathcal{A}$ nous devons prouver que $A \in \mathcal{M}$ et que $\mu'(A) = \nu(A)$. Il existe $B \in \mathcal{F}$ et $N \in \mathcal{N}$ tels que $A = B \cup N$. Vu que N est μ-négligeable, il est également ν-négligeable et donc ν-mesurable parce que ν est complète : $A \in \mathcal{M}$. En ce qui concerne l'égalité $\mu'(A) = \nu(A)$ nous avons

$$\nu(B) \leqslant \nu(B \cup N) \leqslant \nu(B) + \nu(N) = \nu(B), \tag{13.112}$$

donc $\nu(A) = \nu(B \cup N) = \nu(B) = \mu(B)$. La dernière égalité est le fait que ν prolonge μ. Mais par définition de μ' nous avons aussi $\mu'(A) = \mu'(B \cup N) = \mu(B)$. Au final $\mu'(A) = \nu(A) = \mu(B)$.

□

Définition 13.52.
*L'espace mesuré complet $(X, \mathcal{A}, \mu')$ défini par le théorème 13.51 est l'**espace mesuré complétée** de $(X, \mathcal{F}, \mu)$.*

Nous noterons le complété de $(S, \mathcal{F}, \mu)$ par $(S, \hat{\mathcal{F}}, \hat{\mu})$

Théorème 13.53 (Carathéodory[161])**.**
Soit S un ensemble et m^* une mesure extérieure sur S. Alors

(1) l'ensemble $\mathcal{M}$ des parties m^*-mesurables est une tribu,

(2) la restriction de m^* est une mesure sur $(S, \mathcal{M})$,

(3) l'espace mesuré $(S, \mathcal{M}, m^*)$ est complet[8]*.*

Démonstration. Une grosse partie de la preuve sera de prouver la stabilité de $\mathcal{M}$ par union dénombrable quelconque ; cela sera divisé en plusieurs parties.

Tribu : le vide L'ensemble vide est m^*-mesurable.

Tribu : complémentaire Soit $A \in \mathcal{M}$ et $X \in S$. La condition qui dirait $A^c \in \mathcal{M}$ est :

$$m^*(X) = m^*(X \cap A^c) + m^*(X \cap A), \tag{13.113}$$

qui est la même que celle qui dit que A est dans $\mathcal{M}$.

Tribu : union finie Soient $A, B \in \mathcal{M}$ et $X \subset S$. Alors, vu que m^* est une mesure extérieure,

$$\begin{aligned} m^*(X) &\leqslant m^*\big(X \cap (A \cup B)\big) + m^*\big(X \cap (A \cup B)^x\big) && \text{(13.114a)}\\ &= m^*\big((X \cap A) \cup (X \cap B)\big) + m^*\big(X \cap A^c \cap B^c\big). && \text{(13.114b)} \end{aligned}$$

Mais nous pouvons écrire la première union sous forme d'une union disjointe de la façon suivante :

$$(X \cap A) \cup (X \cap B) = (X \cap A) \cup (X \cap B \cap A^c), \tag{13.115}$$

ce qui donne

$$\begin{aligned} m^*(X) &\leqslant m^*(X \cap A) + m^*(X \cap B \cap A^c) + m^*(X \cap A^c \cap B^c) && \text{(13.116a)}\\ &= m^*(X \cap A) + m^*(X \cap A^c) && \text{(13.116b)}\\ &= m^*(X) && \text{(13.116c)} \end{aligned}$$

parce que les deux dernier termes de (13.116a) se somment à $m^*(X \cap A^c)$ parce que $B \in \mathcal{M}$. La dernière ligne est le fait que A soit m^*-mesurable.

Union finie disjointe Soient $\{A_1, \dots, A_n\}$ des éléments deux à deux disjoints de $\mathcal{M}$. Nous allons maintenant prouver par récurrence que

$$m^*\Big(X \cap \big(\bigcup_{k=1}^{n} A_k\big)\Big) = \sum_{k=1}^{n} m^*(X \cap A_k). \tag{13.117}$$

Si $n = 1$ le résultat est évident. Sinon, le fait que A_{n+1} soit m^*-mesurable donne

$$m^*\Big(X \cap \big(\bigcup_{k=1}^{n+1} A_k\big)\Big) = m^*\Big(X \cap \big(\bigcup_{k=1}^{n+1} A_k\big) \cap A_{n+1}\Big) + m^*\Big(X \cap \big(\bigcup_{k=1}^{n+1} A_k\big) \cap A_{n+1}^c\Big). \tag{13.118}$$

Le fait que les A_k soient disjoints implique aussi que

$$X \cap \big(\bigcup_{k=1}^{n+1} A_k\big) \cap A_{n+1} = X \cap A_{n+1} \tag{13.119}$$

8. Définition 13.50.

et

$$X \cap \big(\bigcup_{k=1}^{n+1} A_k\big) \cap A_{n+1}^c = X \cap \big(\bigcup_{k=1}^{n} A_k\big) \tag{13.120}$$

et donc

$$m^*\Big(X \cap \big(\bigcup_{k=1}^{n+1} A_k\big)\Big) = m^*(X \cap A_{n+1}) + m^*\Big(X \cap \big(\bigcup_{k=1}^{n} A_k\big)\Big) \tag{13.121a}$$

$$\stackrel{rec.}{=} m^*(X \cap A_{n+1}) + \sum_{k=1}^{n} m^*(X \cap A_k) \tag{13.121b}$$

$$= \sum_{k=1}^{n+1} m^*(X \cap A_k). \tag{13.121c}$$

La relation (13.117) est prouvée.

Notons qu'en particularisant à $X = S$ nous avons

$$m^*\big(\bigcup_{k=1}^{n} A_k\big) = \sum_{k=1}^{n} m^*(A_k) \tag{13.122}$$

dès que les A_k sont des éléments deux à deux disjoints de $\mathcal{M}$.

Union dénombrable disjointe Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'éléments deux à deux disjoints dans $\mathcal{M}$. Nous allons prouver les choses suivantes :

— $\bigcup_n A_n \in \mathcal{M}$

— $m^*\big(\bigcup_n A_n\big) = \sum_n m^*(A_n)$

où toutes les sommes et union sur n sont entre 1 et ∞.

Nous posons $A = \bigcup_k A_k$ et $B_n = \bigcup_{k=1}^{n} A_k$. Nous savons que $B_n \in \mathcal{M}$ pour tout n par le point précédent. Donc si $X \in S$ nous avons

$$m^*(X) = m^*(X \cap B_n) + m^*(X \cap B_n^c) \tag{13.123a}$$

$$= \sum_{k=1}^{n} m^*(X \cap A_k) + m^*(X \cap B_n^x) \tag{13.123b}$$

$$\geqslant \sum_{k=1}^{n} m^*(X \cap A_k) + m^*(X \cap A^c) \tag{13.123c}$$

où nous avons utilisé la relation (13.117) sur les B_n ainsi que le fait que $A^c \subset B_n^c$ (parce que $B_n \subset A$). L'inégalité (13.123a) étant vraie pour tout n, elle est vraie à la limite :

$$m^*(X) \geqslant \sum_{k=1}^{\infty} m^*(A \cap A_k) + m^*(X \cap A^c) \tag{13.124a}$$

$$\geqslant m^*\Big(\bigcup_k (X \cap A_k)\Big) + m^*(X \cap A^c) \tag{13.124b}$$

$$= m^*\Big(X \cap \big(\bigcup_k A_k\big)\Big) + m^*(X \cap A^c) \tag{13.124c}$$

$$= m^*(X \cap A) + m^*(X \cap A^c), \tag{13.124d}$$

ce qui signifie que $A \in \mathcal{M}$. La première des deux choses que nous voulions montrer est faite. En la particularisant à $X = A$ et en tenant compte des faits que $A \cap A_k = A_k$ et $A \cap A^c = \varnothing$,

$$m^*(A) \geqslant \sum_{k=1}^{\infty} m^*(A \cap A_k) + m^*(A \cap A^c), \tag{13.125}$$

c'est à dire que pour tout n nous avons

$$m^*\big(\bigcup_{k\in\mathbb{N}} A_k\big) \geqslant \sum_{k=1}^{n} m^*(A_k). \tag{13.126}$$

L'inégalité est encore vraie à la limite, et l'inégalité inverse étant toujours vraie pour une mesure extérieure,

$$m^*\big(\bigcup_{k\in\mathbb{N}} A_k\big) = \sum_{k=1}^{\infty} m^*(A_k). \tag{13.127}$$

Union dénombrable quelconque Soit maintenant une suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{M}$ que nous ne supposons plus être disjoints. Nous nous ramenons au cas disjoint en posant

$$\begin{cases} B_1 = A_1 & \text{(13.128a)}\\ B_n = A_n \cap \big(\bigcup_{k=1}^{n-1} A_k\big)^c, & \text{(13.128b)} \end{cases}$$

c'est à dire que nous mettons dans B_n les éléments de A_n qui ne sont dans aucun des A_k précédents. Autrement dit, nous posons $B_0 = \varnothing$ et $B_n = A_n \backslash B_{n-1}$. L'ensemble $\mathcal{M}$ étant stable par réunion finie, par complément et par intersection finie nous avons $B_n \in \mathcal{M}$. De plus les B_n sont disjoints, donc

$$\bigcup_{k=1}^{\infty} A_k = \bigcup_{k=1}^{\infty} B_k \in \mathcal{M}. \tag{13.129}$$

La première égalité se justifie de la façon suivante : si $x \in \bigcup_{k=1}^{\infty} A_k$ alors nous notons n_0 le plus petit n tel que $x \in A_n$ et alors $x \in B_{n_0}$.

Espace complet Nous prouvons à présent que $(S, \mathcal{M}, m^*)$ est un espace mesuré complet. Soit N une partie m^*-négligeable de S et $Y \in \mathcal{M}$ tel que $m^*(Y) = 0$ et $N \subset Y$. D'abord $m^*(N) = 0$ parce que

$$m^*(N) \leqslant m^*(Y) = 0. \tag{13.130}$$

Si $X \subset S$ nous avons

$$X \cap N \subset N \Rightarrow m^*(X \cap N) = 0 \tag{13.131a}$$
$$X \cap N^c \subset X \Rightarrow m^*(X \cap N^c) \leqslant m^*(X). \tag{13.131b}$$

Donc

$$m^*(X \cap N) + m^*(X \cap N^c) \leqslant m^*(X), \tag{13.132}$$

ce qui montre que N est est m^*-mesurable.

□

13.54.
Ce théorème nous pousse à adopter de la notation. Lorsqu'un espace mesuré $(S, \mathcal{F}, \mu)$ est donné, nous noterons

$$(S, \mathcal{M}, \mu^*) \tag{13.133}$$

l'espace mesuré construit de la façon suivante. D'abord μ^* est la mesure extérieure associée à μ par la proposition 13.43. Ensuite $\mathcal{M}$ est la tribu des parties μ^*-mesurables, qui est bien une tribu parce que μ^* est une mesure extérieure (13.53). La proposition (13.46) dit alors que $\mathcal{F} \subset \mathcal{M}$. De plus 13.53 nous explique que si $A \in \mathcal{F}$ alors $\mu(A) = \mu^*(A)$. Tout cela pour dire que

$$(S, \mathcal{F}, \mu) \subset (S, \mathcal{M}, \mu^*). \tag{13.134}$$

Et enfin, 13.53 nous dit que l'espace mesuré $(S, \mathcal{M}, \mu^*)$ est complet.

Exemple 13.55
Montrons un cas dans lequel $(S, \mathcal{M}, \mu^*)$ n'est pas σ-fini. Soit S un ensemble non dénombrable et $\mathcal{F}$ la tribu des parties de S qui sont soit fini ou dénombrables soit de complémentaire fini ou dénombrable. Nous y mettons la mesure

$$\mu(A) = \begin{cases} 0 & \text{si } A \text{ est au plus dénombrable} \\ \infty & \text{sinon.} \end{cases} \tag{13.135}$$

Cette mesure n'est pas σ-finie parce qu'aucune union de dénombrables est non dénombrable. De plus $(S, \mathcal{F}, \mu)$ est complet parce que toute partie contenue dans un ensemble fini ou dénombrable est fini ou dénombrable (2.17).

$\mathcal{F}$ **n'est pas** $\mathcal{P}(S)$ La tribu $\mathcal{F}$ est différente de $\mathcal{P}(S)$. En effet S étant infini, il existe par 2.11 une bijection $\varphi\colon \{1,2\} \times S \to S$. Alors l'ensemble $\varphi\big(\{1\} \times S\big)$ est non dénombrable et son complémentaire

$$\varphi\big(\{1\} \times S\big)^c = \varphi\big(\{2\} \times S\big) \tag{13.136}$$

n'est as dénombrable non plus. Cet ensemble n'est donc pas de $\mathcal{F}$.

$\mathcal{M}$ **est** $\mathcal{P}(S)$ En effet, soit $A \subset S$; il faut prouver que pour tout $X \subset S$ nous avons

$$\mu^*(X) = \mu^*(X \cap A) + \mu^*(X \cap A^c). \tag{13.137}$$

Nous prouvons cela en séparant les cas suivant que X est dénombrable ou non.
Si X est fini ou dénombrable, alors $X \cap A$ et $X \cap A^c$ le sont également et nous avons $\mu^*(X) = \mu(X) = 0$ ainsi que $\mu^*(X \cap A) = \mu^*(X \cap A^c) = 0$.
Si au contraire X n'est pas dénombrable,

$$\mu^*(X) = \inf_{\substack{A \in \mathcal{F} \\ X \subset A}} \mu(A) = \infty, \tag{13.138}$$

parce que X n'étant pas dénombrable, l'ensemble A ne l'est pas non plus et $\mu(A) = \infty$. Mais comme X n'est pas dénombrable, soit $X \cap A$ soit $X \cap A^c$ (soit les deux) n'est pas dénombrable non plus ; par conséquent

$$\mu^*(X \cap A) + \mu^*(X \cap A^c) = \infty. \tag{13.139}$$

Par conséquent $(S, \mathcal{F}, \mu) \neq (S, \mathcal{M}, \mu^*)$. Mais vu que $(S, \mathcal{F}, \mu)$ est complété nous devons avoir $(S, \mathcal{F}, \mu) = (S, \hat{\mathcal{F}}, \hat{\mu})$. Tout cela pour dire que nous avons un exemple avec

$$(S, \mathcal{M}, \mu^*) \neq (S, \hat{\mathcal{F}}, \hat{\mu}). \tag{13.140}$$

△

Nous avons deux façons de créer un espace complet à partir de $(S, \mathcal{F}, \mu)$.

(1) Partir de la mesure extérieure μ^* et construire $(S, \mathcal{M}, \mu^*)$.
(2) Partir des ensembles μ-négligeables, construire $\hat{\mathcal{F}}$ et ensuite $(S, \hat{\mathcal{F}}, \hat{\mu})$.

Ces deux façons ne sont pas équivalentes en général comme le montre l'exemple 13.55. Mais il sera montré par la proposition 13.59 que si $(S, \mathcal{F}, \mu)$ est σ-fini alors les deux sont équivalent.

Lemme 13.56.
Soit $(S, \mathcal{F}, \mu)$ un espace mesuré. Alors pour tout $X \subset S$ tel que $\mu^*(X) < \infty$ il existe $A \in \mathcal{F}$ tel que $X \subset A$ et $\mu^*(X) = \mu(A)$.

C'est à dire que μ^* a beau être défini sur toutes les parties de S, ce qu'il faut rajouter pour être μ-mesurable, c'est pas grand chose.

Démonstration. Par définition de la mesure extérieure associée à μ en tant qu'infimum, pour tout $n \geqslant 1$, il existe $A_n \in \mathcal{F}$ tel que $X \subset A_n$ et $\mu(A_n) \leqslant \mu^*(X) + \frac{1}{2^n}$. Nous posons $A = \bigcap_{n\geqslant 1} A_n$ et nous vérifions que ce A fait l'affaire.

D'abord $A \in \mathcal{F}$ parce qu'une tribu est stable par union dénombrable. Ensuite pour tout $n \geqslant 1$ nous avons

$$\mu(A) \leqslant \mu(A_n) \leqslant \mu^*(X) + \frac{1}{2^n}, \tag{13.141}$$

et à la limite $\mu(A) \leqslant \mu^*(X)$. Mais $X \subset A$ implique $\mu^*(X) \leqslant \mu(A)$ parce que $\mu^*(X)$ l'infimum d'un ensemble contenant $\mu(A)$. □

Corollaire 13.57.
Soit une mesure μ et la mesure extérieure μ^* associée[9]*. Une partie N de X est négligeable si et seulement si $\mu^*(N) = 0$.*

Démonstration. Si μ^* est la mesure extérieure associée à μ et si N est μ-négligeable alors $\mu^*(N) = 0$ parce que

$$\mu^*(N) \leqslant \mu^*(Y) = \mu(Y) = 0 \tag{13.142}$$

pour un certain Y mesurable de mesure nulle contenant N.

D'autre part si $\mu^*(N) = 0$ alors le lemme 13.56 donne une partie mesurable A telle que $N \subset A$ et $\mu(A) = 0$, c'est à dire que N est négligeable. □

Lemme 13.58.
Si l'espace mesuré $(S, \mathcal{F}, \mu)$ est σ-fini alors l'espace mesuré $(S, \mathcal{M}, \mu^*)$ est également σ-fini.

Démonstration. Vu que $(S, \mathcal{F}, \mu)$ est σ-fini, nous avons une suite croissante A_n d'éléments de $\mathcal{F}$ tels que $\bigcup_n A_n = S$ et telle que $\mu(A_n) < \infty$ pour tout n. Étant donné que $\mathcal{F} \subset \mathcal{M}$, cette suite convient également pour montrer que $(S, \mathcal{M}, \mu^*)$ est σ-fini parce que $\mu^*(A_n) = \mu(A_n) < \infty$. □

La proposition suivante montre que si $(S, \mathcal{F}, \mu)$ est σ-finie alors nous avons l'égalité.

Proposition 13.59.
Soit $(S, \mathcal{F}, \mu)$ un espace mesuré σ-fini, μ^* la mesure extérieure associée et $\mathcal{M}$ la tribu des ensembles μ^*-mesurables[10]*. Alors*

$$(S, \mathcal{M}, \mu^*) = (S, \hat{\mathcal{F}}, \hat{\mu}). \tag{13.143}$$

Démonstration. La proposition 13.46 indique que tous les éléments de $\mathcal{F}$ sont μ^*-mesurables, c'est à dire que $\mathcal{F} \subset \mathcal{M}$. Mais l'espace $(S, \mathcal{M}, \mu^*)$ est complet par le théorème de Carathéodory 13.53, donc par minimalité du complété (13.51(7)),

$$(S, \hat{\mathcal{F}}, \hat{\mu}) \subset (S, \mathcal{M}, \mu^*) \tag{13.144}$$

au sens où $\hat{\mathcal{F}} \subset \mathcal{M}$ et si $A \in \hat{\mathcal{F}}$ alors $\hat{\mu}(A) = \mu^*(A)$. Notons que cette inclusion est vraie même si la mesure n'est pas σ-finie.

Nous passons à l'inclusion inverse. Soit $A \in \mathcal{M}$, c'est à dire que pour tout $Y \subset S$ nous avons

$$\mu^*(Y) = \mu^*(Y \cap A) + \mu^*(Y \cap A^c). \tag{13.145}$$

Nous allons montrer que $A \in \hat{\mathcal{F}}$ en séparant les cas suivant que $\mu^*(A) = \infty$ ou non.

Si $\mu^*(A) < \infty$ Par le lemme 13.56, il existe $X \in \mathcal{F}$ tel que $A \subset X$ et $\mu^*(A) = \mu(X)$. Vu que $(S, \mathcal{F}, \mu) \subset (S, \mathcal{M}, \mu^*)$ nous avons alors

$$\mu^*(A) = \mu(X) = \mu^*(X). \tag{13.146}$$

9. Par la proposition 13.43.
10. C'est bien une tribu par 13.53(1).

Nous écrivons la relation (13.145) avec ce X en guise de Y, et en nous souvenant que $X \cap A = A$ et $X \cap A^c = X\backslash A$:

$$\mu^*(X) = \mu^*(A) + \mu^*(X\backslash A). \tag{13.147}$$

En tenant compte de (13.146) et du fait que $\mu^*(A) < \infty$, nous pouvons simplifier et trouver $\mu^*(X\backslash A) = 0$. Le lemme 13.56 nous donne alors $B \in \mathcal{F}$ tel que $X\backslash A \subset B$ et $\mu(B) = \mu^*(X\backslash A) = 0$, c'est à dire que $X\backslash A$ est μ-négligeable. Par conséquent $X\backslash A \in \hat{\mathcal{F}}$. En écrivant

$$A = X\backslash(X\backslash A), \tag{13.148}$$

nous avons écrit A comme différence de deux éléments de $\hat{\mathcal{F}}$ et nous concluons que $A \in \hat{\mathcal{F}}$.

Si $\mu^*(A) < \infty$ Le lemme 13.58 nous indique que $(S, \mathcal{M}, \mu^*)$ est σ-fini et il existe donc une suite $(S_n)_{n\geqslant 1}$ dans $\mathcal{M}$ telle que $\bigcup_n S_n = S$ et $\mu^*(S_n) < \infty$. L'ensemble $A \cap S_n$ est un élément de $\mathcal{M}$ vérifiant

$$\mu^*(A \cap S_n) \leqslant \mu^*(A) < \infty, \tag{13.149}$$

ce qui implique que $A \cap S_n \in \hat{\mathcal{F}}$ par la première partie. Maintenant $A = \bigcup_n (A \cap S_n) \in \hat{\mathcal{F}}$ par union dénombrable d'éléments de la tribu $\hat{\mathcal{F}}$.

□

13.2.5 Prolongement

Théorème 13.60 (Prolongement de Hahn[161])**.**
Soit $\mathcal{A}$ une algèbre de parties d'un ensemble S et μ une mesure sur $(S, \mathcal{A})$. Soit $\mathcal{F} = \sigma(\mathcal{A})$ la tribu engendrée par $\mathcal{A}$. Alors

(1) La mesure μ se prolonge en une mesure m sur $\mathcal{F}$.

(2) Si μ est σ-finie alors le prolongement est unique et m est σ-finie.

(3) Si μ est finie, alors m l'est aussi.

Démonstration. La proposition 13.22 nous donne une mesure extérieure μ^* sur S dont la restriction à $\mathcal{A}$ est μ. Si $\mathcal{M}$ est la tribu des parties μ^*-mesurables de S alors le théorème de Carathéodory 13.53 nous dit que $(S, \mathcal{M}, \mu^*)$ est un espace mesuré.

$\mathcal{A} \subset \mathcal{M}$ Cette partie est une adaptation de ce qui a déjà été fait dans la preuve de la proposition 13.46. Soit $A \in \mathcal{A}$ et $X \in S$; nous devons prouver la relation de la définition 13.44. Vu que μ^* est une mesure extérieure nous avons automatiquement

$$\mu^*(A) \leqslant \mu^*(X \cap A) + \mu^*(X \cap A^c). \tag{13.150}$$

Il reste à prouver l'inégalité inverse. Soit une suite B_k d'éléments de $\mathcal{A}$ telle que $X \subset \bigcup_k B_k$; nous avons alors

$$\mu^*(X \cap A) \leqslant \mu^*\big(\bigcup_{k=1}^{\infty} B_k \cap A\big) \leqslant \sum_{k=1}^{\infty} \mu^*(B_k \cap A) = \sum_k \mu(B_k \cap A) \tag{13.151}$$

où nous avons utilisé la définition 13.17(3) ainsi que le lemme 13.19. De la même façon,

$$\mu^*(X \cap A^c) \leqslant \sum_k \mu(B_k \cap A^c). \tag{13.152}$$

Mettant les deux bouts ensemble, en remarquant que $B_k \cap A \in \mathcal{A}$ et donc que $\mu^*(B_k \cap A) = \mu(B_k \cap A)$,

$$\mu^*(X \cap A) + \mu^*(X \cap A^c) \leqslant \sum_k \mu(B_k \cap A) + \mu(B_k \cap A^c) = \sum_k \mu(B_k). \tag{13.153}$$

La somme $\mu^*(X\cap A)+\mu^*(X\cap A^c)$ est donc inférieure à chacun des éléments de l'ensemble sur lequel on prend l'infimum pour définir [11] $\mu^*(X)$, donc

$$\mu^*(X\cap A)+\mu^*(X\cap A^c)\leqslant\mu^*(X). \tag{13.154}$$

A fortiori nous avons $\sigma(\mathcal{A})\subset\mathcal{M}$ et donc $(S,\sigma(\mathcal{A}),\mu^*)$ est un espace mesuré. Cela prouve l'existence d'une mesure prolongeant μ à $\sigma(\mathcal{A})$.

Unicité Nous supposons à présent que μ est σ-finie. Soient m_1 et m_2 deux mesures prolongeant μ et définies sur une tribu contenant $\mathcal{A}$. Nous posons

$$\mathcal{C}=\{A\in\mathcal{A}\text{ tel que }\mu(A)<\infty\}. \tag{13.155}$$

Dans l'optique d'utiliser le théorème d'unicité des mesures 13.40, nous prouvons que $\sigma(\mathcal{A})=\sigma(\mathcal{C})$. Vu que μ est σ-finie, il existe une suite croissante (S_n) d'éléments de $\mathcal{A}$ telle que $S=\bigcup_n S_n$ et $\mu(S_n)\in\mathcal{C}$. Alors si $A\in\mathcal{A}$ nous avons $A=\bigcup_n(A\cap S_n)$, et donc $A\in\sigma(\mathcal{C})$. Donc $\mathcal{A}\subset\sigma(\mathcal{C})$. Mais étant donné que $\mathcal{C}\subset\mathcal{A}$ nous avons aussi $\sigma(\mathcal{C})\subset\sigma(\mathcal{A})$. Au final $\sigma(\mathcal{A})=\sigma(\mathcal{C})$.

Les mesures m_1 et m_2 sont des mesures sur $\sigma(\mathcal{C})$ coïncidant sur $\mathcal{C}$ (parce que $\mathcal{C}\subset\mathcal{A}$). De plus la classe $\mathcal{C}$ est stable par intersection finie et contient une suite croissante dont l'union est S (parce que μ est σ-finie).

Le théorème 13.40 nous dit alors que m_1 et m_2 coïncident sur $\sigma(\mathcal{C})=\sigma(\mathcal{A})$.

Extension finie et σ-finie Enfin si μ est σ-finie il existe $S_n\in\mathcal{A}$ avec $\mu(S_n)<\infty$ et $\bigcup_n S_n=S$. Ces ensembles vérifient tout autant $m(S_n)=\mu(S_n)<\infty$ pour tout prolongement m de μ.

Idem si μ est finie, tout prolongement est fini.

□

Exemple 13.61([161])
Soit $\mathcal{A}$, l'algèbre de parties de $\mathbb{R}$ formée par les réunions finies d'intervalles de la forme $]-\infty,a[$, $[a,b[$ et $[b,+\infty[$ avec $-\infty<a\leqslant b<+\infty$. Notons que les singletons ne font pas partie de $\mathcal{A}$ parce que $[a,a[=\varnothing$. Nous posons

$$\mu(A)=\begin{cases}0 & \text{si } A=\varnothing\\ \infty & \text{sinon.}\end{cases} \tag{13.156}$$

Cela donne une mesure (non σ-finie) sur $(\mathbb{R},\mathcal{A})$.

Nous allons prouver que la tribu engendrée par $\mathcal{A}$ est la tribu des boréliens et que μ accepte (au moins) deux prolongements distincts à $\sigma(\mathcal{A})$.

D'abord nous avons

$$]a,b[=\big(-\infty,a[\cup[b,+\infty[\big)\cap[a,b[, \tag{13.157}$$

donc toutes les boules ouvertes appartiennent à $\sigma(\mathcal{A})$. Ces dernières comprenant une base dénombrable de la topologie de $\mathbb{R}$ (par la proposition 6.60), tous les ouverts de $\mathbb{R}$ sont dans $\sigma(\mathcal{A})$. Par conséquent $\mathcal{B}\text{or}(\mathbb{R})\subset(\mathbb{R}^d)$. Mais en même temps tous les éléments de $\mathcal{A}$ sont des boréliens, donc $\mathcal{B}\text{or}(\mathbb{R})=\sigma(\mathcal{A})$ parce que la fermeture en tant qu'algèbre de parties est plus petite que la fermeture en tant que tribu.

La mesure de comptage prolonge μ parce qu'à part l'ensemble vide, tous les éléments de $\mathcal{A}$ sont infinis. Notons que les singletons sont dans $\sigma(\mathcal{A})$, donc la mesure de comptage prend d'autres valeurs que 0 et $+\infty$.

Par ailleurs la mesure

$$\mu'(A)=\begin{cases}0 & \text{si } A=\varnothing\\ +\infty & \text{sinon}\end{cases} \tag{13.158}$$

est également une mesure prolongeant μ à $\sigma(\mathcal{A})=\mathcal{B}\text{or}(\mathbb{R})$.

11. Définition 13.28.

La mesure de comptage et μ' sont deux prolongements distincts de μ. △

Exemple 13.62([161])
Nous montrons maintenant une mesure non σ-finie qui se prolonge en deux mesures distinctes, toutes deux σ-finies.

Nous considérons la même algèbre $\mathcal{A}$ de parties que celle donnée dans l'exemple 13.61, mais cette fois vue sur $\mathbb{Q}$ uniquement. La mesure de comptage m sur $(\mathbb{Q}, \mathcal{A})$ n'est pas σ-finie.

Vu que les singletons sont des boréliens nous avons $\sigma(\mathcal{A}) = \mathcal{P}(\mathbb{Q})$, ce qui fait que $(\mathbb{Q}, \sigma(\mathcal{A}), m)$ est un prolongement σ-fini de m. L'espace mesuré $(\mathbb{Q}, \sigma(\mathcal{A}), 2m)$ est également σ-fini et est un prolongement distinct de $(\mathbb{Q}, \mathcal{A}, m)$. △

13.3 Applications mesurables

13.3.1 Propriétés

Définition 13.63 (Fonction mesurable)**.**
Soit $(E, \mathcal{A})$ et $(F, \mathcal{F})$ deux espaces mesurés. Une fonction $f\colon E \to F$ est ***mesurable*** *si pour tout $\mathcal{O} \in \mathcal{F}$, l'ensemble $f^{-1}(\mathcal{O})$ est dans $\mathcal{A}$.*

Définition 13.64 (Fonction borélienne)**.**
Une application $f\colon (\Omega, \mathcal{A}) \to (\mathbb{R}^d, \mathcal{B}or(\mathbb{R}^d))$ est ***borélienne*** *si elle est mesurable, c'est à dire si pour tout $B \in \mathcal{B}or(\mathbb{R}^d)$ nous avons $f^{-1}(B) \in \mathcal{A}$.*

Si rien n'est précisé, une application entre deux espaces topologiques est borélienne lorsqu'elle est mesurable en considérant la tribu borélienne sur les deux *espace.*

Si $\mathcal{A}$ est une tribu sur un ensemble E, nous notons $m(\mathcal{A})$ l'ensemble des fonctions qui sont $\mathcal{A}$-mesurables.

Le plus souvent lorsque nous parlerons de fonctions $f\colon X \to Y$ où Y est un espace topologique, nous considérons la tribu borélienne sur Y. Ce sera en particulier le cas dans la théorie de l'intégration.

Remarque 13.65.
Lorsque nous considérons des fonctions à valeurs réelles $f\colon X \to \mathbb{R}$ nous utiliserons toujours la tribu borélienne sur $\mathbb{R}$. Pour X, cela peut dépendre des contextes. En théorie de l'intégration, nous mettrons sur X la tribu des ensembles mesurables au sens de Lebesgue sur X, *tout en gardant celle des boréliens sur l'ensemble d'arrivée.*

Pour toute la partie sur l'intégration, une fonction $f\colon \mathbb{R}^n \to \mathbb{R}^m$ sera mesurable si pour tout borélien A de $\mathbb{R}^m$ l'ensemble $f^{-1}(A)$ est Lebesgue-mesurable dans $\mathbb{R}^n$.

Étant donné qu'il est franchement difficile de créer des ensembles non mesurables au sens de Lebesgue, il est franchement difficile de créer des fonction non mesurables à valeurs réelles. L'hypothèse de mesurabilité est donc toujours satisfaite dans les cas pratiques.

Proposition 13.66.
Soient $(S_i, \mathcal{F}_i)$ $(i = 1, 2, 3)$ des espaces mesurables et des fonctions mesurables $f\colon S_1 \to S_2$ et $g\colon S_2 \to S_3$. Alors la fonction $g \circ f\colon S_1 \to S_3$ est mesurable.

Démonstration. Soit $B \in \mathcal{F}_3$. Alors

$$(g \circ f)^{-1}(B) = f^{-1}\big(g^{-1}(B)\big) \in f^{-1}(\mathcal{F}_2) \subset \mathcal{F}_1. \tag{13.159}$$

□

13.3.2 D'une tribu à l'autre

Lemme 13.67 ([166])**.**
Soit une application $f\colon S_1 \to S_2$ et une tribu $\mathcal{F}_2$ sur S_2. Alors $f^{-1}(\mathcal{F}_2)$ est une tribu sur S_1

Démonstration. Il faut prouver les trois propriétés de la définition 13.1 d'une tribu.

(1) D'abord f est définit sur tout S_1, donc $f^{-1}(S_2) = S_1$ alors que $S_2 \in \mathcal{F}_2$.

(2) Soit $A \in f^{-1}(\mathcal{F}_2)$, c'est à dire $A = f^{-1}(B)$ pour un certain $B \in \mathcal{F}_2$. En ce qui concerne le complémentaire :

$$A^c = f^{-1}(B)^c = S_1 \backslash f^{-1}(B) = f^{-1}(S_2 \backslash B) = f^{-1}(B^c). \tag{13.160}$$

(3) Si $(A_i)_{i\in\mathbb{N}}$ sont des éléments de $f^{-1}(\mathcal{F}_2)$ avec $A_i = f^{-1}(B_i)$ alors

$$\bigcup_i A_i = \bigcup_i f^{-1}(B_i) = f^{-1}\big(\bigcup_i B_i\big). \tag{13.161}$$

Ce qui est dans la dernière parenthèse est dans $\mathcal{F}_2$ parce que cette dernière est une tribu.

□

Lemme 13.68 ([166])**.**
Soit une application $f\colon S_1 \to S_2$ et $\mathcal{F}$ une tribu de S_1. Alors

(1) L'ensemble

$$\mathcal{F}_f = \{B \subset S_2 \text{ tel que } f^{-1}(B) \in \mathcal{F}\} \tag{13.162}$$

est une tribu sur S_2.

(2) C'est la plus grande tribu de S_2 pour laquelle f est mesurable.

Démonstration. Encore les trois propriétés à vérifier.

(1) $S_2 \in \mathcal{F}$, sont $S_1 = f^{-1}(S_2) \in \mathcal{F}_f$.

(2) Si $A \in \mathcal{F}_f$ alors $A = f^{-1}(B)$ pour un certain $B \in \mathcal{F}$. Nous avons alors aussi $B^c \in \mathcal{F}$ et donc

$$f^{-1}(B^c) = f^{-1}(B)^c = A^c. \tag{13.163}$$

Par conséquent A^c est dans $\mathcal{F}_f$.

(3) Si (A_i) sont des éléments de $\mathcal{F}_f$ avec $A_i = f^{-1}(B_i)$ pour $B_i \in \mathcal{F}$ alors $\bigcup_i B_i \in \mathcal{F}$ et

$$f^{-1}\big(\bigcup_i B_i\big) = \bigcup_i f^{-1}(B_i) \in \mathcal{F}_f. \tag{13.164}$$

En ce qui concerne la maximalité, si $R \subset S_2$ n'est pas dans $\mathcal{F}_f$ alors $f^{-1}(R)$ n'est pas dans $\mathcal{F}$ et donc f ne serait pas mesurable. □

Définition 13.69 (Tribu engendrée)**.**
Soit une application $f\colon S_1 \to S_2$ et $\mathcal{F}$ une tribu de S_1. Alors conformément au lemme 13.68 l'ensemble

$$\mathcal{F}_f = \{B \subset S_2 \text{ tel que } f^{-1}(B) \in \mathcal{F}\} \tag{13.165}$$

*est la **tribu engendrée**.*

Lemme 13.70 (Lemme de transfert)**.**
Soit $f\colon S_1 \to S_2$ une application et une classe $\mathcal{C}$ de parties de S_2. Alors

$$\sigma\big(f^{-1}(\mathcal{C})\big) = f^{-1}\big(\sigma(\mathcal{C})\big). \tag{13.166}$$

Démonstration. Vu que $\sigma(\mathcal{C})$ es tune tribu, dans S_2 alors le lemme 13.68 dit que $f^{-1}\big(\sigma(\mathcal{C})\big)$ est une tribu qui contient en particulier $f^{-1}(\mathcal{C})$. Nous en déduisons que $\sigma\big(f^{-1}(\mathcal{C})\big) \subset f^{-1}\big(\sigma(\mathcal{C})\big)$.

Réciproquement. Dans S_1 nous avons la tribu $\sigma\big(f^{-1}(\mathcal{C})\big)$. Nous pouvons alors considérer la tribu

$$\mathcal{F}_f = \{B \subset S_2 \text{ tel que } f^{-1}(B) \in \sigma\big(f^{-1}(\mathcal{C})\big)\}. \tag{13.167}$$

Montrons que $\mathcal{C} \subset \mathcal{F}_f$. Lorsque $B \in \mathcal{C}$ nous avons $f^{-1}(B) \in f^{-1}(\mathcal{C}) \subset \sigma\big(f^{-1}(\mathcal{C})\big)$. Du coup $B \in \mathcal{F}_f$. Nous avons alors, en passant aux tribus engendrées :

$$\sigma(\mathcal{C}) \subset \sigma(\mathcal{F}_f) = \mathcal{F}_f. \tag{13.168}$$

Si maintenant $B \in \sigma(\mathcal{C})$, nous avons $f^{-1}(B) \in \sigma\big(f^{-1}(\mathcal{C})\big)$, ce qui signifie que

$$f^{-1}\big(\sigma(\mathcal{C})\big) \subset \sigma\big(f^{-1}(\mathcal{C})\big). \tag{13.169}$$

□

Le théorème suivant est important pour prouver qu'une application est mesurable. En effet, il permet de ne tester si une application est mesurable uniquement sur une partie génératrice de la tribu d'arrivé [12].

Théorème 13.71.
Soient des espaces mesurables $(S_1, \mathcal{F}_1)$ et $(S_2, \mathcal{F}_2)$ ainsi qu'une application $f\colon S_1 \to S_2$. S'il existe un ensemble de parties $\mathcal{C}$ de S_2 tel que
— $\sigma(\mathcal{C}) = \mathcal{F}_2$
— $f^{-1}(B) \in \mathcal{F}_1$ *pour tout* $B \in \mathcal{C}$
alors f est mesurable.

Démonstration. Par hypothèse, $\sigma(\mathcal{C}) = \mathcal{F}_2$ et $f^{-1}(\mathcal{C}) \subset \mathcal{F}_1$ et nous pouvons utiliser le lemme de transfert 13.70 :

$$\sigma\big(f^{-1}(\mathcal{C})\big) = f^{-1}\big(\sigma(\mathcal{C})\big) \tag{13.170}$$

qui s'écrit ici

$$\sigma\big(f^{-1}(\mathcal{C})\big) = f^{-1}(\mathcal{F}_2). \tag{13.171}$$

Mais vu que $f^{-1}(\mathcal{C}) \subset \mathcal{F}_1$, nous avons aussi $\sigma\big(f^{-1}(\mathcal{C})\big) \subset \mathcal{F}_1$, ce qui signifie que

$$f^{-1}(\mathcal{F}_2) \subset \mathcal{F}_1. \tag{13.172}$$

Cela est exactement le fait que f soit mesurable. □

Le théorème suivant est très important parce qu'en pratique c'est souvent lui, en conjonction avec la proposition 13.111 qui permet de déduire qu'une fonction est borélienne.

Théorème 13.72 ([166])**.**
Soient X et Y deux espaces topologiques. Alors toute application continue $f\colon X \to Y$ est borélienne [13].

Démonstration. Pour vérifier que f est borélienne, nous devons prouver que $f^{-1}(B)$ est borélien pour tout borélien B de Y. Heureusement, le théorème 13.71 nous permet de limiter la vérification aux B appartenant à une classe engendrant les boréliens de Y.

La classe en question est toute trouvée : ce sont les ouverts. Si $\mathcal{O}$ est un ouvert de Y alors $f^{-1}(\mathcal{O})$ est un ouvert de X et donc un borélien de X. □

Le théorème suivant donne une importante compatibilité entre l'induction de tribu et l'induction de topologie : la tribu induite à partir des boréliens sur un sous-espace topologique est la tribu des boréliens pour la topologie induite.

Théorème 13.73 ([166])**.**
Soit X, un espace topologique et $Y \subset X$ une partie munie de la topologie induite. Alors

$$\mathcal{B}or(Y) = \mathcal{B}or(X)_Y \tag{13.173}$$

où $\mathcal{B}or(X)_Y$ est la tribu sur Y induite de $\mathcal{B}or(X)$ par la définition 13.7.

12. Typiquement les ouverts pour les boréliens.
13. Définition 13.64.

Démonstration. Nous notons τ_X et τ_Y les topologies de X et Y.

$\underline{\mathcal{B}\mathbf{or}(Y) \subset \mathcal{B}\mathbf{or}(X)_Y}$ Si $A \in \tau_Y$ alors $A = Y \cap \Omega$ pour un $\Omega \in \tau_X$. Mais vu que Ω est un ouvert de X, il est un borélien de X, ce qui donne que $Y \cap \Omega$ est un élément de $\mathcal{B}\mathrm{or}(X)_Y$. Cela prouve que $\tau_Y \subset \mathcal{B}\mathrm{or}(X)_Y$, c'est à dire que $\mathcal{B}\mathrm{or}(X)_Y$ est une tribu sur Y contenant les ouverts de Y. Nous avons donc

$$\mathcal{B}\mathrm{or}(X) \subset \mathcal{B}\mathrm{or}(X)_Y. \tag{13.174}$$

$\underline{\textbf{Réciproquement}}$ L'application $\mathrm{Id}\colon (Y, \tau_Y) \to (X, \tau_X)$ est continue parce que si Ω est ouvert de X alors $\mathrm{Id}^{-1}(\Omega) = \Omega \cap Y \in \tau_Y$. Par conséquent l'identité est une application borélienne (théorème 13.72), ce qui signifie que $\mathrm{Id}^{-1}\big(\mathcal{B}\mathrm{or}(X)\big) \subset \mathcal{B}\mathrm{or}(Y)$, ou encore que si $B \in \mathcal{B}\mathrm{or}(X)$, alors $\mathrm{Id}^{-1}(B) = B \cap Y \in \mathcal{B}\mathrm{or}(Y)$. Cela signifie que

$$\mathcal{B}\mathrm{or}(X)_Y \subset \mathcal{B}\mathrm{or}(Y). \tag{13.175}$$

□

Corollaire 13.74.
Si U est un borélien de l'espace topologique X, alors les boréliens de U sont les boréliens de X inclus à U :

$$\mathcal{B}or(U) = \{B \in \mathcal{B}or(X) \text{ tel que } B \subset U\}. \tag{13.176}$$

Démonstration. Si $B' \in \mathcal{B}\mathrm{or}(U)$, le théorème 13.73 donne un borélien $B \in \mathcal{B}\mathrm{or}(X)$ tel que $B' = B \cap U$. Mais U étant borélien de X, l'intersection $B \cap U$ est encore un borélien de X. □

Ce corollaire s'applique en particulier lorsque U est un ouvert.

13.3.3 Mesure image

Le produit d'une mesure par une fonction est définit par la propriété 13.137.

Proposition-définition 13.75 (Mesure image[166])**.**
Soient $(S_1, \mathcal{F}_1)$ et $(S_2, \mathcal{F}_2)$ des espaces mesurables. Soit $\varphi\colon S_1 \to S_2$ une application mesurable. Si m_1 est une mesure positive sur S_1 alors l'application définie par

$$m_2(A_2) = m_1\big(\varphi^{-1}(A_2)\big) \tag{13.177}$$

est une mesure positive sur $(S_2, \mathcal{F}_2)$.

La mesure m_2 ainsi définie est la ***mesure image*** *de m_1 par l'application φ. Elle est notée $\varphi(m_1)$.*

Démonstration. Il y a deux choses à vérifier pour avoir une mesure positive [14]. D'abord pour l'ensemble vide :

$$m_2(\varnothing) = m_1\big(\varphi^{-1}(\varnothing)\big) = m_1(\varnothing) = 0. \tag{13.178}$$

Ensuite pour l'additivité. Soient A_n dans $\mathcal{F}_2$ des parties deux à deux disjointes et telles que $\bigcup_n A_n \in \mathcal{F}_2$. Alors nous avons

$$\begin{aligned} m_2\big(\bigcup_n A_n\big) &= m_1\Big(\varphi^{-1}\big(\bigcup_n A_n\big)\Big) && \text{(13.179a)}\\ &= m_1\big(\bigcup_n \varphi^{-1}(A_n)\big) && \text{(13.179b)}\\ &= \sum_n m_1\big(\varphi(A_n)\big) && \text{(13.179c)}\\ &= \sum_n m_2(A_n). && \text{(13.179d)} \end{aligned}$$

□

14. Définition 13.23

Lemme 13.76.
Soient deux espaces mesurables $(S_1, \mathcal{F}_1)$ et $(S_2, \mathcal{F}_2)$ ainsi que deux mesures μ et ν sur $(S_1, \mathcal{F}_1)$. Si $\varphi\colon S_1 \to S_2$ est mesurable et si $\mu \leqslant \nu$ alors $\varphi(\mu) \leqslant \varphi(\nu)$.

Démonstration. Soit B mesurable dans $(S_2, \mathcal{F}_2)$ (c'est à dire $B \in \mathcal{F}_2$). Alors

$$\varphi(\mu)(B) = \mu\big(\varphi^{-1}(B)\big) \leqslant \nu\big(\varphi^{-1}(B)\big) = \varphi(\nu)(B). \tag{13.180}$$

□

Il est naturel de se demander comment il faut intégrer par rapport à une mesure image. La réponse sera dans le théorème 13.141.

13.3.4 Régularité d'une mesure

Certaines mesures ont de la compatibilité avec la topologie. Nous allons étudier ça.

Théorème 13.77 ([166])**.**
Soit X un espace métrique et m une mesure positive bornée sur $\big(X, \mathcal{B}or(X)\big)$. Alors si B est un borélien,

(1) Régularité extérieure : $m(B) = \inf\{m(\Omega) \text{où } \Omega \text{ est un ouvert contenant } B\}$

(2) Régularité intérieure : $m(B) = \sup\{m(F) \text{où } F \text{ est un fermé, } F \subset B\}$.

Démonstration. Soit $\mathcal{F}$ l'ensemble des $B \in \mathcal{B}\text{or}(X)$ tels que pour tout $\epsilon > 0$, il existe Ω_ϵ ouvert et F_ϵ fermé tels que $F_\epsilon \subset B \subset \Omega_\epsilon$ et $m(\Omega_\epsilon \backslash F_\epsilon) \leqslant \epsilon$. Nous allons montrer que cela est une tribu contenant les ouverts. Comme cela est inclus à la tribu borélienne, nous en déduirons que $\mathcal{F} = \mathcal{B}\text{or}(X)$.

<u>$\mathcal{F}$ **contient les ouverts**</u> Soit Ω un ouvert de X. Alors Ω^c est fermé et $d(x, \Omega^c) = 0$ si et seulement si $x \in \Omega^c$ par la proposition 6.144. Nous pouvons donc écrire

$$\Omega^c = \bigcap_{n \geqslant 1} \{x \in X \text{ tel que } d(x, \Omega^c) < \frac{1}{n}\}. \tag{13.181}$$

En passant au complémentaire et en posant $F_n = \{x \in X \text{ tel que } d(x, \Omega^c) \geqslant \frac{1}{n}\}$ nous avons

$$\Omega = \bigcup_{n \geqslant 1} F_n. \tag{13.182}$$

Chacun des F_n est fermé parce que F_n est l'image réciproque du fermé $[\frac{1}{n}, \infty[$ par l'application $x \mapsto d(x, \Omega^c)$ qui est continue. De plus les F_n forment une suite croissante, donc le lemme 13.27 nous assure que $m(\Omega) = \lim_{n\to\infty} m(F_n)$. Et le lemme 13.26 que $m(\Omega \backslash F_n) = m(\Omega) - m(F_n)$.

Soit $\epsilon > 0$. Il existe alors $n_\epsilon \geqslant 1$ tel que

$$m(\Omega \backslash F_n) = m(\Omega) - m(F_n) \leqslant \epsilon. \tag{13.183}$$

Bref si Ω est ouvert nous considérons $\Omega_\epsilon = \Omega$ et $F_\epsilon = F_{n_\epsilon}$ et nous avons

$$F_\epsilon \subset \Omega \subset \Omega_\epsilon \tag{13.184}$$

avec $m(\Omega_\epsilon \backslash F_\epsilon) \leqslant \epsilon$.

L'ensemble $\mathcal{F}$ contient les ouverts.

<u>$\mathcal{F}$ **est une tribu**</u> Il y a à vérifier les trois conditions de la définition 13.1.

<u>**Les ensembles faciles**</u> Les ensembles X et $\varnothing$ sont dans $\mathcal{F}$ parce qu'ils sont ouverts et fermés.

Complémentaire Soit $B \in \mathcal{F}$, soit $\epsilon > 0$ et les ensembles F_ϵ et Ω_ϵ qui vont avec. Alors en passant au complémentaire nous avons

$$\Omega_\epsilon^c \subset B^c \subset F_\epsilon^c \tag{13.185}$$

De plus

$$F_\epsilon^c \backslash \Omega_\epsilon^c = F_\epsilon^c \cap (\Omega_\epsilon^c)^c = F_\epsilon^c \cap \Omega_\epsilon = \Omega_\epsilon \backslash F_\epsilon. \tag{13.186}$$

Par conséquent

$$m(F_\epsilon^c \backslash \Omega_\epsilon^c) = m(\Omega_\epsilon \backslash F_\epsilon) \leqslant \epsilon. \tag{13.187}$$

Cela montre que $B^c \in \mathcal{F}$.

Union dénombrable Soient (B_n) une suite d'éléments de $\mathcal{F}$ et $\epsilon > 0$. Pour chaque n nous choisissons un ouvert Ω_n et un fermé F_n tels que $F_n \subset B_n \subset \Omega_n$ et

$$m(\Omega_n \backslash F_n) \leqslant \frac{\epsilon}{2^{n+2}}. \tag{13.188}$$

Vu que $\Omega_n \backslash B_n \subset \Omega_n \backslash F_n$ nous avons aussi

$$m(\Omega_n \backslash B_n) \leqslant m(\Omega_n \backslash F_n) \leqslant \frac{\epsilon}{2^{n+2}}. \tag{13.189}$$

Nous posons $\Omega = \bigcup_{n \geqslant 1} \Omega_n$ (un ouvert) et $B = \bigcup_{n \geqslant 1} B_n$ ainsi que $A = \bigcup_{n \geqslant 1} F_n$ (qui n'est pas spécialement fermé).

Le but est de majorer $m(\Omega \backslash F)$ où F est un fermé qui est encore à déterminer. Calculons déjà ceci :

$$\begin{aligned}
\Omega \backslash B &= \bigcup_n \Omega_n \cap \big(\bigcup_k B_k\big)^c && \text{(13.190a)}\\
&= \bigcup_n \Big(\Omega_n \cap \big(\bigcap_k B_k^c\big)\Big) && \text{(13.190b)}\\
&\subset \bigcup_n \big(\Omega_n \cap B_n^c\big) && \text{(13.190c)}\\
&= \bigcup_n (\Omega_n \backslash B_n) && \text{(13.190d)}
\end{aligned}$$

où l'union n'est pas spécialement disjointe. Par conséquent,

$$m(\Omega \backslash B) \leqslant \sum_{n=1}^{\infty} m(\Omega_n \backslash B_n) \leqslant \sum_{n=1}^{\infty} \frac{\epsilon}{2^{n+2}} = \frac{\epsilon}{4}. \tag{13.191}$$

De la même façon nous avons

$$B \backslash A = \big(\bigcup_{n=1}^{\infty} B_n\big) \cap \big(\bigcup_{k=1}^{\infty} F_n\big)^c \subset \bigcup_{n=1}^{\infty} B_n \backslash F_n. \tag{13.192}$$

Nous avons alors le inégalités de mesures

$$\begin{aligned}
m(B \backslash A) &\leqslant \sum_{n=1}^{\infty} m(B_n \backslash F_n) && \text{(13.193a)}\\
&\leqslant \sum_{n=1}^{\infty} m(\Omega_n \backslash F_n) && \text{(13.193b)}\\
&\leqslant \frac{\epsilon}{4}. && \text{(13.193c)}
\end{aligned}$$

C'est vraiment dommage que A ne soit pas en générale un fermé, sinon il répondrait à la question. Nous posons $F_1' = F_1$ et $F_n' = \bigcup_{k=1}^{n} F_k$. En tant qu'unions finies de fermés,

les F'_n sont des fermés (lemme 6.2(2)). De plus la suite (F'_n) est croissante et l'union est A. Par le lemme 13.27(1) nous avons

$$m(A) = m\big(\bigcup_n F'_n\big) = \lim_{n\to\infty} m(F'_n). \tag{13.194}$$

Il existe donc n_ϵ tel que

$$m(A) - m(F'_n) \leqslant \epsilon \tag{13.195}$$

Nous posons $F = F'_{n_\epsilon}$. Vu que $F \subset A$ nous avons aussi $m(A\backslash F) = m(A) - m(F) \leqslant \epsilon$. Et en plus $F \subset A \subset B \subset \Omega$, ce qui donne bien la propriété voulue $F \subset B \subset \Omega$. Il reste à nous assurer de $m(\Omega\backslash F)$. Nous avons d'abord

$$m(B\backslash F) = m\big((B\backslash A) \cup (A\backslash F)\big) = m(B\backslash A) + m(A\backslash F) \leqslant \frac{5\epsilon}{4}. \tag{13.196}$$

Et enfin :

$$m(\Omega\backslash F) = m\big((\Omega\backslash B) \cup (B\backslash F)\big) = m(\Omega\backslash B) + m(B\backslash F) \leqslant \frac{6\epsilon}{4}. \tag{13.197}$$

Et donc à redéfinition près de ϵ c'est d'accord.

Il est donc établi que $\mathcal{F}$ est une tribu. Qui plus est, l'ensemble $\mathcal{F}$ est une tribu incluse aux boréliens et contenant les ouverts. Ergo $\mathcal{F} = \mathcal{B}\text{or}(X)$.

Régularité extérieure Soit B un borélien et $\epsilon > 0$. Alors il existe F_ϵ fermé et Ω_ϵ ouvert tels que $F_\epsilon \subset B \subset \Omega_\epsilon$ et $m(\Omega_\epsilon\backslash F_\epsilon) \leqslant \epsilon$. Vu que $B \subset \Omega_\epsilon$ pour tout ϵ, nous avons aussi

$$m(B) \leqslant \inf_\epsilon m(\Omega_\epsilon). \tag{13.198}$$

Mais comme $\mu(\Omega_\epsilon) \geqslant m(B)$ pour tout ϵ, nous avons en réalité $m(B) = \inf_\epsilon m(\Omega_\epsilon)$.

Soit maintenant un ouvert Ω tel que $B \subset \Omega$. Nous devons prouver l'existence d'un $\epsilon > 0$ tel que $m(\Omega_\epsilon) \leqslant m(\Omega)$. Cela permettra de conclure que l'infimum sur tous les ouverts contenant B est égal à l'infimum sur les ouverts de la forme Ω_ϵ.

Nous posons $m(\Omega) = m(B) + \delta$ et avec $\epsilon \leqslant \delta$ nous avons

$$m(\Omega_\epsilon\backslash B) \leqslant m(\Omega_\epsilon\backslash F_\epsilon) \leqslant \epsilon \tag{13.199}$$

et donc aussi

$$m(\Omega_\epsilon) \leqslant m(B) + \epsilon \leqslant m(B) + \delta = m(\Omega). \tag{13.200}$$

Régularité intérieure Elle se fait de même.

□

Définition 13.78.
Soit X un espace topologique et m une mesure positive sur $\big(X, \mathcal{B}or(X)\big)$.

*(1) m est une **mesure de Borel** si elle est finie sur tout compact.*

*(2) m est **régulière extérieurement** si $\forall B \in \mathcal{B}or(X)$,*

$$m(B) = \inf\{m(\Omega) \text{ tel que } \Omega \text{ est ouvert et } B \subset \Omega\} \tag{13.201}$$

*(3) m est **régulière intérieurement** si $\forall B \in \mathcal{B}or(X)$,*

$$m(B) = \sup\{m(K) \text{ tel que } K \text{ est compact et } K \subset B\} \tag{13.202}$$

*(4) m est une mesure **régulière** si elle est régulière dans les deux sens.*

*(5) m est une **mesure de Radon** si elle est de Borel et régulière.*

Proposition 13.79.
Soit X un espace localement compact et dénombrable à l'infini[15] *Alors toute mesure de Borel sur $\big(X, \mathcal{B}or(X)\big)$ est de Radon.*

Démonstration. Nous avons une suite exhaustive[16] de compacts X_k tels que

$$X = \bigcup_{k\geqslant 1} X_k = \bigcup_{k\geqslant 1} \mathrm{Int}(X_k). \tag{13.203}$$

Régularité intérieure Soit B, un borélien de X ; nous avons $B = \bigcup_{k\geqslant 1}(B \cap X_k)$ et comme cette union est croissante,

$$m(B) = \lim_{k\to\infty} m(B \cap X_k) \tag{13.204}$$

par le lemme 13.27(1). Dans la suite, il va y avoir beaucoup de considérations sur les topologies induites. Nous nommons τ_k la topologie de X_k induite depuis celle de X. Il ne faudra pas confondre les expressions «un compact *de* X_k» et «un compact *dans* X_k». La première parle d'un compact pour la topologie τ_k. La seconde parle d'un compact pour la topologie de X, inclus à X_k.
Si $a < m(B)$ alors il existe $k \geqslant 1$ tel que $a < m(B \cap X_k)$, c'est à dire

$$a < m(B \cap X_k) \leqslant m(B). \tag{13.205}$$

Mais (X_k, m) est un espace mesuré borné parce que m est de Borel et X_k est compact. Par conséquent la (restriction de la) mesure m est régulière sur l'espace mesuré $\big(X_k, \mathcal{B}\mathrm{or}(X_k)\big)$ par le théorème 13.77. De plus l'ensemble $B \cap X_k$ est un borélien de (X_k, τ_k) parce que

$$B \cap X_k \in \mathcal{B}\mathrm{or}(X)_{X_k} = \mathcal{B}\mathrm{or}(X_k) \tag{13.206}$$

où nous avons utilisé la propriété de compatibilité entre topologie induite et tribu des borélien du théorème 13.73. Il existe donc un fermé F_ϵ de (X_k, τ_k) tel que

$$\begin{cases} F_\epsilon \subset B \cap X_k & \text{(13.207a)} \\ m(B \cap X_k) \leqslant m(F_\epsilon) + \epsilon. & \text{(13.207b)} \end{cases}$$

En mettant bout à bout les inégalités nous avons trouvé

$$a < m(B \cap X_k) \leqslant m(F_\epsilon) + \epsilon < m(F_\epsilon), \tag{13.208}$$

et donc en particulier $a < m(F_\epsilon)$. L'ensemble F_ϵ est en plus un compact de (X, τ_X). En effet X_k étant fermé de (X, τ_X), le lemme 6.83 nous dit que F_ϵ est un fermé de (X, τ_X). Mais X_k étant compact, F_ϵ est un fermé inclus à un compact, il est donc compact (lemme 6.75). Pour tout $a < m(B)$ nous avons trouvé un compact F_ϵ inclus à B dont la mesure est plus grande que a. Cela prouve la régularité intérieure de la mesure m.

Régularité extérieure Soit un borélien B de X. Si $m(B) = \infty$ alors tous les ouverts contenant B ont mesure infinie et $m(B)$ en est évidemment le supremum. Nous supposons donc que $m(B) < \infty$.
Nous notons τ_k la topologie induite de X sur $\mathrm{Int}(X_k)$. Nous posons $B_k = B \cap \mathrm{Int}(X_k)$. L'espace $\big(\mathrm{Int}(X_k), m\big)$ est un espace mesuré borné et $B_k \in \mathcal{B}\mathrm{or}\Big(\mathrm{Int}(X_k)\Big)$. Il existe donc un ouvert Ω_k de $\big(\mathrm{Int}(X_k), \tau_k\big)$ tel que $B_k \subset \Omega_k$ et

$$m(\Omega_k \backslash B_k) \leqslant \frac{\epsilon}{2^k}. \tag{13.209}$$

De plus $\mathrm{Int}(X_k)$ est un ouvert de (X, τ_X), donc en réalité Ω_k est un ouvert de X. Nous posons

$$\Omega = \bigcup_{k=1}^{\infty} \Omega_k \tag{13.210}$$

15. Définitions 6.13 et 6.15.
16. Définition 6.156.

qui est encore un ouvert de (X, τ_X).
Il est temps de voir que Ω vérifie $m(\Omega\backslash B) \leqslant \epsilon$. Pour cela,

$$\Omega\backslash B = \big(\bigcup_k \Omega_k\big) \cap \big(\bigcup_l B_l\big)^c \tag{13.211a}$$
$$= \big(\bigcup_k \Omega_k\big) \cap \big(\bigcap_l B_l^c\big) \tag{13.211b}$$
$$\subset \bigcup_k (\Omega_k \cap B_k^c) \tag{13.211c}$$
$$= \bigcup_k (\Omega_k\backslash B_k), \tag{13.211d}$$

ce qui donne au niveau des mesures :

$$m(\Omega\backslash B) \leqslant \sum_{k=1}^{\infty} m(\Omega_k\backslash B_k) \leqslant \sum_{k=1}^{\infty} \frac{\epsilon}{2^k} = \epsilon. \tag{13.212}$$

□

Remarque 13.80.
Exprimé sur $\mathbb{R}^N$, la proposition 13.79 s'exprime en disant que toute mesure de Borel sur $\mathbb{R}^N$ est régulière. Typiquement, l'espace X dont il est question est un ouvert de $\mathbb{R}^N$.

13.3.5 Théorème de récurrence

Soit X un espace mesurable, μ une mesure finie sur X et $\phi\colon X \to X$ une application mesurable préservant la mesure, c'est à dire que pour tout ensemble mesurable $A \subset X$,

$$\mu\big(\phi^{-1}(A)\big) = \mu(A). \tag{13.213}$$

Si $A \subset X$ est un ensemble mesurable, un point $x \in A$ est dit **récurrent** par rapport à A si et seulement si pour tout $p \in \mathbb{N}$, il existe $k \geqslant p$ tel que $\phi^k(x) \in A$.

Théorème 13.81 (Théorème de récurrence de Poincaré.)**.**
Si A est mesurable dans X, alors presque tous les points de A sont récurrents par rapport à A.

Démonstration. Soit $p \in \mathbb{N}$ et l'ensemble

$$U_p = \bigcup_{k=p}^{\infty} \phi^{-k}(A) \tag{13.214}$$

des points qui repasseront encore dans A après p itérations de ϕ. C'est un ensemble mesurable en tant que union d'ensembles mesurables (pour rappel, les tribus sont stables par union dénombrable, comme demandé à la définition 13.1), et nous avons donc

$$\mu(U_p) \leqslant \mu(X) < \infty. \tag{13.215}$$

De plus $U_p = \phi^{-p}(U_0)$, donc $\mu(U_p) = \mu(U_0)$. Vu que $U_p \subset U_p$, nous avons

$$\mu(U_0\backslash U_p) = 0. \tag{13.216}$$

Étant donné que $A \subset U_0$ nous avons a fortiori que

$$\{x \in A \text{ tel que } x \notin U_p\} \subset U_0\backslash U_p, \tag{13.217}$$

et donc

$$\mu\{x \in A \text{ tel que } x \notin U_p\} = 0. \tag{13.218}$$

Cela signifie exactement que l'ensemble des points x de A tels que aucun des $\phi^k(x)$ avec $k \geqslant p$ n'est dans A est de mesure nulle. □

13.4 Mesurabilité des fonctions à valeurs réelles

Nous allons parler de la mesurabilité de fonctions

$$f\colon (S,\mathcal{F}) \to \big(\bar{\mathbb{R}}, \mathcal{B}\text{or}(\bar{\mathbb{R}})\big) \tag{13.219}$$

où $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$.

13.4.1 Quelques mots à propos de $\overline{\mathbb{R}}$

13.82.
Nous convenons que $0 \times \pm\infty = 0$ parce que nous voulons qu'une droite (qui est un rectangle dont une mesure est 0 et l'autre ∞) soit de mesure nulle dans $\mathbb{R}^2$.

Les produits et sommes $\pm\infty \pm \pm\infty$ et $\pm\infty \times \pm\infty$ sont ceux que l'on croit. Sauf bien entendu $+\infty - \infty$ et $1/0$ qui ne sont toujours pas définis.

Définition 13.83 (Topologie sur $\bar{\mathbb{R}}$)**.**
La topologie sur $\bar{\mathbb{R}}$ est celle sur $\mathbb{R}$ à laquelle nous ajoutons les voisinages de $\pm\infty$ de la façon suivante. Une partie V de $\bar{\mathbb{R}}$ est un voisinage de $+\infty$ s'il existe $m > 0$ tel que $]m, +\infty] \subset V$.

Lemme 13.84.
L'ensemble B est un borélien de $\bar{\mathbb{R}}$ si et seulement s'il existe un borélien B_0 de $\mathbb{R}$ tel que B soit B_0 ou $B_0 \cup \{-\infty\}$ ou $B_0 \cup \{-\infty\}$ ou $B_0 \cup \{+\infty, -\infty\}$.

Démonstration. Vu que la topologie usuelle sur $\mathbb{R}$ est la topologie induite de celle sur $\bar{\mathbb{R}}$, la tribu induite l'est aussi par le théorème 13.72. Donc si B est un borélien de $\bar{\mathbb{R}}$, l'ensemble $B \cap \mathbb{R}$ est un borélien de $\mathbb{R}$. □

Lemme 13.85 ([166])**.**
Si $\mathcal{S}_0$ est l'ensemble des intervalles du type

$$]\alpha,\beta[, \quad [-\infty,\beta[, \quad]\alpha,+\infty] \tag{13.220}$$

avec $-\infty < \alpha < \beta < +\infty$ alors $\sigma(\mathcal{S}_0) = \mathcal{B}or(\bar{\mathbb{R}})$.

Démonstration. Les intervalles $]\alpha,\beta[$ engendrent la topologie de $\mathbb{R}$ [17], donc $\mathcal{B}\text{or}(\mathbb{R}) \subset \sigma(\mathcal{S}_0)$. De plus le lemme 13.3 nous autorise à dire que

$$\bigcap_{n\geqslant 1} [n, +\infty] = \{+\infty\} \in \sigma(\mathcal{S}_0). \tag{13.221}$$

Par conséquent tous les ensembles énumérés dans le lemme 13.84 font partie de $\sigma(\mathcal{S}_0)$. Cela implique que $\mathcal{B}\text{or}(\bar{\mathbb{R}}) \subset \sigma(\mathcal{S}_0)$.

Pour l'inclusion inverse, $\sigma(\mathcal{S}_0)$ est engendré par des parties qui font parie de $\mathcal{B}\text{or}(\bar{\mathbb{R}})$, donc $\sigma(\mathcal{S}_0) \subset \mathcal{B}\text{or}(\bar{\mathbb{R}})$. □

Pour la suite nous utilisons la notation (pratique en probabilité)

$$\{f < a\} = \{x \in S \text{ tel que } f(x) < a\}. \tag{13.222}$$

13.4.2 Limite supérieure et inférieure

Définition 13.86.
*Soit (a_n) une suite dans $\bar{\mathbb{R}}$. Nous définissons la **limite supérieure** et la **limite inférieure** par*

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} \big(\sup_{k\geqslant n} a_k\big) \tag{13.223}$$

et

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} \big(\inf_{k\geqslant n} a_k\big). \tag{13.224}$$

17. Parce toutes les boules sont des intervalles de ce type et que les boules forment une base de topologie, proposition 6.60.

13.87.
En ce qui concerne les suites d'ensembles, utiles en théorie des probabilités, nous définissons de même. Si les A_n sont des parties de Ω, nous définissons la **limite supérieure** et la **limite inférieure** de la suite A_n par

$$\limsup_{n\to\infty} A_n = \bigcap_{n\geqslant 1}\bigcup_{k\geqslant n} A_k \tag{13.225}$$

et

$$\liminf_{n\to\infty} A_n = \bigcup_{n\geqslant 1}\bigcap_{k\geqslant n} A_k \tag{13.226}$$

Nous avons

$$\limsup A_n = \{\omega\in\Omega \text{ tel que } \omega\in A_n \text{pour une infinité de } n\}. \tag{13.227}$$

Lemme 13.88.
Nous avons les formules pratiques suivantes :

$$\limsup a_n = \inf_{n\geqslant 1}\big(\sup_{k\geqslant n} a_k\big) \tag{13.228a}$$

$$\liminf a_n = \sup_{n\geqslant 1}\big(\inf_{k\geqslant n} a_k\big). \tag{13.228b}$$

Démonstration. La suite $n\mapsto \sup_{k\geqslant n} a_k$ est une suite décroissante, donc la limite est l'infimum. Même argument pour l'autre. □

Lemme 13.89.
La suite (a_n) dans $\mathbb{R}$ converge si et seulement si

$$\limsup a_n = \liminf a_n. \tag{13.229}$$

Dans ce cas, $\lim a_n = \limsup a_n = \liminf a_n$.

Démonstration. Nous commençons par supposer que $\limsup a_n = \liminf a_n = l$, et nous prouvons que $\lim a_n$ existe et vaut l. Soit $\epsilon > 0$. Il existe N tel que si $n\geqslant N$ nous avons

$$\big|\sup_{k\geqslant n} a_k - l\big| < \epsilon \tag{13.230}$$

et

$$\big|\inf_{k\geqslant n} a_k - l\big| < \epsilon. \tag{13.231}$$

Pour tout $k\geqslant N$ nous avons alors $a_k \leqslant l+\epsilon$ et $a_k\geqslant l-\epsilon$. Cela donne $a_n\in B(l,\epsilon)$, c'est à dire $a_k\to l$ par la proposition 6.90.

Dans l'autre sens, nous supposons que $\lim_n a_n = l$ et nous prouvons que les limites supérieures et inférieures sont toutes deux égales à l. Soit $\epsilon>0$ et N_ϵ tel que $|a_n-l|<\epsilon$ pour tout $n\geqslant N_\epsilon$. Si $n\geqslant N_\epsilon$ nous avons

$$\big|\sup_{k\geqslant n} a_k - l\big| \leqslant \epsilon \tag{13.232}$$

et donc la limite de $\sup_{k\geqslant n} a_k$ lorsque $n\to\infty$ est bien l. □

13.4.3 Fonctions à valeurs réelles sur un espace mesurable

Théorème 13.90.
Soit un espace mesurable $(S,\mathcal{F})$ et une fonction $f\colon S\to\bar{\mathbb{R}}$. Les faits suivants sont équivalents.

(1) La fonction f est mesurable.

(2) L'ensemble $\{f<a\}$ est dans $\mathcal{F}$ pour tout $a\in\mathbb{R}$

(3) L'ensemble $\{f\leqslant a\}$ est dans $\mathcal{F}$ pour tout $a\in\mathbb{R}$

Démonstration. Plusieurs implications à prouver.

(1)⇒(2) Vu que f est mesurable et que $[-\infty,a[\in\mathcal{B}\text{or}(\bar{\mathbb{R}})$, nous avons $f^{-1}\big([-\infty,a[\big)\in\mathcal{F}$.

(2)⇒(1) Nous posons $\mathcal{A}=\{[-\infty,a[\text{ tel que }a\in\mathbb{R}\}$.

Nous avons $\mathcal{A}\subset\mathcal{S}_0$ (le $\mathcal{S}_0$ du lemme 13.85). Et de plus,

$$]\alpha,\beta[=[-\infty,\beta[\backslash[-\infty,\alpha]=[-\infty,\beta[\backslash\bigcap_{n\geqslant 1}[-\infty,\alpha+\frac{1}{n}[. \tag{13.233}$$

Donc $]\alpha,\beta[\in\sigma(\mathcal{A})$.

Et aussi :

$$]\alpha,+\infty]=\bar{\mathbb{R}}\backslash[-\infty,\alpha+\frac{1}{n}[, \tag{13.234}$$

ce qui donne $]\alpha,+\infty]\in\sigma(\mathcal{A})$.

Au final, $\mathcal{S}_0\subset\sigma(\mathcal{A})$ et donc $\sigma(\mathcal{S}_0)\subset\sigma(\mathcal{A})$. Le lemme 13.85 nous dit que $\sigma(\mathcal{S}_0)=\mathcal{B}\text{or}(\bar{\mathbb{R}})$. Nous avons donc bien $\sigma(\mathcal{S}_0)=\sigma(\mathcal{A})=\mathcal{B}\text{or}(\bar{\mathbb{R}})$.

par ailleurs, nous savons que $f^{-1}(\mathcal{A})\subset\mathcal{F}$ parce que les éléments de $\mathcal{A}$ sont de la forme $\{f<a\}$. Cela donne $\sigma\big(f^{-1}(\mathcal{A})\big)=\mathcal{F}$. Mais $\sigma\big(f^{-1}(\mathcal{A})\big)$ peut aussi s'exprimer par le lemme de transfert 13.70 : $\sigma\big(f^{-1}(\mathcal{A})\big)=f^{-1}\big(\sigma(\mathcal{A})\big)$. En combinant les deux,

$$f^{-1}\big(\sigma(\mathcal{A})\big)=\mathcal{F}, \tag{13.235}$$

et en remplaçant $\sigma(\mathcal{A})$ par $\mathcal{B}\text{or}(\bar{\mathbb{R}})$ nous avons ce que nous voulions :

$$f^{-1}\big(\mathcal{B}\text{or}(\bar{\mathbb{R}})\big)\in\mathcal{F}, \tag{13.236}$$

ce qui signifie que f est mesurable.

(3)⇒(2) Nous avons

$$\{f<a\}=\bigcup_{n\geqslant 1}\{f\leqslant a-\frac{1}{n}\}. \tag{13.237}$$

donc cela est une union dénombrable d'éléments de $\mathcal{F}$. Donc $\{f<a\}$ est dans $\mathcal{F}$.

(1)⇒(3) Nous avons

$$\{f\leqslant a\}=\{f<a\}\cup f^{-1}\big([-\infty,a]\big). \tag{13.238}$$

Le premier ensemble est dans $\mathcal{F}$ par (2). Ensuite $[-\infty,a]$ est un fermé de $\bar{\mathbb{R}}$ et donc un borélien de $\bar{\mathbb{R}}$. Son image réciproque est donc un élément de $\mathcal{F}$ parce que f est mesurable. Au final nous avons bien $\{f\leqslant a\}\in\mathcal{F}$.

□

Lemme 13.91 ([167])**.**
Une fonction $f\colon X\to\mathbb{R}$ est mesurable si et seulement si $f^{-1}(I)$ est mesurable pour tout I de la forme $]a,\infty[$.

Démonstration. Nous devons prouver que $f^{-1}(A)$ est mesurable dans X pour tout borélien A de $\mathbb{R}$. Nous posons

$$S=\{A\subset\mathbb{R}\text{ tel que }f^{-1}(A)\text{ est mesurable dans }X\} \tag{13.239}$$

et nous prouvons que cela est une tribu. D'abord $f^{-1}(\mathbb{R})=X$, et X est mesurable, donc $\mathbb{R}\in S$. Ensuite si $A\in S$ alors $f^{-1}(A^c)=f^{-1}(A)^c$. En tant que complémentaire d'un mesurable de X, l'ensemble $f^{-1}(A)^c$ est mesurable dans X. Et enfin si $A_n\in S$ alors $f^{-1}(\bigcup_n A_n)=\bigcup_n f^{-1}(A_n)$ qui est encore mesurable dans X en tant qu'union de mesurables.

Donc S est une tribu qui contient tous les ensembles de la forme $]a,\infty]$. Le lemme 13.10 conclu que S contient tous les boréliens de $\mathbb{R}$. □

Lemme 13.92 ([167])**.**
Soit $f_n\colon X\to\mathbb{R}$ une suite de fonctions mesurables[18]*. Alors $\sup_n f_n$ est mesurable.*

18. Ici X est un espace mesuré et $\mathbb{R}$ est muni des boréliens.

Démonstration. Nous avons

$$(\sup f_n)^{-1}\big(]a,\infty]\big) = \{x \in X \text{ tel que } (\sup f_n)(x) > a\} \tag{13.240a}$$
$$= \bigcup_n \{x \in X \text{ tel que } f_n(x) > a\} \tag{13.240b}$$
$$= \bigcup_n f_n^{-1}\big(]a,\infty]\big). \tag{13.240c}$$

Étant donné que f_n est mesurable et que $]a,\infty]$ est mesurable, chacun des $f_n^{-1}\big(]a,\infty]\big)$ est mesurable dans X. Nous sommes en présence d'une union dénombrable de mesurables, donc $(\sup f_n)^{-1}\big(]a,\infty]\big)$ est mesurable.

Le lemme 13.91 conclu que $\sup f_n$ est mesurable. □

Proposition 13.93.
Si f_n est une suite de fonctions mesurables et positives, alors la fonction $\sum_n f_n$ est mesurable.

Démonstration. Nous considérons les fonctions $s_k(x) = \sum_{n=0}^{k} f_n(x)$ qui vaut éventuellement ∞ en certains points. Nous avons

$$\sum_n f_n(x) = \sup_k s_k(x), \tag{13.241}$$

donc le lemme 13.92 nous donne la mesurabilité de la somme de f_n. □

Définition 13.94.
Soit $(S,\mathcal{F})$ un espace mesurable. Une ***partition mesurable dénombrable*** *de e S est une suite $(S_n)_{n\geqslant 1}$ de parties de S telles que*

(1) $S_n \in \mathcal{F}$ pour tout n,

(2) $S_N \cap S_k = \varnothing$ si $n \neq k$,

(3) $S = \bigcup_{n\geqslant 1} S_n$.

Lemme 13.95 (Lemme de recollement)**.**
Soit (S_n) une partition mesurable dénombrable de l'espace mesurable $(S,\mathcal{F})$. Soit $(S',\mathcal{F}')$ un autre espace mesurable et des fonctions mesurables

$$f_n\colon (S_n,\mathcal{F}_{S_n}) \to (S',\mathcal{F}') \tag{13.242}$$

où $\mathcal{F}_{S_n}$ est la tribu induite[19]*. Alors la fonction*

$$\begin{aligned} f\colon (S,\mathcal{F}) &\to (S',\mathcal{F}') \\ x &\mapsto f_n(x) \text{ si } x \in S_n \end{aligned} \tag{13.243}$$

est mesurable.

Démonstration. Soit $A' \in \mathcal{F}'$; nous devons prouver que $f^{-1}(A') \in \mathcal{F}$. Nous savons que

$$f^{-1}(A') = \bigcup_{n\geqslant 1} f_n^{-1}(A'), \tag{13.244}$$

qui est une union dénombrable d'éléments $f_n^{-1}(A') \in \mathcal{F}_{S_n}$.

Vu que $S_n \in \mathcal{F}$ nous avons $\mathcal{F}_{S_n} \subset \mathcal{F}$ parce qu'un élément de $\mathcal{F}_{S_n}$ est de la forme $S_n \cap B$ avec $B \in \mathcal{F}$. Du coup pour chaque n nous avons

$$f_n^{-1}(A') \in \mathcal{F}_{S_n} \subset \mathcal{F}. \tag{13.245}$$

Au final l'égalité (13.244) écrit $f^{-1}(A')$ comme une union d'éléments de $\mathcal{F}$ et est donc un élément de $\mathcal{F}$. □

19. Définition 13.7.

Proposition 13.96.
Soit $(S, \mathcal{F})$ un espace mesurable et des applications mesurables $f, g\colon S \to \bar{\mathbb{R}}$. Alors les fonctions suivantes sont mesurables :

(1) λf pour tout $\lambda \in \mathbb{R}$

(2) $f + g$ si elle existe.

(3) $1/f$ si elle existe.

(4) fg.

Démonstration. Commençons par clarifier « si elle existe». La fonction $f+g$ n'existe pas au point $x \in S$ si $f(x) = +\infty$ et $g(x) = -\infty$. La fonction $1/f$ n'existe pas au point $x \in S$ si $f(x) = 0$. Voir le point 13.82.

La partie où $f + g$ existe est mesurable La partie de S sur laquelle $f + g$ existe est

$$\{x \in S \text{ tel que } \big(f(x), g(x)\big) \neq (+\infty, -\infty), \big(f(x), g(x)\big) \neq (-\infty, +\infty)\}. \tag{13.246}$$

Nous avons

$$\{(f, g) = (+\infty, -\infty)\} = \{f = \infty\} \cap \{g = -\infty\} \tag{13.247}$$

qui est un ensemble mesurable parce que, par exemple,

$$\{+\infty\} = \bigcap_{n \geqslant 1} [n, +\infty]. \tag{13.248}$$

La cas $(-\infty, +\infty)$ est identique, et au final la partie de S sur laquelle $f + g$ n'existe pas est mesurable. Par complémentarité la partie sur laquelle $f+g$ existe est également mesurable [20].

Idem pour la partie sur laquelle $1/f$ existe Idem.

Mesurabilité de λf Si $\lambda = 0$, nous avons une fonction constante dont la mesurabilité est évidente [21]. Nous supposons $\lambda > 0$. Alors

$$\{\lambda f < a\} = \{f < a/\lambda\} \in \mathcal{F}. \tag{13.249}$$

.Pour $\lambda < 0$ nous avons de la même manière

$$\{\lambda f < a\} = \{f > a/\lambda\} \in \mathcal{F}. \tag{13.250}$$

Ce dernier point est suffisant pour que λf soit mesurable par la théorème 13.90(3) et par complémentarité.

Mesurabilité de $f + g$ Soit $a \in \mathbb{R}$; le théorème 13.90 nous demande d'avoir envie de prouver que $\{f + g < a\} \in \mathcal{F}$. Nous avons

$$f(x) + g(x) < a \tag{13.251}$$

si et seulement si

$$f(x) < a - g(x) \tag{13.252}$$

si et seulement si

$$\exists q \in \mathbb{Q} \text{ tel que } f(x) < q < a - g(x). \tag{13.253}$$

Donc

$$\{f + g < a\} = \bigcup_{q \in \mathbb{Q}} \Big(\{f < q\} \cap \{g < a - r\} \Big), \tag{13.254}$$

qui est une union dénombrable d'éléments de $\mathcal{F}$. Donc $\{f+g < a\} \in \mathcal{F}$ et $f+g$ est mesurable.
Note qu'en toute rigueur il faudrait «$\cap$là où $f + g$ est définie» un peu partout, mais cela ne change rien parce que l'intersection de deux parties mesurables est mesurable.

20. Parfois on a envie de dire que l'affirmation «A est mesurable» ne passe pas le test de Popper.
21. Prenez quand même le temps d'y penser.

Mesurabilité de $1/f$ Soit $a \in \mathbb{R}$. Si $a > 0$ alors

$$\{1/f < a\} = \{f < 0\} \cup \{f > \frac{1}{a}\} \in \mathcal{F}. \tag{13.255}$$

et si $a < 0$ alors

$$\{1/f < a\} = \{f < 0\} \cap \{f > \frac{1}{a}\} \in \mathcal{F}. \tag{13.256}$$

Mesurabilité de fg Nous allons la prouver en plusieurs fois.

Si f est mesurable alors f^2 est mesurable Si $a \leqslant 0$ alors $\{f^2 < a\} = \varnothing$. Si $a > 0$ nous avons

$$\{f^2 < a\} = \{-\sqrt{a} < f < \sqrt{a}\} \in \mathcal{F}. \tag{13.257}$$

$f\mathbb{1}_A$ est mesurable Soit $A \in \mathcal{F}$, et prouvons que $f\mathbb{1}_A$ est mesurable. Par définition,

$$(f\mathbb{1}_A)(x) = \begin{cases} f(x) & \text{si } x \in A \\ 0 & \text{si } x \notin A. \end{cases} \tag{13.258}$$

Nous posons

$$\begin{aligned} f_1 \colon A^c &\to \bar{\mathbb{R}} \\ x &\mapsto 0 \end{aligned} \tag{13.259}$$

et

$$\begin{aligned} f_2 \colon A &\to \bar{\mathbb{R}} \\ x &\mapsto f(x). \end{aligned} \tag{13.260}$$

Alors nous avons

$$(\mathbb{1}_A f)(x) = \begin{cases} f_1(x) & \text{si } x \in A^c \\ f_2(x) & \text{si } x \in A. \end{cases} \tag{13.261}$$

Les ensembles A et A^c forment une partition mesurable dénombrable de S. La fonction f_1 est mesurable ; pour prouver que f_2 est mesurable, nous l'écrivons $f_2 = f \circ j_A$ où $j_A \colon A \to S$ est l'injection canonique. L'application

$$j_A \colon (A, \mathcal{F}_A) \to (S, \mathcal{F}) \tag{13.262}$$

est mesurables parce que si $B \in \mathcal{F}$ alors $j_A^{-1}(B) = A \cap B \in \mathcal{F}_A$. D'autre part l'application

$$f \colon (S, \mathcal{F}) \to \big(\bar{\mathbb{R}}, \mathcal{B}\text{or}(\bar{\mathbb{R}})\big) \tag{13.263}$$

est mesurable par hypothèse. La composée $f_2 = f \circ j_A$ est alors mesurable par la proposition 13.66. Le lemme de recollement 13.95 nous donne alors la mesurabilité de $f\mathbb{1}_A$.

Le produit fg est mesurable Nous posons

$$F = \{x \in S \text{ tel que } |f(x)| < +\infty, |g(x)| < \infty\}. \tag{13.264}$$

En tant qu'intersection de deux ensembles mesurables, F est mesurable. Par la partie précédente, les applications $f_1 = g\mathbb{1}_F$ et $g_1 = g\mathbb{1}_F$ sont mesurables. L'application $f_1 + g_1 \colon S \to \mathbb{R}$ est encore mesurable. Par conséquent l'application

$$f_1 g_1 = \frac{1}{2}\big((f_1 + g_1)^2 - f_1^2 - g_1^2\big) \tag{13.265}$$

est mesurable.

Voyons maintenant ce qui se passe en dehors de F. Nous allons utiliser le lemme de recollement sur la fonction

$$(fg)(x) = \begin{cases} (f_1 f_2)(x) & \text{si } x \in F \\ -\infty & \text{si } x \in \mathcal{U} \\ 0 & \text{si } x \in \mathcal{V} \\ +\infty & \text{si } x \in \mathcal{W} \end{cases} \tag{13.266}$$

où $F,\mathcal{U},\mathcal{V},\mathcal{W}$ forment une partition mesurable dénombrable [22] de S. Pour le sport nous montrons que $\mathcal{U}$ est mesurable :

$$\mathcal{U} = \big(\{f = -\infty\} \cap \{g > 0\}\big) \tag{13.267a}$$
$$\cup \big(\{f = +\infty\} \cap \{g < 0\}\big) \tag{13.267b}$$
$$\cup \big(\{g = -\infty\} \cap \{f > 0\}\big) \tag{13.267c}$$
$$\cup \big(\{g = +\infty\} \cap \{f < 0\}\big). \tag{13.267d}$$

□

Proposition 13.97.
Si $f_n\colon S \to \bar{\mathbb{R}}$ est une suite de fonctions mesurables, alors les fonctions $\inf_n f_n$ et $\sup_n f_n$ sont mesurables.

Démonstration. Nous avons les découpages

$$\{\inf_n f_n < a\} = \bigcup_n \{f_n < a\} \in \mathcal{F} \tag{13.268}$$

et

$$\{\sup_n f_n \leqslant a\} = \bigcap_n \{f_n \leqslant a\} \in \mathcal{F}. \tag{13.269}$$

Le théorème 13.90 permet de conclure. □

Note que pour (13.269) nous ne pouvions pas utiliser les inégalités strictes parce que $\{\sup_n f_n < a\}$ n'est pas spécialement égal à $\bigcap_n \{f_n < a\}$.

13.98.
La proposition 13.97 nous permet de définir les parties positives et négatives de f par $f^+ = \sup(f, 0)$ et $f^- = \sup(-f, 0)$. Ce sont des applications mesurables. Nous avons les décompositions

$$f = f^+ - f^- \tag{13.270a}$$
$$|f| = f^+ + f^-. \tag{13.270b}$$

Corollaire 13.99.
Si $f\colon S \to \bar{\mathbb{R}}$ est mesurable alors les applications f^+, f^- et $|f|$ sont mesurables en tant qu'applications $S \to \bar{\mathbb{R}}^+$.

Démonstration. Nous faisons la preuve pour f^+. Nous savons que $f^+\colon S \to \bar{\mathbb{R}}$ est mesurable par la proposition 13.97. Nous considérons l'injection canonique $f\colon \bar{\mathbb{R}}^+ \to \bar{\mathbb{R}}$ et

$$\begin{aligned} f_1^+\colon S &\to \bar{\mathbb{R}}^+ \\ x &\mapsto f^+(x). \end{aligned} \tag{13.271}$$

Alors $f_1^+ = j \circ f^+$ est mesurable. Et c'est bien cela que nous voulions.

□

Note : f^+ et f_1^+ sont exactement les mêmes fonctions. Elles ne diffèrent que par la tribu que nous considérons sur l'espace d'arrivée. Nous allons à partir de maintenant les noter toutes deux f^+.

Remarque 13.100.
L'application $|f|$ peut être mesurable sans que f le soit. Soit en effet une partie $A \notin \mathcal{F}$, et posons

$$f(x) = \begin{cases} 1 & \text{si } x \in A \\ -1 & \text{si } x \in A^c. \end{cases} \tag{13.272}$$

Alors $f^{-1}(\{1\}) = A$ n'est pas mesurable alors que $|f|(x) = 1$ pour tout x.

22. Définition 13.94.

Il est temps d'aller relire les définitions 13.86.

Proposition 13.101.
Si les fonctions $f_n\colon S \to \bar{\mathbb{R}}$ sont mesurables alors les fonctions $\limsup f_n$ et $\liminf f_n$ sont mesurables.

Démonstration. Par le lemme 13.88 nous écrivons $\limsup_n f_n(x) = \inf_{n\geqslant 1}\sup_{k\geqslant n} f_k(x)$. Pour chaque k nous considérons la fonction $g_k = \sup_{n\geqslant k} f_n$. Par la proposition 13.97, les fonctions g_k sont mesurables. En utilisant encore la même proposition, $\inf_{n\geqslant 1} g_k$ est encore mesurable. □

Proposition 13.102 ([168])**.**
Si $f_n\colon S \to \bar{\mathbb{R}}$ est une suite de fonctions mesurables telle dont la limite ponctuelle existe, alors la limite est mesurable.

Démonstration. Si la limite existe, elle est égale à la limite supérieure par le lemme 13.89. Or la limite supérieure est mesurable par la proposition 13.101. □

13.4.4 Fonction étagée

Définition 13.103 ([169])**.**
Soit $(S,\mathcal{F})$ un espace mesurable et une fonction $f\colon S \to \big(\bar{\mathbb{R}}, \mathcal{B}or(\bar{\mathbb{R}})\big)$. Il serait dommage de confondre les trois concepts suivants.

- *Une **fonction simple** est une fonction dont l'image est constituée d'un nombre fini de valeurs.*
- *Une **fonction étagée** est une fonction simple qui est elle-même une fonction mesurable.*
- *Une **fonction en escalier** est une fonction étagée dont les valeurs sont constantes sur des intervalles : ce sont donc des fonctions constantes par morceaux.*

Dans les trois cas, la fonction f peut être écrite comme somme de fonctions caractéristiques :

$$f(x) = \sum_{j=1}^{p} \alpha_j \mathbb{1}_{A_j}(x) \tag{13.273}$$

où $A_j = f^{-1}(\alpha_j)$. Ce qui change est la nature des A_j.

- Si f est simple, les A_j sont quelconques.
- Si f est étagée, les A_i peuvent être choisis mesurables parce que $\{\alpha_i\}$ est un borélien, ce qui fait que $A_i = f^{-1}(\alpha_i)$ un choix mesurable.
- Si f est en escalier, les A_i sont des intervalles.

La **forme canonique** d'une fonction simple f est la suivante. Soit $\{\alpha_i\}_{i=1,\dots,l}$ les valeurs distinctes prises par f et $A_i = f^{-1}(\alpha_i)$. La forme canonique de f est alors

$$f = \sum_{i=1}^{l} \alpha_i \mathbb{1}_{A_i}. \tag{13.274}$$

Notons que nous avons $S = \bigcup_i A_i$, et que cette union est disjointe dans le cas d'une représentation canonique.

Problèmes et choses à faire

Le lemme 13.104 et le théorème 13.107 disent la même chose alors que la preuve du théorème 13.107 est beaucoup plus compliquée.La démonstration du lemme serait fausse ?

M'est avis que ce que le théorème donne en plus est la convergence uniforme en cas de fonction bornée. La suite (13.275) ne va pas converger uniformément.

Lemme 13.104 (Limite croissante de fonctions étagées[1])**.**
Soit $f\colon (S,\mathcal{F}) \to \bar{\mathbb{R}}$ une fonction positive mesurable. Il existe une suite $f_n\colon S \to \mathbb{R}$ de fonctions étagées positives telles que $f_n \to f$ ponctuellement et $f_n \leqslant f$.

Démonstration. Nous considérons (q_n) une suite parcourant tous les rationnels positifs[23] avec $q_0 = 0$ pour être sûr. Pour $n \in \mathbb{N}$ nous définissons la fonction

$$f_n(x) = \max\{q_i \text{ tel que } i \leqslant n,\ q_i \leqslant f(x)\}. \tag{13.275}$$

L'ensemble sur lequel le maximum est pris n'est pas vide parce que $q_0 = 0$. La fonction f_n est simple parce qu'elle ne prend que n valeurs différentes. Nous avons aussi par construction que $f_n(x) \leqslant f(x)$. Nous avons aussi pour tout $x \in S$ que $f_n(x) \to f(x)$ parce que $\mathbb{Q}$ est dense dans $\mathbb{R}$.

En ce qui concerne le fait que f_n est mesurable, nous notons $\{r_0, \dots, r_n\}$ l'ensemble des $\{q_0, \dots, q_n\}$ classés dans l'ordre croissant. Nous posons en plus $r_{n+1} = +\infty$. Nous avons alors

$$f_n^{-1}(r_k) = \{x \in S \text{ tel que } f(x) \geqslant r_k, f(x) < r_{k+1}\} = \{f \geqslant r_k\} \cap \{f < r_{k+1}\} \tag{13.276}$$

qui est un ensemble mesurable comme intersection deux ensembles mesurables par le théorème 13.90. □

Remarque 13.105.
Pour avoir $f_n < |f|$ nous pouvons poser

$$f_n(x) = \begin{cases} \max\{q_i \text{ tel que } i \leqslant n,\ q_i \leqslant f(x)\} & \text{si } f(x) \geqslant 0 \\ \min\{q_i \text{ tel que } i \leqslant n,\ q_i \geqslant f(x)\} & \text{si } f(x) < 0. \end{cases} \tag{13.277}$$

Proposition 13.106 ([170])**.**
Approximation de fonctions mesurables par des fonctions étagées.

(1) *Une fonction mesurable et positive est limite (simple) d'une suite croissante de fonctions étagées, mesurables et positives.*
(2) *Si $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ est mesurable, alors elle est limite (simple) de fonctions étagées f_n telles que $|f_n| \leqslant |f|$.*

Théorème 13.107 (Théorème fondamental d'approximation[166])**.**
Pas à pas.

(1) *Soit une fonction mesurable $f\colon S \to [0, +\infty]$. Alors il existe une suite croissante de fonctions $\varphi_n\colon S \to [0, +\infty[$ étagées positives dont la limite ponctuelle est f.*
(2) *Si de plus f est bornée, la convergence est uniforme.*
(3) *Idem pour f à valeurs dans $\bar{\mathbb{R}}$ ou $\mathbb{C}$.*

Démonstration. Nous découpons l'intervalle $[0, n]$ en plusieurs morceaux.

$$I_{n,k} = \begin{cases} [\frac{k}{2^n}, \frac{k+1}{2^n}[& \text{si } 0 \leqslant k \leqslant n2^n - 1 \\ [n, \infty] & \text{si } k = n2^n. \end{cases} \tag{13.278}$$

Nous posons $S_{n,k} = f^{-1}(I_{n,k})$. Ce sont des ensembles mesurables parce que f est mesurable. Et de plus pour chaque n, la suite $(S_{n,k})_{k \geqslant 0}$ est une partition mesurable finie de S. Nous posons

$$\varphi_n = \sum_{k=0}^{n2^n} \frac{k}{2^n} \mathbb{1}_{S_{n,k}}. \tag{13.279}$$

C'est à dire que sur chaque $S_{n,k}$ nous approximons f par le bas. La fonction φ_n est étagée et positive : $0 \leqslant \varphi_n(x) \leqslant f(x)$ par construction.

Croissance Nous allons voir que $\varphi_n \leqslant \varphi_{n+1}$. Soit $k \neq n2^n$. Si $x \in S_{n,k}$ alors $\varphi_n(x) = \frac{k}{2^n}$ et nous avons aussi la décomposition

$$S_{n,k} = S_{n+1,2k} \cup S_{n+,2k+1}. \tag{13.280}$$

23. Nous rappelons que $\mathbb{Q}$ est dénombrable et dense dans $\mathbb{R}$ par la proposition 6.58.

Si $x \in S_{n+1,2k}$ alors $\varphi_{n+1}(x) = \frac{2k}{2^{n+1}} = \frac{k}{2^n} = \varphi_n(x)$. Et si $x \in S_{n+1,2k+1}$ alors

$$\varphi_{n+1}(x) = \frac{2k+1}{2^{n+1}} = \frac{k+\frac{1}{2}}{2^n} > \varphi_n(x). \tag{13.281}$$

Il reste à traiter le cas $x \in \{f \geqslant n\}$. Dans ce cas nous avons $\varphi_n(x) = n$. il y a encore deux cas à traiter :

$$\{f \geqslant n\} = \{f \in [n, n+1[\} \cup \{f \in [n+1, \infty]\}. \tag{13.282}$$

Pour plus de simplicité dans les notations, nous notons $\bar{n} = n2^n$, c'est à dire que $I_{n,\bar{n}}$ est le $I_{n,k}$ avec le k le plus grand possible. Nous avons

$$I_{n,\bar{n}} = [n, n+1[\cup [n+1, \infty]. \tag{13.283}$$

Le premier élément se décompose en $I_{n+1,k}$ avec $k < n+1$ (nous préciserons plus tard exactement les valeurs de k) tandis que le second est $[n+1, \infty] = I_{n+1,\overline{n+1}}$.

Pour $x \in S_{n+1,\overline{n+1}}$ nous avons

$$\varphi_{n+1}(x) = \frac{(n+1)2^{n+1}}{2^{n+1}} = n+1 > \varphi_n(x). \tag{13.284}$$

Si au contraire $f(x) \in [n, n+1[$ nous devons précisément voir quels sont les k qui font en sorte que $I_{n+1,k}$ recouvre $[n, n+1[$. Le plus petit k est donné par $\frac{k}{2^{n+1}} = n$, c'est à dire $k = n2^{n+1}$ et le plus grand k est donné par $\frac{k}{2^{n+1}} < n+1$, c'est à dire $k = 2^{n+1}(n+1) - 1$. Donc si $f(x) \in [n, n+1[$ alors $x \in S_{n+1,k}$ avec

$$n2^{n+1} \leqslant k \leqslant (n+1)2^{n+1} - 1 \tag{13.285}$$

Dans ce cas

$$\varphi_{n+1}(x) = \frac{k}{2^{n+1}} \geqslant \frac{n2^{n+1}}{2^{n+1}} = n = \varphi_n(x). \tag{13.286}$$

Convergence ponctuelle Si $f(x) < \infty$ alors il existe [24] $n_0 \in \mathbb{N}$ tel que $f(x) < n_0$. Pour $bn \geqslant n_0$ nous avons $f(x) < n$ et donc $\varphi_n(x)$ se calcule à partir d'un des intervalles de taille $1/2^n$:

$$\varphi_n(x) = \frac{k}{2^n} \leqslant f(x) < \frac{k+1}{2^n}. \tag{13.287}$$

Donc

$$|\varphi_n(x) - f(x)| \leqslant \frac{1}{2^n}, \tag{13.288}$$

ce qui signifie que $\lim_{n\to\infty} \varphi_n(x) = f(x)$.

Si $f(x) = +\infty$ alors $f(x) > n$ pour tout n. Et alors $\varphi_n(x) = n$ pour tout n, ce qui donne bien $\varphi_n(x) \to \infty$.

Convergence uniforme Soit f bornée : $0 \leqslant f(x) < M$ pour tout $x \in S$. Soit aussi $\epsilon > 0$. Nous prenons $n_0 > M$ tel que $\frac{1}{2^{n_0}} < \epsilon$. Alors pour tout $n \geqslant n_0$ nous avons

$$0 \leqslant f(x) - \varphi_n(x) \leqslant \frac{1}{2^n} \leqslant \frac{1}{2^{n_0}} \leqslant \epsilon. \tag{13.289}$$

Note qu'il n'y a pas de valeurs absolues parce que nous savons déjà que la limite est croissante.

□

24. Le vrai snob citera ici le lemme 2.49.

13.4.5 Fonctions réelle à variables réelles

Nous nous particularisons à présent au cas de fonctions

$$f\colon \big(\mathbb{R}, \mathcal{B}\text{or}(\mathbb{R})\big) \to \big(\bar{\mathbb{R}}, \mathcal{B}\text{or}(\bar{\mathbb{R}})\big). \tag{13.290}$$

Remarque 13.108.
En théorie de l'intégration il se parle souvent de fonctions mesurables au sens de la tribu de Lebesgue : $f\colon \big(\mathbb{R}, \mathcal{L}\text{eb}(\mathbb{R})\big) \to \big(\bar{\mathbb{R}}, \mathcal{B}\text{or}(\bar{\mathbb{R}})\big)$ où $\mathcal{L}\text{eb}(\mathbb{R})$ est la tribu de Lebesgue sur $\mathbb{R}$, c'est à dire la tribu complétée de celle des boréliens (définition 13.172).

Le fait est qu'en pratique, c'est déjà assez compliqué de construire des fonctions non mesurables au sens des boréliens ; alors on va s'en contenter. Mais construire des fonctions non mesurables au sens de la tribu de Lebesgue, c'est encore plus compliqué.

Nous allons donc nous contenter de donner des conditions assurant qu'une fonction $f\colon \big(\mathbb{R}, \mathcal{B}\text{or}(\mathbb{R})\big) \to (\bar{\mathbb{R}}, \mathcal{B}\text{or}(\bar{\mathbb{R}}))$ soit mesurable. Ces fonctions seront a fortiori mesurables en les considérant comme fonctions $f\colon \big(\mathbb{R}, \mathcal{L}\text{eb}(\mathbb{R})\big) \to (\bar{\mathbb{R}}, \mathcal{B}\text{or}(\bar{\mathbb{R}}))$.

Corollaire 13.109.
Si I est un intervalle de $\mathbb{R}$, alors toute application monotone $f\colon I \to \mathbb{R}$ est borélienne.

Démonstration. Vu que f est monotone, l'ensemble $\{f < a\}$ est un intervalle. Or tous les intervalles sont boréliens, donc f est mesurable par le théorème 13.90. □

Définition 13.110.
Si I est un intervalle de $\mathbb{R}$, une fonction $f\colon I \to \mathbb{R}$ a une propriété (monotone, mesurable, continue, etc.) ***par morceaux*** *s'il existe une suite strictement croissante de points $(x_I)_{i\in\mathbb{Z}}$ dans I telle que f ait la propriété sur chacun des ouverts $]x_j, x_{j+1}[$..*

Dans cette définition, les points sont numérotés par $\mathbb{Z}$ et non par $\mathbb{N}$ parce que nous nous laissons la liberté d'avoir une infinité de points de chacun des deux côtés.

Proposition 13.111.
Soit I un intervalle de $\mathbb{R}$ et une fonction $f\colon I \to \mathbb{R}$. Si f est continue ou monotone par morceaux sur I alors elle y est borélienne.

Démonstration. L'ensemble $\{]x_j, x_{j+1}[\}_{j\in\mathbb{Z}} \cup \{x_i\}_{i\in\mathbb{Z}}$ forme une partition mesurable dénombrable de I (les singletons sont des boréliens). À une belle redéfinition près de la numérotation (deux fois $\mathbb{Z}$ va dans $\mathbb{N}$), nous les appelons $(I_n)_{n\in\mathbb{N}}$, et nous définissons les fonctions f_n comme étant les restrictions de f aux intervalles I_k.

Toute fonctions sur un singleton est mesurable. Toute fonctions continue sur un ouvert est mesurable (théorème 13.72). Toute fonction monotone sur un ouvert est mesurable (corollaire 13.109).

Le lemme de recollement 13.95 donne alors la mesurabilité de f. □

13.112.
Toutes les fonctions que nous pouvons écrire explicitement sont mesurables ... en tout cas toutes celles que l'on trouve en pratique. En effet nous avons déjà toutes les fonctions continues par morceaux via la proposition 13.111 et ensuite toutes les limites par la proposition 13.102. Cela donne les séries, les dérivées, les primitives, etc.

13.5 Intégrale par rapport à une mesure

Nous avons besoin d'un peu de théorie de l'intégration parce que la définition de la mesure sur un espace mesurable [25] produit passe par une intégrale.

25. Théorème 13.160.

Une mesure μ sur un espace mesurable $(\Omega, \mathcal{A})$ permet de définir une fonctionnelle linéaire sur l'ensemble des fonctions mesurables $\Omega \to \mathbb{R}$. Cette fonctionnelle linéaire est l'intégrale que nous allons définir à présent.

Définition 13.113.
Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. Si $Y \in \mathcal{A}$ et si f est une fonction étagée[26]*, si sa forme canonique est $f = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i}$ alors nous définissons*

$$\int_Y f d\mu = \sum_i \alpha_i \mu(Y \cap A_i). \tag{13.291}$$

Pour une fonction $\mathcal{A}$-mesurable générale $f : \Omega \to [0, \infty]$ nous définissons l'intégrale de f sur Y par

$$\int_Y f d\mu = \sup \Big\{ \int_Y h d\mu \text{ où } h \text{ est une fonction étagée telle que } 0 \leqslant h \leqslant f \Big\}. \tag{13.292}$$

Si f est mesurable à valeurs dans $\bar{\mathbb{R}}$ ou $\mathbb{C}$, l'intégrale se définit séparément pour les parties positives, négatives, réelles et imaginaires.

Remarque 13.114.
Toute fonction mesurable à valeurs dans $\bar{\mathbb{R}}$ est intégrable (l'intégrale vaut éventuellement $+\infty$). Au moment où une fonction commence à prendre des valeurs positives et négatives, nous demandons à pouvoir intégrer séparément les parties positive et négative. C'est pour cela que nous disons qu'une fonction f à valeurs dans $\mathbb{R}$ est intégrable si $|f|$ l'est.

Cela est indépendant du fait que $\int_0^\infty f$ en tant que limite de $\int_0^M f$ peut très bien exister grâce à des compensations, alors que $\int_{[0,\infty[} |f|$ n'existe pas.

Lemme 13.115.
Si $(\Omega, \mathcal{A}, \mu)$ est un espace mesuré et si $B \in \mathcal{A}$ alors

$$\mu(B) = \int_B 1 d\mu = \int_\Omega \mathbb{1}_B. \tag{13.293}$$

Démonstration. La fonction caractéristique d'une partie mesurable est une fonction étagée dont la forme canonique est $\mathbb{1}_B = 1 \cdot \mathbb{1}_B + 0 \times \mathbb{1}_{B^c}$. Son intégrale est donc

$$\int \mathbb{1}_B d\mu = 1 \times \mu(B) + 0 \times \mu(B^c) = \mu(B) \tag{13.294}$$

parce que $0 \times \mu(B^c) = 0$, même si $\mu(B^c) = \infty$, comme nous l'avons convenu en 13.82. □

Définition 13.116.
*Une fonction $f : \Omega \to \mathbb{R}$ est dite **intégrable** au sens de Lebesgue si $\int_\Omega |f| < \infty$. Dans ce cas nous définissons*

$$\int_\Omega f = \int_\Omega f^+ - \int_\Omega f^- \tag{13.295}$$

où f^+ et f^- sont les parties positives et négatives de f ; les deux intégrales à droite dans (13.295) *sont finies dès que f est intégrables.*

Nous verrons comment donner un sens à $\int_\Omega f$ dans certains cas où f n'est pas intégrable sur Ω dans la section 13.11.4 sur les intégrales impropres.

Nous définissons aussi

$$\mu(f) = \int_\Omega f \tag{13.296}$$

si f est une fonction mesurable sur Ω.

26. Définition 13.103.

Remarque 13.117.
Dans $\mathbb{R}^d$, quasiment toutes les fonctions et ensembles sont mesurables. En effet la construction d'ensembles non mesurables demande obligatoirement l'utilisation de l'axiome du choix ; de tels ensembles doivent être construits «exprès pour». Il y a très peu de chances pour que vous tombiez sur un ensemble non mesurable de $\mathbb{R}^d$ sans que vous ne vous en rendiez compte.

Remarque 13.118.
«Mesurable» ne signifie pas «intégrable». Par exemple la fonction

$$\begin{aligned} f\colon \mathbb{R} &\to \bar{\mathbb{R}} \\ \omega &\mapsto \begin{cases} \frac{1}{\omega} & \text{si } \omega \neq 0 \\ \infty & \text{if} \omega = 0. \end{cases} \end{aligned} \tag{13.297}$$

est mesurable, mais non intégrable.

13.5.1 Quelque propriétés

Théorème 13.119 ([161])**.**
Soient f, g des fonctions étagées positives sur $(S, \mathcal{F}, \mu)$. Alors si $\alpha \in [0, \infty]$ nous avons

(1)

$$\int_S (\alpha f) d\mu = \alpha \int_S f d\mu. \tag{13.298}$$

(2)

$$\int_S (f + g) d\mu = \int_S f d\mu + \int_S g d\mu. \tag{13.299}$$

(3) Si $a_k \in \mathbb{R}^+$ et si les f_k sont étagées positives,

$$\int_S \left(\sum_{k=1}^{n} a_k f_k \right) = \sum_{k=1}^{n} a_k \left(\int_S f_k d\mu \right). \tag{13.300}$$

Démonstration. En ce qui concerne le produit par un nombre, tout repose sur le fait que

$$(\alpha f)^{-1}(\alpha a_i) = f^{-1}(a_i), \tag{13.301}$$

ce qui fait que si la représentation canonique de f est $f = \sum_i a_i \mathbb{1}_{A_i}$ alors la représentation canonique de αf est $\alpha f = \sum_i (\alpha a_i) \mathbb{1}_{A_i}$. Donc

$$\int_S \alpha f d\mu = \sum_i \alpha a_i \mu(A_i) = \alpha \sum_i a_i \mu(A_i) = \alpha \int_S f d\mu. \tag{13.302}$$

Pour la somme c'est plus lourd. Soient les formes canoniques

$$f = \sum_i a_i \mathbb{1}_{A_i} \tag{13.303a}$$

$$g = \sum_j b_j \mathbb{1}_{B_i}. \tag{13.303b}$$

Vu que l'union des B_j est S nous avons l'union disjointe $A_i = \bigcup_j A_i \cap B_j$ et donc $\mu(A_i) = \sum_j \mu(A_i \cap B_j)$. Nous avons donc pour les intégrales :

$$\int_S f d\mu = \sum_i a_i \sum_j \mu(A_i \cap B_j) \tag{13.304a}$$

$$\int_S g d\mu = \sum_i b_k \sum_l \mu(B_k \cap A_l). \tag{13.304b}$$

Pour la somme :

$$\int_S f d\mu + \int_S g d\mu = \sum_{k,l} (a_k + b_l)\mu(A_k \cap B_l). \tag{13.305}$$

Nous devons maintenant évaluer $\int_S (f+g)d\mu$. Pour cela nous remarquons que si $c \in (f+g)(S)$ (l'ensemble des valeurs atteintes pas $f+g$), alors nous notons

$$I_c = \{(k,l) \text{ tel que } a_k + b_l = c\} \tag{13.306}$$

et nous avons

$$\{f + g = c\} = \bigcup_{(k,l)\in I_c} (A_k \cap B_l), \tag{13.307}$$

et comme cette union est disjointe, nous pouvons faire la somme des mesures :

$$\mu(f + g = c) = \sum_{(k,l)\in I_c} \mu(A_k \cap B_l). \tag{13.308}$$

Cela nous permet de faire le calcul suivant :

$$\begin{aligned}
\int_S (f+g)d\mu &= \sum_{c\in(f+g)(S)} c\mu(f+g=c) && (13.309a)\\
&= \sum_{c\in(f+g)(S)} c \sum_{(k,l)\in I_c} \mu(A_k \cap B_l) && (13.309b)\\
&= \sum_{c\in(f+g)(S)} \sum_{(k,l)\in I_c} (a_k + b_l)\mu(A_k \cap B_l) && (13.309c)
\end{aligned}$$

Dans cette double somme, tous les couples (k,l) sont tirés une et une seule fois parce qu'ils sont tous dans un et un seul des I_c, donc

$$\begin{aligned}
\int_S (f+g)d\mu &= \sum_{c\in(f+g)(S)} \sum_{(k,l)\in I_c} (a_k + b_l)\mu(A_k \cap B_l) && (13.310a)\\
&= \sum_{(k,l)} (a_k + b_l)\mu(A_k \cap B_l) && (13.310b)\\
&= \int_S f d\mu + \int_S g d\mu. && (13.310c)
\end{aligned}$$

□

Remarque 13.120.
Si $f = \sum_k a_k \mathbb{1}_{A_k}$ n'est pas une décomposition canonique, il n'en reste pas moins que chacun des $\mathbb{1}_{A_k}$ est la forme canonique de lui-même. Donc le théorème 13.119 s'applique et nous avons quand même

$$\int_S f d\mu = \sum_k a_k \mu(A_k). \tag{13.311}$$

Proposition 13.121.
Si f, g sont des fonctions étagées et si $A, B \subset S$ sont disjoints, alors

(1) $\int_A f d\mu = \int_S f \mathbb{1}_A d\mu$

(2) $\int_{A\cup B} f d\mu = \int_A f d\mu + \int_B f d\mu$.

Démonstration. Si $f = \sum_k a_k \mathbb{1}_{B_k}$ alors d'une part

$$\int_A f d\mu = \sum_k a_k \mu(B_k \cap A) \tag{13.312}$$

et d'autre part,

$$f\mathbb{1}_A)d = \sum_k a_k \mathbb{1}_A \mathbb{1}_{B_k} = \sum_k a_k \mathbb{1}_{B_k \cap A}, \tag{13.313}$$

ce qui donne

$$\int f\mathbb{1}_A d\mu = \sum_k a_k \mu(B_k \cap A). \tag{13.314}$$

En ce qui concerne la seconde égalité à prouver, tout repose sur le fait que $\mathbb{1}_{A\cup B=\mathbb{1}_A+\mathbb{1}_B}$. Du coup nous avons, en utilisant le théorème 13.119 :

$$\int_{A\cup B} f d\mu = \int_S f\mathbb{1}_{A\cup B} d\mu \tag{13.315a}$$

$$= \int_S f(\mathbb{1}_A + \mathbb{1}_B) d\mu \tag{13.315b}$$

$$= \int_S f\mathbb{1}_A + \int_S f\mathbb{1}_B \tag{13.315c}$$

$$= \int_A f + \int_B f. \tag{13.315d}$$

□

Proposition 13.122.
Si $A, B \subset \Omega$ sont des ensembles disjoints et si f est intégrable sur $A \cup B$ alors

$$\int_{A\cup B} f = \int_A f + \int_B f. \tag{13.316}$$

Le lemme suivant nous aide à détecter des fonctions presque partout nulles.

Lemme 13.123.
Soit f une fonction mesurable positive ou nulle telle que

$$\int_\Omega f d\mu = 0. \tag{13.317}$$

Alors $f = 0$ μ-presque partout.

Démonstration. L'ensemble des points $x \in \Omega$ tels que $f(x) \neq 0$ peut s'écrire comme une union dénombrable disjointe :

$$\{x \in \Omega \text{ tel que } f(x) \neq 0\} = \bigcup_{i=0}^{\infty} E_i \tag{13.318}$$

avec

$$E_0 = \{x \in \Omega \text{ tel que } f(x) > 1\} \tag{13.319a}$$

$$E_i = \{x \in \Omega \text{ tel que } \frac{1}{i+1} \leqslant f(x) < \frac{1}{i}\}. \tag{13.319b}$$

Si un des ensembles E_i est de mesure non nulle, alors nous pouvons considérer la fonction simple $h(x) = \frac{1}{i+1}\mathbb{1}_{E_i}$ dont l'intégrale sur Ω est strictement positive. Par conséquent le supremum de la définition (13.292) est strictement positif.

Nous savons donc que $\mu(E_i) = 0$ pour tout i. Étant donné que la mesure d'une union disjointe dénombrable est égale à la somme des mesures, nous avons

$$\mu\{x \in \Omega \text{ tel que } f(x) \neq 0\} = 0, \tag{13.320}$$

ce qui signifie que f est nulle μ-presque partout. □

Corollaire 13.124.
Soit f une fonction mesurable sur l'espace mesuré $(\Omega, \mathcal{A}, \mu)$ telle que

$$\int_\Omega f \mathbb{1}_{f>0} d\mu = 0. \tag{13.321}$$

Alors $f \leqslant 0$ presque partout.

Démonstration. Nous avons l'égalité d'ensembles

$$\{f\mathbb{1}_{f>0} \neq 0\} = \{\mathbb{1}_{f>0} \neq 0\}. \tag{13.322}$$

Mais lemme 13.123 implique que $f\mathbb{1}_{f>0}$ est nulle presque partout, c'est à dire que la mesure de l'ensemble du membre de gauche est nulle par conséquent

$$\mu\{\mathbb{1}_{f>0} \neq 0\} = 0. \tag{13.323}$$

Cela signifie que la fonction f est presque partout négative ou nulle. □

Lemme 13.125.
Soit f une fonction telle que $|f(x)| \leqslant g(x)$ pour tout $x \in \Omega$. Si g est intégrable, alors f est intégrable.

Démonstration. Nous décomposons f en parties positives et négatives :

$$A_+ = \{x \in \Omega \text{ tel que } f(x) > 0\} \tag{13.324a}$$
$$A_- = \{x \in \Omega \text{ tel que } f(x) < 0\}. \tag{13.324b}$$

Nous posons $f_+(x) = f(x)\mathbb{1}_{A_+}$ et $f_-(x) = f(x)\mathbb{1}_{A_-}$. Nous avons $f = f_+ - f_-$ et

$$\int_\Omega f = \int_{A_+} f + \int_{A_-} f \tag{13.325}$$

parce que $\Omega = A_+ \cup A_- \cup \{x \in \Omega \text{ tel que } f(x) = 0\}$. Si φ est une fonction simple qui majore f_+ nous avons

$$\varphi(x) = \sum_k a_k \mathbb{1}_{E_k}(x) \leqslant f(x)\mathbb{1}_{A_+}(x) \leqslant g(x). \tag{13.326}$$

Par conséquent le supremum qui définit $\int f_+$ est inférieur au supremum qui définit $\int g$. La fonction f_+ est donc intégrable. La même chose est valable pour la fonction f_-. □

13.5.2 Permuter limite et intégrale

13.5.2.1 Convergence uniforme

Proposition 13.126 (Permuter limite et intégrale)**.**
Soit $f_n \to f$ uniformément sur un ensemble mesuré A de mesure finie. Alors si les fonctions f_n et f sont intégrables sur A, nous avons

$$\lim_{n\to\infty} \int_A f_n = \int_A \lim_{n\to\infty} f_n. \tag{13.327}$$

Démonstration. Notons f la limite de la suite (f_n). Pour tout n nous avons les majorations

$$\left| \int_A f_n d\mu - \int_A f d\mu \right| \leqslant \int_A |f_n - f| d\mu \tag{13.328a}$$
$$\leqslant \int_A \|f_n - f\|_\infty d\mu \tag{13.328b}$$
$$= \mu(A)\|f_n - f\|_\infty \tag{13.328c}$$

où $\mu(A)$ est la mesure de A. Le résultat découle maintenant du fait que $\|f_n - f\|_\infty \to 0$. □

Il existe un résultat considérablement plus intéressant que cette proposition. En effet, l'intégrabilité de f n'est pas nécessaire. Cette hypothèse peut être remplacée soit par l'uniforme convergence de la suite (théorème 13.127), soit par le fait que les normes des f_n sont uniformément bornées (théorème de la convergence dominée de Lebesgue 13.134).

Théorème 13.127 ([171]).
La limite uniforme d'une suite de fonctions intégrables sur un borné est intégrable, et on peut permuter la limite et l'intégrale.

Plus précisément, soit A un ensemble de μ-mesure finie et $f_n\colon A \to \mathbb{R}$ des fonctions intégrables sur A. Si la limite $f_n \to f$ est uniforme, alors f est intégrable sur A et nous pouvons inverser la limite et l'intégrale :

$$\lim_{n\to\infty} \int_A f_n = \int_A \lim_{n\to\infty} f_n. \tag{13.329}$$

Démonstration. Soit $\epsilon > 0$ et n tel que $\|f_n - f\|_\infty \leqslant \epsilon$ (ici la norme uniforme est prise sur A). Étant donné que f_n est intégrable sur A, il existe une fonction simple φ_n qui minore f_n telle que

$$\left| \int_A \varphi_n - \int_A f_n \right| < \epsilon. \tag{13.330}$$

La fonction $\varphi_n + \epsilon$ est une fonction simple qui majore la fonction f. Si ψ est une fonction simple qui minore f, alors

$$\int_A \psi \leqslant \int_A \varphi_n + \epsilon \leqslant \int_A f_n + \epsilon\mu(A). \tag{13.331}$$

Par conséquent le supremum qui définit $\int_A f$ existe, ce qui montre que f est intégrable. Le fait qu'on puisse inverser la limite et l'intégrale est maintenant une conséquence de la proposition 13.126. □

Remarque 13.128.
L'hypothèse sur le fait que A soit de mesure finie est importante. Il n'est pas vrai qu'une suite uniformément convergente de fonctions intégrables est intégrables. En effet nous avons par exemple la suite

$$f_n(x) = \begin{cases} 1/x & \text{si } x < n \\ 0 & \text{sinon} \end{cases} \tag{13.332}$$

qui converge uniformément vers $f(x) = 1/x$ sur $A = [1, \infty[$. Le limite n'est cependant guerre intégrable sur A.

13.5.2.2 Convergence monotone

Théorème 13.129 (Théorème de la convergence monotone ou de Beppo-Levi[172]).
Soit un espace mesuré $(\Omega, \mathcal{A}, \mu)$ et (f_n) une suite croissante de fonctions mesurables à valeurs dans $[0, \infty]$. Alors la limite ponctuelle $\lim_{n\to\infty} f_n$ existe, est mesurable et

$$\lim_{n\to\infty} \int_\Omega f_n d\mu = \int_\Omega \lim_{n\to\infty} f_n d\mu, \tag{13.333}$$

cette intégrable valant éventuellement ∞.

Démonstration. La limite ponctuelle de la suite est la fonction à valeurs dans $[0, \infty]$ donnée par

$$f(x) = \lim_{n\to\infty} f_n(x). \tag{13.334}$$

Ces limites existent parce que pour chaque x la suite $f_n(x)$ est une suite numérique croissante. Nous notons

$$I_0 = \int_\Omega f d\mu. \tag{13.335}$$

Nous posons par ailleurs

$$I_n = \int_\Omega f_n. \tag{13.336}$$

Cela est une suite numérique croissante qui a par conséquent une limite que nous notons $I = \lim_{n\to\infty} I_n$. Notre objectif est de montrer que $I = I_0$. D'abord par croissance de la suite, pour tous n nous avons $I_n \leqslant I_0$, par conséquent $I \leqslant I_0$.

Nous prouvons maintenant l'inégalité dans l'autre sens en nous servant de la définition (13.292). Soit une fonction simple h telle que $h \leqslant f$, et une constante $0 < C < 1$. Nous considérons les ensembles

$$E_n = \{x \in \Omega \text{ tel que } f_n(x) \geqslant Ch(x)\}. \tag{13.337}$$

Ces ensembles vérifient les propriétés $E_n \subset E_{n+1}$ et $\bigcup_{n=1}^{\infty} E_n = \Omega$. Pour chaque n nous avons les inégalités

$$\int_\Omega f_n \geqslant \int_{E_n} f_n \geqslant C \int_{E_n} h. \tag{13.338}$$

Si nous prenons la limite $n \to \infty$ dans ces inégalités,

$$\lim_{n\to\infty} \int_\Omega f_n \geqslant C \lim_{n\to\infty} \int_{E_n} h = C \int_\Omega h. \tag{13.339}$$

Par conséquent $\lim_{n\to\infty} \int f_n \geqslant C \int_\Omega h$. Mais étant donné que cette inégalité est valable pour tout C entre 0 et 1, nous pouvons l'écrire sans le C :

$$\lim_{n\to\infty} \int_\Omega f_n \geqslant \int_\Omega h. \tag{13.340}$$

Par définition, l'intégrale de f est donné par le supremum des intégrales de h où h est une fonction simple dominée par f. En prenant le supremum sur h dans l'équation (13.340) nous avons

$$\lim_{n\to\infty} \int_\Omega f_n \geqslant \int_\Omega f, \tag{13.341}$$

ce qu'il nous fallait. □

Remarque 13.130.
La proposition 13.107 ainsi que le lemme 13.104 montrent qu'une fonction mesurable peut-être écrite comme limite croissante de fonctions simples. Cela permet de démontrer des théorèmes en commençant par prouver sur les fonctions simples et en utilisant Beppo-Levi pour généraliser.

Remarque 13.131.
Une des raisons de demander la positivité des fonctions f_n est de n'avoir pas d'ambiguïté à parler d'intégrales qui valent ∞. Si par exemple nous prenons $\Omega = [0, 1]$ et que nous considérons

$$f_n(x) = \begin{cases} 0 & \text{si } x \leqslant \frac{1}{n} \\ \frac{1}{x} & \text{sinon.} \end{cases} \tag{13.342}$$

Ce sont des fonctions intégrables, mais la limite étant la fonction $1/x$, l'égalité (13.333) est une égalité entre deux intégrales valant ∞.

Corollaire 13.132 (Inversion de somme et intégrales)**.**
Si (u_n) est une suite de fonctions mesurables positives ou nulles, alors

$$\sum_{i=0}^{\infty} \int u_i = \int \sum_{i=0}^{\infty} u_i. \tag{13.343}$$

Démonstration. Nous considérons la suite des sommes partielles de (u_n) : $f_n(x) = \sum_{i=0}^{n} u_n(x)$. Le théorème de la convergence monotone (théorème 13.129) implique que

$$\lim_{n\to\infty} \int f_n = \int \lim_{n\to\infty} f_n. \tag{13.344}$$

Nous remplaçons maintenant f_n par sa valeur en termes des u_i et dans le membre de gauche nous permutons l'intégrale avec la somme finie :

$$\lim_{n\to\infty} \sum_{i=0}^{\infty} \int u_n = \int \sum_{i=0}^{\infty} u_n, \tag{13.345}$$

ce qu'il fallait démontrer. □

Lemme 13.133 (Lemme de Fatou)**.**
Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et $f_n \colon \Omega \to [0,\infty]$ une suite de fonctions mesurables. Alors la fonction $f(x) = \liminf f_n(x)$ est mesurable et

$$\int_\Omega \liminf f_n d\mu \leqslant \liminf \int_\Omega f d\mu. \tag{13.346}$$

Démonstration. Nous posons

$$g_n(x) = \inf_{i\geqslant n} f_i(x). \tag{13.347}$$

Cela est une suite croissance de fonctions positives mesurables telles que, par définition,

$$\lim_{n\to\infty} g_n(x) = \liminf f_n(x). \tag{13.348}$$

Nous pouvons y appliquer le théorème de la convergence monotone,

$$\lim_{n\to\infty} \int g_n(x) = \int \liminf f_n(x). \tag{13.349}$$

Par ailleurs, pour chaque $i \geqslant n$ nous avons

$$\int g_n \leqslant \int f_i, \tag{13.350}$$

en passant à l'infimum nous avons

$$\int g_n \leqslant \inf_{i\geqslant n} \int f_i, \tag{13.351}$$

et en passant à la limite nous avons

$$\int \liminf f_n = \lim_{n\to\infty} \int g_n \leqslant \lim_{n\to\infty} \inf_{i\geqslant n} \int f_i = \liminf_{i\to\infty} \inf f_i. \tag{13.352}$$

□

L'inégalité donnée dans ce lemme n'est en général pas une égalité, comme le montre l'exemple suivant :

$$f_i = \begin{cases} \mathbb{1}_{[0,1]} & \text{si } i \text{ est pair} \\ \mathbb{1}_{[1,2]} & \text{si } i \text{ est impair.} \end{cases} \tag{13.353}$$

Nous avons évidemment $g_n(x) = 0$ tandis que $\int_{[0,2]} f_i = 1$ pour tout i.

13.5.2.3 Convergence dominée de Lebesgue

Théorème 13.134 (Convergence dominée de Lebesgue)**.**
Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions intégrables sur $(\Omega, \mathcal{A}, \mu)$ à valeurs dans $\mathbb{C}$ ou $\mathbb{R}$. Nous supposons que $f_n \to f$ simplement sur Ω presque partout et qu'il existe une fonction intégrable g telle que

$$|f_n(x)| \leqslant g(x) \tag{13.354}$$

pour presque[27] *tout $x \in \Omega$ et pour tout $n \in \mathbb{N}$. Alors*

(1) f est intégrable,
(2) $\lim_{n\to\infty} \int_\Omega f_n = \int_\Omega f$,
(3) $\lim_{n\to\infty} \int_\Omega |f_n - f| = 0$.

Démonstration. La fonction limite f est intégrable parce que $|f| \leqslant g$ et g est intégrable (lemme 13.125). Par hypothèse nous avons

$$-g(x) \leqslant f_n(x) \leqslant g(x). \tag{13.355}$$

En particulier la fonction $g_n = f_n + g$ est positive et mesurable si bien que le lemme de Fatou 13.133 implique

$$\int_\Omega \liminf g_n \leqslant \liminf \int_\Omega g_n. \tag{13.356}$$

Évidement nous avons $\liminf g_n = f + g$, de telle sorte que

$$\int f + \int g \leqslant \liminf \int g_n = \liminf \int f_n + \int g, \tag{13.357}$$

et le nombre $\int g$ étant fini, nous pouvons le retrancher des deux côtés de l'inégalité :

$$\int f \leqslant \liminf \int f_n. \tag{13.358}$$

Afin d'obtenir une minoration de $\int f$ nous refaisons exactement le même raisonnement en utilisant la suite de fonctions $k_n = -f_n \to k = -f$. Nous obtenons que

$$\int k \geqslant \liminf \int k_n = -\limsup \int f_n, \tag{13.359}$$

et par conséquent

$$\liminf \int f_n \leqslant \int f \leqslant \limsup \int f_n. \tag{13.360}$$

La limite supérieure étant plus grande ou égale à la limite inférieure, les trois quantités dans les inégalités (13.360) sont égales.

Nous prouvons maintenant le troisième point. Soit la suite de fonctions

$$h_n(x) = |f_n(x) - f(x)| \tag{13.361}$$

qui tend ponctuellement vers zéro. De plus

$$h_n(x) \leqslant |f_n(x)| + |f(x)| \leqslant 2g(x), \tag{13.362}$$

ce qui prouve que les h_n majorés par une fonction intégrable. Donc

$$\lim_{n\to\infty} \int_\Omega |f_n - f| = \lim_{n\to\infty} \int_\Omega h_n(x)dx = \int_\Omega \lim_{n\to\infty} |f_n(x) - f(x)| = 0 \tag{13.363}$$

□

27. S'il n'y avait pas le «presque» ici, ce théorème serait à peu près inutilisable en probabilité ou en théorie des espaces L^p, comme dans la démonstration du théorème de Fischer-Riesz 24.30 par exemple.

Remarque 13.135.
Lorsque nous travaillons sur des problèmes de probabilités, la fonction g peut être une constante parce que les constantes sont intégrables sur un espace de probabilité.

Corollaire 13.136.
Soit $(a_i)_{i\in\mathbb{N}}$ une suite numérique absolument convergente. Alors elle est convergente. Il en est de même pour les séries de fonctions si on considère la convergence ponctuelle.

Démonstration. L'hypothèse est la convergence de l'intégrale $\int_{\mathbb{N}} |a_i| dm(i)$ où dm est la mesure de comptage. Étant donné que $|a_i| \leqslant |a_i|$, la fonction a_i (fonction de i) peut jouer le rôle de g dans le théorème de la convergence dominée de Lebesgue (théorème 13.134). □

13.5.3 Produit d'une mesure par une fonction (mesure à densité)

Proposition-définition 13.137 ([1, 173])**.**
Soit un espace mesuré $(S, \mathcal{F}, \mu)$ et une fonction mesurable positive $w\colon S \to \bar{\mathbb{R}}^+$. Alors la formule

$$(w \cdot \mu)(A) = \int_A w d\mu \tag{13.364}$$

*pour tout $A \in \mathcal{F}$ définit une mesure positive sur $(S, \mathcal{F})$ appelée **produit** de la mesure μ par la fonction w. La fonction w est la **densité** de la mesure $w \cdot \mu$ par rapport à la mesure μ.*

Démonstration. D'abord $(w \cdot \mu)(\varnothing) = 0$ parce que le lemme 13.115 donne

$$(w \cdot \mu)(\varnothing) = \int_S w \mathbb{1}_\varnothing d\mu = \int_S 0 d\mu = 0 \times \mu(S) = 0 \tag{13.365}$$

où nous avons (éventuellement) utilisé deux fois la convention $0 \times \infty = 0$.

Ensuite si les ensembles A_i sont des éléments deux à deux disjoints de $\mathcal{F}$ alors nous avons $\mathbb{1}_{\bigcup_{i=1}^\infty} = \sum_{i=1}^\infty \mathbb{1}_{A_i}$, et donc

$$(w \cdot \mu)(\bigcup_i A_i) = \int_S w \mathbb{1}_{\bigcup_i A_i} d\mu = \int_S (\sum_{i=1}^\infty w \mathbb{1}_{A_i}) d\mu = \sum_i \int_S w \mathbb{1}_{A_i} d\mu = \sum_i (w \cdot \mu)(A_i). \tag{13.366}$$

Dans ce calcul nous avons utilisé le fait que f était positive pour justifier l'application du théorème de la convergence monotone 13.129. □

En particulier nous parlons souvent de mesure à densité par rapport à la mesure de Lebesgue. C'est alors la construction suivante. Si μ est une mesure sur $\mathbb{R}^d$, une fonction $f\colon \mathbb{R}^d \to \mathbb{R}$ est une **densité** si pour tout $A \subset \mathbb{R}^d$ nous avons

$$\mu(A) = \int_A f(x) dx \tag{13.367}$$

où dx est la mesure de Lebesgue.

Proposition 13.138 ([173])**.**
Soit une fonction mesurable $w\colon (S, \mathcal{F}, \mu) \to \bar{\mathbb{R}}^+$.

(1) Si $f\colon S \to \bar{\mathbb{R}}^+$ est mesurable, alors $f \cdot (w \cdot \mu) = (fg) \cdot \mu$.

(2) Si $f\colon S \to \bar{\mathbb{R}}$ ou $\mathbb{C}$ est mesurable, elle est $w \cdot \mu$-intégrable si et seulement si fw est μ-intégrable. Dans ce cas, nous avons encore $f \cdot (w \cdot \mu) = (fg) \cdot \mu$.

Attention : dans le cas où f est à valeurs dans $\mathbb{C}$, alors il faut que w soit à valeurs finies dans $\mathbb{R}$ parce que nous n'avons pas définit $\infty \times z$ lorsque $z \in \mathbb{C}$.

Démonstration. Nous commençons par prouver le résultat pour la fonction caractéristique de l'ensemble mesurable A. Nous avons : $\mathbb{1}_A \cdot (w \cdot \mu)(B) = \int_B \mathbb{1}_A d(w \cdot \mu)$. Mais par définition, l'intégrale d'une fonction indicatrice est la mesure de l'ensemble indiqué. En passant sur le fait que $\mathbb{1}_A \mathbb{1}_B = \mathbb{1}_{A\cap B}$,

$$\int_B \mathbb{1}_A d(w \cdot \mu) = (w \cdot \mu)(A \cap B) = \int_S \mathbb{1}_{A\cap B} w d\mu = \int_S \mathbb{1}_A \mathbb{1}_B w d\mu = \int_B \mathbb{1}_A w d\mu = (\mathbb{1}_A w) \cdot \mu(B). \tag{13.368}$$

Supposons maintenant que f soit une fonction étagées qui s'écrit $f = \sum_k a_k \mathbb{1}_{A_k}$ où les A_k sont des ensembles mesurables disjoints. Alors le calcul est le suivant, en utilisant le fait que sur A_k, on a $a_k = f(x)$:

$$\begin{aligned}
f \cdot (g \cdot \mu)B &= \int_B f d(g \cdot \mu) && \text{(13.369a)}\\
&= \sum_k a_k (g \cdot \mu)(A_k \cap B) && \text{(13.369b)}\\
&= \sum_k a_k \int_{A_k \cap B} gf\mu && \text{(13.369c)}\\
&= \int_{A_k \cap} f(x)g(x)d\mu(x) && \text{(13.369d)}\\
&= \sum_k (fg \cdot \mu)(A_k \cap B) && \text{(13.369e)}\\
&= (fg \cdot \mu)(B) && \text{(13.369f)}
\end{aligned}$$

parce que les $A_k \cap B$ forment une partition de l'ensemble B (voir le point (2) de la définition 13.23).

Si $f\colon S \to \bar{\mathbb{R}}^+$ est mesurable, le théorème 13.107 donne une suite croissante f_n de fonctions étagées positives convergeant (ponctuellement) vers f. Vu que la fonction w est positive, nous avons aussi la limite positive et croissante $wf_n \to wf$. Ainsi l'utilisation du théorème de la convergence monotone est justifié dans le calcul suivant :

$$\int_S f d(w \cdot \mu) = \lim_{n\to\infty} \int_S f_n d(w \cdot \mu) = \lim_{n\to\infty} \int_S (wf_n) d\mu = \int_S wf d\mu. \tag{13.370}$$

Nous passons maintenant au cas général où f est une fonction à valeurs dans $\bar{\mathbb{R}}$ ou $\mathbb{C}$ (avec w finie dans ce dernier cas). Nous avons la chaîne d'équivalences

$\Leftrightarrow$ f est $(w \cdot \mu)$ intégrable
$\Leftrightarrow$ $|f|$ est $(w \cdot \mu)$-intégrable
$\Leftrightarrow$ $|f|w$ est μ-intégrable
$\Leftrightarrow$ $|fw|$ est μ-intégrable.

Si cela est le cas, la formule se démontre en se ramenant au cas déjà prouvé des fonctions positives en utilisant les $(fw)^+ = f^+w$, $(fw)^- = f^-w$ etc. □

13.5.4 Mesure et topologie

Exemple 13.139(Un compact n'est pas toujours de mesure finie)
Soit l'espace mesurable $(\mathbb{R}, \mathcal{B}\text{or}(\mathbb{R}))$ réel avec ses boréliens et la fonction

$$\begin{aligned}
w\colon \big(\mathbb{R}, \mathcal{B}\text{or}(\mathbb{R})\big) &\to \big(\bar{\mathbb{R}}, \mathcal{B}\text{or}(\bar{\mathbb{R}})\big)\\
x &\mapsto \begin{cases} \frac{1}{|x|} & \text{si } x \neq 0\\ +\infty & \text{si } x = 0. \end{cases}
\end{aligned} \tag{13.371}$$

Essayons d'étudier la mesure de densité w par rapport à la mesure de Lebesgue.

w **est mesurable** Soit un borélien B de $\bar{\mathbb{R}}$. Si B ne contient pas ∞ alors $w^{-1}(B)$ est un borélien de $\mathbb{R}$ par continuité de l'application restreinte $w\colon \mathbb{R}\backslash\{0\} \to \mathbb{R}$. Ici nous avons par exemple appliqué la proposition 13.111 à chacun des deux intervalles $]-\infty, 0[$ et $]0, \infty[$. Si $+\infty \in B$ alors

$$w^{-1}(B) = w^{-1}\big(B\backslash\{0\}\big) \cup w^{-1}(\{\infty\}) = w^{-1}\big(B\backslash\{0\}\big) \cup \{0\}, \tag{13.372}$$

qui est borélien par union de boréliens.

Mesure produit La proposition 13.137 nous assure alors qu'en posant [28]

$$\mu(B) = \int_B \frac{1}{|x|} d\lambda(x) \tag{13.373}$$

où λ est la mesure de Lebesgue, nous avons une mesure.

Mesure du singleton Pour avoir les idées claires, nous pouvons nous demander la mesure $\mu\big(\{0\}\big)$. Nous cela nous devons calculer

$$\int_{\{0\}} \frac{1}{|x|} d\lambda(x) = \int_{\{0\}} w(x) d\lambda(x) \tag{13.374}$$

où là, l'abus de notation n'est plus possible. Mais quelle que soit la fonction étagée $h = \sum_i \alpha_i \mathbb{1}_{A_i}$ considérée,

$$\int_{\{0\}} h(x) d\lambda(x) = \sum_i \alpha_i \lambda\big(A_i \cap \{0\}\big) = 0. \tag{13.375}$$

Attention : ceci n'a rien de particulier à la fonction $x \mapsto 1/|x|$. Lorsqu'une mesure a une densité par rapport à Lebesgue, la mesure d'un singleton sera toujours nulle.

Mesure de la boule compacte Il n'en reste pas moins que $\mu\big([-1,1]\big) = \infty$.

△

13.140.
En réalité, il n'y a pas de liens forts entre mesure et topologie. Un espace topologique est une chose, et y mettre une mesure en est une autre. Bien entendu, une topologie étant donnée, nous pouvons considérer la tribu des boréliens et y mettre une mesure un peu quelconque. Il n'y a pas de choix canonique.

Notons que même dans l'exemple de compact de mesure infinie 13.139, la mesure introduite n'est pas sans lien avec la topologie de $\mathbb{R}$. En effet pour avoir une mesure à densité par rapport à Lebesgue, nous avons dû prendre une application mesurable par rapport à la tribu des boréliens, laquelle est éminemment liée à la topologie. Il y a donc parfaitement moyen de construire des espaces mesurés tenant compte de la topologie, et ayant des propriétés qui ne sont pas celle attendues.

13.6 Propriétés

Théorème 13.141 ([173])**.**
Soient deux espaces mesurables $(S_1, \mathcal{F}_1)$ et $(S_2, \mathcal{F}_2)$ ainsi qu'une application mesurable $\varphi\colon S_1 \to S_2$. Soit encore μ, une mesure positive sur $(S_1, \mathcal{F}_1)$.

Si $f\colon S_2 \to \bar{\mathbb{R}}$ ou $\mathbb{C}$ est mesurable alors,

(1) f est $\varphi(\mu)$-intégrale si et seulement si $f \circ \varphi$ est μ-intégrable.

(2) dans le cas où f est $\varphi(\mu)$-intégrable, nous avons

$$\int_{S_2} f d\big(\varphi(\mu)\big) = \int_{S_1} (f \circ \varphi) d\mu. \tag{13.376}$$

28. Avec un mini abus de notation : si $0 \in B$, cette notation n'est pas tout à fait correcte.

Démonstration. L'intégrabilité est la définition 13.116, et demande que $|f|$ soit intégrable. L'égalité (13.376) a un sens si les deux membres sont infinis. Tant que les fonctions considérées sont positives, le point (1) est immédiat. Ce n'est qu'au moment où les fonctions considérées deviennent à valeurs dans $\mathbb{C}$ ou $\mathbb{R}$ que l'intégrabilité de $|f|$ commence à jouer parce qu'il faut que f^+ et f^- soient séparément intégrables.

Nous allons prouver la formule (13.376) pour des fonctions de plus en plis générales. Pour la suite nous notons $\mu' = \varphi(\mu)$.

Pour $f = \mathbb{1}_B$, B mesurable Soit $B \in \mathcal{F}_2$. Nous avons $\mathbb{1}_B \circ \varphi = \mathbb{1}_{\varphi^{-1}(B)}$. Donc en utilisant le lemme 13.115 nous avons

$$\int_{S_2} \mathbb{1}_B d\mu' = \mu'(B) = \mu\big(\varphi^{-1}(B)\big) = \int_{S_1} \mathbb{1}_{\varphi^{-1}(B)} d\mu = \int_{S_1} (\mathbb{1}_B \circ \varphi) d\mu. \tag{13.377}$$

f est étagée positive La fonction f peut être écrite sous la forme

$$f = \sum_{k=1}^{n} a_k \mathbb{1}_{B_k} \tag{13.378}$$

avec $B_k \in \mathcal{F}_2$ et $a_k \in \mathbb{R}^+$. Nous avons alors, en utilisant la sous-additivité de l'intégrale du théorème 13.119(3),

$$\begin{aligned}
\int_{S_2} f d\mu' &= \sum_k a_k \int_{S_2} \mathbb{1}_{B_k} d\mu' && (13.379a)\\
&= \sum_k a_k \int_{S_1} (\mathbb{1}_{B_k} \circ \varphi) d\mu && (13.379b)\\
&= \int_{S_1} \Big(\sum_k a_k \mathbb{1}_{B_k} \Big) \circ \varphi d\mu && (13.379c)\\
&= \int_{S_1} (f \circ \varphi) d\mu. && (13.379d)
\end{aligned}$$

f à valeurs dans $\bar{\mathbb{R}}^+$ Vu que f est mesurable, par le théorème 13.107 il existe une suite croissante de fonctions étagées positives convergeant vers f. Soit donc cette suite, $f_n \colon S_2 \to \mathbb{R}^+$. Les fonction s$f_n \circ \varphi$ sont étagées et positives et nous avons aussi la limite ponctuelle et croissante $f_n \circ \varphi \to f \circ \varphi$ parce que φ est continue. Le théorème de la convergence monotone (théorème 13.129) permet d'écrire ceci :

$$\int_{S_2} f d\mu' = \lim \int_{S_2} f_n d\mu' = \lim \int_{S_1} (f_n \circ \varphi) d\mu = \int_{S_1} (f \circ \varphi) d\mu. \tag{13.380}$$

Pour $f \colon S_2 \to \bar{\mathbb{R}}$ ou $\mathbb{C}$ C'est maintenant que l'intégrabilité va jouer. Nous avons $|f| \circ \varphi = |f \circ \varphi|$, donc

$$\int_{S_2} |f| d\mu' = \int_{S_1} |f| \circ \varphi d\mu = \int_{S_1} |f \circ \varphi| d\mu, \tag{13.381}$$

ce qui montre que f est μ'-intégrable si et seulement si $f \circ \varphi$ est μ-intégrable.

De plus si $f = f^+ - f^-$ alors $f^+ \circ \varphi = (f \circ \varphi)^+$, $f^- \circ \varphi = (f \circ \varphi)^-$, et de façon similaire pour les parties imaginaires et réelles.

□

13.7 Mesure à densité

13.7.1 Théorème de Radon-Nikodym

Proposition 13.142 (Produit d'une mesure par une fonction)**.**
Si $(S, \mathcal{F}, m_1)$ est un espace mesuré, et si $f \colon S \to \mathbb{R}$ est intégrable, et si B est un ensemble mesurable,

nous définissons fm_1 par

$$m_2(B) = (fm_1)(B) = \int_B f(t)dm_1(t). \tag{13.382}$$

Cela est une mesure positive sur $(S, \mathcal{F})$.

Démonstration. D'abord pour l'ensemble vide : $m_2(\varnothing) = \int_\varnothing f dm_1 = 0$.

Si A_n sont des éléments disjoints de $\mathcal{F}$ tels que $\bigcup_n A_n \in \mathcal{F}$. Alors en utilisant la proposition 13.122, nous avons le calcul suivant :

$$m_2(\bigcup_n A_n) = \int_{\bigcup_n A_n} f(t)dm_1(t) = \sum_n \int_{A_n} f(t)dm_1(t) = \sum_n m_2(A_n). \tag{13.383}$$

□

Définition 13.143 ([174])**.**
*Soient μ et ν deux mesures sur l'espace mesurable $(\Omega, \mathcal{A})$. Nous disons que la mesure μ est **dominée** par ν si pour tout ensemble mesurable A, $\nu(A) = 0$ implique $\mu(A) = 0$.*

*Si ν est une mesure positive et μ une mesure, nous disons que μ est **absolument continue** par rapport à ν si $\nu(A) = 0$ implique $\mu(A) = 0$. On note aussi $\mu \ll \nu$.*

La mesure μ est **portée** par l'ensemble $E \in \mathcal{A}$ si pour tout $A \in \mathcal{A}$,

$$\mu(A) = \mu(A \cap E). \tag{13.384}$$

Nous écrivons que $\mu \perp \nu$ s'il existe un ensemble $E \in \mathcal{A}$ tel que μ soit porté par E et ν soit porté par $\complement E$.

Théorème 13.144 (Radon-Nikodym[175])**.**
Soient μ et ν deux mesures σ-finies sur un espace métrisable $(\Omega, \mathcal{A})$.

(1) Il existe un unique couple de mesures μ_1 et μ_2 telles que
   (a) $\mu = \mu_1 + \mu_2$
   (b) μ_1 est dominé par ν
   (c) $\mu_2 \perp \nu$.
   Dans ce cas, les mesures μ_1 et μ_2 sont positives et σ-finies.

(2) À égalité ν-presque partout près, il existe une unique fonction mesurable positive f telle que pour tout mesurable A,

$$\mu_1(A) = \int_A d\mu_1 = \int_\Omega \mathbb{1}_A f d\nu. \tag{13.385}$$

(3) À égalité ν-presque partout près, il existe une unique fonction positive mesurable h telle que $\mu_1 = h\nu$.

Corollaire 13.145.
Si μ es une mesure σ-finie dominée par la mesure σ-finie m, alors μ possède une unique fonction de densité.

Corollaire 13.146.
Soient μ et m, deux mesures positives σ-finies sur $(\Omega, \mathcal{A})$. Alors m domine μ si et seulement si μ possède une densité par rapport à m.

Démonstration. Si μ est dominée par m, alors la décomposition $\mu = \mu + 0$ satisfait le théorème de Radon-Nikodym. Par conséquent il existe une fonction f telle que

$$\mu(A) = \int_A f dm. \tag{13.386}$$

Cette fonction est alors une densité pour μ par rapport à m.

Pour la réciproque, nous supposons que μ a une densité f par rapport à m, et que A est une ensemble de m-mesure nulle :

$$m(A) = \int_\Omega \mathbb{1}_A dm = 0. \tag{13.387}$$

Cela signifie que la fonction $\mathbb{1}_A$ est m-presque partout nulle. La fonction produit $\mathbb{1}_A f$ est également nulle m-presque partout, et par conséquent

$$\mu(A) = \int_\Omega \mathbb{1}_A f dm = 0. \tag{13.388}$$

□

Problèmes et choses à faire

Est-ce que la démonstration de cela ne demande pas la convergence monotone d'une façon ou d'une autre ?

13.7.2 Mesure complexe

Définition 13.147 (Mesure complexe[176])**.**
*Si $(\Omega, \mathcal{A})$ est un espace mesurable, une **mesure complexe** est une application $\mu\colon \mathcal{A} \to \mathbb{C}$ telle que*

(1) $\mu(\varnothing) = 0$,

(2) ν est sous-additive : si les ensembles $A_i \in \mathcal{A}$, alors $\sum_i \mu(A_i) = \mu(\bigcup_i A_i)$.

Notons que la série $\sum_i \mu(A_i)$ est alors nécessairement absolument convergente. En effet changer l'ordre de la somme ne change pas l'union, et donc ne change pas la valeur de la somme. Si $\sigma\colon \mathbb{N} \to \mathbb{N}$ est une permutation,

$$\sum_i \mu(A_{\sigma(i)}) = \mu\Big(\bigcup_i A_{\sigma(i)}\Big) = \mu\Big(\bigcup_i A_i\Big) = \sum_i \mu(A_i). \tag{13.389}$$

Le théorème 9.176 dit alors que la somme doit être absolument convergente.

Théorème 13.148 (Radon-NikoDym complexe [29])**.**
Soit μ une mesure positive sur $(\Omega, \mathcal{A})$ et ν une mesure complexe. Alors

(1) Il existe un unique couple de mesures complexes ν_a, ν_s sur $(\Omega, \mathcal{A})$ tel que

(a) $\nu = \nu_a + \nu_s$

(b) $\nu_a \ll \mu$

(c) $\nu_s \perp \mu$.

(2) Ces mesures satisfont alors $\nu_a \perp \nu_s$.

(3) Il existe une fonction intégrable $h\colon \Omega \to \mathbb{C}$ telle que $\nu_a = h\mu$.

(4) La fonction h est unique à μ-équivalence près.

(5) Si de plus $\nu \ll \mu$ alors $\nu = h\mu$.

Démonstration. No proof. □

Remarque 13.149.
Le point (5) est souvent utilisé sous la forme

$$\nu(A) = \int_\Omega \mathbb{1}_A(\omega) h(\omega) d\mu(\omega) = \int_A h(\omega) d\mu(\omega). \tag{13.390}$$

29. L'histoire du nom de ce théorème est intéressante. Lorsque monsieur et madame Rèmederdonnukodym apprirent que leurs amis, les Rèmedelaboulechevelue avaient appelé leur fils Théo, ils décidèrent d'en faire autant. C'est en souvenir de ces circonstances que monsieur Nikodym (prénommé Radon) décida de faire des math.

13.7.3 Théorème d'approximation

Théorème 13.150 (Théorème d'approximation[170])**.**
Soit $(X, \mathcal{B}, \mu)$ un espace mesuré où $\mathcal{B}$ sont les boréliens de X. Soit $A \in \mathcal{B}$ tel que $A \subset W$ où W est un ouvert avec $\mu(W) < \infty$. Soit aussi $\epsilon > 0$.

(1) Il existe un fermé F et un ouvert V tels que $\mu(V) < \infty$ et

$$F \subset A \subset V \tag{13.391}$$

et $\mu(V \backslash F) < \epsilon$.

(2) Il existe $f \in C^0(X, \mathbb{R})$ nulle hors de W vérifiant $0 \leqslant f \leqslant 1$ et

$$\int_X |\mathbb{1}_A - f|^p d\mu(x) < \epsilon. \tag{13.392}$$

13.8 Tribu produit, mesure produit

13.8.1 Produit d'espaces mesurables

Définition 13.151.
*Si $\mathcal{A}_1$ et $\mathcal{A}_2$ sont deux tribus sur deux ensembles Ω_1 et Ω_2, nous définissons la **tribu produit** $\mathcal{A}_1 \otimes \mathcal{A}_2$ comme étant la tribu engendrée par*

$$\{X \times Y \text{ tel que } X \in \mathcal{A}_1, Y \in \mathcal{A}_2\}. \tag{13.393}$$

*Ces ensembles sont appelés **rectangles** de $(\Omega_1, \mathcal{A}_1) \otimes (\Omega_2, \mathcal{A}_2)$.*

Proposition 13.152 ([177])**.**
Soient deux espaces mesurables $(S_1, \mathcal{F}_1)$ et $(S_2, \mathcal{F}_2)$. Si $\mathcal{C}_i$ est une classe de parties de S_i avec $\mathcal{F}_i = \sigma(\mathcal{C}_i)$ et $S_i \in \mathcal{C}_i$. Alors

$$\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(\mathcal{C}_1 \times \mathcal{C}_2). \tag{13.394}$$

Démonstration. Nous notons p_1 et p_2 les projections de $S_1 \times S_2$ vers S_1 et S_2. Nous commençons par prouver que

$$\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma\big(\mathrm{P}_1^{-1}(\mathcal{F}_1) \cup p_2^{-1}(\mathcal{F}_2)\big). \tag{13.395}$$

En effet cette union est dans $\mathcal{F}_1 \otimes \mathcal{F}_2$ parce que ce sont tous des produits de la forme $A_1 \times S_2$ et $S_1 \times A_2$ où $A_i \in \mathcal{F}_i$. Inversement, tous les produits de la forme $A_1 \times A_2$ sont dans la tribu engendrée par l'union parce que

$$A_1 \cup A_2 = (A_1 \times S_2) \cap (S_1 \times A_2). \tag{13.396}$$

Par conséquent, la partie $p_1^{-1}(\mathcal{F}_1) \cup p_2^{-1}(\mathcal{F}_2)$ engendre tous les produits qui engendrent la tribu $\mathcal{F}_1 \otimes \mathcal{F}_2$. L'égalité (13.395) est donc correcte.

Si $C_1 \in \mathcal{C}_1$ alors

$$p_1^{-1}(C_1) = C_1 \times S_2 \in \mathcal{C}_1 \times \mathcal{C}_2 \tag{13.397}$$

et donc $p_1^{-1}(\mathcal{C}_1) \subset \mathcal{C}_1 \times \mathcal{C}_2$. En utilisant le lemme de transfert 13.70 nous avons alors

$$p_1^{-1}(\mathcal{F}_1) = p_1^{-1}\big(\sigma(\mathcal{C}_1)\big) = \sigma\big(p_1^{-1}\mathcal{C}_1\big) \subset \sigma(\mathcal{C}_1 \times \mathcal{C}_1) \tag{13.398}$$

et au bout de la même façon,

$$p_2^{-1}(\mathcal{F}_1) \subset \sigma(\mathcal{C}_1 \times \mathcal{C}_2). \tag{13.399}$$

Vu les relations (13.398), (13.399) et (13.395) nous avons

$$\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma\big(\mathrm{P}_1^{-1}(\mathcal{F}_1) \cup p_2^{-1}(\mathcal{F}_2)\big) \subset \sigma(\mathcal{C}_1 \times \mathcal{C}_2). \tag{13.400}$$

Réciproquement, si $C_1 \in \mathcal{C}_1$ et $C_2 \in \mathcal{C}_2$ alors

$$C_1 \times C_2 = (C_1 \times S_1) \cap (S_1 \times C_2) = p_1^{-1}(C_1) \cap p_2^{-1}(C_2) \in \mathcal{F}_1 \otimes \mathcal{F}_2. \tag{13.401}$$

□

13.8.2 Le cas des boréliens

Si X_1 et X_2 sont des espaces topologiques et si nous notons $\mathcal{O}_i$ l'ensemble de leurs ouverts, par définition $\mathcal{B}\text{or}(X_i) = \sigma(\mathcal{O}_i)$. De plus par la proposition 13.152 nous savons que

$$\sigma(\mathcal{O}_1 \times \mathcal{O}_2) = \mathcal{B}\text{or}(X_1) \otimes \mathcal{B}\text{or}(X_2). \tag{13.402}$$

Lemme 13.153.
Si $(X_i, \mathcal{O}_i)$ sont des espaces topologiques, alors

$$\mathcal{B}or(X_1) \otimes \mathcal{B}or(X_2) \subset \mathcal{B}or(X_1 \times X_2) \tag{13.403}$$

Démonstration. Si $A_i \in \mathcal{O}_i$ alors $A_1 \times A_2$ est un ouvert de $X_1 \times X_2$ (voir la définition 6.5). Par conséquent, $\mathcal{O}_1 \times \mathcal{O}_2$ est contenu dans l'ensemble des ouverts de $X_1 \times X_2$ ou encore

$$\mathcal{O}_1 \times \mathcal{O}_2 \subset \mathcal{B}\text{or}(X_1 \times X_2), \tag{13.404}$$

et donc

$$\sigma(\mathcal{O}_1 \times \mathcal{O}_2) \subset \sigma\big(\mathcal{B}\text{or}(X_1 \times X_2)\big) \tag{13.405}$$

finalement, par (13.402)

$$\mathcal{B}\text{or}(X_1) \otimes \mathcal{B}\text{or}(X_2) \subset \mathcal{B}\text{or}(X_1 \times X_2). \tag{13.406}$$

□

Il n'y a en général pas égalité, mais nous allons immédiatement voir que dans (presque) tous les cas raisonnables, les boréliens sur un produit sont le produit des boréliens.

Proposition 13.154 ([177])**.**
Soient (X_1, d_1) et (X_2, d_2) des espace métriques séparables. Alors

$$\mathcal{B}or(X_1 \times X_2) = \mathcal{B}or(X_1) \otimes \mathcal{B}or(X_2). \tag{13.407}$$

Démonstration. Nous savons par le lemme 6.62 que tout ouvert de $X_1 \times X_2$ est une réunion dénombrable d'éléments de $\mathcal{O}_1 \times \mathcal{O}_2$. Donc tout ouvert de $X_1 \times X_2$ est dans $\mathcal{B}\text{or}(X_1) \otimes \mathcal{B}\text{or}(X_2)$. Par conséquent

$$\mathcal{B}\text{or}(X_1 \times X_2) \subset \mathcal{B}\text{or}(X_1) \otimes \mathcal{B}\text{or}(X_2). \tag{13.408}$$

L'inclusion inverse étant déjà acquise par le lemme 13.153, nous avons l'égalité. □

Proposition 13.155.
Les boréliens sur $\mathbb{R}^N$ sont ceux qu'on croit.

(1) $\mathcal{B}or(\mathbb{R}^2) = \mathcal{B}or(\mathbb{R}) \otimes \mathcal{B}or(\mathbb{R})$

(2) $\mathcal{B}or(\mathbb{R}^{N+1}) = \mathcal{B}or(\mathbb{R}^N) \otimes \mathcal{B}or(\mathbb{R})$

Démonstration. Cela n'est rien d'autre que la proposition 13.154. □

Proposition 13.156.
Soit un espace mesurable $(S, \mathcal{F})$ et des applications $f_k \colon S \to \mathbb{R}$ $(k = 1, \dots, N)$. Alors l'application

$$\begin{aligned} f \colon (S, \mathcal{F}) &\to (\mathbb{R}^n, \mathcal{B}or(\mathbb{R}^N)) \\ x &\mapsto \big(f_1(x), \dots, f_N(x)\big) \end{aligned} \tag{13.409}$$

est mesurable si et seulement si chacun des f_i est mesurable.

Démonstration. Division en deux.

Condition nécessaire Nous supposons que les f_i sont mesurables. Nous avons

$$f^{-1}\big(\prod_{k=1}^{N}]a_k,b_k[\big)=\{x\in S \text{ tel que } f_1(x)\in]a_1,b_1[,\cdots f_N(x)\in]a_N,b_N[\} \tag{13.410a}$$

$$=\bigcap_{k=1}^{N}f_k^{-1}\big(]a_k,b_k[\big). \tag{13.410b}$$

Cela est une intersection finie d'éléments de $\mathcal{F}$ et est donc un élément de $\mathcal{F}$. Mais les pavés ouverts engendrent $\mathcal{B}\text{or}(\mathbb{R}^N)$ parce qu'ils sont une base dénombrable de la topologie (proposition 13.9). Le théorème 13.71 nous assure alors que f est mesurable parce que l'image inverse d'une base de la tribu est mesurable.

Condition suffisante Si f est mesurable alors en particulier

$$f_k^{-1}\big(]a,b[\big)=f^{-1}\big(\mathbb{R}\times\ldots\times]a,b[\times\mathbb{R}\times\ldots\times\mathbb{R}\big)\in\mathcal{F}. \tag{13.411}$$

Pour cela nous avons utilisé la proposition 13.155 qui nous indique que le produit dans la parenthèse est un borélien de $\mathbb{R}^N$ en tant que produit de boréliens de $\mathbb{R}$.
Encore une fois f_k^{-1} tombe dans $\mathcal{F}$ pour une base dénombrable de la topologie de $\mathbb{R}$ et est donc mesurable.

□

13.8.3 Produit de mesures

Lemme 13.157 (Propriété des sections[167])**.**
Soient $\mathcal{A}_1$ et $\mathcal{A}_2$ des tribus sur les ensembles Ω_1 et Ω_2. Si $A\in\mathcal{A}_1\otimes\mathcal{A}_2$ alors pour tout $x\in\Omega_1$ et $y\in\Omega_2$, les ensembles

$$A_1(y)=\{x\in\Omega_1 \textit{ tel que } (x,y)\in A\} \tag{13.412a}$$

$$A_2(x)=\{y\in\Omega_2 \textit{ tel que } (x,y)\in A\} \tag{13.412b}$$

sont mesurables.

Démonstration. Soit $y\in\Omega_2$; nous allons prouver le résultat pour $A_1(y)$. Pour cela nous notons

$$S=\{A\in\mathcal{A}_1\otimes\mathcal{A}_2 \text{ tel que } \forall y\in\Omega_2, A_1(y)\in\mathcal{A}_1\}, \tag{13.413}$$

et nous allons noter que S est une tribu contenant les rectangles. Par conséquent, S sera égal à $\mathcal{A}_1\otimes\mathcal{A}_2$.

Les rectangles Considérons le rectangle $A=X\times Y$ et si $y\in\Omega_2$ alors

$$A_1(y)=\{x\in\Omega_1 \text{ tel que } (x,y)\in X\times Y\}. \tag{13.414}$$

Donc soit $y\in Y$ alors $A_1(y)=X\in\mathcal{A}_1$, soit $y\notin Y$ et alors $A_1(y)=\varnothing\in\mathcal{A}_1$.

Tribu : ensemble complet Nous avons $\Omega_1\times\Omega_2\in S$ parce que c'est un rectangle.

Tribu : complémentaire Soit $A\in S$ et montrons que $A^c\in S$. Nous avons d'abord

$$(A^c)_1(y)=\{x\in\Omega_1 \text{ tel que } (x,y)\in A^c\}. \tag{13.415}$$

D'autre part

$$A_1(y)^c=\{x\in\Omega_1 \text{ tel que } (x,y)\notin A\}=\{x\in\Omega_1 \text{ tel que } (x,y)\in A^c\}=(A^c)_1(y). \tag{13.416}$$

Vu que $\mathcal{A}_1$ est une tribu et que par hypothèse $A_1(y)\in\mathcal{A}_1$, nous avons aussi $A_1(y)^c\in S$, et donc $(A^c)_1(y)\in\mathcal{A}_1$, ce qui prouve que $A^c\in S$.

Tribu : union dénombrable Soit une suite $A_n \in S$. Nous avons

$$(\bigcup_n A_n)_1(y) = \{x \in \Omega_1 \text{ tel que } (x,y) \in \bigcup_n A_n\} = \bigcup_n \{x \in \Omega_1 \text{ tel que } (x,y) \in A_n\} = \bigcup_n (A_n)_1(y), \tag{13.417}$$

et ce dernier ensemble est dans $\mathcal{A}_1$ parce que c'est une union dénombrable d'éléments de $\mathcal{A}_1$.

Nous avons donc prouvé que S est une tribu contenant les rectangles, donc S contient au moins $\mathcal{A}_1 \otimes \mathcal{A}_2$. □

Corollaire 13.158.
Si $f\colon \Omega_1 \times \Omega_2 \to \mathbb{R}$ est une fonction mesurable[30] *sur $X \times Y$ alors pour chaque y dans Ω_2, la fonction*

$$\begin{aligned} f_y\colon X &\to \mathbb{R} \\ x &\mapsto f(x,y) \end{aligned} \tag{13.418}$$

est mesurable.

Démonstration. Soit $\mathcal{O}$ un ensemble mesurable de $\mathbb{R}$ (i.e. un borélien), et $y \in \Omega_2$. Nous avons

$$f_y^{-1}(\mathcal{O}) = \{x \in X \text{ tel que } f(x,y) \in \mathcal{O}\} = A_1(y) \tag{13.419}$$

où

$$A = \{(x,y) \in \Omega_1 \times \Omega_2 \text{ tel que } f(x,y) \in \mathcal{O}\} = f^{-1}(\mathcal{O}). \tag{13.420}$$

Ce dernier est mesurable parce que f l'est. □

Théorème 13.159 ([167][31])**.**
Soient $(\Omega_i, \mathcal{A}_i, \mu_i)$ $(i = 1, 2)$ deux espaces mesurés σ-finie. Soit $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$. Alors les fonctions[32]

$$x \mapsto \mu_2\big(A_2(x)\big) \tag{13.421a}$$
$$y \mapsto \mu_1\big(A_1(y)\big) \tag{13.421b}$$

sont mesurables et

$$\int_{\Omega_1} \mu_2\big(A_2(x)\big) d\mu_1(x) = \int_{\Omega_2} \mu_2\big(A_1(y)\big) d\mu_2(y). \tag{13.422}$$

Démonstration. Nous supposons d'abord que μ_1 et μ_2 sont finies et nous notons $\mathcal{D}$ le sous-ensemble de $\mathcal{A}_1 \otimes \mathcal{A}_2$ sur lequel le théorème est correct. Nous allons commencer par prouver que $\mathcal{D}$ est un λ-système.

λ-système : différence ensembliste Soient $A, B \in \mathcal{D}$ avec $A \subset B$. Nous avons

$$(B\backslash A)_1(y) = \{x \in \Omega_1 \text{ tel que } (x,y) \in B\backslash A\} \tag{13.423a}$$
$$= \{x \in \Omega_1 \text{ tel que } (x,y) \in B\}\backslash\{x \in \Omega_1 \text{ tel que } (x,y) \in A\} \tag{13.423b}$$
$$= B_1(y)\backslash A_1(y). \tag{13.423c}$$

Vu que $A_1(y) \subset B_1(y)$ et que les mesure sont finies le lemme 13.26 nous donne

$$\mu_1\big((B\backslash A)_1(y)\big) = \mu_1\big(B_1(y)\big) - \mu_1\big(A_1(y)\big), \tag{13.424}$$

et similairement pour $1 \leftrightarrow 2$. Les deux fonctions (de y) à droite étant mesurables, nous avons la mesurabilité de la fonction $y \mapsto \mu_1\big((B\backslash A)_1(y)\big)$.

30. Définition 13.63.
31. Modèle non contractuel : des notations et la définition de λ-système peuvent varier entre la référence et le présent texte.
32. Voir la notation du lemme 13.412.

Prouvons la formule intégrale en nous rappelant que la formule (13.422) est supposée correcte pour A et B séparément :

$$\int_{\Omega_2} \mu_1\big((B\backslash A)_1(y)\big)d\mu_2(y) = \int_{\Omega_2} \mu_1\big(B_1(y)\big)d\mu_2(y) - \int_{\Omega_2} \mu_1\big(A_1(y)\big)d\mu_2(y) \tag{13.425a}$$

$$= \int_{\Omega_1} \mu_2\big(B_2(x)\big)d\mu_1(x) - \int_{\Omega_1} \mu_2\big(A_2(x)\big)d\mu_1(x) \tag{13.425b}$$

$$= \int_{\Omega_1} \mu_2\big((B\backslash A)_2(x)\big)d\mu_1(x). \tag{13.425c}$$

λ-système : limite de suite croissante Soit (A_n) une suite croissante dans $\mathcal{D}$; nous posons $B_n = A_n\backslash A_{n-1}$ et $A_0 = \varnothing$ de telle sorte à travailler avec une suite d'ensembles disjoints qui satisfait $\bigcup_n A_n = \bigcup_n B_n$. Vu que la suite est croissante nous avons $A_{n-1} \subset A_n$ et donc $B_n \in \mathcal{D}$ par le point déjà fait sur la différence ensembliste. Nous avons :

$$\mu_1\big((\bigcup_n B_n)_1(y)\big) = \{x \in \Omega_1 \text{ tel que } (x,y) \in \bigcup_n B_n\} \tag{13.426a}$$

$$= \bigcup_n \{x \in \Omega_1 \text{ tel que } (x,y) \in B_n\} \tag{13.426b}$$

$$= \bigcup_n (B_n)_1(y). \tag{13.426c}$$

Par conséquent, par la propriété (2) d'une mesure nous avons

$$\mu_1\big((\bigcup_n B_n)_1(y)\big) = \sum_n \mu_1\big((B_n)_1(y)\big). \tag{13.427}$$

En tant que somme de fonctions positives et mesurables, la fonction

$$y \mapsto \sum_n \mu_1\big((B_n)_1(y)\big) \tag{13.428}$$

est mesurable par la proposition 13.93. Il faut encore vérifier la formule intégrale. Le gros du boulot est de permuter une somme et une intégrale par le corollaire 13.132 :

$$\int_{\Omega_2} \sum_n \mu_1\big((B_n)_1(y)\big)d\mu_2(y) = \sum_n \int_{\Omega_2} \mu_1\big((B_n)_1(y)\big)d\mu_2(y) \tag{13.429a}$$

$$= \sum_n \int_{\Omega_1} \mu_2\big((B_n)_2(x)\big)d\mu_1(x) \tag{13.429b}$$

$$= \int_{\Omega_1} \sum_n \mu_2\big((B_n)_2(x)\big)d\mu_1(x) \tag{13.429c}$$

$$= \int_{\Omega_1} \mu_2\big((\bigcup_n B_n)_1(y)\big)d\mu_1(x). \tag{13.429d}$$

Maintenant que $\mathcal{D}$ est un λ-système contenant les rectangles, le lemme 13.39 dit que la tribu engendrée par $\mathcal{D}$ (c'est à dire $\mathcal{A}_1 \otimes \mathcal{A}_2$) est le λ-système $\mathcal{D}$ lui-même.

La preuve est finie dans le cas de mesures finies. Nous commençons maintenant à prouver dans le cas où les mesures μ_1 et μ_2 sont seulement σ-finies. Nous considérons des suites croissantes $\Omega_{i,n} \to \Omega_i$ d'ensembles mesurables et de mesure finie : $\mu_i(\Omega_{i,n}) < \infty$. D'abord remarquons que

$$\mu_2\Big((A \cap \Omega_{1,j} \times E_{2,j})_2(x)\Big) = \mu_2\Big(A_2(x) \cap \Omega_{2,j}\Big)\mathbb{1}_{\Omega_{1,j}}. \tag{13.430}$$

En effet,

$$\heartsuit = (A \cap \Omega_{1,j} \times E_{2,j})_2(x) \tag{13.431a}$$

$$= \{y \in \Omega_2 \text{ tel que } (x,y) \in A \cap \Omega_{1,j} \times E_{2,j}\} \tag{13.431b}$$

$$= \{y \in \Omega_2 \text{ tel que } (x,y) \in A \times \Omega_{2,j}\} \cap \{y \in \Omega_2 \text{ tel que } (x,y) \in \Omega_{1,j} \times \Omega_{2,j}\}. \tag{13.431c}$$

Si $y \in \Omega_{1,j}$ alors $\{y \in \Omega_2 \text{ tel que } (x,y) \in \Omega_{1,j} \times \Omega_{2,j}\} = \Omega_{2,j}$ et dans ce cas

$$\heartsuit = \{y \in \Omega_2 \text{ tel que } (x,y) \in A \times \Omega_{2,j}\} \cap \Omega_{2,j} = A_2(x) \cap E_{2,j}. \tag{13.432}$$

Et inversement, si $x \notin \Omega_{1,j}$ alors $\heartsuit = \varnothing$. Dans les deux cas nous avons (13.430).

Les ensembles $A \cap \Omega_{1,j} \times \Omega_{2,j}$ étant de mesure finie, nous pouvons leur appliquer la première partie :

$$\int_{\Omega_1} \mu_2\Big((A \cap \Omega_{1,j} \times \Omega_{2,j})_2(x)\Big) d\mu_1(x) = \int_{\Omega_2} \mu_1\Big((A \cap \Omega_{1,j} \times \Omega_{2,j})_1(y)\Big) d\mu_2(u), \tag{13.433}$$

ou encore

$$\int_{\Omega_1} \mu_2\Big(A_2(x) \cap \Omega_{2,j}\Big) \mathbb{1}_{\Omega_{1,j}}(x) d\mu_1(x) = \int_{\Omega_2} \mu_1\Big(A_1(y) \cap \Omega_{1,j}\Big) \mathbb{1}_{\Omega_{2,j}}(y) d\mu_2(y). \tag{13.434}$$

Ce que nous avons dans ces intégrales sont (par rapport à j) des suites croissantes de fonction positives ; nous pouvons donc permuter une limite et une intégrale. En sachant que si $k \to \infty$, alors

$$\mathbb{1}_{1,j}(x) \to 1 \tag{13.435a}$$
$$\mu_2\big(A_2(x) \cap \Omega_2, j\big) \to \mu_2\big(A_2(x)\big), \tag{13.435b}$$

nous trouvons le résultat demandé. □

Théorème 13.160 ([178, 179])**.**
Soient μ_i des mesures σ-finies sur $(\Omega_i, \mathcal{A}_i)$ $(i = 1, 2)$.

(1) Il existe une et une seule mesure, notée $\mu_1 \otimes \mu_2$, sur $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$ telle que

$$(\mu_1 \otimes \mu_2)(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2) \tag{13.436}$$

pour tout $A_1 \in \mathcal{A}_1$ et $A_2 \in \mathcal{A}_2$.

(2) Cette mesure est donnée par la formule[33]

$$(\mu_1 \otimes \mu_2)(A) = \int_{\Omega_1} \mu_2\big(A_2(x)\big) d\mu_1(x) = \int_{\Omega_2} \mu_1\big(A_1(y)\big) d\mu_2(y). \tag{13.437}$$

*Cette mesure est la **mesure produit** de μ_1 par μ_2.*

(3) La mesure $\mu_1 \otimes \mu_2$ ainsi définie est σ-finie.

Démonstration. La partie «existence» sera divisée en deux parties : l'une pour prouver que les formules (13.437) donnent une mesure et une pour montrer que cette mesure vérifie la condition (13.436).

Unicité L'ensemble des rectangles de $\Omega_1 \times \Omega_2$ engendre la tribu $\mathcal{A}_1 \otimes \mathcal{A}_2$, est fermé par intersection et contient une suite croissante d'ensembles $P_n \times R_n$ de mesure finie ($\mu(P_n \times R_n) < \infty$) telle que $P_n \times R_n \to \Omega_1 \times \Omega_2$. Cette suite est donné par le fait que μ_1 et μ_2 sont σ-finies. En effet si (X_n) et (Y_n) sont des recouvrements dénombrables de Ω_1 et Ω_2 par des ensembles de mesure finie, en posant $P_n = \bigcup_{k=1}^n X_n$ et $R_n = \bigcup_{k=1}^n Y_n$ nous avons bien une suite croissante de rectangles qui tendent vers $\Omega_1 \times \Omega_2$. Avec ces rectangles en main, le théorème 13.40 donne l'unicité.

Les formules définissent une mesure Le théorème 13.159 dit que ces formules ont un sens et que l'égalité entre les deux intégrales est correcte. Nous prouvons à présent qu'elles déterminent effectivement une mesure sur $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$.

Pour tout $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$, $\mu(A) \geqslant 0$ parce que μ est donnée par l'intégrale d'une fonction positive.

33. Voir les notations du lemme 13.157.

En ce qui concerne la condition d'unions dénombrable disjointe, soient $A^{(i)}$ des éléments disjoints de $\mathcal{A}_1 \otimes \mathcal{A}_2$; nous commençons par remarquer que

$$\left(\bigcup_{i=1}^{\infty} A^{(i)}\right)_2 (x) = \{y \in \Omega_2 \text{ tel que } (x,y) \in \bigcup_{i=1}^{\infty} A^{(i)}\} \tag{13.438a}$$

$$= \bigcup_{i=1}^{\infty} \{y \in \Omega_2 \text{ tel que } (x,y) \in A^{(i)}\} \tag{13.438b}$$

$$= \bigcup_{i=1}^{\infty} A_2^{(i)}(x). \tag{13.438c}$$

Par conséquent,

$$\mu\left(\bigcup_{i=1}^{\infty} A^{(i)}\right) = \int_{\Omega_1} \mu_2\left(\left(\bigcup_{i=1}^{\infty} A^{(i)}\right)_2(x)\right) d\mu_1(x) \tag{13.439a}$$

$$= \int_{\Omega_1} \sum_{i=1}^{\infty} \mu_2\big(A_2^{(i)}(x)\big) d\mu_1(x) \tag{13.439b}$$

$$= \int_{\Omega_1} \lim_{n\to\infty} \sum_{i=1}^{n} \mu_2\big(A_2^{(i)}(x)\big) d\mu_1(x). \tag{13.439c}$$

où nous avons utilisé l'additivité de la mesure μ_2. À ce niveau, il serait commode de permuter la somme et l'intégrale. Pour ce faire nous considérons la suite (croissante) de fonctions

$$f_n(x) = \sum_{i=1}^{n} \mu_2\big(A_2^{(i)}(x)\big). \tag{13.440}$$

Nous pouvons permuter la limite et l'intégrale grâce au théorème de la convergence monotone 13.129 ; ensuite la somme se permute avec l'intégrale en tant que somme finie :

$$\mu\left(\bigcup_{i=1}^{\infty} A^{(i)}\right) = \lim_{n\to\infty} \sum_{i=1}^{n} \int_{\Omega_1} \big(A_2^{(i)}(x)\big) d\mu_1(x) \tag{13.441a}$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} \mu(A^{(i)}) \tag{13.441b}$$

$$= \sum_{i=1}^{\infty} \mu(A^{(i)}). \tag{13.441c}$$

Elles vérifient la condition Prouvons que les formules (13.437) se réduisent à (13.436) dans le cas des rectangles. Soit donc $A = X_1 \times X_2$ avec $X_i \in \mathcal{A}_i$. Alors

$$A_1(y) = \{x \in \Omega_1 \text{ tel que } (x,y) \in X_1 \times X_2\} \tag{13.442}$$

et

$$\mu_1\big(A_1(y)\big) = \mathbb{1}_{X_2}(y)\mu_1(X_1), \tag{13.443}$$

donc

$$(\mu_1 \otimes \mu_2)(A) = \int_{\Omega_2} \mu_1\big(A_1(y)\big) d\mu_2(y) \tag{13.444a}$$

$$= \int_{\Omega_2} \mu_1(X_1)\mathbb{1}_{X_2}(y) d\mu_2(y) \tag{13.444b}$$

$$= \mu_1(X_1) \int_{\Omega_2} \mathbb{1}_{X_2}(y) d\mu_2(y) \tag{13.444c}$$

$$= \mu_1(X_1)\mu_2(X_2). \tag{13.444d}$$

Pour cela nous avons utilisé le fait que l'intégrale de la fonction caractéristique d'un ensemble mesurable est la mesure de cet ensemble.

□

Définition 13.161 (Produit d'espaces mesurés)**.**
*Si $(\Omega_i, \mathcal{A}_i, \mu_i)$ sont deux espaces mesurés, l'**espace produit** est l'ensemble $\Omega_1 \times \Omega_2$ muni de la tribu produit $\mathcal{A}_1 \otimes \mathcal{A}_2$ de la définition 13.151 et de la mesure produit $\mu_1 \otimes \mu_2$ définie par le théorème 13.160.*

Remarque 13.162.
Il n'est pas garantit que la tribu $\mathcal{A}_1 \otimes \mathcal{A}_2$ soit la tribu la plus adaptée à l'ensemble $S_1 \times S_2$. Dans le cas de $\mathbb{R}^N$, il se fait que c'est le cas : en prenant des produits des boréliens sur $\mathbb{R}$ on obtient bien les boréliens sur $\mathbb{R}^N$, voir proposition 13.155.

13.9 Mesure de Lebesgue sur $\mathbb{R}$

Nous notons $\mathcal{S}$ l'ensemble des intervalles [34] de $\mathbb{R}$.

Proposition 13.163.
L'ensemble réunions finies d'éléments de $\mathcal{S}$ est une algèbre de parties de $\mathbb{R}$ que nous allons noter $\mathcal{A}_{\mathcal{S}}$.

Démonstration. Nous devons vérifier la définition 13.18. Les ensembles $\mathbb{R}$ et $\varnothing$ sont des intervalles et font donc partie de $\mathcal{A}_{\mathcal{S}}$.

Si $A \in \mathcal{A}_{\mathcal{S}}$ se décompose en union d'intervalles de la forme (a_k, b_k) avec $k = 1, \dots, n$ (ici nous mettons des parenthèses au lieu de crochets parce qu'a priori nous ne savons pas). Alors

$$A^c = \bigcup_{k=0}^{k} (b_k, a_{k+1}) \tag{13.445}$$

où nous avons posé $b_0 = -\infty$ et $a_{n+1} = +\infty$. Ici encore les parenthèses sont soit fermées soit ouvertes en fonction de ce qu'étaient celles dans la décomposition de A. Quoi qu'il en soit, cette décomposition de A^c montre que $A^c \in \mathcal{A}_{\mathcal{S}}$.

Enfin si $A, B \in \mathcal{A}_{\mathcal{S}}$ alors $A \cup B \in \mathcal{A}_{\mathcal{S}}$. □

Lemme 13.164.
Tout élément de $\mathcal{A}_{\mathcal{S}}$ admet une décomposition minimale unique en réunion finie d'intervalles. Cette décomposition est formée d'intervalles deux à deux disjoints.

Démonstration. Nous allons montrer que si $A \in \mathcal{A}_{\mathcal{S}}$, alors la décomposition minimale consiste en les composantes connexes de A. Pour cela nous rappelons que la proposition 6.119 dit qu'une partie de $\mathbb{R}$ est connexe si et seulement si elle est un intervalle. D'abord cela nous dit immédiatement que les composantes connexes de A forment une décomposition de A en intervalles. Nous devons prouver qu'elle est minimale.

Soit $\{C_k\}_{k=1,\dots,n}$ les composantes connexes de A. Aucun connexe de $\mathbb{R}$ contenu dans A ne peut intersecter plus d'un des C_k, et par conséquent nous ne pouvons pas décomposer A en moins de n intervalles.

Pour l'unicité, soit $\{I_k\}_{k=1,\dots,n}$ un ensemble de n intervalles tels que $\bigcup_{k=1}^{n} I_k = A$. Chacun des I_k intersecte un et un seul des C_k. En effet si $x \in I_k \cap C_i$ et $y \in I_k \cap C_j$, alors $[x, y] \subset I_k$ parce que I_k est un intervalle. Mais C_i étant le plus grand connexe contenant x, $[x, y] \subset C_i$ et de la même façon, $[x, y] \subset C_j$. Par conséquent C_i et C_j sont tous deux la composante connexe de x et y. Nous en déduisons que $C_i = C_j$, c'est à dire $i = j$.

Par ailleurs nous avons $I_k \cap I_l = \varnothing$ dès que $k \neq l$ parce que sinon l'ensemble $I_k \cap I_l$ serait connexe et la décomposition des $\{I_k\}_{k=1,\dots,n}$ ne serait pas minimale : en remplaçant I_k et I_l par $I_k \cup I_l$ on aurait eu une décomposition contenant moins d'éléments. Donc à renumérotation près nous pouvons supposer que I_k intersecte C_l si et seulement si $k = l$.

34. Définition 2.8.

Dans ce cas nous devons avoir $I_k = C_k$, sinon les éléments de $C_k \backslash I_k$ ne seraient pas dans $\bigcup_{i=1}^{n} I_i$. □

Définition 13.165 (longueur d'intervalle[161])**.**
Si I est un intervalle d'extrémités a et b avec $-\infty \leqslant a \leqslant b \leqslant +\infty$ alors nous définissons la ***longueur*** *de I par*

$$\ell(I) = \begin{cases} b-a & si\ -\infty < a \leqslant b < +\infty \\ \infty & si\ a\ ou\ b\ est\ infini \end{cases} \tag{13.446}$$

Si $A \in \mathcal{A}_{\mathcal{S}}$ et si sa décomposition minimale est $A = \bigcup_{k=1}^{n} I_k$, alors on définit

$$\ell(A) = \sum_{k=1}^{n} \ell(I_k). \tag{13.447}$$

Le lemme suivant nous indique que nous pouvons calculer la longueur d'un élément de $\mathcal{A}_{\mathcal{S}}$ sans savoir la décomposition minimale, pourvu que l'on connaisse une décomposition disjointe.

Lemme 13.166 ([161])**.**
Si

$$B = \bigcup_{r=1}^{p} J_r \tag{13.448}$$

est une décomposition de $B \in \mathcal{A}_{\mathcal{S}}$ en intervalles deux à deux disjoints alors

$$\ell(B) = \sum_{r=1}^{p} \ell(J_r). \tag{13.449}$$

Démonstration. Nous prouvons dans un premier temps le résultat dans le cas où $B = I$ est un intervalle. Soit I un intervalle et une décomposition en intervalles disjoints $I = \bigcup_{r=1}^{p} J_r$. Nous montrons qu'alors $\ell(I) = \sum_{r=1}^{p} \ell(J_r)$. Nous verrons ensuite comment passer au cas où B est un élément générique de $\mathcal{A}_{\mathcal{S}}$.

Si $B = I$ est un intervalle infini Si I est infini alors un des J_r soit l'être et donc $\sum_{r=1}^{p} \ell(J_r) = \infty = \ell(I)$.

Si $B = I$ est un intervalle ininfini Pour chaque $r = 1, \ldots, p$ nous notons a_r et b_r les extrémités de J_r. Vu que les J_r sont connexes et disjoints, si $a_k \leqslant a_l$ alors $b_k \leqslant a_l$, sinon l'ensemble (non vide) $]a_l, b_k[$ serait dans l'intersection $I_k \cap I_l$ qui, elle, est vide. Plus généralement, si $x \in J_k$ et $y \in J_l$ avec $x < y$ alors pour tout $x' \in J_k$ et tout $y' \in J_l$ nous avons $x' < y'$. Vu qu'il y a un nombre fini d'ensembles J_r, nous pouvons les classer dans l'ordre croissant :

$$a_1 \leqslant b_1 \leqslant a_2 \leqslant b_2 \leqslant \ldots \leqslant b_{p-1} \leqslant a_p \leqslant b_p. \tag{13.450}$$

Vu que les J_r sont disjoints et que leur union est connexe nous avons en réalité

$$a = a_1 \leqslant b_1 = a_2 \leqslant b_2 = a_3 \leqslant \ldots \leqslant b_{p-1} = a_p \leqslant b_p, \tag{13.451}$$

donc une somme télescopique donne

$$\ell(I) = b - a = \sum_{r=1}^{p} (b_r - a_r) = \sum_{r=1}^{p} \ell(J_r). \tag{13.452}$$

Si B n'est pas un intervalle Soit $\{I_k\}_{k=1,\ldots,n}$ la décomposition minimale de B. Alors

$$\spadesuit = \ell(B) = \sum_{k=1}^{n} \ell(I_k) = \sum_{k=1}^{n} \ell\big(\bigcup_{r=1}^{p} (I_k \cap J_r)\big). \tag{13.453}$$

Mais I_k est un intervalle et s'écrit comme union disjointe $I_k = \bigcup_{r=1}^{p}(I_k \cap J_r)$, donc par la première partie

$$\spadesuit = \sum_{k=1}^{n}\sum_{r=1}^{p} \ell(I_k \cap J_r) = \sum_{r=1}^{p}\sum_{k=1}^{n} \ell(I_k \cap J_r). \tag{13.454}$$

Ici J_r est un intervalle qui se décompose en $J_r = \bigcup_{k=1}^{n}(I_k \cap J_r)$, donc nous pouvons encore utiliser la première partie :

$$\spadesuit = \sum_{r=1}^{p} \ell(J_r), \tag{13.455}$$

ce qu'il fallait.

□

Lemme 13.167.
Si $A, B \in \mathcal{A}_{\mathcal{S}}$ avec $A \subset B$ alors $\ell(A) \leqslant \ell(B)$.

Démonstration. Nous avons évidemment $B = A \cup B\backslash A$. Notons que $B\backslash A \in \mathcal{A}_{\mathcal{S}}$ par le lemme 13.19. Si $\{I_k\}$ est une décomposition disjointe de A et $\{J_i\}$ une de $B\backslash A$ alors $\{I_k\} \cup \{J_i\}$ est une décomposition disjointe de $A \cup B\backslash A$ et le lemme 13.166 nous dit que

$$\ell(B) = \ell(A \cup B\backslash A) = \ell(A) + \ell(B\backslash A). \tag{13.456}$$

Par conséquent $\ell(B) \geqslant \ell(A)$. □

Lemme 13.168.
Si I est un intervalle et s'il se décompose en

$$I = \bigcup_{n\in\mathbb{N}} I_n \tag{13.457}$$

où les I_n sont des intervalles disjoints, alors

$$\ell(I) = \sum_{n=1}^{\infty} \ell(I_n). \tag{13.458}$$

Démonstration. Nous allons encore diviser la preuve en deux parties suivant que I soit de longueur finie ou pas.

Si I est de longueur finie Soient a et b les extrémités de I : $-\infty < a \leqslant b < +\infty$. Pour tout $N \geqslant 1$ nous avons

$$\sum_{n=1}^{N} \ell(I_n) = \ell\Big(\bigcup_{n=1}^{n} I_n\Big) \leqslant \ell(I). \tag{13.459}$$

La première égalité est le lemme dans le cas d'une union finie 13.166. L'inégalité est le lemme 13.167. Cela étant vrai pour tout N, à la limite $N \to \infty$ nous conservons l'inégalité :

$$\sum_{n=1}^{\infty} \ell(I_n) \leqslant \ell(I). \tag{13.460}$$

Nous devons encore voir l'inégalité inverse. Pour cela nous supposons que $a < b$. Sinon $\ell(I) = 0$ et tous les I_n doivent être vide sauf un qui contiendra seulement $\{a\}$ (si I le contient).

Soit $\epsilon > 0$ avec $\epsilon < b - a$ et l'intervalle

$$[a + \frac{\epsilon}{4}, b - \frac{\epsilon}{4}] = [a', b'] \subset I. \tag{13.461}$$

Si les a_n et le b_n sont le extrémités des I_n alors

$$[a', b'] \subset I = \bigcup_{n\geqslant 1} I_n \subset \bigcup_{n\geqslant 1}]a_n - \frac{\epsilon}{2^{n+2}}, b_n + \frac{\epsilon}{2^{n+2}}[= \bigcup_{n\geqslant 1}]a'_n, b'_n[\tag{13.462}$$

où nous avons posé $a'_n = a_n - \epsilon/2^{n+2}$ et $b'_n = b_n + \epsilon/2^{n+2}$. Nous avons donc recouvert le compact [35] $[a', b']$ par des ouverts. Nous pouvons donc en extraire un sous-recouvrement fini (c'est la définition de la compacité), c'est à dire une partie finie F de $\mathbb{N}$ telle que

$$[a', b'] \subset \bigcup_{n\in F}]a'_n, b'_n[. \tag{13.463}$$

Le lemme 13.167 nous dit alors que

$$\heartsuit = b' - a' \leqslant \ell\Big(\bigcup_{n\in F}]a'_n, b'_n[\Big) \leqslant \sum_{n\in F} (b'_n - a'_n). \tag{13.464}$$

La seconde inégalité se prouve en recopiant [36] la preuve de 13.21. Nous continuons le calcul :

$$\heartsuit \leqslant \sum_{n\in F} (b_n - a_n) + \sum_{n\in F} \frac{\epsilon}{2^{n+1}} \leqslant \sum_{n\in F} (b_n - a_n) + \frac{\epsilon}{2}. \tag{13.465}$$

Mais $b' - a' = (b - a) - \frac{\epsilon}{2}$, donc

$$b - a - \frac{\epsilon}{2} \leqslant \sum_{n\in F} (b_n - a_n) + \frac{\epsilon}{2}. \tag{13.466}$$

D'où nous déduisons que

$$\ell(I) = b - a \leqslant \sum_{n\in F} (b_n - a_n) + \epsilon \leqslant \sum_{n\in \mathbb{N}} (b_n - a_n) + \epsilon = \sum_{n\in\mathbb{N}} \ell(I_n) + \epsilon. \tag{13.467}$$

Cela étant valable pour tout ϵ nous déduisons que

$$\ell(I) \leqslant \sum_{n\in\mathbb{N}} \ell(I_n). \tag{13.468}$$

Si I est de longueur infinie Étant donné que I est un intervalle de longueur infinie, il contient au moins un ensemble du type $]-\infty, a]$ ou $[a, +\infty[$; donc pour tout $M > 0$, il existe $N \geqslant 1$ tel que

$$\ell\big(I \cap [-N, N]\big) \geqslant M. \tag{13.469}$$

Mais $I \cap [-N, N]$ est un intervalle et

$$I \cap [-N, N] = \bigcup_{n\in\mathbb{N}} I_n \cap [-N, N] \tag{13.470}$$

qui est une union disjointe. Par conséquent,

$$M \leqslant \ell\big(I \cap [-N, N]\big) = \sum_n \ell\big(I_n \cap [-N, N]\big) \leqslant \sum_n \ell(I_n). \tag{13.471}$$

Cela étant vrai pour tout $M > 0$, nous concluons que

$$\sum_{n\in\mathbb{N}} \ell(I_n) = \infty. \tag{13.472}$$

□

Remarque 13.169.
Pour la preuve de 13.168 nous ne pouvons pas classer les I_n en ordre croissant comme nous l'avons fait dans la preuve de 13.166. En effet si $I = [0, 1]$ et que nous recouvrons $[0, \frac{1}{2}[$ et $]\frac{1}{2}, 1]$ par une infinité d'intervalles chacun, nous ne pouvons plus les classer par ordre croissant.

35. Lemme 6.121.

36. Nous ne pouvons pas invoquer directement le lemme 13.21 parce que nous n'avons pas encore prouvé que ℓ était une mesure sur $(\mathbb{R}, \mathcal{A}_\mathcal{S})$.

Proposition 13.170 ([161]).
La fonction ℓ ainsi définie est une mesure σ-finie sur l'algèbre de parties $\mathcal{A}_{\mathcal{S}}$.

Démonstration. Le fait que ℓ soit σ-finie provient par exemple du fait que $\ell\big(]-n,n[\big) = 2n$ tandis que $\bigcup_n]-n,n[= \mathbb{R}$.

Nous devons à présent prouver que ℓ est additive. Soient $(A_i)_{i\in\mathbb{N}}$ des éléments disjoints de $\mathcal{A}_{\mathcal{S}}$, avec leurs décomposition minimales

$$A_i = \bigcup_{k=1}^{n} I_k^{(i)}. \tag{13.473}$$

Pour chaque $i \in \mathbb{N}$, le lemme 13.168 nous indique que

$$\ell(A_i) = \sum_{k\in\mathbb{N}} \ell(I_k^{(i)}). \tag{13.474}$$

L'ensemble $\mathbb{N}\times\mathbb{N}$ est dénombrable et nous pouvons considérer la décomposition

$$\bigcup_{i\in\mathbb{N}} A_i = \bigcup_{(i,k)\in\mathbb{N}\times\mathbb{N}} I_k^{(i)}. \tag{13.475}$$

Cette décomposition n'est pas spécialement minimale[37] mais elle est disjointe. Le lemme 13.168 donne

$$\ell\Big(\bigcup_i A_i\Big) = \sum_{(i,k)\in\mathbb{N}\times\mathbb{N}} \ell(I_k^{(i)}) = \sum_{i\in\mathbb{N}} \left(\sum_{k\in\mathbb{N}} \ell(I_k^{(i)}) \right) = \sum_{i\in\mathbb{N}} \ell(A_i). \tag{13.476}$$

La décomposition de la somme sur $\mathbb{N}^2$ en deux sommes sur $\mathbb{N}$ est faite en vertu de la proposition 9.183.

□

13.9.1 Mesure et tribu de Lebesgue

Théorème 13.171.
Il existe une unique mesure λ sur $\big(\mathbb{R}, \mathcal{B}or(\mathbb{R})\big)$ telle que

$$\lambda\big(]a,b[\big) = b - a \tag{13.477}$$

pour tout $a \leqslant b$ dans $\mathbb{R}$.

Démonstration. L'existence provient du théorème de prolongement de Hahn 13.60 : la mesure ℓ sur $(\mathcal{A}_{\mathcal{S}})$ se prolonge à $\sigma(\mathcal{A}_{\mathcal{S}}) = \mathcal{B}\text{or}(\mathbb{R})$.

Nous ne pouvons pas prouver l'unicité en invoquant la partie unicité de Hahn (c'est tentant parce que ℓ est σ-finie) parce que dans ce théorème nous ne fixons la valeur de λ que sur une toute petite partie de $\mathcal{A}_{\mathcal{S}}$. Nous allons cependant voir que cette petite partie suffit à garantir l'unicité.

La classe

$$\mathcal{D} = \{]a,b[\text{ tel que } -\infty < a \leqslant b < +\infty\} \tag{13.478}$$

est stable par intersection finie et engendre la tribu borélienne. En effet $\mathcal{D}$ contient toutes les boules et donc une base dénombrable de la topologie de $\mathbb{R}$ (proposition 6.60). Donc tous les ouverts de $\mathbb{R}$ sont dans $\sigma(\mathcal{D})$ et $\sigma(\mathcal{D}) = \mathcal{B}\text{or}(\mathbb{R})$. Nous pouvons donc dire grâce au théorème 13.40 qu'il y a unicité de la mesure sur $\mathcal{B}\text{or}(\mathbb{R})$ lorsque les valeurs sur $\mathcal{D}$ sont fixées. □

Définition 13.172.
*La mesure de l'espace mesuré $\big(\mathbb{R}, \mathcal{B}or(\mathbb{R}), \lambda\big)$ donné par le théorème 13.171 est la **mesure de Lebesgue** sur $\big(\mathbb{R}, \mathcal{B}or(\mathbb{R})\big)$.*

*Nous définissons aussi la **tribu de Lebesgue** par la proposition 13.59 : $\big(\mathbb{R}, \mathcal{L}eb(\mathbb{R}), \lambda\big)$ est l'espace mesuré complété de $\big(\mathbb{R}, \mathcal{B}or(\mathbb{R}), \lambda\big)$.*

37. A_1 pourrait contenir $[0,1]$ et A_2 contenir $]1,2]$.

Remarque 13.173.
Il n'est pas évident que la tribu de Lebesgue soit plus grande que celle des boréliens, ni que la tribu des parties soit plus grande que celle de Lebesgue. Nous mentionnons cependant les faits suivants.

(1) Il existe des ensembles mesurables non-boréliens, et cela ne nécessite pas l'axiome du choix. Un argument classique de cardinalité est donné dans [160]. La construction la plus explicite que j'aie trouvée est dans [3], mais ça a l'air de demander des connaissances précises sur les ordinaux.
(2) Vu que l'ensemble de Cantor C est mesurable de mesure nulle, tout sous-ensemble de Cantor est mesurable de mesure nulle parce que la tribu de Lebesgue est complète par définition. Le cardinal de $\mathcal{P}(C)$ est strictement supérieur à la puissance du continu, alors que le cardinal de l'ensemble des boréliens est au plus égal à la puissance du continu. Donc il existe des non boréliens contenus dans Cantor ; de tels non boréliens sont alors mesurables au sens de Lebesgue.
(3) Si nous admettons l'axiome du choix alors il existe des ensembles non mesurables au sens de Lebesgue. Nous en verrons un dans l'exemple 13.190.

Exemple 13.174(Un ouvert contenant tous les rationnel et de mesure arbitrairement petite)
Il est possible de construire un ouvert de $\mathbb{R}$ contenant $\mathbb{Q}$ et de mesure de Lebesgue plus petite que ϵ. Pour cela si (q_i) est une énumération des rationnels, il suffit de prendre

$$\mathcal{O} = \bigcup_{n=1}^{\infty} B(q_n, \frac{\epsilon}{2^{n+1}}). \tag{13.479}$$

Cela est un ouvert comme union d'ouverts, ça contient tous les rationnels, et sa mesure se majore. En effet le théorème 13.171 donne $\lambda\big(B(q_n, \frac{\epsilon}{2^n})\big) = \frac{\epsilon}{2^n}$. Vu que ces boules ne sont a priori pas disjointes, le lemme 13.26 donne

$$\lambda(\mathcal{O}) \leqslant \sum_{n=1}^{\infty} \frac{\epsilon}{2^n} = \epsilon \tag{13.480}$$

par (10.130) avec $q = \frac{1}{2}$.

Par complémentarité, nous pouvons construire un ensemble fermé de mesure non nulle et ne contenant aucun rationnel. Et même un fermé dans $[0, 1]$, de mesure $1 - \epsilon$ ne contenant aucun rationnel.

Cela peut surprendre parce qu'il existe des tonnes de suites d'irrationnels qui convergent vers des rationnels [38], et il semble difficile de créer un ensemble contenant beaucoup d'irrationnels tout en préservant la propriété de fermeture vis à vis des suites convergentes. △

13.9.2 Propriétés de la mesure de Lebesgue

Proposition 13.175.
Tout ensemble dénombrable de $\mathbb{R}$ est mesurable de mesure nulle.

Démonstration. Un point de $\mathbb{R}$ est un intervalle de mesure nulle. Si D est dénombrable, il est union disjointes et dénombrable de points. Le lemme 13.168 nous dit alors que sa mesure est $\lambda(D) = \sum_{i=1}^{\infty} \lambda(\{a_i\}) = 0$. □

Remarque 13.176.
Il existe cependant des ensembles non dénombrables et tout de même de mesure nulle. Par exemple l'ensemble de Cantor (voir la proposition 13.189).

Proposition 13.177.
La mesure de Lebesgue est invariante par translation, c'est à dire que si A est mesurable alors $\lambda(A) = \lambda(A + \alpha)$ pour tout réel α.

38. Si $q \in \mathbb{Q}$ et $r \in \mathbb{R}\backslash\mathbb{Q}$ alors la suite $(q + r/10^k)_k$ est une suite d'irrationnels convergente vers le rationnel q

Démonstration. Nous commençons par les intervalles ouverts :

$$\lambda\big(]a,b[+\alpha\big)=\lambda\big(]a+\alpha,b+\alpha[\big)=(b+\alpha)-(a+\alpha)=b-a=\lambda\big(]a,b[\big). \tag{13.481}$$

D'après ce qui est dit dans l'exemple 13.41, la mesure de Lebesgue sur les boréliens est invariante par translation.

Si A est mesurable alors il existe un borélien B et un ensemble négligeable N tels que $A=B\cup N$ par la caractérisation 13.97b de la complétion. Alors $A+\alpha=B+\alpha\cup N+\alpha$ et $N+\alpha$ est encore un ensemble négligeable. Donc $\lambda(A+\alpha)=\alpha(B+\alpha)=\lambda(B)$. □

Le mesure ℓ définie sur l'algèbre de parties $\mathcal{A}_{\mathcal{S}}$ (voir proposition 13.170). La proposition 13.22 nous donne donc une mesure extérieure par

$$\lambda^*(X)=\inf\{\sum_n \ell(A_n);A_n\in\mathcal{A}_{\mathcal{S}},X\subset\bigcup_n A_n\}. \tag{13.482}$$

La proposition suivante montre que cette mesure extérieure peut être exprimée seulement avec des intervalles ouverts.

Proposition 13.178.
Nous avons

$$\lambda^*(X)=\inf\{\sum_{n\geqslant 1}\ell(I_n);I_n \text{ sont des intervalles ouverts et } X\subset\bigcup_n I_n\}. \tag{13.483}$$

Démonstration. Nous savons que dans la définition (13.482), chacun des A_n est une réunion disjointe d'intervalles (pas spécialement ouverts) deux à deux disjoints ; donc

$$\lambda^*(X)=\inf\{\sum_n \ell(I_n);I_n\in\mathcal{S},X\subset\bigcup_n I_n\}. \tag{13.484}$$

Soit $\epsilon>0$. Si $A\subset\bigcup_n I_n$, pour chaque $n\geqslant 1$ nous considérons un intervalle ouvert J_n tel que $I_n\subset J_n$ et $\ell(I_n)+\frac{\epsilon}{2^n}\leqslant\ell(J_n)$. Faisant cela pour chacun des découpages de X en intervalles nous trouvons

$$\lambda^*(X)\leqslant\inf\{\sum_n \ell(J_n)\ J_n \text{ est ouvert et } X\subset\bigcup_n J_n\}+\epsilon. \tag{13.485}$$

Étant donné que ϵ est arbitraire nous avons l'égalité. □

Proposition 13.179 ([161])**.**
Si $X\subset\mathbb{R}$ est tel que $\lambda^*(X)<\infty$ alors

(1) Pour tout $\epsilon>0$ il existe un ouvert Ω_ϵ tel que

$$\begin{cases} X\subset\Omega_\epsilon & \text{(13.486a)}\\ \lambda(\Omega_\epsilon)\leqslant\lambda^*(X)+\epsilon. & \text{(13.486b)}\end{cases}$$

(2) Il existe une intersection dénombrable d'ouverts G telle que

$$\begin{cases} X\subset G & \text{(13.487a)}\\ \lambda(G)=\lambda^*(X). & \text{(13.487b)}\end{cases}$$

Démonstration. Pour (1), la proposition 13.178 nous a déjà dit que

$$\lambda^*(X)=\inf\{\sum_n \ell(I_n)\ I_n \text{ est un intervalle ouvert}, X\subset\bigcup_n I_n\}, \tag{13.488}$$

donc si $\epsilon>0$, il existe des intervalles ouverts I_n tels que

$$\begin{cases} X\subset\bigcup_n I_n & \text{(13.489a)}\\ \sum_n \ell(I_n)\leqslant\lambda^*(X)+\epsilon. & \text{(13.489b)}\end{cases}$$

Si nous posons $\Omega_\epsilon = \bigcup_n I_n$, alors nous avons bien

$$\begin{cases} X \subset \Omega_\epsilon & (13.490a) \\ \lambda(\Omega_\epsilon) \leqslant \sum_n \ell(I_n) \leqslant \lambda^*(X) + \epsilon. & (13.490b) \end{cases}$$

En ce qui concerne (2), pour chaque $k \geqslant 1$ nous considérons l'ensemble $\Omega_{1/k}$ obtenu comme précédemment avec $\epsilon = 1/k$ et nous posons $G = \bigcap_{k\geqslant 1} \Omega_{1/k}$. Cela est une intersection dénombrable d'ouverts vérifiant $X \subset G$ (parce que $X \subset \Omega_{1/k}$ pour tout k) et donc $\lambda^*(X) \leqslant \lambda^*(G) = \lambda(G)$. De plus pour tout k nous avons

$$\lambda(G) \leqslant (\Omega_{1/k}) \leqslant \lambda^*(X) + \frac{1}{k} \tag{13.491}$$

pour tout k. En faisant $k \to \infty$ nous avons

$$\lambda(G) \leqslant \lambda^*(X). \tag{13.492}$$

Au final

$$\lambda(G) \leqslant \lambda^*(X) \leqslant \lambda(G), \tag{13.493}$$

d'où l'égalité. □

Corollaire 13.180.
Une partir $N \subset \mathbb{R}$ est négligeable[39] *si et seulement si $\lambda^*(N) = 0$.*

Démonstration. Nous savons que si N est négligeable il existe un borélien Y tel que $N \subset Y$ avec $\lambda(Y) = 0$. Par conséquent [40]

$$\lambda^*(N) \leqslant \lambda^*(Y) = \lambda(Y) = 0. \tag{13.494}$$

Pour l'implication inverse nous supposons que $\lambda^*(N) = 0$ et nous prenons l'ensemble G définit par la proposition 13.179(2) : c'est un borélien contenant N et tel que $\lambda(G) = \lambda^*(N) = 0$. L'ensemble N est donc négligeable. □

Théorème 13.181 (Régularité extérieure de la mesure de Lebesgue)**.**
Pour tout mesurable $A \subset \mathbb{R}$ nous avons

$$\lambda(A) = \inf\{\lambda(\Omega); \Omega \text{ ouvert contenant } A\}. \tag{13.495}$$

Démonstration. Nous commençons par le cas où B est un borélien.

Si B borélien, $\lambda(B) < \infty$ Soit $\epsilon > 0$; par la proposition 13.179(1) il existe un ouvert Ω_ϵ contenant B tel que $\lambda(\Omega_\epsilon) \leqslant \lambda^*(B) + \epsilon$. Vu qu'ici B est borélien, $\lambda^*(B) = \lambda(B)$ et nous concluons que pour tout ϵ il existe un ouvert Ω_ϵ tel que

$$\begin{cases} B \subset \Omega_\epsilon & (13.496a) \\ \lambda(\Omega_\epsilon) \leqslant \lambda(B) + \epsilon, & (13.496b) \end{cases}$$

et donc

$$\lambda(B) = \inf\{\lambda(\Omega);\ \Omega \text{ ouvert contenant } B\ \}. \tag{13.497}$$

Si B borélien, $\lambda(B) = +\infty$ Dans ce cas l'infimum est pris uniquement sur des ouverts Ω tels que $\lambda(\Omega) = \infty$.

39. Définition 13.48.
40. Au péril d'être lourd nous rappelons que λ^* est défini sur toutes les parties de $\mathbb{R}$.

Si A est mesurable non borélien Nous passons maintenant au cas où A est mesurable sans être borélien. Il s'écrit donc $A = B \cup N$ avec B borélien et N négligeable par la proposition 13.51, et par définition $\lambda(A) = \lambda(B)$. Si Y est un borélien tel que $N \subset Y$ et $\lambda(Y) = 0$ alors

$$\begin{aligned}
\lambda(A) = \lambda(B) &= \inf\{\lambda(\Omega) \text{ tel que } \Omega \text{ ouvert}, B \subset \Omega\} && (13.498a)\\
&\leqslant \inf\{\lambda(\Omega) \text{ tel que } \Omega \text{ ouvert}, B \cup N \subset \Omega\} && (13.498b)\\
&\leqslant \inf_{\Omega', Y'}\{\lambda(\Omega' \cup Y') \text{ tel que } \Omega', Y' \text{ ouverts}, B \subset \Omega', Y \subset Y'\} && (13.498c)\\
&\leqslant \inf_{\Omega', Y'}\{\lambda(\Omega') + \lambda(Y') \text{ tel que } \Omega', Y' \text{ ouverts}, B \subset \Omega', Y \subset Y'\} && (13.498d)\\
&\leqslant \inf_{\Omega'}\{\lambda(\Omega') \text{ tel que } \Omega' \text{ ouvert}, B \subset \Omega\} && (13.498e)\\
&= \lambda(B). && (13.498f)
\end{aligned}$$

Justifications :
— (13.498a) Le cas borélien déjà fait.
— (13.498b) Les ouverts Ω tels que $B \cup N \subset \Omega$ vérifient a fortiori $B \subset \Omega$; nous avons donc agrandit l'ensemble sur lequel l'infimum est pris.
— (13.498c) Parmi les ouverts Ω qui recouvrent $B \cup N$, il y a ceux de la forme $\Omega' \cup Y'$ où Ω' recouvre B et Y' est un ouvert contenant Y. Donc nous avons rétréci l'ensemble sur lequel l'infimum est pris et par conséquent agrandit l'infimum.
— (13.498d) Mesure d'une union majorée par la somme des mesures.
— (13.498e) Vu que Y est borélien, $\lambda(Y) = \inf_{Y' \text{ ouvert}}\{\lambda(Y') \text{ tel que } Y \subset Y'\} = 0$. Donc pour tout Ω' et tout $\epsilon > 0$, nous pouvons trouver un Y' vérifiant les conditions tel que $\lambda(\Omega') + \lambda(Y') \leqslant \lambda(\Omega') + \epsilon$.

Toutes les inégalités sont des égalités en en particulier (13.498b) donne

$$\lambda(A) = \inf\{\lambda(\Omega) \text{ tel que } \Omega \text{ ouvert}, B \cup N \subset \Omega\}, \tag{13.499}$$

ce qu'il fallait.

□

Proposition 13.182 ([161])**.**
Si A est mesurable dans $\mathbb{R}$ et si $\epsilon > 0$ alors il existe un ouvert Ω_ϵ et un fermé F_ϵ tels que

$$\begin{cases} F_\epsilon \subset A \subset \Omega_\epsilon & (13.500a)\\ \lambda(\Omega_\epsilon \backslash F_\epsilon) \leqslant \epsilon. & (13.500b) \end{cases}$$

Démonstration. Nous commençons par le cas d'un borélien B.

Première étape Montrons qu'il existe un ouvert U_ϵ tel que

$$\begin{cases} B \subset U_\epsilon & (13.501a)\\ \lambda(U_\epsilon \backslash B) \leqslant \dfrac{\epsilon}{2}. & (13.501b) \end{cases}$$

Si $\lambda(B) < \infty$ alors le théorème 13.181 nous donne un ouvert U_ϵ tel que $B \subset U_\epsilon$ et $\lambda(U_\epsilon) \leqslant \lambda(B) + \frac{\epsilon}{2}$. Nous avons alors

$$\lambda(\Omega_\epsilon \backslash B) = \lambda(\Omega_\epsilon) - \lambda(B) \leqslant \frac{\epsilon}{2}. \tag{13.502}$$

Si par contre $\lambda(B) = \infty$, nous posons $B_n = B \cap [-n, n]$ et $\epsilon_n = \epsilon/2^{n+1}$. Pour chaque n nous avons un ouvert Ω_n tel que

$$\begin{cases} B_n \subset \Omega_n & (13.503a)\\ \lambda(\Omega_n \backslash B_n) \leqslant \dfrac{\epsilon}{2^{n+1}} & (13.503b) \end{cases}$$

Par conséquent en posant $\Omega = \bigcup_{n\geqslant 1}\Omega_n$ nous avons[41]

$$\begin{cases} B \subset \Omega & (13.504a)\\ \lambda(\Omega\backslash B) \leqslant \lambda\big(\bigcup_n(\Omega_n\backslash B_n)\big) \leqslant \sum_{n\geqslant 1}\lambda(\Omega_n\backslash B_n) = \frac{\epsilon}{2}. & (13.504b)\end{cases}$$

La première étape est terminée.

Deuxième étape Nous prouvons à présent qu'il existe un ouvert Ω_ϵ et un fermé F_ϵ tels que

$$\begin{cases} F_\epsilon \subset B \subset \Omega_\epsilon & (13.505a)\\ \lambda(\Omega_\epsilon\backslash B) \leqslant \frac{\epsilon}{2} & (13.505b)\\ \lambda(B\backslash F_\epsilon) \leqslant \frac{\epsilon}{2}. & (13.505c)\end{cases}$$

L'ouvert Ω_ϵ, nous l'avons déjà de l'étape précédente. Pour le fermé, nous appliquons la première étape au borélien B^c ; ce qui nous trouvons est un ouvert G_ϵ tel que

$$\begin{cases} B^c \subset G_\epsilon & (13.506a)\\ \lambda(G_\epsilon\backslash B^c) \leqslant \frac{\epsilon}{2}. & (13.506b)\end{cases}$$

En posant $F_\epsilon = G_\epsilon^c$ nous avons un fermé tel que $F_\epsilon \subset B$ et

$$\lambda(B\backslash F_\epsilon) = \lambda(F_\epsilon^c\backslash B^c) = \lambda(G_\epsilon\backslash B^c) \leqslant \frac{\epsilon}{2}. \tag{13.507}$$

Dernière étape Les ensembles F_ϵ et Ω_ϵ trouvés à la deuxième étape donnent bien les relations (13.500). En effet $\Omega_\epsilon\backslash F_\epsilon = (\Omega_\epsilon\backslash B)\cup(B\backslash F_\epsilon)$, donc

$$\lambda(\Omega_\epsilon\backslash F_\epsilon) \leqslant \lambda(\Omega_\epsilon\backslash B) + \lambda(B\backslash F_\epsilon) = \epsilon. \tag{13.508}$$

Nous passons au cas où $A = B\cup N$ est mesurable. Nous commençons par prendre les Ω_ϵ et F_ϵ qui correspondent à B :

$$\begin{cases} F_\epsilon \subset B \subset \Omega_\epsilon & (13.509a)\\ \lambda(\Omega_\epsilon\backslash F_\epsilon) \leqslant \epsilon. & (13.509b)\end{cases}$$

Soit Y un borélien tel que $N \subset Y$ et $\lambda(Y)$ puis un ouvert Y' tel que $\lambda(Y') \leqslant \epsilon$ et $Y \subset Y'$. L'existence d'un tel Y' est assurée par la proposition 13.181 appliquée à Y. Nous vérifions que les ensembles F_ϵ et $\Omega_\epsilon \cup Y'$ fonctionnent. En effet $\Omega_\epsilon \cup Y'\backslash F_\epsilon \subset (\Omega_\epsilon\backslash F_\epsilon)\cup Y'$, donc

$$\begin{cases} F_\epsilon \subset B\cup N \subset \Omega_\epsilon \cup Y' & (13.510a)\\ \lambda\big((\Omega_\epsilon\backslash F_\epsilon)\big) \leqslant \lambda(\Omega_\epsilon\backslash F_\epsilon) + \lambda(Y') \leqslant 2\epsilon. & (13.510b)\end{cases}$$

Donc en réalité il faut choisir $\Omega_{\epsilon/2}$, $F_{\epsilon/2}$ et $\lambda(Y') \leqslant \epsilon/2$. □

Théorème 13.183 (Régularité intérieure de la mesure de Lebesgue)**.**
Si A est mesurable dans $\mathbb{R}$ alors

$$\lambda(A) = \sup\{\lambda(K); K \text{ compact contenu dans } A\}. \tag{13.511}$$

Démonstration. Par la proposition 13.182 nous avons

$$\lambda(A) = \sup_{F \text{ fermé dans } A} \lambda(F). \tag{13.512}$$

41. Nous utilisons la petite relation ensembliste $\big(\bigcup_n A_n\big)\backslash\big(\bigcup_n B_n\big) \subset \bigcup_n(A_n\backslash B_n)$.

Pour un tel F nous posons $K_n = F \cap [-n, n]$ qui est compact [42] et contenu dans B. De plus le lemme 13.27(2) nous dit que

$$\lambda(F) = \lim_{n\to\infty} \lambda(K_n) \tag{13.513}$$

Donc tous les $\lambda(F)$ peuvent être arbitrairement approchés par un $\lambda(K)$ avec K compact dans A, et le supremum (13.512) n'est pas affecté en nous restreignant à prendre des compacts contenus dans B :

$$\lambda(A) = \sup_{F \text{ fermé dans } A} \lambda(F) = \sup_{K \text{ compact dans } A} \lambda(K). \tag{13.514}$$

□

13.9.3 Ensemble de Cantor

Nous considérons la fonction donnant l'écriture décimale des nombres définie en (10.140).

Définition 13.184 (Ensemble de Cantor)**.**
Soit $K_0 = [0, 1[$ et les ensembles K_n définis par la récurrence

$$K_{n+1} = \big(\frac{1}{3}K_n\big) \cup \big(\frac{1}{3}(K_n + 2)\big). \tag{13.515}$$

L'ensemble

$$K = \bigcup_{n\geqslant 0} K_n \tag{13.516}$$

*est l'**ensemble triadique de Cantor**.*

Les principales propriétés de l'ensemble de Cantor sont qu'il est non dénombrable (proposition 13.188) et borélien de mesure nulle (proposition 13.189).

13.185.
L'idée de base pour prouver que l'ensemble K est non dénombrable est que ses éléments sont les nombres qui s'écrivent en base 3 sans utiliser le chiffre 1. En prenant un nombre sans 1 écrit en base 3, en changeant tous les 2 en 1 et en lisant le résultat en base 2, nous obtenons tous les nombres possibles en base 2 et donc une quantité non dénombrable. L'idée est donc simple et astucieuse. La mise en musique est un peu plus délicate parce qu'il faut faire attention aux queues de suites ; c'est pour cela que nous avons construit l'ensemble de Cantor en partant de $[0, 1[$ et non de $[0, 1]$.

Le lemme suivant dit précisément ce que nous entendons en disant que les éléments de l'ensemble de Cantor sont les nombres qui s'écrivent en base 3 sans utiliser le chiffre 1.

Lemme 13.186 ([1])**.**
Soit $n \in \mathbb{N}$ et $x \in \mathbb{D}_3$; nous avons $\varphi_3(x) \in K_n \in$ si et seulement si $x_1, \dots, x_n \in \{0, 2\}$.

Démonstration. Nous procédons par récurrence en commençant avec $n = 1$. Si $x_1 = 1$ alors

$$\varphi_3(x) = \frac{1}{3} + \sum_{k=2}^{\infty} \frac{x_k}{3^k} \in [\frac{1}{3}, \frac{2}{3}[. \tag{13.517}$$

Notons que $\varphi_3(x) = \frac{2}{3}$ est impossible parce que ça demanderait une queue de suite de 2. Par conséquent $\varphi_3(x) = [0, 1[\backslash[\frac{1}{3}, \frac{2}{3}[= K_1$.

Nous passons à la récurrence.

Sens direct Nous supposons que $x_1, \dots, x_{n+1} \in \{0, 2\}$ et nous montrons que $\varphi_3(x) \in K_{n+1}$. La chose surprenante est que nous n'allons pas considérer deux cas suivant que x_{n+1} vaut 0 ou

42. parce que fermé et borné, théorème de Borel-Lebesgue 6.123.

1 ; nous allons considérer deux cas suivant [43] que x_1 vaut 0 ou 1. Écrivons encore $\varphi_3(x)$:

$$\varphi_3(x) = \sum_{k=1}^{n+1} \frac{x_k}{3^k} + \sum_{k=n+2}^{\infty} \frac{x_k}{3^k}. \tag{13.518}$$

Si $x_1 = 0$ Alors nous avons

$$3\varphi_3(x) = \sum_{k=2}^{\infty} \frac{x_k}{3^{k-1}} = \sum_{k=1}^{\infty} \frac{x_{k+1}}{3^k} = \varphi_3(x_2, \dots, x_n, x_{n+1}, \dots) \tag{13.519}$$

Vu que par hypothèse $x_2, \dots, x_{n+1}$ sont dans $\{0, 2\}$ nous avons $3\varphi_3(x) \in K_n$ par hypothèse de récurrence. Cela implique que $\varphi_3(x) \in K_{n+1}$.

Si $x_1 = 2$ Alors

$$\varphi_3(x) = \frac{2}{3} + \sum_{k=2}^{\infty} \frac{x_k}{3^k}, \tag{13.520}$$

et

$$3\varphi_3(x) - 2 = \sum_{k=1}^{\infty} \frac{x_{k+1}}{3^k} = \varphi(x_2, \dots, x_{n+1}, \dots), \tag{13.521}$$

et donc là nous avons $3\varphi_3(x) - 2 \in K_n$, ce qui implique encore $\varphi_3(x) \in K_{n+1}$.

Sens réciproque Nous devons maintenant prouver que $\varphi_3(x) \in K_{n+1}$ implique $x_1, \dots, x_{n+1} \in \{0, 2\}$. Par le même calcul que précédemment nous avons soit

$$3\varphi_3(x) = \varphi_3(x_2, \dots, x_{n+1}, \dots), \tag{13.522}$$

si $x_1 = 0$, soit

$$3\varphi_3(x) - 2 = \varphi_3(x_2, \dots, x_{n+1}, \dots), \tag{13.523}$$

si $x_1 = 2$. Dans les deux cas, si $x_l = 1$ pour un certain $2 \leqslant l \leqslant n + 1$, alors l'hypothèse de récurrence donne que ces éléments ne sont pas dans K_n et donc $\varphi_3(x)$ pas dans K_{n+1}.

□

Corollaire 13.187 ([1])**.**
En posant $\mathbb{E} = \{x \in \mathbb{D}_3$ tel que $x_i \neq 1 \forall i\}$ nous avons $K = \varphi_3(\mathbb{E})$. Et plus précisément, $\varphi_3 \colon \mathbb{E} \to K$ est une bijection.

Démonstration. Nous divisons la preuve en trois étapes.

Image contenue dans K Si $x \in \mathbb{E}$ et $n \in \mathbb{N}$ nous avons $x_1, \dots, x_n \in \{0, 2\}$ et donc $\varphi_3(x) \in K_n$ par la proposition 13.186. Donc

$$\varphi_3(x) \in \bigcup_{n \geqslant 1} K_n = K. \tag{13.524}$$

Injective L'application $\varphi_3 \colon \mathbb{E} \to K$ est injective parce qu'elle est déjà injective depuis $\mathbb{D}_3$.

Surjective Soit $p \in K \subset [0, 1[$. Vu que $\varphi_3 \colon \mathbb{D}_3 \to [0, 1[$ est surjective (théorème 10.69), il existe $x \in \mathbb{D}_3$ tel que $\varphi_3(x) = p$. Pour tout n nous avons $\varphi_3(x) \in K_n$ et donc $x_1, \dots, x_n \in \{0, 2\}$ et donc au final $x \in \mathbb{E}$.

□

Proposition 13.188 ([1])**.**
L'ensemble de Cantor est non dénombrable.

43. Pour comprendre pourquoi, faire un dessin de comment K_n se transforme en K_{n+1} et remarquer dans K_2, les deux premiers segments ne sont pas une division du premier segment de K_1, mais bien une copie des *deux* segments de K_1.

Démonstration. Nous avons prouvé à la proposition 10.70 que l'ensemble $\mathbb{D}_2$ n'était pas dénombrable. Nous allons à présent prouver que l'application

$$\begin{aligned}\psi\colon \mathbb{D}_2 &\to K\\ c &\mapsto \varphi_3(c \text{ en remplaçant les 1 par des 2})\end{aligned} \tag{13.525}$$

est une bijection. Le fait que ψ soit injective est une conséquence du fait que ce soit la composition de deux applications injectives (le remplacement et φ_3). Il faut par contre montrer que l'image est égale à K, en notant qu'il n'est pas évident a priori que l'image soit contenue dans K.

L'opération qui consiste à remplacer les 1 par des 2 est une bijection $\mathbb{D}_2 \to \mathbb{E}$. Le corollaire 13.187 nous dit aussi que $\varphi_3\colon \mathbb{E} \to K$ est une bijection. En tant que composée de bijections, ψ est une bijection.

Étant en bijection avec $\mathbb{D}_2$ qui n'est pas dénombrable par la proposition 10.70, l'ensemble de Cantor n'est pas dénombrable. □

Proposition 13.189 (Ensemble de Cantor)**.**
L'ensemble de Cantor[44] *est borélien, non dénombrable et de mesure nulle.*

Démonstration. Nous reprenons les notations de la définition 13.184. Le fait que l'ensemble de Cantor soit non dénombrable a été prouvé dans la proposition 13.188.

L'ensemble de Cantor étant une intersection dénombrable de boréliens, il est borélien par le lemme 13.3. Vu que $K_n \subset [0,1[$ nous avons $\frac{1}{3}K_n \leqslant \frac{1}{3}$ et $\frac{1}{3}(K_n+2) \geqslant \frac{2}{3}$, donc K_n est une union disjointe de 2^n intervalles de mesure $2/3^n$. Nous avons donc

$$\lambda(K_n) = \left(\frac{2}{3}\right)^n. \tag{13.526}$$

L'ensemble de Cantor étant contenu dans chacun des K_n, sa mesure est plus petite que la mesure de chacun des K_n (lemme 13.26) et donc $\lambda(K) \leqslant \left(\frac{2}{3}\right)^n$ pour tout n ; ergo $\lambda(K) = 0$. □

13.9.4 Ensemble de Vitali (non mesurable)

Exemple 13.190(Un ensemble non mesurable au sens de Lebesgue)
Nous considérons[180] l'ensemble quotient $\mathbb{R}/\mathbb{Q}$; chaque classe intersecte l'intervalle $[0,1]$. Grâce à l'axiome du choix (voir 2.2) nous pouvons construire un ensemble V contenant un représentant dans $[0,1]$ de chaque classe. Un tel ensemble est un **ensemble de Vitali**. Nous allons prouver que V n'est pas mesurable.

Supposons que V soit mesurable.Alors tous les ensemble de la forme $V+q$ ($q \in \mathbb{Q}$) sont mesurables et ont même mesure par la proposition 13.177. Nous posons

$$A = \bigcup_{\substack{q\in\mathbb{Q}\\ -1\leqslant q\leqslant 1}} (V+q) \subset [-1,2]. \tag{13.527}$$

Cela est une union disjointe d'ensembles mesurables. Donc

$$\lambda(A) = \sum_{\substack{q\in\mathbb{Q}\\ -1\leqslant q\leqslant 1}} \lambda(V+q). \tag{13.528}$$

Vu que $A \subset [-1,2]$ nous avons $\lambda(A) \leqslant 3$ et donc tous les termes de la somme doivent être nuls. Nous avons donc $\lambda(A) = 0$.

Prouvons toutefois que $[0,1] \subset A$, ce qui serait une contradiction. Soit $x \in [0,1]$; il est dans une des classes de $\mathbb{R}/\mathbb{Q}$ et donc il existe $v \in V$ tel que $x - v \in \mathbb{Q}$. De plus $x, v \in [0,1]$, donc

$$-1 \leqslant x - v \leqslant 1. \tag{13.529}$$

Cela fait que $x \in V + (x-v) \subset A$. Nous avons donc $x \in A$ et donc $[0,1] \subset A$. En conséquence de quoi nous aurions $\lambda(A) \geqslant 1$. △

44. Définition 13.184

13.10 Tribu et mesure de Lebesgue sur $\mathbb{R}^d$

Quelques liens internes :
— Le produit de tribus est donné par la définition 13.151,
— le produit d'espaces mesurés est donné par la définition 13.161.

La mesure de Lebesgue sur $\mathbb{R}^d$, notée λ_N est d'abord la mesure définie sur

$$\big(\mathbb{R}^d, \mathcal{B}\text{or}(\mathbb{R}) \otimes \ldots \otimes \mathcal{B}\text{or}(\mathbb{R})\big) \tag{13.530}$$

comme le produit $\lambda \otimes \ldots \otimes \lambda$. Ensuite nous nous souvenons du corollaire 13.155 qui donne λ_N comme une mesure sur

$$\big(\mathbb{R}^N, \mathcal{B}\text{or}(\mathbb{R}^N)\big). \tag{13.531}$$

Et enfin λ_N est la complétion de cette mesure.

Proposition 13.191 ([181])**.**
Tout ouvert de $\mathbb{R}^n$ est une union dénombrable et disjointe de cubes semi-ouverts.

Démonstration. Nous allons même montrer que ces cubes peuvent être choisis sur un quadrillage.

Soit G un ouvert de $\mathbb{R}^n$. Soit $\{Q_i^1\}_{i\in\mathbb{N}}$ un découpage de $\mathbb{R}^n$ en cubes semi-ouverts de côté 1 et dont les sommets sont en les coordonnées entières. Ils sont de la forme

$$\prod_{i=1}^{n} [n_i, n_i + 1[\tag{13.532}$$

où les n_i sont des entiers. Ce sont des cubes disjoints. Nous considérons ensuite pour chaque $k > 1$ le découpage $\{Q_i^{(k)}\}_{i\in\mathbb{N}}$ de $\mathbb{R}^n$ en cubes de côtés 2^{-k} qui consiste à découper en 2 les côtés des cubes du découpage $Q^{(k-1)}$. Ces cubes forment encore un découpage dénombrable de $\mathbb{R}^n$ en des cubes disjoints. Ils sont de la forme

$$\prod_{i=1}^{n} [\frac{n_i}{2^k}, \frac{n_i + 1}{2^k}[\tag{13.533}$$

où les n_i sont encore entiers. Ensuite nous considérons $\mathcal{E}$ l'union de tous les $Q_i^{(k)}$ contenus dans G.

Montrons que $\mathcal{E} = G$. D'abord $\mathcal{E} \subset G$ parce que $\mathcal{E}$ est une union d'ensembles contenus dans G. Ensuite si $x \in G$, il existe une boule de rayon r autour de x contenue dans G ; alors un des ensembles $Q_i^{(k)}$ avec $2^{-j} < \frac{r}{2}$ est contenue dans $B(x, r)$ et donc dans $\mathcal{E}$.

Bien entendu l'union qui donne $\mathcal{E}$ n'est pas satisfaisante par ce que les $Q_i^{(k+1)}$ sont contenus dans les $Q_i^{(k)}$; les intersections sont donc loin d'être vides.

Nous faisons ceci :

$$R^{(0)} = \{Q_i^{(1)} \text{contenu dans } G\} \tag{13.534a}$$

$$R^{(k+1)} = \{Q_i^{(k+1)} \text{contenus dans } G \text{ et pas dans } R^{(k)}\}. \tag{13.534b}$$

En fin de compte l'union de tous les ensembles contenus dans les $R^{(k)}$ forment encore $\mathbb{R}^n$, mais sont d'intersection vide. □

Les cubes dont il est question dans cette preuve, de côtés 2^{-k} sont souvent appelés des cubes **dyadiques**.

Corollaire 13.192 ([181])**.**
Tout ouvert de $\mathbb{R}^n$ est une union dénombrable de cubes presque disjoints[45]*.*

Démonstration. Il suffit de prendre les cubes de la proposition 13.191 et de les fermer. Ce que l'on ajoute est de mesure nulle. □

45. «presque» au sens où les intersections éventuelles sont de mesure de Lebesgue nulle.

Remarque 13.193.
La proposition 13.191 est une propriété seulement de la topologie de $\mathbb{R}^n$ alors que le corollaire fait intervenir la mesure de Lebesgue parce qu'il faut bien dire que les intersections sont de mesure (de Lebesgue) nulle.

13.10.1 Ensembles négligeables

Lemme 13.194 ([182])**.**
L'image d'une partie négligeables de $\mathbb{R}^N$ par une application Lipschitz est négligeable.

Démonstration. Soit N une partie négligeable de $\mathbb{R}^N$ et une application Lipschitz $f\colon N \to \mathbb{R}^N$. Soit $Q \subset \mathbb{R}^N$ un cube borné de côté r. Pour tout $x, x' \in N \cap Q$ nous avons

$$\|f(x) - f(x')\| \leqslant C\|x - x'\| \leqslant Cr. \tag{13.535}$$

Donc $f(N \cap Q)$ est dans une boule de rayon Cr. Mais comme toutes les normes sont équivalentes [46] sur $\mathbb{R}^N$ nous pouvons tout aussi bien prendre la norme $\|.\|_1$ au lieu de la norme $\|.\|_2$ (qui est touours la norme prise implicitement lorsqu'on parle de $\mathbb{R}^n$), de telle sorte que les boules soient des cubes. Quoi qu'il en soit, $f(N \cap Q)$ est contenu dans un cube de coté $2Cr$ et au niveau de la mesure extérieure,

$$m^*\big(f(N \cap Q)\big) \leqslant (2Cr)^N = (2C)^N r^N, \tag{13.536}$$

ou encore

$$m\big(f(N \cap Q)\big) \leqslant (2C)^N m(Q) \tag{13.537}$$

parce que r^N est la mesure du cube Q.

Soit maintenant $\epsilon > 0$; vu que N est négligeable, il existe un ouvert U contenant N et tel que $m(U) < \epsilon$. Ce U est une union presque disjointe de cubes dyadiques (Q_n) par le corollaire 13.192. Nous avons alors

$$\begin{aligned}
m^*\big(f(N)\big) &= m^*\big(f(\bigcup_n N \cap Q_n)\big) && \text{(13.538a)}\\
&= m^*\big(\bigcup_n f(N \cap Q_n)\big) && \text{(13.538b)}\\
&\leqslant \sum_n m^*(f(N \cap Q_n)) && \text{(13.538c)}\\
&\leqslant \sum_n (2C)^N m(Q_n) && \text{(13.538d)}\\
&= (2C)^N m(U) && \text{(13.538e)}\\
&< (2C)^d \epsilon. && \text{(13.538f)}
\end{aligned}$$

Au final, $m^*\big(f(N)\big) \leqslant (2C)^N \epsilon$. L'ensemble N est donc négligeable parce que le lemme 13.57 le dit : $m^*(N) = 0$. □

Corollaire 13.195.
Un sous-espace vectoriel strict de $\mathbb{R}^N$ est négligeable.

Démonstration. Un sous-espace vectoriel strict de $\mathbb{R}^N$ de dimension $k < N$ est l'image de

$$A = \{t_1 e_1 + \cdots + t_k e_k \text{ tel que } t_i \in \mathbb{R}\} \tag{13.539}$$

par une application linéaire. Ce A est un pavé de mesure de Lebesgue nulle. Donc l'image est négligeable par le lemme 13.194. □

46. Proposition 10.2

13.10.2 Parties et fonctions mesurables

Pour rappel, la notion d'application de classe C^1 est donnée par la définition 12.215.

Proposition 13.196.
Soient U et V des ouverts de $\mathbb{R}^N$ et $\phi\colon U \to V$ un C^1-difféomorphisme. Si $E \subset U$ est mesurable, alors $\phi(E)$ est mesurable[47].

Démonstration. Si E est mesurable, il existe un borélien B et un ensemble négligeable N tels que $E = B \cup N$. Vu que ϕ est un homéomorphisme, l'application ϕ^{-1} est borélienne parce que continue (théorème 13.72). Nous avons

$$\phi(B) = (\phi^{-1})^{-1}(B), \tag{13.540}$$

c'est à dire que $\phi(B)$ est l'image inverse de B par ϕ^{-1}. L'ensemble $\phi(B)$ est donc borélien.

Il reste à voir que $\phi(N)$ est négligeable. Soit $Q \subset U$ une cube compact. L'application $d\phi\colon Q \to \mathcal{L}(\mathbb{R}^N, \mathbb{R}^N)$ est continue et donc bornée (par la remarque 12.216) sur le compact Q. Par les accroissements finis (théorème 12.229), l'application ϕ est donc Lipschitz sur Q. La partie $\phi(N \cap Q)$ est alors négligeable par le lemme 13.194. Pour conclure,

$$\phi(N) = \bigcup_i \phi(N \cap Q_i) \tag{13.541}$$

où les Q_i sont tous des cubes compacts. Donc $\phi(N)$ est une union dénombrable d'ensembles négligeables ; ergo négligeable lui-même par le lemme 13.49. □

Proposition 13.197.
Soient U et V des ouverts de $\mathbb{R}^N$ et $\phi\colon U \to V$ un C^1-difféomorphisme. Si $f\colon V \to \mathbb{C}$ est mesurable, alors $f \circ \phi\colon U \to \mathbb{C}$ l'est.

Démonstration. Soit A une partie mesurable de $\mathbb{C}$. Il nous faut prouver que

$$(f \circ \phi)^{-1}(A) = \phi^{-1}\big(f^{-1}(A)\big) \tag{13.542}$$

soit mesurable. Par hypothèse , $f^{-1}(A)$ est mesurable. Vu que ϕ est un C^1-difféomorphisme, elle et son inverse sont mesurables par la proposition 13.196. Donc l'image du mesurable $f^{-1}(A)$ par ϕ^{-1} est encore mesurable. □

13.10.3 Propriétés d'unicité

Corollaire 13.198.
La mesure λ_N est l'unique mesure sur $(\mathbb{R}^N, \mathcal{B}or(\mathbb{R}^N))$ à satisfaire

$$\mu\Big(\prod_{i=1}^{N}[a_i, b_i]\Big) = \prod_{i=1}^{n}|a_i - b_i| \tag{13.543}$$

Démonstration. Par définition de la mesure produit, λ_N est l'unique mesure sur $(\mathbb{R}^N, \mathcal{B}\text{or}(\mathbb{R}) \otimes \ldots \otimes \mathcal{B}\text{or}(\mathbb{R}))$ à satisfaire la condition. La proposition 13.155 conclut. □

Vu que les compacts de $\mathbb{R}^n$ sont les fermés bornés (théorème 6.123), et que tout borné est dans un tel produit d'intervalle, la mesure de Lebesgue est une mesure de Borel (définition 13.78(1)).

Théorème 13.199 ([183])**.**
La mesure de Lebesgue est invariante par translation. Autrement dit si A est mesurable dans $\mathbb{R}^n$ et si $a \in \mathbb{R}^n$ alors $A + a$ est mesurable et

$$\lambda_N(A + a) = \lambda_N(A). \tag{13.544}$$

47. Ici «mesurable» parle de mesurabilité au sens de la tribu de Lebesgue, c'est à dire pas seulement les boréliens.

Démonstration. Nous supposons que A est borélien ; sinon il l'est à ensemble négligeable près. Nous notons t_a la translation et nous nommons μ la mesure donnée par

$$\mu(A) = \lambda_N(A + a). \tag{13.545}$$

Vu que

$$\mu\big(\prod_{n=1}^{N}[r_n, s_n[\big) = \lambda_N\big(\prod_i [r_n + a_n, s_n + a_n[\big) = \prod_i |s_n - r_n|. \tag{13.546}$$

Vu qu'il y a unicité de la mesure vérifiant cette propriété (corollaire 13.198), nous avons $\mu = \lambda_N$. □

Pour la suite nous notons Q_0 le cube unité de $\mathbb{R}^N$: $Q_0 = \big([0,1[\big)^N$.

Théorème 13.200 ([183])**.**
Soit μ une mesure positive sur $\mathbb{R}^N$ telle que

(1) μ soit invariante par translation (des boréliens),

(2) $\mu(Q_0) = 1$.

Alors $\mu = \lambda_N$.

Démonstration. Pour simplifier l'écriture nous faisons $N = 2$. Notre but est de prouver que $\mu([0, r[\times [0, r'[) = rr'$ pour tout $r, r' \in \mathbb{R}$.

Longueur $=1/J$ Soient J, K des entiers. Nous pouvons diviser le cube Q_0 en rectangles de cotés $1/J$ et A/K :

$$Q_0 = \bigcup_{\substack{1\leqslant j\leqslant J\\ 1\leqslant k\leqslant K}} [\frac{j-1}{J}, \frac{j}{J}[\times [\frac{k-1}{K}, \frac{k}{K}[\tag{13.547}$$

où l'union est disjointe. En ce qui concerne la mesure nous commençons par utiliser la sous-additivité :

$$\mu(Q_0) = \sum_{\substack{1\leqslant j\leqslant J\\ 1\leqslant k\leqslant K}} \mu\left([\frac{j-1}{J}, \frac{j}{J}[\times [\frac{k-1}{K}, \frac{k}{K}[\right). \tag{13.548}$$

Nous utilisons ensuite, sur chacun des termes séparément l'invariance par translation selon les vecteurs $(\frac{j-1}{J}, 0)$ et $(0, \frac{k-1}{K})$:

$$1 = \mu(Q_0) = \sum_{\substack{1\leqslant j\leqslant J\\ 1\leqslant k\leqslant K}} \mu\left([0, \frac{1}{J}[\times [0, \frac{1}{K}[\right) = JK\mu\mu\left([0, \frac{1}{J}[\times [0, \frac{1}{K}[\right), \tag{13.549}$$

et donc

$$\mu\left([0, \frac{1}{J}[\times [0, \frac{1}{K}[\right) = \frac{1}{J} \times \frac{1}{K}. \tag{13.550}$$

Longueur L/K Soient L, M des entiers et calculons :

$$\mu\left([\frac{0}{J}, \frac{L}{J}[\times [\frac{0}{K}, \frac{M}{K}[\right) = \sum_{\substack{0\leqslant l\leqslant L-1\\ 0\leqslant m\leqslant M-1}} \mu\left([\frac{l}{J}, \frac{l+1}{J}[\times [\frac{m}{K}, \frac{m+1}{K}[\right) \tag{13.551a}$$

$$= LM\mu\left([\frac{0}{J}, \frac{1}{J}[\times [\frac{0}{K}, \frac{1}{K}[\right) \tag{13.551b}$$

$$= LM \times \frac{1}{J} \times \frac{1}{K}. \tag{13.551c}$$

Nous avons donc, pour tout J, K, L, M :

$$\mu\left([0, \frac{L}{J}[\times [0, \frac{M}{K}[\right) = \frac{L}{J} \times \frac{M}{K}, \tag{13.552}$$

c'est à dire que pour tout $r, s \in \mathbb{Q}^+$ nous avons

$$\mu\big([0, r[\times [0, s[\big) = rs. \tag{13.553}$$

Longueur réelle Nous passons au cas de longueur réelle. Soit $a > 0$ et une suite croissante de rationnels $r_n \to a$. Une telle suite existe par la proposition 6.59. Nous avons $[0, a[= \bigcup_{n\geqslant 1}[0, r_n[$ où l'union n'est pas disjointe mais croissante, ce qui permet d'utiliser le lemme 13.27(1) pour écrire

$$\mu\big([0, a[\big) = \mu\left(\bigcup_{n\geqslant 1}[0, r_n[\right) = \lim_{n\to\infty} \mu\big([0, r_n[\big) = \lim_{n\to\infty} r_n = a. \tag{13.554}$$

Enfin, si $a, a' \in \mathbb{R}$, l'invariance par translation donne

$$\mu\big([a, a'[\big) = \mu\big([0, a' - a[\big) = a' - a. \tag{13.555}$$

Par unicité de la mesure ayant cette propriété, nous avons $\mu = \lambda_N$. □

Corollaire 13.201.
Si μ est une mesure positive sur $\mathbb{R}^N$ invariante par translation et telle que $\mu(Q_0) = C < \infty$ alors $\mu = C\lambda_N$.

Démonstration. Si $C > 0$ nous considérons la mesure $\frac{1}{C}\mu$ qui vérifie $(\frac{1}{C}\mu)(Q_0) = 1$. En conséquence du théorème 13.200, $\frac{1}{C}\mu = \lambda_N$ et $\mu = C\lambda_N$.

Si au contraire $C = 0$ alors nous pouvons paver $\mathbb{R}^N$ avec des cubes Q_i de côté 1 qui ont tous mesure 0. Par conséquent, $\mathbb{R}^N = \bigcup_{i=1}^{\infty} Q_i$, donc $\mu(\mathbb{R}^N) = \sum_i \mu(Q_i) = 0$. Par conséquent $\mu = 0$ parce que toute partie de $\mathbb{R}^N$ a une mesure au maximum égale à celle de $\mathbb{R}^N$. □

13.10.4 Régularité

Les différentes notions de régularité pour une mesure sont données dans la définition 13.78. Ce sont essentiellement des questions de compatibilité entre la mesure et la topologie.

Proposition 13.202.
La mesure de Lebesgue est une mesure de Radon sur tout ouvert de $\mathbb{R}^N$.

Démonstration. Soit V un ouvert de $\mathbb{R}^N$. C'est localement compact et dénombrable à l'infini. Il suffit de prouver que λ_N est de Borel sur V pour que le théorème 13.79 conclue à la régularité de la mesure de Lebesgue.

Soit K un compact de V. Par la proposition 6.76 c'est également un compact de $\mathbb{R}^N$. Par conséquent K est dans un pavé fermé de $\mathbb{R}^N$ du type

$$K \subset \prod_{n=1}^{N}[a_n, b_n] \tag{13.556}$$

et donc en passant par le corollaire 13.198,

$$\lambda_N(K) \leqslant \prod_{i=1}^{N}(b_n - a_n) < \infty. \tag{13.557}$$

Nous avons démontré que λ_N reste fini sur tout compact de V. □

13.11 Propriétés de l'intégrale de Lebesgue

13.11.1 Théorème de la moyenne

Théorème 13.203 ([1]).
Soit Q un compact connexe par arcs et une fonction continue $f\colon Q \to \mathbb{R}$. Si λ est la mesure de Lebesgue, alors il existe $a \in Q$ tel que

$$f(a) = \frac{1}{\lambda(Q)}\int_Q f d\lambda \tag{13.558}$$

Démonstration. En posant $I = \int_Q f d\lambda$ nous avons immédiatement

$$\min(f)\lambda(Q) \leqslant I \leqslant \max(f)\lambda(Q) \tag{13.559}$$

où le minimum et le maximum existent parce que f est continue sur un compact. Si une des deux inégalités est une égalité alors la fonction est constante. En effet supposons que la première inégalité soit une égalité ; si la fonction n'était pas constante, il existerait une boule sur laquelle f serait strictement supérieure à $\min(f)$. En intégrant d'abord sur cette boule et ensuite sur le complémentaire nous obtenons une intégrale plus grande que $\min(f)\lambda(Q)$.

Soit $\epsilon > 0$. Il existe $\alpha, \beta \in Q$ tels que $f(\alpha) \leqslant \min(f) + \epsilon$ et $f(\beta) \geqslant \max(f) - \epsilon$. Soit $\gamma\colon [0,1] \to Q$ un chemin continu tel que $\gamma(0) = \alpha$ et $\gamma(1) = \beta$. La fonction $f \circ \gamma\colon [0,1] \to \mathbb{R}$ est alors continue et vérifie $(f \circ \gamma)(0) \leqslant \min(f) + \epsilon$ et $(f \circ \gamma)(1) \geqslant \max(f) - \epsilon$.

Si ϵ est assez petit et vu que les inégalités (13.559) sont strictes,

$$\lambda(Q)(f \circ \gamma)(0) \leqslant \min(f)\lambda(Q) + \epsilon\lambda(Q) < I < \max(f)\lambda(Q) - \epsilon\lambda(Q) \leqslant \lambda(Q)(f \circ \gamma)(1). \tag{13.560}$$

Par le théorème des valeurs intermédiaires 12.21, il existe $t_0 \in [0,1]$ tel que $\lambda(Q)(f \circ \gamma)(t_0) = I$. Le point $a = \gamma(t_0)$ vérifie

$$f(a) = \frac{1}{\lambda(Q)} \int_Q f d\lambda. \tag{13.561}$$

□

13.11.2 Primitives et intégrales

En termes de notations, si $a < b$ nous posons

$$\int_a^b f(t)dt = \int_{[a,b]} f. \tag{13.562}$$

Si par contre $a > b$ nous posons $\int_a^b f = -\int_b^a f$.

Proposition 13.204 (Primitive et intégrale)**.**
Soit f une fonction intégrable sur $[a,b]$ et continue sur $]a,b[$. Alors la fonction

$$\begin{aligned} F\colon [a,b] &\to \mathbb{R} \\ x &\mapsto \int_{[a,x]} f(t)dt. \end{aligned} \tag{13.563}$$

est l'unique primitive de f sur $]a,b[$ s'annulant en $x = a$.

Démonstration. Nous devons prouver que F est dérivable et que pour tout $x_0 \in]a,b[$ nous ayons $F'(x_0) = f(x_0)$. Soit $\epsilon > 0$. Par continuité de f en x_0, il existe une fonction $\alpha\colon \mathbb{R} \to \mathbb{R}$ telle que

$$f(x_0 + h) = f(x_0) + \alpha(h) \tag{13.564}$$

avec $\lim_{h\to 0} \alpha(h) = 0$. De plus il existe un $\delta > 0$ tel que $|\alpha(h)| < \epsilon$ pour tout $h < \delta$. À partir de maintenant nous ne considérons plus que de tels h.

Nous calculons la dérivée de F en x_0. Pour cela,

$$F(x_0 + h) - F(x_0) = \int_{x_0}^{x_0+h} f(t)dt \tag{13.565a}$$

$$= \int_0^h f(x_0 + t)dt \tag{13.565b}$$

$$= \int_0^h \big[f(x_0) + \alpha(t)\big]dt \tag{13.565c}$$

$$= hf(x_0) + \int_0^h \alpha(t)dt. \tag{13.565d}$$

Nous avons donc, pour tout $h < \delta$,

$$hf(x_0) - h\epsilon \leqslant F(x_0 + h) - F(x_0) \leqslant hf(x_0) + h\epsilon. \tag{13.566}$$

En divisant par h et en prenant la limite $h \to 0$,

$$F'(x_0) \in B\big(f(x_0), \epsilon\big). \tag{13.567}$$

Cela étant valable pour tout $\epsilon > 0$ nous en déduisons que

$$F'(x_0) = f(x_0). \tag{13.568}$$

Le fait que F s'annule en $x = a$ est par sa définition. L'unicité provient du corollaire 12.136. □

Le théorème suivant est à utiliser pour calculer des intégrales des fonctions réelle lorsqu'on a des primitives sur un domaine strictement plus large que le domaine sur lequel nous voulons intégrer. Une version pour les intégrales impropres sera donnée au corollaire 13.219.

Théorème 13.205 (Théorème fondamental du calcul intégral)**.**
Soit f une fonction continue sur un intervalle ouvert I contenant strictement l'intervalle $[a, b] \subset \mathbb{R}$ et F une primitive de f sur I. Alors

$$\int_a^b f(t)dt = F(b) - F(a). \tag{13.569}$$

Démonstration. Nous avons vu par la proposition 13.204 que la fonction

$$\begin{aligned} \tilde{F}\colon [a, b] &\to \mathbb{R} \\ x &\mapsto \int_a^x f(t)dt \end{aligned} \tag{13.570}$$

était une primitive de f ; c'est même l'unique[48] primitive de f sur $[a, b]$ à s'annuler pour $x = a$. Nous avons évidemment

$$\int_a^b f(t)dt = \tilde{F}(b). \tag{13.571}$$

Si F est une primitive quelconque, il suffit de soustraire sa valeur en $x = a$: $\tilde{F}(x) = F(x) - F(a)$ et donc

$$\int_a^b f(t)dt = \tilde{F}(b) = F(b) - F(a), \tag{13.572}$$

comme il fallait le prouver. □

Remarque 13.206.
Le lien entre primitive et intégrale est fondamentalement lié à l'invariance par translation de la mesure de Lebesgue, et non à la construction précise de cette mesure. Mais en même temps, la mesure de Lebesgue est l'unique à être invariante par translation.

Remarque 13.207.
Une primitive est forcément une fonction continue parce qu'une primitive est dérivable.

Ce petit résultat nous donne une façon «pratique» de calculer des intégrales en cherchant des primitives. Nous rappelons qu'en vertu du corollaire 12.136, une fonction ne possède qu'une seule primitive à constante près.

Corollaire 13.208 (Théorème fondamental du calcul intégral)**.**
Soit une fonction $f\colon I \to \mathbb{R}$ où I est un intervalle. Nous supposons que f est continue sur I. Soient $a < b$ dans I et une primitive F de f sur I.

Alors

$$\int_a^b f = F(b) - F(a). \tag{13.573}$$

48.

Démonstration. Nous savons par la proposition 13.204 que la fonction donnée par

$$G(x) = \int_a^x f \tag{13.574}$$

est une primitive de f. Vu que F est également une primitive nous avons $F(x) = G(x)+C$ pour une certaine constante $C \in \mathbb{R}$ (c'est le corollaire 12.136). Vu que $G(a) = 0$, l'égalité $F(a) = G(a) + C$ donne $C = F(a)$.

Ce que l'on nous demande de calculer dans (13.573) est $G(b)$, c'est à dire l'égalité

$$F(b) = G(b) + C = G(b) + F(a), \tag{13.575}$$

et donc $G(b) = F(b) - F(a)$, ce qu'il fallait. □

13.11.3 Exemples et applications

Si f est une fonction définie sur un intervalle I et y admettant des primitives, nous notons

$$\int f(x)dx \tag{13.576}$$

l'ensemble des primitives de f sur I :

$$\int f(x)dx = \{F(x) + C \text{ tel que } C \in \mathbb{R}\} \tag{13.577}$$

où F est une quelconque primitive de f.

Exemple 13.209
Une primitive bien connue de $f\colon x \mapsto x^2$ est la fonction $F\colon x \to \frac{x^3}{3}$. Nous écrivons donc

$$\int x^2 dx = \frac{x^3}{3} + C. \tag{13.578}$$

Cela est un abus de notations terrible pour dire en réalité

$$\{x \mapsto \frac{x^3}{3} + C \text{ tel que } C \in \mathbb{R}\}. \tag{13.579}$$

△

En termes de notations, nous posons

$$\int_a^b f(t)dt = \Big[F(t)\Big]_{t=a}^{t=b} = F(b) - F(a). \tag{13.580}$$

Remarque 13.210.
La valeur de l'intégrale ne dépend pas de la primitive qu'on choisi pour le calculer, car si F_1 et F_2 sont deux primitives de f alors $F_1 = F_2 + C$ et $F_1(b) - F_1(a) = (F_2(b) + C) - (F_2(a) + C) = F_2(b) - F_2(a)$.

Remarque 13.211.
Si l'intervalle d'intégration est réduit à un seul point alors la valeur de l'intégrale est zéro, car $\int_a^a f(t)dt = F(a) - F(a) = 0$.

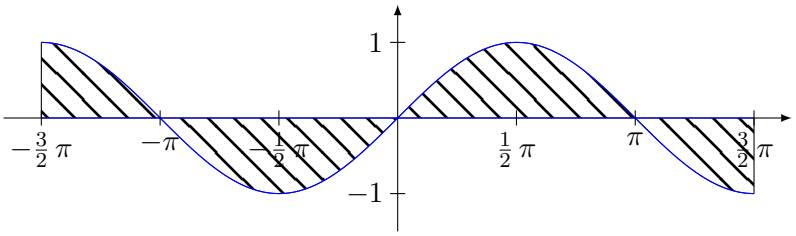
Remarque 13.212.
Conformément à ce que nous montre la figure 14.16, si une fonction continue est positive sur l'intervalle $[a, b]$, alors le nombre $\int_a^b f(t)dt$ est l'aire de la portion de plan comprise entre les droites verticales $x = a$, $x = b$, la courbe représentant la fonction f et l'axe des abscisses.

Si la fonction est négative : l'aire est comptée négativement.

Exemple 13.213
Comme nous le voyons sur le dessin suivant,

$$\int_{-3\pi/2}^{3\pi/2} \sin(x)\,dx = 0 \tag{13.581}$$

parce que les deux parties bleues s'annulent avec les deux parties rouges (qui sont comptées comme des aires négatives).



△

Remarque 13.214.
Toute intégrale d'une fonction impaire sur un intervalle symétrique par rapport à l'origine est nulle.

13.11.4 Intégrales impropres

Définition 13.215 ([120]).
*Une fonction $f\colon D \subset \mathbb{R} \to \mathbb{R}$ est **localement intégrable** sur un intervalle I si f est intégrable sur tout intervalle compact contenu dans I.*

Proposition 13.216.
Soit $f\colon [a,b] \to \mathbb{R}$ une fonction intégrable. Alors

$$\int_{[a,b]} f = \lim_{x\to b^-} \int_a^x f. \tag{13.582}$$

Démonstration. Notons que la valeur de f en b n'a strictement aucune importance parce que l'intégrale de Lebesgue ne dépend pas du choix de la valeur de la fonction en un ensemble de mesure nulle ; et en même temps la limite à gauche de (13.582) ne dépend pas non plus de la valeur de f en b. Bref si f n'est pas définie en b, nous pouvons poser $f(b) = 42$.

Notons de plus que du point de vue de l'intégrale de Lebesgue, $\int_{[a,b]}$ et $\int_{[a,b[}$ sont identiques et valent toutes les deux $\int_a^b$ (lorsque ça existe).

Supposons d'abord que f est positive. Alors nous posons $f_n = f\mathbb{1}_{[a,b-\frac{1}{n}]}$. Ponctuellement nous avons la limite croissante $f_n \to f$ et de plus

$$\lim_{x\to b^-} \int_{[a,x]} f = \lim_{n\to\infty} \int_{[a,b]} f_n. \tag{13.583}$$

Chacun des f_n est intégrable sur $[a,b]$. Le théorème de Beppo-Levi 13.129 implique que f est intégrable sur $[a,b]$ et que

$$\lim_{n\to\infty} \int_a^b f_n = \int_a^b f. \tag{13.584}$$

Cela montre que dans le cas d'une fonction f positive nous avons bien (13.582).

Si f n'est pas positif, alors nous la décomposons en partie positive et négative $f = f^+ - f^-$ et par définition de l'intégrale d'une fonction non positive,

$$\lim_{x\to b^-} \int_{[a,x[} f = \lim \int f^+ - \lim \int f^-. \tag{13.585}$$

□

Il peut cependant arriver que la limite $\lim_{x\to b}\int_a^b f$ existe alors que f n'est pas intégrable sur $[a,b]$. C'est l'ennui des fonctions non positives. Un exemple classique est

$$\int_0^\infty \frac{\sin(t)}{t}dt \tag{13.586}$$

Définition 13.217 ([184])**.**
Si

$$\lim_{x\to b}\int_a^b f \tag{13.587}$$

existe alors nous disons que l'intégrale est ***convergente*** *en b. Ce procédé de limite est l'intégrale* ***impropre*** *de f sur $[a,b]$.*

Exemple 13.218(Intégale impropre)
Nous considérons la fonction $f\colon [0,\infty[\to\mathbb{R}$ définie par

$$f(x)=\begin{cases}\frac{1}{n} & \text{si } x\in[2n-2,2n-1[\\ -\frac{1}{n} & \text{si } x\in[2n-1,2n[.\end{cases} \tag{13.588}$$

Par la divergence de la série harmonique, $\int_0^\infty |f|$ n'existe pas. La fonction f n'est donc pas intégrable au sens de Lebesgue (définition 13.116).

Cependant pour tout n pair nous avons

$$\int_0^n f=0. \tag{13.589}$$

Du coup pour tout $x\geqslant 0$ nous avons

$$\int_0^x f=\int_{2n}^x f \tag{13.590}$$

où $2n$ est le plus grand nombre pair inférieur à x. Nous avons $|x-2n|\leqslant 2$ et $|f(x)|\leqslant\frac{1}{n}$ pour $x\in[2n,x]$. Donc

$$\int_{2n}^x f\leqslant\frac{2}{n}. \tag{13.591}$$

Nous avons par conséquent

$$\lim_{x\to\infty}\int_0^x f=0, \tag{13.592}$$

ce qui signifie que l'intégrale de f sur $[0,\infty[$ converge au sens des intégrales impropres. △

L'intégrale (13.586) est une intégrale convergente mais la fonction n'est pas intégrable (parce que pour être intégrale il faut que $|f|$ soit intégrable). Nous pouvons ainsi dire que cette intégrale converge mais n'existe pas.

Le corollaire suivant nous autorise à utiliser le théorème fondamental du calcul intégral 13.205 même dans les cas limites.

Corollaire 13.219.
Si f est localement intégrable sur $[a,b]$ et si F est une primitive de f sur tout ouvert de $[a,b]$ alors

$$\int_a^b f=\lim_{x\to b^-}F(x)-F(a). \tag{13.593}$$

Démonstration. Pour chaque x dans $[a,b[$ nous avons

$$\int_a^x f=F(x)-F(b). \tag{13.594}$$

La proposition 13.216 nous explique que la limite $x \to b^-$ du membre de gauche existe et vaut $\int_a^b f$. Donc également le membre de droite :

$$\int_a^b f = \lim_{x\to b^-} \int_a^x f = \lim_{x\to b^-} F(x) - F(b). \tag{13.595}$$

□

La convergence des intégrales de fonctions $\frac{1}{x^\alpha}$ en 0 et ∞ est une question classique de l'intégration. De plus ces fonctions servent souvent à utiliser une théorème de comparaison (type intégrale dominée de Lebesgue).

Proposition 13.220.
Deux intégrales remarquables.

(1) Nous avons

$$\int_0^1 \frac{1}{x^\alpha} = \infty \tag{13.596}$$

si et seulement si $\alpha \geqslant 1$.

(2) Nous avons

$$\int_1^\infty \frac{1}{x^\alpha} = \infty \tag{13.597}$$

si et seulement si $\alpha \leqslant 1$.

Démonstration. La fonction $\frac{1}{x^\alpha}$ admet la primitive $F(x) = \frac{1}{1-\alpha}\frac{1}{x^{\alpha-1}}$ sur tout compact de $]0,\infty[$. Le corollaire 13.219 nous permet [49] de dire que $\int_0^1 \frac{1}{x^\alpha}$ vaudra

$$\lim_{x\to 0-^+} \frac{1}{1-\alpha}\frac{1}{x^{\alpha-1}}. \tag{13.598}$$

Cela est strictement plus petit que ∞ si et seulement si $\alpha < 1$. □

13.12 Changement de variables dans une intégrale multiple

Dans ce qui suit, U et V sont des ouverts de $\mathbb{R}^N$ et $\phi\colon U \to V$ est un C^1-difféomorphisme. Nous notons $\mathcal{Q}$ l'ensemble des cubes fermés dans U dont les cotés sont parallèles aux axes.

13.12.1 Des lemmes

Lemme 13.221 ([183])**.**
Soient μ et ν deux mesures de Borel sur l'ouvert U de $\mathbb{R}^N$. Si $\mu(Q) \leqslant \nu(Q)$ pour tout $Q \in \mathcal{Q}$ alors $\mu(B) \leqslant \nu(B)$ pour tout borélien B.

Démonstration. Si Q est un cube semi-ouvert, c'est à dire de la forme

$$Q = \prod_{i=1} N[a_n, a_n + h[\subset U \tag{13.599}$$

alors Q est une réunion croissante de cubes fermés du type $[a_n+\epsilon, a_n+h-\epsilon]$, et donc $\mu(Q) \leqslant \nu(Q)$ par le lemme 13.27(1). La propriété est donc vraie pour les cubes semi-ouverts.

Si Ω est un ouvert, alors il est réunion disjointe dénombrable de cubes semi-ouverts par la proposition 13.191. Donc pour tout ouvert $\Omega \subset U$ nous avons $\mu(\Omega) \leqslant \nu(\Omega)$. En vertu de la proposition 13.79 et de la remarque 13.80, les mesures μ et ν sont régulières, et l'inégalité au niveau des ouverts se répercute en inégalité pour tout boréliens de U :

$$\mu(B) \leqslant \nu(B) \tag{13.600}$$

pour tout $B \in \mathcal{B}\text{or}(U)$. Notons que U étant ouvert dans $\mathbb{R}^N$, les boréliens de U sont exactement les boréliens de $\mathbb{R}^N$ inclus à U par le corollaire 13.74. □

49. Tout ce que nous avons fait avec la borne b de l'intégrale $\int_a^b$ reste valable avec la borne a.

Lemme 13.222 ([183]).
Soit une application $\theta\colon U \to \mathbb{R}^N$ de classe C^1 où U est ouvert dans $\mathbb{R}^N$. Pour tout $Q \in \mathcal{Q}$ nous avons

$$\lambda_N\big(\theta(Q)\big) \leqslant \sup_{s\in Q} \|d\theta_s\|^N \lambda_N(Q). \tag{13.601}$$

Démonstration. Nous notons h la longueur du côté du cube. Le théorème des accroissements finis 12.241, pour la composante θ_i donne, pour $u, v \in Q$:

$$\big|\theta_i(u) - \theta_i(v)\big| \leqslant \sup_{s\in Q} \|(d\theta_i)_s\| \|u - v\| \leqslant \sum_{s\in Q} \|(d\theta_i)_s\| h. \tag{13.602}$$

D'autre part nous avons (nous écrivons pour $N = 2$ pour être plus court) :

$$d\theta_s(u) = \frac{d}{dt}\Big[\theta_1(s + tu)e_1 + \theta_2(s + tu)e_2\Big]_{t=0} = (d\theta_1)_s(u)e_1 + (d\theta_2)_s(u)e_2. \tag{13.603}$$

Donc pour chaque i : $\|d\theta_s\| \geqslant \|(d\theta_i)_s\|$, et nous continuons la majoration (13.602) :

$$\big|\theta_i(u) - \theta_i(v)\big| \leqslant\leqslant \sum_{s\in Q} \|(d\theta_i)_s\| h \leqslant \sup_{s\in Q} \|d\theta_s\| h. \tag{13.604}$$

Les points $\theta(u)$ et $\theta(v)$ sont donc dans un cube de côté $\sup_{s\in Q} \|d\theta_s\| h$, ce qui permet de majorer $\lambda_N\big(\theta(Q)\big)$ par

$$\lambda_N\big(\theta(Q)\big) \leqslant \left(\sup_{s\in Q} \|d\theta_s\| h\right)^N = \left(\sup_{s\in Q} \|d\theta_s\|\right)^N \lambda_N(Q) \tag{13.605}$$

où le dernier facteur provient de l'égalité $h^N = \lambda_N(Q)$. □

13.12.2 Déterminant et mesure de Lebesgue

Dans la suite, Q_0 désigne le cube unité : $Q_0 = \big([0, 1[\big)^N$.

Théorème 13.223 (Interprétation géométrique du déterminant[183]).
Soit une application linéaire $T\colon \mathbb{R}^N \to \mathbb{R}^N$. Alors pour tout borélien B de $\mathbb{R}^N$,

$$\lambda_N\big(T(B)\big) = |\det(T)| \lambda_N(B). \tag{13.606}$$

Démonstration. Nous considérons la mesure positive μ donnée par $\mu(B) = \lambda_N\big(T(B)\big)$, qui est bien une mesure par la proposition 13.75. Cette mesure est invariante par translation parce que λ_N l'est :

$$\mu(B + a) = \lambda_N\big(T(B) + a\big) = \lambda_N\big(T(B)\big) = \mu(B). \tag{13.607}$$

De plus, $T(Q_0)$ est borné et nous notons $\mu(Q_0) = C$. Nous avons $\mu = C\lambda_N$ par le corollaire 13.201.

Underline{$C(T_1T_2) = C(T_1)C(T_2)$} Par définition,

$$C(T_1T_2)\lambda_N(B) = \lambda_N\big((T_1T_2)(B)\big) = \lambda_N\big(T_1(T_2B)\big) = C(T_1)\lambda_N\big(T_2(B)\big) = C(T_1)C(T_2)\lambda_N(B). \tag{13.608}$$

Par conséquent la fonction C est multiplicative :

$$C(T_1T_2) = C(T_1)C(T_2). \tag{13.609}$$

Et en plus, $C(\mathrm{Id}) = 1$.

Matrice diagonale Nous considérons pour $T = D$ l'application linéaire diagonale $D = \mathrm{diag}(d_1, \ldots, d_N)$ qui fait

$$T(Q_0) = [0, d_1[\times \ldots \times [0, d_N[\tag{13.610}$$

La mesure de cela est $|d_1 \cdots d_N|$, ce qui nous donne

$$C(D) = |d_1 \ldots d_N| = |\det(D)|. \tag{13.611}$$

Matrice orthogonale Nous considérons maintenant $T = U$ où U est une matrice orthogonale ($UU^t = 1$). Une matrice orthogonale est une isométrie [50] qui conserve donc la boule unité : $UB(0,1) = B(0,1)$. Nous avons

$$\lambda_N\big(B(0,1)\big) = \lambda_N\big(UB(0,1)\big) = C(U)\lambda_N\big(B(0,1)\big) \tag{13.612}$$

par conséquent $C(U) = 1$, et 1 est justement le déterminant de U.

Matrice quelconque Nous savons par le corollaire 12.348 de la décomposition polaire que toute matrice peut être écrite sous la forme $T = U_1DU_2$ où U_i sont orthogonales et D est diagonale. Donc $C(T) = C(U_1)C(D)C(U_2) = \det(U_1)\det(D)\det(U_2) = \det(U_2DU_2) = \det(T)$ parce que le déterminant est multiplicatif (proposition 9.10(1)).

□

Ce théorème donne une interprétation géométrique du déterminant en tant que facteur de dilatation des volumes lors de l'utilisation d'une application linéaire. Si T est une application linéaire quelconque,

$$\lambda_N\big(T(Q_0)\big) = |\det(T)|\lambda_N(Q_0) = |\det(T)|. \tag{13.613}$$

Le déterminant de T est le volume de l'image du cube unité par l'application T.

De la même façon, en utilisant l'application linéaire $T(x) = ax$ nous avons pour tout borélien B :

$$\lambda_N(aB) = a^N\lambda_N(B). \tag{13.614}$$

Une dilatation d'un facteur a des longueurs provoque une multiplication par a^N des volumes.

13.12.3 Le théorème et sa démonstration

Théorème 13.224 (Changement de variable[182, 183])**.**
Soient U et V des ouverts de $\mathbb{R}^N$ ainsi qu'un C^1-difféomorphisme $\phi\colon U \to V$.

(1) Si $E \subset U$ est borélien, alors $\phi(E)$ est borélien et

$$\lambda_N\big(\phi(E)\big) = \int_E |J_\phi| d\lambda_N, \tag{13.615}$$

c'est à dire $\phi^{-1}(\lambda_N) = |J_\phi| \cdot \lambda_N$.

(2) Si $f\colon V \to [0,+\infty]$ est mesurable alors la fonction

$$(f \circ \phi) \times |J_\phi|\colon U \to [0,\infty] \tag{13.616}$$

l'est également et [51]

$$\int_V f d\lambda_N = \int_U (f \circ \phi)(x)|J_\phi(x)| d\lambda_N(x). \tag{13.617}$$

(3) Si $f\colon V \to \mathbb{C}$ est mesurable alors elle est intégrable si et seulement si $(f \circ \phi) \times |J_\phi|\colon U \to \mathbb{C}$ est intégrable. Si c'est le cas, alors nous avons encore la formule de changement de variables :

$$\int_V f d\lambda_N = \int_U (f \circ \phi)|J_\phi| d\lambda_N. \tag{13.618}$$

Démonstration. Attention : la preuve va être longue.

(1) Le fait que $\phi(E)$ soit borélien lorsque E l'est est la proposition 13.196. En ce qui concerne la formule annoncée, il faut travailler.

50. Proposition 9.50.

51. L'intégrabilité d'une fonction est la définition 13.116 qui stipule que l'intégrale de $|f(x)|$ est finie. L'égalité proposée a un sens si les deux membres sont infinis. Il n'y a donc pas d'hypothèses d'intégrabilité obligatoire pour écrire une intégrale lorsque la fonction a des valeurs positives.

Inégalité dans un sens (cubes) Nous commençons par prouver l'inégalité

$$\lambda_N\big(\phi(Q)\big) \leqslant \int_Q |J_\phi(x)|dx \tag{13.619}$$

pour tout $Q \in \mathcal{Q}$. On peut diviser le côté du cube Q en k éléments de longueurs égales. Le cube est alors divisé en k^N petits cubes d'intérieurs disjoints. Nous les nommons Q_i ($i = 1, \dots, k^N$) Nous avons alors

$$\sum_i \lambda_N(Q_i) = \sum_i \lambda_N\big(\operatorname{Int}(Q_i)\big) = \lambda_N\big(\bigcup_i \operatorname{Int}(Q_i)\big) \leqslant \lambda_N(Q) \leqslant \sum_i \lambda_N(Q_i). \tag{13.620}$$

La dernière inégalité est le fait que les intersections ne sont pas disjointes. Toutes ces inégalités sont en réalité des égalités et en particulier : $\lambda_N(Q) = \sum_i \lambda_N(Q_i)$.
Soit $a \in Q_i$ et posons

$$\begin{aligned} &\theta\colon U && \to U \\ &\theta = (d\phi_a)^{-1} \circ \phi \end{aligned} \tag{13.621}$$

Cela appelle deux commentaires. D'abord l'application $d\phi_a\colon U \to V$ est inversible parce que ϕ est un difféomorphisme (lemme 12.223). Ensuite, l'application θ est la composée de $(d\phi_a)$ (qui est linéaire) et de ϕ qui est de classe C^1 ; donc θ est de classe C^1. Donc le lemme 13.222 s'applique. La différentielle de θ n'est pas trop compliquée à écrire parce que nous avons la formule de différentielle d'une composée (théorème 12.220) et le fait que $(d\phi_a)^{-1}$ qui est linéaire et donc sa propre différentielle (lemme 12.217). Nous avons donc $d\theta = (d\phi_a)^{-1} \circ d\phi$, et le lemme donne

$$\lambda_N\left((d\phi_a)^{-1}\phi(a)\right) \leqslant \sup_{s\in Q_i} \|(d\phi_a)^{-1} \circ d\phi_s\|^N \lambda_N(Q_i) \tag{13.622}$$

Étant donné que $(d\phi_a)^{-1}$ est une application linéaire, la proposition 13.223 s'applique, et donc

$$\lambda_N\left((d\phi_a)^{-1}\phi(a)\right) = |\det(d\phi_a)^{-1}|\lambda_N\big(\phi(a)\big). \tag{13.623}$$

Le déterminant d'une application réciproque est donné par la proposition 9.10(4) :

$$\det\left((d\phi_a)^{-1}\right) = \frac{1}{\det\left(d\phi_a\right)} = \frac{1}{J_\phi(a)}. \tag{13.624}$$

Recollant les morceaux,

$$\lambda_N\big(\phi(Q_i)\big)\frac{1}{J_\phi(a)} \leqslant \sup_{s\in Q_i} \|(d\phi_a)^{-1} \circ d\phi_s\|^N \lambda_N(Q_), \tag{13.625}$$

ou encore :

$$\lambda_N\big(\phi(Q_i)\big) \leqslant |J_\phi(a)| \sup_{s\in Q_i} \|(d\phi_a)^{-1} \circ d\phi_s\|^N \lambda_N(Q_i). \tag{13.626}$$

Vu que a et s sont proches l'un de l'autre (on peut choisir encore la taille du cube), nous pouvons espérer que $(d\phi_a)^{-1}$ ne soit pas loin d'être l'inverse de $d\phi_s$. Et c'est en effet le cas. Pour s'en assurer, remarquons que l'application

$$d\phi\colon Q_i \to \mathcal{L}(\mathbb{R}^N, \mathbb{R}^N) \tag{13.627}$$

est continue et même uniformément continue parce que Q_i est compact. De plus la composition de différentielles étant un produit de matrices nous pouvons permuter la limite dans le calcul suivant :

$$\lim_{s\to a}(d\phi_a)^{-1} \circ d\phi_s = (d\phi_a)^{-1} \circ \lim_{s\to a} d\phi_s = \mathbb{1}. \tag{13.628}$$

Donc si $\epsilon > 0$ est donné, il existe δ tel que pour tout $s \in B(a, \delta)$, $\|(d\phi_a)^{-1} \circ d\phi_s - \mathbb{1}\| \leqslant \epsilon$. En ce qui concerne les normes, si $\|A - \mathbb{1}\| \leqslant \epsilon$ alors $\|A\| \leqslant \|A - \mathbb{1}\| + \|\mathbb{1}\| \leqslant \epsilon + 1$.

Cela étant dit, nous nous souvenons que nous avions découpé U en un nombre fini de cubes Q_i d'égales dimensions ; il suffit de prendre k suffisamment grand pour que la diagonale des cubes sot plus petite que le minimum des δ_i. Avec un tel découpage,

$$\sup_{s\in Q_i} \|(d\phi_a)^{-1} \circ d\phi_s\| \leqslant 1 + \epsilon \tag{13.629}$$

et par conséquent

$$\lambda_N\big(\phi(Q_i)\big) \leqslant (1+\epsilon)^N |J_\phi(a_i)| \lambda_N(Q_i) \tag{13.630}$$

où nous avons ajouté un indice i au point a pour nous rappeler que nous avons choisit $a \in Q_i$.

Le théorème de la moyenne 13.203 appliqué à l'intégrale $\int_{Q_i} |J_\phi(t)| d\lambda_N(t)$ donne l'existence d'un $a_i \in Q_i$ tel que

$$|J_\phi(a_i)| = \frac{1}{\lambda_N(Q_i)} \int_{Q_i} |J_\phi| d\lambda_N. \tag{13.631}$$

Ce point a_i vérifie l'inégalité (13.630) comme tout point de Q_i. Nous sommons ces inégalités sur tous les i :

$$\begin{aligned}
\lambda_N\big(\phi(Q)\big) &\leqslant \sum_i \lambda_N\big(\phi(Q_i)\big) && \text{(13.632a)}\\
&\leqslant (1+\epsilon^N \sum_i \left(\frac{1}{\lambda_N(Q_i) \int_{Q_i} |J_\phi| d\lambda_N} \right) \lambda_N(Q_i) && \text{(13.632b)}\\
&= (1+\epsilon)^N \sum_i \int_{Q_i} |J_\phi| d\lambda_N && \text{(13.632c)}\\
&= (1+\epsilon)^N \int_Q |J_\phi| d\lambda_N && \text{(13.632d)}
\end{aligned}$$

où nous avons utilisé le fait que $\mathbb{1}_Q = \sum_i \mathbb{1}_{Q_i}$ presque partout. En prenant le limite $\epsilon \to 0$ nous trouvons

$$\lambda_N\big(\phi(Q)\big) \leqslant \int_Q |J_\phi| d\lambda_N. \tag{13.633}$$

L'inégalité (13.619) est prouvée.

Inégalité pour les boréliens Soit B un borélien de U. Vu que U et V sont des ouverts de $\mathbb{R}^N$, les mesures de Lebesgue sur U et sur V sont les mêmes que celles sur $\mathbb{R}^n$ par le corollaire 13.74.

Par les définitions 13.137 et 13.75, les applications μ et n définies par $\mu = \phi^{-1}(\lambda_N)$ et $\nu = |J_\phi|\lambda_N$ sont des mesures positives sur U (de Borel, qui plus est). L'inégalité (13.619) à peine prouvée s'écrit $\mu(Q) \leqslant \nu(Q)$ pour tout cube Q. Le lemme 13.221 nous dit alors que l'inégalité tient pour tout borélien.

Inégalité dans l'autre sens En utilisant la notation de la mesure image et du produit d'une mesure par une fonction [52], nous pouvons écrire l'inégalité prouvée sous la forme $\phi^{-1}(\lambda_N) \leqslant |J_\phi|\lambda_N$. En inversant les rôles de U et V (et donc de ϕ et ϕ^{-1}) nous avons aussi

$$\phi(\lambda_N) \leqslant |J_{\phi^{-1}}|\lambda_N. \tag{13.634}$$

En y appliquant ϕ^{-1} et le lemme 13.75,

$$\lambda_N \leqslant \phi^{-1}\big(|J_{\phi^{-1}}|\lambda_N\big). \tag{13.635}$$

52. Définition 13.75 et 13.137

Nous prouvons à présent que $\phi^{-1}\big(|J_{\phi^{-1}}|\cdot\lambda_N\big)=\Big(|J_{\phi^{-1}}|\circ\phi\Big)\cdot\phi^{-1}(\lambda_N)$ en appliquant à un borélien B de U. D'une part

$$\phi^{-1}\big(|J_{\phi^{-1}}|\cdot\lambda_N\big)(B)=\big(|J_{\phi^{-1}}|\cdot\lambda_N\big)\phi(B) \tag{13.636a}$$

$$=\int_{\phi(B)}|J_{\phi^{-1}}|d\lambda_N, \tag{13.636b}$$

et d'autre part,

$$\Big(|J_{\phi^{-1}}|\circ\phi\Big)\cdot\phi^{-1}(\lambda_N)B=\int_{\mathbb{R}^N}\mathbb{1}_B(x)\big(|J_{\phi^{-1}}|\circ\phi\big)(x)d\big(\phi^{-1}(\lambda_N)\big)(x) \tag{13.637a}$$

$$=\int_{\mathbb{R}^N}\mathbb{1}_B\big(\phi^{-1}(x)\big)\big(|J_{\phi^{-1}}|\circ\phi\big)\big(\phi^{-1}(x)\big)d\lambda_N(x) \tag{13.637b}$$

$$=\int_{\mathbb{R}^N}\mathbb{1}_{\phi(B)}|J_{\phi^{-1}}| \tag{13.637c}$$

$$=\int_B|J_{\phi^{-1}}|d\lambda_N. \tag{13.637d}$$

Justification :
— Pour (13.637b), le théorème 13.141(2).
L'équation (13.635) devient alors

$$\lambda_N\leqslant\Big(|J_{\phi^{-1}}|\circ\phi\Big)\cdot\phi^{-1}(\lambda_N). \tag{13.638}$$

Nous allons faire le produit de cette mesure par $|J_\phi|$ en nous souvenant que $J_\phi(x)=\det\big(d\phi_x\big)$. Par le lemme 12.223 nous avons aussi $(d\phi_x)^{-1}=d\phi^{-1}_{\phi(x)}$ et donc, par la propriété 9.10(3) du déterminant,

$$J_\phi(x)=\frac{1}{\det\big(d\phi^{-1}_{\phi(x)}\big)}=\frac{1}{J_{\phi^{-1}}\big(\phi(x)\big)}. \tag{13.639}$$

Nous avons

$$|J_\phi|\cdot\lambda_N\leqslant|J_\phi|\cdot\Big(|J_{\phi^{-1}}|\circ\phi\Big)\cdot\phi^{-1}(\lambda_N). \tag{13.640}$$

En utilisant la proposition 13.138, il s'agit de multiplier la mesure $\phi^{-1}(\lambda_N)$ par la fonction

$$x\mapsto|J_\phi(x)J_{\phi^{-1}}\big(\phi(x)\big)|=1. \tag{13.641}$$

Nous avons donc bien

$$|J_\phi|\cdot\lambda_N\leqslant\phi^{-1}(\lambda_N), \tag{13.642}$$

et donc l'égalité

$$|J_\phi|\cdot\lambda_N=\phi^{-1}(\lambda_N), \tag{13.643}$$

c'est à dire le point (1).

(2) Le fait que la fonction proposée soit mesurable est le fait que la mesurabilité n'est pas affectée par produit et composition (propositions 13.96 et 13.66), et le fait que pour les mêmes raisons, l'application $J_\phi\colon U\to\mathbb{R}$ est également mesurable. En ce qui concerne la formule nous allons la démontrer dans le cas de fonctions de plus en plus générales.

Pour les fonctions indicatrices Soit B un borélien de U, et considérons la fonction $f=\mathbb{1}_{\phi(B)}$. Alors

$$\int_V fd\lambda_N=\int_{\mathbb{R}^N}\mathbb{1}_{\phi(B)}(y)\mathbb{1}_V(y)d\lambda_N(y)=\int_{\mathbb{R}^N}\mathbb{1}_{\phi(B)}d\lambda_N=\lambda_N\big(\phi(B)\big). \tag{13.644}$$

parce que $V=\phi(U)$ et $B\subset U$, donc $\mathbb{1}_{\phi(B)}\mathbb{1}_{\phi(U)}=\mathbb{1}_{\phi(B)}$. D'autre part, pour calculer l'autre membre de (13.617) nous remarquons que $f=\mathbb{1}_{\phi(B)}=\mathbb{1}_B\circ\phi^{-1}$, ce qui donne

$$\int_U f\big(\phi(x)\big)|J_\phi(x)|d\lambda_N(x)=\int_U\mathbb{1}_B|J_\phi|d\lambda_N=\int_B|J_\phi|d\lambda_N. \tag{13.645}$$

L'ensemble B étant borélien, il est extrêmement mesurable, ce qui fait que le point (1) s'applique : les expressions (13.644) et (13.645) sont égales.

Pour les fonctions étagées Soit $f\colon V \to \mathbb{R}^+$ une fonction étagée :

$$f(x) = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}(x) \tag{13.646}$$

Nous pouvons faire le calcul suivant :

$$\begin{aligned}
\int_V f d\lambda_N &= \int_V \sum_i a_i \mathbb{1}_{A_i} d\lambda_N && \text{(13.647a)}\\
&= \sum_i a_i \int_V \mathbb{1}_{A_i} d\lambda_N && \text{(13.647b)}\\
&= \sum_i \int_U (\mathbb{1}_{a_i} \circ \phi)(x) |J_\phi(x)| d\lambda_N(x) && \text{(13.647c)}\\
&= \sum_i a_i \int_U \mathbb{1}_{\phi^{-1}(A_i)} |J_\phi(x)| d\lambda_N(x) && \text{(13.647d)}\\
&= \int_V \underbrace{\sum_i a_i \mathbb{1}_{\phi^{-1}(A_i)}(x)}_{=(f\circ\phi)(x)} |J_\phi(x)| d\lambda_N(x) && \text{(13.647e)}\\
&= \int_V (f \circ \phi)|J_\phi| d\lambda_N. && \text{(13.647f)}
\end{aligned}$$

Justifications :
— Pour (13.647b) : linéarité de l'intégrale, théorème 13.119(2) [53]
— Pour (13.647c) : le cas des fonctions indicatrices est utilisé pour chaque i entre 1 et n.

Fonction mesurable positive Soit $f\colon V \to [0,\infty]$. Par le théorème fondamental d'approximation 13.107, il existe une suite croissante de fonctions étagées et mesurables $\varphi_n\colon V \to [0,\infty[$ dont la limite ponctuelle est f. Nous avons alors le calcul suivant :

$$\begin{aligned}
\int_V f d\lambda_N &= \lim_{n\to\infty} \int_V \varphi_n d\lambda_N && \text{(13.648a)}\\
&= \lim_{n\to\infty} \int_U (\varphi_n \circ \phi)|J_\phi| d\lambda_N && \text{(13.648b)}\\
&= \int_U \lim_{n\to\infty} (\varphi_n \circ \phi)|J_\phi| d\lambda_N && \text{(13.648c)}\\
&= \int_U (f \circ \phi)|J_\phi| d\lambda_N. && \text{(13.648d)}
\end{aligned}$$

Justifications :
— Pour (13.648a), c'est le théorème de la convergence monotone 13.129.
— Pour (13.648b), c'est le présent théorème pour la fonction étagée φ_n.
— Pour (13.648c), c'est encore la convergence dominée, justifiée par le fait que $\varphi_n \circ \phi$ est également une suite croissante : si $x \in U$ alors $\varphi_{n+1}\big(\phi(x)\big) \geqslant \varphi_n\big(\phi(x)\big)$.
— Pour (13.648d), c'est la limite ponctuelle $\varphi_n\big(\phi(x)\big) \to f\big(\phi(x)\big)$.

(3) La partie sur l'intégrabilité repose sur le fait que $|f| \circ \phi = |f \circ \phi|$. Ici $|.|$ est le module et non une valeur absolue. Les faits suivants sont équivalents :
— la fonction $f\colon V \to \mathbb{C}$ est intégrable
— la fonction $|f|\colon V \to \mathbb{R}$ est intégtrable
— la fonction $(|f| \circ \phi)|J_\phi|\colon U \to \mathbb{R}$ est intégrable (par le point (2)).
— la fonction $(f \circ \phi)|J_\phi|\colon U \to \mathbb{R}$ est intégrable.

53. Il est remarquable que nous n'utilisons cette linéarité que pour les fonction étagées.

En ce qui concerne la formule, il s'agit seulement d'appliquer le point (2) aux parties positives, négatives, imaginaires et réelles de f.

□

Notons que la formule peut être écrite sous la forme

$$\langle f, g\rangle_V = \langle f \circ \phi, (g \circ \phi)|J|\rangle_U, \tag{13.649}$$

qui est plus pratique lorsqu'on parle de produits scalaires. Pour rappel, $\phi\colon U \to C$ est un C^1-difféomorphisme.

13.225.
La formule de changement de variables peut être comprise de la façon suivante. Si ϕ est linéaire alors le facteur $|J_\phi|$ est la mesure de l'image par ϕ d'une portion de $\mathbb{R}^p$ de mesure 1, sinon $|J_\phi|$ est le rapport entre la mesure de l'image d'un élément infinitésimale de volume de $\mathbb{R}^p$ et sa mesure originale.

Soit $\phi(u, v) = g(u, v)e_1 + h(u, v)e_2$ un difféomorphisme dans $\mathbb{R}^2$. Soit (x_0, y_0) l'image par ϕ de (u_0, v_0). On considère le petit rectangle R de sommets (u_0, v_0), $(u_0 + \Delta u, v_0)$, $(u_0 + \Delta u, v_0 + \Delta v)$ et $(u_0, v_0 + \Delta v)$. L'image de R n'est pas un rectangle en général, mais peut être bien approximée par le rectangle de sommets (x_0, y_0), $(x_0, y_0) + \phi_u\Delta u$, $(x_0, y_0) + \phi_u\Delta u + \phi_v\Delta v$ et $(x_0, y_0) + \phi_v\Delta v$ et son aire est $\|\phi_u \times \phi_v\|\Delta u\Delta v$. La valeur $|\phi_u \times \phi_v|$ est exactement $|J_\phi|$

13.12.4 Exemples

Exemple 13.226
Soit V la région trapézoïdale de sommets $(0, -1)$, $(1, 0)$, $(2, 0)$, $(0, -2)$, comme à la figure 13.1(a). Calculons ensemble l'intégrale double

$$\int_V e^{\frac{x+y}{x-y}}\, dV,$$

avec le changement de variable $\psi(x, y) = (x + y, x - y)$. C'est à dire que nous considérons les nouvelles variables

$$\begin{cases} u = x + y & \text{(13.650a)} \\ v = x - y. & \text{(13.650b)} \end{cases}$$

Il faut remarquer d'abord que le changement de variable proposé est dans le mauvais sens. On écrit alors $\phi(u, v) = \psi^{-1}(u, v) = \big((u + v)/2, (u - v)/2\big)$, c'est à dire

$$\begin{cases} x = \dfrac{u + v}{2} & \text{(13.651a)} \\ y = \dfrac{u - v}{2}. & \text{(13.651b)} \end{cases}$$

La région qui correspond à V est U, le trapèze de sommets $(-1, 1)$, $(1, 1)$, $(2, 2)$ et $(-2, 2)$, qu'on voit sur la figure 13.1(b) et qu'on décrit par

$$U = \{(u, v) \in \mathbb{R}^2 \,|\, 1 \leqslant v \leqslant 2,\ -v \leqslant u \leqslant v\}.$$

On observe que U est une région du premier type tandis que V n'est pas du premier ou du deuxième type. Le déterminant de la matrice jacobienne de ψ^{-1} est $J_{\psi^{-1}}$,

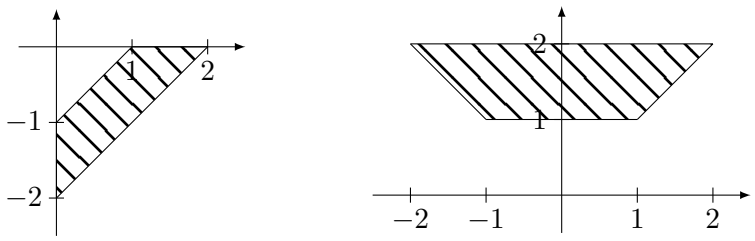
$$J_{\psi^{-1}}(u, v) = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}. \tag{13.652}$$

On a alors

$$\int_V e^{\frac{x+y}{x-y}}\, dV = \int_U e^{\frac{u}{v}}\, \frac{1}{2}\, dV = \int_1^2 \int_{-v}^{v} e^{\frac{u}{v}}\, \frac{1}{2}\, du\, dv = \frac{3}{4}(e - e^{-1}).$$

△

Énormément d'exemples sont disponibles avec les coordonnées polaires et toutes leurs variations. Cependant les fonctions trigonométriques ne seront vues que plus tard. Ce sera pour 16.7.2.



(a) La région V (b) La région $U = \phi^{-1}(V)$

FIGURE 13.1 – Avant et après le changement de variables

13.13 Théorème de Fubini-Tonelli et de Fubini

Nous rappelons que $\mathbb{R}^n$ muni de la mesure de Lebesgue est un espace mesuré σ-fini, conformément à la définition 13.23.

Le théorème de Fubini-Tonelli parle de fonctions réelles et non complexes, et même positives. Le truc est que ce théorème va servir de base pour construire les autres. Si nous avons une fonction à valeurs complexes, elle se décompose en parties réelles et imaginaires qui elles-mêmes se décomposent en parties positives et négatives. Au final, les preuves pour $f : \Omega \to \mathbb{C}$ se ramènent à appliquer quatre fois le théorème pour $f : \Omega \to \bar{\mathbb{R}}^+$.

Théorème 13.227 (Fubini-Tonelli[167])**.**
Soient $(\Omega_i, \mathcal{A}_i, \mu_i)$ deux espaces mesurés σ-finis, et $(\Omega, \mathcal{A}, \mu)$ l'espace produit. Soit une fonction $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ une fonction mesurable et positive (valant éventuellement ∞ à certains endroits) Alors :

(1) Les fonction

$$F_1 : x \mapsto \int_{\Omega_2} f(x,y) d\mu_2(y) \tag{13.653}$$

et

$$F_2 : y \mapsto \int_{\Omega_1} f(x,y) d\mu_1(x) \tag{13.654}$$

sont mesurables.

(2) Toutes les intégrales imaginables existent et sont égales :

$$\int_{\Omega_1 \times \Omega_2} f(x,y) d(\mu_1 \otimes \mu_2)(x,y) = \int_{\Omega_1} \left[\int_{\Omega_2} f(x,y) d\mu_2(y) \right] d\mu_1(x) \tag{13.655a}$$

$$= \int_{\Omega_2} \left[\int_{\Omega_1} f(x,y) d\mu_1(x) \right] d\mu_2(y) \tag{13.655b}$$

où tous les membres de l'égalité valent éventuellement $+\infty$.

Démonstration. Commençons par prouver le théorème dans le cas d'une fonction caractéristique d'un ensemble mesurable : $f(x,y) = \mathbb{1}_A(x,y)$ pour un certain ensemble $A \subset \Omega_1 \times \Omega_2$. Dans ce cas,

$$F_1(x) = \int_{\Omega_2} \mathbb{1}_A(x,y) d\mu_2(y) = \int_{\Omega_2} \mathbb{1}_{A_1(y)}(x) d\mu_2(y) = \mu_2\big(A_1(x)\big), \tag{13.656}$$

et nous avons déjà vu au théorème 13.159 que cette fonction F_1 était alors mesurable. En utilisant

maintenant les égalités (13.437) ainsi que le fait que $\mathbb{1}_A(x,y) = \mathbb{1}_{A_2(x)}(y)$ nous avons

$$\int_{\Omega_1\times\Omega_2} \mathbb{1}_A(x,y)d(\mu_1\otimes\mu_2)(x,y) = (\mu_1\otimes\mu_2)(A) \tag{13.657a}$$
$$= \int_{\Omega_1} \mu_2\big(A_2(x)\big)d\mu_1(x) \tag{13.657b}$$
$$= \int_{\Omega_1}\left[\int_{\Omega_2} \mathbb{1}_{A_2(x)}(y)d\mu_2(y)\right]d\mu_1(x) \tag{13.657c}$$
$$= \int_{\Omega_1}\left[\int_{\Omega_2} \mathbb{1}_{A}(x,y)d\mu_2(y)\right]d\mu_1(x). \tag{13.657d}$$

Le théorème étant valable pour les fonctions caractéristiques, il est valable pour les fonctions simples (définition 13.103) par linéarité de l'intégrale.

Si f n'est pas une fonction simple, alors la proposition 13.107 nous donne une suite croissante de fonctions simples et positives convergeant ponctuellement vers f. La partie du théorème sur les fonctions simples dit que pour chaque n l'intégrale

$$\int_{\Omega_1\times\Omega_2} f_n(x,y)d(\mu_1\otimes\mu_2)(x,y) \tag{13.658}$$

peut être décomposée comme il faut en suivant la formule (13.655). Il faut pouvoir permuter la limite et l'intégrale dans chacun de cas. D'abord le théorème de la convergence monotone 13.129 appliqué à l'espace $\Omega_1\times\Omega_2$ dit que

$$\lim_{n\to\infty}\int_{\Omega_1\times\Omega_2} f_n(x,y)d(\mu_1\otimes\mu_2)(x,y) = \int_{\Omega_1\times\Omega_2} f(x,y)d(\mu_1\otimes\mu_2)(x,y). \tag{13.659}$$

Ensuite, pour chaque $x\in\Omega_1$, les fonctions

$$\sigma_n(y) = \int_{\Omega_1} f_n(x,y)d\mu_1(x) \tag{13.660}$$

forment une suite croissante de fonctions mesurables ; nous leur appliquons encore le théorème de la convergence monotone :

$$\lim_{n\to\infty}\int_{\Omega_2}\left[\int_{\Omega_1} f_n(x,y)d\mu_1(x)\right]d\mu_2(y) = \lim_{n\to\infty}\int_{\Omega_2}\sigma_n(y)d\mu_2(y) \tag{13.661a}$$
$$= \int_{\Omega_2}\left[\lim_{n\to\infty}\int_{\Omega_1} f_n(x,y)d\mu_1(x)\right]d\mu_2(y) \tag{13.661b}$$
$$= \int_{\Omega_2}\left[\int_{\Omega_1} f(x,y)d\mu_1(x)\right]d\mu_2(y) \tag{13.661c}$$

où nous avons utilisé une seconde fois Beppo-Levi. □

Remarque 13.228.
Les formules (13.655) sont bien, mais ne garantissent en aucun cas que $f\in L^1(\Omega_1\times\Omega_2)$: il faut encore que ces intégrales soient finies.

Corollaire 13.229 ([179])**.**
Soient $(\Omega_i,\mathcal{A}_i,\mu_i)$ deux espaces mesurés σ-finis, et $(\Omega,\mathcal{A},\mu)$ l'espace produit[54]*. Soit une fonction mesurable $f\colon\Omega\to\mathbb{R}$ ou $\mathbb{C}$. Alors les conditions suivantes sont équivalentes*

(1) $f\in L^1(\Omega_1\times\Omega_2)$,

(2)

$$\int_{\Omega_1}\left[\int_{\Omega_2}|f|d\mu_2\right]d\mu_1 < \infty, \tag{13.662}$$

54. Définition 13.161.

(3)

$$\int_{\Omega_2}\left[\int_{\Omega_1}|f|d\mu_1\right]d\mu_2<\infty. \tag{13.663}$$

Démonstration. Nous commençons par supposer que f est à valeurs dans $\mathbb{R}$. La notation $|f|$, pour l'instant, dénote donc bien la valeur absolue et non le module.

La fonction $|f|$ est mesurable et positive par hypothèse et par le fait que si f est mesurable, alors $|f|$ l'est également par le corollaire 13.99. Le théorème 13.227(2) nous dit alors que les intégrales suivantes existent et sont égales :

$$\int_{\Omega_1\times\Omega_2}|f|d(\mu_1\otimes\mu_2)=\int_{\Omega_1}\left[\int_{\Omega_2}|f(x,y)|d\mu_2(y)\right]d\mu_1(x)=\int_{\Omega_2}\left[\int_{\Omega_1}|f(x,y)|d\mu_1(x)\right]d\mu_2(y). \tag{13.664}$$

Attention : rien ne dit encore que ces intégrales sont finies.

(1) implique (2) et (3) Si $f\in L^1(\Omega_1\times\Omega_2)$ alors $|f|$ y est également. Cela implique que le membre de droite de (13.664) est fini. Les deux autres sont alors également finis.

(2) ou (3) implique (1) Les expressions à droite de (13.664) sont finies. Donc celle de gauche également. Cele signifie que $|f|\in L^1(\Omega_1\times\Omega_2)$. Par conséquent f est également dans $L^1(\Omega_2\times\Omega_2)$.

Nous passons maintenant au cas où f est à valeurs dans $\mathbb{C}$. Nous décomposons

$$f=f_R+if_I \tag{13.665}$$

où f_R et f_I sont des fonctions réelles. Nous avons

$$\int_\Omega|f|\leqslant\int_\Omega|f_R|+\int_\Omega|f_I|. \tag{13.666}$$

Donc si f_R et f_I sont dans $L^1(\Omega)$, la fonction f le sera aussi. De même,

$$\int_\Omega|f_R|\leqslant\int_\Omega|f|, \tag{13.667}$$

qui donne l'inverse : si $f\in L^1(\Omega)$ alors $f_R,f_I\in L^1(\Omega)$. Bref, f est intégrable sur Ω si et seulement si f_R et f_I le sont.

Supposons que $f\in L^1(\Omega_1\times\Omega_2)$. Alors

$$\int_{\Omega_1}\left[\int_{\Omega_2}|f|d\mu_2\right]d\mu_1\leqslant\int_{\Omega_1}\left[\int_{\Omega_2}|f_R|\right]+\int_{\Omega_1}\left[\int_{\Omega_2}|f_I|\right]<\infty \tag{13.668a}$$

où nous avons appliqué (1) implique (2) aux fonctions f_R et f_I qui sont dans $L^1(\Omega_1\times\Omega_2)$ parce que f y est.

Dans l'autre sens, si

$$\int_{\Omega_1}\left[\int_{\Omega_2}|f|\right]<\infty, \tag{13.669}$$

alors en remplaçant $|f|$ par $|f_R|$ ou par $|f_I|$ nous restons fini. En appliquant alors «(2) implique (1)» nous trouvons que f_R et f_I sont dans $L^1(\Omega_1\times\Omega_2)$. Et cela implique que $f\in L^1(\Omega_1\times\Omega_2)$. □

Théorème 13.230 (Fubini[179])**.**
Soient $(\Omega_i,\mathcal{A}_i,\mu_i)$ deux espaces mesurés σ-finis, et $(\Omega,\mathcal{A},\mu)$ l'espace produit. Soit

$$f\in L^1\big((\Omega,\mathcal{A}),\mathbb{C}\big), \tag{13.670}$$

c'est à dire une fonction à valeurs mesurable et intégrable sur Ω. Alors :

(1) Pour presque tout $x\in\Omega_1$, la fonction $y\mapsto f(x,y)$ est $L^1(\Omega_2)$.

(2) Si nous posons

$$\varphi_f(x) = \int_{\Omega_2} f(x,y)d\mu_2(y); \tag{13.671}$$

alors $\varphi_f \in L^1(\Omega_1)$.

(3) Nous avons la formule d'inversion d'intégrale

$$\int_\Omega f d(\mu_1 \otimes \mu_2) = \int_{\Omega_1} \varphi_f d\mu_1 \tag{13.672a}$$

$$= \int_{\Omega_1} \left[\int_{\Omega_2} f(x,y)d\mu_2(y) \right] d\mu_1(x) \tag{13.672b}$$

$$= \int_{\Omega_2} \left[\int_{\Omega_1} f(x,y)d\mu_1(x) \right] d\mu_2(y). \tag{13.672c}$$

Démonstration. Nous commençons par supposer que f est à valeurs réelles : $f \in L^1\big((\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2), \mathbb{R}\big)$. Nous décomposons la fonction f en parties positives et négatives : $f = f^+ - f^-$ avec f^+ et f^- positives ou nulles. Nous avons évidemment

$$\int_{\Omega_1 \times \Omega_2} |f^+| \leqslant \int_{\Omega_1 \times \Omega_2} |f| < \infty. \tag{13.673}$$

Donc f^+ et f^- sont des éléments de $L^1(\Omega_1 \times \Omega_2)$.

Pour (1) Nous posons

$$\varphi_{f^+}(x) = \int_{\Omega_2} f^+(x,y)d\mu_2(y) \tag{13.674}$$

pour tous les $x \in \Omega_1$ pour lesquels cette intégrale est bien définie. Vu que f^+ est positive et mesurable, le théorème de Fubini-Tonelli 13.227(1) s'applique donc pour nous dire que φ_{f^+} est mesurable.

De plus le résultat (13.655) appliqué à f^+ donne

$$\int_{\Omega_1} \varphi_{f^+} d\mu_1 = \int_{\Omega_1 \times \Omega_2} f^+ d(\mu_1 \otimes \mu_2) < \infty. \tag{13.675}$$

Le fait que le tout soit fini est une conséquence du fait déjà mentionné que $f^+ \in L^1(\Omega_1 \times \Omega_2$. Vu que φ_{f^+} est une fonction positive, l'inégalité (13.675) signifie que $\varphi_{f^+} \in L^1(\Omega_1, \mu_1)$.

En particulier, $\varphi_{f^+}(x) < \infty$ pour presque tout $x \in \Omega_1$. C'est à dire pour presque tout $x \in \Omega_1$:

$$\int_{\Omega_2} f^+(x,y)d\mu_2(y) < \infty, \tag{13.676}$$

et sachant que $f^+ \geqslant 0$ nous avons $f^+(x, \ \cdot\ ) \in L^1(\Omega_2)$ pour presque tout x.

Pour (2) Partout où φ_{f^+} et φ_{f^-} sont finies nous avons

$$\varphi_f = \varphi_{f^+} - \varphi_{f^-}, \tag{13.677}$$

et comme cela a lieu presque partout, nous pouvons considérer une partie mesurable $A \subset \Omega_1$ telle que $\mu_1(A) = 0$ et $\varphi_f(x) = \varphi_{f^+}(x) - \varphi_{f^-}(x)$ pour tout x hors de A. Bref, nous posons

$$g(x) = \begin{cases} \varphi_{f^+} - \varphi_{f^-}(x) & \text{si } x \in A^c \\ 0 & \text{si } x \in A. \end{cases} \tag{13.678}$$

Cette fonction g est mesurable et $g = \varphi_f$ presque partout. De plus

$$\int_{\Omega_1} |g| d\mu_1 = \int_{A^c} |g| \leqslant \int_{A^c} \varphi_{f^+} + \int_{A^c} \varphi_{f^-} < \infty. \tag{13.679}$$

La dernière inégalité est le fait que $\varphi_{f^\pm}$ sont dans $L^1(\Omega_1)$. Et notons au passage que nous aurions pu laisser toutes les intégrales sur Ω_1 sans faire de précisions sur la distinction entre Ω_1 et A^c parce que la partie de Ω_1 sur laquelle $\varphi_{f^\pm}$ sont infinies est trop petite pour changer la valeur de l'intégrale.

Nous avons donc $g \in L^1(\Omega_1)$, et par conséquent également $\varphi_f \in L^1(\Omega_1)$ parce que ces deux fonctions sont égales presque partout (les classes sont égales).

Pour (3) En utilisant l'équation (13.675) nous avons

$$\int_{\Omega_1} \varphi_f d\mu_1 = \int g d\mu_1 = \int_{\Omega_1} \varphi_{f^+} - \int_{\Omega_1} \varphi_{f^-} \tag{13.680a}$$

$$= \int_{\Omega_1 \times} f^+ d\mu - \int_{\Omega_1 \times \Omega_2} f^- d\mu \tag{13.680b}$$

$$= \int_{\Omega_1 \times \Omega 2} f d\mu. \tag{13.680c}$$

Et toutes ces intégrales sont finies.

Et c'est maintenant que nous considérons le cas complexe. Nous décomposons $f = f_R + if_I$ avec des fonctions réelles f_R et f_I. Comme déjà mentionné autour de (13.666), les fonctions f_R et f_I sont intégrables. Nous leur appliquons le théorème.

Les valeurs de x pour lesquelles $f_R(x, \cdot\,)$ et $f_I(x, \cdot\,)$ ne sont pas dans $L^1(\Omega_2)$ forment un ensemble de mesure nulle, nommons le A. En posant

$$g(x,y) = \begin{cases} f_R(x,y) + if_I(x,y) & \text{si } x \in A^c \\ 0 & \text{si } x \in A, \end{cases} \tag{13.681}$$

nous avons que $g(x, \cdot\,)$ est intégrable pour tout $x \in A^c$. Vu que pour ces valeurs de x nous avons $g(x,y) = f(x,y)$ nous en déduisons que pour $x \in A^c$ nous avons aussi $f(x, \cdot\,) \in L^1(\Omega_2)$.

Les autres points se traitent de la même façon [55] . □

13.231.
En pratique, il n'est pas toujours évident qu'une fonction soit intégrable sur $\Omega_1 \times \Omega_2$. Pour permuter des intégrales sur une fonction à deux paramètres nous faisons comme suit.

(1) Nous testons l'intégrabilité en chaîne de $|f|$, et si c'est bon, le corollaire 13.229 nous donne $f \in L^1(\Omega_1 \times \Omega_2)$.
(2) Nous utilisons le théorème de Fubini 13.230 pour séparer et permuter les intégrales comme des ingénieurs.

Si la fonction $(x,y) \mapsto f(x)g(y)$ satisfait aux hypothèse du théorème de Fubini alors

$$\int_{\Omega_1 \times \Omega_2} f(x)g(y) dx \otimes dy = \left(\int_{\Omega_1} f(x)dx \right) \left(\int_{\Omega_2} g(y)dy \right). \tag{13.682}$$

Le théorème de Fubini est souvent utilisé sous cette forme.

Exemple 13.232(Nécessité d'avoir des mesures σ-finies)
Nous montrons que le théorème ne tient pas si une des deux mesures n'est pas σ-finie. Soit $I = [0,1]$. Nous considérons l'espace mesuré

$$(I, \mathcal{B}\text{or}(I), \lambda) \tag{13.683}$$

où $\mathcal{B}\text{or}(I)$ est la tribu des boréliens sur I et λ est la mesure de Lebesgue (qui est σ-finie). D'autre part nous considérons l'espace mesuré

$$(I, \mathcal{P}(I), m) \tag{13.684}$$

55. Attention : je n'ai pas vérifié explicitement. C'est juste une intuition.

où $\mathcal{P}(I)$ est l'ensemble des parties de I et m est la mesure de comptage. Cette dernière n'est pas σ-finie parce que les seuls ensembles de mesure finie pour la mesure de comptage sont des ensembles finis, or une union dénombrable d'ensemble finis ne peut pas recouvrir l'intervalle I.

Nous allons montrer que dans ce cadre, l'intégrale de la fonction indicatrice de la diagonale sur I^2 ne vérifie pas le théorème de Fubini. Étant donné que $\mathcal{B}\text{or}(I) \subset \mathcal{P}(I)$ nous avons

$$\mathcal{B}\text{or}(I^2) \subset \mathcal{B}\text{or}(I) \otimes \mathcal{P}(I). \tag{13.685}$$

Soit $\Delta = \{(x,x) \text{ tel que } x \in I\}$. La fonction

$$\begin{aligned} g\colon I^2 &\to \mathbb{R} \\ (x,y) &\mapsto x - y \end{aligned} \tag{13.686}$$

est continue et $\Delta = g^{-1}(\{0\})$ est donc fermé dans I^2. L'ensemble Δ est donc un borélien de I^2 et par conséquent un élément de la tribu $\mathcal{B}\text{or}(I)\otimes\mathcal{P}(I)$. La fonction indicatrice $\mathbb{1}_\Delta$ est alors mesurable pour l'espace mesuré

$$(I \times I, \mathcal{B}\text{or}(I) \otimes \mathcal{P}(I), \lambda \otimes m). \tag{13.687}$$

Pour x fixé nous avons

$$\mathbb{1}_\Delta(x,y) = \begin{cases} 1 & \text{si } y = x \\ 1 & \text{si } y \neq x \end{cases} = \mathbb{1}_{\{x\}}(y), \tag{13.688}$$

et donc

$$\begin{aligned} A_1 &= \int_I \left(\int_I \mathbb{1}_\Delta(x,y) dm(y) \right) d\lambda(x) && \text{(13.689a)} \\ &= \int_I \left(\int_I \mathbb{1}_{\{x\}}(y) dm(y) \right) d\lambda(x) && \text{(13.689b)} \\ &= \int_I \Big(m(\{x\}) \Big) d\lambda(x) && \text{(13.689c)} \\ &= \int_I 1 d\lambda(x) && \text{(13.689d)} \\ &= 1. && \text{(13.689e)} \end{aligned}$$

Par contre le support de $\mathbb{1}_\Delta$ étant de mesure nulle pour la mesure de Lebesgue, nous avons

$$\int_I \mathbb{1}_\Delta(x,y) d\lambda(x) = 0 \tag{13.690}$$

et par conséquent

$$A_2 = \int_I \left(\int_I \mathbb{1}_\Delta(x,y) d\lambda(x) \right) dm(y) = 0. \tag{13.691}$$

Nous voyons donc que le théorème de Fubini ne s'applique pas. △

Exemple 13.233
Nous nous proposons de calculer l'intégrale suivante en utilisant le théorème de Fubini :

$$G = \int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi} \tag{13.692}$$

alors que la fonction $x \mapsto e^{-x^2}$ n'a pas de primitives parmi les fonctions élémentaires.

Nous allons le faire de deux façons. Une première directe en utilisant le théorème de Fubini sur un domaine non borné, et une seconde en utilisant Fubini sur un domaine borné, et en passant à la limite ensuite.

Fubini, domaine non borné Par symétrie nous pouvons nous contenter de calculer

$$G_+ = \int_0^\infty e^{-x^2} dx. \tag{13.693}$$

L'astuce est de passer par l'intermédiaire

$$\begin{aligned} H &= \int_{\mathbb{R}^+\times\mathbb{R}^+} e^{-(x^2+y^2)} dxdy && \text{(13.694a)}\\ &= \int_{\mathbb{R}^+} \left(\int_{\mathbb{R}^+} e^{-x^2} e^{-y^2} dx \right) dy && \text{(13.694b)}\\ &= \left(\int_{\mathbb{R}^+} e^{-x^2} dx \right)^2 && \text{(13.694c)}\\ &= G_+^2 && \text{(13.694d)} \end{aligned}$$

L'intégrale (13.694a) se calcule en passant aux coordonnées polaires et le résultat est $H = \frac{\pi}{4}$. Nous avons alors $G = \frac{\sqrt{\pi}}{2}$ et

$$\int_{\mathbb{R}} e^{-x^2} = \sqrt{\pi}. \tag{13.695}$$

Fubini, domaine borné, puis limite Une variante, qui n'applique pas Fubini sur un domaine non borné. Nous commençons par écrire

$$I = \int_{-\infty}^{+\infty} e^{-x^2} dx := \lim_{R\to+\infty} \int_{-R}^{+R} e^{-x^2} dx \tag{13.696}$$

et puis nous faisons le calcul

$$\begin{aligned} I^2 &= \lim_{R\to+\infty} \left((\int_{-R}^{+R} e^{-x^2} dx)(\int_{-R}^{+R} e^{-y^2} dy) \right)\\ &= \lim_{R\to+\infty} \left(\iint_{K_R} e^{-(x^2+y^2)} dxdy \right)\\ &= \lim_{R\to+\infty} \left(\iint_{C_R} e^{-(x^2+y^2)} dxdy \right) \end{aligned} \tag{13.697}$$

où K est le carré de demi côté R centré à l'origine et de côtés parallèles aux axes et C_R est le cercle de rayon R centré à l'origine.

La première étape à justifier est simplement l'application de Fubini. Pour le passage de l'intégrale du carré vers le cercle, définissons

$$I_K(r) = \int_{K_r} f, \quad I_C(r) = \int_{C_r} f \tag{13.698}$$

où K_r est la carré de demi côté r et C_r est le cercle de rayon r. Le demi côté du carré inscrit à C_r est $\sqrt{2}$, donc pour tout r nous avons

$$I_K(\sqrt{2}r) \leqslant I_C(r) < I_K(r), \tag{13.699}$$

et en prenant la limite, nous avons évidement

$$\lim_{r\to\infty} I_K(\sqrt{2}r) = \lim_{r\to\infty} I_K(r), \tag{13.700}$$

et donc cette limite est également égale à $\lim_{r\to\infty} I_C(t)$.

Il ne reste qu'à calculer la dernière intégrale sur le cercle en passant aux coordonnées polaires :

$$\iint_{C_R} e^{-(x^2+y^2)}dxdy = \int_0^{2\pi} d\theta \int_0^R re^{-r^2}dr = \pi(1-e^{-R^2}). \tag{13.701}$$

La limite donne π, nous en déduisons que

$$\int_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi}. \tag{13.702}$$

△

Le théorème de Fubini-Tonelli nous permet également d'inverser des sommes et des séries. En effet une somme n'est rien d'autre qu'une intégrale pour la mesure de comptage :

$$\sum_{n=0}^{\infty} a_n = \int_{\mathbb{N}} a_n dm(n). \tag{13.703}$$

La proposition suivante montre comment il faut faire.

Proposition 13.234.
Soient les espaces mesurés $(\mathbb{N}, \mathcal{P}(\mathbb{N}), m)$, $(\mathbb{R}^n, \mathcal{B}or(\mathbb{R}^n), \lambda)$ où λ est la mesure de Lebesgue ainsi qu'une suite de fonctions positives $f_n\colon \mathbb{R}^d \to \mathbb{R}$. Nous supposons de plus que la fonction f_n soit intégrable pour tout n et que les résultats forment une suite sommable. Alors

$$\sum_{n=0}^{\infty} \int_{\mathbb{R}^n} f_n(x)dx = \int_{\mathbb{R}^d} \sum_{n\in\mathbb{N}} f_n(x)dx. \tag{13.704}$$

Démonstration. Nous pouvons la récrire le membre de gauche sous la forme

$$\int_{\mathbb{N}} \left(\int_{\mathbb{R}^n} f(n,x)dx \right) dm(n) \tag{13.705}$$

avec la notation évidente $f(n,x) = f_n(x)$. Prouvons que la fonction $f\colon \mathbb{N}\times\mathbb{R}^d \to \mathbb{R}$ ainsi définie est une fonction mesurable pour l'espace mesuré

$$\big(\mathbb{N}\times\mathbb{R}^d, \mathcal{P}(\mathbb{N})\otimes\mathcal{B}\mathrm{or}(\mathbb{R}^d), m\otimes\lambda\big). \tag{13.706}$$

Si $A\subset\mathbb{R}$, nous avons

$$f^{-1}(A) = \bigcup_{n\in\mathbb{N}} \{n\}\times f_n^{-1}(A). \tag{13.707}$$

Chacun des ensembles dans l'union appartient à la tribu $\mathcal{P}(\mathbb{N})\times\mathcal{B}\mathrm{or}(\mathbb{R}^d)$ tandis que les tribus sont stables sous les unions dénombrables. La fonction f est donc mesurable. Comme nous avons supposé que f était positive, le théorème de Fubini-Tonelli s'applique et nous avons

$$\int_{\mathbb{R}^d} \left(\int_{\mathbb{N}} f(n,x)dm(n) \right) dx = \int_{\mathbb{R}^d} \sum_{n\in\mathbb{N}} f_n(x)dx. \tag{13.708}$$

□

Théorème 13.235 (Fubini)**.**
Soit $(x,t)\mapsto f(x,y)\in\bar{\mathbb{R}}$ une fonction intégrable sur $B_n\times B_m\subset\mathbb{R}^{n+m}$ où B_n et B_m sont des ensembles mesurables de $\mathbb{R}^n$ et $\mathbb{R}^m$. Alors :

(1) pour tout $x\in B_n$, sauf éventuellement en les points d'un ensemble $G\subset B_n$ de mesure nulle, la fonction $y\in B_m\mapsto f(x,y)\in\bar{\mathbb{R}}$ est intégrable sur B_m

(2) la fonction

$$x \in B_n \backslash G \mapsto \int_{B_m} f(x,y)dy \in \mathbb{R} \tag{13.709}$$

est intégrable sur $B_n \backslash G$

(3) On a

$$\int_{B_n \times B_m} f(x,y)dxdy = \int_{B_n} \left(\int_{B_m} f(x,y)dy \right) dx. \tag{13.710}$$

Notons en particulier que si $f(x,y) = \varphi(x)\phi(y)$, alors $\int_{B_m} \varphi(y)dy$ est une constante qui peut sortir de l'intégrale sur B_n, et donc

$$\int_{B_n \times B_m} \varphi(x)\phi(y)dxdy = \int_{B_n} \varphi(x)dx \int_{B_m} \phi(y)dy. \tag{13.711}$$

Chapitre 14

Intégration

14.1 Intégrale sur une variété

14.1.1 Mesure sur une carte

Nous considérons dans cette section uniquement des variétés M de dimension 2 dans $\mathbb{R}^3$. Une particularité de $\mathbb{R}^3$ (par rapport aux autres $\mathbb{R}^n$) est qu'il existe le produit vectoriel.

Si v, $w \in \mathbb{R}^3$, alors le vecteur $v \times w$ est une vecteur normal au plan décrit par v et w qui jouit de l'importante propriété suivante :

$$\text{aire du parallélogramme} = \|v \times w\|. \tag{14.1}$$

L'aire du parallélogramme construit sur v et w est donnée par la norme du produit vectoriel. Afin de donner une mesure infinitésimale en un point $p \in M$, nous voudrions prendre deux vecteurs tangents à M en p, et puis considérer la norme de leur produit vectoriel. Cette idée se heurte à la question du choix des vecteurs tangents à considérer.

Dans $\mathbb{R}^2$, le choix est évident : nous choisissons e_x et e_y, et nous avons $\|e_x \times e_y\| = 1$. L'idée est donc de choisir une carte $F\colon W \to F(w)$ autour du point $p = F(w)$, et de choisir les vecteurs tangents qui correspondent à e_x et e_y via la carte, c'est à dire les vecteurs

$$\frac{\partial F}{\partial x}(w), \quad \text{et} \quad \frac{\partial F}{\partial y}(w). \tag{14.2}$$

L'**élément infinitésimal de surface** sur M au point $p = F(w)$ est alors défini par

$$d\sigma_F = \|\frac{\partial F}{\partial x}(w) \times \frac{\partial F}{\partial y}(w)\| dw, \tag{14.3}$$

et si la partie $A \subset M$ est entièrement contenue dans $F(W)$, nous définissons la **mesure** de A par

$$\mu_2(A) = \int_{F^{-1}(A)} d\sigma_F = \int_{F^{-1}(A)} \|\frac{\partial F}{\partial x}(w) \times \frac{\partial F}{\partial y}(w)\| dw. \tag{14.4}$$

Remarque 14.1.
Afin que cette définition ait un sens, nous devons prouver qu'elle ne dépend pas du choix de la carte F. En effet, les vecteurs $\partial_x F$ et $\partial_y F$ dépendent de la carte F, donc leur produit vectoriel (et sa norme) dépendent également de la carte F choisie. Il faudrait donc un petit miracle pour que le nombre $\mu_2(A)$ donné par (14.4) soit indépendant du choix de F. Nous allons bientôt voir comme cas particulier du théorème 14.4 que c'est en fait le cas. C'est à dire que si F et $\tilde{F}$ sont deux cartes qui contiennent A, alors

$$\int_{F^{-1}(A)} d\sigma_F = \int_{\tilde{F}^{-1}(A)} d\sigma_{\tilde{F}}. \tag{14.5}$$

14.1.1.1 Exemple : la mesure de la sphère

Nous nous proposons maintenant de calculer la surface de la sphère $S^2 = x^2 + y^2 + z^2 = R^2$. L'application $F\colon B((0,0),R) \to R^3$ donnée par

$$F(x,y) = \begin{pmatrix} x \\ y \\ \sqrt{R^2 - x^2 - y^2} \end{pmatrix} \tag{14.6}$$

est une carte pour une demi sphère. Ses dérivées partielles sont

$$\frac{\partial F}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ -\frac{x}{\sqrt{R^2-x^2-y^2}} \end{pmatrix}, \quad \frac{\partial F}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ -\frac{y}{\sqrt{R^2-x^2-y^2}} \end{pmatrix}. \tag{14.7}$$

Le produit vectoriel de ces deux vecteurs tangents donne

$$\frac{\partial F}{\partial x}(x,y) \times \frac{\partial F}{\partial y}(x,y) = \frac{x}{\alpha}e_1 + \frac{y}{\alpha}e_2 + e_3 \tag{14.8}$$

où $\alpha = \sqrt{R^2 - x^2 - y^2}$. En calculant la norme, nous trouvons

$$\|\frac{\partial F}{\partial x}(x,y) \times \frac{\partial F}{\partial y}(x,y)\| = \sqrt{\frac{R^2}{R^2 - x^2 - y^2}}, \tag{14.9}$$

et en passant aux coordonnées polaires, nous écrivons l'intégrale (14.4) sous la forme

$$\int_B \|\partial_x F \times \partial_y F\| = \int_0^{2\pi} d\theta \int_0^R r\sqrt{\frac{R^2}{R^2 - x^2 - y^2}}dr = 2\pi R^2, \tag{14.10}$$

qui est bien la mesure de la demi sphère.

14.1.2 Intégrale sur une carte

Nous pouvons maintenant définir l'intégrale d'une fonction sur une carte de la variété M.

Définition 14.2.
*Soit $F\colon W \subset \mathbb{R}^2 \to \mathbb{R}^3$, une carte pour une variété M. Soit A, une partie de $F(W)$ telle que $A = F(B)$ où $B \subset W$ est mesurable. Soit encore $f\colon A \to \mathbb{R}$, une fonction continue. L'**intégrale** de f sur A est le nombre*

$$\int_A f = \int_A f d\sigma_F = \int_{F^{-1}(A)} (f \circ F)(w)\|\frac{\partial F}{\partial x}(w) \times \frac{\partial F}{\partial y}(w)\|dw \tag{14.11}$$

Remarque 14.3.
L'intégrale (14.11) n'est pas toujours bien définie. Étant donné que F est C^1 et que f est continue, l'intégrante est continue. L'intégrale sera donc bien définie par exemple lorsque B est borné et si la fermeture $\bar{A}$ est un compact contenu dans $F(w)$.

Le théorème suivant montre que le travail que nous avons fait jusqu'à présent ne dépend en fait pas du choix de carte F effectué.

Théorème 14.4.
Soient $F\colon W \to F(w)$ et $\tilde{F}\colon \tilde{W} \to \tilde{F}(\tilde{W})$, deux cartes de la variété M. Soit une partie $A \subset F(W) \cap \tilde{F}(\tilde{W})$ telle que $A = F(B)$ avec $B \subset W$ mesurable. Alors $A = \tilde{F}(\tilde{B})$ avec $\tilde{B} \subset \tilde{W}$ mesurable.

Si f est une fonction continue, et si $\int_A f d\sigma_F$ existe, alors $\int_A f d\sigma_{\tilde{F}}$ existe et

$$\int_A f d\sigma_F = \int_A f d\sigma_{\tilde{F}}. \tag{14.12}$$

14.1.3 Exemples

Intégrons la fonction $f(x,y,z)$ sur le carré $K =]0,1[\times]0,2[\times \{1\}$. La première carte que nous pouvons utiliser est

$$\begin{aligned} F\colon]0,1[\times]0,2[&\to K \\ (x,y) &\mapsto (x,y,1). \end{aligned} \tag{14.13}$$

Nous trouvons aisément les vecteurs tangents qui forment l'élément de surface :

$$\frac{\partial F}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \frac{\partial F}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \tag{14.14}$$

donc $d\sigma_F = 1 \cdot dxdy$, et

$$\int_K f d\sigma_F = \int_{]0,1[\times]0,2[} f(x,y,1) \cdot 1 \cdot dxdy. \tag{14.15}$$

Nous pouvons également utiliser la carte

$$\begin{aligned} \tilde{F}\colon]0,\frac{1}{2}[\times]0,6[&\to K \\ (\tilde{x},\tilde{y}) &\mapsto (2\tilde{x},\frac{\tilde{y}}{3},1). \end{aligned} \tag{14.16}$$

Les vecteurs tangents sont maintenant

$$\frac{\partial \tilde{F}}{\partial \tilde{x}} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \quad \frac{\partial \tilde{F}}{\partial \tilde{y}} = \begin{pmatrix} 0 \\ 1/3 \\ 0 \end{pmatrix}, \tag{14.17}$$

et nous avons donc $d\sigma_{\tilde{F}} = \|\frac{2}{3}e_3\| = \frac{2}{3}$. Cette fois, l'intégrale de f sur K s'écrit

$$\int_K f d\sigma_{\tilde{F}} = \int_{]0,\frac{1}{2}[\times]0,6[} f\big(2\tilde{x},\frac{\tilde{y}}{3},1\big) \cdot \frac{2}{3} \cdot d\tilde{x}s\tilde{y}. \tag{14.18}$$

Conformément au théorème 14.4, cette dernière intégrale est égale à l'intégrale (14.15) parce qu'il s'agit juste d'un changement de variable.

14.1.4 Orientation

Soient $F\colon W \to F(w)$ et $\tilde{F}\colon \tilde{W} \to \tilde{F}(\tilde{W})$, deux cartes de la variété M. Nous pouvons considérer la fonction $h = \tilde{F}^{-1} \circ F$, définie uniquement sur l'intersection des cartes :

$$h\colon F^{-1}\big(F(W) \cap \tilde{F}(\tilde{W})\big) \to \tilde{F}^{-1}\big(F(W) \cap \tilde{F}(\tilde{W})\big). \tag{14.19}$$

Nous disons que F et $\tilde{F}$ ont même **orientation** si

$$J_h(w) > 0 \tag{14.20}$$

pour tout $w \in F^{-1}\big(F(W) \cap \tilde{F}(\tilde{W})\big)$.

Considérons les deux carte suivantes pour le même carré :

$$\begin{aligned} F\colon]0,1[\times]0,1[&\to \mathbb{R}^3 \\ (x,y) &\mapsto (x,y,0) \end{aligned} \tag{14.21}$$

et

$$\begin{aligned} \tilde{F}\colon]0,\frac{1}{2}[\times]0,\frac{1}{3}[&\to \mathbb{R}^3 \\ (x,y) &\mapsto (2x,3y,0) \end{aligned} \tag{14.22}$$

Ici, $h(x,y) = \left(\frac{x}{2}, \frac{y}{3}\right)$ et nous avons $J_h = \frac{1}{6} > 0$. Ces deux cartes ont même orientation. Notez que

$$\frac{\partial F}{\partial x} \times \frac{\partial F}{\partial y} = e_3, \tag{14.23}$$

tandis que

$$\frac{\partial \tilde{F}}{\partial x} \times \frac{\partial \tilde{F}}{\partial y} = 6e_3. \tag{14.24}$$

Les vecteurs normaux à la paramétrisation pointent dans le même sens.

Si par contre nous prenons la paramétrisation

$$\begin{aligned} G\colon]0,1[\times]0,1,[&\to \mathbb{R}^2 \\ (x,y) &\mapsto (x,(1-y),0), \end{aligned} \tag{14.25}$$

nous avons

$$\frac{\partial G}{\partial x} \times \frac{\partial G}{\partial y} = -e_3, \tag{14.26}$$

et si $g = G^{-1} \circ F$, alors $J_g = -1$.

L'orientation d'une carte montre donc si le vecteur normal à la surface pointe d'un côté ou de l'autre de la surface.

Définition 14.5.
*Une variété M est **orientable** s'il existe un atlas de M tel que deux cartes quelconques ont toujours même orientation. Une variété est **orientée** lorsque qu'un tel choix d'atlas est fait.*

Proposition 14.6.
Soit M, une variété orientable et un atlas orienté $\{F_i\colon W_i \to \mathbb{R}^3\}$. Alors le vecteur unitaire

$$\frac{\frac{\partial F}{\partial x}(x,y) \times \frac{\partial F}{\partial y}(x,y)}{\left\| \frac{\partial F}{\partial x}(x,y) \times \frac{\partial F}{\partial y}(x,y) \right\|} \tag{14.27}$$

ne dépend pas du choix de F parmi les F_i.

Démonstration. Considérons deux cartes F_1 et F_2, ainsi que l'application $h = F_2^{-1} \circ F_1$. Écrivons le vecteur $\partial_x F_1 \times \partial_y F_1$ en utilisant $F_1 = F_2 \circ h$. D'abord, par la règle de dérivation de fonctions composées,

$$\frac{\partial (F_2 \circ h)}{\partial x} = \frac{\partial F_2}{\partial x}\frac{\partial h_1}{\partial x} + \frac{\partial F_2}{\partial y}\frac{\partial h_2}{\partial x}. \tag{14.28}$$

Après avoir fait le même calcul pour $\frac{\partial (F_2 \circ h)}{\partial y}$, nous pouvons écrire

$$\partial_x(F_2 \circ h) \times \partial_y(F_2 \circ h) = (\partial_x h_1 \partial_x F_2 + \partial_x h_2 \partial_y F_2) \times (\partial_y h_1 \partial_x F_2 + \partial_y h_2 \partial_y F_2). \tag{14.29}$$

Dans cette expression, les facteurs $\partial_i h_j$ sont des nombres, donc ils se factorisent dans les produits vectoriels. En tenant compte du fait que $\partial_x F_2 \times \partial_x F_2 = 0$ et $\partial_y F_2 \times \partial_y F_2 = 0$, ainsi que de l'antisymétrie du produit vectoriel, l'expression se réduit à

$$\left(\frac{\partial F_2}{\partial x} \times \frac{\partial F_2}{\partial y}\right)(\partial_x h_1 \partial_y h_2 - \partial_x h_2 \partial_y h_2). \tag{14.30}$$

Par conséquent,

$$\frac{\partial F_1}{\partial x} \times \frac{\partial F_1}{\partial y} = \frac{\partial (F_2 \circ h)}{\partial x} \times \frac{\partial (F_2 \circ h)}{\partial y} = \left(\frac{\partial F_2}{\partial x} \times \frac{\partial F_2}{\partial y}\right) \det J_h. \tag{14.31}$$

Donc, tant que J_h est positif, les vecteurs unitaires correspondants au membre de gauche et de droite sont égaux. □

Corollaire 14.7.
Si nous avons choisit un atlas orienté pour la variété M, nous avons une fonction continue $G\colon M \to \mathbb{R}^3$ telle que $\|G(p)\| = 1$ pour tout $p \in M$. Cette fonction est donnée par

$$G(F(x,y)) = \frac{\frac{\partial F}{\partial x}(x,y) \times \frac{\partial F}{\partial y}(x,y)}{\|\frac{\partial F}{\partial x}(x,y) \times \frac{\partial F}{\partial y}(x,y)\|} \tag{14.32}$$

sur l'image de la carte F.

Démonstration. La fonction G est construite indépendamment sur chaque carte $F(W)$ en utilisant la formule (14.32). Cette fonction est une fonction bien définie sur tout M parce que nous venons de démontrer que sur $F_1(W_1) \cap F_2(W_2)$, les fonctions construites à partir de F_1 et à partir de F_2 sont égales. □

Il est possible que prouver, bien que cela soit plus compliqué, que la réciproque est également vraie.

Proposition 14.8.
Une variété M de dimension 2 dans $\mathbb{R}^3$ est orientable si et seulement s'il existe une fonction continue $G\colon M \to \mathbb{R}^3$ telle que pour tout $p \in M$, le vecteur $G(p)$ soit de norme 1 et normal à M au point p.

14.1.5 Formes différentielles

Nous allons donner une toute petite introduction aux formes différentielles sur des variétés compactes.

Lemme 14.9 ([185])**.**
Soit ω une k-forme sur $\mathbb{R}^n$ et f, une fonction C^∞ sur $\mathbb{R}^n$. Alors $d(f^*\omega) = f^*d\omega$.

Démonstration. Nous effectuons la preuve par récurrence sur le degré de la forme. Soit d'abord une 0-forme, c'est à dire une fonction $g\colon \mathbb{R}^n \to \mathbb{R}$. Nous avons

$$d(d^*g)X = d(g \circ f)X = (dg \circ df)X = dg\big(df X\big) = (f^*dg)(X). \tag{14.33}$$

Supposons maintenant que le résultat soit exact pour toute les $p-1$ formes et montrons qu'il reste valable pour les p-formes. Par linéarité de la différentielle nous pouvons nous contenter de considérer la forme différentielle

$$\omega = g\, dx^1 \wedge \dots dx^p \tag{14.34}$$

où g est une fonction C^∞. Pour soulager les notations nous allons noter $dx^I = dx^1 \wedge \dots dx^{p-1}$. Nous avons

$$\begin{aligned}
d(f^*\omega) &= d\big(f^*(g dx^I \wedge dx^p)\big) && \text{(14.35a)}\\
&= d\big(f^*(g dx^I) \wedge f^*dx^p\big) && \text{(14.35b)}\\
&= d\big(f^*(g dx^I)\big) \wedge f^*dx^p + (-1)^{p-1} f^*(g dx^I) \wedge (f^*dx^p) && \text{(14.35c)}\\
&= f^*\big(d(g dx^I)\big) \wedge f^*dx^p && \text{(14.35d)}\\
&= f^*\big(d(g dx^I) \wedge dx^p\big) && \text{(14.35e)}\\
&= f^*d\omega && \text{(14.35f)}
\end{aligned}$$

Justifications : (14.35c) est la formule de Leibnitz. (14.35d) est parce que le second terme est nul : $d(f^*dx^p) = f^*(d^2x^p) = 0$. Nous avons utilisé l'hypothèse de récurrence et le fait que $d^2 = 0$. L'étape (14.35f) est une utilisation à l'envers de la règle de Leibnitz en tenant compte que $d^2x^p = 0$. □

Soit M une variété de dimension n et ω une n-forme différentielle

$$\omega_p = f(p) dx_1 \wedge \dots \wedge dx_n. \tag{14.36}$$

Si (U, φ) est une carte ($U \subset \mathbb{R}^n$ et $\varphi\colon U \to M$) alors nous définissons

$$\int_{\varphi(U)} \omega = \int_U f\big(\varphi(x)\big) dx_1 \dots dx_n. \tag{14.37}$$

Lorsque nous voulons intégrer sur une partie plus grande qu'une carte nous utilisons une partition de l'unité.

Lemme 14.10.
Soit $\{U_i\}$ un recouvrement de M par un nombre fini d'ouverts[1]*. Alors il existe une famille de fonctions $f_i \in C^\infty(M)$ telle que*

(1) $\operatorname{supp} f_i \subset U_i$,

(2) pour tout i, nous avons $f_i \geqslant 0$,

(3) pour tout $p \in M$ nous avons $\sum_i f_i(p) = 1$.

La famille (f_i) est une **partition de l'unité** subordonnée au recouvrement $\{U_i\}$.

Définition 14.11 (Intégrale d'une forme sur une variété)**.**
Si $\{f_i\}$ est une partition de l'unité subordonnée à un atlas de M nous définissons

$$\int_M \omega = \sum_i \int_{U_i} f\omega. \tag{14.38}$$

Il est possible de montrer que cette définition ne dépend pas du choix de la partition de l'unité.

Remarque 14.12.
Nous ne définissons pas d'intégrale de k-forme différentielle sur une variété de dimension $n \neq k$. Le seul cas où cela se fait est le cas de 0-formes (les fonctions), mais cela n'est pas vraiment un cas particulier vu que les 0-formes sont associées aux n-formes de façon évidente.

Remarque 14.13.
La définition 14.11 permet d'intégrer des formes sur des variétés, pas sur des sous-variétés. Autrement dit, ce n'est pas cette définition là qu'il faut utiliser pour comprendre des objets comme

$$\int_\gamma f \tag{14.39}$$

où γ est un chemin dans $\mathbb{R}^n$. Nous définirons ce genre de choses plus bas.

14.1.6 Intégrale d'une fonction sur une sous-variété

Nous allons nous restreindre au cas d'une sous-variété de $\mathbb{R}^n$. C'est à dire que nous considérons l'espace $\mathbb{R}^n$ comme espace ambiant et nous allons intégrer sur des parties de $\mathbb{R}^n$. Ces parties peuvent être de dimension plus basses que n, et c'est justement ça qui fait la différence entre ce que nous faisons maintenant et la définition 14.11.

L'exemple typique est l'intégrale sur une surface dans $\mathbb{R}^3$, ou de volumes.

Nous supposons à présent que M est une variété compacte de dimension 2 dans $\mathbb{R}^3$. La compacité fait que M possède un atlas contenant un nombre fini de cartes $F_i\colon W_i \to F_i(W_i)$.

Si $A \subset M$ est tel que pour chaque i, $A \cap F_i(W_i) = F_i(V_i)$ pour une ensemble V_i mesurable dans $\mathbb{R}^2$, alors nous considérons

$$A_1 = A \cap F_1(W_2) = F_1(V_1). \tag{14.40}$$

Ensuite, nous construisons A_2 en considérant $F_A(W_2)$ et en lui retranchant A_1 :

$$A_2 = \big(A \cap F_2(W_2)\big) \cap F_1(V_1). \tag{14.41}$$

1. Si M n'est pas compacte, alors il faut utiliser une version un peu plus élaborée du lemme[185].

En continuant de la sorte, nous construisons la décomposition

$$A = A_1 \cup \ldots \cup A_p \tag{14.42}$$

de A en ouverts disjoints, chacun de ouverts A_p étant compris dans une carte.

Il est possible de prouver que dans ce cas, la définition suivante a un sens et ne dépend pas du choix de l'atlas effectué.

Définition 14.14.
*Si $f\colon A \to \mathbb{R}$ est une fonction continue, alors l'**intégrale** est le nombre*

$$\int_A f = \sum_{i=1}^{p} \int_{A_i} f d\sigma_{F_i}. \tag{14.43}$$

14.2 Intégrales curvilignes

14.2.1 Chemins de classe C^1

Définition 14.15.
*Soit $p, q \in \mathbb{R}^n$. Un **chemin** C^1 par morceaux joignant p à q est une application continue*

$$\gamma : [a, b] \to \mathbb{R}^n \qquad \gamma(a) = p, \gamma(b) = q \tag{14.44}$$

pour laquelle il existe une subdivision $a = t_0 < t_1 < \ldots < t_{r-1} < t_r = b$ telle que :

(1) la restriction de γ sur chaque ouvert $]t_i, t_{i+1}[$ est de classe C^1 ;

(2) pour tout $0 \leqslant i \leqslant r$, γ' possède une limite à gauche (sauf pour $i = 0$) et une limite à droite (sauf pour $i = r$) en t_i.

*Le **chemin γ est (globalement) de classe** C^1 si la subdivision peut être choisie de « longueur » $r = 1$.*

Remarque 14.16.
Si a et b sont des points de $\mathbb{R}^n$, on peut créer le chemin particulier

$$\gamma : [0, 1] \to \mathbb{R}^n : t \mapsto (1 - t)a + tb \tag{14.45}$$

qui relie ces points par un segment de droite.

14.2.2 Intégrer une fonction

Définition 14.17.
*Soit $f : D \subset \mathbb{R}^n \to \mathbb{R}$ une fonction continue, et $\gamma : [a, b] \to D$ un chemin C^1. On définit **l'intégrale de f sur γ** par*

$$\int_\gamma f ds = \int_\gamma f = \int_a^b f(\gamma(t)) \left\|\gamma'(t)\right\| dt. \tag{14.46}$$

Exemple 14.18
Soit l'hélice

$$\begin{aligned} \sigma\colon [0, 2\pi] &\to \mathbb{R}^3 \\ t &\mapsto \begin{pmatrix} \cos(t) \\ \sin(t) \\ t \end{pmatrix}, \end{aligned} \tag{14.47}$$

et la fonction $f(x,y,z)=x^2+y^2+z^2$. L'intégrale de f sur σ est

$$\begin{aligned}\int_\sigma f &= \int_0^{2\pi}(\cos^2 t+\sin^2 t+t^2)\|\sigma'(t)\|dt\\ &= \int_0^{2\pi}(1+t^2)\sqrt{2}dt\\ &= \sqrt{2}\left[t+\frac{t^3}{3}\right]_0^{2\pi}\\ &= \sqrt{2}\left(2\pi+\frac{8\pi^3}{8}\right).\end{aligned} \tag{14.48}$$

△

Remarque 14.19.
Si $f=1$, alors nous tombons sur

$$\int_\gamma ds=\int_a^b\|\gamma'(t)\|dt, \tag{14.49}$$

Nous verrons par le théorème 19.9 que cette dernière intégrale est la longueur de la courbe. Il est un fait général que l'intégrale de la fonction 1 sur un ensemble en donne la «mesure». Cela est à mettre en rapport avec le lemme 13.115 en gardant en tête que $\int_\gamma 1$ n'est pas la mesure de l'image de γ dans $\mathbb{R}^2$.

Proposition 14.20 (Indépendence en la paramétrisation)**.**
La valeur de l'intégrale de f sur γ ne dépend pas du paramétrage (équivalent ou pas) choisi.

Démonstration. Soit donc un chemin $\gamma\colon[c,d]\to\mathbb{R}^3$ ainsi que $\varphi\colon[c,d]\to[a,b]$, une reparamétrisation de classe C^1, strictement monotone et le chemin σ définit par $\gamma(s)=\sigma\big(\varphi(s)\big)$ avec $s\in[c,d]$. En supposant que $\varphi'(s)\geqslant 0$, nous avons

$$\begin{aligned}I=\int_\gamma f &= \int_c^d f\big(\gamma(s)\big)\|\gamma'(s)\|ds\\ &= \int_c^d f\Big(\sigma\big(\varphi(s)\big)\Big)\|\sigma'\big(\varphi(s)\big)\||\varphi'(s)|ds.\end{aligned} \tag{14.50}$$

Pour cette intégrale, nous posons $t=\varphi(s)$, et par conséquent $dt=\varphi'(s)ds$. Étant donné que $\varphi'(s)\geqslant 0$, nous pouvons supprimer les valeurs absolues, et obtenir

$$\begin{aligned}I &= \int_{\varphi(c)}^{\varphi(d)} f\big(\sigma(t)\big)\|\sigma'(t)\|dt\\ &= \int_a^b f\big(\sigma(t)\big)\|\sigma'(t)\|dt\\ &= \int_\sigma f.\end{aligned} \tag{14.51}$$

Essayez de faire le cas $\varphi'(s)\leqslant 0$. □

Remarque 14.21.
Attention : les intégrales sur des chemins dans $\mathbb{C}$ ne sont la même chose. En effet $\mathbb{C}$ doit être souvent plutôt traité comme $\mathbb{R}$ que comme $\mathbb{R}^2$. Si γ est un chemin dans $\mathbb{C}$, l'intégrale

$$\int_\gamma f \tag{14.52}$$

doit être comprise comme une généralisation de $\int_a^b f(x)dx$ et non comme l'intégrale sur un chemin. La différence est qu'en retournant les bornes d'une intégrale usuelle sur $\mathbb{R}$ on change le signe, alors

qu'en retournant un chemin dans $\mathbb{R}^2$, on ne change pas. Bref, la définition est que si $\gamma\colon [a,b]\to\mathbb{C}$ est un chemin, alors

$$\int_\gamma f = \int_\gamma f(z)dz = \int_a^b f\big(\gamma(t)\big)\gamma'(t)dt. \tag{14.53}$$

14.2.3 Intégrer un champ de vecteurs

Définition 14.22.
*Un **champ de vecteur** est une application $G : \mathbb{R}^n \to \mathbb{R}^n$. On définit l'intégrale de G sur un chemin $\gamma : [a,b] \to \mathbb{R}^n$ par*

$$\int_\gamma G \stackrel{def}{=} \int_a^b \big\langle G(\gamma(t)), \gamma'(t)\big\rangle\, dt.$$

Remarque 14.23.
Cette définition ne dépend pas de la paramétrisation choisie, mais le signe change selon le sens du chemin.

Si $\sigma'(t)\neq 0$, nous pouvons considérer le vecteur unitaire tangent à la courbe :

$$T(t) = \frac{\sigma'(t)}{\|\sigma'(t)\|}. \tag{14.54}$$

Si F est un champ de vecteurs sur $\mathbb{R}^3$, la circulation de F le long de σ sera donnée par

$$\int_\sigma F\cdot ds = \int_a^b F\big(\sigma(t)\big)\cdot\sigma'(t)dt = \int_a^b F\big(\sigma(t)\big)\cdot\frac{\sigma'(t)}{\|\sigma'(t)\|}dt = \int_\sigma F\cdot T ds \tag{14.55}$$

où dans la dernière expression, $F\cdot T$ est vu comme fonction $(x,y,z)\mapsto F(x,y,z)\cdot T(x,y,z)$. L'intégrale d'un champ de vecteurs sur une courbe n'est donc rien d'autre que l'intégrale de la composante tangentielle du champ de vecteurs.

14.2.4 Intégrer une forme différentielle sur un chemin

La formule d'intégration d'un champ de vecteur [2],

$$\int_\gamma G = \int_{[a,b]} \langle G(\gamma(t)), \gamma'(t)\rangle dt, \tag{14.56}$$

contient quelque chose d'intéressant : la combinaison $\langle G(\gamma(t)), \gamma'(t)\rangle$. Cette combinaison sert à transformer le vecteur tangent $\gamma'(t)$ en un nombre en utilisant le produit scalaire avec le vecteur $G(\gamma(t))$.

Si G est un champ de vecteur sur $\mathbb{R}^n$, et si $x\in\mathbb{R}^n$, nous pouvons utiliser l'isomorphisme musical (définition 12.366)

$$\begin{aligned} G_x^\flat\colon \mathbb{R}^n &\to \mathbb{R} \\ v &\mapsto \langle G(x), v\rangle \end{aligned} \tag{14.57}$$

pour écrire de façon plus compacte :

$$\int_\gamma G = \int_{[a,b]} G^\flat_{\gamma(t)}\big(\gamma'(t)\big)dt. \tag{14.58}$$

L'intégrale d'une forme différentielle sur un chemin est définie par

$$\int_\gamma \omega = \int_a^b \omega_{\gamma(t)}\gamma'(t)dt \tag{14.59}$$

Remarque 14.24.
Cette définition ne dépend pas de la paramétrisation choisie, mais le signe change selon le sens du chemin.

2. Définition 14.22.

14.2.5 Intégration d'une forme différentielle sur un chemin

Les formes intégrales que nous avons déjà vues sont celles de fonctions et de champs de vecteur sur des chemins. Si $\gamma\colon [a,b]\to\mathbb{R}^n$ est le chemin, les formules sont

$$\begin{aligned}\int_\gamma f &= \int_{[a,b]} f\big(\gamma(t)\big)\|\gamma'(t)\|dt\\ \int_\gamma G &= \int_{[a,b]}\langle G\big(\gamma(t)\big),\gamma'(t)\rangle dt.\end{aligned}\tag{14.60}$$

Dans les deux cas, le principe est que nous disposons de quelque chose (la fonction f ou le vecteur G), et du vecteur tangent $\gamma'(t)$, et nous essayons d'en tirer un nombre que nous intégrons. Lorsque nous avons une 1-forme, la façon de l'utiliser pour produire un nombre avec le vecteur tangent est évidement d'appliquer la 1-forme au vecteur tangent. La définition suivante est donc naturelle.

Définition 14.25.
*Soit $\gamma\colon [a,b]\to\mathbb{R}^n$, un chemin de classe C^1 tel que son image est contenue dans le domaine D. Si ω es une 1-forme différentielle sur D, nous définissons l'**intégrale de ω le long de** γ le nombre*

$$\begin{aligned}\int_\gamma\omega &= \int_a^b \omega_{\gamma(t)}\big(\gamma'(t)\big)dt\\ &= \int_a^b\Big[a_1\big(\gamma(t)\big)\gamma_1'(t)+\cdots+a_n\big(\gamma(t)\big)\gamma_n'(t)\Big]dt.\end{aligned}\tag{14.61}$$

Cette définition est une bonne définition parce que si on change la paramétrisation du chemin, on ne change pas la valeur de l'intégrale, c'est la proposition suivante.

Proposition 14.26.
Si γ et β sont des chemins équivalents, alors

$$\int_\gamma\omega=\int_\beta\omega,\tag{14.62}$$

c'est à dire que l'intégrale est invariante sous les reparamétrisation du chemin.

Démonstration. Deux chemins sont équivalents quand il existe un difféomorphisme C^1 $h\colon[a,b]\to[c,d]$ tel que $\gamma(t)=(\beta\circ h)(t)$. En remplaçant γ par $(\beta\circ h)$ dans la définition de $\int_\gamma\omega$, nous trouvons

$$\int_a^b\omega_{\gamma(t)}\big(\gamma'(t)\big)dt=\int_a^b\omega_{(\beta\circ h)(t)}\big((\beta\circ h)'(t)\big)dt.\tag{14.63}$$

Un changement de variable $u=h(t)$ transforme cette dernière intégrale en $\int_\beta\omega$, ce qui prouve la proposition. □

Remarque 14.27.
Si γ est une somme de chemins, $\gamma=\gamma^{(1)}+\cdots+\gamma^{(n)}$, où chacun des $\gamma^{(i)}$ est un chemin, alors

$$\int_\gamma\omega=\sum_{i=1}^n\int_{\gamma_i}\omega\tag{14.64}$$

parce que ω est linéaire.

Remarque 14.28.
Si $-\gamma$ est le chemin

$$\begin{aligned}-\gamma\colon[a,b]&\to\mathbb{R}^n\\ t&\mapsto\gamma\big(b-(t-a)\big),\end{aligned}\tag{14.65}$$

alors

$$\int_{-\gamma}\omega=-\int_\gamma\omega,\tag{14.66}$$

c'est à dire que si l'on parcours le chemin en sens inverse, alors on change le signe de l'intégrale.

L'intégrale d'une forme différentielle sur un chemin est compatible avec l'intégrale déjà connue d'un champ de vecteur sur le chemin parce que si G est un champ de vecteurs,

$$\int_\gamma G^\flat = \int_\gamma G. \tag{14.67}$$

En effet,

$$\begin{aligned}\int_{\gamma G^\flat} &= \int_a^b G^\flat_{\gamma(t)}(\gamma'(t)) \\ &= \int_a^b \big[G_1(\gamma(t))dx_1 + \dots G_n(\gamma(t))dx_n\big]\big(\gamma_1'(t), \dots, \gamma_n'(t)\big) \\ &= \int_a^b \langle G(\gamma(t)), \gamma'(t)\rangle \\ &= \int_\gamma G.\end{aligned} \tag{14.68}$$

Proposition 14.29.
Soit $\omega = df$, une 1-forme exacte et continue sur le domaine D. Alors la valeur de $\int_\gamma df$ ne dépend que des valeurs de f aux extrémités de γ.

Démonstration. Nous avons

$$\begin{aligned}\int_\gamma \omega = \int_\gamma df &= \int_a^b \sum_{i=1} n \frac{\partial f}{\partial x_i}\big(\gamma(t)\big)\gamma_i'(t)dt \\ &= \int_a^b \frac{d}{dt}\Big((f\circ\gamma)(t)\Big)dt \\ &= (f\circ\gamma)(b) - (f\circ\gamma(a)).\end{aligned} \tag{14.69}$$

□

14.2.6 Interprétation physique : travail

Définition 14.30.
*Une force $F\colon D \subset \mathbb{R}^n \to \mathbb{R}^n$ est **conservative** si elle dérive d'un potentiel, c'est à dire s'il existe une fonction $V \in C^1(D, \mathbb{R})$ telle que*

$$F(x) = (\nabla V)(x). \tag{14.70}$$

Étant donné que F est un champ de vecteurs, nous avons une forme différentielle associée $F^\flat$,

$$F^\flat_x\colon x \mapsto \langle F(x), v\rangle. \tag{14.71}$$

Lemme 14.31.
Le champ F est conservatif si et seulement si la 1-forme différentielle $F^\flat$ est exacte.

Démonstration. Supposons que la force F soit conservative, c'est à dire qu'il existe une fonction V telle que $F = \nabla V$. Dans ce cas, il est facile de prouver que $F^\flat$ est exacte et est donnée par $F^\flat_x = dV(x)$. En effet,

$$\begin{aligned}F^\flat_x(v) &= \langle F(x), v\rangle \\ &= F_1(x)v_1 + \dots + F_n(x)v_n \\ &= \frac{\partial V}{\partial x_1}(x)v_1 + \dots \frac{\partial V}{\partial x_n}(x)v_n \\ &= dV(x)v.\end{aligned} \tag{14.72}$$

Pour le sens inverse, supposons que $F^\flat$ soit exacte. Dans ce cas, nous avons une fonction V telle que $F^\flat = dV$. Il est facile de prouver qu'alors, $F = \nabla V$. □

En résumé, nous avons deux façons équivalentes d'exprimer que la force F dérive du potentiel V : soit nous disons $F = \nabla V$, soit nous disons $F^\flat = dV$.

Proposition 14.32.
Si F est une force conservative, alors le travail de F lors d'un déplacement ne dépend pas du chemin suivit.

Démonstration. Le travail d'une force le long d'un chemin n'est autre que l'intégrale de la force le long du chemin, et le calcul est facile :

$$W_\gamma(F) = \int_\gamma F = \int_\gamma dV = V\big(\gamma(b)\big) - V\big(\gamma(a)\big). \tag{14.73}$$

Donc si β est un autre chemin tel que $\beta(a) = \gamma(a)$ et $\beta(b) = \gamma(b)$, nous avons $W_\beta(F) = W_\gamma(F)$. □

14.2.7 Intégrer un champs de vecteurs sur un bord en $2D$

Si $D \subset \mathbb{R}^2$ est tel que ∂D est une variété de dimension 1 et tel que D accepte un champ de vecteur normal extérieur unitaire ν. Si nous voulons définir

$$\int_{\partial D} G, \tag{14.74}$$

le mieux est de prendre une paramétrisation $\gamma\colon [0,1] \to \mathbb{R}^2$ et de calculer

$$\int_0^1 \langle G_{\gamma(t)}, \frac{\dot\gamma(t)}{\|\dot\gamma(t)\|}\rangle dt. \tag{14.75}$$

Hélas, cette définition ne fonctionne pas parce que son signe dépend du sens de la paramétrisation γ. Si la paramétrisation tourne dans l'autre sens, il y a un signe de différence.

Nous allons définir

$$\int_{\partial D} G = \int_0^1 \langle G_{\gamma(t)}, T(t)\rangle dt \tag{14.76}$$

où $T(t) = \dot\gamma(t)/\|\dot\gamma(t)\|$ et où γ est choisit de telle façon que la rotation d'angle $\frac{\pi}{2}$ amène ν sur T. Cela fixe le choix de sens.

Ce choix de sens aura des répercussions dans l'application de la formule de Green et du théorème de Stokes.

14.2.8 Intégrer une forme différentielle sur un bord en $2D$

Nous n'allons pas chercher très loin :

$$\int_{\partial D} \omega = \int_{\partial D} \omega^\sharp, \tag{14.77}$$

c'est à dire que l'intégrale de la forme différentielle est celle du champ de vecteur associé. Le membre de droite est définit par (14.76), avec le choix d'orientation qui va avec.

14.2.9 Intégrer une forme différentielle sur un bord en $3D$

Nous allons maintenant intégrer une forme différentielle sur certains chemins fermés dans $\mathbb{R}^3$. Soit $F(D) \subset \mathbb{R}^3$, une variété de dimension 2 dans $\mathbb{R}^3$ où $F\colon D \subset \mathbb{R}^2 \to \mathbb{R}^3$ est la carte. Nous supposons que D vérifie les hypothèses de la formule de Green. Alors nous définissons

$$\int_{F(\partial D)} \omega = \int_{\partial D} F^*\omega \tag{14.78}$$

où $F^*\omega$ est la forme différentielle définie sur ∂D par $(F^*\omega)(v) = \omega\big(dF(v)\big)$.

Cette définition est très abstraite, mais nous n'allons, en pratique, jamais l'utiliser, grâce au théorème de Stokes.

14.2.10 Intégrer d'un champ de vecteurs sur un bord en $3D$

Encore une fois, nous n'allons pas chercher bien loin :

$$\int_{F(\partial D)} G = \int_{F(\partial D)} G^\flat \tag{14.79}$$

où $G^\flat$ est la forme différentielle associée au champ de vecteur. Le membre de droite est définit par l'équation (14.78).

14.2.11 Dérivées croisées et forme différentielle exacte

Nous considérons le problème suivant : trouver une fonction $f\colon \mathbb{R}^2 \to \mathbb{R}$ telle que

$$\begin{cases} \dfrac{\partial f}{\partial x} = a(x,y) & \text{(14.80a)} \\ \dfrac{\partial f}{\partial y} = b(x,y) & \text{(14.80b)} \end{cases}$$

où a et b sont des fonctions supposées suffisamment régulières. Nous savons que ce problème n'a pas de solutions lorsque

$$\frac{\partial a}{\partial y} \neq \frac{\partial b}{\partial x} \tag{14.81}$$

parce que cela impliquerait $\partial^2_{xy} f \neq \partial^2_{yx} f$. Nous sommes en droit de nous demander si la condition

$$\frac{\partial a}{\partial y} = \frac{\partial b}{\partial x} \tag{14.82}$$

impliquerait qu'il existe une solution au problème (14.80). La réponse est oui, et nous allons brièvement la justifier. Pour plus de détails nous vous demandons de chercher un peu sur internet les mots-clefs *forme différentielles exacte*. Vous consulterez également avec profit [186].

Proposition 14.33.
Si a et b sont des fonctions qui satisfont à la condition

$$\frac{\partial a}{\partial y} = \frac{\partial b}{\partial x}, \tag{14.83}$$

alors la fonction

$$f(x,y) = \int_0^x a(t,0)dt + \int_0^y b(x,t)dt \tag{14.84}$$

répond au problème

$$\begin{cases} \dfrac{\partial f}{\partial x} = a(x,y) & \text{(14.85a)} \\ \dfrac{\partial f}{\partial y} = b(x,y) & \text{(14.85b)} \end{cases}$$

La preuve qui suit n'en est pas complètement une parce qu'il manque des justification, notamment au moment de permuter la dérivée et l'intégrale.

Démonstration. La clef de la preuve est le théorème fondamental de l'analyse :

$$\int_0^x \frac{\partial f}{\partial x}(t,y)dt = f(x,y) \tag{14.86}$$

et son pendant par rapport à y :

$$\int_0^y \frac{\partial f}{\partial y}(x,t)dt = f(x,y). \tag{14.87}$$

En appliquant ces version du théorème fondamental, nous obtenons immédiatement.

$$\frac{\partial f}{\partial y} = b(x, y). \tag{14.88}$$

En ce qui concerne la dérivée par rapport à y,

$$\begin{aligned}
\frac{\partial f}{\partial x} &= a(x, 0) + \int_0^y \frac{\partial b}{\partial x}(x, t)dt && \text{(14.89a)}\\
&= a(x, 0) + \int_0^y \frac{\partial a}{\partial y}(x, t)dt && \text{(14.89b)}\\
&= a(x, 0) + [a(x, t)]_{t=0}^{t=y} && \text{(14.89c)}\\
&= a(x, y). && \text{(14.89d)}
\end{aligned}$$

□

En ce qui concerne l'unicité, supposons que f et g soient deux solutions au problème. L'équation

$$\frac{\partial f}{\partial x} = a(x, y) = \frac{\partial g}{\partial x} \tag{14.90}$$

implique que

$$f(x, y) = g(x, y) + C(y) \tag{14.91}$$

où C est une fonction seulement de y. L'autre équation implique

$$f(x, y) = g(x, y) + D(x) \tag{14.92}$$

où D est seulement une fonction de x. En égalisant nous voyons que les fonctions C et D doivent être des constantes.

Par conséquent la fonction f est donnée à une constante près et en réalité la fonction (14.84) est suffisante pour répondre au problème de trouver toutes les fonctions dont les dérivées partielles sont données par les fonctions a et b.

La fonction f ainsi créée est un **potentiel** pour le champ de force

$$F(x, y) = \begin{pmatrix} a(x, y) \\ b(x, y) \end{pmatrix}. \tag{14.93}$$

Notez que ce champ de vecteurs est le gradient de f. La question initiale aurait donc pu être posée en les termes suivants : trouver une fonction f dont le gradient est donné par

$$\nabla f = \begin{pmatrix} a(x, y) \\ b(x, y) \end{pmatrix}. \tag{14.94}$$

14.3 Surfaces paramétrées

De la même façon qu'un chemin dans $\mathbb{R}^3$ est décrit comme une application $\sigma\colon \mathbb{R} \to \mathbb{R}^3$, une surface dans $\mathbb{R}^3$ sera vue comme une application $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^3$. Une **surface paramétrée** dans $\mathbb{R}^3$ est une application

$$\begin{aligned}
&\varphi\colon D \subset \mathbb{R}^2 \to \mathbb{R}^3\\
&(u, v) \mapsto \varphi(u, v) = \begin{pmatrix} x(u, v) \\ y(u, v) \\ z(u, z) \end{pmatrix}.
\end{aligned} \tag{14.95}$$

Nous allons parler de la «surface φ» pour désigner l'image de φ dans $\mathbb{R}^3$.

Si on fixe le paramètre $u - u_0$, alors l'application

$$v \mapsto \varphi(u_0, v) \tag{14.96}$$

est un chemin dans la surface. Un vecteur tangent à ce chemin sera tangent à la courbe :

$$\frac{\partial\varphi}{\partial v}(u_0,v_0)=\begin{pmatrix}\frac{\partial x}{\partial v}(u_0,v_O)\\ \frac{\partial y}{\partial v}(u_0,v_O)\\ \frac{\partial z}{\partial v}(u_0,v_O)\end{pmatrix}. \tag{14.97}$$

De même, en fixant v_0, on considère le chemin

$$u\mapsto\varphi(u,v_0). \tag{14.98}$$

Le vecteur tangent à ce chemin est égalent tangent à la surface :

$$\frac{\partial\varphi}{\partial u}=\begin{pmatrix}\frac{\partial x}{\partial u}(u_0,v_O)\\ \frac{\partial y}{\partial u}(u_0,v_O)\\ \frac{\partial z}{\partial u}(u_0,v_O)\end{pmatrix} \tag{14.99}$$

Définition 14.34.
Nous disons que la surface est ***régulière*** *si les vecteurs $\partial_u\varphi(u_0,v_0)$ et $\partial_v\varphi(u_0,v_0)$ sont non nuls et non colinéaires.*

Si la surface est régulière, les vecteurs tangents à la paramétrisation forment le plan tangent à la surface au point $\varphi(u_0,v_0)$.

Un vecteur orthogonal à la surface (et donc au plan tangent) est donc donné par le produit vectoriel :

$$n(u_0,v_0)=\frac{\partial\varphi}{\partial u}(u_0,v_0)\times\frac{\partial\varphi}{\partial v}(u_0,v_0). \tag{14.100}$$

L'équation du plan tangent est alors obtenue par

$$\begin{pmatrix}x-x_0\\ y-y_0\\ z-z_0\end{pmatrix}\cdot n(u_0,v_0)=0 \tag{14.101}$$

où $x_0=x(u_0,v_0)$, $y_0=y(u_0,v_0)$, $z_0=z(u_0,v_0)$.

14.3.1 Graphe d'une fonction

Soit la fonction $f\colon D\subset\mathbb{R}^2\to\mathbb{R}$. Le graphe de f est l'ensemble des points de la forme

$$\big(x,y,f(x,y)\big) \tag{14.102}$$

tels que $(x,y)\in D$. Cela est une surface paramétrée par

$$\begin{aligned}&\varphi\colon D\to\mathbb{R}^3\\ &(x,y)\mapsto\begin{pmatrix}x\\ y\\ f(x,y)\end{pmatrix}.\end{aligned} \tag{14.103}$$

Les vecteurs tangents sont

$$\frac{\partial\varphi}{\partial x}=\begin{pmatrix}1\\ 0\\ \frac{\partial f}{\partial x}\end{pmatrix},\quad\frac{\partial\varphi}{\partial y}=\begin{pmatrix}0\\ 1\\ \frac{\partial\varphi}{\partial y}\end{pmatrix}. \tag{14.104}$$

La surface est donc partout régulière parce que ces deux vecteurs ne sont jamais nuls ou colinéaires. Un vecteur normal à cette surface au point $(x_0,y_0,f(x_0,y_0))$ est donné par le produit vectoriel

$$n=\begin{vmatrix}e_x & e_y & e_z\\ 1 & 0 & \partial_x f(x_0,y_0)\\ 0 & 1 & \partial_y f(x_0,y_0)\end{vmatrix}=-\frac{\partial f}{\partial x}(x_0,y_0)e_x-\frac{\partial f}{\partial y}(x_0,y_0)e_y+e_z. \tag{14.105}$$

En suivant l'équation (14.101), nous avons l'équation suivante pour le plan :

$$\begin{pmatrix} x-x_0 \\ y-y_0 \\ z-z_0 \end{pmatrix} \cdot \begin{pmatrix} -\frac{\partial f}{\partial x}(x_0,y_0) \\ -\frac{\partial f}{\partial y}(x_0,y_0) \\ 1 \end{pmatrix} = 0, \tag{14.106}$$

c'est à dire

$$-(x-x_0)\frac{\partial f}{\partial x}(x_0,y_0) - (y-y_0)\frac{\partial f}{\partial y}(x_0,y_0) + z - f(x_0,y_0) = 0, \tag{14.107}$$

ce qui revient à

$$z - f(x_0,y_0) = \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0). \tag{14.108}$$

Nous retrouvons donc l'équation du plan tangent à un graphe.

Exemple 14.35
La sphère de rayon R peut être paramétrée par les angles sphériques :

$$\phi(\theta,\varphi) = \begin{pmatrix} R\sin\theta\cos\varphi \\ R\sin\theta\sin\varphi \\ R\cos\theta \end{pmatrix} \tag{14.109}$$

avec $(\theta,\varphi) \in [0,\pi] \times [0,2\pi]$.

Tentons d'en trouver le plan tangent au point $(x,y,z) = (R,0,0)$. Un petit dessin nous montre que c'est un plan vertical d'équation $x = R$. Montrons cela en utilisant la théorie que nous venons de découvrir. D'abord le point $(R,0,0)$ correspond à $\theta_0 = \frac{\pi}{2}$ et $\varphi = 0$. Les vecteurs tangents sont

$$T_\theta = \frac{\partial \phi}{\partial \theta}(R,\frac{\pi}{2},0) = \begin{pmatrix} R\cos\theta\cos\varphi \\ R\cos\theta\sin\varphi \\ -R\sin\theta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -R \end{pmatrix}, \tag{14.110}$$

et

$$T_\varphi = \frac{\partial \phi}{\partial \varphi}(R,\frac{\pi}{2},0) = \begin{pmatrix} -R\sin\theta\sin\varphi \\ R\sin\theta\cos\varphi \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ R \\ 0 \end{pmatrix}. \tag{14.111}$$

Cela sont de toute évidence bien les deux vecteurs tangents à la sphère au point $(x,y,z) = (R,0,0)$. Le vecteur normal est

$$\begin{vmatrix} e_x & e_y & e_z \\ 0 & 0 & -R \\ 0 & R & 0 \end{vmatrix} = R^2 e_x. \tag{14.112}$$

Ici encore, nous avons le vecteur que nous attendions sur un dessin. L'équation du plan tangent est maintenant

$$\begin{pmatrix} x-R \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} R^2 \\ 0 \\ 0 \end{pmatrix} = 0, \tag{14.113}$$

c'est à dire $R^2(x-R) = 0$ et donc $x = R$. △

14.3.2 Intégrale sur une partie de $\mathbb{R}^m$

Soit M une variété de dimension n dans $\mathbb{R}^m$. Soit $F : W \subset \mathbb{R}^n \to M$ une paramétrisation d'un ouvert relatif de M.

Si f est une fonction définie sur un sous-ensemble $A \subset F(W)$ tel que $F^{-1}(A)$ est mesurable, l'*intégrale de f sur A* est définie par

$$\int_A f = \int_{F^{-1}(A)} f(F(w))\sqrt{\det(J_F(w)^t J_F(w))}dw$$

où l'intégrale est l'intégration usuelle (de Lebesgue) sur $F^{-1}(A) \subset \mathbb{R}^n$. On écrit parfois cette intégrale $\int_{F^{-1}(A)} f(F(w))d\sigma$ où

$$d\sigma = \sqrt{\det(J_F(w)^t J_F(w))}dw \tag{14.114}$$

est l'*élément infinitésimal de volume* de la variété.

Si $m = 3$ et $n = 2$, l'élément infinitésimal de volume vaut

$$d\sigma = \left\| \frac{\partial F}{\partial w_1} \times \frac{\partial F}{\partial w_2} \right\| dw$$

où $\times$ représente le produit vectoriel dans $\mathbb{R}^3$, et (w_1, w_2) sont les coordonnées sur $W \subset \mathbb{R}^2$. Dans la suite, nous ne regarderons plus que ce cas.

14.4 Intégrales de surface

14.4.1 Intégrale d'un champ de vecteurs

Dans l'intégration curviligne, on a noté que si l'intégrale d'une fonction ne dépendait pas de l'orientation du chemin, l'intégrale d'un champ de vecteurs ou d'une forme différentielle en dépendait. Ce problème d'orientation apparait également dans l'intégration sur des surfaces de l'espace.

Une *orientation* sur une surface $S \subset \mathbb{R}^3$ est le choix d'un champ de vecteurs continu $\nu : S \to \mathbb{R}^3$ dont la norme en tout point de S vaut 1. On remarque qu'ayant fait un tel choix d'orientation $\nu(x)$ en un point x, le seul autre choix possible en x est $-\nu(x)$. Si S est le bord d'un ouvert $D \subset \mathbb{R}^3$, l'*orientation induite par D sur S* est, si elle existe, l'orientation qui pointe hors de D en tout point de S. Plus précisément, il faut que pour tout $x \in D$ il existe $\epsilon > 0$ vérifiant, pour tout $0 < t < \epsilon$, la relation $t\nu(x) \notin D$. Dans ce cas, le champ de vecteurs ν est appelé le *vecteur normal unitaire extérieur* à D et il est forcément unique.

Soit G un champ de vecteurs défini sur une surface orientée par un champ ν. L'intégrale de G sur S, aussi appelée le *flux de G à travers S*, est

$$\iint_S G \cdot dS \stackrel{\text{def}}{=} \iint_S \langle G, \nu \rangle \, d\sigma. \tag{14.115}$$

Si on suppose que la surface est paramétrée par une application

$$F : W \subset \mathbb{R}^2 \to \mathbb{R}^3 : (u, v) \mapsto (F_1(u, v), F_2(u, v), F_3(u, v))$$

alors un vecteur unitaire ν peut s'écrire sous la forme

$$\nu = \frac{\frac{\partial F}{\partial u} \times \frac{\partial F}{\partial v}}{\left\| \frac{\partial F}{\partial u} \times \frac{\partial F}{\partial v} \right\|}$$

et grâce à cette paramétrisation l'intégrale (14.115) devient

$$\iint_S G \cdot dS = \iint_W \left\langle G(F(u, v)), \frac{\partial F}{\partial u} \times \frac{\partial F}{\partial v} \right\rangle du dv.$$

où on utilise l'expression de $d\sigma$ obtenue précédemment dans le cas qui nous intéresse (surface dans l'espace).

14.5 Intégrales de surface

14.5.1 Aire d'une surface paramétrée

Lorsque nous avions vu la longueur d'une courbe paramétrée, nous avions pris comme «élément de longueur» la norme du vecteur tangent. Il est donc naturel de prendre comme «élément de surface» une petite surface que l'on peut construire à partir des deux vecteurs tangents à la surface.

Au point $\varphi(u_0, v_0)$, nous avons les deux vecteurs tangents

$$T_u = \frac{\partial\varphi}{\partial u}(u_0, v_0) \quad T_v = \frac{\partial\varphi}{\partial v}(u_0, v_0). \tag{14.116}$$

L'élément de surface que nous pouvons construire à partir de ces deux vecteurs est la surface du parallélogramme, donnée par la norme du produit vectoriel :

$$dS = \|T_u \times T_v\|. \tag{14.117}$$

L'aire de la surface donné par $\varphi\colon D \subset \mathbb{R}^2 \to \mathbb{R}^3$ sera donc donnée par

$$Aire\big(\varphi(D)\big) = \iint_D \|T_u \times T_v\| du\, dv. \tag{14.118}$$

Exemple 14.36
Calculons l'aire de la sphère. Les vecteurs tangents ont déjà été calculés aux équations (14.110) et (14.111) :

$$T_\theta = \begin{pmatrix} R\cos\theta\cos\varphi \\ R\cos\theta\sin\varphi \\ -R\sin\theta \end{pmatrix}, \quad T_\varphi = \begin{pmatrix} -R\sin\theta\sin\varphi \\ R\sin\theta\cos\varphi \\ 0 \end{pmatrix}. \tag{14.119}$$

Le produit vectoriel vaut

$$\begin{aligned} T_\theta \times T_\varphi &= \begin{vmatrix} e_x & e_y & e_z \\ R\cos\theta\cos\varphi & R\cos\theta\sin\varphi & -R\sin\theta \\ -R\sin\theta\sin\varphi & R\sin\theta\cos\varphi & 0 \end{vmatrix} \\ &= (R^2\sin^2\theta\cos\varphi)e_x + (R^2\sin^2\theta\sin\varphi)e_y \\ &\quad + (R^2\cos\theta\sin\theta\cos^2\varphi + R^2\sin\theta\cos\theta\sin^2\varphi)e_z. \end{aligned} \tag{14.120}$$

La norme demande quelques calculs et mises en évidences. Le résultat est :

$$\|T_\theta \times T_\varphi\| = R^2\sin\theta. \tag{14.121}$$

L'aire de la sphère est donc donnée par

$$Aire = \int_0^{2\pi} d\varphi \int_0^{\pi} R^2\sin\theta d\theta = 2\pi R^2[-\cos\theta]_0^{\pi} = 4\pi R^2. \tag{14.122}$$

Il est bon de se souvenir que, en coordonnées sphériques,

$$\|T_\theta \times T_\varphi\| = R^2\sin\theta. \tag{14.123}$$

Or nous savons que ce vecteur est dirigé dans le sens de e_r parce que ce dernier est le vecteur qui est constamment dirigé radialement. En coordonnées sphériques nous avons donc

$$T_\theta \times T_\varphi = R^2\sin(\theta)e_r. \tag{14.124}$$

△

Remarque 14.37.
L'équation (14.121) donne l'élément de surface pour la sphère. Notez que cela est justement l'expression du jacobien des coordonnées sphériques. Cela n'est évidemment pas une coïncidence.

Exemple 14.38
Nous pouvons donner l'aire du graphe d'une fonction quelconque. La surface est paramétrée par

$$\varphi(x,y) = \begin{pmatrix} x \\ y \\ f(x,y) \end{pmatrix}. \tag{14.125}$$

Les vecteurs tangents sont

$$T_x = \begin{pmatrix} 1 \\ 0 \\ \partial_x f \end{pmatrix}, \quad T_y = \begin{pmatrix} 0 \\ 1 \\ \partial_y f \end{pmatrix}. \tag{14.126}$$

Le produit vectoriel est donné par

$$T_x \times T_y = \begin{vmatrix} e_x & e_y & e_z \\ 1 & 0 & \partial_x f \\ 0 & 1 & \partial_y f \end{vmatrix} = (-\partial_x f)e_x - (\partial_y f)e_y + e_z. \tag{14.127}$$

L'élément de surface est par conséquent

$$dS = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + 1}, \tag{14.128}$$

et la surface du graphe sera

$$Aire = \iint_D \sqrt{\left(\frac{\partial f}{\partial x}(x,y)\right)^2 + \left(\frac{\partial f}{\partial y}(x,y)\right)^2 + 1}\, dx\, dy \tag{14.129}$$

△

14.5.2 Intégrale d'une fonction sur une surface

Si S est une surface dans $\mathbb{R}^3$ paramétrée par

$$\begin{aligned} &\varphi\colon D \to \mathbb{R}^3 \\ &(u,v) \mapsto \varphi(u,v) \in S, \end{aligned} \tag{14.130}$$

et si f est une fonction $f\colon \mathbb{R}^3 \to \mathbb{R}$ définie au moins sur S, l'intégrale de f sur S est logiquement définie par

$$\int_S f\, dS = \iint_D f\big(\varphi(u,v)\big)\|T_u(u,v) \times T_v(u,v)\| du dv \tag{14.131}$$

où $T_u = \frac{\partial \varphi}{\partial u}$ et $Y_v = \frac{\partial \varphi}{\partial v}$. La quantité

$$\|T_u(u,v) \times T_v(u,v)\| du dv \tag{14.132}$$

est appelé **élément de surface**.

Encore une fois, si on prend $f = 1$, alors on retrouve la surface de S :

$$\int_S dS = Aire(S). \tag{14.133}$$

Remarque 14.39.
Le nombre $\int_S f dS$ ne dépend pas de la paramétrisation choisie pour S.

14.5.3 Aire d'une surface de révolution

Soit γ une courbe dans le plan xy, paramétrée par

$$\gamma(u) = \begin{pmatrix} x(u) \\ y(u) \\ 0 \end{pmatrix} \tag{14.134}$$

avec $u \in [a, b]$. Nous supposons que la courbe est toujours positive, c'est à dire $y(u) > 0$ pour tout u.

Nous voulons considérer la surface obtenue en effectuant une rotation de cette ligne autour de l'axe X. Chaque point de la courbe va parcourir un cercle de rayon $y(u)$ dans le plan YX et centré en $(x(u), 0, 0)$. La surface est donc donnée par

$$\varphi(u, \theta) = \begin{pmatrix} x(u) \\ y(u)\cos\theta \\ y(u)\sin\theta \end{pmatrix} \tag{14.135}$$

avec $(u, \theta) \in [a, b] \times [0, 2\pi]$. Notez que la courbe de départ correspond à $\theta = 0$.

Les vecteurs tangents à la surface pour cette paramétrisation sont

$$T_u = \frac{\partial\varphi}{\partial u} = \begin{pmatrix} x'(u) \\ y'(u)\cos\theta \\ y'(u)\sin\theta \end{pmatrix} \quad T_\theta = \frac{\partial\varphi}{\partial\theta} = \begin{pmatrix} 0 \\ -y(u)\sin\theta \\ y(u)\cos\theta \end{pmatrix}. \tag{14.136}$$

Le produit vectoriel de ces deux vecteurs vaut

$$\begin{aligned} T_u \times T_\theta &= \begin{vmatrix} e_x & e_y & e_z \\ x' & y'\cos\theta & y'\sin\theta \\ 0 & -y\sin\theta & y\cos\theta \end{vmatrix} \\ &= y'(u)y(u)\,e_x - x'(u)y(u)\cos\theta\,e_y + x'(u)y(u)\sin\theta\,e_z. \end{aligned} \tag{14.137}$$

En ce qui concerne la norme :

$$dS = \|T_u \times T_\theta\| = \sqrt{(y'y)^2 + (x'y)^2} = |y(u)|\sqrt{y'(u)^2 + x'(u)^2}. \tag{14.138}$$

Étant donné que nous avons supposé que $y(u) > 0$, nous pouvons supprimer les valeurs absolues, et l'aire de la surface de révolution devient :

$$\begin{aligned} Aire(S) &= \int_0^{2\pi} d\theta \int_a^b y(u)\sqrt{x'(u)^2 + y'(u)^2}du \\ &= 2\pi \int_a^b y(u)\sqrt{x'(u)^2 + y'(u)^2}du. \end{aligned} \tag{14.139}$$

Exemple 14.40
Calculons la surface du cône de révolution de rayon (à la base) R et de hauteur h. La courbe de départ est le segment droite qui part de $(0, 0)$ et qui termine en (R, h) de la figure 14.1.

Ce segment peut être paramétré par

$$\gamma(u) = \begin{pmatrix} Ru \\ hu \\ 0 \end{pmatrix} \tag{14.140}$$

avec $u \in [0, 1]$. Cela donne $x(u) = Ru$, $y(u) = hu$ et par conséquent

$$Aire = 2\pi \int_0^1 hu\sqrt{R^2 + h^2} = \pi h\sqrt{R^2 + h^2}. \tag{14.141}$$

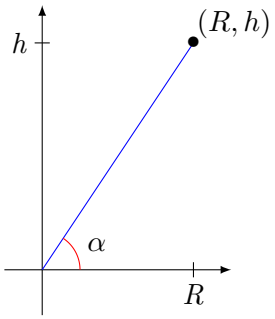


FIGURE 14.1 – En faisant tourner cette droite autour de l'axe X, nous obtenons un cône.

Nous pouvons aussi exprimer ce résultat en fonction de l'angle, en sachant que $h = \sqrt{h^2 + R^2}\sin(\alpha)$:

$$Aire = \pi(R^2 + h^2)\sin(\alpha). \tag{14.142}$$

△

Exemple 14.41
Calculons la surface latérale du tore obtenu par révolution du cercle de la figure 14.2.

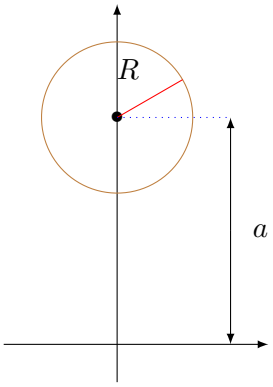


FIGURE 14.2 – Si nous tournons ce cercle autour de l'axe X, nous obtenons un tore de rayon «externe» a et de rayon «interne» R.

Le chemin qui détermine le cercle de départ est

$$\gamma(u) = \begin{pmatrix} R\cos(u) \\ a + R\sin(u) \\ 0 \end{pmatrix}, \tag{14.143}$$

c'est à dire $x(u) = R\cos(u)$, $y(u) = a + R\sin(u)$ avec $u \in [0, 2\pi]$. Nous avons donc l'aire

$$\begin{aligned} Aire &= 2\pi \int_0^{2\pi} \big(a + R\sin(u)\big) R\, du \\ &= 2\pi R\big(2\pi a + R[-\cos(u)]_0^{2\pi}\big) \\ &= 4\pi^2 aR. \end{aligned} \tag{14.144}$$

△

14.5.4 Intégrale d'une 2-forme

Nous considérons ω, une 2-forme différentielle sur $\mathbb{R}^2$.

Définition 14.42.
Si $\omega_{(x,y)} = u(x,y)dx \wedge dy$ et si D est un ouvert de $\mathbb{R}^2$ alors nous définissons

$$\int_D \omega = \int_D u(x,y)dx\,dy. \tag{14.145}$$

Nous voulons maintenant intégrer une 2-forme sur une surface dans $\mathbb{R}^3$. Soit $S \subset \mathbb{R}^3$, une surface orientée (c'est à dire que nous avons un choix continu d'un vecteur normal unitaire n). Nous supposons de plus avoir un paramétrage $\phi\colon D \to S$ de S avec D ouvert dans $\mathbb{R}^2$ compatible avec l'orientation, c'est à dire que pour tout $(t,s) \in D$,

$$n\big(\phi(t,s)\big) = \frac{\partial\phi}{\partial t}(t,s) \times \frac{\partial\phi}{\partial s}(t,s). \tag{14.146}$$

Définition 14.43.
Pour intégrer ω sur S nous faisons

$$\int_S \omega = \int_D \phi^*\omega \tag{14.147}$$

où $\phi^*\omega$ est de la forme $F(t,s)dt \wedge ds$.

Montrons ce que cela fait. Soient u, v des vecteurs de D et calculons

$$(\phi^*\omega)(u,v) = \omega\big(d\phi(u), d\phi(v)\big) \tag{14.148a}$$

$$= \omega\left(u_1\frac{\partial\phi}{\partial x_1} + u_2\frac{\partial\infty}{\partial x_2}, v_1\frac{\partial\phi}{\partial x_1} + v_2\frac{\partial\infty}{\partial x_2}\right). \tag{14.148b}$$

Les termes en u_1v_1 et u_2v_2 sont nuls ; par exemple :

$$\omega\left(u_1\frac{\partial\phi}{\partial x_1}, v_1\frac{\partial\phi}{\partial x_1}\right) = u_1v_1\omega\left(\frac{\partial\infty}{\partial x_1}, \frac{\partial\phi}{\partial x_1}\right) = 0 \tag{14.149}$$

parce que ω est antisymétrique. Il nous reste donc

$$(\phi^*\omega)(u,v) = (u_1v_2 - u_2v_1)\omega\left(\frac{\partial\phi}{\partial x_1}, \frac{\partial\phi}{\partial x_2}\right) \tag{14.150a}$$

$$= \omega\left(\frac{\partial\phi}{\partial x_1}, \frac{\partial\phi}{\partial x_2}\right)(dt \wedge ds)(u,v). \tag{14.150b}$$

Cette dernière ligne est bien de la forme $\phi^*\omega = F(t,s)dt \wedge ds$.

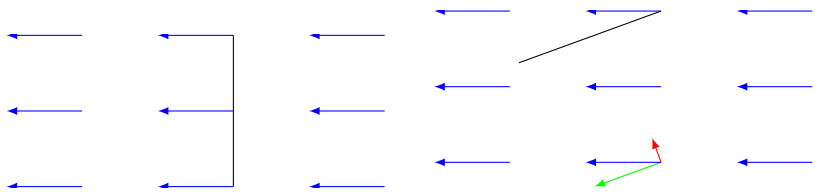
14.6 Flux d'un champ de vecteurs à travers une surface

Nous voulons construire un moulin à eau. Comment placer les pales pour maximiser le travail de la pression de l'eau ? On n'a pas attendu l'invention du calcul intégral pour répondre à cette question. Trois paramètres rentrent en ligne de compte :

(1) plus il y a d'eau, plus ça pousse ;
(2) plus la surface de la palle est grande, plus on va utiliser d'eau ;
(3) plus la palle est perpendiculaire au courant, plus on va en profiter.

Nous voyons sur la figure 14.3 que lorsque la palle du moulin est inclinée, non seulement elle prend moins d'eau sur elle, mais qu'en plus elle la prend avec un moins bon angle : une partie de la force ne sert pas à la faire tourner.

L'idée du flux d'un champ de vecteurs à travers une surface est de savoir quelle est la quantité «utile» de vecteurs qui traverse la surface. Ce sera simplement l'intégrale sur la surface de la composante du champ de vecteurs normale à la surface. Il reste deux problèmes à régler : le premier est de savoir quel est le vecteur normal à la surface, et le second est de savoir comment «sélectionner» la composante normale d'un champ de vecteurs F.



(a) L'eau pousse bien perpendiculairement à la palle du moulin.

(b) Lorsque l'eau ne pousse pas perpendiculairement, une partie de la force est perdue. Ici la partie normale (en rouge) est très petite.

FIGURE 14.3 – La partie rouge de la force est perdue si l'eau ne pousse pas perpendiculairement. De plus lorsque la palle est inclinée, elle prend moins d'eau sur elle.

Le problème de trouver un vecteur normal est résolu par le produit vectoriel des vecteurs tangents. Si la surface est donnée par $\varphi\colon D \subset \mathbb{R}^2 \to \mathbb{R}^3$, les vecteurs tangents sont $T_u = \partial_u\varphi(u,v)$ et $T_v = \partial_v\varphi(u,v)$. Le normal de norme 1 est donné par :

$$n(u,v) = \frac{T_u \times T_v}{\|T_u \times T_v\|}. \tag{14.151}$$

Si p est un point de la surface $\varphi(D)$, la composante de $F(p)$ qui est normale à la surface au point p est donnée par le produit scalaire

$$F(p)_\perp = F(p) \cdot n(p). \tag{14.152}$$

C'est ce nombre là que nous intégrons sur la surface.

Définition 14.44.
*Le **flux du champ de vecteurs** à travers la surface $S = \varphi(D)$ est*

$$\int F \cdot dS = \int F \cdot n \, dudv. \tag{14.153}$$

Une petite simplification se produit lorsqu'on veut calculer effectivement cette intégrale. En effet $F \cdot n$ est, en soi, une fonction sur S. Pour l'intégrer, il faut donc la multiplier par $\|T_u \times T_v\|$ (c'est la définition de l'intégrale d'une fonction sur une surface). Donc, étant donné que $n = (T_u \times T_v)/\|T_u \times T_v\|$, nous avons

$$\int F \cdot dS = \iint_D F\big(\varphi(u,v)\big) \cdot (T_u \times T_v) \, dudv \tag{14.154}$$

où $T_u = \frac{\partial\phi}{\partial u}$ et $T_v = \frac{\partial\varphi}{\partial v}$.

Exemple 14.45
Soit le champ de vecteurs

$$F = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix}. \tag{14.155}$$

Calculons son flux au travers de la sphère de rayon R.

Nous choisissons de paramétrer la sphère en coordonnées sphériques avec $\phi(\theta,\varphi)$. Nous pouvons reprendre le résultat (14.124) :

$$T_\theta \times T_\varphi = R^2 \sin(\theta). \tag{14.156}$$

Nous savons aussi que

$$F\big(\phi(\theta,\varphi)\big) = 2e_r. \tag{14.157}$$

L'intégrale à calculer est donc

$$I = \int_0^{\pi} d\theta \int_0^{2\pi} d\varphi \, 2e_r \cdot \big(R^2 \sin(\theta) e_r\big). \tag{14.158}$$

Vu que le produit scalaire $e_r \cdot e_r$ vaut 1, nous calculons

$$I = 4\pi R^2 \int_0^{\pi} \sin(\theta) d\theta = 8\pi R^2. \tag{14.159}$$

△

Exemple 14.46
Calculons le flux du champ de force de gravitation d'une masse au travers de la sphère de centre R centrée autour la masse. À un coefficient constant près, le champ vaut

$$G(r, \theta, \varphi) = \frac{1}{r^2} e_r. \tag{14.160}$$

Sur la sphère de rayon R, nous avons

$$G\big(\phi(\theta, \varphi)\big) = \frac{1}{R^2} e_r. \tag{14.161}$$

L'intégrale est donc

$$\int_0^{\pi} d\theta \int_0^{2\pi} \frac{1}{R^2} e_r \cdot \big(R^2 \sin(\theta) e_r\big) d\varphi = 8\pi. \tag{14.162}$$

Ce flux ne dépend pas de R. △

Exemple 14.47
Soit S le disque de rayon 5 placé horizontalement à la hauteur 12. Calculer le flux du champ de vecteurs

$$F(x, y, z) = xe_x + ye_y + ze_z. \tag{14.163}$$

Les équations de la surface sont $z = 12$, $x^2 + y^2 \leqslant 25$. Nous prenons le paramétrage en coordonnées cylindriques :

$$\varphi(r, \theta) = \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \\ 12 \end{pmatrix}. \tag{14.164}$$

Les vecteurs tangents sont

$$T_r = \frac{\partial \varphi}{\partial r} = \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix} \quad T_\theta = \frac{\partial \varphi}{\partial \theta} = \begin{pmatrix} -r\sin\theta \\ r\cos\theta \\ 0 \end{pmatrix}. \tag{14.165}$$

Le vecteur normal est alors

$$T_r \times T_\theta = re_z. \tag{14.166}$$

Sur la surface, le champ de vecteurs s'écrit

$$F\big(\varphi(r, \theta)\big) = r\cos(\theta)e_x + r\sin(\theta)e_y + 12e_z. \tag{14.167}$$

Par conséquent

$$F \cdot (T_r \times T_\theta) = 12r. \tag{14.168}$$

L'intégrale à calculer est

$$\begin{aligned} \int_0^5 dr \int_0^{2\pi} 12r \, d\theta &= 12 \cdot 2\pi \int_0^5 r \, dr \\ &= \frac{25}{2} 24\pi \\ &= 300\pi. \end{aligned} \tag{14.169}$$

△

14.7 Divergence, Green, Stokes

Le théorème de Stokes (et ses variations) peut se voir comme une généralisation du théorème fondamental du calcul différentiel et intégral qui stipule que

$$\int_a^b f'(x)dx = f(b) - f(a)$$

c'est-à-dire qui relie l'intégrale de f' sur $I = [a, b]$ aux valeurs de f sur le bord $\partial I = \{a, b\}$. Le signe $-$ qui apparait vient de l'orientation ; celle-ci requiert de la prudence dans l'utilisation des théorèmes.

Voici, pour votre culture générale, un énoncé général :

Théorème 14.48.
Si M est une variété orientable de dimension n avec un bord noté ∂M, alors pour toute forme différentielle ω de degré $n-1$ on a

$$\int_M d\omega = \int_{\partial M} \omega.$$

où $d\omega$ désigne la différentielle extérieure de ω.

Nous allons maintenant voir quelques cas particuliers.

Une des nombreuses formes du théorème de Stokes (théorème 14.48) est que si la forme différentielle ω est exacte alors son intégrale est facile.

Théorème 14.49.
Si γ est une chemin de classe C^1 dans un ouvert Ω et si ω est la forme différentielle exacte $\omega = df$, alors

$$\int_\gamma df = f\big(\gamma(1)\big) - f\big(\gamma(0)\big). \tag{14.170}$$

Cela est également une extension du théorème fondamental du calcul différentiel.

14.7.1 Théorème de la divergence

Si nous considérons une surface dans $\mathbb{R}^n$ et un champ de vecteurs, il est bon de se demander quelle « quantité de vecteurs » traverse la surface. Soit D, un ouvert borné de $\mathbb{R}^n$ telle que ∂D soit une variété de dimension $n-1$, et G, un champ de vecteurs défini sur $\bar{D}$. Afin de compter combien de G traverse ∂D, il faudra faire en sorte de ne considérer que la composante de G normale à ∂D : pas question d'intégrer par exemple la norme de G sur ∂D.

Comme nous le savons, la composante du vecteur v dans la direction w est le produit scalaire $v \cdot 1_w$ où 1_w est le vecteur de norme 1 dans la direction w. Nous allons donc introduire le concept de vecteur normal extérieur.

Définition 14.50.
*Soit $x \in \partial D$ et $\nu \in \mathbb{R}^n$, nous disons que ν est un **vecteur normal extérieur** de ∂D si*

(1) $\langle \nu, v \rangle = 0$ pour tout vecteur tangent v à ∂D au point x. Pour rappel, ∂D étant une variété de dimension $n-1$, il y a $n-1$ tels vecteurs v linéairement indépendants.

(2) Il existe un $\delta > 0$ tel que $\forall t \in]0, \delta[$, nous avons $c + t\nu \notin \bar{D}$ et $x - t\nu \in D$.

Nous pouvons maintenant définir le concept de flux. Soit $D \subset \mathbb{R}^n$ tel que ∂D soit une variété de dimension $n-1$ qui admette un vecteur normal extérieur $\nu(x)$ en chaque point. Soit aussi

$G\colon \bar{D} \to \mathbb{R}^n$, un champ de vecteur de classe C^1. Le **flux** de G au travers de ∂D est le nombre

$$\int_{\partial D} \langle G(x), \nu(x) \rangle d\sigma(x). \tag{14.171}$$

Cette intégrale est en général très compliquée à calculer parce qu'il faut trouver le champ de vecteur normal, puis une paramétrisation de la surface ∂D et ensuite appliquer la méthode décrite au point 14.4.

Heureusement, il y a un théorème qui nous permet de calculer plus facilement : sans devoir trouver de vecteurs normaux.

Il n'est pas plus contraignant d'énoncer ce théorème dans le cadre d'une hypersurface de $\mathbb{R}^n$, ce que nous faisons donc :

Théorème 14.51 (Formule de la divergence)**.**
Soit D un ouvert borné de $\mathbb{R}^n$ dont le bord est « assez régulier par morceaux », c'est-à-dire :

$$\partial D = A_1 \cup \dots A_p \cup N \tag{14.172}$$

où

(1) $A_1, \dots, A_p, N$ sont deux à deux disjoints,

(2) pour tout $i \leqslant p$, A_i est un ouvert relatif d'une certaine variété M_i de dimension $(n-1)$

(3) $\bar{A}_i \subset M_i$

(4) N est un compact contenu dans une réunion finie de variétés de dimensions $(n-2)$.

Supposons également qu'en chaque point de $A_1 \cup \dots \cup A_p$ il existe un vecteur normal extérieur ν.

Si G est un champ de vecteurs de classe C^1 sur $\bar{D}$ alors

$$\int_D \nabla \cdot G = \sum_{i=1}^{p} \int_{A_i} \langle G, \nu \rangle. \tag{14.173}$$

L'intégrale du membre de gauche est l'intégrale sur un ouvert d'une simple fonction.

14.7.2 Formule de Green

Pour rappel, une chemin $\gamma\colon [0,1] \to \mathbb{R}^n$ est **régulier** s'il est C^1 et si $\gamma(t) \neq 0$ pour tout r. Le chemin est de **Jordan** si $\gamma(1) = \gamma(0)$ et si $\gamma\colon [a, b[\to \mathbb{R}^n$ est injective.

La formule de Green est un cas particulier du théorème de la divergence dans le cas $n = 2$, légèrement reformulé :

Théorème 14.52.
Soit $D \subset \mathbb{R}^2$ ouvert borné tel que son bord est est la réunion finie d'un certain nombre de chemins de classe C^1 de Jordan réguliers. Supposons qu'en chaque point de son bord, D possède un vecteur normal unitaire extérieur ν. Soient P et Q deux fonctions réelles de classe C^1 sur $\bar{D}$. Alors

$$\iint_D (\partial_x Q - \partial_y P) dx\, dy = \oint_{\partial D} P dx + Q dy \tag{14.174}$$

où chaque chemin γ formant le bord de D est orienté de sorte que $T\nu = \frac{\dot{\gamma}}{\|\dot{\gamma}\|}$ où T représente la rotation d'angle $+\frac{\pi}{2}$.

Justifions le fait que cela soit un cas particulier de la formule de Stokes du théorème 14.48. Nous considérons la forme différentielle

$$\omega = P dx + Q dy, \tag{14.175}$$

et sa différentielle

$$d\omega = \sum_i d\omega_i \wedge dx_i \tag{14.176a}$$
$$= \left(\frac{\partial P}{\partial x}dx + \frac{\partial P}{\partial y}dy\right) \wedge dx + \left(\frac{\partial Q}{\partial x}dx + \frac{\partial Q}{\partial y}dy\right) \wedge dy \tag{14.176b}$$
$$= \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx \wedge dy. \tag{14.176c}$$

Intégrons cette forme $d\omega$ sur le domaine ouvert D que nous paramétrons de façon triviale par

$$\begin{aligned} \varphi\colon D &\to \mathbb{R}^2 \\ (u,v) &\mapsto (u,v). \end{aligned} \tag{14.177}$$

Ce que nous avons est

$$\iint_D d\omega = \iint_D d\omega_{(u,v)} \left(\frac{\partial \varphi}{\partial u}, \frac{\partial \varphi}{\partial v}\right) dudv \tag{14.178}$$

Nous avons aussi $T_u = \frac{\partial \varphi}{\partial u} = \begin{pmatrix}1\\0\end{pmatrix}$ et $T_v = \frac{\partial \varphi}{\partial v} = \begin{pmatrix}0\\1\end{pmatrix}$ et donc

$$(dx \wedge dy)(T_u, T_v) = dx(T_u)dy(T_v) - dx(T_v)dy(T_u) = 1 - 0 = 1. \tag{14.179}$$

L'intégrale (14.178) se développe donc en

$$\iint_D d\omega = \iint_D \left(\frac{\partial Q}{\partial x}(u,v) - \frac{\partial P}{\partial y}(u,v)\right)(dx \wedge dy)(T_u, T_v)dudv = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dudv. \tag{14.180}$$

Par conséquent la formule de Stokes nous donne la formule (14.174).

La formule de Green nous permet de calculer l'aire de la surface délimitée par une courbe fermée en termes de l'intégrale d'une forme bien choisie le long du contour. Pour cela nous prenons la forme

$$\omega = -\frac{y}{2}dx + \frac{x}{2}dy, \tag{14.181}$$

de telle sorte que $\partial_x Q - \partial_y Q = 1$ et que

$$\iint_D d\omega = \iint_D ddudv = S, \tag{14.182}$$

et au final l'aire est donnée par

$$S = \int_{\partial D} \left(-\frac{y}{2}dx + \frac{x}{2}dy\right). \tag{14.183}$$

Lorsque le bord de D est paramétré par

$$\begin{aligned} \gamma\colon [a,b] &\to \mathbb{R}^2 \\ u &\mapsto \begin{pmatrix}x(u)\\y(u)\end{pmatrix}, \end{aligned} \tag{14.184}$$

nous avons

$$(Pdx + Qdy)\gamma'(u) = Px' + Qy', \tag{14.185}$$

et alors

$$\int_{\partial D} Pdx + Qdy = \int_a^b P\big(x(u), y(u)\big)x'(u) + Q\big(x(u), y(u)\big)y'(u)du. \tag{14.186}$$

En ce qui concerne l'aire de la surface, nous prenons les P et Q de la forme 14.181 :

$$S = \frac{1}{2}\int_a^b \Big(- y(u)x'(u) + x(u)y'(u)\Big)du. \tag{14.187}$$

14.7.3 Formule de Stokes

La formule de Stokes est la version classique, qui permet d'exprimer la circulation d'un champ de vecteur le long d'une courbe de $\mathbb{R}^3$ comme le flux de son rotationnel à travers n'importe quel surface dont le bord est la courbe. La version présentée ici suppose que la surface peut se paramétrer en un seul morceau :

Théorème 14.53 (Formule de Stokes)**.**
Soit $F : W \subset \mathbb{R}^2 \to \mathbb{R}^3$ une paramétrisation (carte) d'une surface dans $\mathbb{R}^3$, supposée de classe C^2. Soit D un ouvert de $\mathbb{R}^2$ vérifiant les hypothèses de la formule de Green, et tel que $\bar{D} \subset W$. Soit G un champ de vecteurs de classe C^1 défini sur $F(\bar{D})$, et soit N le champ normal unitaire donné par la paramétrisation

$$N = \frac{\frac{\partial F}{\partial u} \wedge \frac{\partial F}{\partial v}}{\left\| \frac{\partial F}{\partial u} \wedge \frac{\partial F}{\partial v} \right\|} \tag{14.188}$$

alors

$$\iint_{F(D)} \langle \nabla \times G, N \rangle \, d\sigma_F = \int_{F(\partial D)} G \tag{14.189}$$

où les chemins formant le bord ∂D sont orientés comme dans le théorème de Green.

Notons, juste pour avoir une bonne nouvelle de temps en temps, que

$$d\sigma_F = \left\| \frac{\partial F}{\partial u} \times \frac{\partial F}{\partial v} \right\| du dv, \tag{14.190}$$

mais cette norme apparaît exactement au dénominateur de N. Il ne faut donc pas la calculer parce qu'elle se simplifie.

Sous forme un peu plus physicienne [3], la formule (14.189) s'écrit

$$\int_{F(D)} \langle \nabla \times G, N(x) \rangle \, d\sigma_F(x) = \int_{F(\gamma)} \langle G, T \rangle \, ds \tag{14.191}$$

où T est le vecteur unitaire tangent à $F(\gamma)$.

14.7.3.1 Quelle est la bonne orientation ?

Le signe du vecteur normal N dépend du choix de l'ordre des coordonnées dans la carte. Supposons que je veuille paramétrer la surface $x^2 + y^2 = 1$, $z = 1$. Nous prenons naturellement comme carte le cercle C de rayon 1 dans $\mathbb{R}^2$ et la carte

$$F(r, \theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ 1 \end{pmatrix}. \tag{14.192}$$

Mais nous aurions aussi pu mettre les coordonnées r et θ dans l'autre ordre :

$$\tilde{F}(\theta, r) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ 1 \end{pmatrix}. \tag{14.193}$$

Les vecteurs normaux ne sont pas les même : la carte F donnera $\partial_r F \times \partial_\theta F$, tandis que l'autre donnera $\partial_\theta \tilde{F} \times \partial_r \tilde{F}$. Le signe change !

Il faut savoir laquelle choisir. Le cercle $C \subset \mathbb{R}^2$ a une orientation donnée par le théorème de Green. Nous choisissons l'ordre des coordonnées pour que 1_θ et 1_r soient dans la même orientation que les vecteurs ν et T tels que donnés par le théorème de Green, et tels que dessinés sur la figure 14.4.

Plus généralement, nous choisissons l'ordre des coordonnées u et v pour que la base $(1_u, 1_v)$ ait la même orientation que (ν, T) où T a le sens convenu dans le théorème de Green.

3. et surtout plus explicite.

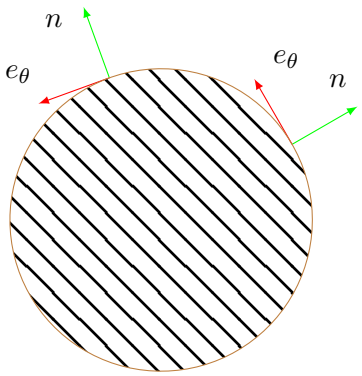


FIGURE 14.4 – L'orientation sur le cercle. Si nous les prenons dans l'ordre, les vecteurs $(1_r, 1_\theta)$ ont la même orientation que celle donnée par les vecteurs (ν, T) donnés par la convention de Green.

14.8 Résumé des intégrales vues

Nous sommes maintenant capables de revoir tous les types d'intégrales vues jusqu'ici de façon très cohérentes. Nous commencerons par les intégrales de fonctions et nous ferons ensuite les intégrales de champs de vecteurs.

14.8.1 L'intégrale d'une fonction sur les réels

Si $f\colon [a,b] \subset \mathbb{R} \to \mathbb{R}$ est une fonction usuelle, sont intégrale est

$$\int_a^b f(x)dx = F(b) - F(a) \tag{14.194}$$

où F est une primitive de f.

14.8.2 Intégrale d'une fonction sur un chemin

Si f est une fonction sur $\mathbb{R}^3$ et si $\sigma\colon [a,b] \to \mathbb{R}^3$ est un chemin dans $\mathbb{R}^3$, l'intégrale de f sur σ est, par définition,

$$\int f\, d\sigma = \int_a^b f\big(\sigma(t)\big)\|\sigma'(t)\|dt. \tag{14.195}$$

14.8.3 Intégrale d'une fonction sur une surface

Nous devons paramétrer la surface S par une application $\varphi\colon D \subset \mathbb{R}^2 \to \mathbb{R}^3$. À partir d'une telle paramétrisation, nous construisons un élément de surface en prenant le produit vectoriel des deux vecteurs tangents :

$$dS = \frac{\partial\varphi}{\partial u} \times \frac{\partial\varphi}{\partial v} dudv. \tag{14.196}$$

L'intégrale est

$$\int f\, dS = \iint_D f\big(\varphi(u,v)\big) \left\| \frac{\partial\varphi}{\partial u} \times \frac{\partial\varphi}{\partial v} \right\| dudv. \tag{14.197}$$

Il ne faut pas rajouter de jacobien : la norme du produit vectoriel *est* le jacobien.

Remarque 14.54.
La formule (14.197) est autant valable pour des surfaces dans $\mathbb{R}^2$ que dans $\mathbb{R}^3$. Si nous considérons une surface dans $\mathbb{R}^2$, nous la voyons dans $\mathbb{R}^3$ en ajoutant un zéro comme troisième composante.

Exemple 14.55
Les coordonnées polaires sont données par

$$\varphi(r,\theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ 0 \end{pmatrix}. \tag{14.198}$$

Les vecteurs tangents à cette paramétrisation sont

$$T_r = \frac{\partial\varphi}{\partial r} = \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix}, \quad T_\theta = \frac{\partial\varphi}{\partial\theta} = \begin{pmatrix} -r\sin\theta \\ r\cos\theta \\ 0 \end{pmatrix}. \tag{14.199}$$

Le vecteur normal est

$$\frac{\partial\varphi}{\partial r} \times \frac{\partial\varphi}{\partial\theta} = \begin{vmatrix} e_x & e_y & e_z \\ \cos\theta & \sin\theta & 0 \\ -r\sin\theta & r\cos\theta & 0 \end{vmatrix} = re_z. \tag{14.200}$$

Nous trouvons donc que l'élément de surface est la norme de re_z, c'est à dire r, le jacobien connu. △

14.8.4 Intégrale d'une fonction sur un volume

Si V est un volume dans $\mathbb{R}^3$, nous effectuons la même procédure : nous trouvons une paramétrisation, et nous formons un élément de volume avec les vecteurs tangents de la paramétrisation. Nous avons donc un volume déterminé par l'application

$$\varphi\colon D \subset \mathbb{R}^3 \to \mathbb{R}^3, \tag{14.201}$$

et ses trois vecteurs tangents

$$\begin{aligned} T_u &= \frac{\partial\varphi}{\partial u} \\ T_v &= \frac{\partial\varphi}{\partial v} \\ T_w &= \frac{\partial\varphi}{\partial w}. \end{aligned} \tag{14.202}$$

Comment former un volume avec trois vecteurs ? Réponse : le produit mixte. L'intégrale de f sur V sera

$$\int f\,dV = \iiint_D f\big(\varphi(u,v,w)\big) \left\| \frac{\partial\varphi}{\partial u} \cdot \left(\frac{\partial\varphi}{\partial v} \times \frac{\partial\varphi}{\partial w} \right) \right\| du dv. \tag{14.203}$$

Encore une fois, le produit mixte *est* le jacobien. Prenons les coordonnées sphériques :

$$\begin{aligned} x(r,\theta,\varphi) &= r\sin(\theta)\cos(\varphi) \\ y(r,\theta,\varphi) &= r\sin(\theta)\sin(\varphi) \\ z(r,\theta,\varphi) &= r\cos(\theta) \end{aligned} \tag{14.204}$$

Les trois vecteurs tangents seront

$$T_r = \begin{pmatrix} \frac{\partial x}{\partial r} \\ \frac{\partial y}{\partial r} \\ \frac{\partial z}{\partial r} \end{pmatrix} = \begin{pmatrix} \sin(\theta)\cos(\varphi) \\ \sin(\theta)\sin(\varphi) \\ \cos(\theta) \end{pmatrix} \tag{14.205a}$$

$$T_\theta = \begin{pmatrix} \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial \theta} \end{pmatrix} = \begin{pmatrix} r\cos(\theta)\cos(\varphi) \\ r\cos(\theta)\sin(\varphi) \\ -r\sin(\theta) \end{pmatrix} \tag{14.205b}$$

$$T_\varphi = \begin{pmatrix} \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial \varphi} \\ \frac{\partial z}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} -r\sin(\theta)\sin(\varphi) \\ r\sin(\theta)\cos(\varphi) \\ 0 \end{pmatrix} \tag{14.205c}$$

Nous avons vu que le produit mixte revient à mettre toutes les composantes dans une matrice. Ici nous avons donc

$$\frac{\partial\phi}{\partial r}\cdot\left(\frac{\partial\phi}{\partial\theta}\times\frac{\partial\phi}{\partial\varphi}\right)=\begin{vmatrix}\frac{\partial x}{\partial r}&\frac{\partial y}{\partial r}&\frac{\partial z}{\partial r}\\\frac{\partial x}{\partial\theta}&\frac{\partial y}{\partial\theta}&\frac{\partial z}{\partial\theta}\\\frac{\partial x}{\partial\varphi}&\frac{\partial y}{\partial\varphi}&\frac{\partial z}{\partial\varphi}\end{vmatrix} \tag{14.206}$$

Cela est précisément le jacobien dont nous parlions plus haut.

14.8.5 Conclusion pour les fonctions

Lorsque nous intégrons une fonction sur un chemin, une surface ou un volume, la technique est toujours la même :

(1) Trouver une paramétrisation à une, deux ou trois variables.

(2) Dériver la paramétrisation par rapport à ses variables.

(3) Construire un élément de longueur, surface ou volume à partir des vecteurs que l'on a. Cela se fait en prenant la norme, le produit vectoriel ou le produit mixte.

14.8.6 Circulation d'un champ de vecteurs

Pour les champs de vecteurs, nous faisons la même chose, mais au lieu de *multiplier* par l'élément de longueur ou de surface, nous prenons le produit scalaire. Si nous considérons la courbe paramétrée $\sigma\colon[a,b]\to\mathbb{R}^3$ et le champ de vecteurs F, nous avons donc

$$\int_\sigma F=\int F\cdot d\sigma=\int_a^b F\big(\sigma(t)\big)\cdot\sigma'(t)dt. \tag{14.207}$$

14.8.7 Flux d'un champ de vecteurs

Si la surface $S\subset\mathbb{R}^3$ est paramétrée par

$$\begin{aligned}\phi\colon D\subset\mathbb{R}^2&\to\mathbb{R}^3\\(u,v)&\mapsto\phi(u,v),\end{aligned} \tag{14.208}$$

et si F est un champ de vecteurs, alors on a

$$\int_S F=\int_S F\cdot dS=\iint_D F\big(\phi(u,v)\big)\cdot\left(\frac{\partial\phi}{\partial u}\times\frac{\partial\phi}{\partial v}\right)dudv. \tag{14.209}$$

14.8.8 Conclusion pour les champs de vecteurs

La circulation et le flux ne représentent pas tout à fait la même chose. En effet pour la circulation, nous sélectionnons la composante *tangente* à la courbe, c'est à dire la partie du vecteurs qui «circule» le long de la courbe. Une force perpendiculaire au mouvement ne travaille pas.

La situation est exactement le contraire pour le flux. Étant donné que le vecteur

$$\frac{\partial\phi}{\partial u}\times\frac{\partial\phi}{\partial v} \tag{14.210}$$

est normal à la surface, le fait de prendre le produit scalaire du champ de vecteurs avec lui sélectionne la composante *normale* à la surface, c'est à dire la partie du vecteur qui traverse la surface.

14.8.9 Attention pour les surfaces fermées !

Si nous considérons une surface fermée, il faut faire attention à choisir une *orientation*. Les vecteurs normaux doivent soit tous pointer vers l'intérieur soit tous vers l'extérieur. En effet, en tant que vecteur normal, nous avons choisi de prendre

$$T_u \times T_v. \tag{14.211}$$

Mais le vecteur $T_v \times T_u$ est tout aussi normal ! Il n'y a pas a priori de façon standard pour choisir l'un ou l'autre. Il faut juste être cohérent : il faut que si on divise la surface en plusieurs morceaux, tous les vecteurs pointent dans le même sens.

Notez que si vous faites un choix et que votre voisin fait le choix inverse, vous obtiendrez des réponses qui diffèrent d'un signe. Sans plus de précisions [4], les deux réponses sont correctes.

Un exemple de ce problème est donné dans l'exemple 14.56.

Exemple 14.56

Calculer le flux du champ de vecteurs

$$F(x, y, z) = e_x \tag{14.212}$$

au travers du cylindre de rayon R et de hauteur h autour de l'axe z.

Même question si le cylindre est autour de l'axe x.

Remarque : ces cylindres sont considérés *avec* leur «couvercles».

Une paramétrisation du cylindre autour de l'axe z est

$$\phi(\theta, z) = \begin{pmatrix} R\cos\theta \\ R\sin\theta \\ z \end{pmatrix}. \tag{14.213}$$

Les vecteurs tangents sont

$$T_\theta = \begin{pmatrix} -R\sin\theta \\ R\cos\theta \\ 0 \end{pmatrix}, T_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{14.214}$$

Le vecteur normal est donc

$$T_\theta \times T_z = R\cos(\theta)e_x + R\sin(\theta)e_y. \tag{14.215}$$

C'est un vecteur dirigé vers l'extérieur.

Le champ de vecteurs considéré est constant : $F(\theta, z) = e_x$. Nous avons donc

$$F(\theta, z) \cdot (T_\theta \times T_z) = R\cos(\theta) \tag{14.216}$$

et le flux vaut

$$\Phi = \int_0^{2\pi} d\theta \int_0^h R\cos(\theta) dz = 0. \tag{14.217}$$

En ce qui concerne les couvercles haut au bas, ils sont paramétrés par

$$\phi_1(r, \theta) = \begin{pmatrix} R\cos(\theta) \\ R\sin(\theta) \\ h \end{pmatrix}, \phi_2(r, \theta) = \begin{pmatrix} R\cos(\theta) \\ R\sin(\theta) \\ 0 \end{pmatrix}. \tag{14.218}$$

Les vecteurs normaux correspondants sont dans la direction de e_z, de façon que le produit scalaire avec $F(r, \theta)$ soit nul. Le flux total est donc nul.

4. Il faudrait définir ce qu'est une surface *orientable* et choisir une orientation.

Regardons maintenant le cylindre le long de l'axe x. Une paramétrisation est

$$\phi(\theta, x) = \begin{pmatrix} x \\ R\cos(\theta) \\ R\sin(\theta) \end{pmatrix}, \tag{14.219}$$

et le vecteurs tangents sont

$$T_\theta = \begin{pmatrix} 0 \\ -R\sin\theta \\ R\cos\theta \end{pmatrix}, T_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{14.220}$$

Le vecteur normal est alors donné par

$$T_\theta \times T_x = R\cos(\theta)e_y + R\sin(\theta)e_z. \tag{14.221}$$

Nous avons par conséquent $F(\theta, x) \cdot (T_\theta \times T_x) = 0$. Pas de flux par le côté du cylindre.

Regardons les «couvercles». Le premier est donné par la paramétrisation

$$\phi_1(r, \theta) = \begin{pmatrix} 0 \\ r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}. \tag{14.222}$$

Le vecteur normal serait $T_r \times T_\theta = re_x$, et le flux

$$\Phi = \int_0^{2\pi} d\theta \int_0^R r\, dr = \pi R^2. \tag{14.223}$$

Le second couvercle est donné par

$$\phi_2(r, \theta) = \begin{pmatrix} h \\ r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}. \tag{14.224}$$

Le vecteur normal est encore re_x, et le flux est à nouveau πR^2.

Le flux total serait donc $2\pi R^2$.

Cela n'est pas possible parce que tous les vecteurs qui «rentrent» d'un côté doivent «sortir» de l'autre côté. L'erreur est le le premier vecteur normal est un vecteur qui pointe vers l'intérieur du cylindre, tandis que le second pointe vers l'extérieur. Si nous choisissons, par convention, de prendre uniquement les vecteurs extérieurs, il faut changer le vecteur normal du premier couvercle en $-re_x$. Le premier flux vaudra donc

$$-\pi R^2, \tag{14.225}$$

de telle sorte que le flux total sera nul. △

14.9 Intégrales convergeant uniformément

14.9.1 Définition et propriété

Définition 14.57.
Soit (Ω, μ) un espace mesuré. Nous disons que l'intégrale

$$\int_\Omega f(x, \omega)d\mu(\omega) \tag{14.226}$$

***converge uniformément** en x si pour tout $\epsilon > 0$, il existe un compact K_ϵ tel que pour tout compact K tel que $K_\epsilon \subset K$ nous avons*

$$\left| \int_{\Omega \setminus K} f(x, \omega)d\mu(\omega) \right| \leqslant \epsilon. \tag{14.227}$$

Le point important est que le choix de K_ϵ ne dépend pas de x.

Lemme 14.58.
Soit

$$F(x) = \int_\Omega f(x,\omega)d\mu(\omega), \tag{14.228}$$

une intégrale uniformément convergente. Pour chaque $k \in \mathbb{N}$ nous considérons un compact K_k tel que

$$\left| \int_{\Omega\setminus K_k} f(x,\omega)d\mu(\omega) \right| \leqslant \frac{1}{k}. \tag{14.229}$$

Alors la suite de fonctions F_k définie par

$$F_k(x) = \int_{K_k} f(x,\omega)d\mu(\omega) \tag{14.230}$$

converge uniformément vers F.

Démonstration. Nous avons

$$\begin{aligned} \big|F_k(x) - F(x)\big| &= \left| \int_{K_k} f(x,\omega)d\mu(\omega) - \int_\Omega f(x,\omega)d\mu(\omega) \right| && (14.231a)\\ &= | \int_{\Omega\setminus K_k} f(x,\omega)d\mu(\omega)| && (14.231b)\\ &\leqslant \frac{1}{k}. && (14.231c) \end{aligned}$$

□

14.9.2 Critères de convergence uniforme

Afin de tester l'uniforme convergence d'une intégrale, nous avons le **critère de Weierstrass** :

Théorème 14.59.
Soit $f(x,t)\colon [\alpha,\beta]\times[a,\infty[\to\mathbb{R}$, une fonction dont la restriction à toute demi-droite $x = cst$ est mesurable. Si $|f(x,t)| < \varphi(t)$ et $\int_a^\infty \varphi(t)dt$ existe, alors l'intégrale

$$\int_0^\infty f(x,t)dt \tag{14.232}$$

est uniformément convergente.

Le théorème suivant est le **critère d'Abel** :

Théorème 14.60.
Supposons que $f(x,t) = \varphi(x,t)\psi(x,t)$ où φ et ψ sont bornée et intégrables en t au sens de Riemann sur tout compact $[a,b]$, $b \geqslant a$. Supposons que :

(1) $|\int_a^T \varphi(x,t)dt| \leqslant M$ où M est indépendant de T et de x,

(2) $\psi(x,t) \geqslant 0$,

(3) pour tout $x \in [\alpha,\beta]$, $\psi(x,t)$ est une fonction décroissante de t,

(4) les fonctions $x \mapsto \psi(x,t)$ convergent uniformément vers 0 lorsque $t \to \infty$.

Alors l'intégrale

$$\int_a^\infty f(x,t)dt \tag{14.233}$$

est uniformément convergente.

Remarque 14.61.
Étant donné que la fonction sinus est bornée, il est tentant de l'utiliser comme φ dans le critère d'Abel (théorème 14.60). Hélas,

$$\int_0^T \sin(xt) = -\frac{1}{x}\big(\cos(xT) - \cos(x)\big), \tag{14.234}$$

qui n'est pas bornée pour tout x ! Poser $\varphi(x,t) = \sin(xt)$ *ne fonctionne pas* pour assurer la convergence uniforme sur un intervalle qui contient des x arbitrairement proches de 0. Le critère d'Abel avec $\varphi(x,t) = \sin(xt)$ ne permet que de conclure à l'uniforme convergence *sur tout compact* ne contenant pas 0. Cela est toutefois souvent suffisant pour étudier la continuité ou la dérivabilité en se servant du fameux coup du compact.

14.10 Fonctions définies par une intégrale

Soit (Ω, μ) un espace mesuré. Nous nous demandons dans quel cas l'intégrale

$$F(x) = \int_\Omega f(x,\omega)d\omega \tag{14.235}$$

définit une fonction F continue, dérivable ou autre.

Dans la suite nous allons considérer des fonctions f à valeurs réelles. Quitte à passer aux composantes, nous pouvons considérer des fonctions à valeurs vectorielles. Par contre le fait que x soit dans $\mathbb{R}$ ou dans $\mathbb{R}^n$ n'est pas spécialement une chose facile à traiter.

14.10.1 Continuité sous l'intégrale

Nous allons présenter deux théorèmes donnant la continuité de F.

(1) Si f est majorée par une fonction ne dépendant pas de x, nous avons le théorème 14.62,

(2) si l'intégrale est uniformément convergente, nous avons le théorème 14.63.

Théorème 14.62.
Soit (Ω, μ) est un espace mesuré, soit $x_0 \in \mathbb{R}^m$ et $f\colon U \times \Omega \to \mathbb{R}$ où U est ouvert dans $\mathbb{R}^m$. Nous supposons que

(1) La fonction $f(x,.)$ est dans $L^1(\Omega,\mu)$ pour tout $x \in \mathbb{R}^m$.

(2) La fonction $f(.,\omega)$ est continue en x_0 pour tout $\omega \in \Omega$.

(3) Il existe une fonction $G \in L^1(\Omega)$ telle que

$$|f(x,\omega)| \leqslant G(\omega) \tag{14.236}$$

pour tout $x \in U$.

Alors la fonction

$$\begin{aligned} F\colon U &\to \mathbb{R} \\ x &\mapsto \int_\Omega f(x,\omega)d\mu(\omega) \end{aligned} \tag{14.237}$$

est continue en x_0.

Démonstration. Soit (x_n) une suite convergente vers x_0. Nous considérons la suite de fonctions $f_n\colon \Omega \to \mathbb{R}$ définies par

$$f_n(\phi) = f(x_n,\omega). \tag{14.238}$$

sur qui nous pouvons utiliser le théorème de la convergence dominée (théorème 13.134) pour obtenir

$$\lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} \int_\Omega f(x_n,\omega)d\mu(\omega) \tag{14.239a}$$
$$= \int_\Omega \lim_{n\to\infty} f(x_n,\omega)d\mu(\omega) \tag{14.239b}$$
$$= \int_\Omega f(x,\omega)d\mu(\omega) \tag{14.239c}$$
$$= F(x). \tag{14.239d}$$

Nous avons utilisé la continuité de $f(.,\omega)$. □

Si nous avons un peu de compatibilité entre la topologie et la mesure, alors nous pouvons utiliser l'uniforme convergence d'une intégrale pour obtenir la continuité d'une fonction définie par une intégrale.

Théorème 14.63.
Soit (Ω,μ) un espace topologique mesuré tel que tout compact est de mesure finie. Soit une fonction $f\colon \mathbb{R}\times\Omega\to\mathbb{R}$ telle que

(1) Pour chaque $x\in\mathbb{R}$, la fonction $f(x,.)$ est $L^1(\Omega,\mu)$.

(2) Pour chaque $\omega\in\Omega$, la fonction $f(.,\omega)$ est continue en x_0.

(3) L'intégrale

$$F(x) = \int_\Omega f(x,\omega)d\mu(\omega) \tag{14.240}$$

est uniformément convergente[5].

Alors la fonction F est continue en x_0.

Démonstration. Nous reprenons les notations du lemme 14.58. Les fonctions

$$F_k(x) = \int_{K_k} f(x,\omega)d\mu(\omega) \tag{14.241}$$

existent parce que les fonctions $f(x,.)$ sont dans $L^1(\Omega)$. Montrons que les fonctions F_k sont continues. Soit une suite $x_k\to x_0$ nous avons

$$\lim_{n\to\infty} F_k(x_n) = \lim_{n\to\infty} \int_{K_k} f(x_n,\omega)d\mu(\omega). \tag{14.242}$$

Nous pouvons inverser la limite et l'intégrale en utilisant le théorème de la convergence dominée. Pour cela, la fonction $f(x_n,\omega)$ étant continue sur le compact K_k, elle y est majorée par une constante. Le fait que les compacts soient de mesure finie (hypothèse) implique que les constantes soient intégrales sur K_k. Le théorème de la convergence dominée implique alors que

$$\lim_{n\to\infty} F_k(x_n) = \int_{K_k} \lim_{n\to\infty} f(x_n,\omega)d\mu(\omega) = \int_{K_k} f(x_0,\omega)d\mu(\omega) = F_k(x_0). \tag{14.243}$$

Nous avons utilisé le fait que $f(.,\omega)$ était continue en x_0.

Le lemme 14.58 nous indique alors que la convergence $F_k\to F$ est uniforme. Les fonctions F_k étant continues, la fonction F est continue. □

Pour finir, citons ce résultat concernant les fonctions réelles.

Théorème 14.64.
Nous considérons $F(x)=\int_a^\infty f(x,t)dt$. Si f est continue sur $[\alpha,\beta]\times[a,\alpha[$ et l'intégrale converge uniformément, alors $F(x)$ est continue.

5. Définition 14.57.

14.10.2 Le coup du compact

Nous avons vu des fonctions définies par toute une série de processus de limite (suites, séries, intégrales). Une des questions centrales est de savoir si la fonction limite est continue, dérivable, intégrale, etc. étant donné que les fonctions sont continues.

Pour cela, nous inventons le concept de *convergence uniforme*. Si la limite (série, intégrale) est uniforme, alors la fonction limite sera continue. Il arrive qu'une limite ne soit pas uniforme sur un intervalle ouvert $]0,1]$, et que nous voulions quand même prouver la continuité sur cet intervalle. C'est à cela que sert la notion de convergence uniforme *sur tout compact*. En effet, la notion de continuité est une notion locale : savoir ce qu'il se passe dans un petit voisinage autour de x est suffisant pour savoir la continuité en x (idem pour sa dérivée).

Si nous avons uniforme convergence sur tout compact de $]0,1]$, mais pas uniforme convergence sur cet intervalle, la limite sera quand même continue sur $]0,1]$. En effet, si $x \in]0,1]$, il existe un ouvert autour de x contenu dans un compact contenu dans $]0,1]$. L'uniforme convergence sur ce compact suffit à prouver la continuité en x.

Déduire la continuité sur un ouvert à partir de l'uniforme convergence sur tout compact de l'ouvert est appelé faire le **coup du compact**.

14.10.3 Dérivabilité sous l'intégrale

Nous traitons à présent de la dérivabilité de la fonction F définie comme intégrale de f.

Théorème 14.65 (Dérivation sous le signe intégral[179])**.**
Soit (Ω,μ) un espace mesuré et une fonction $f\colon \mathbb{R}\times\Omega \to \mathbb{R}$ dont nous voulons étudier la dérivabilité en $a \in \mathbb{R}$. Nous supposons qu'il existe $\delta > 0$, A mesurable de mesure nulle dans Ω tels que

(1) $f(x,.)$ soit dans $L^1(\Omega)$.

(2) L'application $x \mapsto f(x,\omega)$ est dérivable pour tout $x \in B(a,\delta)$ et pour tout $\omega \in \complement A$.

(3) Il existe une fonction G intégrable sur Ω telle que

$$\left|\frac{\partial f}{\partial x}(x,\omega)\right| \leqslant G(\omega) \tag{14.244}$$

pour tout $x \in B(a,\delta)$ et pour tout $\omega \in \complement A$.

Alors la fonction

$$F(x) = \int_\Omega f(x,\omega)d\mu(\omega) \tag{14.245}$$

est dérivable en a et nous pouvons permuter la dérivée et l'intégrale :

$$F'(a) = \int_\Omega \frac{\partial f}{\partial x}(a,\omega)d\mu(\omega). \tag{14.246}$$

Démonstration. Soit une suite (x_n) dans $B(a,\delta)$ telle que $x_n \neq a$ et $x_n \to a$. Si la limite

$$\lim_{n\to\infty} \frac{F(a)-F(x_n)}{a-x_n} \tag{14.247}$$

existe et ne dépend pas de la suite choisie, alors la fonction F est dérivable en a et sa dérivée vaut cette limite. Par linéarité de l'intégrale, nous devons étudier la limite

$$\lim_{n\to\infty} \int_\Omega \frac{f(a,\omega)-f(x_n,\omega)}{a-x_n}d\omega, \tag{14.248}$$

montrer qu'elle existe, ne dépend pas de la suite choisie et vaut $\int_\Omega \partial_x f(a,\omega)d\omega$. Nous sommes donc dans un problème d'inversion de limite et de dérivée pour lequel nous allons utiliser le théorème de la convergence dominée de Lebesgue. D'abord nous posons

$$g_n(\omega) = \frac{f(x_n,\omega)-f(a,\omega)}{x_n-a}. \tag{14.249}$$

Cela est une suite de fonctions dans $L^1(\Omega)$ parce qu'à la fois a et x_n sont dans $B(a,\delta)$. De plus nous avons

$$\lim_{n\to\infty} g_n(\omega) = \frac{\partial f}{\partial x}(a,\omega) \tag{14.250}$$

parce que nous savons que f est dérivable en a pour tout $\omega \in \complement A$. En ce qui concerne la majoration de g_n, nous utilisons le théorème des accroissements finis (théorème 12.129) sur le numérateur de (14.249). Pour tout n et pour tout $\omega \in \complement A$, il existe un $\theta_{n,\omega}$ dans $]a,x_n[$ tel que

$$f(x_n,\omega) - f(a,\omega) = \frac{\partial f}{\partial x}(\theta_{n,\omega},\omega)(x_n - a), \tag{14.251}$$

donc

$$|g_n(\omega)| = \left|\frac{\partial f}{\partial x}(\theta_{n,\omega},\omega)\right| \leqslant G(\omega). \tag{14.252}$$

La dernière inégalité provient des hypothèses. Le théorème de la convergence dominée de Lebesgue (théorème 13.134) nous permet alors de calculer la limite (14.248) :

$$\lim_{n\to\infty}\int_\Omega g_n(\omega)d\omega = \int_\Omega \lim_{n\to\infty} g_n(\omega)d\omega = \int_\Omega \frac{\partial f}{\partial x}(a,\omega)d\omega. \tag{14.253}$$

Notons que l'existence de la dernière intégrale fait partie du théorème de la convergence dominée.

Nous avons donc prouvé que la limite de gauche existait et ne dépendant pas de la suite choisie. Donc F est dérivable en a et la dérivée vaut cette limite :

$$F'(a) = \int_\Omega \frac{\partial f}{\partial x}(a,\omega)d\mu(\omega). \tag{14.254}$$

□

Théorème 14.66.
Supposons f continue et sa dérivée partielle $\frac{\partial f}{\partial x}$ continue sur $[\alpha,\beta]\times[a,\alpha[$. Supposons que $F(x) = \int_a^\infty f(x,t)dt$ converge et que $\int_a^\infty \frac{\partial f}{\partial x}dt$ converge uniformément. Alors F est C^1 sur $[\alpha,\beta]$ et

$$\frac{dF}{dx} = \int_a^\infty \frac{\partial f}{\partial x}dt. \tag{14.255}$$

En ce qui concerne les fonctions dans $\mathbb{R}^n$, il y a les propositions 14.72 et 14.73 qui parlent de différentiabilité sous l'intégrale.

14.10.4 Absolue continuité

Définition 14.67.
*Une fonction $F\colon \mathbb{R}\to\mathbb{R}$ est **absolument continue** sur $[a,b]$ s'il existe une fonction f sur $[a,b]$ telle que*

$$F(x) = \int_a^x f(t)dt \tag{14.256}$$

pour tout $x \in [a,b]$.

Théorème 14.68.
Soit A un ouvert de $\mathbb{R}$ et Ω, un espace mesuré. Soit une fonction $f\colon A\times\Omega\to\mathbb{R}$ et

$$F(x) = \int_\Omega f(x,\omega)d\omega. \tag{14.257}$$

Nous supposons les points suivants.

1. *La fonction f est mesurable en tant que fonction $A\times\Omega\to\mathbb{R}$. Pour chaque $x\in A$, la fonction $f(x,\cdot)$ est intégrable sur Ω.*
2. *Pour presque tout $\omega\in\Omega$, la fonction $f(x,\omega)$ est une fonction absolument continue de x.*

(3) La fonction $\frac{\partial f}{\partial x}$ est localement intégrable, c'est à dire que pour tout $[a,b] \subset A$,

$$\int_a^b \int_\Omega \left| \frac{\partial f}{\partial x}(x,\omega) \right| d\omega\, dx < \infty. \tag{14.258}$$

Alors la fonction F est absolument continue et pour presque tout $x \in A$, la dérivée est donné par

$$\frac{d}{dx} \int_\Omega f(x,\omega) d\omega = \int_\Omega \frac{\partial f}{\partial x}(x,\omega) d\omega. \tag{14.259}$$

La proposition suivante sera utilisée entre autres pour montrer que sous l'hypothèse d'une densité continue, la loi exponentielle est sans mémoire, proposition 34.98.

Proposition 14.69.
Soit $f(x,t)$ une fonction continue sur $[\alpha,\beta] \times [a,b]$, telle que $\frac{\partial f}{\partial x}$ existe et soit continue sur $]\alpha,\beta[\times[a,b]$. Soient $\varphi(x)$ et $\psi(x)$, des fonctions continues de $[\alpha,\beta]$ dans $\mathbb{R}$ et admettant une dérivée continue sur $]\alpha,\beta[$. Alors la fonction

$$F(x) = \int_{\varphi(x)}^{\psi(x)} f(x,t) dt \tag{14.260}$$

admet une dérivée continue sur $]\alpha,\beta[$ et

$$\frac{dF}{dx} = \int_{\varphi(x)}^{\psi(x)} \frac{\partial f}{\partial x}(x,t) dt + f\big(x,\psi(x)\big) \cdot \frac{d\psi}{dx} - f\big(x,\varphi(x)\big) \cdot \frac{d\varphi}{dx}. \tag{14.261}$$

L'exemple qui suit devrait pouvoir être rendu rigoureux en utilisant des distributions correctement.

Exemple 14.70
Si g est une fonction continue, la fonction suivante est une primitive de g :

$$\int_0^x f(t) dt = \int_0^\infty f(t) \mathbb{1}_{t<x}(t) dt. \tag{14.262}$$

Nous nous proposons de justifier *de façon un peu heuristique* le fait que ce soit bien une primitive de g en considérant la fonction

$$f(t,x) = g(t) \mathbb{1}_{t<x}(t). \tag{14.263}$$

Nous posons

$$F(x) = \int_0^\infty f(x,t) dt, \tag{14.264}$$

et nous calculons F' en permutant la dérivée et l'intégrale [6]. D'abord,

$$f(t,x) = \begin{cases} g(t) & \text{si } t \in [0,x] \\ 0 & \text{sinon.} \end{cases} \tag{14.265}$$

La dérivée de f par rapport à x est donnée par la distribution

$$\frac{\partial f}{\partial x}(t_0,x_0) = g(t_0)\delta(t_0 - x_0). \tag{14.266}$$

Donc

$$F'(x_0) = \int_0^\infty \frac{\partial f}{\partial x}(t,x_0) dt = \int_0^\infty g(t)\delta(t-x_0) = g(x_0), \tag{14.267}$$

comme attendu. △

Cet exemple est rendu rigoureux par la proposition suivante.

6. Ceci n'est pas rigoureux : il faudrait avoir un théorème à propos de distributions qui permet de le faire.

Proposition 14.71.
Si $f \in L^1(\mathbb{R})$, alors la fonction

$$F(x) = \int_{-\infty}^{x} f(t)dt \tag{14.268}$$

est presque partout dérivable et pour les points où elle l'est nous avons $F'(x) = f(x)$.

14.10.5 Différentiabilité sous l'intégrale

Le théorème suivant est restrictif sur l'ensemble d'intégration (qui doit être compact), mais accepte des fonctions de plusieurs variables, ce qui est un premier pas vers la différentiabilité.

Proposition 14.72 (Dérivation sous l'intégrale)**.**
Supposons $A \subset \mathbb{R}^m$ ouvert et $B \subset \mathbb{R}^n$ compact. Nous considérons une fonction $f\colon A \times B \to \mathbb{R}$. Si pour un $i \in \{i, \dots, n\}$, la dérivée partielle $\frac{\partial f}{\partial x_i}$ existe dans $A \times B$ et est continue, alors la fonction

$$F(x) = \int_B f(x,t)dt \tag{14.269}$$

admet une dérivée partielle dans la direction x_i sur A. Cette dérivée partielle y est continue et

$$\frac{\partial F}{\partial x_i}(a) = \int_B \frac{\partial f}{\partial x_i}(a,t)dt, \tag{14.270}$$

pour tout a dans l'ouvert A.

Démonstration. Nous procédons en plusieurs étapes.

<u>F **est dérivable**</u> Nous voulons prouver que $\frac{\partial F}{\partial x_i}(a,t)$ existe. Pour cela nous posons

$$g_l(t) = \frac{f(a_1, \dots, a_i + \epsilon_l, \dots, a_n, t) - f(a_1, \dots, a_i, \dots, a_n, t)}{\epsilon_l} \tag{14.271}$$

où ϵ_l est une suite de nombres tendant vers zéro. La fonction f est dérivable dans la direction x_i si et seulement si $\lim_{l\to\infty} g_l(t)$ existe et ne dépend pas du choix de la suite. À ce moment, la valeur de la dérivée partielle sera cette limite. Dans notre cas, nous savons que f admet une dérivée partielle dans la direction x_i et donc nous avons

$$\frac{\partial f}{\partial x_i}(a,t) = \lim_{l\to\infty} g_l(t). \tag{14.272}$$

De la même façon pour F nous avons

$$\frac{\partial F}{\partial x_i} = \lim_{l\to\infty} \int_B g_l(t)dt. \tag{14.273}$$

Sous-entendu : si la limite de droite ne dépend pas de la suite choisie, alors $\frac{\partial F}{\partial x_i}$ existe et vaut cette limite.

Vu la continuité de f, le seul point à vérifier pour le théorème de la convergence dominée de Lebesgue est l'existence d'une fonction intégrable de t majorant g_l. Pour cela le théorème de accroissements finis (théorème 12.129) appliqué à la fonction $\epsilon \mapsto f(a_n, \dots, a_i + \epsilon, \dots, a_n)$ nous dit que

$$f(a_1, \dots, a_i + \epsilon_l, \dots, a_n, t) - f(a_1, \dots, a_i, \dots, a_n, t) = \epsilon_l \frac{\partial f}{\partial x_i}(a_1, \dots, \theta, \dots, a_n, t) \tag{14.274}$$

pour un certain $\theta \in B(a_i, \epsilon_l)$. Notons que ce θ dépend de t mais pas de l. Vu que $\partial_i f$ est continue par rapport à ses deux variables, si K est un voisinage compact autour de a, il existe $M > 0$ tel que

$$\left| \frac{\partial f}{\partial x_i}(x,t) \right| < M \tag{14.275}$$

pour tout $x \in K$ et tout $t \in B$. La valeur de $\frac{\partial f}{\partial x_i}(a_1, \dots, \theta, \dots, a_n, t)$ est donc bien majorée par rapport à θ et par rapport à t en même temps par une constante qui n'a pas de mal à être intégrée sur le compact B.

Le théorème de la convergence dominée (théorème 13.134) s'applique donc bien et nous avons

$$\lim_{l\to\infty} \int_B g_l(t)dt = \int_B \lim_{l\to\infty} g_l(t) = \int_B \frac{\partial f}{\partial x_i}(a,t)dt. \tag{14.276}$$

Le membre de droite ne dépendant pas de la suite ϵ_l choisie, le membre de gauche est bien la dérivée de F par rapport à x_i et nous avons

$$\frac{\partial F}{\partial x_i}(a) = \int_B \frac{\partial f}{\partial x_i}(a,t)dt. \tag{14.277}$$

Cela prouve la première partie de la proposition.

La dérivée est continue Soit K un voisinage compact autour de a et U' un ouvert tel que $a \in U' \subset K$. Nous avons encore la majoration (14.275) sur U' et donc le théorème de continuité sous l'intégrale 14.62 nous indique que la fonction

$$\begin{aligned} U' &\to \mathbb{R} \\ x &\mapsto \int_B \frac{\partial f}{\partial x_i}(x,t)dt \end{aligned} \tag{14.278}$$

est continue en a.

□

Une conséquence de la proposition 14.72 est que si elle fonctionne pour tous les i, alors F est différentiable et même de classe C^1, et la différentielle de F s'obtient comme intégrale de la différentielle de f.

Proposition 14.73.
Supposons $A \subset \mathbb{R}^m$ ouvert et $B \subset \mathbb{R}^n$ compact. Si pour tout $i \in \{i, \dots, n\}$, la dérivée partielle $\frac{\partial f}{\partial x_i}$ existe dans $A \times B$ et est continue, alors F est de classe C^1 et

$$(dF)_a = \int_B (df_t)_a dt \tag{14.279}$$

où $f_t(x) = f(x,t)$.

Démonstration. En vertu de la proposition 14.72, toutes les dérivées partielles de F sont continues. Cela implique que F est de classe C^1 par la proposition 12.238 et que la différentielle s'écrive en terme des dérivées partielles avec la formule usuelle. Nous avons alors

$$(dF)_a(u) = \sum_k \frac{\partial F}{\partial x_k}(a)u_k \tag{14.280a}$$

$$= \int_B \sum_k \frac{\partial f}{\partial x_k}(a,t)dt \tag{14.280b}$$

$$= \int_B \sum_k \frac{\partial f_t}{\partial x_k}(a)u_k dt \tag{14.280c}$$

$$= \int_B (df_t)_a(u)dt. \tag{14.280d}$$

Cela est la formule annoncée. □

Un autre théorème tourne autour du pot, et me semble inutile.

Théorème 14.74.
Soit (Ω, μ) un espace mesuré, une fonction $f\colon \mathbb{R}^n \times \Omega \to \mathbb{R}$ et $a \in \mathbb{R}^n$. Nous considérons la fonction

$$F(x) = \int_\Omega f(x, \omega) d\mu(\omega). \tag{14.281}$$

Pour chaque $k = 1, \dots, n$ nous supposons avoir

$$\frac{\partial F}{\partial x_k}(a) = F'_{|_k}(a) = \int_\Omega \frac{\partial f_{|_k}}{\partial t}(a_k, \omega) d\mu(\omega) \tag{14.282}$$

où $F_{|_k}(t) = F(a_1, \dots, t, \dots, a_n)$ et $f_{|_k}$ est définie de façon similaire.
Nous supposons de plus que les fonctions $\partial_{x_k} F$ sont continues.
Alors F est de classe C^1 et sa différentielle est donnée par

$$df_a = \int_\Omega (df_\omega)_a d\omega \tag{14.283}$$

où f_ω est définie par $f_\omega(x) = f(x, \omega)$.

Démonstration. Étant donné que les dérivées partielles de F en a existent et sont continues, la proposition 12.238 dit que F est différentiable et que

$$dF_a(u) = \sum_{k=1}^n \frac{\partial F}{\partial x_k}(a) u_k. \tag{14.284}$$

La linéarité de l'intégrale et les hypothèses nous donnent alors

$$\begin{aligned} df_a(u) &= \sum_{k=1}^n \frac{\partial F}{\partial x_k}(a) u_k && \text{(14.285a)} \\ &= \int_\Omega \sum_k \frac{\partial f_{|_k}}{\partial t}(a_k; \omega) u_k d\mu(\omega) && \text{(14.285b)} \\ &= \int_\Omega \sum_k \frac{\partial f}{\partial x_k}(a; \omega) u_k d\mu(\omega) && \text{(14.285c)} \\ &= \int_\Omega (df_\omega)_a(u) d\mu(\omega), && \text{(14.285d)} \end{aligned}$$

et donc $df_a = \int_\Omega (df_\omega)_a d\mu(\omega)$. □

Notons qu'en passant aux composantes, ce théorème fonctionne tout aussi bien pour des fonctions à valeurs dans un espace vectoriel normé de dimension finie plutôt que dans $\mathbb{R}$.

Lemme 14.75 (Hadamard[187])**.**
Soit une fonction $f\colon \mathbb{R}^n \to \mathbb{R}$ de classe C^p avec $p \geqslant 1$. Pour tout $a \in \mathbb{R}^n$ il existe des fonctions $g_1, \dots, g_n$ de classe C^{p-1} telles que

$$f(x) = f(a) + \sum_{i=1}^n (x_i - a_i) g_i(x). \tag{14.286}$$

Démonstration. Vu que f est de classe C^1, le théorème fondamental de l'analyse 13.205 fonctionne et

$$f(x) - f(a) = \int_0^1 \frac{d}{dt}\Big[f\big(a + t(x-a)\big)\Big] dt = \int_0^1 \sum_{i=1}^n \frac{\partial f}{\partial x_i}\big(a + t(x-a)\big)(x_i - a_i). \tag{14.287}$$

Plus de détails : la fonction $t \mapsto \frac{d}{dt}\Big[f\big(a + t(x-a)\big)\Big]$ possède comme primitive la fonction $F(t) = f\big(a + t(x-a)\big)$.

Nous posons

$$g_i(x) = \int_0^1 \frac{\partial f}{\partial x_i}\big(a + t(x-a)\big)dt \tag{14.288}$$

Le fait que l'intégrale existe est simplement le fait qu'il s'agit d'une fonction continue sur un compact et donc majorée par une constante. Pour voir que g_i est de classe C^{p-1} nous pouvons calculer $\frac{\partial g_i}{\partial x_k}$ en permutant dérivée et intégrale par la proposition 14.72 :

$$\frac{\partial g_i}{\partial x_k}(x) = \int_0^1 \frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_i}\big(a + t(x-a)\big)\right)dt = \int_0^1 t\frac{\partial^2 f}{\partial x_k \partial x_i}\big(a + t(x-a)\big). \tag{14.289}$$

Nous pouvons ainsi permuter $p-1$ dérivées tout en gardant une fonction continue dans l'intégrale. Le théorème 14.62 nous donne alors une fonction continue. Ainsi toutes les fonctions

$$\frac{\partial^{p-1} g_i}{\partial x_{i_1} \dots \partial x_{i_{p-1}}} \tag{14.290}$$

sont continues et g_i est de classe C^{p-1} par la proposition 12.262.

En repartant de (14.287) nous avons alors bien ce qui était annoncé :

$$f(x) = f(a) + \sum_{i=1}^{n} g_i(x)(x_i - a_i). \tag{14.291}$$

□

Corollaire 14.76.
Soit $\phi \in \mathscr{D}(\mathbb{R})$ tel que $\phi^{(k)}(x_0) = 0$ pour tout $k \leqslant n$. Alors il existe une fonction $\psi \in \mathscr{D}(\mathbb{R})$ telle que

$$\phi(x) = (x - x_0)^{n+1}\psi(x) \tag{14.292}$$

pour tout $x \in \mathbb{R}$.

Démonstration. En utilisant le lemme de Hadamard 14.75 avec $a = x_0$, $n = 1$ et $f(x_0) = 0$, nous avons une fonction g_1 à support compact telle que

$$\phi(x) = \phi(x_0) + (x - x_0)g_1(x). \tag{14.293}$$

Alors $\phi'(x) = g_1(x) + (x - x_0)g_1'(x)$, ce qui donne immédiatement $g_1(x_0) = 0$ et donc une fonction g_2 telle que $g_1(x) = (x - x_0)g_2(x)$. En injectant dans (14.293) nous avons

$$\phi(x) = (x - x_0)^2 g_2(x). \tag{14.294}$$

Il suffit de continuer ainsi tant que les dérivées de ϕ s'annulent. □

14.11 Formes différentielles exactes et fermées

Nous avons déjà parlé de formes différentielles et de leurs intégrales sur un chemin dans la section 12.18.

Définition 14.77.
*La forme différentielle ω est **exacte** s'il existe une fonction f telle que $\omega = df$; elle est dite **fermée** si $d\omega = 0$.*

Dire que la forme différentielle $\omega = f dx + g dy$ est fermée, c'est dire que

$$\frac{\partial g}{\partial x} = \frac{\partial f}{\partial y}. \tag{14.295}$$

Il est naturel de se demander si toutes les formes différentielles sont des différentielles de fonctions. Une réponse complète est délicate à établir, mais a d'innombrables conséquences en physique,

notamment en ce qui concerne l'existence d'un potentiel vecteur pour le champ magnétique dans les équations de Maxwell.

Le fait qu'une forme exacte soit fermée est relativement facile à établir ; c'est la proposition suivante. La question plus délicate est la réciproque : sous quelles conditions une forme fermée est-elle exacte ?

Proposition 14.78.
Si ω est une 1-forme exacte de classe C^1, alors ω est fermée.

Démonstration. Le fait que ω soit exacte implique l'existence d'une fonction f telle que $\omega = df$, c'est à dire

$$\omega_x = \sum_i a_i(x)dx_i = \sum_i \frac{\partial f}{\partial x_i}(x)dx_i, \tag{14.296}$$

c'est à dire que $a_i(x) = \frac{\partial f}{\partial x_i}(x)$. L'hypothèse que ω est C^1 implique que f est C^2, et donc que nous pouvons inverser l'ordre de dérivation pour les dérivées secondes $\partial^2_{ij} f = \partial^2_{ji} f$. Nous pouvons donc faire le calcul suivant :

$$\frac{\partial a_i}{\partial x_j} = \frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i} = \frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j} = \frac{\partial a_j}{\partial x_i}, \tag{14.297}$$

ce qu'il fallait démontrer. □

La réciproque est vraie sur un ouvert simplement connexe.

Théorème 14.79.
Supposons que $D \subset \mathbb{R}^n$ soit un ouvert simplement connexe. Alors toute forme différentielle de degré 1 et de classe C^1 sur D qui est fermée est exacte.

Nous allons prouver ce théorème dans un cas un peu moins général : celui d'un domaine étoilé de $\mathbb{R}^2$ plutôt que simplement connexe de $\mathbb{R}^n$.

Théorème 14.80.
Soit $D \subset \mathbb{R}^2$, une ouvert étoilé, et ω, une 1-forme fermée de classe C^1. Alors ω est exacte.

Démonstration. Soit $D \subset \mathbb{R}^2$, un ouvert étoilé par rapport à l'origine. Soient $f\colon D \to \mathbb{R}$, $g\colon D \to \mathbb{R}$, des fonctions de classe C^1 telles que

$$\frac{\partial f}{\partial y} = \frac{\partial g}{\partial x} \tag{14.298}$$

sur D, et

$$F(x,y) = \int_0^1 \big[f(tx,ty)x + g(tx,ty)y\big]dt \tag{14.299}$$

pour tout $(x,y) \in D$.

Étant donné que nous ne définissons $F(x,y)$ que pour des $(x,y) \in D$, la fonction $t \mapsto f(tx,ty)$ est C^1 sur tout le compact $[0,1]$ et aucune divergence de l'intégrale n'est à craindre. Nous sommes donc dans le cadre de la proposition 14.72, et nous pouvons dériver sous le signe intégral.

Nous calculons, en utilisant la règle de dérivation de fonctions composées

$$\begin{aligned}\frac{\partial F}{\partial x}(x,t) &= \int_0^1 \left[f\frac{\partial f}{\partial x}(tx,ty)x + f(tx,ty) + t\frac{\partial g}{\partial x}(tx,ty)y\right]dt \\ &= \int_0^1 \left[t\Big(x\frac{\partial f}{\partial x}(tx,ty) + y\frac{\partial f}{\partial y}(tx,ty)\Big) + f(tx,ty)\right]dt\end{aligned} \tag{14.300}$$

où nous avons utilisé l'hypothèse $\partial_y f = \partial_x g$. Ce qui se trouve dans la parenthèse n'est autre que $\partial_t\big(f(tx,ty)\big)$, plus précisément, si nous posons $\mathcal{F}(x,y,t) = f(tx,ty)$, nous avons

$$\frac{\partial \mathcal{F}}{\partial t}(x,y,t) = x\frac{\partial f}{\partial x}(tx,ty) + y\frac{\partial f}{\partial y}(tx,ty). \tag{14.301}$$

En recopiant le résultat (14.300) en termes de $\mathcal{F}$, nous avons

$$\begin{aligned}\frac{\partial F}{\partial x}(x,t) &= \int_0^1 \left(t\frac{\partial \mathcal{F}}{\partial t}(x,y,t) + \mathcal{F}(x,y,t)\right)dt \\ &= \int_0^1 \partial_t\big(t\mathcal{F}(x,y,t)\big)dt \\ &= \big[f\mathcal{F}(x,y,t)\big]_0^1 \\ &= \mathcal{F}(x,y,1) \\ &= f(x,y).\end{aligned} \tag{14.302}$$

Le résultat correspondant pour $\frac{\partial F}{\partial y}(x,y) = g(x,y)$ s'obtient de la même manière. Nous avons donc obtenu que

$$\frac{\partial F}{\partial x} = f, \quad \text{et} \quad \frac{\partial F}{\partial y} = g. \tag{14.303}$$

En ayant prouvé cela, nous avons prouvé que si $\omega = fdx + gdy$ avec $\partial_y f = \partial_x g$, alors $\omega = dF$ où F est définie par (14.299). □

Démonstration alternative du théorème 14.80. Nous posons $u = tx$ et $v = ty$, ainsi que $\mathcal{F}(x,y,t) = f(u,v)$ et $\mathcal{G}(x,y,t) = g(u,v)$. Avec cette notation, nous avons $F(x,y) = \int_0^1 \big(x\mathcal{F}(x,y,t)+y\mathcal{G}(x,y,t)\big)dt$, et

$$\begin{aligned}\frac{\partial \mathcal{F}}{\partial x} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} = t\frac{\partial f}{\partial u}, \\ \frac{\partial \mathcal{G}}{\partial x} &= t\frac{\partial g}{\partial u}.\end{aligned} \tag{14.304}$$

Ainsi,

$$\begin{aligned}\frac{\partial F}{\partial x} &= \int_0^1 \left(x\frac{\partial \mathcal{F}}{\partial x} + \mathcal{F} + y\frac{\partial G}{\partial x}\right)dt \\ &= \int_0^1 \left(xt\frac{\partial f}{\partial u} + \mathcal{F} + yt\frac{\partial g}{\partial u}\right)dt \\ &= \int_0^1 \left[t\left(x\frac{\partial f}{\partial u} + y\frac{\partial f}{\partial v}\right) + \mathcal{F}\right]dt.\end{aligned} \tag{14.305}$$

où nous avons utilisé le fait que, par hypothèse, $\frac{\partial g}{\partial u} = \frac{\partial f}{\partial v}$. Nous calculons par ailleurs que

$$\frac{\partial F}{\partial t} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial t} = x\frac{\partial f}{\partial u} + y\frac{\partial f}{\partial v}. \tag{14.306}$$

Donc, nous avons

$$\frac{\partial F}{\partial x} = \int_0^1 \left(t\frac{\partial \mathcal{F}}{\partial t} + \mathcal{F}\right)dt = \int_0^1 \frac{\partial}{\partial t}(t\mathcal{F})dt. \tag{14.307}$$

Par conséquent,

$$\frac{\partial F}{\partial x} = [t\mathcal{F}]_0^1 = \mathcal{F}(x,y,1) = f(x,y). \tag{14.308}$$

Le même genre de calculs fournit $\frac{\partial F}{\partial y} = g(x,y)$. □

14.12 Théorème d'Abel angulaire

Théorème 14.81 (Abel angulaire[2])**.**
Soit $\sum_n a_n z^n$ une série entière de rayon de convergence plus grand ou égal à 1 et de somme f. Soit $\theta_0 \in [0, \frac{\pi}{2}[$. Nous posons

$$\Delta_{\theta_0} = \{z = 1 - \rho e^{i\varphi} \text{ tel que } \rho > 0, \varphi \in [\theta_0, \theta_0], |z| < 1\}. \tag{14.309}$$

Nous supposons de plus que $\sum_n a_n$ converge. Alors

$$\lim_{\substack{z\to 1\\ z\in\Delta_0}} f(z) = \sum_{k=0}^{\infty} a_k. \tag{14.310}$$

Démonstration. Le résultat de ce théorème est que l'on peut calculer la limite $z \to 1$ avec des chemins contenus dans un domaine de la forme de celui dessiné à la figure 14.5.

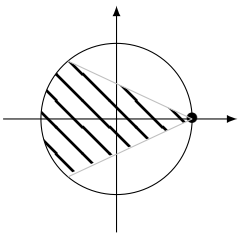


FIGURE 14.5 – La zone dans laquelle peut être le chemin qui va vers $z = 1$.

De façon très classique nous posons

$$S = \sum_{k=0}^{\infty} a_k \quad S_n = \sum_{k=0}^{n} a_k, \tag{14.311}$$

et $R_n = S - S_n$. En particulier $a_n = R_{n-1} - R_n$.

Le but du théorème est de montrer que $\sum a_n z^n$ converge vers S lorsque z converge vers 1 à l'intérieur de Δ_{θ_0}. Pour cela nous calculons pour un N donné la différence $\sum_{n=0}^{N} a_n z^n - S_N$ en triant les termes par ordre de R_n, en isolant le terme R_0 et le terme R_N :

$$\begin{aligned}
\sum_{n=0}^{N} a_n z^n - S_N &= \sum_{n=1}^{N} a_n(z^n - 1) && \text{(14.312a)}\\
&= \sum_{n=1}^{N} (R_{n-1} - R_n)(z^n - 1) && \text{(14.312b)}\\
&= R_0(z-1) + \sum_{n=1}^{N-1} R_n(z^{n+1} - 1 - z^n + 1) + R_N(z^N - 1) && \text{(14.312c)}\\
&= R_0(z-1) + \sum_{n=1}^{N-1} R_n z^n(z-1) + R_N(z^N - 1) && \text{(14.312d)}\\
&= (z-1)\sum_{n=0}^{N-1} R_n z^n + R_N(z^N - 1). && \text{(14.312e)}
\end{aligned}$$

Cela est valable pour tout N et $|z| < 1$. Nous avons donc

$$\sum_{n=0}^{N} a_n z^n - S_N = (z-1)\sum_{n=0}^{N-1} R_n z^n + R_N(z^N - 1). \tag{14.313}$$

Par hypothèse nous avons $\lim_{N\to\infty} R_N = 0$. Et de plus le membre de gauche converge parce que chacun des deux termes converge séparément. En passant à la limite nous avons pour tout $|z| < 1$:

$$f(z) - S = (z-1)\sum_{n=0}^{\infty} R_n z^n. \tag{14.314}$$

Nous voudrions étudier le comportement de la différence $f(z) - S$ lorsque z tend vers 1. Pour cela nous nous fixons $\epsilon > 0$ et $N \geqslant 1$ tel que $|R_n| < \epsilon$ dès que $n \geqslant N$. Alors pour tout $|z| < 1$ nous

avons

$$|f(z)-S| \leqslant |z-1|\left(\sum_{n=0}^{N}|R_n|\underbrace{|z^n|}_{\leqslant 1}+\sum_{n=N+1}^{\infty}\underbrace{|R_n|}_{\leqslant\epsilon}|z^n|\right) \tag{14.315a}$$

$$\leqslant |z-1|\sum_{n=0}^{N}|R_n|+\epsilon\frac{|z-1|}{1-|z|} \tag{14.315b}$$

où nous avons utilisé la somme de la série géométrique (10.128) et l'égalité $|z^n|=|z|^n$. Avant de nous particulariser à $z\in\Delta_{\theta_0}$ nous devons anticiper un problème au dénominateur en multipliant par le binôme conjugué :

$$\frac{|z-1|}{1-|z|}=\frac{|z-1|(1+|z|)}{1-|z|^2}. \tag{14.316}$$

C'est maintenant que nous nous particularisons à $z\in\Delta_{\theta_0}$ en posant $z=\rho e^{i\varphi}$ et en remarquant que $|z|^2=1-2\rho\cos(\varphi)+\rho^2$. Nous avons le calcul suivant :

$$\frac{|z-1|}{1-|z|}=\frac{\rho(1+|z|)}{2\rho\cos(\varphi)-\rho^2} \tag{14.317a}$$

$$=\frac{1+|z|}{2\cos(\varphi)-\rho} \tag{14.317b}$$

$$\leqslant\frac{2}{2\cos(\varphi)-\rho} \tag{14.317c}$$

$$\leqslant\frac{2}{2\cos(\varphi)-\cos(\theta_0)} \tag{14.317d}$$

$$\leqslant\frac{2}{2\cos(\theta_0)-\cos(\theta_0)} \tag{14.317e}$$

$$=\frac{2}{\cos(\theta_0)}. \tag{14.317f}$$

Quelque justifications.

— Vu que nous avons dans l'idée de faire $\rho\to 0$ nous supposons que $\rho<\cos(\theta_0)$.

— Nous avons $\cos(\varphi)>\cos(\theta_0)$ parce que z est dans Δ_{θ_0}.

Nous avons donc, pour tout $z\in\Delta_{\theta_0}$ que

$$|f(z)-S|\leqslant|z-1|\sum_{n=0}^{N}|R_n|+\epsilon\frac{2}{\cos(\theta_0)}. \tag{14.318}$$

Il suffit de prendre ρ assez petit pour que

$$|z-1|\sum_{n=0}^{N}|R_n|<\epsilon \tag{14.319}$$

et nous avons

$$|f(z)-S|\leqslant\epsilon\left(1+\frac{2}{\cos(\theta_0)}\right). \tag{14.320}$$

Nous avons donc bien $\lim_{\substack{z\to 1\\ z\in\Delta_0}}f(z)=S$, comme nous le voulions. □

La réciproque du théorème d'Abel angulaire est que si $f(z)=\sum_n a_nz^n$ sur $B(0,1)$ se prolonge par continuité en $z=1$ alors cette prolongation se fait par $f(1)=\sum_n a_n$. Cela est faux comme le montre l'exemple suivant.

Exemple 14.82

Nous considérons la série entière $\sum_{n=0}^{\infty}(-1)^nz^n$ qui converge [7] vers

$$f(z)=\frac{1}{1+z} \tag{14.321}$$

7. C'est la série géométrique de raison $-z$.

sur $B(0,1)$. De plus nous avons

$$\lim_{\substack{z\to 1\\ |z|<1}} \frac{1}{1+z} = \frac{1}{2}. \tag{14.322}$$

Donc la fonction converge bien vers quelque chose lorsque z tend vers 1. La fonction f se prolonge par continuité en 1. Pourtant la série es coefficients $\sum_n (-1)^n$ ne converge pas. △

Le théorème suivant donne une espèce d'inverse au théorème d'Abel angulaire. En effet il dit que si la série converge en allant vers 1 le long de l'axe réel, alors ça converge vers la somme des coefficients. Il faut cependant une hypothèse en plus sur les a_n.

Théorème 14.83 (Théorème taubérien faible[2])**.**
Soit $\sum_n a_n z^n$ une série entière de rayon de convergence 1 et de somme f. Nous supposons

(1) Il existe $S \in \mathbb{C}$ tel que $\lim_{\substack{x\to 1\\ x\in]-1,1[}} f(x) = S$.

(2) $\lim_{n\to\infty} na_n = 0$.

Alors la série $\sum_{n=0}^{\infty} a_n$ converge et vaut S.

Démonstration. Nous notons $S_n = \sum_{k=0} a_k$ et $M = \sup_{k\geqslant 1} k|a_k|$, qui est fini par hypothèse. Pour $x \in]0,1[$ et $n \geqslant 0$ nous avons

$$S_n - f(x) = \sum_{k=1}^{n} a_k - \sum_{k=1}^{n} a_k x^k - \sum_{k=n+1}^{\infty} a_k x^k = \sum_{k=1}^{n} a_k(1-x^k) - \sum_{k=n+1}^{\infty} a_k x^k. \tag{14.323}$$

Nous utilisons la série géométrique sous la forme $1 - x^k = (1-x)\sum_{i=0}^{n} x^i$ pour écrire

$$S_n - f(x) = \sum_{k=1}^{n} a_k(1-x) \underbrace{\sum_{i=0}^{k-1} x^i}_{\leqslant k} - \sum_{k=n+1}^{\infty} a_k x^k \tag{14.324a}$$

$$\leqslant \sum_{k=1}^{n} k a_k (1-x) - \sum_{k=n+1}^{\infty} a_k x^k, \tag{14.324b}$$

donc en passant à la norme

$$\big|S_n - f(x)\big| \leqslant (1-x)Mn + \sum_{k=n+1} |a_k| x^k \tag{14.325a}$$

$$\leqslant (1-x)Mn + \sum_{k=n+1}^{\infty} \underbrace{\frac{k}{n}|a_k|}_{\leqslant M/n} x^k \tag{14.325b}$$

$$\leqslant (1-x)Mn + \frac{M}{n}\sum_{k=n+1}^{\infty} x^k \tag{14.325c}$$

$$\leqslant (1-x)Mn + \frac{M}{n}\frac{1}{1-x}. \tag{14.325d}$$

Ce que nous cherchons à étudier est le comportement $x \to 1$ et montrer que $S_n \to S$, ce qui nous incite à calculer $|S_n - f(1-\frac{\epsilon}{n})|$ avec $0 < \epsilon < 1$:

$$\big|S_n - f\big(1-\frac{\epsilon}{n}\big)\big| \leqslant \epsilon M + \epsilon. \tag{14.326}$$

Nous choisissons N_1 tel que $\frac{M}{n} \leqslant \epsilon^2$ dès que $n \geqslant N_1$. En sus nous savons que

$$\lim_{\epsilon\to 0} f(1-\epsilon) = S. \tag{14.327}$$

Nous choisissons N_2 de telle sorte à avoir

$$\left|f\left(1-\frac{\epsilon}{n}\right)-S\right|<\epsilon, \tag{14.328}$$

et en prenant $n \geqslant \max(N_1, N_2)$ nous avons

$$|S_n - S| \leqslant \left|S_n - f\left(1-\frac{\epsilon}{n}\right)\right| + \left|f\left(1-\frac{\epsilon}{n}\right)-S\right| \leqslant \epsilon M + 2\epsilon. \tag{14.329}$$

Il suffit de choisit ϵ suffisamment petit (en particulier pour que ϵM soit petit) pour montrer que $|S_n - S|$ est borné par un nombre arbitrairement petit. □

14.12.1 Passage à la limite sous le signe intégral

Un autre résultat très important pour l'étude de l'intégrabilité est le théorème de la **convergence dominée de Lebesgue** :

Théorème 14.84.
Soit $E \subset \mathbb{R}^n$ un ensemble mesurable et $\{f_k\}$, une suite de fonctions intégrables sur E qui converge simplement vers une fonction $f\colon E \to \overline{\mathbb{R}}$. Supposons qu'il existe une fonction g intégrable sur E telle que pour tout k,

$$|f(x)| \leqslant g(x) \tag{14.330}$$

pour tout $x \in E$. Alors f est intégrable sur E et

$$\int_E f = \lim_{k\to\infty} \int_E f_k. \tag{14.331}$$

14.12.2 Intégrale en dimension un

Proposition 14.85 (Critère de comparaison)**.**
Soit f mesurable sur $]a,\infty[$ et bornée sur tout $]a,b]$, et supposons qu'il existe un $X_0 \geqslant a$, tel que sur $]X_0,\infty[$,

$$|f(x)| \leqslant g(x) \tag{14.332}$$

où $g(x)$ est intégrable. Alors $f(x)$ est intégrable sur $]a,\infty[$.

Corollaire 14.86 (Critère d'équivalence)**.**
Soient f et g des fonctions mesurables et positives ou nulles sur $]a,\infty[$, bornées sur tout $]a,b]$, telles que

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = L \tag{14.333}$$

existe dans $\bar{\mathbb{R}}$.

(1) Si $L \neq \infty$ et $\int_a^\infty g(x)$ existe, alors $\int_a^\infty f(x)dx$ existe,
(2) Si $L \neq 0$ et si $\int_a^\infty f(x)dx$ existe, alors $\int_a^\infty g(x)dx$ existe,

Corollaire 14.87 (Critère des fonctions test)**.**
Soit $f(x)$ une fonction mesurable et positive ou nulle sur $]a,\infty[$ et bornée pour tout $]a,b]$. Nous posons

$$L(\alpha) = \lim_{x\to\infty} x^\alpha f(x), \tag{14.334}$$

et nous supposons qu'elle existe.

(1) Si il existe $\alpha > 1$ tel que $L(\alpha) \neq \infty$, alors $\int_a^\infty f(x)dx$ existe,
(2) Si il existe $\alpha \leqslant 1$ et $L(\alpha) \neq 0$, alors $\int_a^\infty f(x)dx$ n'existe pas.

Corollaire 14.88.
Soit $f\colon]a,b] \to \mathbb{R}$ une fonction mesurable, positive ou nulle, et bornée sur $[a+\epsilon, b]$ $\forall \epsilon > 0$. Si $\lim_{x\to a}(x-a)^\alpha f(x) = L$ existe, alors

(1) Si $\alpha < 1$ et $L \neq \infty$, alors $\int_a^b f(x)dx$ existe,
(2) Si $\alpha \geqslant 1$ et $L \neq 0$, alors $\int_a^b f(x)dx$ n'existe pas.

14.12.3 Intégrales convergentes

Définition 14.89.
Soit f, une fonction mesurable sur $[a,\infty[$, bornée sur tout intervalle $[a,b]$. On dit que l'intégrale

$$\int_a^\infty f(x)dx \tag{14.335}$$

converge *si la limite*

$$\lim_{X\to\infty}\int_a^X f \tag{14.336}$$

existe et est finie.

14.12.4 La méthode de Rothstein-Trager

Mes sources pour parler d'intégration de fractions rationnelles : [188].

Théorème 14.90 (Rothstein-Trager)**.**
Soient $P,Q\in\mathbb{Q}[X]$ premiers entre eux avec $\operatorname{pgcd}(P,Q)=1$ et $\deg(P)<\deg(Q)$. Nous supposons que Q est unitaire et sans facteurs carrés. Supposons que nous puissions écrire, dans un extension $\mathbb{K}$ de $\mathbb{Q}$ la primitive de P/Q de la façon suivante :

$$\int\frac{P}{Q}=\sum_{i=1}^n c_i\ln(P_i) \tag{14.337}$$

où les c_i sont des constantes non nulles et deux à deux distinctes et où les P_i sont des polynômes unitaires non constants sans facteurs carrés et premiers deux à deux entre eux dans $\mathbb{K}[X]$.

Alors les c_i sont les racines distinctes du polynôme

$$R(Y)=\operatorname{res}_X(P-YQ',Q)\in\mathbb{K}[Y] \tag{14.338}$$

et

$$P_i=\operatorname{pgcd}(P-c_iQ',Q). \tag{14.339}$$

Démonstration. Nous posons

$$U_i=\prod_{j\neq i}P_j. \tag{14.340}$$

Question de division Ensuite nous dérivons formellement l'équation (14.337) et nous multiplions les deux côtés du résultat par $\prod_{j=1}^n P_j$:

$$P\prod_{j=1}^n P_j=Q\sum_{i=1}^n c_i\frac{P'i}{P_i}\prod_{j=1}^n P_j=Q\sum_{i=1}^n c_iP_i'U_i. \tag{14.341}$$

Une première chose que nous en tirons est que Q divise le produit $P\prod_{j=1}^n P_j$; mais P et Q étant premiers entre eux,

$$Q\mid\prod_{j=1}^n P_j \tag{14.342}$$

par le théorème de Gauss 5.44.

Une seconde chose que nous tirons de (14.341) est que P_j divise $Q\sum_{i=1}^n c_iP_i'U_i$. De cette somme, à cause du U_i qui est divisé par P_j pour tout i sauf $i=j$, le polynôme P_j divise tous les termes sauf peut-être un. Donc il les divise tous et en particulier

$$P_j\mid Qc_jP_j'U_j \tag{14.343}$$

En nous souvenant que les P_k sont premiers entre eux, P_j ne divise pas U_j. De plus P_j étant sans facteurs carrés, les polynômes P_j et P_j' sont premiers entre eux. Il ne reste que Q. Nous en déduisons que

$$P_j \mid Q \tag{14.344}$$

pour tout $1 \leqslant j \leqslant n$. Et vu que les P_i sont premiers entre eux, le fait que chacun divise Q implique que leur produit divise Q, c'est à dire

$$\prod_{j=1}^{n} P_j \mid Q. \tag{14.345}$$

Or nous avions déjà prouvé la division contraire. Du fait que les deux polynômes sont unitaires nous en déduisons qu'ils sont en réalité égaux :

$$Q = \prod_{j=1}^{n} P_j. \tag{14.346}$$

Nous pouvons simplifier les deux membres de (14.341) par cela :

$$P = \sum_{i=1}^{n} c_i P_i' U_i. \tag{14.347}$$

Encore un peu de division En dérivant (14.346) nous trouvons

$$Q' = \sum_{j=1}^{n} P_j' U_j, \tag{14.348}$$

et en écrivant P sous sa forme (14.347),

$$P - c_i Q' = \sum_{j=1}^{n} c_j P_j' U_j - \sum_{j=1}^{n} c_i P_j' U_j = \sum_{j=1}^{n} (c_j - c_i) P_j' U_j. \tag{14.349}$$

Le terme $i = j$ de la somme est nul ; en ce qui concerne les autres termes, ils sont divisés par P_i parce que $P_i \mid U_j$. Donc P_i divise tous les termes de la somme et nous avons

$$P_i \mid P - c_i Q'. \tag{14.350}$$

Un pgcd pour continuer Nous montrons à présent que $P_i = \operatorname{pgcd}(P - c_i Q', Q)$. Pour cela nous utilisons la multiplicativité du PGCD lorsque les facteurs sont premiers entre eux :

$$\operatorname{pgcd}(P - c_i Q', Q) = \operatorname{pgcd}(P - c_i Q', \prod_{j=1}^{n} P_j) = \prod_{j=1}^{n} \operatorname{pgcd}(P - c_i Q', P_j). \tag{14.351}$$

Nous remplaçons $P - c_i Q'$ par son expression (14.349) et nous écrivons un des facteurs du produit :

$$\operatorname{pgcd}(P - c_i Q', P_j) = \operatorname{pgcd}(\sum_{k=1}^{n} (c_k - c_i) P_k' U_k, P_j) \tag{14.352}$$

Le polynôme P_j divise tous les U_k sauf celui avec $k = j$. Donc le lemme 5.43 nous permet de dire

$$\operatorname{pgcd}(P - c_i Q', P_j) = \operatorname{pgcd}\big((c_j - c_i) P_j' U_j, P_j\big) = \begin{cases} 1 & \text{si } i \neq j \\ P_j & \text{si } i = j. \end{cases} \tag{14.353}$$

La seconde ligne provient du fait que nous ayons déjà montré que $P_j \mid P - c_j Q'$. En fin de compte,

$$\operatorname{pgcd}(P - c_i Q', Q) = P_i. \tag{14.354}$$

Une histoire de résultant Les nombres c_i sont tels que les polynômes $P - c_iQ'$ et Q ne sont pas premiers entre eux. Vu que les P_i sont non nuls, la proposition 9.18 nous dit que le résultant

$$\text{res}_X(P - c_iQ', Q) = 0. \tag{14.355}$$

Donc les c_i sont des racines du polynôme (en Y)

$$R(Y) = \text{res}_X(P - YQ', Q). \tag{14.356}$$

Nous n'avons pas prouvé qu'ils étaient *toutes* les racines [8].

Toutes les racines Nous allons maintenant montrer que les c_i étaient toutes les racines imaginables du polynôme (14.356) dans toutes les extensions de $\mathbb{Q}$. Soit donc c une racine de R dans une extension $\hat{\mathbb{K}}$ de $\mathbb{K}$ qui ne soit pas parmi les c_i de la formule (14.337). Étant donné que c est racine du résultat, les polynômes $P - cQ'$ et Q ont un PGCD non trivial, c'est à dire non constant. Donc

$$\text{pgcd}(P - cQ', Q) = s \in \hat{\mathbb{K}}[X] \tag{14.357}$$

est un polynôme non constant. Si T un facteur irréductible de S, alors T divise $P - cQ'$ et Q, mais $Q = \prod_{i=1}^{n} P_i$ avec les P_i premiers entre eux. Donc T ne peut diviser que l'un (et exactement un) d'entre eux [9]. Soit P_{i_0} celui qui est divisé par T. La relation (14.349) dans ce contexte donne :

$$P - cQ' = \sum_{j=1}^{n} (c_j - c)P_j'U_j \tag{14.358}$$

Le polynôme T divise tous les U_j avec $j \neq i_0$, mais comme en plus il divise $P - cQ'$, il divise aussi le dernier terme de la somme :

$$T \mid (c_{i_0} - c)P_{i_0}'U_{i_0}. \tag{14.359}$$

Le polynôme T ne divisant pas U_{i_0} et $(c_{i_0} - c)$ étant non nul, nous concluons que T divise P_{i_0}'. Mais cela n'est pas possible parce que nous avons supposé que P_{i_0} était sans facteur carré, ce qui voulait entre autres dire que P_{i_0} et P_{i_0}' n'ont pas de facteurs communs.

□

Ce théorème suggère la méthode suivante pour trouver la primitive de la fraction rationnelle P/Q (si elle vérifie les hypothèses)

(1) Écrire le résultant $R(y) = \text{res}_X(P - yQ', Q)$ et en trouver les racines $\{c_i\}_{i=1,\ldots,n}$.

(2) Calculer les $P_p = \text{pgcd}(P - c_iQ', Q)$.

(3) Écrire la réponse :

$$\int \frac{P}{Q} = \sum_{i=1}^{n} c_i \ln(P_i). \tag{14.360}$$

Notons que le polynôme $R(Y)$ est de degré $\deg(Q)$ (pour le voir, faire un peu de comptage de lignes et colonnes dans la matrice de Sylvester), donc il n'est a priori pas pire à factoriser que Q lui-même [10]. Mais il se peut que nous ayons de la chance et que R soit plus facile que Q.

À part qu'on a peut-être plus de chance avec R qu'avec Q, l'avantage de la méthode est qu'elle permet d'éviter de passer par des extensions de $\mathbb{Q}$ non nécessaires [11] .

8. De plus, nous n'avons pas de garanties que ces racines soient dans $\mathbb{Q}$, et en fait il y a des cas dans lesquels les c_i n'y sont pas.

9. On ne peut pas diviser deux trucs qui sont premiers entre eux ; c'est une question de cohérence, madame !

10. C'est de la factorisation de Q qu'on a besoin pour utiliser la méthode de décomposition en fractions simples.

11. J'imagine que pour un ordinateur, c'est plus facile d'éviter les extensions.

Exemple 14.91
Prenons la fraction rationnelle $\frac{x}{x^2-3}$. L'intégration via les fraction simples est :

$$\int \frac{x}{x^2-3}dx = \frac{1}{2}\int \frac{1}{x-\sqrt{3}} + \frac{1}{2}\int \frac{1}{x+\sqrt{3}} = \frac{1}{2}\ln(x-\sqrt{3}) + \frac{1}{2}\ln(x+\sqrt{3}) = \frac{1}{2}\ln(x^2-3). \quad (14.361)$$

Nous voyons que dans la réponse, il n'y a pas de racines. Passer par l'extension $\mathbb{Q}[\sqrt{3}]$ est par conséquent peut-être un effort inutile. Voyons comment les choses se mettent avec la méthode Rothstein-Trager.

D'abord

$$\begin{aligned} R(Y) &= \operatorname{res}_X(X - 2YX, X^2 - 3) && (14.362a)\\ &= \operatorname{res}_X\big((1-2Y)X, X^2-3\big) && (14.362b)\\ &= \det\begin{pmatrix} 1-2Y & 0 & 0 \\ 0 & 1-2Y & 0 \\ 1 & 0 & -3 \end{pmatrix} && (14.362c)\\ &= (1-2Y)\big(-3(1-2Y)\big) && (14.362d)\\ &= -3(1-2Y)^2, && (14.362e) \end{aligned}$$

dont les solutions sont faciles : il n'y a que la racine double $y = \frac{1}{2}$. La somme (14.337) sera donc réduite à un seul terme avec $c_1 = \frac{1}{2}$. Nous calculons P_1 :

$$P_1 = \operatorname{pgcd}(X - \frac{1}{2}2X, X^2 - 3) = \operatorname{pgcd}(0, X^2 - 3) = X^2 - 3, \quad (14.363)$$

et par conséquent

$$\int \frac{X}{X^2-3} = \frac{1}{2}\ln(X^2-3). \quad (14.364)$$

À aucun moment nous ne sommes sortis de $\mathbb{Q}$. △

Comme vu sur cet exemple, l'intérêt du théorème de Rothstein-Trager est de permettre, lorsqu'on a de la chance, d'en profiter, et non de nous en rendre compte à la fin en remarquant bêtement que la réponse pouvait s'écrire dans $\mathbb{Q}[X]$.

Remarque 14.92.
Afin d'utiliser cette méthode, il faut s'assurer que Q soit sans facteurs carrés. Si nous devons intégrer un $\frac{P}{Q}$ quelconque, nous devons commencer par écrire

$$Q = Q_1Q_2^2Q_3^3\dots Q_r^r, \quad (14.365)$$

et ensuite il y a moyen de ramener l'intégrale de P/Q à des intégrales de $\frac{P}{Q_1\dots Q_r}$. Cela ne demande pas de factoriser complètement Q, mais seulement de trouver ses facteurs irréductibles Q_i dans $\mathbb{Q}[X]$.

Dans l'exemple donné plus haut, $Q = X^2-3$ a des facteurs irréductibles autres que Q lui-même dans $\mathbb{R}[X]$, mais nous n'en avons pas besoin.

Voici une exemple où nous évitons de passer par les complexes.

Exemple 14.93
À calculer : $\int \frac{1}{x^3+x}$. La décomposition en fractions simples donne :

$$\frac{1}{x^3+x} = \frac{1}{x} - \frac{1/2}{x-i} - \frac{1/2}{x+i}. \quad (14.366)$$

Déjà cette décomposition passe par l'extension $\mathbb{Q}[i]$, et le calcul de la primitive de $\frac{1}{x+i}$ demande le logarithme complexe qui ne sera vu que dans la proposition 25.65 au prix d'un peu de sang. Nous verrons dans l'exemple 25.66 comment ça se passe en passant par les complexes.

En ce qui concerne la méthode de Rothstein-Trager, nous commençons par calculer le résultant (qui est tout de même un peu de calcul) :

$$P(y) = \operatorname{res}_X\big(-3yX^2 - y + 1, X^3 - X\big) \tag{14.367a}$$

$$= \det\begin{pmatrix} -3y & 0 & 1-y & 0 & 0 \\ 0 & -3y & 0 & 1-y & 0 \\ 0 & 0 & -3y & 0 & 1-y \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix} \tag{14.367b}$$

$$= -(y-1)^2(2y+1)^2 \tag{14.367c}$$

```
------------------------------------------------------------------------
| Sage Version 5.7, Release Date: 2013-02-19                           |
| Type "notebook()" for the browser-based notebook interface.          |
| Type "help()" for help.                                              |
------------------------------------------------------------------------
sage: y=var('y')
sage: R=matrix(5,5,[-3*y,0,1-y,0,0,0,-3*y,0,1-y,0,0,0,-3*y,0,1-y,1,0,1,0,0,0,1,0,1,0])
sage: R.determinant().factor()
-(y - 1)*(2*y + 1)^2
```

Les solutions sont $c_1 = 1$ et $c_2 = -\frac{1}{2}$. Nous pouvons alors calculer les P_i :

$$P_1 = \operatorname{pgcd}(-3X^2, X^3 + X) = X \tag{14.368}$$

et

$$P_2 = \operatorname{pgcd}(\frac{3}{2}X^2 + \frac{3}{2}, X^3 + X) = X^2 + 1, \tag{14.369}$$

et finalement

$$\int \frac{1}{X^3 + X} = \ln(X) - \frac{1}{2}\ln(X^2 + 1). \tag{14.370}$$

△

Notons qu'il n'y a pas de miracles : lorsque la réponse contient des racines, nous ne pouvons pas couper à passer par des extensions et factoriser un peu à la dure.

Exemple 14.94

Nous voulons calculer

$$\int \frac{1}{X^2 + 1}. \tag{14.371}$$

Nous posons donc $P = 1$ et $Q = X^2 + 1$. Le résultant à calculer est

$$P(y) = \operatorname{res}_X(-2yX + 1, X^2 + 1) = \det\begin{pmatrix} -2y & 1 & 0 \\ 0 & -2y & 1 \\ 1 & 0 & 1 \end{pmatrix} = 4y^2 + 1. \tag{14.372a}$$

Les racines de cela sont complexes et il n'y a donc pas d'échappatoires : $c_1 = \frac{i}{2}$, $c_2 = -\frac{i}{2}$. Ensuite, étant donné que $X^2 + 1 = (X + i)(X - i) = i(-iX + 1)(X - 1)$ nous avons

$$P_1 = \operatorname{pgcd}(-iX + 1, X^2 + 1) = X + i. \tag{14.373}$$

Notons que de façon naturelle, nous aurions écrit $P_1 = 1 - iX$, mais par convention nous considérons le PGCD unitaire. Cela ne change rien à la réponse parce que changer P_i en kP_i ne fait que rajouter une constante $\ln(k)$ à la primitive trouvée.

De la même façon,

$$P_2 = \text{pgcd}(1 + iX, X^2 + 1) = X - i. \tag{14.374}$$

Au final nous écrivons

$$\int \frac{1}{X^2+1} = \frac{i}{2}\ln(X+i) - \frac{i}{2}\ln(X-i). \tag{14.375}$$

△

Remarque 14.95.
Tout cela est si nous voulons absolument écrire la primitive avec des logarithmes de polynômes. Pour celui de l'exemple 14.94, nous avons trouvé

$$\int \frac{1}{X^2+1} = \frac{i}{2}\ln(X+i) - \frac{i}{2}\ln(X-i). \tag{14.376}$$

Mais

```
sage: f(x)=1/(x**2+1)
sage: f.integrate(x)
x |--> arctan(x)
```

Si nous acceptons de passer aux fonctions trigonométriques (inverses), la primitive prend un tour très différent et bien réel. Ces deux visions de l'univers sont bien entendu [12] compatibles. En effet, afin de tomber juste, nous allons prendre la primitive

$$f(x) = \frac{i}{2}\ln(ix-1) - \frac{i}{2}\ln(ix+1) \tag{14.377}$$

au lieu de (14.376). Il s'agit seulement de multiplier l'intérieur des logarithmes, ce qui ne donne qu'une constante de différence. Ensuite nous passons à la forme trigonométrique des nombres complexes : $ix - 1 = \sqrt{x^2+1}e^{i\arctan(-x)}$ et $ix + 1 = \sqrt{x^2+1}e^{i\arctan(x)}$. Avec un peu de calcul,

$$f(x) = -\frac{1}{2}\Big(\arctan(-x) - \arctan(x)\Big) = \arctan(x). \tag{14.378}$$

14.13 Rappel sur les intégrales usuelles

Soit une fonction

$$\begin{aligned} f\colon [a,b] \subset \mathbb{R} &\to \mathbb{R}^+ \\ x &\mapsto f(x). \end{aligned} \tag{14.379}$$

L'intégrale de f sur le segment $[a,b]$, notée $\int_a^b f(x)dx$ est le nombre égal à l'aire de la surface située entre le graphe de f et l'axe des x, comme indiqué à la figure 14.6.

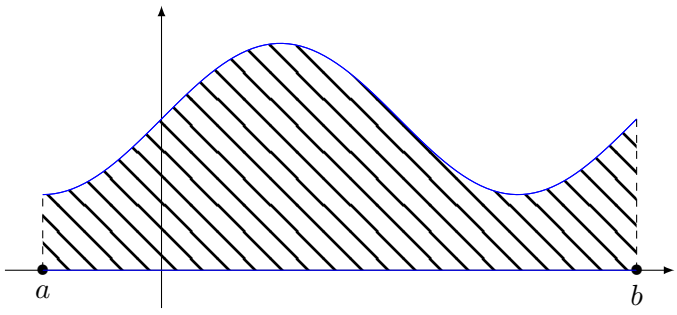


FIGURE 14.6 – L'intégrale de f entre a et b représente la surface sous la fonction.

12. Si on croit que la mathématique est cohérente.

Définition 14.96.
Si f est une fonction de une variable à valeurs réelles, une ***primitive*** *de f est une fonction F telle que $F' = f$.*

Toute fonction continue admet une primitive.

Théorème 14.97 (Théorème fondamental du caclul intégral)**.**
Si f est une fonction positive et continue, et si F est une primitive de f, alors

$$\int_a^b f(x)dx = F(b) - F(a). \tag{14.380}$$

Remarque 14.98.
Si f est une fonction continue par morceaux, l'intégrale de f se calcule comme la somme des intégrales de ses morceaux. Plus précisément si nous avons $a = x_0 < x_1 < \ldots < x_n = b$ et si f est continue sur $]x_i, x_{i+1}[$ pour tout i, alors nous posons

$$\int_a^b f(x)dx = \int_{x_0}^{x_1} f(x)dx + \int_{x_1}^{x_2} f(x)dx + \cdots + \int_{x_{n-1}}^{n_n} f(x)dx. \tag{14.381}$$

Sur chacun des morceaux, l'intégrale se calcule normalement en passant par une primitive.

14.14 Intégrales le long de chemins

14.14.1 Circulation d'un champ de vecteur

Définition 14.99.
Soit $F\colon \mathbb{R}^3 \to \mathbb{R}^3$ un champ de vecteurs et un chemin $\sigma\colon [a,b] \to \mathbb{R}^3$. On appelle ***circulation*** *de F le long du chemin σ le scalaire*

$$\int_a^b F\big(\sigma(t)\big) \cdot \sigma'(t)dt. \tag{14.382}$$

Il existe de nombreuses notations pour cela ; entre autres :

$$\int_\sigma F = \int_\sigma F \cdot ds. \tag{14.383}$$

En physique, la circulation de la force le long d'un chemin est la travail de la force.

Exemple 14.100
À la surface de la Terre, le champ de gravitation est donné par

$$G(x,y,z) = -mg \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{14.384}$$

Si nous considérons un mobile qui monte à vitesse constante jusqu'à la hauteur h, c'est à dire le chemin

$$\sigma(t) = \begin{pmatrix} 0 \\ 0 \\ t \end{pmatrix} \tag{14.385}$$

avec $t \in [0,h]$. Le travail de la gravitation est alors donné par

$$W = \int_0^h G\big(\sigma(t)\big) \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = -mg \int_0^h \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = -mgh. \tag{14.386}$$

Cela est bien le résultat usuel de l'énergie potentielle. Nous allons voir bientôt que nous nommons la fonction mgh énergie *potentielle* précisément parce que la force dérive de ce potentiel. △

Exemple 14.101
Soit le chemin

$$\begin{aligned}\sigma\colon [0,2\pi] &\to \mathbb{R}^3\\ t &\mapsto \begin{pmatrix}\sin(t)\\ \cos(t)\\ t\end{pmatrix}.\end{aligned} \tag{14.387}$$

et le champ de vecteurs

$$F\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{pmatrix}x\\ y\\ z\end{pmatrix}. \tag{14.388}$$

La circulation de ce champ de vecteur le long de l'hélice σ est

$$\begin{aligned}\int_\sigma F\cdot ds &= \int_0^{2\pi}(F\circ\sigma)(t)\cdot\sigma'(t)dt\\ &= \int_0^{2\pi}\begin{pmatrix}\sin(t)\\ \cos(t)\\ t\end{pmatrix}\cdot\begin{pmatrix}\cos(t)\\ \sin(t)\\ 1\end{pmatrix}dt\\ &= \int_0^{2\pi} t dt\\ &= \left[\frac{t^2}{2}\right]_0^{2\pi}\\ &= 2\pi^2.\end{aligned} \tag{14.389}$$

△

Proposition 14.102.
La circulation d'un champ de vecteurs le long d'un chemin ne dépend pas de la paramétrisation. En d'autres termes, si σ_1 et σ_2 sont deux chemins équivalents, alors

$$\int_{\sigma_1} F = \int_{\sigma_2} F. \tag{14.390}$$

Démonstration. Soient deux chemins $\sigma_1\colon [a,b]\to\mathbb{R}^3$ et $\sigma_2\colon [c,d]\to\mathbb{R}^3$ équivalents, c'est à dire tels que

$$\sigma_1(t) = \sigma_2\big(\varphi(t)\big) \tag{14.391}$$

où $\varphi\colon [a,b]\to[c,d]$ strictement croissante. En utilisant le fait que $\sigma_1(t) = \varphi'(t)\sigma_2'\big(\varphi(t)\big)$, nous avons

$$\begin{aligned}\int_{\sigma_1} F\cdot ds &= \int_a^b F\big(\sigma_1(t)\big)\cdot\sigma_1'(t)dt\\ &= \int_a^b F\Big(\sigma_2\big(\varphi(t)\big)\Big)\cdot\sigma_2'\big(\varphi(t)\big)\varphi'(t)dt\\ &= \int_{\varphi(a)}^{\varphi(b)} F\big(\sigma_2(s)\big)\cdot\sigma_2(s)ds\\ &= \int_c^d F\big(\sigma_2(s)\big)\cdot\sigma_2'(s)ds\\ &= \int_{\sigma_2} F\cdot ds.\end{aligned} \tag{14.392}$$

où nous avons effectué le changement de variables $s=\varphi(t)$, $ds=\varphi'(t)dt$. □

Remarque 14.103.
Si σ_2 est le chemin opposé de σ, alors

$$\int_{\sigma_2} F = -\int_{\sigma_1} F. \tag{14.393}$$

14.15 Circulation d'un champ conservatif

Si nous avons une fonction scalaire $V\colon \mathbb{R}^3 \to \mathbb{R}$, nous pouvons construire un champ de vecteur en prenant le gradient :

$$F(x) = \nabla V(x). \tag{14.394}$$

On dit que le champ de vecteur F **dérive** de V, et on dit que V est le **potentiel** de F. Nous posons la définition suivante :

Définition 14.104.
*Un champ de vecteurs $F\colon \mathbb{R}^3 \to \mathbb{R}^3$ est un champ **conservatif** s'il existe une fonction $V\colon \mathbb{R}^3 \to \mathbb{R}$ telle que*

$$F(x) = \nabla V(x). \tag{14.395}$$

Nous disons aussi parfois que le champ V dérive d'un potentiel *ou bien qu'il s'agit d'un* champ de gradient*.*

Les champs de vecteurs conservatifs sont particulièrement importants parce que presque toutes les forces connues en physiques dérivent d'un potentiel. Nous verrons que la terminologie «conservatif» provient du fait que les forces de ce type conservent l'énergie associée.

Proposition 14.105.
Considérons une fonction $V\colon \mathbb{R}^3 \to \mathbb{R}$ (que nous appellerons potentiel*) et le champ de vecteur qui en dérive :*

$$F = \nabla V. \tag{14.396}$$

Alors

$$\int_{\sigma} F \cdot ds = V\big(\sigma(b)\big) - V\big(\sigma(a)\big). \tag{14.397}$$

Autrement dit, le travail nécessaires pour déplacer un objet d'un point à un autre dans un champ de force conservatif vaut la différence de potentiel entre le point de départ et le point d'arrivée.

Démonstration. Par définition,

$$\int_{\sigma} F \cdot ds = \int_a^b F\big(\sigma(t)\big) \cdot \sigma'(t)dt. \tag{14.398}$$

Nous pouvons transformer l'intégrante de la façon suivante :

$$\begin{aligned} F\big(\sigma(t)\big) \cdot \sigma'(t) &= \nabla V\big(\sigma(t)\big) \cdot \sigma'(t) \\ &= \frac{\partial V}{\partial x}\big(\sigma(t)\big)\sigma_x'(t) + \frac{\partial V}{\partial y}\big(\sigma(t)\big)\sigma_y'(t) + \frac{\partial V}{\partial z}\big(\sigma(t)\big)\sigma_z'(t) \\ &= \frac{d}{dt}\Big[V\big(\sigma(t)\big)\Big] \end{aligned} \tag{14.399}$$

où nous avons posé

$$\sigma(t) = \begin{pmatrix} \sigma_x(t) \\ \sigma_y(t) \\ \sigma_z(t) \end{pmatrix} \tag{14.400}$$

et utilisé à l'envers la formule de dérivation de fonction composée pour

$$\frac{d}{dt}\Big[V\big(\sigma(t)\big)\Big] = \Big((V \circ \sigma)(t)\Big)'. \tag{14.401}$$

En remettant ces expressions dans l'intégrale (14.398),

$$\int_\sigma F \cdot ds = \int_a^b \frac{d}{dt}\Big[V\big(\sigma(t)\big)\Big]dt = V\big(\sigma(b)\big) - V\big(\sigma(a)\big). \tag{14.402}$$

□

Exemple 14.106
Nous savons que le champ de gravitation dérive d'un potentiel. À la surface de la Terre, le potentiel de gravitation vu par une masse m est donné par la fonction $V(x,y,z) = mgz$. Si nous voulons soulever cette masse d'une hauteur h, cela demandera toujours une énergie mgh, quel que soit le chemin suivit : en ligne droite vertical, en diagonal, en hélice, ... △

Exemple 14.107
À plus grande échelle, le champ de gravitation est encore un champ qui dérive d'un potentiel. En coordonnées sphériques,

$$V(\rho,\theta,\varphi) = k\frac{m}{\rho} \tag{14.403}$$

Lorsqu'un satellite a une orbite de rayon R autour la Terre, il reste sur la sphère $\rho = R$. Donc il reste sur une surface sur laquelle V est constante. Il n'y a donc pas de travail de la force de gravitation ! C'est pour cela qu'un satellite peut tourner pendant des siècles sans apport énergétique. △

Exemple 14.108
Soit le champ de vecteurs

$$F\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}y\\x\end{pmatrix} \tag{14.404}$$

et le chemin

$$\sigma(t) = \begin{pmatrix}t^4/4\\ \sin^3(t\frac{\pi}{2}).\end{pmatrix} \tag{14.405}$$

Nous voulons calculer la circulation de F le long du chemin σ entre $t = 0$ et $t = 1$.

La première chose à voir est que $F = \nabla V$ avec $V(x,y) = xy$. Donc la circulation sera donnée par

$$\int_\sigma F \cdot ds = V\big(\sigma(1)\big) - V\big(\sigma(0)\big) = V\big(\frac{1}{4},1\big) - V(0,0) = \frac{1}{4} - 0 = \frac{1}{4}. \tag{14.406}$$

Nous n'avons pas réellement calculé l'intégrale. △

14.16 Intégration de fonction à deux variables

14.16.1 Intégration sur un domaine rectangulaire

Soit une fonction positive

$$\begin{aligned} f\colon [a,b]\times[c,d] &\to \mathbb{R}^+ \\ (x,y) &\mapsto f(x,y). \end{aligned} \tag{14.407}$$

L'intégrale de f sur le rectangle $[a,b]\times[c,d]$ est le volume sous le graphe de la fonction. C'est à dire le volume de l'ensemble

$$\{(x,y,z) \text{ tel que } (x,y)\in[a,b]\times[c,d], z \leqslant f(x,y)\}. \tag{14.408}$$

Théorème 14.109 (Théorème de Fubini)**.**
Soit une fonction $f\colon \mathbb{R}^2 \to \mathbb{R}$ une fonction continue par morceaux sur $\mathcal{R} = [a,b]\times[c,d]$. Alors

$$\int_{\mathcal{R}} f(x,y)dxdy = \int_a^b \left[\int_c^d f(x,y)dy\right]dx = \int_c^d \left[\int_a^b f(x,y)dx\right]dy. \tag{14.409}$$

En pratique, nous utilisons le théorème de Fubini pour calculer les intégrales sur des rectangles.

Exemple 14.110
Nous voudrions intégrer la fonction $f(x,y) - 4 + x^2 + y^2$ sur le rectangle de la figure 14.7.

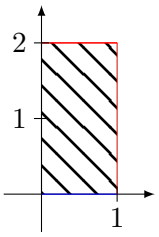


FIGURE 14.7 – Intégration sur un rectangle

L'ensemble sur lequel nous intégrons est donné par le produit cartésien d'intervalles $E = [0,1] \times [0,2]$. Le théorème de Fubini montre que nous pouvons intégrer séparément sur l'intervalle horizontal et vertical :

$$\int_{E=[0,1]\times[0,2]} f = \int_{[0,1]} \left(\int_{[0,2]} (4 - x^2 - y^2) dy \right) dx. \tag{14.410}$$

Ces intégrales sont maintenant des intégrales usuelles qui s'effectuent en calculant des primitives :

$$\begin{aligned} \int_0^1 \int_0^2 (4 - x^2 - y^2) dy\, dx &= \int_0^1 \left[4y - x^2 y - \frac{y^3}{3} \right]_0^2 dx \\ &= \int_0^1 (8 - 2x^2 - \frac{8}{3}) dx \\ &= \left[\frac{16x}{3} - \frac{2x^3}{3} \right]_0^1 \\ &= \frac{14}{3}. \end{aligned} \tag{14.411}$$

Avec Sage, on peut faire comme ceci :

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y)=4-x**2-y**2
sage: f.integrate(y,0,2).integrate(x,0,1)
(x, y) |--> 14/3
```

△

14.16.2 Intégration sur un domaine non rectangulaire

Nous voulons maintenant intégrer la fonction $f(x,y) = x^2 + y^2$ sur le triangle de la figure 14.8.

Étant donné que y varie de 0 à 2 et que *pour chaque* y, la variable x varie de 0 à y, nous écrivons l'intégrale sur le triangle sous la forme :

$$\int_{\text{triangle}} (x^2 + y^2) dx dy = \int_0^2 \left(\int_0^y (x^2 + y^2) dx \right) dy. \tag{14.412}$$

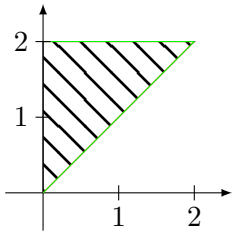
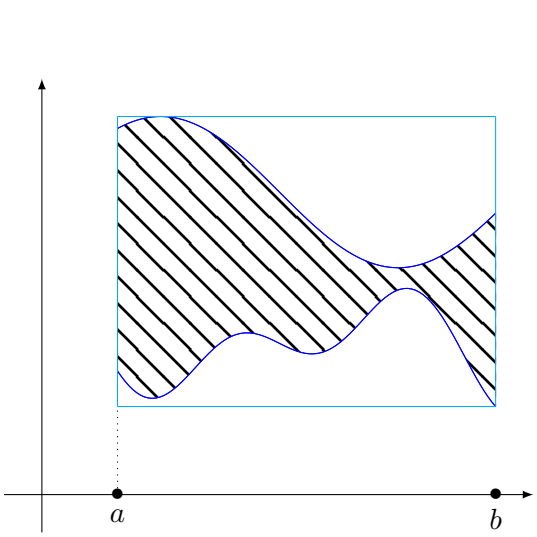
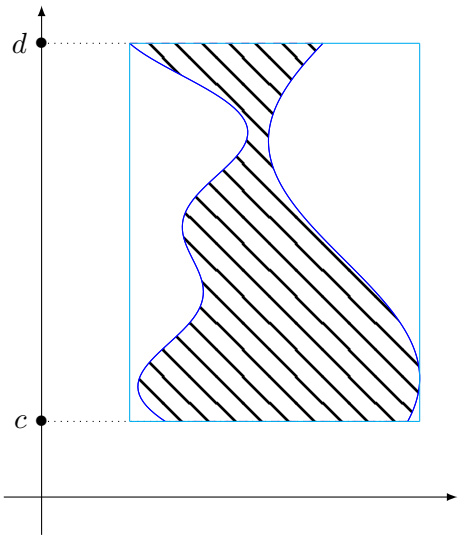


FIGURE 14.8 – Intégration sur un triangle

Il existe principalement deux types de domaines non rectangulaires : les «horizontaux» et les «verticaux», voir figure 14.9.



(a) Un domaine horizontal.



(b) Un domaine vertical.

FIGURE 14.9 – Deux types de surfaces. Nous avons tracé un rectangle qui contient chacune des deux surfaces. L'intégrale sur un domaine sera l'intégrale sur le rectangle de la fonction qui vaut zéro en dehors du domaine.

Les surfaces horizontales sont de la forme

$$D = \{(x,y) \text{ tel que } x \in [a,b], \varphi_1(x) \leqslant y \leqslant \varphi_2(x)\} \tag{14.413}$$

où φ_1 et φ_2 sont les deux fonctions qui bornent le domaine. Le domaine D est la région comprise entre les graphes de φ_1 et φ_2. Pour un tel domaine nous avons

$$\iint_D f(x,y)dxdy = \int_a^b dx \int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)dy. \tag{14.414}$$

Les surfaces verticales sont de la forme

$$D = \{(x,y) \text{ tel que } y \in [c,d], \psi_1(y) \leqslant x \leqslant \psi_2(y)\} \tag{14.415}$$

où φ_1 et φ_2 sont les deux fonctions qui bornent le domaine. Le domaine D est la région comprise entre les graphes de φ_1 et φ_2. Dans ces cas nous avons

$$\iint_D f = \int_c^d dy \int_{\psi_1(y)}^{\psi_2(y)} f(x,y)dx. \tag{14.416}$$

Proposition 14.111.
L'aire du domaine D vaut l'intégrable de la fonction $f(x,y) = 1$ sur D :

$$Aire(D) = \iint_D dxdy. \tag{14.417}$$

Démonstration. Supposons que le domaine soit du type «horizontal». En utilisant le théorème de Fubini avec $f(x,y)=1$ nous avons

$$\iint_D dxdy = \int_a^b \left[\int_{\varphi_1(x)}^{\varphi_2(x)} dy\right] dx = \int_a^b \big[\varphi_2(x)-\varphi_1(x)\big]. \tag{14.418}$$

Cela représente l'aire sous φ_2 moins l'aire sous φ_1, et par conséquent l'aire contenue entre les deux. □

Exemple 14.112
Cherchons la surface du disque de centre $(0,0)$ et de rayon 1 dessinée à la figure 14.10.

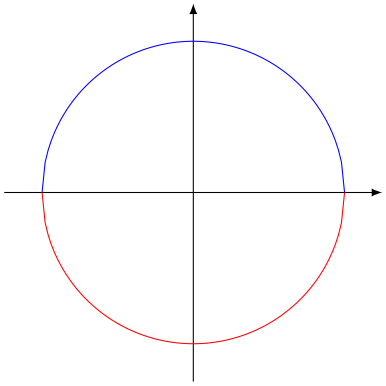


FIGURE 14.10 – En bleu, la fonction $\sqrt{r^2-x^2}$ et en rouge, la fonction $-\sqrt{r^2-x^2}$.

Le domaine est donné par $\varphi_1(x) \leqslant y \leqslant \varphi_2(x)$ et $x \in [-r,r]$ où $\varphi_1(x) = -\sqrt{r^2-x^2}$ et $\varphi_2(x) = \sqrt{r^2-x^2}$. L'aire est donc donnée par

$$A = \int_{-r}^{r} \big[\varphi_2(x)-\varphi_1(x)\big]dx = 2\int_{-r}^{r} \sqrt{r^2-x^2}dx = 4\int_0^r \sqrt{r^2-x^2}. \tag{14.419}$$

Nous effectuons le premier changement de variables $x=ru$, donc $dx=rdu$. En ce qui concerne les bornes, si $x=0$, alors $u=0$ et si $x=r$, alors $u=1$. L'intégrale à calculer devient

$$A = 4\int_0^1 \sqrt{r^2-r^2u^2}rdu = 4r^2\int_0^1 \sqrt{1-u^2}du. \tag{14.420}$$

Cette dernière intégrale se calcule en posant

$$\begin{aligned} &u=\sin(t) && du=\cos(t)dt\\ &u=0 && t=0\\ &u=1 && t=\pi/2. \end{aligned} \tag{14.421}$$

Nous avons

$$A = 4r^2\int_0^{\pi/2} \sqrt{1-\sin^2(t)}\cos(t)dt = 4r^2\int_0^{\pi/2}\cos^2(t)dt. \tag{14.422}$$

En utilisant la formule $2\cos^2(x) = 1+\cos(2x)$, nous avons

$$A = 4r^2\int_0^{\pi/2}\frac{1+\cos(2t)}{2}dt = \pi r^2. \tag{14.423}$$

△

14.16.3 Changement de variables

Comme dans les intégrales simples, il y a souvent moyen de trouver un changement de variables qui simplifie les expressions. Le domaine $E = \{(x, y) \in \mathbb{R}^2 \text{ tel que } x^2 + y^2 < 1\}$ par exemple s'écrit plus facilement $E = \{(r, \theta) \text{ tel que } r < 1\}$ en coordonnées polaires. Le passage aux coordonnées polaire permet de transformer une intégration sur un domaine rond à une intégration sur le domaine rectangulaire $]0, 2\pi[\times]0, 1[$. La question est évidement de savoir si nous pouvons écrire

$$\int_E f = \int_0^{2\pi} \int_0^1 f(r\cos\theta, r\sin\theta) dr d\theta. \tag{14.424}$$

Hélas ce n'est pas le cas. Il faut tenir compte du fait que le changement de base dilate ou contracte certaines surfaces.

Soit $\varphi\colon D_1 \subset \mathbb{R}^2 \to D_2 \subset \mathbb{R}^2$ une fonction bijective de classe C^1 dont l'inverse est également de classe C^1. On désigne par x et y ses composantes, c'est à dire que

$$\varphi(u, v) = \begin{pmatrix} x(u, v) \\ y(u, v) \end{pmatrix} \tag{14.425}$$

avec $(u, v) \in D_1$.

Théorème 14.113.
Soit une fonction continue $f\colon D_2 \to \mathbb{R}$. Alors

$$\iint_{\varphi(D_1)} f(x, y) dx dy = \iint_{D_1} f\big(x(u, v), y(u, v)\big) |J_\varphi(u, v)| du dv \tag{14.426}$$

où J_φ est le Jacobien de φ.

Pour rappel,

$$J_\varphi(u, v) = \det \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial u}{\partial v} \end{pmatrix}. \tag{14.427}$$

Ne pas oublier de prendre la valeur absolue lorsqu'on utilise le Jacobien dans un changement de variables.

14.16.3.1 Le cas des coordonnées polaires

La fonction qui donne les coordonnées polaires est

$$\begin{aligned} \varphi\colon \mathbb{R}^+ \times]0, 2\pi[&\to \mathbb{R}^2 \\ (r, \theta) &\mapsto \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}. \end{aligned} \tag{14.428}$$

Son Jacobien vaut

$$J_\varphi(r, \theta) = \det \begin{pmatrix} \frac{\partial x(r,\theta)}{\partial r} & \frac{\partial x(r,\theta)}{\partial \theta} \\ \frac{\partial y(r,\theta)}{\partial r} & \frac{\partial y(r,\theta)}{\partial \theta} \end{pmatrix} = \begin{vmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{vmatrix} = r. \tag{14.429}$$

Exemple 14.114
Calculons la surface du disque D de rayon R. Nous devons calculer

$$\iint_D dx dy. \tag{14.430}$$

Pour passer au polaires, nous savons que le disque est décrit par

$$D = \{(r, \theta) \text{ tel que } 0 \leqslant r \leqslant R, 0 \leqslant \theta \leqslant 2\pi\}. \tag{14.431}$$

Nous avons donc

$$\iint_D dxdy = \iint_D r\,drd\theta = \int_0^{2\pi}\int_0^R r\,drd\theta = 2\pi\int_0^R r\,dr = \pi R^2. \tag{14.432}$$

△

Exemple 14.115
Montrons comment intégrer la fonction $f(x,y) = \sqrt{1-x^2-y^2}$ sur le domaine délimité par la droite $y = x$ et le cercle $x^2 + y^2 = y$, représenté sur la figure 14.11. Pour trouver le centre et le rayon du cercle $x^2+y^2 = y$, nous commençons par écrire $x^2+y^2-y = 0$, et ensuite nous reformons le carré : $y^2 - y = (y - \frac{1}{2})^2 - \frac{1}{4}$.

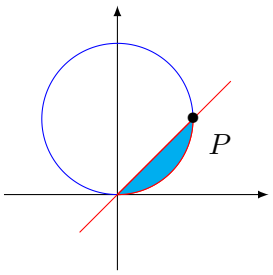


FIGURE 14.11 – Passage en polaire pour intégrer sur un morceau de cercle.

Le passage en polaire transforme les équations du bord du domaine en

$$\begin{aligned}\cos(\theta) &= \sin(\theta)\\ r^2 &= r\sin(\theta).\end{aligned} \tag{14.433}$$

L'angle θ parcours donc $]0,\pi/4[$, et le rayon, pour chacun de ces θ parcours $]0,\sin(\theta)[$. La fonction à intégrer se note maintenant $f(r,\theta) = \sqrt{1-r^2}$. Donc l'intégrale à calculer est

$$\int_0^{\pi/4}\left(\int_0^{\sin(\theta)}\sqrt{1-r^2}r\,rd\right). \tag{14.434}$$

Remarquez la présence d'un r supplémentaire pour le jacobien.

Notez que les coordonnées du point P sont $(1,1)$. △

En pratique, lors du passage en coordonnées polaires, le «$dxdy$» devient «$r\,drd\theta$».

14.16.3.2 Les coordonnées cylindriques

En ce qui concerne les coordonnées cylindriques, le Jacobien est donné par

$$J(r,\theta,z) = \begin{vmatrix}\frac{\partial x}{\partial r} & \frac{\partial x}{\partial\theta} & \frac{\partial x}{\partial z}\\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial\theta} & \frac{\partial y}{\partial z}\\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial\theta} & \frac{\partial z}{\partial z}\end{vmatrix} = \begin{vmatrix}\cos\theta & -r\sin\theta & 0\\ \sin\theta & r\cos\theta & 0\\ 0 & 0 & 1\end{vmatrix} = r. \tag{14.435}$$

Nous avons donc $dx\,dy\,dz = r\,dr\,d\theta\,dz$.

14.16.3.3 Coordonnées sphériques

Le calcul est un peu plus long :

$$\begin{aligned} J(\rho,\theta,\varphi) &= \begin{vmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \varphi} \\ \frac{\partial z}{\partial \rho} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \varphi} \end{vmatrix} \\ &= \begin{vmatrix} \sin\theta\cos\varphi & \rho\cos\theta\cos\varphi & -\rho\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & \rho\cos\theta\sin\varphi & -\rho\sin\theta\cos\varphi \\ \cos\theta & -\rho\sin\theta & 0 \end{vmatrix} \\ &= \rho^2\sin\theta. \end{aligned} \tag{14.436}$$

Donc

$$dx\,dy\,dz = \rho^2\sin(\theta)\,d\rho\,d\theta\,d\varphi. \tag{14.437}$$

14.16.3.4 Un autre système utile

Un changement de variables que l'on voit assez souvent est

$$\begin{cases} u = x + y & \text{(14.438a)} \\ v = x - y. & \text{(14.438b)} \end{cases}$$

Afin de calculer son jacobien, il faut d'abord exprimer x et y en fonctions de u et v :

$$\begin{cases} x = (u+v)/2 & \text{(14.439a)} \\ y = (u-v)/2. & \text{(14.439b)} \end{cases}$$

La matrice jacobienne est

$$\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}. \tag{14.440}$$

Le déterminant vaut $-\frac{1}{2}$. Nous avons donc

$$dxdy = \frac{1}{2}dudv. \tag{14.441}$$

Nous insistons sur le fait que c'est $\frac{1}{2}$ et non $-\frac{1}{2}$ qui intervient parce que que la formule du changement de variable demande d'introduire la *valeur absolue* du jacobien.

Exemple 14.116
Calculer l'intégrale de la fonction $f(x,y) = x^2 - y^2$ sur le domaine représenté sur la figure 14.12.

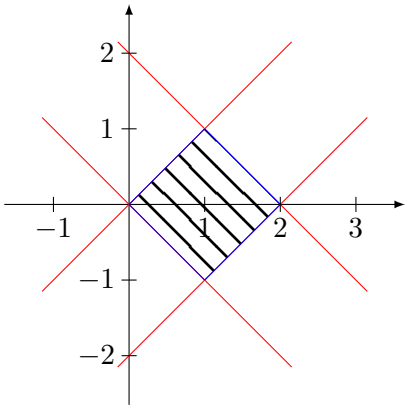


FIGURE 14.12 – Un domaine qui s'écrit étonnament bien avec un bon changement de coordonnées.

Les droites qui délimitent le domaine d'intégration sont

$$\begin{aligned} y &= -x + 2 \\ y &= x - 2 \\ y &= x \\ y &= -x \end{aligned} \tag{14.442}$$

Le domaine est donc donné par les équations

$$\begin{cases} y + x < 2 & \text{(14.443a)} \\ y - x > -2 & \text{(14.443b)} \\ y - x < 0 & \text{(14.443c)} \\ y + x > 0. & \text{(14.443d)} \end{cases}$$

En utilisant le changement de variables $u = x + y$, $v = x - y$ nous trouvons le domaine $0 < u < 2$, $0 < v < 2$. En ce qui concerne la fonction, $f(x, y) = (x + y)(x - y)$ et par conséquent

$$f(u, v) = uv. \tag{14.444}$$

L'intégrale à calculer est simplement

$$\int_0^2 \int_0^2 uv \, dudv = \int_0^2 u \, du \left[\frac{v^2}{2} \right]_0^2 = 2 \int_0^2 u \, du = 4. \tag{14.445}$$

△

14.17 Les intégrales triples

Les intégrales triples fonctionnent exactement de la même manière que les intégrales doubles. Il s'agit de déterminer sur quelle domaine les variables varient et d'intégrer successivement par rapport à x, y et z. Il est autorisé de permuter l'ordre d'intégration [13] à condition d'adapter les domaines d'intégration.

Exemple 14.117
Soit le domaine parallélépipédique rectangle

$$R = [0, 1] \times [1, 2] \times [0, 4]. \tag{14.446}$$

Pour intégrer la fonction $f(x, y, z) = x^2 y \sin(z)$ sur R, nous faisons

$$\begin{aligned} I &= \iiint_R x^2 y \sin(z) \, dxdydz \\ &= \int_0^1 dx \int_1^2 dy \int_0^4 x^2 y \sin(z) dz \\ &= \int_0^1 dx \int_1^2 x^2 y (1 - \cos(4)) dy \\ &= \int_0^1 \frac{3}{2} (1 - \cos(4)) x^2 dx \\ &= \frac{1}{2} \big(1 - \cos(4)\big). \end{aligned} \tag{14.447}$$

13. En toute rigueur, cela n'est pas vrai, mais nous ne considérons seulement des cas où cela est autorisé.

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y,z)=x**2*y*sin(z)
sage: f.integrate(x,0,1).integrate(y,1,2).integrate(z,0,4)
(x, y, z) |--> -1/2*cos(4) + 1/2
```

△

Exemple 14.118
Soit D la région délimitée par le plan $x = 0$, $y = 0$, $z = 2$ et la surface d'équation

$$z = x^2 + y^2. \tag{14.448}$$

Cherchons à calculer $\iiint_D x\,dx\,dy\,dz$. Ici, un dessin indique que le volume considéré est $z \geqslant x^2 + y^2$. Il y a plusieurs façon de décrire cet ensemble. Une est celle-ci :

$$\begin{aligned} z&\colon 0 \to 2\\ x&\colon 0 \to \sqrt{z}\\ y&\colon 0 \to \sqrt{z - x^2}. \end{aligned} \tag{14.449}$$

Cela revient à dire que z peut prendre toutes les valeurs de 0 à 2, puis que pour chaque z, la variable x peut aller de 0 à $\sqrt{z}$, mais que pour chaque z et x fixés, la variable y ne peut pas dépasser $\sqrt{z - x^2}$. En suivant cette méthode, l'intégrale à calculer est

$$\int_0^2 dz \int_0^{\sqrt{z}} dx \int_0^{\sqrt{z-x^2}} f(x, y, z)dy. \tag{14.450}$$

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y,z)=x
sage: assume(z>0)
sage: assume(z-x**2>0)
sage: f.integrate(y,0,sqrt(z-x**2)).integrate(x,0,sqrt(z)).integrate(z,0,2)
(x, y, z) |--> 8/15*sqrt(2)
```

Notez qu'il a fallu aider Sage en lui indiquant que $z > 0$ et $z - x^2 > 0$.

Une autre paramétrisation serait

$$\begin{aligned} x&\colon 0 \to \sqrt{2}\\ y&\colon 0 \to \sqrt{2 - x^2}\\ z&\colon x^2 + y^2 \to 2. \end{aligned} \tag{14.451}$$

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y,z)=x
sage: assume(2-x**2>0)
sage: f.integrate(y,0,sqrt(z-x**2)).integrate(x,0,sqrt(z)).integrate(z,0,2)
(x, y, z) |--> 8/15*sqrt(2)
```

Écrivons le détail de cette dernière intégrale :

$$\begin{aligned} I &= \int_0^{\sqrt{2}} dx \int_0^{\sqrt{2-x^2}} dy \int_{x^2+y^2}^{2} x dz \\ &= \int_0^{\sqrt{2}} dx \int_0^{\sqrt{2-x^2}} x(2-x^2-y^2)dy \\ &= \int_0^{\sqrt{2}} dx\, x \left[(2-x^2)y - \frac{y^3}{3} \right]_0^{\sqrt{2-x^2}} \\ &= \int_0^{\sqrt{2}} \frac{2}{3} x(2-x^2)^{3/2} dx. \end{aligned} \tag{14.452}$$

Ici nous effectuons le changement de variable $u = x^2$, $du = 2xdx$. Ne pas oublier de changer les bornes de l'intégrale :

$$I = \frac{1}{3} \int_0^2 (2-u)^{3/2} du. \tag{14.453}$$

Le changement de variable $t = 2 - u$, $dt = -du$ fait venir (attention aux bornes ! !)

$$I = -\frac{1}{3} \int_2^0 t^{3/2} dt = \frac{1}{3} \left[\frac{t^{5/2}}{5/2} \right]_0^2 = \frac{8}{15}\sqrt{2}. \tag{14.454}$$

△

14.17.1 Volume

Parmi le nombreuses interprétations géométriques de l'intégrale triple, notons celle-ci :

Proposition 14.119.
Soit $D \subset \mathbb{R}^3$. Le volume de D est donné par

$$Vol(D) = \iiint_D dxdydz. \tag{14.455}$$

C'est à dire l'intégrale de la fonction $f(x, y, z) = 1$ sur D.

Suivant les points de vue, cette proposition peut être considérée comme une *définition*] du volume.

Exemple 14.120
Calculons le volume de la sphère de rayon R. Le domaine de variation des variables x, y et z pour la sphère est

$$\begin{aligned} &x\colon\ -R \to R \\ &y\colon\ -\sqrt{R^2-x^2} \to \sqrt{R^2-x^2} \\ &z\colon\ -\sqrt{R^2-x^2-y^2} \to \sqrt{R^2-x^2-y^2}. \end{aligned} \tag{14.456}$$

Par conséquent nous devons calculer l'intégrale

$$V = \int_{-R}^{R} dx \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} dy \int_{-\sqrt{R^2-x^2-y^2}}^{\sqrt{R^2-x^2-y^2}} dz. \tag{14.457}$$

La première intégrale est simple :

$$V = 2 \int_{-R}^{R} dx \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \sqrt{R^2-x^2-y^2} dy. \tag{14.458}$$

Afin de simplifier la notation, nous posons $a = R^2 - x^2$. Ceci n'est pas un changement de variables : juste une notation provisoire le temps d'effectuer l'intégration sur y. Étudions donc

$$I = \int_{-\sqrt{a}}^{\sqrt{a}} \sqrt{a - y^2} dy, \tag{14.459}$$

ce qui est la surface du demi-disque de rayon $\sqrt{a}$. Nous avons donc

$$I = \frac{\pi a}{2} = \frac{\pi}{2}(R^2 - x^2), \tag{14.460}$$

et

$$V = 2\int_{-R}^{R} \frac{\pi}{2}(R^2 - x^2)dx = \pi \left[R^2 x - \frac{x^3}{3} \right]_{-R}^{R} = \frac{4}{3}\pi R^3. \tag{14.461}$$

△

Exemple 14.121
Nous pouvons calculer le volume de la sphère en utilisant les coordonnées sphériques. Les bornes des variables pour la sphère de rayon R sont

$$\begin{aligned} \rho &: 0 \to R \\ \theta &: 0 \to \pi \\ \varphi &: 0 \to 2\pi. \end{aligned} \tag{14.462}$$

En n'oubliant pas le jacobien $\rho^2 \sin(\theta)$, l'intégrale à calculer est

$$V = \int_0^R d\rho \int_0^{2\pi} d\varphi \int_0^{\pi} \rho^2 \sin(\theta) d\theta \tag{14.463}$$

L'intégrale sur φ fait juste une multiplication par 2π. Celle sur ρ vaut

$$\int_0^R \rho^2 d\rho = \frac{R^3}{3}. \tag{14.464}$$

L'intégrale sur θ donne

$$\int_0^{\pi} \sin(\theta) d\theta = [-\cos(\theta)]_0^{\pi} = 2. \tag{14.465}$$

Le tout fait par conséquent

$$V = \frac{4}{3}\pi R^3. \tag{14.466}$$

Sans contestes, le passage aux coordonnées sphériques a considérablement simplifié le calcul par rapport à celui de l'exemple 14.120. △

14.18 Un petit peu plus formel

14.18.1 Intégration sur un domaine non rectangulaire

La méthode de Fubini ne fonctionne plus sur un domaine non rectangulaire tel que celui de la figure 14.13. Nous allons donc utiliser une astuce. Considérons le domaine

$$E = \{(x, y) \in \mathbb{R}^2 \text{ tel que } a < x < b \text{ et } \alpha(x) < y < \beta(x)\} \tag{14.467}$$

représenté sur la figure 14.13. Nous considérons la fonction

$$\tilde{f}(x, y) = \begin{cases} f(x, y) & \text{si } (x, y) \in E \\ 0 & \text{sinon.} \end{cases} \tag{14.468}$$

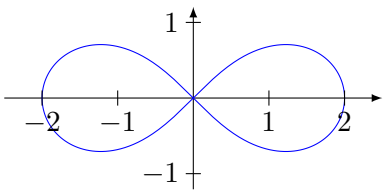


FIGURE 14.13 – Intégrer sur des domaines plus complexes.

Ensuite intégrons $\tilde{f}$ sur un rectangle qui englobe la surface à intégrer à l'aide de Fubini. Étant donné que $\tilde{f} = f$ sur la surface et que $\tilde{f}$ est nulle en dehors, nous avons

$$\int_E f = \int_E \tilde{f} = \int_{\text{rectangle}} \tilde{f} = \int_a^b \left(\int_{\alpha(x)}^{\beta(x)} f(x,y)dy \right) dx. \tag{14.469}$$

Dans le cas de l'intégrale de $f(x,y) = x^2 + y^2$ sur le triangle de la figure 14.8, nous avons

$$\int_{\text{triangle}} (x^2 + y^2)dxdy = \int_0^2 \left(\int_0^y (x^2 + y^2)dx \right) dy. \tag{14.470}$$

Remarque 14.122.
Le nombre $\iint_D f(x,y)dxdy$ ne dépend pas du choix du rectangle englobant D.

En pratique, nous calculons l'intégrale en utilisant une extension du théorème de Fubini :

Théorème 14.123.
Soit $f\colon D \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction continue où D est un domaine de type vertical ou horizontal.

(1) Si D est vertical, alors

$$\iint_D f = \int_a^b \left[\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)dy \right] dx. \tag{14.471}$$

(2) Si D est horizontal, alors

$$\iint_D f = \int_c^d \left[\int_{\psi_1(y)}^{\psi_2(y)} f(x,y)dx \right] dy. \tag{14.472}$$

14.18.2 Changement de variables

Le théorème du changement de variable est le suivant.

Théorème 14.124.
Soit $g\colon A \to B$ un difféomorphisme. Soient $F \subset B$ un ensemble mesurable et borné et $f\colon F \to \mathbb{R}$ une fonction bornée et intégrable. Supposons que $g^{-1}(F)$ soit borné et que Jg soit borné sur $g^{-1}(F)$. Alors

$$\int_F f(x)dy = \int_{g^{-1}(F)} f\big(g(x)\big)|Jg(x)|dx \tag{14.473}$$

Pour rappel, Jg est le déterminant de la matrice jacobienne (aucun lien de parenté) donnée par

$$Jg = \det \begin{pmatrix} \partial_x g_1 & \partial_y g_1 \\ \partial_x g_2 & \partial_t g_2 \end{pmatrix}. \tag{14.474}$$

Un **difféomorphisme** est une application $g\colon A \to B$ telle que g et $g^{-1}\colon B \to A$ soient de classe C^1.

14.18.2.1 Coordonnées polaires

Les coordonnées polaires sont données par le difféomorphisme

$$\begin{aligned} g\colon]0,\infty[\times]0,2\pi[&\to \mathbb{R}^2\backslash D \\ (r,\theta) &\mapsto \big(r\cos(\theta), r\sin(\theta)\big) \end{aligned} \tag{14.475}$$

où D est la demi droite $y = 0$, $x \geqslant 0$. Le fait que les coordonnées polaires ne soient pas un difféomorphisme sur tout $\mathbb{R}^2$ n'est pas un problème pour l'intégration parce que le manque de difféomorphisme est de mesure nulle dans $\mathbb{R}^2$. Le jacobien est donné par

$$Jg = \det\begin{pmatrix} \partial_r x & \partial_\theta x \\ \partial_r y & \partial_\theta y \end{pmatrix} = \det\begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix} = r. \tag{14.476}$$

14.18.2.2 Coordonnées sphériques

Les coordonnées sphériques sont données par

$$\left\{\begin{array}{ll} x = r\cos\theta\sin\varphi & r \in]0,\infty[\\ y = r\sin\theta\sin\varphi \quad \text{avec} & \theta \in]0,2\pi[\\ z = r\cos\varphi & \phi \in]0,\pi[. \end{array}\right. \tag{14.477}$$

Le jacobien associé est $Jg(r,\theta,\varphi) = -r^2\sin\varphi$. Rappelons que ce qui rentre dans l'intégrale est la valeur absolue du jacobien.

Si nous voulons calculer le volume de la sphère de rayon R, nous écrivons donc

$$\int_0^R dr \int_0^{2\pi} d\theta \int_0^{\pi} r^2\sin(\phi)d\phi = 4\pi R = \frac{4}{3}\pi R^3. \tag{14.478}$$

Ici, la valeur absolue n'est pas importante parce que lorsque $\phi \in]0,\pi,[$, le sinus de ϕ est positif.

Des petits malins pourraient remarquer que le changement de variable (14.477) est encore une paramétrisation de $\mathbb{R}^3$ si on intervertit le domaine des angles :

$$\begin{aligned} \theta &\colon 0 \to \pi \\ \phi &\colon 0 \to 2\pi, \end{aligned} \tag{14.479}$$

alors nous paramétrons encore parfaitement bien la sphère, mais hélas

$$\int_0^R dr \int_0^{\pi} d\theta \int_0^{2\pi} r^2\sin(\phi)d\phi = 0. \tag{14.480}$$

Pourquoi ces «nouvelles» coordonnées sphériques sont-elles mauvaises ? Il y a que quand l'angle ϕ parcours $]0,2\pi[$, son sinus n'est plus toujours positif, donc la *valeur absolue* du jacobien n'est plus $r^2\sin(\phi)$, mais $r^2\sin(\phi)$ pour les ϕ entre 0 et π, puis $-r^2\sin(\phi)$ pour ϕ entre π et 2π. Donc l'intégrale (14.480) n'est pas correcte. Il faut la remplacer par

$$\int_0^R dr \int_0^{\pi} d\theta \int_0^{\pi} r^2\sin(\phi)d\phi - \int_0^R dr \int_0^{\pi} d\theta \int_{\pi}^{2\pi} r^2\sin(\phi)d\phi = \frac{4}{3}\pi R^3 \tag{14.481}$$

14.19 Aire et primitive

Soit $f\colon \mathbb{R} \to \mathbb{R}$, une fonction continue, et $x \in \mathbb{R}$. Pour chaque $x \in \mathbb{R}$, nous pouvons considérer le nombre $F(x)$ défini par

$$F(x) = \int_a^x f(t)dt. \tag{14.482}$$

La fonction F ainsi définie a deux importantes propriétés :

(1) C'est une primitive de f,

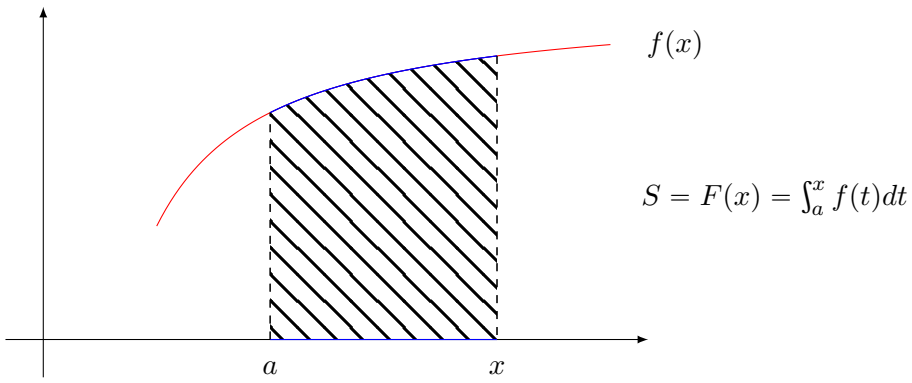


FIGURE 14.14 – Surface sous une courbe

(2) Elle donne la surface en dessous de f entre les points a et x, voir la figure 14.14.

Notons que tant que f est positive, la surface est croissante.
La manière de calculer la surface comprise entre deux fonctions est dessinée à la figure 14.15.
La surface entre les deux fonctions $y_1(x)$ et $y_2(x)$ se calcule comme suit.

(1) On calcule les intersections entre $y1$ et y_2. Notons a et b les ordonnées obtenues.

(2) La surface demandée est la différence entre la surface sous la fonction y_1 (la plus grande) et la surface sous la fonction y_2 (la plus petite), donc

$$S = \int_a^b y_1 - \int_a^b y_1. \tag{14.483}$$

14.19.1 Longueur d'arc de courbe

La longueur de l'arc de courbe de la fonction $y = f(x)$ entre les abscisses x_0 et x_1 est donné par la formule

$$l(x_0, x_1) = \int_{x_0}^{x_1} \sqrt{1 + y'(t)^2}dt. \tag{14.484}$$

Lorsque la courbe est donnée sous forme paramétrique

$$\begin{cases} x = x(t) & \text{(14.485a)} \\ y = y(t), & \text{(14.485b)} \end{cases}$$

alors la formule devient

$$l(t_1, t_2) = \int_{t_1}^{t_2} \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}dt, \tag{14.486}$$

où $\dot{x}(t) = x'(t)$.

14.19.2 Aire de révolution

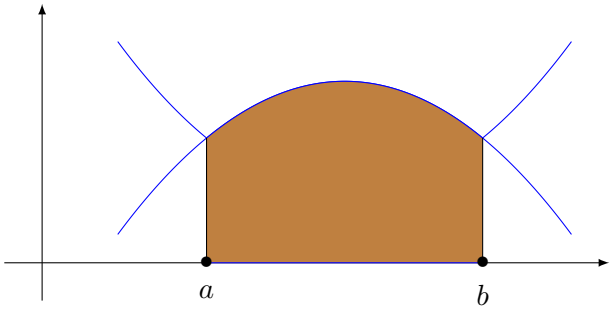
Pour savoir l'aire engendrée par la ligne $y = f(x)$ entre a et b autour de l'axe Ox, on utilise la formule

$$S = 2\pi \int_a^b \sqrt{1 + f'(x)^2} f(x)dx. \tag{14.487}$$

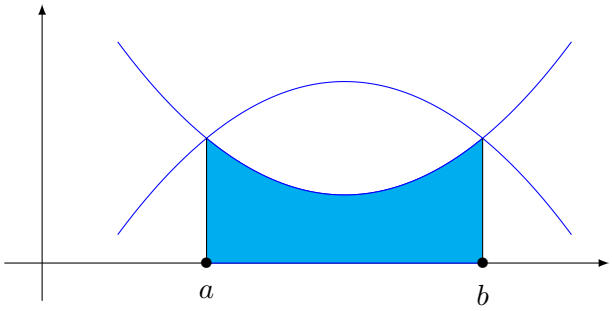
14.20 L'aire en dessous d'une courbe

Soit f une fonction à valeurs dans $\mathbb{R}^+$.

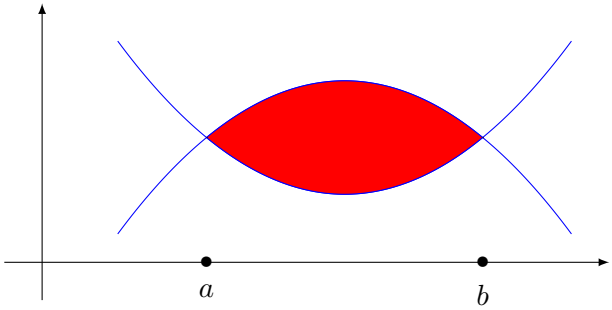
Nous voudrions pouvoir calculer l'aire au-dessous du graphe de la fonction f. Nous notons $S_f(x)$ l'aire là-dessous de la fonction f entre l'abscisse 0 et x, c'est à dire l'aire bleue de la figure 14.16.



(a) La grande surface.



(b) La surface à soustraire.



(c) La surface entre les deux fonctions.

FIGURE 14.15 – Le calcul de la surface comprise entre deux fonctions.

Si la fonction f est continue et que Δx est assez petit, la fonction ne varie pas beaucoup entre x et $x+\Delta x$. L'augmentation de surface entre x et $x+\Delta x$ peut donc être approximée par le rectangle de surface $f(x)\Delta x$. Ce que nous avons donc, c'est que quand Δx est très petit,

$$S_f(x+\Delta x) - S_f(x) = f(x)\Delta x, \tag{14.488}$$

ou encore

$$f(x) = \frac{S_f(x+\Delta x) - S_f(x)}{\Delta x}. \tag{14.489}$$

Nous formalisons la notion de «lorsque Δx est très petit» par une limite :

$$f(x) = \lim_{\Delta x \to 0} \frac{S_f(x+\Delta x) - S_f(x)}{\Delta x}. \tag{14.490}$$

Donc, la fonction f est la dérivée de la fonction qui représente l'aire là-dessous de f. Calculer des surfaces revient donc au travail inverse de calculer des dérivées.

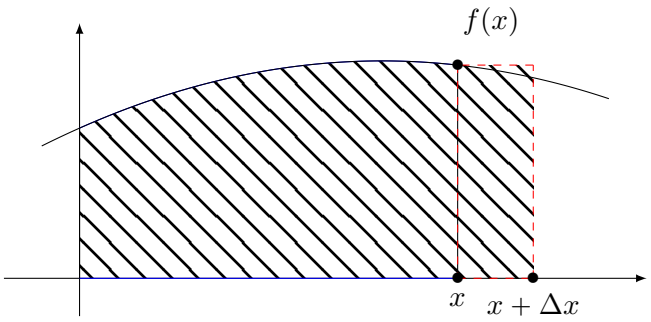


FIGURE 14.16 – L'aire en dessous d'une courbe. Le rectangle rouge d'aire $f(x)\Delta x$ approxime de combien la surface augmente lorsqu'on passe de x à $x + \Delta x$.

14.21 Propriétés des intégrales

Lemme 14.125.
Pour toute fonction $u\colon [a,b] \to \mathbb{R}^n$, nous avons

$$\| \int_a^b u(t)dt \| \leqslant \int_a^b \|u(t)\| dt \tag{14.491}$$

pourvu que le membre de gauche ait un sens.

Démonstration. Étant donné que $\int_a^b u(t)dt$ est un élément de $\mathbb{R}^n$, par la proposition 9.267, il existe un $\xi \in \mathbb{R}^n$ de norme 1 tel que

$$\| \int_a^b u(t)dt \| = \xi \cdot \int_a^b u(t)dt = \int_a^b u(t) \cdot \xi dt \leqslant \int_a^b \|u(t)\|\|\xi\| = \int_a^b \|u(t)\|dt. \tag{14.492}$$

□

Proposition 14.126 (Relations de Chasles)**.**
Soit f une fonction continue sur l'intervalle I. Si $a, b, c \in I$ nous avons

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx. \tag{14.493}$$

Sur la figure 14.17, la surface de a à c est évidemment égale à la somme des surfaces de a à b et de b à c.

Corollaire 14.127.

$$\int_a^b f(x)dx = -\int_b^a f(x)dx. \tag{14.494}$$

Proposition 14.128 (Linéarité de l'intégrale)**.**
Si f et g sont deux fonction continues sur $I \subset \mathbb{R}$, a, $b \in I$ et $\lambda \in \mathbb{R}$ nous avons

$$\int_a^b \big(f(x) + g(x)\big)dx = \int_a^b f(x)dx + \int_a^b g(x)dx, \tag{14.495}$$

et

$$\int_a^b \lambda f(x)dx = \lambda \int_a^b f(x)dx. \tag{14.496}$$

Proposition 14.129 (L'intégrale est monotone)**.**
Soient $a, b \in I$ avec $a < b$. Si $f \geqslant g$ sur $[a,b]$ alors

$$\int_a^b f(x)dx \geqslant \int_a^b g(x)dx. \tag{14.497}$$

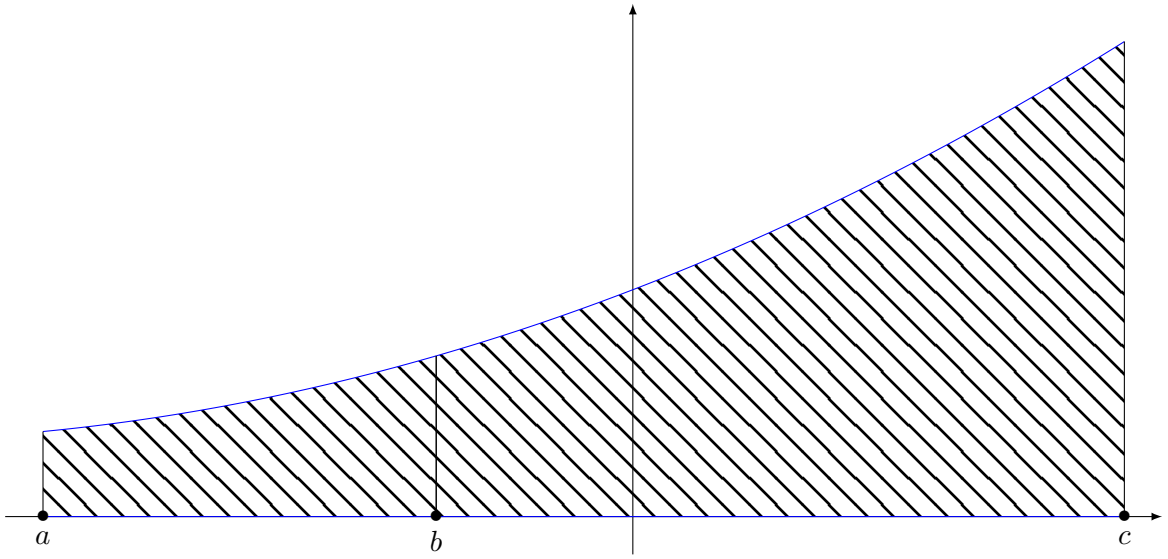


FIGURE 14.17 – Illustration pour les relations de Chasles.

Corollaire 14.130 (Positivité)**.**
Si $a < b$ et $f \geqslant 0$ sur $[a, b]$ alors

$$\int_a^b f(x)dx \geqslant 0. \tag{14.498}$$

Ce résultat n'est qu'une application de la proposition 14.129 car il consiste à prendre comme fonction g la fonction nulle.

14.22 Techniques d'intégration

Par le corollaire 13.208, la calcul d'une intégrale consiste essentiellement à trouver une primitive de la fonction à intégrer. Il est donc indispensable de bien connaître les dérivées des fonctions usuelles.

Voici un tableau des primitives à connaître.

Fonction $f(x)$	Primitive $\int f(x)\,dx$	Ensemble de définition de f	Remarques
x^α	$\frac{x^{\alpha+1}}{\alpha+1} + C$	dépend de α	$\alpha \in \mathbb{R}\backslash\{-1\}$
$\frac{1}{x}$	$\ln\big(\lvert x\rvert\big) + C$	$x \neq 0$	
$\frac{1}{1+x^2}$	$\arctan(x) + C$	$\mathbb{R}$	
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x) + C$	$]-1, 1[$	
$\frac{-1}{\sqrt{1-x^2}}$	$\arccos(x) + C$	$]-1, 1[$	
e^x	$e^x + C$	$\mathbb{R}$	
$\sin(x)$	$-\cos(x) + C$	$\mathbb{R}$	
$\cos(x)$	$\sin(x) + C$	$\mathbb{R}$	
$1 + \tan^2(x)$	$\tan(x) + C$	in intervalle de la forme $]-\frac{\pi}{2}, \frac{\pi}{2}[+ k\pi$	

Notez que au signe près, les fonctions arcsin et arccos ont la même dérivée.

Si la fonction à intégrer est une combinaison linéaire de fonctions usuelles alors sa primitive peut être calculée en utilisant la proposition 14.128. Dans les sections suivantes on abordera deux autres cas où la fonction à intégrer peut s'écrire en termes de fonctions dont on connaît une primitive.

14.22.1 Intégration par parties

Proposition 14.131.
Si u et v sont deux fonctions dérivables de dérivées continues sur l'intervalle $[a, b]$ alors

$$\int_a^b u(x)v'(x)dx = \big[u(x)v(x)\big]_a^b - \int_a^b u'(x)v(x)dx. \tag{14.499}$$

Démonstration. Il s'agit d'utiliser à l'envers la formule de dérivation d'un produit :

$$uv' = (uv)' - u'v. \tag{14.500}$$

Les fonctions à gauche et à droite étant égales, elles ont même intégrale sur $[a, b]$ et par linéarité, voir proposition 14.128, on a :

$$\int_a^b u(x)v'(x)dx = \int_a^b (uv)'(x) - \int_a^b u'(x)v(x)dx. \tag{14.501}$$

La fonction uv est évidemment une primitive de $(uv)'$, de telle sorte que l'on puisse un peu simplifier cette expression :

$$\int_a^b u(x)v'(x)dx = \Big[u(x)v(x)\Big]_a^b - \int_a^b u'(x)v(x)dx, \tag{14.502}$$

ce qu'il fallait démontrer. □

Exemple 14.132
Un cas typique d'utilisation de l'intégrale par parties est le suivant. Soit à calculer

$$\int_0^\pi x\cos(x)dx. \tag{14.503}$$

Nous devons écrire $x\cos(x)$ comme un produit $u(x)v'(x)$. Il y a (au moins) deux moyens de le faire :

$$\begin{cases} u = x & \text{(14.504a)} \\ v' = \cos(x). & \text{(14.504b)} \end{cases}$$

ou

$$\begin{cases} u = \cos(x) & \text{(14.505a)} \\ v' = x. & \text{(14.505b)} \end{cases}$$

Nous allons choisir le premier [14]. Nous avons donc

$$\begin{array}{ll} u = x, & v' = \cos(x) \\ u' = 1 & v = \sin(x). \end{array} \tag{14.506}$$

En utilisant la formule d'intégration par parties,

$$\int_0^\pi x\cos(x)dx = \Big[x\sin(x)\Big]_0^\pi - \int_0^\pi 1\times\sin(x)dx = \pi\sin(\pi) - \Big[-\cos(x)\Big]_0^\pi = -2. \tag{14.507}$$

△

Le plus souvent, pour alléger les notations, il est plus pratique d'utiliser l'intégration par parties pour déterminer une primitive. Nous utilisons pour cela la formule (sans doute plus simple à retenir)

$$\int uv' = uv - \int u'v. \tag{14.508}$$

14. Mais nous conseillons vivement au lecteur d'essayer le deuxième pour se rendre compte qu'il ne fonctionne pas.

Exemple 14.133
Nous reprenons l'exemple 14.132 en déterminant cette fois une primitive de $x\cos(x)$:

$$\int x\cos(x)dx = x\sin(x) - \int \sin(x)dx = x\sin(x) + \cos(x) + C, \qquad C \in \mathbb{R}. \tag{14.509}$$

Nous retrouvons le résultat numérique de l'exemple précédent en ajoutant les extrêmes d'intégration

$$\int_0^\pi x\cos(x)dx = \big[x\sin(x) + \cos(x)\big]_0^\pi = -2. \tag{14.510}$$

△

Remarque 14.134.
Lorsqu'on calcule des intégrales, il est bon de passer par la primitive (c'est à dire en suivant l'exemple 14.133 et non 14.132) parce qu'il est alors facile de vérifier le résultat en calculant la dérivée de la primitive trouvée.

Par exemple pour vérifier si (14.509) est correct, il suffit de dériver $x\sin(x) + \cos(x)$:

$$\big(x\sin(x) + \cos(x)\big)' = \sin(x) + x\cos(x) - \sin(x) = x\cos(x). \tag{14.511}$$

La fonction $x\sin(x) + \cos(x)$ est donc bien une primitive de $x\cos(x)$.

14.22.2 Changement de variables – pour trouver des primitives

De la même manière que l'utilisation «à l'envers» de la formule de dérivation du produit avait donné la méthode d'intégration par parties, nous allons voir que que l'utilisation «à l'envers» de la formule de dérivation d'une fonction composée donne lieu à la méthode d'intégration par changement de variables.

Proposition 14.135.
Soit I et J des intervalles de $\mathbb{R}$ et une fonction $u\colon I \to J$ qui est dérivable de dérivée continue. Soit $f\colon J \to \mathbb{R}$ une fonction admettant une primitive F. Alors la fonction

$$x \mapsto F\big(u(x)\big) \tag{14.512}$$

est une primitive de

$$f\big(u(x)\big)u'(x). \tag{14.513}$$

Démonstration. Cela est une utilisation immédiate de la formule de dérivée des fonctions composées. □

Exemple 14.136
Soit à calculer

$$\int x\sqrt{1-x^2}dx. \tag{14.514}$$

La fonction $g(x) = x\sqrt{1-x^2}$ est le produit de x et de $\sqrt{1-x^2}$. On remarque que la dérivée de $1-x^2$ est $-2x$: nous avons alors, à un facteur -2 près, une expression de la forme (14.513) où la racine carré joue le rôle de f, $f(t) = \sqrt{t}$, et $1-x^2$ le rôle de u. Une primitive de la fonction $f(t) = \sqrt{t}$ est $F(t) = 2t^{3/2}/3$.

Donc la fonction $\frac{2u(x)^{3/2}}{3} = \frac{2}{3}(1-x^2)^{3/2}$ est primitive de $-2x\sqrt{1-x^2} = -2g(x)$. Autrement dit,

$$\int -2x\sqrt{1-x^2}\,dx = \frac{2(1-x^2)^{3/2}}{3} + C, \tag{14.515}$$

et en divisant par -2 nous trouvons la primitive demandée :

$$\int x\sqrt{1-x^2}\,dx = -\frac{(1-x^2)^{3/2}}{3} + C. \tag{14.516}$$

△

L'exemple suivant donne une façon plus économe de retenir la méthode du changement de variables.

Exemple 14.137
Soit à calculer

$$\int \cos(x)e^{\sin(x)}dx. \tag{14.517}$$

Vu qu'il y a beaucoup de fonctions trigonométriques dans la fonction à intégrer, nous allons poser $u(x) = \sin(x)$, et remplacer élément par élément tout ce qui contient du «x» dans l'intégrale demandée par la quantité correspondante en termes de u.

La difficulté est de savoir ce que nous allons faire du «dx» dans l'intégrale. Ce dx marque une variation (infinitésimale) de x. La formule des accroissements finis dit que si x augmente de la valeur dx, alors $u(x)$ augmente de $u'(x)dx$, c'est à dire que

$$du = \cos(x)dx. \tag{14.518}$$

Nous avons donc les substitutions suivantes à faire :

$$\sin(x) = u \tag{14.519a}$$
$$du = \cos(x)dx \tag{14.519b}$$
$$dx = \frac{du}{\cos(x)}. \tag{14.519c}$$

La chose «magique» est que le $\cos(x)$ se trouvant dans la fonction se simplifie avec le cosinus qui arrive lorsqu'on remplace dx par $\frac{du}{\cos(x)}$. Les substitutions faites nous restons avec

$$\int \cos(x)e^{\sin(x)}dx = \int e^u du = e^u + C, \qquad \text{où } u = \sin(x). \tag{14.520}$$

Attention : la réponse doit être impérativement donnée en termes de x et non de u. Nous écrivons donc

$$\int \cos(x)e^{\sin(x)} = e^{\sin(x)} + C. \tag{14.521}$$

△

14.22.3 Changement de variables – pour calculer des intégrales

Le corollaire 13.208 fixe la relation entre la recherche des primitives de f et la calcul de l'intégrale de f sur l'intervalle d'extrêmes a et b. On a vu dans la section précédente comment utiliser le changement de variable pour trouver une primitive de f. Il faut maintenant comprendre comment appliquer ce qu'on a vu dans le calcul d'une intégrale.

En effet nous avons le choix entre
- trouver une primitive de f comme dans la section précédente et appliquer ensuite la formule du corollaire 13.208 ;
- écrire une intégrale pour la nouvelle variable $u = u(x)$ sur l'intervalle entre $u(a)$ et $u(b)$.

Nous allons voir ce deux méthodes dans des exemples.

Exemple 14.138
Soit à calculer

$$\int_{1/3}^{1/2} x\sqrt{1-x^2}dx. \tag{14.522}$$

Les primitives $\int x\sqrt{1-x^2}dx$ ont été trouvé dans l'exemple 14.137. Une primitive est

$$F(x) = \int x\sqrt{1-x^2}dx = -\frac{(1-x^2)^{3/2}}{3}. \tag{14.523}$$

Nous pouvons maintenant calculer l'intégrale de $x\sqrt{1-x^2}$ sur l'intervalle $[1/3, 1/2]$ par la définition

$$\int_{1/3}^{1/2} x\sqrt{1-x^2}dx = F\left(\frac{1}{2}\right) - F\left(\frac{1}{3}\right) = -\frac{\sqrt{3}}{8} + \frac{16\sqrt{2}}{81}. \tag{14.524}$$

△

Remarque 14.139.
Pour que le calcul d'intégrale donne quelque chose de sensé il faut absolument que la primitive soit écrite en tant que fonction de x et non comme fonction de u. La méthode que nous allons voir dans l'exemple suivant réduit grandement la probabilité d'oublier ce détail, d'où le fait qu'elle soit de loin la plus utilisée.

Exemple 14.140
Calculons à nouveau

$$\int_{1/3}^{1/2} x\sqrt{1-x^2}dx. \tag{14.525}$$

Cette fois nous allons toucher à l'intervalle d'intégration en même temps que faire le changement de variables. Nous savons déjà les substitutions

$$\begin{cases} u = 1 - x^2 & (14.526a) \\ du = -2xdx & (14.526b) \\ dx = \dfrac{du}{-2x}. & (14.526c) \end{cases}$$

En ce qui concerne les extrêmes d'intégration, si $x = 1/3$ alors $u = 1 - \frac{1}{9} = \frac{8}{9}$ et si $x = \frac{1}{2}$ alors $u = \frac{3}{4}$. Nous avons donc encore les substitutions suivantes :

$$\begin{cases} x = 1/3 \to u = 8/9 & (14.527a) \\ x = 1/2 \to u = 3/4 & (14.527b) \end{cases}$$

Le calcul est alors

$$\int_{1/3}^{1/2} x\sqrt{1-x^2}dx = -\frac{1}{2}\int_{8/9}^{3/4} \sqrt{u}du = -\frac{1}{2}\left[\frac{u^{3/2}}{3/2}\right]_{8/9}^{3/4} = -\frac{\sqrt{3}}{8} + \frac{16\sqrt{2}}{81}. \tag{14.528}$$

Attention : la dernière égalité n'est pas immédiate ; elle demande quelque calculs et une bonne utilisation des règles de puissances. △

La deuxième méthode est plus utilisée et, avec un peu d'exercice, plus rapide à mettre en place que la première.

Jusqu'à présent nous avons utilisé des changement de variables dans lesquels nous exprimions u en termes de x. Comme le montre l'exemple suivant, il est parfois fructueux d'utiliser le changement de variable dans le sens inverse : avec x exprimé en termes d'un paramètre.

Exemple 14.141
À calculer :

$$\int_{1/2}^{\sqrt{3}/2} \sqrt{1-x^2}dx. \tag{14.529}$$

Nous posons $x = \sin(\theta)$ parce que nous savons que $1 - \sin^2(\theta) = \cos^2(\theta)$; nous espérons que le changement de variables simplifie l'expression [15]. Les substitutions à faire dans l'intégrale sont :

$$\begin{cases} x = \sin(\theta) & (14.530a) \\ dx = \cos(\theta)d\theta, & (14.530b) \end{cases}$$

15. Lorsqu'on fait un changement de variables, il s'agit toujours d'*espérer* que l'expression se simplifie. Il n'y a pas moyen de savoir a priori si tel changement de variable va être utile. Il faut essayer.

et en ce qui concerne les bornes, si $x = 1/2$ alors $\sin(\theta) = \frac{1}{2}$, c'est à dire $\theta = \frac{\pi}{6}$. Si $x = \sqrt{3}/2$ alors $\theta = \frac{\pi}{3}$. Donc

$$\int_{1/2}^{\sqrt{3}/2} \sqrt{1-x^2}dx = \int_{\pi/6}^{\pi/3} \sqrt{1-\sin^2(\theta)}\cos(\theta)dt. \tag{14.531}$$

Nous avons $1-\sin^2(\theta) = \cos^2(\theta)$ et vu que $\theta \in [\frac{\pi}{6}, \frac{\pi}{3}]$ nous avons toujours $\cos(\theta) > 0$, ce qui donne $\sqrt{\cos^2(\theta)} = \cos(\theta)$. Nous devons donc calculer

$$\int_{\pi/6}^{\pi/3} \cos^2(\theta)d\theta. \tag{14.532}$$

Pour celle-là, il faut utiliser une formule de trigonométrie [16] :

$$\cos^2(\theta) = \frac{1+\cos(2\theta)}{2}. \tag{14.533}$$

Donc

$$\int_{\pi/6}^{\pi/3} \cos^2(\theta)d\theta = \int_{\pi/6}^{\pi/3} \frac{1+\cos(2\theta)}{2}d\theta = \left[\frac{\theta}{2}\right]_{\pi/6}^{\pi/3} + \int_{\pi/6}^{\pi/3} \frac{\cos(2\theta)}{2}d\theta, \tag{14.534}$$

Pour calculer proprement la dernière intégrale nous effectuons un autre changement de variable (facile) en posant $t = 2\theta$, $dt = 2d\theta$, $t(\pi/6) = \pi/3$ et $t(\pi/3) = 2\pi/3$, nous avons alors

$$\int_{\pi/6}^{\pi/3} \cos^2(\theta)d\theta = \left[\frac{\theta}{2}\right]_{\pi/6}^{\pi/3} + \int_{\pi/3}^{2\pi/3} \frac{\cos(t)}{4}dt = \left[\frac{\theta}{2}\right]_{\pi/6}^{\pi/3} = \frac{\pi}{6} - \frac{\pi}{12} = \frac{\pi}{12}, \tag{14.535}$$

parce que $\sin\left(\frac{2\pi}{3}\right) = \sin\left(\frac{\pi}{3}\right)$. Au final,

$$\int_{1/2}^{\sqrt{3}/2} \sqrt{1-x^2}dx = \frac{\pi}{12}. \tag{14.536}$$

△

14.22.4 Intégrations des fractions rationnelles réduites

Définition 14.142.
*Une **fraction rationnelle** est un quotient de deux polynômes à coefficients réels ou complexes.*

Par exemple

$$\frac{x^5 + 7x^4 - \frac{x^3}{2} + x}{x^2 - 1} \tag{14.537}$$

est une fraction rationnelle.

Il sera expliqué dans le cours d'algèbre que toute fraction rationnelle peut être écrite sous forme d'une somme d'éléments simples, c'est à dire de fractions rationnelles d'un des deux types suivants :

$$\frac{\alpha}{(x-a)^m}, \qquad \alpha, a \in \mathbb{R}, m \in \mathbb{N} \tag{14.538a}$$

$$\frac{\alpha x + \beta}{(x^2+ax+b)^m}; \qquad \alpha, \beta, a, b \in \mathbb{R}, m \in \mathbb{N}, a^2 - 4b < 0. \tag{14.538b}$$

Nous allons nous contenter de donner un exemple de chaque type.

16. En fait, il y a moyen de terminer le calcul en intégrant deux fois par parties, mais c'est plus compliqué.

(1) En ce qui concerne le cas (14.538a) avec $m = 1$, nous avons par exemple

$$\int \frac{1}{x-3}dx = \ln\big(|x-3|\big) + C. \tag{14.539}$$

Si vous voulez en être tout à fait sûr, effectuez d'abord le changement de variables $u = x-3$ qui donne $dx = du$.

(2) En ce qui concerne le cas (14.538a) avec $m \neq 1$, nous avons par exemple

$$\int \frac{1}{(x-1)^4}dx = -\frac{1}{3(x-1)^3} + C. \tag{14.540}$$

Encore une fois, pour s'en convaincre, utiliser le changement de variables $u = x-1$, $dx = du$:

$$\int \frac{1}{(x-1)^4}dx = \int \frac{1}{u^4}du = \int u^{-4}du = -\frac{u^{-3}}{3} + C = -\frac{1}{3}\frac{1}{(x-1)^3} + C. \tag{14.541}$$

(3) En ce qui concerne le cas (14.538b) avec $\alpha \neq 0$, nous avons par exemple

$$\int \frac{x}{x^2+4}dx = \frac{1}{2}\ln(x^2+4) + C. \tag{14.542}$$

Pour ce faire, il faut faire le changement de variables $u = x^2+4$, $du = 2xdx$, $dx = \frac{du}{2x}$ qui donne

$$\int \frac{x}{x^2+4}dx = \frac{1}{2}\int \frac{du}{u} = \frac{1}{2}\ln(|u|) + C = \frac{1}{2}\ln(|x^2+4|) + C. \tag{14.543}$$

Dans ce cas nous pouvons oublier d'écrire la valeur absolue dans le logarithme parce que de toutes façons, x^2+4 est toujours positif.

(4) En ce qui concerne le cas (14.538b) avec $\alpha = 0$, nous avons par exemple

$$\int \frac{dx}{x^2+4} = \frac{1}{4}\int \frac{dx}{(\frac{x}{2})^2+1} = \frac{1}{2}\arctan(\frac{x}{2}) + C. \tag{14.544}$$

où nous avons utilisé la primitive $\int \frac{dx}{x^2+1}dx = \arctan(x)$ du tableau de la page 867. Pour vous en convaincre vous pouvez faire la dernière étape avec le changement de variables $u = x/2$, $dx = 2du$.

14.22.5 Quelques formules à connaître

À retenir 14.143

$$\int \big(\alpha f(x) + \beta g(x)\big)\, dx = \alpha \int f(x)\, dx + \beta \int g(x)\, dx. \tag{14.545a}$$

$$\int f(x)g'(x)\, dx = f(x)g(x) - \int f'(x)g(x)\, dx. \tag{14.545b}$$

$$\int f'(u(x))u'(x)\, dx = \int f(t)\, dt, \qquad \text{avec } t = u(x). \tag{14.545c}$$

$$\int \frac{f'(x)}{f(x)}\, dx = \log|f(x)| + C, \qquad \text{c'est un cas particulier de la formule précédente.} \tag{14.545d}$$

14.23 Trucs et astuces de calcul d'intégrales

Afin d'alléger le texte de calculs parfois un peu longs, nous regroupons ici les intégrales à une variable que nous devons utiliser dans les autres parties du cours.

14.23.1 Quelques intégrales «usuelles»

(1) L'intégrale

$$\boxed{I = \int x\ln(x)dx = \frac{x^2}{2}\big(\ln(x) - \frac{1}{2}\big)} \tag{14.546}$$

se fait par partie en posant

$$\begin{aligned} u &= \ln(x), \quad dv = x\,dx \\ du &= \frac{1}{x}\,dx, \qquad v = \frac{x^2}{2}, \end{aligned} \tag{14.547}$$

et ensuite

$$I = \ln(x)\frac{x^2}{2} - \int \frac{x}{2} = \frac{x^2}{2}\big(\ln(x) - \frac{1}{2}\big). \tag{14.548}$$

(2) L'intégrale

$$\boxed{I = \int x\ln(x^2)dx = x^2\ln(x) - \frac{x^2}{2}.} \tag{14.549}$$

En utilisant le fait que $\ln(u^2) = 2\ln(u)$, nous retombons sur une intégrale du type (1) :

$$I = x^2\ln(x) - \frac{x^2}{2}. \tag{14.550}$$

(3) L'intégrale

$$\boxed{I = \int x\ln(1+x^2)dx = \frac{1}{2}\ln(x^2+1)(x^2+1) - x^2 - \frac{1}{2}} \tag{14.551}$$

se traite en posant $v = 1 + x^2$ de telle sorte à avoir $dx = \frac{dv}{2x}$ et donc

$$I = \frac{1}{2}\ln(x^2+1)(x^2+1) - x^2 - \frac{1}{2}. \tag{14.552}$$

(4) L'intégrale

$$I = \int \cos(\theta)\sin(\theta)\ln\left(1 + \frac{1}{\cos^2(\theta)}\right)d\theta \tag{14.553}$$

demande le changement de variable $u = \cos(\theta)$, $d\theta = -\frac{du}{\sin(\theta)}$. Nous tombons sur l'intégrale

$$I = -\int u\ln\left(\frac{1+u^2}{u^2}\right) = -\int u\ln(1+u^2) + \int u\ln(u^2), \tag{14.554}$$

qui sont deux intégrales déjà faites. Nous trouvons

$$I = -\frac{1}{2}\ln\left(\frac{\sin^2(\theta)-1}{\sin^2(\theta)-2}\right)\sin^2(\theta) - \ln\big(\sin^2(\theta)-2\big) + \frac{1}{2}\ln\big(\sin^2(\theta)-1\big) \tag{14.555}$$

(5) L'intégrale

$$\boxed{\int \frac{r^3}{1+r^2}dr = \frac{r^2}{2} - \frac{1}{2}\ln(r^2+1).} \tag{14.556}$$

commence par faire la division euclidienne de r^3 par $r^2 + 1$; ce que nous trouvons est $r^3 = (r^2+1)r - r$. Il reste à intégrer

$$\int \frac{r^3}{1+r^2}dr = \int r\,dr - \int \frac{r}{1+r^2}dr. \tag{14.557}$$

La fonction dans la seconde intégrale est $\frac{r}{1+r^2} = \frac{1}{2}\frac{f'(r)}{f(r)}$ où $f(r) = 1 + r^2$, et donc $\int \frac{r}{1+r^2} = \frac{1}{2}\ln(1+r^2)$. Au final,

$$I = \frac{1}{2}r^2 - \frac{1}{2}\ln(r^2+1). \tag{14.558}$$

(6) L'intégrale

$$\boxed{I = \int \cos(\theta)\sin(\theta)d\theta = \frac{\sin^2(\theta)}{2}} \tag{14.559}$$

se traite par le changement de variable $u = \sin(\theta)$, $du = \cos(\theta)d\theta$, et donc

$$\int \cos(\theta)\sin(\theta)d\theta = \int udu = \frac{u^2}{2} = \frac{\sin^2(\theta)}{2}. \tag{14.560}$$

(7) L'intégrale

$$\boxed{\int \sqrt{1+x^2}dx = \frac{x}{2}\sqrt{1+x^2} + \frac{1}{2}\operatorname{arcsinh}(x)} \tag{14.561}$$

s'obtient en effectuant le changement de variable $u = \sinh(\xi)$.

(8) L'intégrale

$$\boxed{\int \cos^2(x)\sin^2(x)dx = \frac{x}{8} - \frac{\sin(4x)}{32}} \tag{14.562}$$

s'obtient à coups de formules de trigonométrie. D'abord, $\sin(t)\cos(t) = \frac{1}{2}\sin^2(2t)$ fait en sorte que la fonction à intégrer devient

$$f(x) = \frac{1}{4}\sin^2(x). \tag{14.563}$$

Ensuite nous utilisons le fait que $\sin^2(t) = (1-\cos(2t))/2$ pour transformer la formule à intégrer en

$$f(x) = \frac{1-\cos(4x)}{8}. \tag{14.564}$$

Cela s'intègre facilement en posant $u = 4x$, et le résultat est

$$\int f(x)dx = \frac{x}{8} - \frac{\sin(4x)}{32}. \tag{14.565}$$

(9) La fonction

$$\operatorname{sinc}(x) = \frac{\sin(x)}{x} \tag{14.566}$$

est le **sinus cardinal** de x. Nous allons montrer que

$$\boxed{\int_0^\infty \big|\operatorname{sinc}(x)\big|dx = \infty}. \tag{14.567}$$

D'abord nous avons

$$\int_{(n-1)\pi}^{n\pi} \frac{\big|\sin(t)\big|}{t}dt \geqslant \int_{(n-1)\pi}^{n\pi} \frac{\big|\sin(t)\big|}{n\pi}dt, \tag{14.568}$$

mais par périodicité,

$$\int_{(n-1)\pi}^{n\pi} \big|\sin(t)\big|dt = \int_0^\pi \sin(t)dt = 2. \tag{14.569}$$

Par conséquent

$$\int_0^{n\pi} \big|\operatorname{sinc}(t)\big|dt \geqslant \frac{2}{\pi}\sum_{k=1}^{n}\frac{1}{k}, \tag{14.570}$$

ce qui diverge lorsque $n \to \infty$.

(10) Les intégrales, pour $\epsilon > 0$,

$$\boxed{\int_0^\infty \cos(kx)e^{-\epsilon x}dx = \frac{\epsilon}{k^2+\epsilon^2}} \tag{14.571}$$

et

$$\boxed{\int_0^\infty \sin(kx)e^{-\epsilon x}dx = \frac{k}{k^2+\epsilon^2}} \tag{14.572}$$

se calculent deux fois par partie. Nous posons

$$I = \int_0^\infty \cos(kx)e^{-\epsilon x}dx \tag{14.573a}$$

$$J = \int_0^\infty \sin(kx)e^{-\epsilon x}dx. \tag{14.573b}$$

L'intégrale I s'effectue par partie en posant $u = \cos(kx)$ et $v' = e^{-\epsilon x}$. Un peu de calcul montre que

$$I = \frac{1}{\epsilon} - \frac{k}{\epsilon}J. \tag{14.574}$$

Par ailleurs l'intégrale J se fait également par partie pour obtenir

$$J = \frac{k}{\epsilon}I. \tag{14.575}$$

En résolvant pour I et J les deux équations déduites, nous trouvons

$$I = \frac{\epsilon}{k^2+\epsilon^2} \tag{14.576a}$$

$$J = \frac{k}{k^2+\epsilon^2}. \tag{14.576b}$$

14.23.2 Reformer un carré au dénominateur

Lorsqu'on a un second degré au dénominateur, le bon plan est de reformer un carré parfait. Par exemple :

$$x^2 + 2x + 2 = (x+1)^2 + 1. \tag{14.577}$$

Ensuite, le changement de variable $t = x + 1$ est pratique parce que cela donne $t^2 + 1$ au dénominateur.

Cherchons

$$I = \int \frac{1-x}{x^2+2x+2}dx = \int \frac{1-x}{(x+1)^2+1}dx = \int \frac{1-(t-1)}{t^2+1} \tag{14.578}$$

où nous avons fait le changement de variable $t = x + 1$, $dt = dx$. L'intégrale se coupe maintenant en deux parties :

$$I = \int \frac{-t}{t^2+1} + \int \frac{2}{t^2+1}. \tag{14.579}$$

La seconde est dans les formulaires et vaut

$$2\arctan(t) = 2\arctan(x+1), \tag{14.580}$$

tandis que la première est presque de la forme f'/f :

$$\int \frac{t}{t^2+1} = \frac{1}{2}\int \frac{2t}{t^2+1} = \frac{1}{2}\ln(t^1+1) = \frac{1}{2}\ln(u^2+2u+2). \tag{14.581}$$

14.23.3 Décomposer en fractions simples

Il y a la méthode de Rothstein-Trager décrite 14.12.4 qui permet de l'éviter lorsqu'on a de la chance.

Il est possible de décomposer une fraction rationnelle en fractions dites «simples». Si $|z| < 1$ nous avons par exemple la décomposition

$$\frac{1}{1-z^r} = \sum_{\omega \in U_r} \frac{A_\alpha}{\omega - z} \tag{14.582}$$

où U_r est le groupe des racines r^{e}de l'unité définit en (3.166). Les nombres A_ω peuvent alors être déterminés en effectuant la somme. Le dénominateur commun sera $1 - z^2$ tandis que les A_ω sont déterminés en égalant le numérateur à 1.

Exemple 14.144
Pour décomposer la fraction $\frac{1}{1-x^2}$ nous savons que les racines sont ± 1. Donc nous écrivons

$$\frac{1}{1-x^2} = \frac{A}{1-x} + \frac{B}{1+x}. \tag{14.583}$$

Nous trouvons les valeurs de A et B en effectuant la somme :

$$\frac{A(1+x) + B(1-x)}{1-x^2} = \frac{A + B + (A-B)x}{1-x^2}. \tag{14.584}$$

Les coefficients A et B doivent donc vérifier $A + B = 1$ et $A - B = 0$. Au final,

$$\frac{1}{1-x^2} = \frac{1}{2(1-x)} + \frac{1}{2(1+x)}. \tag{14.585}$$

△

14.24 Constructions plus naïves de l'intégrale dans le cas réel

Les sections 13.2 et 13.9 ont donné une construction très complète de la mesure de Lebesgue, et nous avons définit la théorie de l'intégration sur un espace mesuré quelconque dans la définition 13.113.

Dans cette section nous allons donner différentes choses plus rapides qui servent souvent de définition dans les cours moins avancés.

14.24.1 Mesure de Lebesgue, version rapide

Nous construisons à présent la mesure de Lebesgue sur $\mathbb{R}^n$. Un **pavé** dans $\mathbb{R}^n$ est un ensemble de la forme

$$B = \prod_{i=1}^{n} [a_i, b_i]; \tag{14.586}$$

le volume d'un tel pavé est défini par $\text{Vol}(B) = \prod_i (b_i - a_i)$. Soit maintenant $A \subset \mathbb{R}^n$. La **mesure externe** de A est le nombre

$$m^*(A) = \inf\{ \sum_{B \in \mathcal{F}} \text{Vol}(B) \text{ où } \mathcal{F} \text{ est un ensemble dénombrable de pavés dont l'union recouvre } A.\} \tag{14.587}$$

Définition 14.145.
*Nous disons que A est **mesurable** au sens de Lebesgue si pour tout ensemble $S \subset \mathbb{R}^n$ nous avons l'égalité*

$$m^*(S) = m^*(A \cap S) + m^*(S \backslash A). \tag{14.588}$$

Dans ce cas nous disons que la mesure de Lebesgue de A est $m(A) = m^*(A)$.

Proposition 14.146.
Deux fonctions continue égales presque partout pour la mesure de Lebesgue[17] *sont égales.*

Démonstration. Soient f et g deux fonctions continues telles que $f(x) = g(x)$ pour presque tout $x \in D$. La fonction $h = f - g$ est alors presque partout nulle et nous devons prouver qu'elle est nulle sur tout D. La fonction h est continue ; si $h(a) \neq 0$ pour un certain $a \in D$ alors h est non nulle sur un ouvert autour de a par continuité et donc est non nulle sur un ensemble de mesure non nulle. □

14.24.2 Pavés et subdivisions

Définition 14.147.
Nous appelons ***pavé*** *de $\mathbb{R}^p$ toute partie de $\mathbb{R}^p$ obtenue comme produit de p intervalles de $\mathbb{R}$. Plus explicitement, une partie R est un pavé de $\mathbb{R}^p$ s'il s'écrit sous la forme*

$$R = \left\{(x_1, \ldots, x_p) \in \mathbb{R}^p \,\middle|\, x_i \in \mathcal{I}_i, i = 1, \ldots, p\right\},$$

où $\mathcal{I}_i$ est un intervalle de $\mathbb{R}$ pour tout $i = 1, \ldots, p$.

On appelle pavé fermé de $\mathbb{R}^p$ le produit de p intervalles fermés

$$R = \prod_{i=1}^{p} [a_i, b_i].$$

On définit de même le pavé ouvert

$$S = \prod_{i=1}^{p}]a_i, b_i[.$$

Un pavé $R = \prod_{i=1}^{p} \mathcal{I}_i$ est dit borné si tous les intervalles $\mathcal{I}_i$ sont bornés dans $\mathbb{R}$. Les pavés non bornés sont des produits d'intervalles où un (ou plusieurs) des intervalles n'est pas borné. Par exemple,

$$N =]-\infty, 5] \times [0, 13].$$

L'espace $\mathbb{R}^p$, lui-même, est un pavé de $\mathbb{R}^p$.

Définition 14.148.
Une partie A de $\mathbb{R}^p$ est dite ***pavable*** *s'il existe une famille finie de pavés bornés R_j, $j = 1, \ldots, n$, et deux à deux disjoints tels que*

$$A = \bigcup_{j=1}^{n} R_j.$$

Un exemple d'ensemble pavable dans $\mathbb{R}^2$ est donné à la figure 14.18. Il existe beaucoup d'ensembles dans $\mathbb{R}^2$ qui ne sont pas pavables, par exemple les ellipses.

Le complémentaire d'un pavé est un ensemble pavable et, en particulier, tout complémentaire d'un pavé borné est une réunion de pavés non bornés. Toute union finie et toute intersection d'ensemble pavables est pavable.

Définition 14.149.
Soit R un pavé borné de $\mathbb{R}^p$, pour fixer les idées on peut penser $R = \prod_{i=1}^{p} [a_i, b_i]$. On appelle ***longueur*** *de l'i-ème arrête de R le nombre $b_i - a_i$. La* ***mesure p-dimensionnelle de*** *R, $m(R)$, est le produit des longueurs*

$$m(R) = \prod_{i=1}^{p} (b_i - a_i).$$

17. Définition 14.145.

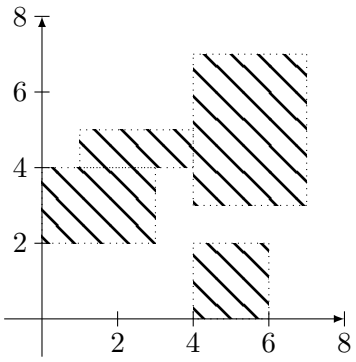


FIGURE 14.18 – Un ensemble pavable.

Exemple 14.150
Dans $\mathbb{R}^3$, l'ensemble $R = [-1, 1] \times [3, 4] \times [0, 2]$ est un pavé fermé de mesure

$$m(R) = (1 + 1) \cdot (4 - 3) \cdot (2 - 0) = 4.$$

Il s'agit du volume usuel du parallélépipède rectangle. △

Exemple 14.151
L'ensemble $R =]-1, 1[\times [3, 4] \times [0, 2]$ est un pavé de $\mathbb{R}^3$. Il n'est ni fermé ni ouvert, sa mesure est encore 4. △

Si R est un pavé non borné on peut encore définir sa mesure. La notion de mesure se généralise en deux étapes. D'abord on dit que la longueur d'une arête non bornée est ∞. Ensuite, on adopte la convention $0 \cdot \infty = 0$. Il faut remarquer que avec cette généralisation tout point et toute droite dans $\mathbb{R}^2$ ont mesure nulle.

Afin de définir les intégrales, nous allons intensivement faire appel à la notion de subdivision d'intervalles, voir définition 19.3 et la discussion qui suit.

Lorsqu'on considère un pavé borné $R = \prod_{i=1}^{p} \mathcal{I}_i$ de $\mathbb{R}^p$, on note $\mathcal{S}_i$ l'ensemble des subdivisions de l'intervalle $\mathcal{I}_i$. La notion de subdivision de généralise au cas des pavés.

Définition 14.152.
Soir R un pavé fermé borné de $\mathbb{R}^p$, pour fixer les idées on peut penser à $R = \prod_{i=1}^{p}[a_i, b_i]$. On appelle ***subdivision*** *finie de R les éléments de l'ensemble $\mathcal{S} = \prod_{i=1}^{p} \mathcal{S}_i$,*

$$\mathcal{S} = \left\{ (Y_1, \dots, Y_p) \,\middle|\, Y_i = (y_{i,j})_{j=1}^{n_i} \in \mathcal{S}_i,\ i = 1, \dots, p \right\}.$$

On peut définir de même l'ensemble des subdivisions d'un pavé non borné.

Souvent, une subdivision d'un pavé $R = \prod_{i=1}^{p} \mathcal{I}_i$ sera noté $\sigma = (y_{i,j})_{j=1}^{n_i}$. Dans cette notation, on sous-entend que pour chaque i fixé, les nombres $y_{i,j}$ (il y en a n_i) forment une subdivision de l'intervalle $\mathcal{I}_i$. Afin de vous familiariser avec ces notations, repérez bien tous les éléments de la figure 14.19.

Définition 14.153.
Si σ est une subdivision d'un pavé R, un ***raffinement*** *de σ est une subdivision de R obtenue en fixant plus de points dans chaque intervalle.*

La subdivision σ de R détermine $n_1 n_2 \dots n_p$ pavés fermés de la forme

$$R_{(k_1,\dots,k_p)} = \{(x_1, \dots, x_p) \in \mathbb{R}^p \,|\, y_{i,k_{i-1}} \leqslant x_i \leqslant y_{i,k_i}\},$$

où k_i est dans $\{1, \dots, n_i\}$ et i dans $\{1, \dots, p\}$. On les appelles **cellules** de σ. On remarque que les cellules de σ sont toujours deux à deux disjointes (sauf au plus sur leurs bords).

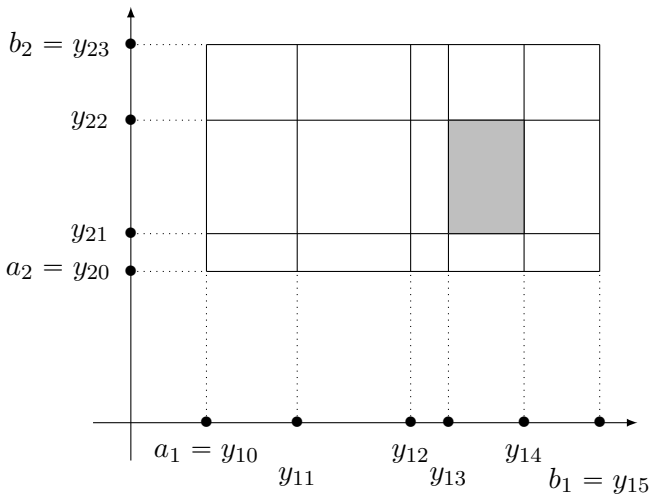


FIGURE 14.19 – Une cellule d'une subdivision d'un pavé de $\mathbb{R}^2$. La cellule grisée est $R_{(4,2)}$.

Lemme 14.154.
Soit R un pavé borné de $\mathbb{R}^p$ et soit $\sigma = (y_{i,j})_{j=1}^{n_i}$ une subdivision de R. On a

$$m(R) = \sum_{(k_1,\ldots,k_p)\in K} m(R_{(k_1,\ldots,k_p)}),$$

où $K = \{1,\ldots,n_1\} \times \{1,\ldots,n_2\} \times \ldots \times \{1,\ldots,n_p\}$.

Le lemme 14.154 suggère de définir la mesure d'un ensemble borné pavable $P = \cup_{j=1}^{n} R_j$ comme la somme des mesures des pavés disjoints R_j, $j = 1,\ldots,n$.

Définition 14.155.
*Une application $f : \mathbb{R}^p \to \mathbb{R}$ est dite **application en escalier** sur $\mathbb{R}^m$ si*
- *f est une application bornée,*
- *il existe une subdivision σ de $\mathbb{R}^p$ telle que la restriction de f est une application constante sur toute cellule R_k de σ*

$$f_{|_{R_k}} = C_k, \qquad C_k \in \mathbb{R},$$

*Une telle subdivision σ est dite **associée** à f.*

Exemple 14.156
La fonction f de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$f(x,y) = \begin{cases} 1 & \text{si } (x,y) \in [0,3] \times [-1,2], \\ 2 & \text{sinon.} \end{cases} \tag{14.589}$$

est une application en escalier. Exercice : donner une subdivision de $\mathbb{R}^2$ associée à cette fonction. △

Exemple 14.157
La fonction f de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$f(x,y) = \begin{cases} \frac{1}{m^2+n^2}, & \text{si } (x,y) \in [m,m+1] \times [n,n+1], \quad m,\, n \in \mathbb{N}_0, \\ 0, & \text{sinon} \end{cases} \tag{14.590}$$

est une application en escalier. Observez que, dans ce cas, il n'existe pas une subdivision finie de $\mathbb{R}^2$ associée à f. △

Remarque 14.158.
Si la subdivision σ est associée à f alors tout raffinement de σ (c'est à dire, toute subdivision obtenue en fixant plus de points dans chaque intervalle) a la même propriété.

Si f et g sont deux application en escalier sur R et σ_f et σ_g sont des subdivisions de R associées respectivement à f et g, alors on peut construire une troisième subdivision de R qui est associée à f et à g en même temps. Soient $\sigma_f = (Y_1, \dots, Y_p)$ et $\sigma_g = (Z_1, \dots, Z_p)$, où $Y_i = (y_{i,j})_{j=1}^{m_i}$ et $Z_i = (z_{i,j})_{j=1}^{n_i}$ sont des subdivision de l'intervalle $[a_i, b_i]$, pour $i = 1, \dots, p$. La subdivision de $[a_i, b_i]$ obtenue par l'union de Y_i et Z_i est encore une subdivision finie, qu'on appellera $\bar{Y}_i$. La subdivision $\bar{\sigma} = (\bar{Y}_1, \dots, \bar{Y}_p)$ de R est un raffinement de σ_f et de σ_g, donc elle est associée à la fois à f et à g.

Cela nous permet de prouver que si f et g sont des application en escalier, alors $f + g$, fg, $\min\{f, g\}$, $\max\{f, g\}$ et $|f|$ sont des applications en escalier.

14.24.3 Intégrale d'une fonction en escalier

Définition 14.159.
*Soit f une fonction de $\mathbb{R}^m$ dans $\mathbb{R}^n$. Le **support** de f est la fermeture de l'ensemble des points x tels que $f(x) \neq 0$.*

Définition 14.160.
*Une application en escalier f est dite **intégrable** si son support est compact.*

Soit f une application en escalier sur $\mathbb{R}^p$. Soit σ une subdivision de $\mathbb{R}^p$ associée à f et appelons R_k les cellules de σ, avec $k = (k_1, \dots, k_p)$ dans $K = \{1, \dots, n_1\} \times \{1, \dots, n_2\} \times \dots \times \{1, \dots, n_p\}$. Alors

$$f_{|_{R_k}} = C_k, \qquad C_k \in \mathbb{R}.$$

Définition 14.161.
*On définit l'**intégrale** de f sur $\mathbb{R}^p$ par*

$$\int_{\mathbb{R}^p} f \, dV = \sum_{k \in K} C_k m(R_k).$$

L'intégrale ainsi définie est un nombre réel. La proposition suivante nous dit que l'intégrale est «bien définie», au sens que sa valeur ne dépend pas de la subdivision associée à f qu'on utilise dans le calcul.

Proposition 14.162.
Soit f une application en escalier intégrable sur $\mathbb{R}^p$. Soient σ_1 et σ_2 deux subdivisions de $\mathbb{R}^p$ associées à f. L'intégrale de f ne dépend pas de la subdivision choisie.

On ne donne pas une preuve complète de cette proposition. En fait elle est une conséquence de la formule de réduction introduite dans la suite de ce chapitre.

14.24.4 Intégrales partielles

Soit f de $\mathbb{R}^p$ dans $\mathbb{R}$ une fonction continue, nulle hors du pavé borné R. Posons $R = \prod_{i=1}^{p} [a_i, b_i]$, pour fixer les idées. Pour chaque i dans $\{1, \dots, p\}$ fixé, on peut associer à f la fonction F_i de $p-1$ variables définie par

$$F_i(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p) = \int_{a_i}^{b_i} f(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_p) \, dy.$$

La fonction F_i est l'intégrale partielle de f par rapport à la i-ème variable. En particulier, si $f(x_1, \dots, x_p) = g(x_i)h(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p)$ on obtient

$$F_i = \int_{a_i}^{b_i} g(y)h(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p) \, dy = h \cdot \int_{a_i}^{b_i} g \, dy.$$

La fonction d'une seule variable qu'on obtient à partir de f en fixant $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_p$ et qui associe à x_i la valeur $f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_p)$, est appelée x_i-ème section de f en $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_p$.

Exemple 14.163
Soit f la fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$f(x,y) = \begin{cases} x+3y & \text{si } (x,y) \in [9,10] \times]\pi, 5] \\ 0 & \text{sinon.} \end{cases} \tag{14.591}$$

Les intégrales partielles de f sont

$$F_1(y) = \int_9^{10} x + 3y \, dx = \left[\frac{x^2}{2} + 3xy \right]_{x=9}^{x=10} = \frac{19}{2} + 3y,$$

$$F_2(x) = \int_\pi^5 x + 3y \, dy = \left[xy + \frac{3y^2}{2} \right]_{y=\pi}^{y=5} = x(5-\pi) + \frac{3}{2}(25 - \pi^2).$$

△

14.24.5 Réduction d'une intégrale multiple

Soit $R = [a,b] \times [c,d]$ un pavé fermé et borné de $\mathbb{R}^2$ et soit f une application en escalier intégrable sur $\mathbb{R}^2$ telle que le support de f soit contenu dans R. On considère la subdivision σ de R définie par les subdivisions

$$a = x_0 \leqslant x_1 \leqslant \ldots \leqslant x_m = b,$$
$$c = y_0 \leqslant y_1 \leqslant \ldots \leqslant y_n = d.$$

Les cellules de σ sont

$$R_{i,j} = [x_i, x_{i+1}] \times [y_j, y_{j+1}], \qquad i = 0, \ldots, m-1, \quad j = 0, \ldots, n-1.$$

La mesure de R est la somme des mesures des $R_{i,j}$

$$\begin{aligned} m(R) = & \sum_{(i,j) \in \{0,\ldots,m-1\} \times \{0,\ldots,n-1\}} m(R_{i,j}) = \\ & = \sum_{j=0}^{n-1} \sum_{i=0}^{m-1} (x_{i+1} - x_i) \cdot (y_{i+1} - y_i) = \\ & = \sum_{i=0}^{m-1} (x_{i+1} - x_i) \cdot \sum_{j=0}^{n-1} (y_{i+1} - y_i) = \\ & = (b-a) \cdot (d-c). \end{aligned} \tag{14.592}$$

Si f est constante sur chaque cellule de σ on peut écrire f de la forme suivante

$$f(x,y) = \sum_{j=0}^{n-1} \sum_{i=0}^{m-1} C_{i,j} \, \chi_{R_{i,j}}$$

où les $C_{i,j}$ sont des constantes réelles et $\chi_{R_{i,j}}$ est la **fonction caractéristique** de $R_{i,j}$

$$\chi_{R_{i,j}}(x,y) = \begin{cases} 1, & \text{si } (x,y) \in R_{i,j}, \\ 0, & \text{sinon.} \end{cases} \tag{14.593}$$

Comme (x,y) est dans $R_{i,j}$ si et seulement si $x \in [x_i, x_{i+1}]$ et $y \in [y_j, y_{j+1}]$, on vérifie que la fonction $\chi_{R_{i,j}}$ est égal au produit des fonctions caractéristiques des intervalles $[x_i, x_{i+1}]$ et $[y_j, y_{j+1}]$

$$\chi_{R_{i,j}}(x,y) = \chi_{[x_i, x_{i+1}]}(x) \cdot \chi_{[y_j, y_{j+1}]}(y).$$

On peut donc écrire la fonction f de la façon suivante

$$f(x,y) = \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} C_{i,j}\,\chi_{[x_i,x_{i+1}]}(x)\,\cdot\,\chi_{[y_j,y_{j+1}]}(y).$$

Comme on suppose que le support de f est une partie de R, l'intégrale de f sur $\mathbb{R}^2$ est

$$\int_{\mathbb{R}^2} f\,dV = \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} C_{i,j}\,m(R_{i,j}) = \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} C_{i,j}\,(x_{i+1}-x_i)\,\cdot\,(y_{j+1}-y_j). \tag{14.594}$$

Cette intégrale peut être réduite à la composition de deux intégrales partielles. Il suffit de remarquer que la valeur de l'intégrale de la fonction caractéristique d'un intervalle est la longueur de l'intervalle,

$$\begin{aligned} C_{i,j}(x_{i+1}-x_i)\,\cdot\,(y_{j+1}-y_j) &= \\ &= C_{i,j}\left(\int_{x_i}^{x_{i+1}} \chi_{[x_i,x_{i+1}]}(x)\,dx\right)\,\cdot\,\left(\int_{y_j}^{y_{j+1}} \chi_{[y_j,y_{j+1}]}(y)\,dy\right) = \\ &= C_{i,j}\left(\int_{a}^{b} \chi_{[x_i,x_{i+1}]}(x)\,dx\right)\,\cdot\,\left(\int_{c}^{d} \chi_{[y_j,y_{j+1}]}(y)\,dy\right), \end{aligned} \tag{14.595}$$

et utiliser les propriétés de linéarité de l'intégrale

$$\begin{aligned} \int_{\mathbb{R}^2} f\,dV &= \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} C_{i,j}\,\left(\int_{a}^{b} \chi_{[x_i,x_{i+1}]}(x)\,dx\right)\,\cdot\,\left(\int_{c}^{d} \chi_{[y_j,y_{j+1}]}(y)\,dy\right) = \\ &= \int_c^d\int_a^b \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} C_{i,j}\,\chi_{[x_i,x_{i+1}]}(x)\,\cdot\,\chi_{[y_j,y_{j+1}]}(y)\,dxdy = \\ &= \int_c^d\int_a^b f\,dxdy. \end{aligned} \tag{14.596}$$

De même on obtient

$$\begin{aligned} \int_{\mathbb{R}^2} f\,dV &= \int_a^b\int_c^d \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} C_{i,j}\,\chi_{[x_i,x_{i+1}]}(x)\,\cdot\,\chi_{[y_j,y_{j+1}]}(y)\,dxdy = \\ &= \int_a^b\int_c^d f\,dxdy. \end{aligned} \tag{14.597}$$

En général, on prouve la proposition suivante

Proposition 14.164.
Soit f une application en escalier intégrable sur $\mathbb{R}^p$ et soit R un pavé borné dans $\mathbb{R}^p$ qui contient le support de f. Comme d'habitude, pour fixer les idées nous écrivons $= \prod_{i=1}^p [a_i,b_i]$. Alors

$$\begin{aligned} \int_{\mathbb{R}^p} f(x_1,\ldots,x_p)\,dV &= \int_{a_p}^{b_p}\int_{a_{p-1}}^{b_{p-1}}\cdots\int_{a_1}^{b_1} f(x_1,\ldots,x_p)\,dx_1\cdots dx_p = \\ &= \int_{a_{s_p}}^{b_{s_p}}\int_{a_{s_{p-1}}}^{b_{s_{p-1}}}\cdots\int_{a_{s_1}}^{b_{s_1}} f(x_1,\ldots,x_p)\,dx_1\cdots dx_p, \end{aligned} \tag{14.598}$$

pour toute permutation $(s_1,\ldots,s_p)$ de l'ensemble $\{1,\ldots p\}$.

14.24.6 Propriétés de l'intégrale

Soient f et g deux fonctions en escalier intégrables de $\mathbb{R}^p$ dans $\mathbb{R}$, et soient a et b dans $\mathbb{R}$.

Linéarité de l'intégrale :

— Additivité : $f+g$ est intégrable et

$$\int_{\mathbb{R}^p}(f+g)\,dV=\int_{\mathbb{R}^p}f\,dV+\int_{\mathbb{R}^p}g\,dV,$$

— Homogénéité : λf est intégrable pour tout réel λ

$$\int_{\mathbb{R}^p}\lambda f\,dV=\lambda\int_{\mathbb{R}^p}f\,dV,$$

Monotonie Si $f\leqslant g$ alors

$$\int_{\mathbb{R}^p}f\,dV\leqslant\int_{\mathbb{R}^p}g\,dV,$$

Inégalité fondamentale

$$\left|\int_{\mathbb{R}^p}f\,dV\right|\leqslant\int_{\mathbb{R}^p}|f|\,dV.$$

Cette dernière inégalité s'obtient de la façon suivante :

$$\left|\int_{\mathbb{R}^p}f\,dV\right|=\left|\sum_{k\in K}C_k m(R_k)\right|\leqslant\sum_{k\in K}|C_k|m(R_k)=\int_{\mathbb{R}^p}|f|\,dV.$$

Inégalité de Čebičeff Si f est une application en escalier alors pour tout $a>0$ dans $\mathbb{R}$ l'ensemble $\{x\in\mathbb{R}^p\,:\,|f(x)|\geqslant a\}$ est pavable et borné, et l'inégalité suivante est satisfaite

$$m\left(\{x\in\mathbb{R}^p\,:\,|f(x)|\geqslant a\}\right)\leqslant\frac{1}{a}\int_{\mathbb{R}^p}|f|\,dV.$$

14.24.7 Intégrales multiples, cas général

Nous voulons généraliser la définition d'intégrale multiple au cas des domaines non pavables et de fonctions qui ne sont pas en escalier. Il y a plusieurs méthodes de le faire et ici on ne considère qu'une seule, introduite par Riemann.

Définition 14.165.
Soit $f:\mathbb{R}^p\to\mathbb{R}$ une fonction.

— *Pour toute application en escalier intégrable f_* telle que $f_*\leqslant f$, l'intégrale de f_* est dit une **somme inférieure** de f.*

— *Pour toute application en escalier intégrable f^* telle que $f_*\geqslant f$, l'intégrale de f^* est dit une **somme supérieure** de f.*

Soient $\sum_* f$ et $\sum^* f$ les ensembles des sommes inférieures et supérieures de f. Grâce à la propriété de monotonie de l'intégrale on sait que si a est dans $\sum_* f$ et b est dans $\sum^* f$ alors $a\leqslant b$.

Définition 14.166.
La fonction f est intégrable (au sens de Riemann) si $\sum_* f$ et $\sum^* f$ ne sont pas vides et

$$\inf\Sigma^* f=I=\sup\Sigma_* f.$$

Dans ce cas, la valeur I est appelée intégrale de f sur $\mathbb{R}^p$.

Remarque 14.167.
Toute fonction intégrable est bornée et à support compact. En effet, si le support de la fonction n'est pas compact alors soit $\sum_* f$ soit $\sum^* f$ doit être vide !

L'intégrale qu'on vient de définir possède toutes les propriétés de l'intégrale pour les fonctions en escalier. Le produit de deux fonctions intégrables est intégrable.

Il y a des cas où l'intégrabilité d'une fonction n'est pas évidente. Cependant, dans la plupart des exercices et des exemples de ce cours, nous nous aidons avec le critère suivant

Proposition 14.168.
Toute fonction continue à support compact est intégrable.

Cette proposition n'est a priori pas étonnante, vu qu'une fonction continue sur un support compact est bornée (théorème de Weierstrass 6.149).

14.24.8 Réduction d'une intégrale multiple

On n'utilise jamais la définition pour calculer la valeur d'une intégrale multiple. La méthode plus efficace, en pratique, est de réduire l'intégrale à la composition de plusieurs intégrales d'une variable.

Théorème 14.169 (de Fubini)**.**
Soit f une fonction intégrable de $\mathbb{R}^2$ dans $\mathbb{R}$. Si pour tout x dans $\mathbb{R}$ la section $f(x, \cdot\,)$ est intégrable par rapport à y, alors

$$\int_{\mathbb{R}^2} f(x,y)\,dV = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x,y)\,dx \right) dy.$$

De même, si pour tout y dans $\mathbb{R}$ la section $f(\,\cdot\,, y)$ est intégrable par rapport à x, alors

$$\int_{\mathbb{R}^2} f(x,y)\,dV = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x,y)\,dy \right) dx.$$

En général, on ne peut pas dire que les sections d'une fonction intégrable sont intégrables, donc il faut vraiment se souvenir des hypothèses du théorème 14.169. En dimension plus haute, on a le même résultat

Théorème 14.170.
Soit f une fonction intégrable de $\mathbb{R}^p$ dans $\mathbb{R}$. Si pour tout $(p-1)$-uple $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p)$ dans $\mathbb{R}^{p-1}$ la section $f(x_1, \dots, x_{i-1}, \cdot\,, x_{i+1}, \dots, x_p)$ est intégrable par rapport à x_i, alors

$$\int_{\mathbb{R}^p} f\,dV = \int_{\mathbb{R}} \left(\int_{\mathbb{R}^{p-1}} f\,dV \right) dx_i.$$

Si f est une fonction positive et intégrable de $\mathbb{R}^2$ dans $\mathbb{R}$ on peut interpréter l'intégrale de f comme le volume du solide au-dessous du graphe de f. Avec cette interprétation, l'intégrale partielle par rapport à x pour $y = y_0$ fixé est l'aire de la tranche qu'on obtient en coupant le solide par le plan $y = y_0$.

Exemple 14.171
Le premier exemple à faire est celui d'une fonction en escalier intégrable et positive. Soit $f\colon \mathbb{R}^2 \to \mathbb{R}$ la fonction

$$f(x,y) = \begin{cases} 1 & \text{si } (x,y) \in R_1 =]-1,3] \times [4,5] \\ 3 & \text{si } (x,y) \in R_2 =]13,15[\times [0,2[\\ 0 & \text{dans les autres cas.} \end{cases} \tag{14.599}$$

L'intégrale de f sur $\mathbb{R}^2$ est $1 \cdot m(R_1) + 3 \cdot m(R_2) = 16$. On voit tout de suite qu'il s'agit de la somme du volume des deux parallélépipèdes de hauteurs respectives 1 et 3 et bases R_1 et R_2. △

Exemple 14.172
On veut calculer le volume du solide S, borné par le paraboloïde elliptique $x^2 + 2y^2 + z = 16$ et le plans $x = 2$, $x = 0$, $y = 2$ $y = 0$, $z = 0$. On observe que la portion de paraboloïde elliptique qui nous intéresse est le graphe de la fonction $f(x,y) = 16 - x^2 - 2y^2$ pour (x,y) dans $R = [0,2] \times [0,2]$. La fonction f est continue ainsi que ses sections, donc on peut appliquer le théorème 14.169 et

décomposer l'intégrale double en deux intégrales simples :

$$\begin{aligned}\int_R 16 - x^2 - 2y^2\, dV &= \int_0^2 \int_0^2 f(x,y)\, dxdy = \\ &= \int_0^2 \left[(16 - 2y^2)x - \frac{x^3}{3}\right]_{x=0}^{x=2} dy = \\ &= \left[\left(32 - \frac{8}{3}\right) y - \frac{4y^3}{3}\right]_{x=0}^{x=2} = 64 - \frac{16+32}{3} = 48.\end{aligned} \tag{14.600}$$

Vérifiez, comme exercice, qu'on obtient le même résultat en intégrant d'abord par rapport à y et puis par rapport à x. △

Exemple 14.173
Dans les hypothèses du théorème 14.169 l'ordre des intégrations partielles ne change pas la valeur de l'intégrale. En fait, si les calculs sont faites par des êtres humains l'ordre d'intégration peut faire une certaine différence comme dans cet exemple. On veut évaluer la valeur de l'intégrale

$$\int_{\mathbb{R}^2} f(x,y)\, dV$$

où

$$f(x,y) = \begin{cases} y\sin(x,y) & \text{si } (x,y) \in [1,2,] \times [0,\pi] \\ 0 & \text{sinon.} \end{cases} \tag{14.601}$$

Les deux section de $f(x,y) = y\sin(xy)$ sont continues. Si on intègre d'abord par rapport à y on obtient

$$-\int_1^2 \frac{\pi\cos(\pi x)}{x} dx + \int_1^2 \frac{\sin(\pi x)}{x^2} dx,$$

qui n'est pas du tout immédiat, alors que, si on intègre d'abord par rapport à x on obtient

$$\int_0^\pi \cos y - \cos(2y)\, dy.$$

△

14.24.9 Intégrales sur des parties de $\mathbb{R}^2$

On veut évaluer l'intégrale de la fonction $f(x,y) = \sqrt{1-x^2}$ sur son domaine, la boule unité $B((0,0),1)$. La théorie introduite jusqu'ici n'est pas suffisante pour résoudre ce problème, parce que $B((0,0),1)$ n'est pas pavable. Les parties bornées de $\mathbb{R}^p$ sur lesquelles on peut intégrer des fonction sont dites mesurables (au sens de Riemann) parce que, comme on verra dans la suite, la mesure d'une partie de $\mathbb{R}^p$ est l'intégrale (s'il existe) de sa fonction caractéristique.

On peut dire que une partie de $\mathbb{R}^p$ est mesurable si son bord est «assez régulier». Dans $\mathbb{R}^2$ il est suffisant que le bord de A soit une réunion finie de courbes paramétrées continues. En particulier, on est très souvent dans un des deux cas suivantes

Régions du premier type A est borné et contenu entre les graphes de deux fonctions continues de x

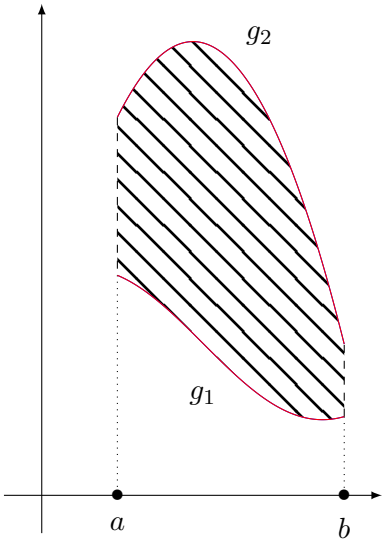
$$A = \{(x,y) \in \mathbb{R}^2 \,:\, a \leqslant x \leqslant b,\, g_1(x) \leqslant y \leqslant g_2(x)\},$$

avec g_1 et g_2 continues.

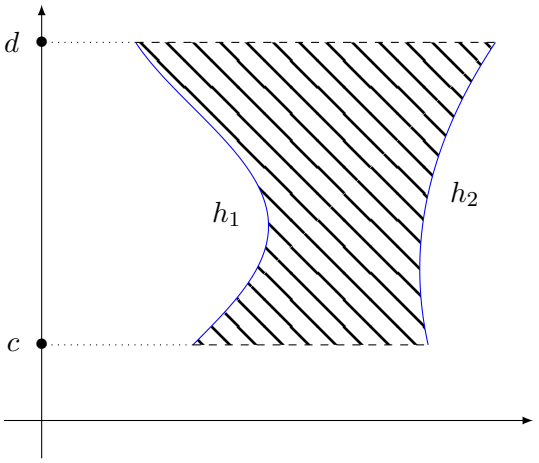
Régions du deuxième type A est borné et contenu entre les graphes de deux fonctions continues de y

$$A = \{(x,y) \in \mathbb{R}^2 \,:\, c \leqslant y \leqslant d,\, h_1(y) \leqslant x \leqslant h_2(y)\},$$

avec h_1 et h_2 continues.



(a) Une region du premier type



(b) Une region du deuxième type

FIGURE 14.20 – Régions du premier et du deuxième type

Exemple 14.174
Il y a des régions qui sont des deux types au même temps, comme les boules centrées à l'origine, le triangle de sommets $(0,0)$, $(0,a)$ et $(b,0)$, ou la région C délimité par les courbes $y = 2x$ et $y = x^2$. Cette dernière admets les représentations suivantes

$$C = \{(x,y) \in \mathbb{R}^2 \,:\, 0 \leqslant x \leqslant 1,\, x^2 \leqslant y \leqslant 2x\},$$

et

$$C = \{(x,y) \in \mathbb{R}^2 \,:\, 0 \leqslant y \leqslant 1,\, y/2 \leqslant x \leqslant \sqrt{y}\}.$$

△

Définition 14.175.
Soit f une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ dont le support A est une région du premier ou du deuxième type. On définit la fonction $\bar{f}$ comme

$$\bar{f}(x,y) = \begin{cases} f(x,y), & \text{si } (x,y) \in A, \\ 0, & \text{sinon.} \end{cases} \tag{14.602}$$

*La fonction f est dite **intégrable** si $\bar{f}$ est intégrable, et la valeur de son intégrale est*

$$\int_A f\,dV = \int_{\mathbb{R}^2} \bar{f}\,dV.$$

Une fonction continue définie sur une région du premier ou du deuxième type est toujours intégrable.

Pour fixer les idées on suppose ici que A est du premier type et contenue dans le pavé borné $R = [a, b] \times [c, d]$. En suivant la définition on obtient

$$\begin{aligned}\int_A f\,dV &= \int_{\mathbb{R}^2} \bar{f}\,dV = \\ &= \int_a^b \int_c^d \bar{f}\,dydx = \\ &= \int_a^b \left(\int_c^{g_1(x)} \bar{f}\,dy + \int_{g_1(x)}^{g_2(x)} \bar{f}\,dy + \int_{g_2(x)}^d \bar{f}\,dy \right) dx = \\ &= \int_a^b \int_{g_1(x)}^{g_2(x)} f\,dydx.\end{aligned} \tag{14.603}$$

De même, si A est du deuxième type on obtient

$$\int_A f\,dV = \int_c^d \int_{h_1(y)}^{h_2(y)} f\,dxdy. \tag{14.604}$$

Exemple 14.176

On peut maintenant résoudre notre problème de départ, évaluer l'intégrale de la fonction $f(x, y) = \sqrt{1 - x^2}$ sur $B((0, 0), 1)$. Nous choisissons de décrire la boule unité de $\mathbb{R}^2$ comme une région du premier type : $B((0, 0), 1) = \{(x, y) : x \in [-1, 1],\ -\sqrt{1 - x^2} \leqslant y \leqslant \sqrt{1 - x^2}\}$.

$$I = \int_B \sqrt{1 - x^2}\,dV = \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \sqrt{1 - x^2} dydx \tag{14.605}$$

La première intégrale à effectuer, par rapport à y, est l'intégrale d'une fonction constante. Ne pas oublier que l'on intègre $\sqrt{1 - x^2}$ par rapport à y ; c'est bien une constante et l'intégrale consiste seulement à multiplier par y :

$$I = \int_{-1}^1 \left[y\sqrt{1 - x^2} \right]_{y=-\sqrt{1-x^2}}^{y=\sqrt{1-x^2}} dx = 2 \int_{-1}^1 (1 - x^2)dx. \tag{14.606}$$

Cela est à nouveau une intégrale simple à effectuer. Le résultat est

$$2 \int_{-1}^1 (1 - x^2)dx = 2 \left[x - \frac{x^3}{3} \right]_{x=-1}^{x=1} = \frac{8}{3}. \tag{14.607}$$

△

Remarque 14.177.
Toutes les techniques d'intégration à une variable restent valables. Par exemple, lorsqu'une des intégrales est l'intégrale d'une fonction impaire sur un intervalle symétrique par rapport à zéro, l'intégrale vaut zéro.

14.178.
Par le lemme 13.115 nous savons que la mesure d'une région bornée de $\mathbb{R}^2$ est l'intégrale de sa fonction caractéristique, si elle existe.

La mesure d'une région bornée de $\mathbb{R}^2$ est dite son **aire**, et celle d'une région bornée de $\mathbb{R}^3$ est son **volume**. Voir aussi la remarque 19.10.

Exemple 14.179
On veut calculer l'aire de la région de la figure 14.21 définie par

$$A = \{(x, y) \in \mathbb{R}^2 \,|\, 0 \leqslant x \leqslant 1, x^3 - 1 \leqslant y \leqslant x\}.$$

On considère l'intégrale

$$\int_{\mathbb{R}^2} \chi_A \, dV = \int_0^1 \int_{x^3+1}^{x} 1 \, dy \, dx = \int_0^1 -x^3 + x + 1 \, dx = -\frac{1}{4} + \frac{1}{2} + 1 = \frac{5}{4}.$$

△

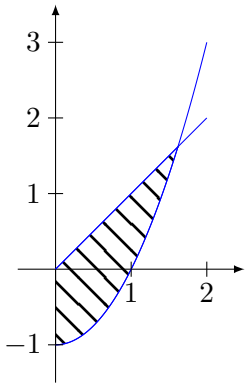


FIGURE 14.21 – La région A de l'exemple 14.179

Exercice 1

Parfois la région sur laquelle on veut intégrer peut être décrite indifféremment en deux façons, mais la fonction à intégrer nous force a choisir un ordre particulier. Vérifiez que la fonction $f(x, y) = \sin(y^2)$ sur la région triangulaire de sommets $(0, 0)$, $(0, 2)$, $(2, 2)$ doit être intégrée d'abord par rapport à x.

Si une région bornée n'est pas de premier ou de deuxième type on peut normalement la découper en morceaux plus faciles à décrire. On utilise alors la propriété suivante.

Lemme 14.180.
Soit A un sous-ensemble borné de $\mathbb{R}^2$ et soient B_1 et B_2 deux parties de A telles que $B_1 \cap B_2 = \varnothing$ et $B_1 \cup B_2 = A$. Alors, pour toute fonction f intégrable sur A (et en particulier pour sa fonction caractéristique) on a

$$\int_A f \, dV = \int_{B_1} f \, dV + \int_{B_2} f \, dV.$$

Exemple 14.181

La région D que nous voyons sur la figure 14.22 est bornée par la parabole $y^2 = 2x + 6$ et la droite $y = x - 1$. La région D est une région du deuxième type. Nous pouvons aussi la décrire comme l'union de deux régions du premier type D_1 et D_2,

$$D_1 = \{(x, y) \, : \, -3 \leqslant x \leqslant -1, \, -\sqrt{2x + 6} \leqslant y \leqslant \sqrt{2x + 6}\},$$

et

$$D_2 = \{(x, y) \, : \, -3 \leqslant x \leqslant -1, \, x - 1 \leqslant y \leqslant \sqrt{2x + 6}\}.$$

△

14.24.10 Intégrales sur des parties de $\mathbb{R}^3$

Dans ces notes nous n'avons pas l'ambition de traiter d'une façon rigoureuse l'étude des ensemble mesurables de $\mathbb{R}^3$. Comme dans la section précédente on se limitera à considérer des cas particuliers.

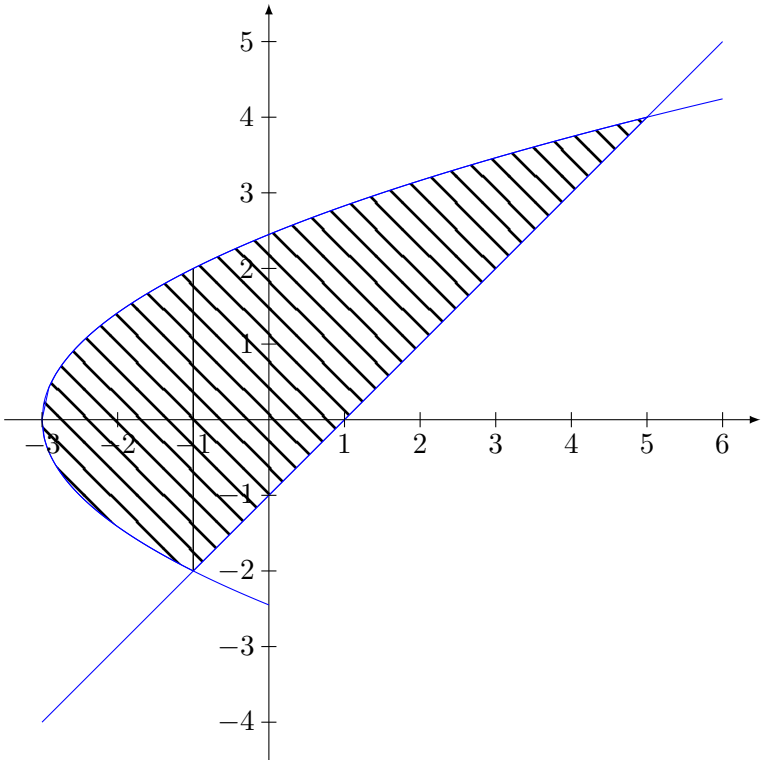


FIGURE 14.22 – La région D de l'exemple 14.181

Définition 14.182.
Soit E une région de $\mathbb{R}^3$. On dit que E est une ***région solide de premier type*** *si E est contenue entre les graphes de deux fonctions continues de x et y.*

$$E = \{(x, y, z) \in \mathbb{R}^3 \,|\, (x, y) \in A \subset \mathbb{R}^2, u_1(x, y) \leqslant z \leqslant u_2(x, y)\}.$$

Le sous-ensemble de A de $\mathbb{R}^2$ qui apparaît dans la définition 14.182 est la projection (ou l'ombre) de E sur le plan x-y.

Exemple 14.183
La région E donnée par une portion de sphère collée à un cône est une région solide de premier type

$$E = \{(x, y, z) \in \mathbb{R}^3 \,|\, (x, y) \in \overline{B\big((0, 0), 1\big)}, \sqrt{x^2 + y^2} \leqslant z \leqslant \sqrt{1 - x^2 - y^2}\}. \tag{14.608}$$

L'ombre de E est la boule unité de $\mathbb{R}^2$. L'ensemble $\sqrt{x^2 + y^2} \leqslant z$ est un cône posé sur sa pointe tandis que l'ensemble $z \leqslant \sqrt{1 - x^2 - y^2}$ est la demi-sphère. L'ensemble E contient les points entre les deux, voir la figure 14.23.

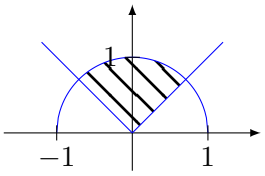


FIGURE 14.23 – Il faut voir ça en trois dimensions.

△

Si la fonction f, à intégrer sur E, et ses sections sont intégrables alors on peut réduire l'intégrale

$$\begin{aligned}\int_E f(x,y,z)\,dV &= \int_A \left(\int_{u_1(x,y)}^{u_2(x,y)} f(x,y,z)\,dz \right) dV = \\ &= \int_A \left(F(x,y,u_2(x,y)) - F(x,y,u_1(x,y)) \right) dV,\end{aligned} \tag{14.609}$$

où F est une primitive de f par rapport à la variable z, c'est à dire en considérant x et y comme des constantes. Il faut ensuite évaluer la partie qui reste comme dans la section précédente. Comme le calcul des aires dans $\mathbb{R}^2$, le calcul des volumes dans $\mathbb{R}^3$ est fait par des intégrales. En fait le **volume** d'une région solide dans $\mathbb{R}^3$ est sa mesure.

Définition 14.184.
La mesure d'une région de $\mathbb{R}^3$ est l'intégrale de sa fonction caractéristique.

Soit E une région solide du premier type, nous pouvons évaluer son volume par l'intégrale

$$\int_A \left(u_2(x,y) - u_1(x,y) \right) dV.$$

Parfois c'est plus intéressant de calculer le volume avec la formule de réduction contraire : l'intégrale double d'abord et puis l'intégrale simple par rapport à z. On parle alors de calcul de volume «par tranche».

Exemple 14.185
On veut calculer le volume de la boule de rayon a, centrée à l'origine $B = \{(x,y,z) \in \mathbb{R}^3 \,|\, x^2 + y^2 + z^2 \leqslant a^2\}$. On peut décrire B par

$$B = \left\{ (x,y,z) \in \mathbb{R}^3 \,|\, (x,y) \in D_a, -\sqrt{a^2 - x^2 - y^2} \leqslant z \leqslant \sqrt{a^2 - x^2 - y^2} \right\},$$

où D_a est le disque de rayon a centré en $(0,0)$, donc le volume B sera

$$2 \int_{D_a} \sqrt{a^2 - x^2 - y^2} dV.$$

Cet intégrale est un peu ennuyeuse à calculer. On peut simplifier le calcul en observant que pour $\bar{z}$ fixé dans l'intervalle $[-a,a]$ la section de la boule au niveau $\bar{z}$ est un disque de rayon $\sqrt{a^2 - z^2}$. L'aire d'un tel disque est $\pi(a^2 + z^2)$. Si on réduit l'intégrale de volume de la façon

$$\int_B 1\,dV = \int_{-a}^{a} \sqrt{a^2 - z^2}\,dz,$$

on obtient tout de suite la valeur cherchée : le volume de B est $4/3\pi a^3$. △ **Exemple 14.186**

On calcul l'intégrale de $f(x,y,z) = z$ sur la pyramide P bornée par le plans $x = 0$, $y = 0$, $x + y + z = 1$, $x + y + z/2 = 1$. On remarque tout de suite que le plans $x + y + z = 1$, $x + y + z/2 = 1$ se coupent en la droite $x + y = 1$, $z = 0$ (on se souvient qu'*une* droite dans $\mathbb{R}^3$, c'est *deux* équations). Cela veut dire que la projection de P sur le plan x-y est le triangle T borné par les droites $x = z = 0$, $y = z = 0$ et $x + y = 1$, $z = 0$. On décrit donc P par

$$P = \{(x,y,z) \in \mathbb{R}^3 \,|\, (x,y) \in T,\, 1 - 2x - 2y \leqslant z \leqslant 1 - x - y\}$$

et T par

$$T = \{(x,y) \in \mathbb{R}^2 \,|\, 0 \leqslant x \leqslant 1,\, 0 \leqslant y \leqslant 1 - x\},$$

donc l'intégrale de f sur P est

$$\int_p f(x,y,z)\,dV = \int_0^1 \int_0^{1-x} \int_{1-2x-2y}^{1-x-y} z\,dz\,dy\,dx = -\frac{1}{24}.$$

Notez que lorsque x et y sont entre 0 et 1, nous avons bien $1 - 2x - 2y < 1 - x - y$, d'où le fait que nous mettons $1 - 2x - 2y$ dans la borne inférieure de l'intégrale. △

De façon analogue on définit les régions solides du deuxième et du troisième type.

14.24.11 Intégrales de fonctions non bornées sur des ensembles non bornés

Soit $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$, une fonction positive. On dit qu'elle est **intégrable** sur $E \subset \mathbb{R}^n$ si

(1) $\forall r > 0$, la fonction $f_r(x) = f(x)\mathbb{1}_{f<r}$ est intégrable sur E_r ;

(2) la limite $\lim_{r\to\infty} \int_{E_r} f_r$ est finie.

Dans ce cas, on pose

$$\int_E f = \lim_{r\to\infty} \int_{E_r} f_r. \tag{14.610}$$

Théorème 14.187.
Soit E mesurable dans $\mathbb{R}^n$ et $f\colon E \to \overline{\mathbb{R}}$. Si f est mesurable et s'il existe $g\colon E \to \overline{\mathbb{R}}$ intégrable sur E telle que $|f(x)| \leqslant g(x)$ pour tout $x \in E$, alors f est intégrable sur E.

Réciproquement, si f est intégrable sur E, alors f est mesurable.

Lemme 14.188.
Si f est une fonction sur $[a, \infty[$, alors nous avons la formule

$$\lim_{b\to\infty} \int_a^b f(x)dx = \int_a^\infty f(x)dx \tag{14.611}$$

au sens où si un des deux membres existe, alors l'autre existe et est égal.

Démonstration. Supposons que le membre de gauche existe. Cela signifie que la fonction

$$\psi(x) = \int_a^x f \tag{14.612}$$

est bornée. Soit M, un majorant. Pour toute fonction simple φ dominant f, on a que $\int \varphi \leqslant M$, donc l'ensemble sur lequel on prend le supremum pour calculer $\int_a^\infty f$ est majoré par M et possède donc un supremum. Nous avons donc

$$\int_a^\infty f \leqslant \lim_{b\to\infty} \int_a^b f. \tag{14.613}$$

□

Chapitre 15

Suites et séries de fonctions

Nous commençons par donner quelques éléments à propos de dérivée et de différentielle pour des fonctions $\mathbb{C} \to \mathbb{C}$ parce que les séries entières vont souvent être des fonctions complexes. Le gros du chapitre sur les fonctions holomorphes est le chapitre 25.

15.1 Différentielle et dérivée complexe

Nous identifions $\mathbb{R}^2$ à $\mathbb{C}$ par l'application

$$\begin{aligned} \varphi\colon \mathbb{R}^2 &\to \mathbb{C} \\ (x,y) &\mapsto x+iy. \end{aligned} \tag{15.1}$$

Dans cette partie, nous désignons par Ω un ouvert de $\mathbb{C}$.

Définition 15.1.
*Une fonction $f\colon \Omega \to \mathbb{C}$ est $\mathbb{C}$-**dérivable** si la limite*

$$\lim_{\substack{h\to 0\\ h\in\mathbb{C}}} \frac{f(a+h)-f(a)}{h} \tag{15.2}$$

existe. Dans ce cas, cette limite est la dérivée de f et est notée f'.

Définition 15.2.
*Soit Ω un ouvert dans $\mathbb{C}$. Une fonction $f\colon \Omega \to \mathbb{C}$ est **holomorphe** si elle est $\mathbb{C}$-dérivable sur Ω.*

Définition 15.3.
Une matrice de la forme

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \tag{15.3}$$

*avec $\alpha, \beta \in \mathbb{R}$ est une **similitude**.*

Lemme 15.4.
En tant qu'application linéaire $\mathbb{C} \to \mathbb{C}$, l'opération de multiplication par $\alpha + \beta i$ est la matrice

$$\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}. \tag{15.4}$$

Démonstration. Cela est vite remarqué en calculant explicitement $(\alpha+\beta i)(u_1+iu_2)$. □

Lemme 15.5.
Une application $A\colon \mathbb{C} \to \mathbb{C}$ est $\mathbb{C}$-linéaire si et seulement si elle est une similitude en tant qu'application $\mathbb{R}^2 \to \mathbb{R}^2$.

Dans ce cas, il existe $z_0 \in \mathbb{C}$ tel que $A(z) = z_0 z$ pour tout $z \in \mathbb{C}$.

Démonstration. Commençons par considérer l'application A sur $\mathbb{R}^2$. Elle est en particulier une application $\mathbb{R}$-linéaire et par conséquent il existe une matrice $\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$ telle que

$$A\begin{pmatrix}x\\ y\end{pmatrix} = \begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}. \tag{15.5}$$

Nous voulons maintenant imposer la $\mathbb{C}$-linéarité, c'est à dire que nous voulons

$$A\big((a+bi)(x+iy)\big) = (a+bi)A(x+iy) \tag{15.6}$$

pour tout $a, b, x, y \in \mathbb{R}$. À gauche nous avons

$$A\big(ax - by + i(bx+ay)\big) \tag{15.7}$$

et à droite nous avons

$$(a+bi)\big(\alpha x + \beta y + i(\gamma x + \delta y)\big). \tag{15.8}$$

En égalant les deux expressions nous obtenons les équations

$$\begin{cases} \beta b = -b\gamma & \text{(15.9a)}\\ -\alpha b + \beta a = a\beta - b\delta & \text{(15.9b)}\\ \delta b = b\alpha & \text{(15.9c)}\\ -\gamma b + \delta a = b\beta + a\delta, & \text{(15.9d)}\end{cases}$$

dont nous tirons immédiatement que $\gamma = -b\beta$ et $\delta = \alpha$. La matrice de A est donc de la forme demandée.

Inversement nous devons prouver que la fonction

$$f(x+iy) = \alpha x + \beta y + i(-\beta x + \alpha y) \tag{15.10}$$

est $\mathbb{C}$-linéaire, c'est à dire qu'elle vérifie $f(z_0 z) = z_0 f(z)$ pour tout $z_0, z \in \mathbb{C}$. Cela est un simple calcul que nous confions à Sage : le code suivant affiche «0».

```
#! /usr/bin/sage -python
# -*- coding: utf8 -*-

from sage.all import *

def f(z):
    var('alpha,beta')
    x=z.real_part()
    y=z.imag_part()
    return alpha*x+beta*y+I*( -beta*x+alpha*y  )


var('a,b,x,y')

A=a+b*I
Z=x+y*I

z1=f( A*Z  )
z2=A*f( Z )

rep=z1-z2
print(rep.full_simplify())
```

tex/frido/code_sage3.py

Pour conclure, notons que la fonction (15.10) est la fonction de multiplication par $\alpha - i\beta$. □

15.6.
Soit une fonction $f\colon \mathbb{C} \to \mathbb{C}$ et l'isomorphisme canonique $\varphi\colon \mathbb{C} \to \mathbb{R}^2$. La fonction f définit une la fonction

$$F = \varphi^{-1} \circ f \circ \varphi\colon \mathbb{R}^2 \to \mathbb{R}^2. \tag{15.11}$$

Cela est la fonction $\mathbb{R}^2 \to \mathbb{R}^2$ associée à f. Il serait tentant de croire que tout ce qui est vrai pour F est également vrai pour f. Eh bien non.

Par exemple, F peut être différentiable sans que f le soit. La proposition suivant donne une condition sur dF pour que f soit différentiable.

Proposition 15.7.
Une fonction $f\colon \Omega \to \mathbb{C}$ est $\mathbb{C}$-dérivable en $a \in \Omega$ si et seulement si elle est différentiable en a et si df_a est une similitude.

Plus précisément avec les notations de 15.6, la fonction f est $\mathbb{C}$-dérivable (donc holomorphe) au point $z_0 = x_0 + iy_0$ si et seulement si la fonction F est différentiable en (x_0, y_0) et si la matrice de dF est de la forme

$$dF = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}, \tag{15.12}$$

c'est à dire si $dF_{(x_0,y_0)}$ fournit une application $\mathbb{C}$-linéaire.

Dans ce cas, le lien entre $\mathbb{C}$-dérivée et différentielle est donné par

$$(df_{z_0})(z) = f'(z_0)z. \tag{15.13}$$

Démonstration. Nous décomposons f en parties réelles et imaginaires :

$$f(x + iy) = P(x, y) + iQ(x, y) \tag{15.14}$$

où P et Q sont des fonctions réelles. La jacobienne de F est la matrice

$$\begin{pmatrix} \frac{\partial P}{\partial x} & \frac{\partial P}{\partial y} \\ \frac{\partial Q}{\partial x} & \frac{\partial Q}{\partial y} \end{pmatrix}, \tag{15.15}$$

et la condition dont nous parlons s'écrit comme le système

$$\begin{cases} \dfrac{\partial P}{\partial x} = \dfrac{\partial Q}{\partial y} & \text{(15.16a)} \\ \dfrac{\partial P}{\partial y} = -\dfrac{\partial Q}{\partial x}. & \text{(15.16b)} \end{cases}$$

Si F est différentiable en (x_0, y_0) alors nous avons

$$F\big((x_0, y_0) + (h, k)\big) = F(x_0, y_0) + dF_{(x_0,y_0)} \begin{pmatrix} h \\ k \end{pmatrix} + s(|h| + |k|) \tag{15.17}$$

où s est une fonction vérifiant $\lim_{t\to 0} \frac{s(t)}{t} = 0$. Soit

$$dF_{(x_0,y_0)} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}. \tag{15.18}$$

Si nous posons $\sigma = \alpha - i\beta$ et $w = h + ik$, l'équation (15.17) s'écrit dans $\mathbb{C}$ sous la forme

$$f(z_0 + w) = f(z_0) + \sigma w + s(|w|), \tag{15.19}$$

ce qui implique que f est $\mathbb{C}$-dérivable en z_0.

Supposons maintenant que f soit $\mathbb{C}$-dérivable en z_0. Alors nous avons

$$f'(z_0) = \lim_{w\to 0} \frac{f(z_0+w)-f(z_0)}{w} = \sigma \in \mathbb{C}, \tag{15.20}$$

ce qui se récrit sous la forme

$$\lim_{w\to 0} \frac{f(z_0+w)-f(z_0)-\sigma w}{w} = 0. \tag{15.21}$$

Si nous posons $z_0 = x_0 + iy_0$, $w = h + ik$ et $\sigma = \alpha - i\beta$ nous avons

$$\lim_{(h,k)\to(0,0)} \left| \frac{F\big((x_0,y_0)+(h,k)\big) - F(x_0,y_0) - \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix}}{|w|} \right| = 0, \tag{15.22}$$

ce qui signifie que F est différentiable et que sa différentielle est la matrice

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}. \tag{15.23}$$

La matrice (15.23) est, vue dans $\mathbb{R}^2$, la matrice de multiplication dans $\mathbb{C}$ par $\alpha - i\beta = f'(z_0)$. En d'autre termes, dans $\mathbb{C}$ nous avons

$$df_{z_0}(z) = f'(z_0)z, \tag{15.24}$$

et en particulier la différentielle est donnée par

$$df_{z_0} = f'(z_0)dz. \tag{15.25}$$

□

Exemple 15.8(Une application C^∞ mais pas $\mathbb{C}$-dérivable)
Nous considérons la fonction

$$\begin{aligned} f\colon \mathbb{C} &\to \mathbb{C} \\ x+iy &\mapsto x. \end{aligned} \tag{15.26}$$

Vu que c'est une application linéaire, elle est différentiable une infinité de fois et sa différentielle est elle-même. C'est donc une application C^∞.

Elle n'est cependant pas $\mathbb{C}$-dérivable. En effet le quotient différentiel est, pour $\epsilon \in \mathbb{C}$:

$$\frac{f(x+iy+\epsilon_x+i\epsilon_y)-f(x+iy)}{\epsilon} = \frac{\epsilon_x}{\epsilon}. \tag{15.27}$$

Cela n'a pas de limite lorsque $\epsilon \to 0$. Pour voir cela nous invoquons la méthode des chemins du corollaire 12.98 avec les chemins $\epsilon_1(t) = t$ et $\epsilon_2(t) = it$. Dans le premier cas, le quotient différentiel vaut 1 pour tout t, tandis que dans le second il vaut zéro pour tout t. △

15.2 Série de fonctions

Les séries de fonctions sont des cas particuliers de suites, étant donné que, par définition,

$$\sum_{n=1}^{\infty} f_n = \lim_{N\to\infty} \sum_{n=1}^{N} f_n. \tag{15.28}$$

Il est bon de se rappeler des convergences absolues (définition 10.45) et uniforme (équation 10.99).

Lemme 15.9.
Soient des fonctions $u_n\colon \Omega \to \mathbb{C}$. Si il existe une suite réelle positive $(a_n)_{n\in\mathbb{N}}$ telle que

(1) pour tout $z \in \Omega$ et pour tout $n \in \mathbb{N}$ nous avons $|u_n(z)| \leqslant a_n$ (c'est à dire $a_n \geqslant \|u_n\|_\infty$),

(2) la somme $\sum_n a_n$ converge,

alors la série de fonctions $\sum_{n=0}^{\infty} u_n$ converge normalement.

Démonstration. Découle du lemme de comparaison 10.55. □

Théorème 15.10.
Soit $\sum_{k=1}^{\infty} g_k(x)$, une série de fonctions complexes où $g_k(x) = \varphi_k(x)\psi_k(x)$. Supposons que

(1) $\varphi_k\colon A \to \mathbb{C}$ et $|\sum_{k=1}^{K} \varphi_k(x)| \leqslant M$ où M est indépendant de x et K,

(2) $\psi_k\colon A \to \mathbb{R}$ avec $\psi_k(x) \geqslant 0$ et pour tout x dans A, $\psi_{k+1}(x) \leqslant \psi_k(x)$, et enfin supposons que $\psi_k(x)$ converge uniformément vers 0.

Alors $\sum_{k=1}^{\infty} g_k$ est uniformément convergente.

Théorème 15.11.
Si la série de puissances (réelle) converge en $x = x_0 + R$, alors elle converge uniformément sur $[x_0 - R + \epsilon, x_0 + R]$ ($\epsilon > 0$) vers une fonction continue.

Proposition 15.12.
Soit (u_n) une suite de fonctions continues $u_n\colon \Omega \subset \mathbb{C} \to \mathbb{C}$. Si la série $\sum_n u_n$ converge normalement alors la somme est continue.

Démonstration. Nous posons $u(z) = \lim_{N\to\infty} \sum_{n=0}^{N} u_n(z)$, et nous vérifions que la fonction ainsi définie sur Ω est continue. Soit $z \in \Omega$ et prouvons la continuité de u au point z. Pour tout z' dans un voisinage de z nous avons

$$\begin{aligned} \big|u(z) - u(z')\big| &= \left|\sum_{n=0}^{N} u_n(z) - \sum_{n=0}^{N} u_n(z') + \sum_{n=N+1}^{\infty} u_n(z) - \sum_{n=N+1}^{\infty} u_n(z')\right| && (15.29a)\\ &\leqslant \left|\sum_{n=0}^{N} u_n(z) - \sum_{n=0}^{N} u_n(z')\right| + \sum_{n=N+1}^{\infty} |u_n(z)| + \sum_{n=N+1}^{\infty} |u_n(z')|. && (15.29b) \end{aligned}$$

Étant donné que les sommes partielles sont continues, en prenant N suffisamment grand, le premier terme peut être rendu arbitrairement petit. Si N est suffisamment grand, le second terme est également petit. Par contre, cet argument ne tient pas pour le troisième terme parce que nous souhaitons une majoration pour tout z' dans une boule autour de z. Nous devons donc écrire

$$\sum_{n=N}^{\infty} |u_n(z)| \leqslant \sum_{n=N+1}^{\infty} \|u_n\|_\infty. \tag{15.30}$$

Ce dernier est arbitrairement petit lorsque N est grand. Notons que nous avons utilisé l'hypothèse de convergence normale. □

La même propriété, avec la même démonstration, tient dans le cas d'espaces vectoriels normée.

Proposition 15.13.
Soient E et F, deux espaces vectoriels normés, Ω une partie ouverte de E et une suite de fonctions $u_n\colon \Omega \to F$ convergeant normalement sur Ω, c'est à dire que $\sum_n \|u_n\|_\infty$ converge, la norme $\|.\|_\infty$ devant être comprise comme la norme supremum sur Ω. Alors la fonction $u = \sum_n u_n$ est continue sur Ω.

Démonstration. Soit $x, x' \in \Omega$ en supposant que $\|x - x'\|$ est petit. Soit encore $\epsilon > 0$. Nous allons montrer la continuité en x. Pour cela nous savons que pour tout N l'inégalité suivante est correcte :

$$\|u(x) - u(x')\| \leqslant \left\|\sum_{n=0}^{N} u_n(x) - \sum_{n=0}^{N} u_n(x')\right\| + \sum_{n=N+1}^{\infty} \|u_n(x)\| + \sum_{n=N+1}^{\infty} \|u_n(x')\|. \tag{15.31}$$

Les deux derniers termes sont majorés par $\sum_{n=N+1}^{\infty} \|u_n\|_\infty$ qui, par hypothèse, peut être rendu aussi petit que souhaité en choisissant N assez grand. Nous choisissons donc un N tel que ces deux termes soient plus petits que ϵ. Ce N étant fixé, la fonction $\sum_{n=0}^{N} u_n$ est continue et nous pouvons choisir x' assez proche de x pour que le premier terme soit majoré par ϵ. □

Théorème 15.14.
Si les g_k sont continues et si $\sum g_k$ converge uniformément, alors $\sum g_k$ est continue.

Le corollaire suivant permet de considérer des séries de fonctions indexées par exemple par $\mathbb{Z}$ plutôt que par $\mathbb{N}$.

Corollaire 15.15.
Une famille dénombrable de fonctions continues convergeant normalement converge vers une fonction continue.

Démonstration. Soit I dénombrable et considérons une famille de fonctions continues $(f_n)_{n\in I}$ telles que la famille $(\|f_i\|_\infty)_{i\in I}$ soit sommable. Le proposition 9.184 nous permet d'utiliser une bijection entre I et $\mathbb{N}$. Le théorème 15.12 s'applique alors. □

Théorème 15.16 (Critère de Weierstrass)**.**
Soit une suite de fonctions $f_k\colon A \to \mathbb{C}$ telles que $|f_k(x)| \leqslant M_k \in \mathbb{R}$, $\forall x \in A$. Si $\sum_{k=1}^{\infty} M_k$ converge, alors $\sum_{k=1}^{\infty} f_k$ converge absolument et uniformément.

Démonstration. La convergence normale est facile : l'hypothèse dit que $\|f_k\|_\infty \leqslant M_k$, et donc que

$$\sum_{k=1}^{\infty} \|f_k\|_\infty \leqslant \sum_k M_k < \infty. \tag{15.32}$$

La convergence uniforme est à peine plus subtile. Nous nommons F la fonction somme. Pour tout x et pour tout N, nous avons

$$\left\| \sum_{n=1}^{N} f_n(x) - F(x) \right\| = \| \sum_{n=N}^{\infty} f_n(x) \| \tag{15.33a}$$

$$\leqslant \sum_{n=N}^{\infty} \|f_k(x)\| \tag{15.33b}$$

$$\leqslant \sum_{n=N}^{\infty} \|f_n\|_\infty. \tag{15.33c}$$

La convergence normale étant assurée, la série $\sum_{n_1}^{\infty} \|f_n\|_\infty$ est finie, ce qui implique que la queue de somme $\sum_{n=N}^{\infty} \|f_n\|_\infty$ tend vers zéro lorsque $N \to \infty$. Pour tout ϵ, il existe donc un N (non dépendant de x) tel que

$$\| \sum_{n=1}^{N} f_n(x) - F(x) \| \leqslant \epsilon. \tag{15.34}$$

En prenant le supremum sur $x \in A$ nous trouvons la convergence uniforme. □

Remarque 15.17.
Il n'y a pas de critère correspondant pour les suites. Il n'est pas vrai que si $\lim_{n\to\infty} \|f_n\|$ existe, alors $\lim_{n\to\infty} f_n$ existe, comme le montre l'exemple

$$f_n(x) = \begin{cases} 1 & \text{si } x \in [0,1] \text{ et } n \text{ est pair} \\ 1 & \text{si } x \in [1,2] \text{ et } n \text{ est impair} \\ 0 & \text{sinon.} \end{cases} \tag{15.35}$$

15.2.1 Intégration de séries de fonctions

Théorème 15.18.
La somme uniforme de fonctions intégrables sur un ensemble de mesure fini est intégrable et on peut permuter la somme et l'intégrale.

En d'autres termes, supposons que $\sum_{n=0}^{\infty} f_n$ converge uniformément vers F sur A avec $\mu(A) < \infty$. Si F et f_n sont des fonctions intégrables sur A alors

$$\int_A F(x)d\mu(x) = \sum_{n=0}^{\infty} \int_A f_n(x)d\mu(x). \tag{15.36}$$

Démonstration. Ce théorème est une conséquence du théorème 13.127. En effet nous définissons la suite des sommes partielles

$$F_N = \sum_{n=0}^{N} f_n. \tag{15.37}$$

La limite $\lim_{N\to\infty} F_N = F$ est uniforme. Par conséquent la fonction F est intégrable et

$$\int_A F = \lim_{N\to\infty} \int_A F_N = \lim_{N\to\infty} \int_A \sum_{n=0}^{N} f_n = \lim_{N\to\infty} \sum_{n=0}^{N} \int_A f_n = \sum_{n=0}^{\infty} \int_A f_n. \tag{15.38}$$

La première égalité est le théorème 13.127, les autres sont de simples manipulations rhétoriques. □

Le théorème suivant est une paraphrase du théorème de la convergence dominée de Lebesgue (13.134).

Théorème 15.19.
Soient des fonctions $(f_n)_{n\in\mathbb{N}}$ telles que $\sum_{n=0}^{N} f_n$ soit intégrable sur $(\Omega, \mathcal{A}, \mu)$ pour chaque N. Nous supposons que la somme converge simplement vers

$$f(x) = \sum_{n=0}^{\infty} f_n(x) \tag{15.39}$$

et qu'il existe une fonction g telle que

$$\left| \sum_{n=0}^{N} f_n \right| < g \tag{15.40}$$

pour tout $N \in \mathbb{N}$. Alors

(1) $\sum_{n=0}^{\infty} f_n$ est intégrable,

(2) on peut permuter somme et intégrale :

$$\lim_{N\to\infty} \int_\Omega \sum_{n=0}^{N} f_n d\mu = \int_\Omega \sum_{n=0}^{\infty} f_n, \tag{15.41}$$

(3)

$$\lim_{N\to\infty} \int_\Omega \left| \sum_{n=0}^{N} f_n - \sum_{n=0}^{\infty} f_n \right| = \lim_{N\to\infty} \int_\Omega \left| \sum_{n=N}^{\infty} f_n \right| = 0. \tag{15.42}$$

Théorème 15.20.
Soit f_n des fonctions $C^1[a,b]$ telles que

(1) la série $\sum_n f_n(x_0)$ converge pour un certain $x_0 \in [a,b]$,

(2) la série des dérivées $\sum_n f'_n$ converge uniformément sur $[a,b]$.

Alors la série $\sum_n f_n$ converge vers une fonction F et

(1) La convergence est uniforme sur $[a,b]$.

(2) La fonction F est dérivable

(3) $F'(x) = \sum_n f'_n(x)$.

15.2.2 Différentiabilité

Lemme 15.21.
Soient E et F deux espaces vectoriels normés. Si la suite (T_n) converge vers T dans $\mathcal{L}(E,F)$, alors pour tout $v \in E$ nous avons

$$\left(\sum_{n=0}^{\infty} T_n\right)(v) = \sum_{n=0}^{\infty} T_n(v). \tag{15.43}$$

Théorème 15.22 ([189])**.**
Soit E et F, deux espaces vectoriels normés, Ω un ouvert connexe par arcs de E. Soit (u_n) une suite de fonctions $u_n\colon \Omega \to F$ telle que

(1) pour tout n, la fonction u_n est de classe C^1 sur Ω,

(2) la série $\sum_n u_n$ converge simplement sur Ω,

(3) la série des différentielles $\sum_n(du_n)$ converge normalement sur tout compact de Ω.

Alors la somme $u = \sum_n u_n$ est de classe C^1 sur Ω et sa différentielle est donnée par

$$du = \sum_{n=0}^{\infty} du_n. \tag{15.44}$$

Démonstration. Pour chaque n, la fonction $du_n\colon \Omega \to \mathcal{L}(E,F)$ est une fonction continue parce que u_n est de classe C^1. La série convergeant normalement, la fonction $\sum_{n=0}^{\infty} du_n$ est également continue par la proposition 15.13. La difficulté de ce théorème est donc de prouver que cela est bien la différentielle de la fonction $\sum_n u_n$.

Soit $a, x \in \Omega$ et $\gamma\colon [0,1] \to \Omega$ un chemin joignant a à x. Nous considérons ce chemin en coordonnées normales et nous notons l sa longueur. Par définition (14.59),

$$\clubsuit = \int_\gamma \sum_{n=0}^{\infty} du_n = \int_0^l \sum_n (du_n)_{\gamma(t)}\big(\gamma'(t)\big)dt \tag{15.45}$$

Si nous notons $f_n(t) = (du_n)_{\gamma(t)}\big(\gamma'(t)\big)$, sachant que la paramétrisation est normale ($\|\gamma'(t)\| = 1$) nous avons [1]

$$\|f_n(t)\| \leqslant \|(du_n)_{\gamma(t)}\| \leqslant \|du_n\|_\infty. \tag{15.46}$$

Or la série des $\|du_n\|_\infty$ converge par hypothèse. L'intervalle $[0,l]$ étant compact, les fonctions f_n sont uniformément (en n) bornées par le nombre $\sum_n \|du_n\|_\infty$ qui est intégrable sur $[0,1]$. Par la convergence dominée (théorème 13.134) nous permutons la somme et l'intégrale :

$$\clubsuit = \sum_{n=0}^{\infty} \int_0^l (du_n)_{\gamma(t)}\big(\gamma'(t)\big)dt = \sum_{n=0}^{\infty} \big(u_n(x) - u_n(a)\big) = u(x) - u(a) \tag{15.47}$$

où nous avons utilisé le théorème 14.49. Jusqu'à présent nous avons montré que

$$u(x) = u(a) + \int_\gamma \sum_{n=0}^{\infty} du_n = u(a) + \int_0^l \sum_{n=0}^{\infty} (du_n)_{\gamma(t)}\big(\gamma'(t)\big)dt. \tag{15.48}$$

Nous allons utiliser cela pour calculer $du_x(v)$ selon la bonne vieille formule

$$du_x(v) = \frac{d}{ds}\Big[u(x+sv)\Big]_{s=0}. \tag{15.49}$$

Cela sera fait en considérant à nouveau un chemin γ_s joignant a à $x+sv$ en paramétrisation normale ; nous notons l_s sa longueur. Dans le calcul suivant, nous inversons la somme et l'intégrale

1. Histoire de ne pas s'embrouiller, il faut se rendre compte que $\|du_n\|_\infty = \sup_{x\in\Omega} \|(du_n)_x\|$.

de la même façon qu'avant. En piste maestro

$$du_x(v) = \frac{d}{ds}\int_0^{l_s}\sum_{n=0}^{\infty}(du_n)_{\gamma_s(t)}\big(\gamma_s'(t)\big)dt\Big|_{s=0} \tag{15.50a}$$

$$= \frac{d}{ds}\sum_{n=0}^{\infty}\int_{\gamma_s}du_n\Big|_{s=0} \tag{15.50b}$$

$$= \frac{d}{ds}\sum_{n=0}^{\infty}\Big[u_n\big(\gamma_s(l_s)\big) - u_n\big(\gamma_s(0)\big)\Big]_{s=0} \tag{15.50c}$$

$$= d\left(\sum_{n=0}^{\infty}u_n\right)_x(v). \tag{15.50d}$$

□

15.3 Séries entières

Dans cette section nous allons parler de séries complexes autant que de séries réelles. L'étude des propriétés à proprement parler complexes des séries entières (holomorphie) sera effectuée dans le chapitre dédié, voir le théorème 25.15 et ses conséquences.

15.3.1 Disque de convergence

Une **série de puissance** est une série de la forme

$$\sum_{k=0}^{\infty}c_k(z-z_0)^k \tag{15.51}$$

où $z_0 \in \mathbb{C}$ est fixé, (c_k) est une suite complexe fixée, et z est un paramètre complexe. Nous disons que cette série est *centrée* en z_0.

Définition 15.23.
*Une **série entière** est une somme de la forme*

$$\sum_{n=0}^{\infty}a_nz^n \tag{15.52}$$

avec $a_n, z \in \mathbb{C}$.

Une série entière peut définir une fonction

$$f(z) = \sum_n a_nz^n. \tag{15.53}$$

Le but de cette section est d'étudier des conditions sur la suite (a_n) qui assurent la continuité de f ou la possibilité de dériver ou intégrer la série terme à terme.

Définition 15.24.
*Soit $\sum_{n\in\mathbb{N}}a_nz^n$ une série entière. Le **rayon de convergence** de cette série est le nombre*

$$R = \sup\{r \in \mathbb{R}^+ \text{ tel que la suite } (a_nr^n) \text{ est bornée}\} \in [0,\infty]. \tag{15.54}$$

*La boule $B(0,R)$ est le **disque de convergence** de la série.*

Le rayon de convergence d'une série ne dépend que des réels $|a_n|$, même si à la base $a_n \in \mathbb{C}$.

Remarque 15.25.
Si pour tout n nous avons $|b_n| \geqslant |a_n|$ alors le rayon de convergence de la série $\sum_n a_n z^n$ est au moins aussi grand que celui de la série $\sum_n b_n z^n$. Cela y compris lorsque l'un ou l'autre des rayons de convergences est infini.

Lemme 15.26 (Critère d'Abel)**.**
Soit $R > 0$ le rayon de convergence de la somme $\sum_n a_n z^n$ et $z \in \mathbb{C}$.

(1) Si $|z| < R$ alors la série converge absolument.

(2) Si $R < \infty$ et si $|z| > R$ alors la série diverge.

Démonstration. Démonstration en deux parties.

(1) Si $|z| < R$ alors la suite $(a_n z^n)$ est bornée et il existe un nombre $M \in \mathbb{R}$ tel que $|a_n| r^n \leqslant M$ pour tout n. Nous considérons alors un r tel que $|z| < r < R$ et nous pouvons calculer :

$$|a_n z^n| = |a_n| r^n \big(\frac{|z|}{r}\big)^n \leqslant M \left(\frac{|z|}{r}\right)^n \tag{15.55}$$

Vu que $|z| < r$ nous tombons sur la série géométrique (10.127) qui converge. Par le critère de comparaison [2] la série $\sum_{n=0}^{\infty} |a_n z^n|$ converge.

(2) Par définition du rayon de convergence, la suite $(a_n z^n)$ n'est donc pas bornée et la série ne peut pas converger à cause de la proposition 10.53.

□

Le critère d'Abel parle bien de convergence absolue, et non de convergence normale. Pour chaque t, la série $\sum_k |a_n t^k|$ converge. Si par contre nous posons $u_k(t) = a_k t^k$, nous n'avons a priori pas la convergence normale $\sum_k \|u_k\|_\infty$, même pas si la norme est la norme supremum sur $B(0, R)$ [3] . Prenons comme exemple simplement $a_k = 1$ pour tout k. Pour tout $|t| < 1$, la série $\sum_k t^k$ converge absolument (série géométrique), mais nous aurions $\|u_k\|_\infty = 1$ et donc divergence évidente de $\sum_k \|u_k\|_\infty$.

La proposition suivante sera surtout utile lorsqu'on parlera de dérivée.

Proposition 15.27 ([190])**.**
Quel que soit le nombre $\alpha \in \mathbb{R}$, les séries $\sum_n a_n z^n$ et $\sum_n n^\alpha a_n z^n$ ont même rayon de convergence.

Démonstration. Nous posons

$$E = \{r \in \mathbb{R}^+ \text{ tel que } (a_n r^n) \text{ est borné }\} E' = \{r \in \mathbb{R}^+ \text{ tel que } (n^\alpha a_n r^n) \text{ est borné }\} \tag{15.56a}$$

Et aussi $R = \sup(E)$, $R' = \sup(E')$. Le fait que $E' \geqslant E$ est facile. Nous supposons $R > 0$ et nous considérons $r < R$ (c'est à dire $r \in E$). Nous allons montrer que $r \in E'$. Pour cela nous prenons un nombre s tel que $r < s < R$. Nous avons

$$n^\alpha a_n r^n = n^\alpha a_n \left(\frac{r}{s}\right)^n s^n = n^\alpha \left(\frac{r}{s}\right)^n a_n s^n. \tag{15.57}$$

Mais $r/s < 1$, donc le lemme 6.94 dit que $n^\alpha (r/s)^n \to 0$. Cela est donc borné par une constante M. Donc

$$n^\alpha a_n r^n \leqslant M a_n s^n. \tag{15.58}$$

Mais la suite $(a_n s^n)$ est bornée. Donc la suite $n^\alpha a_n r^n$ est également bornée, ce qui prouve que $r \in E'$. □

Remarque 15.28.
Au fond, cette proposition n'est rien d'autre que dire que dans $n^\alpha r^n$, l'effet «convergent» est r^n qui est une décroissance exponentielle tandis que l'effet «divergent» est n^α qui a une croissance seulement polynômiale.

2. Lemme 10.55.
3. Il y aurait par contre bien convergence sur tout compact ? Cher lecteur, dites moi ce que vous en pensez

Théorème 15.29 (Formule de Hadamard)**.**
Le rayon de convergence de la série entière $\sum_n c_n z^n$ est donné par une des deux formules

$$\frac{1}{R} = \limsup \sqrt[k]{|a_k|} \tag{15.59}$$

ou

$$\frac{1}{R} = \lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| \tag{15.60}$$

lorsque a_k est non nul à partir d'un certain k.

Le disque $|z - z_0| \leqslant R$ est le **disque de convergence** de la série $\sum_n a_n(z - z_0)^n$. Notons que le critère d'Abel ne dit rien pour les points tels que $|z - z_0| = R$. Il faut traiter ces points au cas par cas. Et le pire, c'est qu'une série donnée peut converger pour certain des points sur le bord du disque, et diverger en d'autres. Le théorème d'Abel radial (théorème 15.37) nous donnera quelques informations sur le sujet.

Il y a un dessin à la figure 15.1.

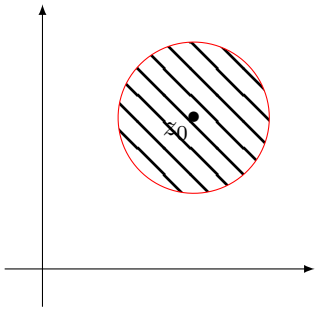


FIGURE 15.1 – À l'intérieur du disque de convergence, la convergence est absolue. En dehors, la série diverge. Sur le cercle proprement dit, tout peut arriver.

Si les suites a_n et b_n sont équivalentes, alors les séries correspondantes auront le même rayon de convergence. Cela ne signifie pas que sur le bord du disque de convergence, elles aient même comportement. Par exemple nous avons

$$\frac{1}{\sqrt{n}} \sim \frac{1}{\sqrt{n}} + \frac{(-1)^n}{n}. \tag{15.61}$$

En même temps, en $z = -1$ la série

$$\sum_{n\geqslant 1} \frac{z^n}{\sqrt{n}} \tag{15.62}$$

converge par le critère des séries alternées (corollaire 10.60). Par contre la série

$$\sum_{n\geqslant 1} \left(\frac{1}{\sqrt{n}} + \frac{(-1)^n}{n}\right) z^n \tag{15.63}$$

ne converge pas pour $z = -1$.

Exemple 15.30
Soit $\alpha \in \mathbb{R}$ et considérons la série $\sum_{n\geqslant 1} a_n z^n$ où a_n est la n-ième décimale de α. Si α est un nombre décimal limité, la suite (a_n) est finie et le rayon de convergence est infini. Sinon, pour tout N il existe un $n > N$ tel que $a_n \neq 0$ et la suite (a_n) ne tend pas vers zéro. Par conséquent la série

$$\sum_n a_n z^n \tag{15.64}$$

diverge pour $z = 1$ et le rayon de convergence satisfait $R \leqslant 1$. Nous avons aussi $|a_n| \leqslant 9$, de telle manière à ce que la série soit bornée et par conséquent majorée en module par $9z^n$, ce qui signifie que $R \geqslant 1$.

Nous déduisons alors $R = 1$. △

15.3.2 Propriétés de la somme

Théorème 15.31.
Soient $\sum_n a_n z^n$ et $\sum b_n z^n$ deux séries de rayon de convergences respectivement R_a et R_b.

(1) Si R_s est le rayon de convergence de $\sum_n (a_n + b_n)z^n$, nous avons

$$R_s \geqslant \min\{R_a, R_b\} \tag{15.65}$$

et nous avons l'égalité si pour tout $|z| \leqslant \min\{R_a, R_b\}$, $\sum(a_n + b_n)z^n = \sum_n a_n z^n + \sum_n b_n z^n$.

(2) Si $\lambda \neq 0$ la série $\sum_n (\lambda a_n)z^n$ a le même rayon de convergence que la série $\sum_n a_n z^n$ et si $|z| < R_a$ nous avons

$$\sum_{n=0}^{\infty} (\lambda a_n)z^n = \lambda \sum_{n=0}^{\infty} a_n z^n. \tag{15.66}$$

*(3) Le **produit de Cauchy** des deux séries est donné par*

$$\left(\sum_n a_n z^n\right)\left(\sum_k b_k z^k\right) = \sum_{n=0}^{\infty}\left(\sum_{i+j=n} a_i b_j\right) z^n = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_k b_{n-k}\right) z^n. \tag{15.67}$$

Si R_p est le rayon de convergence de ce produit nous avons

$$R_p \geqslant \min\{R_a, R_b\} \tag{15.68}$$

et si $|z| < \min\{R_a, R_b\}$ alors

$$\sum_{n=0}^{\infty}\left(\sum_{i+j=n} a_i b_j\right) z^n = \left(\sum_{n=0}^{\infty} a_n z^n\right)\left(\sum_{n=0}^{\infty} b_n z^n\right). \tag{15.69}$$

Démonstration. Nous prouvons la partie sur le produit de Cauchy. En utilisant la propriété du produit de la somme par un scalaire nous avons

$$\left(\sum_{n=0}^{\infty} a_n z^n\right)\left(\sum_{m=0}^{\infty} b_m z^m\right) = \sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty} b_m a_n z^{m+n}\right) \tag{15.70a}$$

$$= \lim_{N\to\infty}\lim_{M\to\infty}\sum_{n=0}^{N}\sum_{m=0}^{M} b_m a_n z^{m+n} \tag{15.70b}$$

$$= \lim_{N\to\infty}\lim_{M\to\infty}\sum_{k=0}^{N+M}\sum_{i+k=k} b_i a_j z^k \tag{15.70c}$$

$$= \lim_{N\to\infty}\sum_{k=0}^{\infty}\sum_{i+k=k} b_i a_j z^k \tag{15.70d}$$

$$= \sum_{k=0}^{\infty}\sum_{i+j=k} b_i a_j z^k. \tag{15.70e}$$

□

Exemple 15.32
Montrons un produit de Cauchy dont le rayon de convergence est strictement plus grand que le minimum. D'abord nous considérons

$$A = 1 - z, \tag{15.71}$$

c'est à dire $a_0 = 1$, $a_1 = -1$, $a_{n\geqslant 2} = 0$ avec $R_a = \infty$. Ensuite nous considérons

$$B = \sum_n z^n, \tag{15.72}$$

c'est à dire $B = (1-z)^{-1}$ et $R_b = 1$. Le produit de Cauchy de ces deux séries valant 1, le rayon de convergence est infini. △

Exemple 15.33
Nous montrons que

$$\sum_{n=0}^{\infty}(n+1)x^n = \frac{1}{(1-x)^2} \tag{15.73}$$

pour $x \in]-1,1[$.

Étant donné que pour tout r dans $]-1,1[$ la suite $(n+1)r^n$ est bornée, le rayon de convergence est correct. Pour les x dans ce domaine nous avons

$$\frac{1}{(1-x)^2} = \frac{1}{(1-x)}\frac{1}{(1-x)} = \left(\sum_{n=0}^{\infty}x^n\right)\left(\sum_{m=0}^{\infty}z^m\right). \tag{15.74}$$

Nous devons expliciter ce produit de Cauchy en utilisant le théorème 15.31. Pour tout i nous avons $a_i = b_i = 1$. Par conséquent le produit (15.74) devient

$$\sum_{n=0}^{\infty}\sum_{i+j=n}x^n = \sum_{n=0}^{\infty}(n+1)x^n. \tag{15.75}$$

△

Théorème 15.34.
Une série entière converge normalement sur tout disque fermé inclus au disque de convergence.

Démonstration. Toute boule fermée incluse à $B(0,R)$ est incluse à la boule $\overline{B(0,r)}$ pour un certain $r < R$. Nous nous concentrons donc sur une telle boule fermée.

Pour chaque n nous posons $u_n(z) = a_n z^n$ que nous voyons comme une fonction sur $\overline{B(0,r)}$. Pour tout $n \in \mathbb{N}$ et tout $z \in \overline{B(0,r)}$ nous avons

$$\|u_n\|_\infty \leqslant |a_n z^n| \leqslant |a_n| r^n. \tag{15.76}$$

Étant donné que $r < R$ la série $\sum_n |a_n| r^n$ converge et la série $\sum_n \|u_n\|$ est convergente. La série $\sum_n a_n z^n$ est alors normalement convergente. □

Exemple 15.35
Encore une fois nous n'avons pas d'informations sur le comportement au bord. Par exemple la série $\sum_n z^n$ a pour rayon de convergence $R = 1$, mais $\sup_{z\in B(0,1)}|z^n| = 1$ et nous n'avons pas de convergence normale sur la boule fermée. △

La convergence normale n'est donc pas de mise sur tout l'intérieur du disque de convergence. La continuité, par contre est effective sur la boule. En effet si $z_0 \in B(0,R)$ alors il existe un rayon $0 < r < R$ tel que $B(z_0,r) \subset B(0,R)$. Sur $B(z_0,r)$ nous avons convergence normale et donc continuité en z_0.

La différence est que la continuité est une propriété locale tandis que la convergence normale est une propriété globale.

Proposition 15.36.
Soit $f(z) = \sum_n a_n z^n$ avec un rayon de convergence R. Si $\sum |a_n| R^n$ converge alors

(1) la série $\sum_n a_n z^n$ converge normalement sur $\overline{B(0,R)}$,

(2) f est continue sur $\overline{B(0,R)}$.

Démonstration. La conclusion est claire dans l'intérieur du disque de convergence. En ce qui concerne le bord, chacune des sommes partielles est une fonction continue. De plus nous avons $\|u_n\| \leqslant |a_n| R^n$, dont la série converge. Par conséquent nous avons convergence normale sur le disque fermé. □

Le théorème suivant permet de donner, dans le cas de fonctions réelle, des informations sur la convergence en une des deux extrémités de l'intervalle de convergence.

Théorème 15.37 (Convergene radiale de Abel)**.**
Soit $f(x) = \sum_n a_n x^n$ une série réelle de rayon de convergence $0 < R < \infty$.

(1) Si $\sum a_n R^n$ converge, alors f est continue sur $[0, R]$.

(2) Si $\sum_n a_n(-R)^n$ converge, alors f est continue sur $[-R, 0]$.

L'exemple 15.63 donnera un exemple d'utilisation pour la série de $\ln(1-x)$ (qui n'est pas encore définie à ce moment).

Le résultat suivant permet d'identifier deux séries complexes lorsque leurs valeurs sur $\mathbb{R}$ sont identiques.

Proposition 15.38.
Soient les séries $f(z) = \sum a_n z^n$ et $g(z) = \sum b_n z^n$ convergentes dans $B(0, R)$. Si $f(x) = g(x)$ pour $x \in [0, R[$ alors $a_n = b_n$.

Démonstration. Soit n_0 le plus petit entier tel que $a_{n_0} \neq b_{n_0}$. Pour tout $z \in B(0, R)$ nous avons

$$f(z) - g(z) = \sum_{n=n_0}^{\infty} (a_n - b_n) z^n = z^{n_0} \varphi(z) \tag{15.77}$$

où

$$\varphi(z) = \sum_{n \geqslant 0} (a_{n+n_0} - b_{n+n_0}) z^n. \tag{15.78}$$

Par le théorème 15.31(1) le rayon de convergence de φ est plus grand que R et la fonction φ est continue en 0. Étant donné que $\varphi(0) = a_{n_0} - b_{n_0} \neq 0$ et que φ est continue nous avons un ρ tel que $\varphi \neq 0$ sur $B(0, \rho)$. Or cela n'est pas possible parce que au moins sur la partie réelle de cette dernière boule, φ doit être nulle. □

Proposition 15.39 ([191, 1])**.**
Si la série entière $\sum_{n \geqslant 0} a_n z^n$ a un rayon de convergence R alors

(1) La somme est une fonction holomorphe dans le disque de convergence.

(2) La somme est différentiable et

$$du_{z_0}(z) = \sum_{n=1}^{\infty} n a_n z_0^{n-1} z. \tag{15.79}$$

(3) De plus pour tout $z_0 \in B(0, R)$, on pose[4]

$$S(z) = \sum_{n \geqslant 0} a_n z^n \tag{15.80a}$$

$$T(z) = \sum_{n \geqslant 1} n a_n z^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} z^n. \tag{15.80b}$$

Alors nous avons

$$\lim_{z \to z_0} \frac{S(z) - S(z_0)}{z - z_0} = T(z_0). \tag{15.81}$$

Démonstration. Nous allons prouver, en utilisant le théorème 15.22, que la somme est une fonction différentiable et que la différentielle est $\mathbb{C}$-linéaire. La proposition 15.7 nous dira alors que la somme est $\mathbb{C}$-dérivable.

4. Pour rappel, dans tout ce texte, $B(a, r)$ est une boule *ouverte*.

Nous posons $u_n(z) = a_n z^n$, qui est une fonction de classe C^1. En ce qui concerne sa différentielle nous considérons $z_0 \in B(0,R)$ et nous avons (si $n = 0$ alors la différentielle est nulle)

$$(du_n)_{z_0}(z) = \frac{d}{dt}\Big[u_n(z_0 + tz)\Big]_{t=0} \tag{15.82a}$$
$$= \frac{d}{dt}\Big[a_n(z_0 + tz)^n\Big]_{t=0} \tag{15.82b}$$
$$= \frac{d}{dt}\Big[na_n(z_0^{n-1}tz)\Big]_{t=0} \tag{15.82c}$$
$$= na_n z_0^{n-1} z. \tag{15.82d}$$

En cours de calcul nous avons développé $(z_0 + tz)^n$ et gardé seulement les termes de degré 1 en t. Il y en a n et ils sont tous égaux à $z_0^{n-1}tz$.

La convergence simple $\sum_n u_n$ est dans les hypothèses. Il reste à prouver que la somme des différentielles converge uniformément sur tout compact autour de z_0 ne débordant pas du disque ouvert de convergence. Soit K un compact autour de z_0. Dans le calcul suivant nous utilisons une première fois la norme uniforme de du_n vu comme fonction de K vers $\mathcal{L}(\mathbb{C},\mathbb{C})$ et une fois la norme opérateur [5] de $(du_n)_{z_0}$ comme application linéaire $\mathbb{C} \to \mathbb{C}$:

$$\|du_n\|_k = \sup_{z_0 \in K} \|(du_n)_{z_0}\| \tag{15.83a}$$
$$= \sup_{z_0 \in K} \sup_{|z|=1} \tag{15.83b}$$
$$= \sup_{z_0 \in K} \sup_{|z|=1} |na_n z_0^{n-1} z| \tag{15.83c}$$
$$= \sup_{z_0 \in K} n|a_n||z_0|^{n-1}. \tag{15.83d}$$

Vu que $z \mapsto |z|^{n-1}$ est une application continue sur le compact K, elle atteint son maximum (théorème 6.149.), nous considérons z_K, un point qui réalise le supremum. Ce nombre est dans le disque de convergence parce que K est un compact autour de z_0.

Nous devons prouver que $\sum_n n|a_n||z_K|^{n-1}$ converge. Vu que $|z_K|$ est une constante (par rapport à n) nous pouvons étudier la convergence en écrivant $|z_K|^n$ au lieu de $|z_K|^{n-1}$.

La suite $(a_n|z_K|^n)$ est une suite bornée. Soit M tel que $|a_n||z_K|^n < M$ pour tout n. Nous considérons de plus r de telle sorte que $K \subset B(0,r) \subset B(0,R)$. En particulier $|z_K| < r$ et nous avons

$$n|a_n||z_K|^n \leqslant n|a_n|r^n \left(\frac{|z_K|}{r}\right)^n \leqslant nM\left(\frac{|z_K|}{r}\right)^n. \tag{15.84}$$

Nous savons que ce qui est dans la parenthèse est plus petit que 1, mais que $\sum_n nx^n$ converge dès que $|x| < 1$. Par conséquent

$$\sum_n \|du_n\|_K \tag{15.85}$$

converge et le théorème 15.22 fonctionne : $du = \sum_{n=1}^{\infty} du_n$ et la somme $\sum_n u_n$ est de classe C^1.

La différentielle de $\sum_n u_n$ s'exprime explicitement par

$$du_{z_0}(z) = \sum_{n=1}^{\infty} na_n z_0^{n-1} z. \tag{15.86}$$

Cette forme montre que du_{z_0} est une application $\mathbb{C}$-linéaire et donc la somme est $\mathbb{C}$-dérivable par la proposition 15.7. Ergo holomorphe sur le disque de convergence par définition 15.2.

En ce qui concerne la formule (15.81), elle provient de la formule (15.13) : $f'(z_0)$ est donné par la facteur multiplicatif de du_{z_0}. En l'occurrence la formule (15.86) nous donne

$$f'(z_0) = \sum_{n \geqslant 1} na_n z_0^{n-1}. \tag{15.87}$$

□

5. Définition 10.5.

15.3.3 Dérivation

Lemme 15.40.
Soit une série entière $\sum a_n z^n$ de rayon de convergence R. Les séries

$$\sum \frac{a_n}{n+1} z^{n+1} \tag{15.88}$$

et

$$\sum_{n\geqslant 1} n a_n z^{n-1} \tag{15.89}$$

ont même rayon de convergence R.

Notons toutefois que nonobstant ce lemme, les séries dont il est question peuvent se comporter différemment sur le bord du disque de convergence. En effet la série

$$\sum \frac{1}{n} z^n \tag{15.90}$$

diverge pour $z = 1$ alors que

$$\sum \frac{1}{n(n+1)} z^{n+1} \tag{15.91}$$

converge pour $z = 1$.

Les théorèmes de dérivation et d'intégration de séries de fonctions (théorèmes 15.18 et 15.20) fonctionnent bien dans le cas des séries entières. Ils donnent la proposition 15.41 pour la dérivation et 15.46 pour l'intégration.

Proposition 15.41.
Soit la série entière

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \tag{15.92}$$

de rayon de convergence R. Alors la fonction f est C^1 sur $]-R, R[$ et se dérive terme à terme :

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} \tag{15.93}$$

pour tout $x \in]-R, R[$.

Démonstration. Nous savons que la série $\sum_{n=1}^{\infty} n a_n x^{n-1}$ a le même rayon de convergente que celui de la série f. En particulier cette série des dérivées converge normalement sur tout compact dans $]-R, R[$ et la somme est continue. Le théorème 15.20 conclu. □

Remarque 15.42.
À part lorsqu'on parle de fonction $\mathbb{R} \to \mathbb{R}$, la notion de classe C^k s'entend au sens de la différentielle, et non de la dérivée, voir les définitions 12.251. C'est cela qui explique la structure de la démonstration de la proposition 15.39.

Corollaire 15.43 ([191, 1])**.**
La somme d'une série entière est de classe C^∞ sur le disque ouvert de convergence.

Démonstration. La proposition 15.39 a démontré en réalité nettement plus : sur le disque ouvert de convergence, la somme est une fonction holomorphe. Il est n'est cependant pas possible de conclure ainsi parce que le fait qu'une fonction holomorphe est C^∞ ne sera démontré qu'au coût de nombreux efforts dans le théorème 25.15(3).

Cas réel Nous considérons la série entière $\sum_n a_n x^n$ pour $x \in \mathbb{R}$ de rayon de convergence R. Une simple récurrence sur la proposition 15.41 donne le résultat.

Cas complexe Attention : le fait d'être de classe C^k est le fait d'être k fois *différentiable*. Rien à voir avec la $\mathbb{C}$-dérivabilité.

En ce qui concerne la différentiabilité nous avons la proposition 15.39 qui dit que dans le disque de convergence, la fonction $u(z) = \sum_n a_n z^n$ a pour différentielle l'application $du\colon \mathbb{C} \to \mathcal{L}_{\mathbb{C}}(\mathbb{C}, \mathbb{C})$,

$$du_{z_0}(z) = \big(\sum_{n=0}^{\infty}(n+1)a_{n+1}z_0^n\big)z. \tag{15.94}$$

Nous allons éviter de considérer la différentielle seconde comme une application

$$d^2u\colon \mathbb{C} \to \mathcal{L}\big(\mathbb{C}, \mathcal{L}(\mathbb{C}, \mathbb{C})\big) \tag{15.95}$$

parce que ça nous mènerait trop loin pour parler de la différentielle k^{e}. Au lieu de cela nous allons considérer l'isomorphisme d'espace vectoriel

$$\begin{aligned}\psi\colon \mathbb{C} &\to \mathcal{L}_{\mathbb{C}}(\mathbb{C}, \mathbb{C})\\ z_0 &\mapsto \psi(z_0)z = z_0 z.\end{aligned} \tag{15.96}$$

Dans cette optique nous écrivons :

$$du_{z_0} = \psi\big(\sum_{n=0}^{\infty}(n+1)a_{n+1}z_0^n\big) \tag{15.97}$$

ou encore :

$$(\psi^{-1}\circ d)u(z_0) = \sum_{n\geqslant 0}(n+1)a_{n+1}z_0^n. \tag{15.98}$$

Nous allons prouver par récurrence que l'égalité suivante est vraie (y compris le fait que la somme converge) :

$$(\psi^{-1}\circ d)^k u(z_0) = \sum_{n=0}^{\infty}\frac{(n+k)!}{n!}a_{n+k}z_0^n. \tag{15.99}$$

Prouvons d'abord que cette somme converge pour tout k. Nous avons $(n+k)!/n! < (n+k)^k$ et donc il suffit de prouver que la série de coefficients $n^k a_n$ converge. C'est le cas par la proposition 15.27.

Nous pouvons calculer la différentielle de $(\psi^{-1}\circ d)^k u$ en dérivant terme à terme en utilisant (encore) la proposition 15.39(2) :

$$d\big((\psi^{-1}\circ d)^k u\big)_{z_0}(z) = \sum_{n=1}^{\infty}\frac{(n+k)!}{n!}a_{n+k}na_0^{n-1}z \tag{15.100a}$$

$$= \sum_{n=0}^{\infty}\frac{(n+k+1)!}{n!}a_{n+k+1}z_0^n z. \tag{15.100b}$$

Nous appliquons ψ^{-1} à cela :

$$(\psi^{-1}\circ d)^{k+1}u(z_0) = \sum_{k=0}^{\infty}\frac{(n+k+1)!}{n!}a_{n+k+1}z_0^n. \tag{15.101}$$

Dérouler à l'envers Nous allons maintenant utiliser la proposition 12.203 pour montrer que u est de classe C^k pour tout k. Nous avons démontré que $(\psi^{-1}\circ d)^k u$ était différentiable. Par conséquent, $d\big((\psi^{-1}\circ d)^{k-1}u\big)$ est différentiable et donc $(\psi^{-1}\circ d)^{k-1}$ est de classe C^1. En continuant ainsi, $(\psi^{-1}\circ d)^{k-l}u$ est de classe C^l et u est de classe C^k.

□

Le lemme suivant est encore essentiellement valable dans un espace de Banach (proposition 10.52).

Lemme 15.44.
La série entière $\sum_{n\geqslant 0} z^{nk}$ a un rayon de convergence 1 et converge vers la fonction

$$\sum_{n\geqslant 0} z^{nk} = \frac{1}{1-z^k}. \tag{15.102}$$

Lorsque $|\omega| = 1$ nous avons aussi un rayon de convergence 1 pour la série

$$\frac{1}{\omega - z} = \sum_{k\geqslant 0} \omega^{-k-1} z^k. \tag{15.103}$$

Sous les mêmes hypothèses sur ω nous avons encore la série

$$\frac{1}{(\omega - z)^k} = \frac{1}{(k-1)!} \sum_{s=0}^{\infty} \omega^{-s-1-k} \frac{(s+k-1)!}{s!} z^s \tag{15.104}$$

Démonstration. Les coefficients de la série sont $a_n = 1$ lorsque n est multiple de k et $a_n = 0$ autrement. Donc pour $r = 1$ la suite $r^n a_n$ reste bornée[6]. Cela prouve que le rayon de convergence est au moins 1. Par ailleurs si $r > 1$ alors clairement la suite $(a_n r^n)$ n'est pas bornée. Cela prouve le rayon de convergence égal à 1.

Soit donc $z \in B(0,1)$. Nous avons

$$\left(\sum_{n\geqslant 0} z^{nk} \right) (1 - z^k) = \sum_{n\geqslant 0} z^{nk} - \sum_{n\geqslant 0} z^{(n+1)k}. \tag{15.105}$$

Le premier terme de la première somme vaut 1 tandis que tous les autres termes s'annulent deux à deux.

En ce qui concerne la série (15.103), elle s'obtient facilement :

$$\frac{1}{\omega - z} = \frac{1}{\omega} \frac{1}{1 - \frac{z}{\omega}} = \frac{1}{\omega} \sum_{s=0}^{\infty} \left(\frac{z}{\omega} \right)^s = \sum_s \omega^{-s-1} z^s. \tag{15.106}$$

La troisième série s'obtient en dérivant la seconde, ce qui est permis dans le disque de convergence par la proposition 15.41. □

Remarque 15.45.
Sur le bord du disque de convergence, la série $\sum_n z^{nk}$ ne converge pas. En effet le rayon étant 1, sur le bord nous avons la série $\sum_n e^{ink\theta}$ dont la norme du terme général ne tend pas vers zéro.

15.3.4 Intégration

Proposition 15.46.
Soit la série entière $\sum a_n x^n$ de rayon de convergence R.

(1) Pour tout segment $[a,b] \subset]-R,R[$ nous pouvons intégrer terme à terme :

$$\int_a^b \left(\sum_{n=0}^{\infty} a_n x^n \right) dx = \sum_{n=0}^{\infty} a_n \int_a^b x^n dx. \tag{15.107}$$

(2) La série entière obtenue en intégrant terme à terme a le même rayon de convergence que celui de la série de départ.

Démonstration. La première assertion est un cas particulier du théorème général 15.18. Pour le rayon de convergence, le lemme 15.40 fait le travail. □

6. Utilisation directe de la définition 15.24.

15.4 Dénombrement des solutions d'une équation diophantienne

Théorème 15.47 ([34, 157]).
Soient des entiers naturels premiers dans leur ensemble[7] $\alpha_1, \dots, \alpha_p$ *et l'équation*

$$\alpha_1 n_1 + \cdots + \alpha_p n_p = n \tag{15.108}$$

pour les naturels n_i où n est un naturel donné. Nous notons S_n le nombre de solutions de cette équation. Alors :

(1) Il existe un algorithme (en temps fini) pour calculer S_n en fonction des α_i et de n.

(2) Nous avons le comportement asymptotique

$$S_n \sim \frac{1}{\alpha_1 \dots \alpha_p} \frac{n^{p-1}}{(p-1)!}. \tag{15.109}$$

Démonstration. Pour $|z| < 1$ dans $\mathbb{C}$, utilisant le lemme 15.44, nous écrivons le développement

$$F(z) = \prod_{i=1}^{p} \frac{1}{1 - z^{\alpha_i}} = \prod_{i=1}^{p} \sum_{n \geqslant 0} z^{n\alpha_i}. \tag{15.110}$$

Nous allons maintenant à la pêche au terme de degré k dans ce produit de sommes en utilisant p fois le produit de Cauchy de la formule (15.67). Nous avons

$$F(z) = \sum_{k \geqslant 0} \left(\sum_{n_1\alpha_1 + \cdots + n_p\alpha_p = n} 1 \right) z^k = \sum_{k \geqslant 0} S_k z^k. \tag{15.111}$$

La technique pour déterminer la valeur de S_n est alors de développer $F(z)$ en série de façon un peu explicite et d'identifier le coefficient de z^n parce que nous venons de voir que ce coefficient est S_n. Nous commençons par une décomposition en éléments simples, expliquée autour de l'équation (14.582) :

$$\frac{1}{1 - z^{\alpha_i}} = \sum_{\alpha \in U_{\alpha_i}} \frac{A_{\omega,i}}{\omega - z}. \tag{15.112}$$

où U_{α_i} est le groupe des racines $\alpha_i$ede l'unité décrit en 3.14. La raison de ce développement est que, comme mentionné dans le lemme 5.117, $\prod_{\omega \in \mathrm{U}_{\alpha_i}} (z - \omega) = z^{\alpha_1} - 1$. Lorsque nous effectuons la somme, le dénominateur commun est donc bien[8] $1 - z^{\alpha_i}$. En récrivant le produit :

$$F(z) = \prod_{i=1}^{p} \frac{1}{1 - z^{\alpha_i}} = \prod_{i=1}^{p} \sum_{\omega \in U_{\alpha_i}} \frac{A_{\omega,i}}{\omega - z} \tag{15.113}$$

Les coefficients $A_{\omega,i}$ sont calculables explicitement, en temps fini.

Vu que 1 est dans tous les U_{α_i}, le produit fait intervenir au dénominateur des puissances de $(1-z)$ jusqu'à la puissance p. Les autres racines de l'unité appartiennent au maximum à $p-1$ des groupes U_{α_i} parce que les nombres α_i sont premiers dans leur ensemble, voir la proposition 3.102.

La fonction F peut alors s'écrire sous la forme

$$F(z) = \frac{A}{(1-z)^p} + G(z) \tag{15.114}$$

où $G(z)$ est une somme de termes de la forme

$$\frac{a_{i,1}}{1 - \omega_i} + \cdots + \frac{a_{i,p}}{(1 - \omega_i)^{p-1}} \tag{15.115}$$

7. Définition 3.30.

8. Pour le signe, c'est ajustable avec le signe de $A_{\omega,i}$.

où les ω_i sont les racines $\alpha_i$eme de l'unité et $a_{k,r}$ sont des nombres complexes. Trouvons A. D'abord grâce au lemme 4.125(2) nous avons

$$F(z)(1-z)^p = \prod_{l=1}^{p} \frac{1-z}{1-z^{\alpha_i}} = \prod_{i=1}^{p} \frac{1}{1+z+\cdots+z^{\alpha_i-1}}, \tag{15.116}$$

et donc

$$\lim_{z\to 1} F(z)(1-z)^p = \prod_{i=1}^{p} \frac{1}{\alpha_i}. \tag{15.117}$$

Mais vu ce que contient $G(z)$, nous avons aussi $\lim_{z\to 1} F(z) = A$. Nous avons donc déjà déterminé $A = \frac{1}{\alpha_1 \ldots \alpha_p}$.

Pour la suite nous avons besoin des développements du lemme 15.44. Nous utiliserons en particulier celle-ci :

$$\frac{1}{(\omega - z)^k} = \frac{1}{(k-1)!} \sum_{s=0}^{\infty} \omega^{-s-1-k} \frac{(s+k-1)!}{s!} z^s. \tag{15.118}$$

En particulier le module du coefficient de z^n là dedans est : $\frac{(n+k-1)!}{n!(k-1)!}$. Dans la partie G de la décomposition (15.114), k est majoré par $p-2$ et la dépendance en n est donc au maximum du type

$$\frac{(n+p-2)!}{n!(p-2)!} \sim \frac{n^{n+p-2}}{n^n(p-2)!} = \frac{n^{p-2}}{(p-2)!}. \tag{15.119}$$

Dans le premier terme par contre, il y a des termes jusqu'à $k = p$. Le terme dominant est alors en $\frac{n^{p-1}}{(p-1)!}$ et son coefficient est A qui est déjà calculé. Au final le terme dominant du coefficient de z^n dans $F(z)$ est

$$S_n \sim \frac{A}{(p-1)!} n^{p-1} = \frac{1}{\alpha_1 \ldots \alpha_p} \frac{n^{p-1}}{(p-1)!}. \tag{15.120}$$

□

Exemple 15.48
Pour $p = 1$, l'équation est $\alpha x = n$, qui possède au maximum une solution, quel que soit n. Et de plus pour avoir une solution il faut et suffit que α divise n, c'est à dire que n soit un multiple de α. Il n'y a que un nombre sur α à être multiple de α. D'où le comportement en $\frac{1}{\alpha}$.

Pour $p = 2$, c'est l'équation (3.74) déjà étudiée. Il y a une famille à un paramètre de solutions dont seulement un certain nombre sont positives. A priori, le nombre de solutions positives croît linéairement en n. △

15.5 Exponentielle et logarithme

La méthode adoptée ici est la suivante :
— L'exponentielle est définie par la série.
— Nous démontrons qu'elle vérifie l'équation différentielle $y' = y$, $y(0) = 1$.
— Nous démontrons l'unicité de la solution à cette équation différentielle.
— Nous démontrons qu'elle est égale à $x \mapsto y(1)^x$. Cela donne la définition du nombre e comme valant $y(1)$.
— Nous définissons le logarithme comme l'application réciproque de l'exponentielle (définition 15.57).
— Les fonctions trigonométriques (sinus et cosinus) sont définies par leurs séries. Il est alors montré que $e^{ix} = \cos(x) + i\sin(x)$.

Théorème 15.49 (Existence de l'exponentielle)**.**
La série entière

$$y(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \tag{15.121}$$

définit une fonction dérivable solution de

$$\begin{cases} y' = y & \text{(15.122a)} \\ y(0) = 1. & \text{(15.122b)} \end{cases}$$

Démonstration. La formule de Hadamard (théorème 15.29) donne le rayon de convergence de la série (15.121) par

$$\frac{1}{R} = \lim_{k\to\infty} \frac{\frac{1}{(k+1)!}}{\frac{1}{k!}} = \lim_{k\to\infty} \frac{1}{k+1} = 0. \tag{15.123}$$

Donc nous avons un rayon de convergence infini. La fonction y est définie sur $\mathbb{R}$ et la proposition 15.41 nous dit que y est dérivable. Nous pouvons aussi dériver terme à terme :

$$y'(x) = \sum_{k=0}^{\infty} \frac{kx^{k-1}}{k!} = \sum_{k=1}^{\infty} \frac{kx^{k-1}}{k!} = \sum_{k=1}^{\infty} \frac{x^{k-1}}{(k-1)!} = \sum_{k=0}^{\infty} \frac{x^k}{k!} = y(x). \tag{15.124}$$

Notez le petit jeu d'indice de départ de k. Dans un premier temps, nous remarquons que $k = 0$ donne un terme nul et nous le supprimons, et dans un second temps nous effectuons la simplification des factorielles (qui ne fonctionne pas avec $k = 0$). □

Pour la suite nous notons y une solution de l'équation $y' = y$, $y(0) = 1$, et nous allons en donner des propriétés indépendamment de l'existence, donnée par le théorème 15.49.

Proposition 15.50.
Quelques propriétés de y (si elle existe) :

(1) Pour tout $x \in \mathbb{R}$ nous avons $y(x)y(-x) = 1$.

(2) $y(x) > 0$ pour tout x.

(3) y est strictement croissante.

Démonstration. Nous posons $\varphi(x) = y(x)y(-x)$ et nous dérivons :

$$\varphi'(x) = y'(x)y(-x) - y(x)y'(-x) = 0. \tag{15.125}$$

Donc φ est constante[9]. Vu que $\varphi(0) = 1$ nous avons automatiquement $y(x)y(-x) = 1$ pour tout x.

Les deux autres allégations sont simples : si $y(x_0) < 0$ alors il existe $t \in]x_0, 1[$ tel que $y(t) = 0$, ce qui est impossible parce que $y(t)y(-t) = 1$. La stricte croissance de y s'ensuit. □

Proposition 15.51 (Unicité de l'exponentielle)**.**
Si elle existe, la solution au problème

$$\begin{cases} y' = y & \text{(15.126a)} \\ y(0) = 1 & \text{(15.126b)} \end{cases}$$

est unique.

9. Proposition 12.125.

Démonstration. Soient y et g deux solutions et considérions la fonction $h(x) = g(x)y(-x)$. Un calcul immédiat donne

$$h'(x) = 0 \tag{15.127}$$

et donc h est constante. Vu que $h(0) = 1$ nous avons $g(x)y(-x) = 1$ pour tout x, c'est à dire

$$g(x) = \frac{1}{y(-x)} = y(x). \tag{15.128}$$

□

Proposition 15.52.
Quelques formules pour tout $a, b \in \mathbb{R}$ et $n \in \mathbb{Z}$:

(1) $y(a + b) = y(a)y(b)$
(2) $y(na) = y(a)^n$
(3) $y\left(\frac{a}{n}\right) = \sqrt[n]{y(a)}$.

Démonstration. Nous posons $h(x) = y(a + b - x)y(x)$ et nous avons encore $h'(x) = 0$ dont nous déduisons que h est constante. De plus

$$h(0) = y(a + b)y(0) = y(a + b) \tag{15.129}$$

et

$$h(b) = y(a)y(b). \tag{15.130}$$

Vu que h est constante, ces deux expressions sont égales : $y(a + b) = y(a)y(b)$.

Forts de cette relation, une récurrence donne $y(na) = y(a)^n$ pour tout $n \in \mathbb{N}$. De plus

$$y(a) = y\left(\frac{a}{n} \times n\right) = y\left(\frac{a}{n}\right)^n, \tag{15.131}$$

ce qui donne $y(a) = y(a/n)^n$ ou encore $y(a/n) = \sqrt[n]{y(a)}$.

Enfin pour les négatifs, si $n \in \mathbb{N}$,

$$y(-na) = \frac{1}{y(na)} = \frac{1}{y(a)^n} = y(a)^{-n}. \tag{15.132}$$

Et de la même façon,

$$y\left(-\frac{a}{n}\right) = \frac{1}{y\left(\frac{a}{n}\right)} = \sqrt[n]{\frac{1}{y(a)}} = \sqrt[-n]{y(a)}. \tag{15.133}$$

□

Proposition 15.53.
Pour tout $x \in \mathbb{R}$, nous avons

$$y(x) = y(1)^x. \tag{15.134}$$

Démonstration. Si $q \in \mathbb{Q}$ alors $q = a/b$ et

$$y(q) = y\left(\frac{a}{b}\right) = y\left(a \times \frac{1}{b}\right) = y\left(\frac{1}{b}\right)^a = \left(\sqrt[b]{y(1)}\right)^a = y(1)^{a/b} = y(1)^q. \tag{15.135}$$

Par ailleurs si $a \in \mathbb{R}$, alors la fonction $x \mapsto a^x$ est continue. Les fonctions y et $x \mapsto y(1)^x$ sont deux fonctions continues égales sur $\mathbb{Q}$. Elles sont donc égales par la proposition 6.142. □

Nous notons $y(1) = e$, le **nombre de Néper**, de telle sorte que

$$y(x) = e^x. \tag{15.136}$$

Une conséquence est que

$$\lim_{x\to-\infty} e^x = 0 \tag{15.137a}$$

$$\lim_{x\to+\infty} e^x = +\infty, \tag{15.137b}$$

et en particulier,

$$\begin{aligned}\exp\colon \mathbb{R} &\to]0,\infty[\\ x &\mapsto e^x\end{aligned} \tag{15.138}$$

est une bijection.

Nous donnons maintenant quelque approximations numériques de e, particulièrement inefficaces.

Lemme 15.54.
Nous avons

$$2 < e < 3. \tag{15.139}$$

Démonstration. Nous savons que $y(0) = 1$ et $y'(0) = 1$. La fonction y est strictement croissante (et donc sa dérivée aussi). Nous avons donc $y'(x) > 1$ pour tout $x \in]0, 1]$, et donc

$$y(1) > 1 + 1 \times 1 = 2. \tag{15.140}$$

Sachant que $2 > y'(x)$ pour tout $x \in]0, 1[$ nous pouvons refaire le coup de l'approximation affine, cette fois en majorant :

$$y(1) < 1 + 2 \times 1 = 3. \tag{15.141}$$

□

De la même façon nous savons que

$$y(\frac{1}{n}) > 1 + \frac{1}{n} \tag{15.142}$$

parce que y' est minoré par 1 sur $]0, \frac{1}{n}[$. Avec cela nous avons aussi la majoration

$$y(\frac{1}{n}) < 1 + \frac{1}{n} \times \left(1 + \frac{1}{n}\right) = 1 + \frac{1}{n} + \frac{1}{n^2}. \tag{15.143}$$

Et enfin nous pouvons donner l'encadrement, valable pour tout n :

$$\left(1 + \frac{1}{n}\right)^n < y(1) < \left(1 + \frac{1}{n} + \frac{1}{n^2}\right)^n. \tag{15.144}$$

Pour $n = 10$ nous trouvons

$$2.50 < e < 2.83. \tag{15.145}$$

Bien que ce soit à mon avis humainement pas possible à faire à la main nous avons, pour $n = 100$:

$$2.70 < e < 2.7317 \tag{15.146}$$

Cela reste un encadrement très modeste.

Une méthode plus efficace consiste à calculer directement le développement de définition

$$e = \exp(1) = \sum_{k=0}^{\infty} \frac{1}{n!}. \tag{15.147}$$

```sage
def u(k):
    """
    return the kth term in the expansion of 'e'
    """
    return 1/factorial(k)

def sum_u(n):
    """
    return the sum of the 'n' first terms, that is with
    k froim 0 to n-1.
    """
    L=[  u(k) for k in range(0,n)  ]
    return sum(L)

s = sum_u(5)          # This is a fraction
print(s)
print( numerical_approx(s)  )
```

tex/sage/sageSnip013.sage

Problèmes et choses à faire

Comment trouver, avec cette méthode, un *encadrement pour e ?*

Ce petit programme, avec 5 termes donne $e \simeq 65/24 \simeq 2.708$. Avouez que c'est déjà bien mieux.

Théorème 15.55 (Définition de l'exponentielle)**.**
Les choses que nous savons sur l'exponentielle :

(1) Il y a unicité de la solution à l'équation différentielle

$$\begin{cases} y' = y & \text{(15.148a)} \\ y(0) = 1. & \text{(15.148b)} \end{cases}$$

(2) L'équation différentielle (15.148) *possède une solution donnée par la série entière*

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \tag{15.149}$$

(3) Cette solution est une bijection $y\colon \mathbb{R} \to]0, \infty[$.

(4) La fonction y ainsi définie est de classe C^∞.

(5) Elle est également donnée par la formule

$$\exp(x) = e^x \tag{15.150}$$

où e est définit par $e = \exp(1)$.

(6) Elle vérifie

$$e^{a+b} = e^a e^b \tag{15.151}$$

Nous nommons **exponentielle** cette fonction.

Démonstration. (1) C'est la proposition 15.51.

(2) C'est le théorème 15.49.

(3) Le rayon de convergence de la série (15.149) est infini (théorème 15.49) ; elle est donc définie sur $\mathbb{R}$. Le fait que ce soit une bijection est dû au fait qu'elle est strictement croissante (proposition 15.50) ainsi qu'aux limite (15.137).

(4) Vu que $y = y'$, y est dérivable. Mais comme y' est alors égale à une fonction dérivable, y' est dérivable. En dérivant l'égalité $y' = y$ nous obtenons $y'' = y'$ et le jeu continue.

(5) C'est la proposition 15.53.

(6) C'est la proposition 15.52(1).

□

Exemple 15.56(Un endomorphisme sans polynôme annulateur[97])
l'exponentielle permet de donner un exemple d'un endomorphisme n'ayant pas de polynôme annulateur [10] : l'endomorphisme de dérivation

$$\begin{aligned} D\colon C^\infty(\mathbb{R},\mathbb{R}) &\to C^\infty(\mathbb{R},\mathbb{R}) \\ f &\mapsto f' \end{aligned} \tag{15.152}$$

n'a pas de polynôme annulateur. En effet supposons que $P = \sum_{k=0}^{p} a_k X^k$ en soit un, et considérons les fonction $f_\lambda\colon t \mapsto e^{\lambda t}$. Nous avons

$$0 = P(D)f_\lambda = \sum_k a_k D^k(f_\lambda) = \sum_k a_k \lambda^k f_\lambda = P(\lambda)f_\lambda. \tag{15.153}$$

Par conséquent λ est une racine de P pour tout $\lambda \in \mathbb{R}$. Cela implique que $P = 0$.

D'ailleurs si on y pense bien, cet exemple n'est qu'un habillage de l'exemple 9.109. △

Proposition-définition 15.57.
L'application $\exp\colon \mathbb{R} \to]0,\infty[$ *est une bijection. L'application réciproque*

$$\ln\colon]0,\infty[\to \mathbb{R} \tag{15.154}$$

est le ***logarithme****.*

Démonstration. Le fonction exponentielle est dérivable, toujours strictement positive, donc strictement croissante. Les limites en $\pm\infty$ sont 0 et $+\infty$. Le théorème des valeurs intermédiaires 12.21 nous dit que c'est une bijection. En effet, l'injectivité est la stricte croissance. En ce qui concerne la surjection, soit $y \in]0,\infty[$. Vu que la limite en $-\infty$ est zéro, il existe $A \in \mathbb{R}$ tel que $\exp(x) < y$ pour tout $x < A$, et de la même façon, il existe $B \in \mathbb{R}$ tel que $\exp(x) > y$ pour tout $x > B$. Si $a < A$ et $b > B$ alors $\exp(a) < y$ et $\exp(b) > y$, donc y est dans l'image de $[a,b]$ par l'exponentielle. □

Proposition 15.58.
Quelque propriétés du logarithme.

(1) Le logarithme est une application dérivable et strictement croissante.

(2) Le logarithme est la primitive de $x \mapsto \frac{1}{x}$ *qui s'annule en* $x = 1$*.*

Démonstration. Elle est donc bijective, d'inverse continue et dérivable par le théorème 12.41 et la proposition 12.42.

La dérivée de la fonction logarithme peut être calculée en utilisant la formule (12.54), mais aussi de façon plus piettone en écrivant l'expression suivante, valable pour tout $x \in \mathbb{R}$:

$$\ln\big(\exp(x)\big) = x, \tag{15.155}$$

que nous pouvons dériver en utilisant le théorème de dérivation des fonctions composées :

$$\ln'\big(\exp(x)\big)\exp'(x) = 1. \tag{15.156}$$

Mais $\exp'(x) = x$, donc

$$\ln'(y) = \frac{1}{y} \tag{15.157}$$

10. Voir la définition 9.105 et ce qui suit.

pour tout y dans l'image de exp, c'est à dire pour tout y dans l'ensemble de définition de ln.

Par ailleurs, $\exp(0) = 1$ donc

$$\ln(1) = \ln\big(\exp(0)\big) = 0. \tag{15.158}$$

En ce qui concerne l'unicité d'une primitive s'annulant en $x = 1$, c'est le corollaire 12.136. □

Lemme 15.59.
Si $u\colon \mathbb{R} \to]0, \infty[$ *est dérivable alors* $\ln(u)' = \dfrac{u'}{u}$.

Démonstration. Cela est une conséquence du théorème de dérivation des fonctions composées : si $g(x) = \ln(u(x))$ alors

$$g'(x) = \ln'\big(u(x)\big)u'(x) = \frac{1}{u(x)}u'(x). \tag{15.159}$$

□

Exemple 15.60(Primitive du logarithme)
La primitive de la fonction logarithme définie en 15.57 nous offre un bon moment d'intégration par partie.

Trouver la primitive de la fonction $x \mapsto \ln(x)$. Pour calculer

$$\int \ln(x)dx \tag{15.160}$$

nous écrivons $\ln(x) = 1 \times \ln(x)$ et nous posons $u' = 1$ et $v = \ln(x)$, c'est à dire

$$\begin{aligned} u' = 1 \quad & v = \ln(x) \\ u = x \quad & v' = \frac{1}{x}. \end{aligned} \tag{15.161}$$

La formule d'intégration par parties (14.499) donne donc

$$\int \ln(x) = x\ln(x) - \int x \times \frac{1}{x} = x\ln(x) - \int 1 = x\ln(x) - x + C, \qquad C \in \mathbb{R}. \tag{15.162}$$

Il est facile de vérifier par un petit calcul que

$$\big(x\ln(x) - x\big)' = \ln(x). \tag{15.163}$$

△

Lemme 15.61.
Si $a, b \in]0, \infty[$ *alors*

$$\ln(ab) = \ln(a) + \ln(b) \tag{15.164}$$

et

$$\ln\left(\frac{1}{b}\right) = -\ln(b). \tag{15.165}$$

Démonstration. Nous posons $f(x) = \ln(ax)$ qui est une fonction dérivable. Alors $f'(x) = \frac{a}{ax} = \frac{1}{x}$. Cette fonction f est donc une primitive de $\frac{1}{x}$ et il existe une constante K telle que

$$f(x) = \ln(x) + K. \tag{15.166}$$

Vu que $\ln(1) = 0$ nous avons $K = f(1) = \ln(a)$. Donc

$$\ln(ax) = \ln(x) + \ln(a). \tag{15.167}$$

En ce qui concerne la seconde formule à démontrer, nous avons

$$\ln(1) = \ln\left(\frac{1}{b}b\right) = \ln\left(\frac{1}{b}\right) + \ln(b). \tag{15.168}$$

Étant donné que $\ln(1) = 0$ nous en déduisons la formule (15.165). □

Exemple 15.62
Montrons que la fonction [11]

$$\begin{aligned} f\colon \mathbb{R}_+\backslash\{0,1\} &\to \mathbb{R} \\ x &\mapsto \frac{\ln(x)}{x-1} \end{aligned} \tag{15.169}$$

admet un prolongement C^∞ sur $\mathbb{R}_+\backslash\{0\}$.

Nous allons étudier la fonction

$$f(x) = \frac{\ln(1+x)}{x} \tag{15.170}$$

autour de $x = 0$. Le logarithme ne pose pas de problèmes à développer dans un voisinage :

$$f(x) = \frac{1}{x}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}x^n \tag{15.171a}$$

$$= \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}x^{n-1} \tag{15.171b}$$

$$= \sum_{n=0}^{\infty}\frac{(-1)^{k}}{k+1}x^{k}. \tag{15.171c}$$

Cette série a un rayon de convergence égal à 1, et donc définit sans problèmes une fonction C^∞ dans un voisinage de $x = 0$. Notons que par convention $x^0 = 1$ même si $x = 0$. △

Exemple 15.63
Montrons que pour tout $x \in]-1, 1[$ nous avons

$$-\ln(1-x) = \sum_{n=1}^{\infty}\frac{x^n}{n}. \tag{15.172}$$

Nous calculerons ensuite la valeur de la série

$$\sum_{n=1}^{\infty}\frac{(-1)^n}{n}. \tag{15.173}$$

La série (15.173) serait $f(-1) = -\ln(2)$ où f est la série de fonctions (15.172). Nous utilisons le théorème de convergence radiale d'Abel (théorème 15.37) pour justifier cette réponse :

$$\sum_{n}\frac{(-1)^n}{n} \tag{15.174}$$

converge. △

15.6 Vitesses de x^α, de l'exponentielle et du logarithme

Lemme 15.64.
Pour tout $\alpha > 0$, il existe N tel que $\ln(n) \leqslant n^\alpha$ pour tout $n \geqslant N$.

Démonstration. En effet, nous avons

$$\lim_{x\to\infty}\frac{x^\alpha}{\ln(x)} = \lim_{x\to\infty}\frac{\alpha x^{\alpha-1}}{1/x} = \lim_{x\to\infty}\alpha x^\alpha = \infty \tag{15.175}$$

quand $\alpha > 0$. □

11. Pour la définition du logarithme, c'est la définition 15.57.

Cela tient également lorsque nous considérons $\ln(x)^p$ au lieu de $\ln(x)$. De cela, nous disons que le logarithme croit moins vite que n'importe quel polynôme.

Lemme 15.65.
L'exponentielle croit plus vite que tout polynôme, et plus vite que que logarithme :

$$\lim_{t\to\infty} e^{-t}(\ln t)^n t^\alpha = 0 \tag{15.176}$$

pour tout n et pour tout α.

Lemme 15.66.
Nous avons aussi la limite utile suivante

$$\lim_{n\to\infty} n^\alpha a^n \tag{15.177}$$

pour tout $\alpha > 0$ et $a < 1$.

Démonstration. En passant à l'exponentielle, pour chaque n nous avons

$$n^\alpha a^n = e^{\alpha\ln(n)+n\ln(a)}. \tag{15.178}$$

Ce qui est dans l'exponentielle est

$$\alpha\ln(n) + n\ln(a) = n\big(\alpha\frac{\ln(n)}{n} + \ln(a)\big). \tag{15.179}$$

Dans la parenthèse, $\ln(a) < 0$ et $\frac{\ln(n)}{n} \to 0$. Donc ce qui est dans l'exponentielle (15.178) tend vers $-\infty$ et au final l'expression demandée tend vers zéro. □

Proposition 15.67.
Pour tout polynôme P et pour tout $a > 0$ la fonction $f(x) = P(x)e^{-ax}$ est intégrable[12] *sur $[0,\infty[$.*

Démonstration. Nous avons $f(x) = P(x)e^{-ax/2}e^{-ax/2}$, et par la vitesse comparée des exponentielles et polynômes, pour un certain $M > 0$ nous pouvons affirmer que $P(x)e^{-ax/2} < 1$ sur $[M, 0[$. Dès lors

$$|f(x)| < e^{-ax/2}, \tag{15.180}$$

qui est intégrable. □

Théorème 15.68.
Soit P, l'ensemble des nombres premiers. Alors la somme $\sum_{p\in P}\frac{1}{p}$ diverge et plus précisément,

$$\sum_{\substack{p\leqslant x\\ p\in P}}\frac{1}{p} \geqslant \ln(\ln(x)) - \ln(2). \tag{15.181}$$

Démonstration. Nous posons

$$S_x = \{q \leqslant x \text{ avec } q \text{ sans facteurs carrés}\} \tag{15.182}$$

et

$$P_x = \{p \in P \text{ tel que } p \leqslant x\}. \tag{15.183}$$

Si

$$K_x = \{(q,m) \text{ tels que } q \text{ n'a pas de facteurs carrés et } qm^2 \leqslant x\}, \tag{15.184}$$

alors nous avons

$$K_x = \bigcup_{q\in S_x}\bigcup_{m\leqslant\sqrt{x/q}}(q,m). \tag{15.185}$$

12. Définition 13.116.

Par définition et par le lemme 3.45 nous avons aussi

$$\{n \leqslant x\} = \{qm^2 \text{ tel que } (q,m) \in K_x\}. \tag{15.186}$$

Tout cela pour décomposer la somme

$$\sum_{n\leqslant x} \frac{1}{n} = \sum_{q\in S_x} \sum_{m\leqslant\sqrt{x/q}} \frac{1}{m^2} \leqslant \sum_{q\in S_x} \frac{1}{q} \underbrace{\sum_{m\geqslant 1} \frac{1}{m^2}}_{=C}. \tag{15.187}$$

Nous avons aussi

$$\prod_{p\in P_x} \left(1+\frac{1}{p}\right) = 1 + \sum_{p\in P_x} \frac{1}{p} + \sum_{\substack{p,q\in P_x\\ p<q}} \frac{1}{pq} + \sum_{\substack{p,q,r\in P_x\\ p<q<r}} \frac{1}{pqr} + \dots \tag{15.188a}$$

$$\geqslant 1 + \sum_{p\in P_x} \frac{1}{p} + \sum_{\substack{p,q\in P_x\\ pq\leqslant x}} \frac{1}{pq} + \sum_{\substack{p,q,r\in P_x\\ pqr\leqslant x}} \frac{1}{pqr} + \dots \tag{15.188b}$$

Les sommes sont finies. Les sommes s'étendent sur toutes les façons de prendre des produits de nombres premiers distincts de telle sorte de conserver un produit plus petit que x ; c'est à dire que les sommes se résument en une somme sur les éléments de S_x :

$$\exp\left(\sum_{p\in P_x} \frac{1}{p}\right) \geqslant \prod_{p\in P_x} \left(1+\frac{1}{p}\right) \geqslant \sum_{q\in S_x} \frac{1}{q}. \tag{15.189}$$

La première inégalité est simplement le fait que $1+u \leqslant e^u$ si $u \geqslant 0$ (directe de la définition 15.55). Les inégalités suivantes proviennent du fait que le logarithme est une primitive de la fonction inverse (proposition 15.58) :

$$\ln(x) \leqslant \sum_{n\geqslant x} \int_n^{n+1} \frac{dt}{t} \leqslant \sum_{n\geqslant x} \frac{1}{n}. \tag{15.190}$$

Nous prolongeons ces inégalités avec les inégalités (15.187) et (15.189) :

$$\ln(x) \leqslant \sum_{n\geqslant x} \frac{1}{n} \leqslant C \sum_{q\in S_x} \frac{1}{q} \leqslant C \leqslant \exp\left(\sum_{p\in P_x} \frac{1}{p}\right). \tag{15.191}$$

En passant au logarithme,

$$\ln\big(\ln(x)\big) \leqslant \ln(C) + \sum_{p\in P_x} \frac{1}{p}. \tag{15.192}$$

Ceci montre la divergence de la série de droite. Nous cherchons maintenant une borne pour C. Pour cela nous écrivons

$$\sum_{n=1}^{N} \frac{1}{n^2} \leqslant 1 + \sum_{n=2} \frac{1}{n(n-1)} \tag{15.193a}$$

$$= 1 + \sum_{n=2}^{N} \left(\frac{1}{n-1} - \frac{1}{n}\right) \tag{15.193b}$$

$$= 1 + 1 - \frac{1}{N} \tag{15.193c}$$

$$\leqslant 2. \tag{15.193d}$$

Donc $C \leqslant 2$. □

Ce théorème prend une nouvelle force en considérant le théorème de Müntz 18.16 qui dit qu'alors l'ensemble $\text{Span}\{x^p \text{ tel que } p \text{ est premier}\}$ est dense dans les fonctions continues sur $[0,1]$ muni de la norme uniforme ou $\|.\|_2$.

15.7 Nombres de Bell

Théorème 15.69 (Nombres de Bell[2])**.**
Soit $n \geqslant 1$ et B_n le nombre de partitions distinctes de l'ensemble $\{1, \dots, n\}$ avec la convention que $B_0 = 0$. Alors

(1) La série entière

$$\sum_{n=0}^{\infty} \frac{B_n}{n!} x^n \tag{15.194}$$

a un rayon de convergence $R > 0$ et sa somme est donnée par

$$f(x) = e^{e^x - 1} \tag{15.195}$$

pour tout $x \in]-R, R[$.

(2) Pour tout $k \in \mathbb{N}$,

$$B_n = \frac{1}{e} \sum_{k=0}^{\infty} \frac{k^n}{k!}. \tag{15.196}$$

(3) Le rayon de convergence de la série (15.194) *est en réalité infini : $R = \infty$.*

Démonstration. (1) Soit $n \geqslant 1$ et $0 \leqslant k \leqslant n$. Nous notons E_k l'ensemble des partitions de $\{1, \dots, n+1\}$ pour lesquelles le «paquet» contenant $n+1$ soit de cardinal $k+1$. Calculons le cardinal de E_k.

Pour construire un élément de E_k, il faut d'abord prendre le nombre $n+1$ et lui adjoindre k éléments choisis dans $\{1, \dots, n\}$, ce qui donne $\binom{n}{k}$ possibilités. Ensuite il faut trouver une partition des $(n+1)-(k+1) = n-k$ éléments restants, ce qui fait B_{n-k} possibilités. Donc

$$\operatorname{Card}(E_k) = \binom{n}{k} B_{n-k}. \tag{15.197}$$

L'intérêt des ensembles E_k est que $\{E_0, \dots, E_n\}$ est une partition de l'ensemble des partitions de $\{1, \dots, n+1\}$, c'est à dire que $B_{n+1} = \sum_{k=0}^{n} \operatorname{Card}(E_k)$, ce qui va nous donner une relation de récurrence pour les B_n :

$$B_{n+1} = \sum_{k=0}^{n} \operatorname{Card}(E_k) = \sum_{k=0}^{n} \binom{n}{k} B_{n-k} = \sum_{l=0}^{n} \binom{n}{n-l} B_l = \sum_{l=0}^{n} \binom{n}{l} B_l. \tag{15.198}$$

où nous avons utilisé un petit changement de variables $l = n - k$. Afin d'étudier la convergence de la série (15.194), nous allons montrer par récurrence que pour tout n, $B_n < n!$. D'abord pour $n = 0$ c'est bon : $B_1 = 1$ parce que la seule partition de $\{1\}$ est $\{1\}$. Supposons que l'inégalité soit vraie pour une certaine valeur k, et montrons qu'elle est vraie pour la valeur $k+1$:

$$B_{k+1} = \sum_{l=0}^{n} \binom{n}{k} B_k \leqslant \sum_{l=0}^{n} \binom{n}{k} k! = k! \sum_{l=0}^{k} \underbrace{\frac{1}{(n-k)!}}_{\leqslant 1} \leqslant n!(n+1) = (n+1)! \tag{15.199}$$

où nous avons utilisé la formule $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

Donc pour tout $x \in \mathbb{R}$ nous avons

$$0 \leqslant \frac{B_n}{n!} |x^n| \leqslant |x|^n, \tag{15.200}$$

et donc la série a un rayon de convergence au moins aussi grand que celui de la série géométrique, c'est à dire que 1. Donc $R \geqslant 1$. Nous nommons R ce rayon de convergence.

(2) Soit $x \in]-R,R[$. Pour une telle valeur de x à l'intérieur du disque de convergence, la proposition 15.41 nous permet de dériver terme à terme la série [13]

$$f(x) = \sum_{k=0}^{\infty} \frac{B_k}{k!}x^k = 1 + \sum_{k=0}^{\infty} \frac{B_{k+1}}{(k+1)!}x^{k+1}, \tag{15.201}$$

pour obtenir

$$f'(x) = \sum_{k=0}^{\infty} \frac{B_{k+1}}{(k+1)}(k+1)x^k = \sum_{k=0}^{\infty} \frac{B_{k+1}}{k!}x^k = \sum_{k=0}^{\infty} \left(\sum_{l=0}^{k} \binom{k}{l} B_l \right) \frac{x^k}{k!} = \sum_{k=0}^{\infty} \left(\sum_{l=0}^{k} \frac{B_l}{l!(l-k)!} \right) x^k. \tag{15.202}$$

En cette expression, nous reconnaissons un produit de Cauchy (proposition 15.31) avec $a_l = \frac{B_l}{l!}$ et $b_n = \frac{1}{n!}$. Vu que ce sont deux séries ayant un rayon de convergence plus grand que zéro, le produit a encore un rayon de convergence plus grand que zéro et nous pouvons prendre le produit des séries :

$$f'(x) = \left(\sum_{l=0}^{\infty} \frac{B_l}{l!}x^l \right) \left(\sum_{k=0}^{\infty} \frac{1}{k!}x^k \right) = f(x)e^x. \tag{15.203}$$

Étudions l'équation différentielle $y' = ye^x$. D'abord par un argument en lacet de chaussure [14], une solution est de classe C^∞. Ensuite si une solution est non nulle, elle est de signe constant. En effet si $y(x_0) < 0$ et $y(x_1) = 0$ (on choisit x_1 minimum pour cette propriété parmi les nombres plus grands que x_0) alors il existe [15] un $t \in]x_0, x_1[$ tel que $y'(t) > 0$, ce qui donnerait $y(t) > 0$, ce qui contredirait la minimalité de x_1.

Nous prétendons [16] que cette équation différentielle a un espace de solutions de dimension 1. En effet, si $y' = ye^x$ et $g' = ge^x$ alors en posant $\varphi = y/g$ nous obtenons tout de suite $\varphi' = 0$, ce qui signifie que φ est constante, ou encore que y et g sont multiples l'un de l'autre.

Si nous en trouvons une non nulle par n'importe quel moyen, c'est bon. Une solution étant dérivable est continue, donc l'équation $f' = fe^x$ nous indique que f' est continue. Une solution non nulle va automatiquement accepter un petit voisinage sur lequel la manipulation suivante a un sens :

$$\frac{f'(x)}{f(x)} = e^x, \tag{15.204}$$

donc $\ln\big(|f(x)|\big) = e^x + C$ et $f(x) = Ke^{e^x}$ pour une certaine constante. Il est vite vérifié que cette fonction est une solution de l'équation différentielle $y'(x) = y(x)e^x$ et par unicité, toutes les solutions sont de cette forme. Autrement dit, l'espace des solution est l'espace vectoriel $\text{Span}\{x \mapsto e^{e^x}\}$. Étant donné que $f(0) = 0$, nous devons choisir $K = \frac{1}{e}$ et donc

$$f(x) = \frac{1}{e}e^{e^x} = e^{e^x - 1}. \tag{15.205}$$

(3) Nous commençons par écrire la fonction f comme une série de puissance. La partie simple du calcul : pour $x \in]-R,R[$, nous avons

$$e^{e^x} = \sum_{k=0}^{\infty} \frac{(e^x)^k}{k!} = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{l=0}^{\infty} \frac{(kx)^l}{l!} = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{k^l}{k!} \frac{x^l}{l!}. \tag{15.206}$$

Notons que cela n'est pas une série de puissance en x parce qu'il y a la double somme. Nous allons inverser les sommes au moyen du théorème de Fubini sous la forme du corollaire 13.229. Pour cela nous considérons la fonction

$$\begin{aligned} a\colon \mathbb{N} \times \mathbb{N} &\to \mathbb{R} \\ (k,l) &\mapsto \frac{(kx)^l}{k!l!} \end{aligned} \tag{15.207}$$

13. C'est ici qu'on utilise la convention $B_0 = 0$ et ça aura une influence sur le choix de la constante K plus bas.
14. Genre ce qui est fait pour prouver 15.55(4).
15. Théorème de Rolle 12.127.
16. Ou alors on utilise le théorème 30.14 avec $M(x) = e^x$ dans les cas $n = 1$ et $I =]-R,R[$.

et nous mettons la mesure de comptage [17] sur $\mathbb{N}$ et $\mathbb{N}^2$. Nous commençons donc à vérifier l'intégrabilité variable par variable de $|a|$:

$$\int_{\mathbb{N}} \left(\int_{\mathbb{N}} |a(k,l)| dm(l) \right) dm(k) = \sum_{k=0}^{\infty} \frac{1}{k!} \frac{(k|x|)^l}{l!} \tag{15.208a}$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} e^{k|x|}. \tag{15.208b}$$

Nous devons montrer que cette dernière somme va bien. Pour cela nous posons $u_k = \frac{e^{k|x|}}{k!}$ et nous remarquons que $\frac{u_{k+1}}{u_k} \to 0$. Donc la double intégrale (15.208) converge, ergo $a \in L^1(\mathbb{N} \times \mathbb{N})$, ce qui nous permet d'utiliser le théorème de Fubini 13.230 pour inverser les ~~sommes~~ ~~intégrales~~ sommes dans l'équation (15.206) :

$$\frac{1}{e} e^{e^x} = \frac{1}{e} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{1}{k!} \frac{1}{l!} (kx)^l = \sum_{l=0} \frac{1}{e} \frac{1}{l!} \left(\sum_{k=0}^{\infty} \frac{k^l}{k!} \right) x^l. \tag{15.209}$$

Cela est un développement en série entière pour la fonction $\frac{1}{e} e^{e^x}$, dont nous savions déjà le développement (15.194) ; par unicité du développement nous pouvons identifier les coefficients :

$$B_l = \frac{1}{e} \sum_{k=0}^{\infty} \frac{k^l}{k!}. \tag{15.210}$$

(4) Le développement (15.206) étant en réalité valable pour tout x et tous les calculs subséquents l'étant aussi, le développement

$$e^{e^x - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} x^n \tag{15.211}$$

est en fait valable pour tout x, ce qui donne à la série entière un rayon de convergence infini.

□

15.8 Trigonométrie hyperbolique

Définition 15.70.
Les fonction ***sinus hyperbolique*** *et* ***cosinus hyperbolique*** *sont les fonctions définies sur* $\mathbb{R}$ *par les formules suivantes :*

$$\cosh(x) = \frac{e^x + e^{-x}}{2} \tag{15.212a}$$

$$\sinh(x) = \frac{e^x - e^{-x}}{2} \tag{15.212b}$$

Leurs principales propriétés sont :

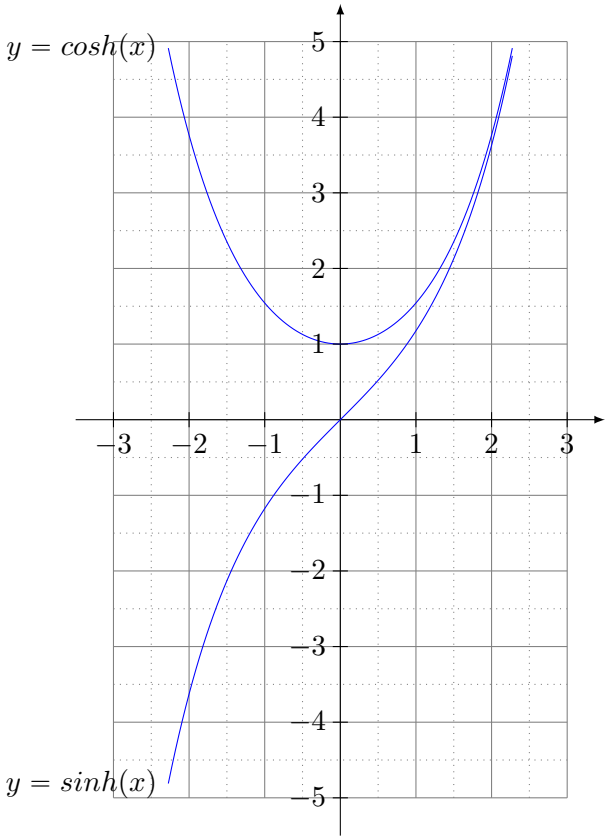
(1) $\cosh^2(x) - \sinh^2(x) = 1$

(2) $\cosh'(x) = \sinh(x)$

(3) $\sinh'(x) = \cosh.$

Les représentations graphiques sont ceci :

17. Nous passons outre les avertissements et menaces de Arnaud Girand.



La **tangente hyperbolique** est donnée par le quotient

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)}. \tag{15.213}$$

15.9 Séries entières de matrices

15.9.1 Différentiabilité

Proposition 15.71.
Soit (a_n) une suite dans $\mathbb{C}$ de rayon de convergence R et la fonction

$$\begin{aligned} f\colon \mathbb{M}(n,\mathbb{R}) &\to \mathbb{M}(n,\mathbb{R}) \\ A &\mapsto \sum_{k=0}^{\infty} a_k A^k \end{aligned} \tag{15.214}$$

Alors

(1) La différentielle de f sur $B(0,R)$ est

$$df_A(U) = \sum_{k=0}^{\infty} a_k \sum_{l=0}^{k-1} A^l U A^{k-1-l}, \tag{15.215}$$

c'est à dire que l'on peut différentier terme à terme. (Ici c'est A qui est dans $B(0,R)$)

(2) La convergence de la somme 15.215 est absolue.

(3) La convergence de la somme 15.215 est normale sur tout compact.

(4) La fonction f est de classe C^1 sur $B(0,R)$, c'est à dire que la fonction $A \mapsto df_A$ est continue.

Notons que df_A n'est pas tout à fait une série entière. Cependant, en ce qui concerne les normes, c'est tout comme si ça l'était.

Démonstration. Nous posons $u_k(A) = a_k A^k$, qui est une fonction de classe C^∞ et dont la différentielle est donnée par

$$(du_k)_A(U) = \frac{d}{dt}\Big[u_k(A+tU)\Big]_{t=0} = a_k \frac{d}{dt}\Big[(A+tU)^k\Big]_{t=0}; \tag{15.216}$$

en distribuant le produit nous trouvons tout un tas de termes dont seuls ceux contenant exactement une fois tU ne vont pas s'annuler. Étant donné que U et A ne commutent pas nous avons l'expression un peu moche

$$(du_k)_A(U) = \sum_{l=0}^{k-1} a_k A^l U A^{k-1-l}. \tag{15.217}$$

En ce qui concerne la norme, nous regardons celle de $(du_k)_A$ pour un A fixé ; c'est à dire que nous en regardons la norme opérateur :

$$\|(du_k)_A\| = \sup_{\|U\|=1} \| \sum_{l=0}^{k-1} a_k A^l U A^{k-1-l}\| \leqslant \sum_{l=0}^{k-1} |a_k| \|A\|^l \|A\|^{k-1-l} \leqslant k|a_k|\|A\|^{k-1}. \tag{15.218}$$

Pour donner la convergence nous considérons un nombre r tel que $\|A\| < r < R$, de telle sorte que la suite $(a_n r^n)$ soit bornée par un nombre M et que nous puissions écrire

$$\|(du_k)_A\| \leqslant k|a_k|\|A\|^{k-1} = \frac{k|a_k|\|A\|^k}{\|A\|} = \frac{k|a_k|}{\|A\|} r^k \left(\frac{\|A\|}{r}\right)^k \leqslant \frac{M}{\|A\|} k \left(\frac{\|A\|}{r}\right)^k, \tag{15.219}$$

dont la série converge. Nous avons donc convergence absolue de la série

$$\sum_{k=0}^{\infty} (du_k)_A. \tag{15.220}$$

Passons à la convergence normale sur tout compact. Nous nous fixons $r < R$ et nous nous intéressons à la norme de du_k sur $\overline{B(0,r)}$, c'est à dire

$$\|du_k\|_\infty = \sum_{x \in \overline{B(0,r)}} \|(du_k)_A\|. \tag{15.221}$$

Vu que $\overline{B(0,r)}$ est compact, ce supremum est un maximum et nous pouvons noter A_k la matrice qui le réalise. Nous réalisons alors les mêmes manipulations que pour (15.219) :

$$\|du_k\|_\infty = \|(du_k)_{A_k}\| \leqslant k|a_k|\|A_k\|^{k-1} \leqslant k|a_k|r^{k-1} = \frac{1}{r}k|a_k|r^k. \tag{15.222}$$

Nous prenons maintenant $r < r_0 < R$ et M, un majorant de $(a_n r_0^n)$, de telle sorte qu'en multipliant et divisant par r_0^k,

$$\|du_k\|_\infty \leqslant \frac{k|a_k|r_0^k}{r}\frac{r^k}{r_0^k} \leqslant \frac{kM}{r}\left(\frac{r}{r_0}\right)^k, \tag{15.223}$$

dont la série converge. Nous avons donc convergence normale sur tout compact. Par voie de ~~fait~~ conséquences nous avons continuité de la série

$$\sum_{k=0}^{\infty} (du_k)_A \tag{15.224}$$

et convergence vers df_A par le théorème 15.22. □

Proposition 15.72.
Si le rayon de convergence de la série $u(A) = \sum_{k=0}^{\infty} a_k A^k$ est R, alors

(1) elle converge normalement sur tout compact de $B(0,R)$;

(2) la fonction u y est de classe C^∞.

Démonstration. Nous posons

$$\begin{aligned} u_k\colon \mathbb{M}(n,\mathbb{R}) &\to \mathbb{M}(n,\mathbb{R}) \\ A &\mapsto a_k A^k \end{aligned} \tag{15.225}$$

qui est évidemment une fonction de classe C^∞. Nous étudions la j^{e} différentielle en m, pour $k > j$ (dans une série, nous ne nous intéressons pas aux premiers termes). La j^{e} différentielle appliquée à v_1 appliquée à v_2, etc s'exprime de la façon suivante :

$$(d^j u_k)_m(v_1,\dots,v_j) = \frac{d}{dt_1}\dots\frac{d}{dt_j}\Big(u_k(m+t_1v_1+\dots+t_jv_j)\Big)_{t_i=0}. \tag{15.226}$$

Dans le produit $(m+t_1v_1+\dots+t_jv_j)^k$, seuls les termes contenant exactement une fois chacun des t_i ne s'annulera pas après avoir fait la dérivée et évalué en $t_i = 0$. Combien de termes cela fait ? Parmi les k facteurs, il faut en placer j qui ne sont pas m (cela fait $\binom{k}{j}$ possibilités), et puis il faut ordonner ces j termes, cela fait encore $j!$ possibilités. Au final,

$$\|(d^j u_k)_m\| \leqslant |a_k|\binom{k}{j}j!\|m\|^{k-j} = |a_k|P(k)\|m\|^{k-j} \tag{15.227}$$

où $P(k) = \frac{k!}{(k-j)!}$ est un polynôme de degré j.

Afin d'étudier la convergence normale sur tout compact de la série des $d^j u_k$, nous considérons $r < r_0 < R$ et nous allons prouver la convergence normale sur $\overline{B(0,r)}$. Vu que c'est un compact, il existe une matrice $m_k \in \overline{B(0,r)}$ telle que

$$\begin{aligned} \|d^j u_k\|_\infty &= \|(d^j u_k)_{m_k}\| && \text{(15.228a)}\\ &\leqslant |a_k|P(k)\|m_k\|^{k-j} && \text{(15.228b)}\\ &\leqslant |a_k|P(k)r^{k-j} && \text{(15.228c)}\\ &= \frac{|a_k|P(k)}{r^j}r^k && \text{(15.228d)}\\ &= \frac{|a_k|r_0^k P(k)}{r^j}\left(\frac{r}{r_0}\right)^k && \text{(15.228e)}\\ &\leqslant \frac{M}{r^j}P(k)\left(\frac{r}{r_0}\right)^k && \text{(15.228f)} \end{aligned}$$

où M est un majorant de $a_n r^n$. Vu que $r_0/r < 1$, la somme sur k converge et nous avons convergence normale sur tout compact de

$$d^j\sum_{k=0}^{\infty}a_kA^k = \sum_{k=0}^{\infty}d^j(a_kA^k) \tag{15.229}$$

avec un peu d'abus de notation. □

15.10 Étude d'asymptote

Lorsqu'une fonction tend vers l'infini pour $x \to \infty$, une question qui peut venir est : à quelle vitesse tend-t-elle vers l'infini ?

Il est «visible» que la fonction logarithme ne tend pas très vite vers l'infini : certes

$$\lim_{x\to\infty}\ln(x) = +\infty, \tag{15.230}$$

mais par exemple $\ln(100000) \simeq 11.5$ tandis que $e^{100000} \simeq 10^{43429}$. Sans contestations possibles, l'exponentielle croit plus vite que le logarithme.

Soient f et g deux fonctions dont la limite $x \to \infty$ est ∞. Si

$$\lim_{x\to\infty}\frac{f(x)}{g(x)} = 0 \tag{15.231}$$

nous disons que g tend vers ∞ plus vite que f ; si

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \infty \tag{15.232}$$

nous disons que f tend vers ∞ plus vite que g, et si

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = a \in \mathbb{R} \tag{15.233}$$

avec $a \neq 0$ alors nous disons que f tend vers l'infini à la même vitesse que $ag(x)$.

Exemple 15.73
La fonction $x \mapsto x^2$ tend vers l'infini plus vite que la fonction $x \mapsto \sqrt{x}$. △

Dans cette section nous allons nous contenter de déterminer les fonctions qui tendent vers l'infini aussi vite qu'une droite oblique, que nous appellons asymptote et que nous voulons déterminer.

Exemple 15.74
Déterminer les asymptotes obliques (s'ils existent) de la fonction

$$f(x) = e^{1/x}\sqrt{1+4x^2}. \tag{15.234}$$

Tout d'abord nous remarquons que $\lim_{x\to\infty} f(x) = \infty$. Nous sommes donc en présence d'une branche du graphe qui tend vers l'infini. Ensuite,

$$\lim_{x\to\infty} \frac{f(x)}{x} = \lim_{x\to\infty} e^{1/x}\sqrt{\frac{1}{x^2}+4} = 2. \tag{15.235}$$

Donc le graphe de f tend vers l'infini à la même vitesse que le graphe de la fonction $y = 2x$. Nous aurons donc une asymptote oblique de coefficient directeur 2. De façon imagée, nous pouvons penser que le graphe de f et celui de $y = 2x$ sont presque parallèles si x est assez grand. Afin de déterminer l'ordonnée à l'origine de l'asymptote, il nous reste à voir quelle est la «distance» entre le graphe de f et celui de $y = 2x$:

$$\lim_{x\to\infty} f(x) - 2x = \lim_{x\to\infty} e^{1/x}\sqrt{1+4x^2} - 2x. \tag{15.236}$$

Cette limite a été calculée dans l'exemple 12.294 et vaut 2.

Nous concluons que le graphe de la fonction f admet l'asymptote

$$y = 2x + 2. \tag{15.237}$$

△

15.11 Développement en série

15.11.1 Série génératrice d'une suite

Avant de commencer, une petite formule de dérivation toute simple que nous allons utiliser souvent :

$$(z^k)^{(l)} = \begin{cases} 0 & \text{si } l > k \\ \frac{k!}{(k-l)!} z^{k-l} & \text{sinon.} \end{cases} \tag{15.238}$$

Soit u_n une suite telle que le rayon de convergence de

$$f(z) = \sum_{n=0}^{\infty} u_n z^n \tag{15.239}$$

soit strictement positif. Alors la série f est la **série génératrice** de la suite (u_n).

Grâce au théorème 15.41 nous pouvons la dériver terme à terme autour de $z = 0$. En utilisant la petite formule (15.238) nous trouvons

$$f^{(l)}(z) = \sum_{n=l}^{\infty} u_n \frac{n!}{(n-l)!} z^{n-l}, \tag{15.240}$$

et donc

$$u_l = \frac{f^{(l)}(0)}{l!}. \tag{15.241}$$

D'où le nom de série génératrice. Cela est évidemment intéressant seulement si nous connaissons une autre forme pour f par ailleurs.

Nous en utiliserons une pour déterminer les partitions d'un nombre en parts fixes, proposition 25.42.

15.11.2 Développement en série et Taylor

Définition 15.75.
*Soit une fonction $f : \mathbb{C} \to \mathbb{C}$ et $z_0 \in \mathbb{C}$. Nous disons que f est **développable en série entière** dans un voisinage de z_0 s'il existe une série $\sum_n a_n z^n$ de rayon de convergence $R > 0$ et $r \leqslant R$ tel que*

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \tag{15.242}$$

pour tout $z \in B(z_0, r)$.

Proposition 15.76.
Si V est un ouvert dans $\mathbb{C}$ alors l'ensemble des fonctions $V \to \mathbb{C}$ développables en série entière forme une $\mathbb{C}$-algèbre.

Démonstration. Les séries entières passent aux sommes et aux produits en gardant des rayons de convergence non nuls. □

Proposition 15.77.
*Si f est développable en série entière à l'origine alors elle est C^∞ sur un voisinage de l'origine et le développement est celui de **Taylor** :*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \tag{15.243}$$

pour tout x dans un voisinage de 0.

Démonstration. Si $f(x) = \sum a_n x^n$, nous savons que f est C^1 et que nous pouvons dériver terme à terme (au moins dans un voisinage). De plus le fait de dériver ne change pas le domaine. Par récurrence, la fonction est C^∞ sur le voisinage. En dérivant k fois la série $\sum a_n x^n$ nous trouvons

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\dots(n-k+1) a_n x^{n-k}. \tag{15.244}$$

En calculant en $x = 0$ nous trouvons

$$f^{(k)}(0) = k! a_k, \tag{15.245}$$

d'où le terme général

$$a_k = \frac{f^{(k)}(0)}{k!}. \tag{15.246}$$

□

Si f est une fonction et si la série

$$T_f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \tag{15.247}$$

converge, alors cette série est la **série de Taylor** de f.

Remarque 15.78.
La série de Taylor d'une fonction n'est pas liée à sa fonction de façon aussi raide qu'on pourrait le croire. Même dans le cas d'une fonction C^∞ il peut arriver que $T_f(x) \neq f(x)$.

Il peut aussi arriver que f ne soit pas développable en série entières.

Exemple 15.79
Nous considérons la fonction

$$f(x) = \begin{cases} e^{-1/x^2} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0. \end{cases} \tag{15.248}$$

Nous avons

$$f'(x) = \begin{cases} \frac{2}{x^3} e^{-1/x^2} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0. \end{cases} \tag{15.249}$$

Note : pour la seconde ligne nous devons faire explicitement le calcul

$$f'(0) = \lim_{t\to 0} \frac{f(t) - f(0)}{t} = \lim_{t\to 0} \frac{1}{t} e^{-1/t^2} = 0. \tag{15.250}$$

Plus généralement nous avons $f^{(k)}(0) = 0$, et par conséquent la série de Taylor converge (trivialement) vers la fonction identiquement nulle.

Cette fonction n'est donc pas développable en série entière vu qu'il n'existe aucun voisinage de zéro sur lequel la série de f coïncide avec f. △

Exemple 15.80
Développement de $f(x) = \arctan(x)$. Nous savons que

$$f'(x) = \frac{1}{1+x^2}, \tag{15.251}$$

alors que nous connaissons le développement

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n \tag{15.252}$$

pour tout $x \in B(0,1)$. Nous avons donc successivement

$$\frac{1}{1+x} = \sum_{n=0} (-x)^n \tag{15.253a}$$

$$\frac{1}{1+x^2} = \sum_{n=0} (-1)^n x^{2n} \tag{15.253b}$$

$$\arctan(x) = \sum_{n=1}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} + C. \tag{15.253c}$$

Notons que dans la dernière nous avons évité d'écrire la somme depuis $n = 0$ (qui serait un terme constant) et nous avons écris explicitement «$+C$». Étant donné que $\arctan(0) = 0$, nous devons poser $C = 0$ et donc

$$\arctan(x) = \sum_{n=1}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}. \tag{15.254}$$

△

15.11.3 Resommer une série

Nous avons vu comment trouver la série correspondant à une fonction donnée. Un exercice difficile consiste à trouver la fonction qui correspond à une somme donnée.

15.11.3.1 Les sommes du type $\sum_n P(n)x^n$

Pour calculer

$$\sum_{n=0}^{\infty} P(n)x^n \tag{15.255}$$

où P est un polynôme de degré m nous commençons par écrire

$$P(n) = \alpha_0 + \alpha_1(n+1) + \alpha_2(n+1)(n+2) + \cdots + \alpha_m(n+1)\dots(n+m). \tag{15.256}$$

Nous décomposons alors la somme en m sommes de la forme

$$\sum_{n=0}^{\infty} \alpha_k \frac{(n+k)!}{n!}x^n = \alpha_k \left(\sum_{n=0}^{\infty} x^{n+k}\right)^{(k)}. \tag{15.257}$$

Effectuons par exemple

$$\sum_{n=0}^{\infty} x^{n+3} = \frac{1}{1-x} - 1 - x - x^2 \tag{15.258}$$

Notons que dans un usage pratique, ce terme devra être ensuite dérivé trois fois, de telle manière que les termes «correctifs» n'interviennent pas. Cette méthode ne demande donc que de calculer les dérivées successives de $1/(1-x)$.

Exemple 15.81
Calculons la fonction

$$f(x) = \sum_{n=0}^{\infty} n^3x^n. \tag{15.259}$$

D'abord nous écrivons

$$n^3 = -1 + 7(n+1) - 6(n+1)(n+2) + (n+1)(n+2)(n+3). \tag{15.260}$$

Nous avons

$$\sum_{n=0}^{\infty}(n+1)x^n = \left(\sum_{n=0}^{\infty} x^{n+1}\right)' = \left(\frac{1}{1-x} - 1\right)' = \frac{1}{(x-1)^2}. \tag{15.261}$$

De la même façon,

$$\sum_n (n+1)(n+2)x^n = \left(\sum x^{n+2}\right)'' = \frac{-2}{(x-1)^3} \tag{15.262a}$$

$$\sum_n (n+1)(n+2)(n+3)x^n = \frac{6}{(x-1)^4}. \tag{15.262b}$$

En remettant tout ensemble nous obtenons

$$\sum_{n=0}^{\infty} n^3x^n = -\frac{1}{1-x} + \frac{7}{(x-1)^2} + \frac{12}{(x-1)^3} + \frac{6}{(x-1)^4}. \tag{15.263}$$

Nous pouvons vérifier ce résultat en traçant les deux courbes et en remarquant qu'elles coïncident.

```
----------------------------------------------------------------------
| Sage Version 4.7.1, Release Date: 2011-08-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: n=var('n')
sage: S(x)=sum(  [ n**3*x**n for n in range(0,30)  ]   )
sage: f(x)=-1/(1-x)+7/((x-1)**2)+12/((x-1)**3)+6/( (x-1)**4  )
sage: S(0.1)
0.214906264288980
sage: f(0.1)
0.214906264288981
sage: f.plot(-0.5,0.5)+S.plot(-0.5,0.5)
```

△

15.11.3.2 Les sommes du type $\sum_n x^n/P(n)$

Si $P(n)$ a des racines entières, nous pouvons le décomposer en fractions simples et utiliser la somme

$$\sum_{n=1}^{\infty} \frac{x^n}{n} = -\ln(1-x). \tag{15.264}$$

Nous avons par exemple

$$\sum_{n=0}^{\infty} \frac{x^n}{n+1} = \frac{1}{x} \sum_{n=0} \frac{x^{n+1}}{n+1} \tag{15.265a}$$

$$= \frac{1}{x} \sum_{n=1}^{\infty} \frac{x^n}{n} = -\frac{\ln(1-x)}{x}. \tag{15.265b}$$

Notez le changement de point de départ de la somme au passage.

Autre exemple :

$$\sum_{n=0}^{\infty} \frac{x^n}{n+3} = \frac{1}{x^3} \left(\sum_{n=1}^{\infty} \frac{x^n}{n} - x - \frac{x^2}{2} \right) \tag{15.266a}$$

$$= -\frac{\ln(x-1)}{x^3} - \frac{1}{x^2} - \frac{1}{2x}. \tag{15.266b}$$

Si le polynôme possède des racines non entières, les choses se compliquent.

Exemple 15.82
Calculons

$$\sum_{n=0}^{\infty} \frac{x^n}{2n+1}. \tag{15.267}$$

Si $x \geqslant$, en posant $t = \sqrt{x}$ nous trouvons

$$\sum_{n=0}^{\infty} \frac{x^n}{2n+1} = \frac{1}{t} \sum_{n=0}^{\infty} \frac{t^{2n+1}}{2n+1}. \tag{15.268}$$

Étudions

$$H(t) = \sum_{n=0}^{\infty} \frac{t^{2n+1}}{2n+1}. \tag{15.269}$$

Nous avons

$$H'(t) = \sum_{n=0}^{\infty} t^{2n} = \sum_{n=0} (t^2)^n = \frac{1}{1-t^2}. \tag{15.270}$$

Une primitive de cette fonction est

$$\frac{1}{2}\ln\left|\frac{t+1}{t-1}\right|. \tag{15.271}$$

En $t=0$, cette fonction vaut 0 qui est la bonne valeur. Donc nous avons bien

$$H(t)=\frac{1}{2}\ln\left|\frac{t+1}{t-1}\right|. \tag{15.272}$$

Notons que ce que l'équation (15.270) nous dit est que $H(t)$ est une primitive de $1/(1-t^2)$. Il faut choisir la bonne primitive en fixant une valeur.

Nous avons donc

$$\sum_{n=0}^{\infty}\frac{x^n}{2n+1}=\frac{1}{2\sqrt{x}}\ln\left|\frac{\sqrt{x}+1}{\sqrt{x}-1}\right| \tag{15.273}$$

pour $x>0$. Nous devons encore trouver ce que cela vaut pour $x<0$.

Nous posons successivement $X=-x$ puis $g(X)=f(-X)$. Ce que nous devons calculer est

$$g(t)=\frac{1}{t}\sum_{n=0}^{\infty}\frac{(-1)^n t^{2n+1}}{2n+1}. \tag{15.274}$$

Si nous posons

$$h(t)=\sum\frac{(-1)^n t^{2n+1}}{2n+1}, \tag{15.275}$$

alors

$$h'(t)=\sum(-1)^n t^{2n}=\sum(-t^2)^n=\frac{1}{1+t^2}, \tag{15.276}$$

par conséquent $h(t)=\arctan(t)$ (cela avait déjà été déduit à l'envers dans l'exemple 15.80).

Au final

$$f(x)=\sum_{n=0}^{\infty}\frac{x^n}{2n+1}=\begin{cases}\frac{1}{2\sqrt{x}}\ln\left|\frac{\sqrt{x}+1}{\sqrt{x}-1}\right| & \text{si } x>0\\ \frac{\arctan(\sqrt{-x})}{\sqrt{-x}} & \text{si } x<0\\ 1 & \text{si } x=0.\end{cases} \tag{15.277}$$

Notons qu'elle est continue en zéro à gauche et à droite.

△

Exemple 15.83

Nous considérons l'exemple suivant :

$$f(x)=\sum_{n=0}^{\infty}\frac{x^n}{3n+2}. \tag{15.278}$$

Nous posons $t=\sqrt[3]{x}$, et nous substituons :

$$\frac{x^n}{3n+2}=\frac{t^{3n}}{3n+2}=\frac{1}{t^2}\frac{t^{3n+2}}{3n+2}. \tag{15.279}$$

Nous devons étudier la fonction

$$g(t)=\sum_{n=0}^{\infty}\frac{t^{3n+2}}{3n+2} \tag{15.280}$$

Nous avons

$$g'(t)=\sum_{n=0}t^{3n+1}=t\sum_{n=0}t^{3n}=\frac{t}{1-t^3}. \tag{15.281}$$

Notons que $g(0)=0$. △

Exemple 15.84
Calculer le nombre

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}. \tag{15.282}$$

Nous aurions envie de dire que cela est $f(-1)$ pour la fonction f donnée en (15.277). Le problème est que le rayon de convergence de f étant 1, rien n'est garantit quand au fait que la fonction y soit continue en $x = -1$. En particulier nous devons justifier le fait que

$$\lim_{x\to -1} \sum_n \frac{x^n}{2n+1} = \lim_{x\to -1} \frac{1}{\sqrt{-x}} \arctan(\sqrt{-x}). \tag{15.283}$$

Ce qui nous sauve est le critère d'Abel radial (théorème 15.37). En effet la série

$$\sum \frac{r^n}{2n+1} \tag{15.284}$$

étant convergente avec $r = -1$, la série correspondante est continue sur $[-1, 0]$. Nous pouvons donc calculer la série (15.282) en posant $x = -1$ dans (15.277) :

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4}. \tag{15.285}$$

Note : la série (15.284) ne converge pas avec $r = 1$. La fonction f n'est pas continue en $x = 1$.
△

Exemple 15.85
Nous avons

$$\sum_{n=1}^{\infty} nx^{n-1} = \frac{1}{(1-x)^2}. \tag{15.286}$$

En effet si nous désignons par f la somme à gauche, nous trouvons que $f = g'$ avec

$$g(x) = \sum_{n=1}^{\infty} x^n. \tag{15.287}$$

Nous savons par ailleurs que $g(x) = 1/(1-x)$. Par conséquent

$$f(x) = \left(\frac{1}{1-x}\right)' = \frac{1}{(1-x)^2}. \tag{15.288}$$

△

15.11.3.3 Sage, primitives et logarithme complexe

15.86.
Attention : Sage pourrait nous induire en erreur si nous n'y prenions pas garde. En effet ce que vous ne savez pas mais que Sage sait, c'est que

$$\ln(-1) = i\pi. \tag{15.289}$$

Par conséquent Sage se permet de donner des primitives sans valeurs absolues dans le logarithme :

```
sage: f(x)=1/x
sage: f.integrate(x)
x |--> log(x)
```

La primitive à laquelle on s'attend d'habitude est $\ln(|x|)$. Ici la réponse est correcte parce que si x est négatif nous avons

$$\ln(x) = \ln\big((-1)|x|\big) = \ln(-1) + \ln(|x|). \tag{15.290}$$

Cette fonction est donc décalée de la primitive usuelle seulement de la constante $\ln(-1)$.

Un exemple plus élaboré :

```
sage: h(x)=1/(1-x**2)
sage: H=h.integrate(x)
sage: H
x |--> -1/2*log(x - 1) + 1/2*log(x + 1)
sage: H(0)
-1/2*I*pi
```

Exemple 15.87
Encore une fois il faut faire attention en demandant la primitive à Sage :

```
----------------------------------------------------------------------
| Sage Version 4.7.1, Release Date: 2011-08-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x)=x/(1-x**3)
sage: F=f.integrate(x)
sage: F(0)
-1/3*I*pi - 1/3*sqrt(3)*arctan(1/3*sqrt(3))
```

Cette fois la primitive proposée diffère de celle qu'on cherche de la constante complexe

$$-\frac{\pi}{3}i. \tag{15.291}$$

Mais il y a pire si nous voulons tracer. Nous voudrions définir la fonction $F_2(x) = F(x) - F(0)$. Mathématiquement c'est bien de cette fonction que nous parlons, mais :

```
sage: F2(x)=F(x)-F(0)
sage: F2(x)
1/3*I*pi - 1/3*sqrt(3)*arctan(1/3*(2*x + 1)*sqrt(3)) +
    +1/3*sqrt(3)*arctan(1/3*sqrt(3)) - 1/3*log(x - 1) + 1/6*log(x^2 + x + 1)
sage: F2.plot(x,-0.1,0.1)
verbose 0 (4101: plot.py, generate_plot_points) WARNING: When plotting, failed to evaluate
verbose 0 (4101: plot.py, generate_plot_points) Last error message: 'unable to simplify to
```

Il refuse de tracer. Pourquoi ? La partie complexe de l'expression de F_2 est mathématiquement nulle, mais elle est en deux parties :

$$\frac{\pi}{3} + \text{la partie imaginaire de} -\frac{1}{3}\ln(x-1). \tag{15.292}$$

Lorsque Sage tente de tracer, il donne à x un certain nombre de valeurs et calcule une *valeur approchée* de $\ln(x-1)$. Cette dernière ne se simplifie pas avec le nombre *exact* $\pi/3$. Sage reste donc avec une partie imaginaire qu'il ne peut pas tracer.

Notez la nuance :

```
sage: ln(-0.1)
-2.30258509299405 + 3.14159265358979*I
sage: ln(-1/10)
I*pi + log(1/10)
```

Du coup nous avons aussi

```
sage: F2(-0.1)
1/3*I*pi - 1/3*sqrt(3)*arctan(0.266666666666667*sqrt(3))
    + 1/3*sqrt(3)*arctan(1/3*sqrt(3)) - 0.0474885065133152 - 1.04719755119660*I
```

△

15.11.3.4 Nombres de Bell

Ici nous montrerions bien le théorème 15.69 sur les nombres de Bell parce que c'est essentiellement un résultat sur les séries entières et leurs manipulations. Hélas, il demande un tout petit peu d'équation différentielle (presque rien). Donc il est postposé jusqu'en page 922.

15.12 Séries entières de matrices

Nous nous proposons d'étudier des séries de la forme

$$\sum_{k=0}^{\infty} a_k A^k \tag{15.293}$$

où A est une matrice. L'essentiel de la théorie va rester. Nous considérons une norme algébrique (définition 10.9), c'est à dire $\|AB\| \leqslant \|A\|\|B\|$.

15.12.1 Rayon de convergence

La notion de rayon de convergence de cette série reste la même : c'est la définition 15.24 qui ne dépend que des coefficients a_k et pas du tout de ce qu'on met à côté dans la somme. Évidemment il faudra montrer que dans le cas des matrices, le nom «rayon de convergence» n'est pas usurpé.

Proposition 15.88.
Soit (a_n) une suite dans $\mathbb{C}$ de rayon de convergence R et $A \in \mathbb{M}(n,\mathbb{R})$ une matrice vérifiant $\|A\| < R$. Alors la série

$$\sum_{k=0}^{\infty} a_k A^k \tag{15.294}$$

converge absolument, c'est à dire que $\sum_k \|a_k A^k\| < \infty$.

Démonstration. Nous avons les majorations

$$\|a_n A^n\| \leqslant |a_n|\|A^n\| \leqslant |a_n|\|A\|^n. \tag{15.295}$$

Par hypothèse $\|A\| < R$ et R est un supremum, donc il existe r tel que $\|A\| < r < R$ avec $(a_n r^n)$ borné. Nommons M un majorant de la suite $(a_n r^n)$. Alors nous avons

$$\|A_n A^n\| \leqslant |a_n| r^n \frac{\|A\|^n}{r^n} \leqslant M \left(\frac{\|A\|}{r}\right)^n. \tag{15.296}$$

La série du membre de droite converge parce que c'est une série géométrique de raison plus petite que 1 ; voir l'exemple 10.62. □

15.12.2 Convergence et rayon spectral

Le concept de rayon spectral permet aussi de donner des informations sur la convergence de séries de matrices. Pour rappel le rayon spectral d'une matrice est le maximum du module de ses valeurs propres (définition 10.10). Le rayon spectral de la matrice A est noté $\rho(A)$.

La proposition suivante sera redémontrée indépendamment dans le théorème 15.90.

Proposition 15.89 ([34]).
Si $A \in \mathbb{M}(n,\mathbb{C})$ est telle que $\rho(A) < 1$, alors $A^n \to 0$.

Démonstration. Nous nous plaçons dans une base des espaces caractéristiques [18] de A, c'est à dire que nous supposons que la matrice A a la forme

$$A = \begin{pmatrix} \lambda_1\mathbb{1} + N_1 & & \\ & \ddots & \\ & & \lambda_s\mathbb{1} + N_s \end{pmatrix} \tag{15.297}$$

où les λ_i sont les valeurs propres de A et les N_i sont nilpotentes. En effet nous savons que l'espace caractéristique F_{λ_i} est l'espace de nilpolence de $A - \lambda_i\mathbb{1}$. Si nous notons A_i la restriction de A à cet espace, la matrice $N_i = A_i - \lambda_i\mathbb{1}$ est nilpotente. Du coup $A_i = \lambda_I\mathbb{1} + N_i$ et nous avons bien la décomposition (15.297).

Nous avons donc $A^n \to 0$ si et seulement si $(N_i + \lambda_i\mathbb{1})^n \to 0$ pour tout i. Soit donc N nilpotente et $\lambda < 1$ (parce que nous savons que toutes les valeurs propres de A sont inférieures à un). Nous avons

$$(\lambda\mathbb{1} + N)^n = \sum_{k=0}^{n} \binom{n}{k} \lambda^{n-k} N^k = \sum_{k=0}^{r-1} \binom{n}{k} \lambda^{n-k} N^k. \tag{15.298}$$

Nous voyons que le nombre de termes dans la somme ne dépend pas de n. De plus pour chacun de termes, la puissance de N ne dépend pas non plus de n. Le terme

$$\binom{n}{k} \lambda^{n-k} \leqslant P(n)\lambda^{n-k} \tag{15.299}$$

où P est un polynôme tend vers zéro lorsque n devient grand parce que c'est une cas polynôme fois exponentielle. □

Théorème 15.90 (Thème 40[123]).
Soit $A \in \mathbb{M}(n,\mathbb{K})$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). Les affirmations suivantes sont équivalentes.

(1) $\lim_{k\to\infty} A^k = 0$
(2) $\rho(A) < 1$
(3) $\sum_{k=0}^{\infty} A^k$ *converge.*

Dans le cas où ces conditions sont vérifiées, nous avons aussi
— $\mathbb{1} - A$ *est inversible,*
— $\sum_{k=0}^{\infty} A^k = (\mathbb{1} - A)^{-1}$

Démonstration. Nous supposons qu'une norme est donnée sur $\mathbb{K}^n$ et nous considérons sur $\mathbb{M}(n,\mathbb{K})$ la topologie associée à la norme subordonnée [19]. Nous subdivisons la preuves en différentes implications.

(1) implique (2) Si $\rho(A) \leqslant 1$, en combinant la proposition 10.15 avec la proposition 10.17, nous avons

$$\|A^m\| \geqslant \big(\rho(A)\big)^m \geqslant 1 \tag{15.300}$$

Mais la limite $A^k \xrightarrow{\mathbb{M}(n,\mathbb{K})} 0$ signifie la limite $\|A^k\| \xrightarrow{\mathbb{R}} 0$. Le fait que tous les éléments de la suite soient plus grand que 1 empêche cette limite.

(2) implique (1) Vu que $\rho(A) < 1$, il existe $\epsilon > 0$ tel que $\rho(A) + \epsilon < 1$. Par le lemme 10.16 il existe une norme N sur $\mathbb{M}(n,\mathbb{K})$ telle que $N(A) \leqslant \rho(A) + \epsilon < 1$. Notons que cette norme N dépend de A et de ϵ.

Avec cette norme nous avons

$$N(A^k) \leqslant N(A)^k \xrightarrow{\mathbb{R}} 0. \tag{15.301}$$

18. Voir le théorème 9.193
19. Si on parle de convergence d'une suite, c'est qu'il y a une topologie quelque part.

Cela signifie que $A^k \xrightarrow{N} 0$. L'équivalence entre toutes les normes sur $\mathbb{M}(n,\mathbb{K})$ donne alors la convergence $A^k \xrightarrow{\|.\|} 0$.

Pour une preuve alternative de cette implication, voir la proposition 15.89.

(3) implique (1) La convergence d'une série implique que la norme du terme général converge vers zéro par la proposition 10.53. Nous avons donc $\|A^k\| \to 0$, ce qui signifie $A^k \to 0$, et donc $\rho(A) < 1$ parce que (1) implique (2).

$\rho(A) < 1$ implique $\mathbb{1} - A$ est inversible Si μ est une valeur propre de $\mathbb{1} - A$ alors

$$\det\big((\mathbb{1} - A) - \mu\mathbb{1}\big) = \det\big(A - (1-\mu)\mathbb{1}\big), \tag{15.302}$$

donc $1 - \mu$ est une valeur propre de A. Donc les valeurs propres de $\mathbb{1} - A$ sont les nombres $1 - \lambda_i$ où les λß sont les valeurs propres de A. Par hypothèse, nous avons $\lambda_i < 1$ pour tout i, donc les valeurs propres de $\mathbb{1} - A$ sont toutes non nulles. Donc $\mathbb{1} - A$ est inversible (pas de noyau).

Le reste Nous montrons à présent que si $\rho(A) < 1$ alors $\sum_{k=0}^{\infty} A^k$ converge vers $\mathbb{1} - A$. Pour cela nous savons déjà que $\mathbb{1} - A$ est inversible. Nous posons

$$B_m = \mathbb{1} + A + \ldots + A^m, \tag{15.303}$$

ce qui donne immédiatement $AB_m = A + A^2 + \ldots + A^{m+1}$. Nous avons donc

$$(\mathbb{1} - A)B_m = \mathbb{1} - A^{m+1}. \tag{15.304}$$

Nous savons que $\lim_{m\to 0} A^m = 0$, donc

$$(\mathbb{1} - A)\sum_{k=0}^{\infty} A^k = \lim_{k\to\infty}(\mathbb{1} - A)B_k = \lim_{k\to\infty}(\mathbb{1} - A^{k+1}) = \mathbb{1}. \tag{15.305}$$

Notez au passage que nous avons permuté la somme avec le produit matriciel (voir 10.8.3).

□

15.12.3 Exponentielle et logarithme de matrice

Proposition 15.91.
L'application

$$\begin{aligned}\exp\colon \mathbb{M}(n,\mathbb{R}) &\to \mathbb{M}(n,\mathbb{R})\\ A &\mapsto \sum_{k=0}^{\infty}\frac{A^k}{k!}\end{aligned} \tag{15.306}$$

est une application de classe C^∞. Sa différentielle en zéro est l'identité : $(d\exp)_0 = \mathrm{Id}$.

Démonstration. En ce qui concerne la continuité, nous savons que le rayon de convergence de la suite $\frac{1}{k!}$ est infini ; la proposition 15.72 conclu.

Pour la différentielle, c'est la proposition 15.71 qui nous permet d'écrire

$$d\exp_0(U) = \frac{d}{dt}\Big[\exp(tU)\Big]_{t=0} = \frac{d}{dt}\Big[\sum_{k=0}^{\infty}\frac{t^kU^k}{k!}\Big]_{t=0} = \sum_{k=0}^{\infty}\frac{kt^{k-1}U^k}{k!}\Big|_{t=0} = U \tag{15.307}$$

parce que seul le terme $k = 1$ n'est pas nul. □

Nous avons vu par la proposition 9.196 que toute matrice complexe inversible a un logarithme. Nous allons maintenant parler de logarithme de matrices réelles avec une condition sur la norme. La formule ci-dessous montre explicitement que le logarithme est réel.

$$\begin{aligned}\ln\colon \{A \in \mathbb{M}(n,\mathbb{R}) \text{ tel que } \|A - \mathbb{1}\| < 1\} &\to \mathbb{M}(n,\mathbb{R})\\ A &\mapsto \sum_{k=0}^{\infty}(-1)^k\frac{(A-\mathbb{1})^{k+1}}{k+1}.\end{aligned} \tag{15.308}$$

Lemme 15.92.
Si $\|m\| < 1$ dans $\mathbb{M}(n, \mathbb{R})$, alors nous posons

$$\ln(\mathbb{1} + m) = \sum_{k=0}^{\infty} (-1)^k \frac{m^{k+1}}{k+1}. \tag{15.309}$$

Cette fonction a les propriétés suivantes.

(1) Elle est de classe C^∞.

(2) Elle est un bon logarithme au sens où

$$e^{\ln(\mathbb{1}+m)} = \mathbb{1} + m. \tag{15.310}$$

(3) Elle vérifie l'approximation

$$\ln(1 + m) = m + \sigma(m) \tag{15.311}$$

où σ a la propriété que

$$\lim_{k\to\infty} k\sigma\left(\frac{m}{k}\right) = 0. \tag{15.312}$$

Démonstration. Le rayon de convergence de la suite $a_k = \frac{(-1)^k}{k+1}$ est 1. Donc l'application donnée est C^∞ sur $B(0, 1)$ par le théorème 15.72.

D'après la formule (15.309) nous avons

$$\sigma(m) = \sum_{l=1}^{\infty} (-1)^l \frac{m^{l+1}}{l+1}. \tag{15.313}$$

Nous avons alors

$$k\sigma(\frac{m}{k}) = \sum_{l=1}^{\infty} (-1)^l \frac{m^{l+1}}{k^l(l+1)}, \tag{15.314}$$

et donc

$$\|k\sigma(\frac{m}{k})\| \leqslant \sum_{l=1}^{\infty} \frac{\|m\|^{l+1}}{k^l(l+1)} \leqslant \frac{1}{k} \sum_{l=1}^{\infty} \frac{\|m\|^{l+1}}{l+1} \overset{k\to\infty}{\to} 0 \tag{15.315}$$

Cela prouve la dernière assertion. □

Proposition 15.93.
Soit V un espace vectoriel de dimension finie et $A \in \mathrm{End}(V)$. Nous considérons la fonction

$$\begin{aligned} f\colon \mathbb{R} &\to \mathrm{End}(V) \\ t &\mapsto e^{tA}. \end{aligned} \tag{15.316}$$

Cette fonction vérifie

$$f'(t) = \left(e^{tA}\right)' = Ae^{tA}. \tag{15.317}$$

Démonstration. Si nous posons $f_k(t) = \frac{t^k A^k}{k!}$ alors la fonction f est la somme : $f = \sum_{k=0}^{\infty} f_k$. Nous allons permuter la somme et la dérivation à l'aide du théorème 15.22. Vu que

$$f_k'(t) = \frac{kt^{k-1}A^k}{k!}, \tag{15.318}$$

la suite suite des dérivées converge normalement sur $\mathbb{R}$, nous pouvons dériver terme à terme pour obtenir

$$\Big(\sum_{k=0}^{\infty} t^k \frac{A^k}{k!}\Big) = \sum_{k=0}^{\infty} kt^{k-1}\frac{A^k}{k!} = \sum_{k=1}^{\infty} kt^{k-1}\frac{A^k}{k!} = A\sum_{k=1}^{\infty} \frac{A^{k-1}t^{k-1}}{(k-1)!} = Ae^{tA}. \tag{15.319}$$

Notez le jeu au niveau du point départ de la somme : elle passe de 0 à 1 parce que le terme zéro est nul, mais la simplification $\frac{k}{k!} = \frac{1}{(k-1)!}$ n'a pas de sens pour $k = 0$. □

Lemme 15.94 ([192])**.**
Soit $A \in \mathrm{End}(V)$ où V est un espace vectoriel réel de dimension finie. Si nous notons λ_i ($i = 1, \dots, r$) les valeurs propres distinctes de A alors il existe un polynôme $P \in \mathbb{R}[X]$ tel que

$$\|e^{tA}\| \leqslant P\big(|t|\big) \sum_{i=1}^{r} e^{t \operatorname{Re}(\lambda_i)}. \tag{15.320}$$

Démonstration. Le polynôme caractéristique de A se note, d'après le corollaire 9.157 de la façon suivante :

$$\chi_A(X) = \prod_{i=1}^{r} (X - \lambda_i)^{m_i} \tag{15.321}$$

où m_i est la multiplicité de la valeur propre λ_i. Le lemme des noyaux 9.102 nous dit qu'en posant

$$V_i = \ker(A - \lambda_i \mathbb{1})^{m_i} \tag{15.322}$$

nous avons $V = \bigoplus_{i=1}^{r} V_i$. Nous nommons $p_i \colon V \to V$ la projection canonique de E sur V_i ainsi que x_i la composante de $x \in V$ dans l'espace caractéristique V_i et nous posons $A_i = p_i \circ A$. Les espaces caractéristiques sont stables par A (lemme 9.191), donc $(Ax_i)_i = Ax_i$. Par conséquent $\sum_i A_i p_i = A$ parce que

$$\big(\sum_i p_i A p_i\big)(x) = \sum_i (Ax_i)_i = \sum_i Ax_i = A \sum_i x_i = Ax. \tag{15.323}$$

En ce qui concerne les puissances de A nous avons de même

$$A_i^n x_I = A_i \underbrace{A_i^{n-1} x_i}_{\in V_i} = A A_i^{n-1} x_i = A^n x_i, \tag{15.324}$$

et donc

$$\sum_{i=1}^{r} A_i^n p_i = A^n. \tag{15.325}$$

En particulier,

$$e^{tA} = \sum_i e^{tA_i} p_i. \tag{15.326}$$

C'est de cette exponentielle de matrice que nous devons étudier la norme.

La décomposition de Dunford du théorème 9.194 est toujours un bon plan pour traiter avec les exponentielles : nous avons $A = s + n$ avec

$$s = \sum_k \lambda_k p_k, \quad n = \sum_k (A - \lambda_k \mathbb{1}) p_k. \tag{15.327}$$

Nous montrons que la décomposition de Dunford de $p_i A$ est $p_i A = p_i s + p_i n$. Nous avons

$$p_i s = \sum_k \lambda_k p_i p_k = \lambda_i p_i \tag{15.328}$$

qui est bien diagonalisable. De plus les espaces caractéristiques sont stables par n, donc $p_i n$ est nilpotent. Enfin ils commutent :

$$[p_i s, p_i n] = \lambda_i (p_i n - p_i n p_i). \tag{15.329}$$

Vu que n préserve les espaces caractéristiques, lorsque $v \in V_k$ avec $k \neq i$ nous avons $p_i n p_i v = 0$ et $p_i n v = 0$. Mais si $v \in V_i$ alors

$$p_i n p_i v = p_i n v = n v \tag{15.330}$$

et $p_i n v = n v$, donc les opérateurs $p_i n$ et $p_i n p_i$ sont égaux et (15.329) donne bien zéro. En ce qui concerne l'exponentielle de A_i nous avons

$$e^{p_i A} = e^{p_i s} e^{p_i n} = e^{\lambda_i p_i} \exp\big((A - \lambda_i \mathbb{1}) p_i\big). \tag{15.331}$$

Nous pouvons maintenant sérieusement nous attaquer à la norme de e^{tA} de l'équation (15.326). D'abord nous avons $\|p_i\| = 1$ parce que l'opérateur p_i est l'identité sur au moins un vecteur (en fait tout ceux de l'espace caractéristique V_i). En utilisant les propriétés de la norme opérateur [20], nous trouvons dans un premier temps [21] :

$$\|e^{tA}\| \leqslant \sum_{i=1}^{r} \|e^{tA_i}\| \leqslant \sum_{i=1}^{r} |e^{t\lambda_i}| \underbrace{\sum_{k=0}^{m_i} \frac{|t|^k}{k!} \|A - \lambda_i \mathbb{1}_i\|^k}_{=P_i(|t|}} \tag{15.332}$$

où $\mathbb{1}_i$ est l'opérateur identité sur V_i. Petit détail dans le calcul :

$$\|e^{\lambda_i p_i}\| \leqslant \sum_{l=0}^{\infty} \frac{\lambda_i^l}{k!} \|p_i\|^l = e^{\lambda_i}. \tag{15.333}$$

Notons que tous les termes de $P_i(|t|)$ et $P_i\big(|t|\big)$ sont positifs, de telle sorte que nous pouvons majorer en ajoutant des termes partout. À la place d'avoir $P_i(|t|)$ comme coefficient de $|e^{t\lambda_i}|$ nous majorons en mettant $\sum_{j=1}^{r} P_j(|t|)$ comme coefficient :

$$\|e^{tA}\| \leqslant \sum_{i=1}^{r} |e^{t\lambda_i}| P_i\big(|t|\big) = \sum_{i=1}^{r} |e^{t\lambda_i}| \sum_{j=1}^{r} P_j\big(|t|\big) = P\big(|t|\big) \sum_{i=1}^{r} e^{t\operatorname{Re}(\lambda_i)}. \tag{15.334}$$

L'arrivée de la partie réelle est une égalité usuelle pour les nombres complexes : $|e^{a+bi}| = e^a|e^{bi}| = e^a$. □

15.13 Lemme de Borel

15.13.1 Fonctions plateaux

Soient $a < b < c < d$ dans $\mathbb{R}$. Nous voulons trouver une fonction $f \in C^\infty(\mathbb{R})$ à valeurs positives telle que

(1) $f(x) = 1$ si $x \in [b, c]$

(2) $\operatorname{supp}(f) \subset [a, d]$.

Nous commençons par l'exemple classique de fonction C^∞ qui n'est pas nulle partout :

$$\varphi(x) = \begin{cases} e^{-1/x} & \text{si } x > 0 \\ 0 & \text{sinon.} \end{cases} \tag{15.335}$$

Il est facile de vérifier que φ est de classe C^∞ parce que $\lim_{x\to 0^+} e^{-1/x} P(x) = 0$ pour tout polynôme P. De plus c'est une fonction qui vaut zéro sur $]-\infty, 0]$. Ensuite nous construisons la fonction

$$\psi_m(x) = 1 - \frac{\int_0^x \varphi(t)dt}{\int_0^m \varphi(t)dt} = \begin{cases} 1 & \text{si } x < 0 \\ 0 & \text{si } x > m \\ \text{positive} & \text{si } x \in [0, m]. \end{cases} \tag{15.336}$$

Cette fonction est encore de classe C^∞. À partir de là nous considérons les fonctions

$$f_1(x) = \psi_{d-c}(x - c) = \begin{cases} 1 & \text{si } x < c \\ 0 & \text{si } x > d \\ \text{positive} & \text{si } x \in [c, d]. \end{cases} \tag{15.337a}$$

$$f_2(x) = \psi(b - a)(b - x) = \begin{cases} 0 & \text{si } x < a \\ 1 & \text{si } x > b \\ \text{positive} & \text{si } x \in [a, b]. \end{cases}, \tag{15.337b}$$

20. Surtout le fait que ce soit une norme d'algèbre, lemme 10.14.

21. Si les valeurs propres de A sont λ_i, celles de tA sont $t\lambda_i$.

et finalement la fonction suivante répond à la question des fonctions plateaux sur $\mathbb{R}$:

$$f(x) = f_1(x)f_2(x). \tag{15.338}$$

Une variation sur le même thème est l'existence de fonctions infiniment dérivables à support compact, c'est à dire des fonctions dans $C_c^\infty(\mathbb{R}^d) = \mathscr{D}(\mathbb{R}^d)$. Ces espaces ne sont pas vides, par exemple nous avons la fonction

$$\xi(x) = \begin{cases} e^{-1/(1-|x|^2)} & \text{si } x \in B(0,1) \\ 0 & \text{sinon.} \end{cases} \tag{15.339}$$

15.13.2 Le lemme de Borel

Lemme 15.95 (Lemme de Borel[2])**.**
Soit (a_n) une suite dans $\mathbb{R}$. Il existe une fonction $u \in C^\infty(\mathbb{R})$ telle que $u^{(k)}(0) = a_k$ pour tout $k \geqslant 0$.

Démonstration. Soit $\varphi \in C_c^\infty(\mathbb{R})$ une fonction telle que $\varphi(x) = 1$ si $|x| \leqslant \frac{1}{2}$ et telle que $\text{supp}(\varphi) \subset]-1,1[$.

Nous commençons par considérer une suite de réels strictement positifs (λ_k) dont nous fixerons une valeur précise plus tard, et nous posons

$$f_k(x) = \varphi(\lambda_k x)\frac{a_k}{k!}x^k. \tag{15.340}$$

Nous allons étudier la convergence et les propriétés de $u(x) = \sum_{k=0}^\infty f_k(x)$.

Calculons (formellement) la m^{e}dérivée de f_k :

$$f_k^{(m)}(x) = \frac{a_k}{k!}\sum_{l=0}^m \binom{m}{l}\lambda_k^{m-l}\varphi^{(m-l)}(\lambda_k x)(x^k)^{(l)} \tag{15.341a}$$

$$= a_k \sum_{l=0}^m \binom{l}{m}\lambda_k^{m-l}\varphi^{(m-l)}(\lambda_k x)\frac{x^{k-l}}{(k-l)!}. \tag{15.341b}$$

Notons que nous travaillons à m fixé et que nous ne nous intéressons qu'aux termes avec k assez grand ; nous pouvons donc supposer $k \geqslant m$. De toutes façons pour $\sum_{k=0}^m f_k$, on a la classe C^∞, et la permutation de la somme avec tout ce qu'on veut. Vu que φ est continue à support compact nous pouvons poser

$$M_m = \max_{0\leqslant j\leqslant m}\|\varphi^j\|_\infty = \max_{0\leqslant j\leqslant m}\max_{x\in\mathbb{R}}|\varphi^{(j)}(x)|. \tag{15.342}$$

Nous continuons en nous fixant un $x \in \mathbb{R}$ et un $k \geqslant m$.

Si $|x| > \frac{1}{\lambda_k}$, alors $\varphi^{(m)}(\lambda_k x) = 0$ parce que $\lambda_k x$ est strictement hors du support de φ qui est $]-1,1[$. Donc pour $|x| > \frac{1}{\lambda_k}$.

Si par contre $|x| \leqslant \frac{1}{\lambda_k}$, nous avons les majorations

$$|f_k^{(m)}(x)| \leqslant |a_k|\sum_{l=0}^m \binom{l}{m}|\lambda_k|^{m-l}\underbrace{\varphi^{(m-l)}(\lambda_k x)}_{\leqslant M_m}\frac{1}{(k-l)!}\underbrace{|x|^{k-l}}_{\leqslant(1/\lambda_k)^{k-l}} \tag{15.343a}$$

$$\leqslant |a_k|M_m|\lambda_k|^{m-k}\frac{1}{(k-m)!}\sum_{l=0}^m\binom{m}{l} \tag{15.343b}$$

$$\leqslant \frac{|a_k|M_m|\lambda_k|^{m-k}2^m}{(k-m)!} \tag{15.343c}$$

$$= \frac{|a_k|M_m 2^m}{(k-m)!|\lambda_k|^{k-m}} \tag{15.343d}$$

où pour faire disparaître la somme de coefficients binomiaux, nous avons remarqué que $\sum_{l=0}^{m}\binom{m}{l}$ est le nombre total de termes dans le développement de $(a+b)^m$, c'est à dire 2^m. Nous voulons, pour m fixé, étudier la convergence de la somme de cela. Notons que le 2^m n'a en particulier strictement aucune importance parce qu'on travaille à m fixé.

Nous fixons maintenant la valeur des λ_k :

$$\lambda_k = \max\{|a_k|, 1\}. \tag{15.344}$$

Avec cela, en nous souvenant que nous n'étudions que les termes $k > m$, le dénominateur de (15.343d) est réellement croissant en k, donc nous avons la majoration

$$|f_k^{(m)}(x)| \leqslant \frac{M_m 2^m}{(k-m)!}. \tag{15.345}$$

Au final nous avons

$$\|f_k^{(m)}\|_\infty \leqslant \frac{2^m M_m}{(k-m)!}. \tag{15.346}$$

Et la somme de cela converge sans difficultés. Donc la série

$$u(x) = \sum_{k=0}^{\infty} f_k^{(m)}(x) \tag{15.347}$$

converge normalement et donc uniformément sur $\mathbb{R}$. Nous pouvons alors permuter la somme et la dérivation par le théorème 15.20. Donc

$$u^{(m)} = \sum_{k=0}^{\infty} f_k^{(m)} \tag{15.348}$$

est continue. En particulier, pour évaluer en zéro, on peut faire

$$u^{(m)}(0) = \sum_{k=0}^{\infty} f_k^{(m)}(0). \tag{15.349}$$

Nous avons

$$f_k(x) = \varphi(\lambda_k x)\frac{a_k}{k!}x^k. \tag{15.350}$$

Pour calculer la dérivée en zéro, il suffit de la calculer sur un voisinage sur lequel $\varphi(\lambda_k x)$ est la constante 1 ; un tel voisinage existe pour tout k. À ce moment le calcul est classique :

$$f_k^{(m)}(x) = \begin{cases} a_k & \text{si } k = m \\ 0 & \text{sinon.} \end{cases} \tag{15.351}$$

Finalement nous avons bien

$$u^{(m)}(0) = \sum_{k=0}^{\infty} f_k^{(m)}(0) = a_k. \tag{15.352}$$

□

Remarque 15.96.
Pour prouver le lemme de Borel, la première chose qui passe par la tête est la fonction toute simple

$$u(x) = \sum_{k=0}^{\infty} \frac{a_k}{k!}x^k. \tag{15.353}$$

Évidemment si on calcule les dérivées successives de cette fonction, nous trouvons les bons résultats. Le problème est la convergence. Rien qu'en prenant $a_k = k!k^k$, la série ne converge pour aucun x positif. L'idée de multiplier chacun de f_k par une fonction plateau sur un petit intervalle autour de zéro a plusieurs avantages. D'abord on conserve les dérivées correctes parce qu'on ne touche pas à la valeur des f_k sur un petit voisinage. Ensuite cela ne modifie pas la continuité ; et enfin en multipliant par $\varphi(\lambda_k x)$, ça calme méchamment les divergences parce que $\lambda_k x$ passe vite au dessus de 1 (et donc en dehors du support de φ) si λ_k est grand. D'où le fait qu'il soit normal que les λ_k soient de l'ordre des a_k.