Le Frido 2017,
volume 3
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”.

ISBN : 979-10-97085-03-2

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

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

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

12.10Continuité et dérivabilité . . . 557
12.11Dérivation 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
12.12Opérations sur les dérivées . . . 561
12.12.1 Développement limité au premier ordre . . . 562
12.13Dérivation à une variable . . . 563
12.13.1 Définition . . . 563
12.13.2 Quelques formules à connaître . . . 563
12.14Dérivation 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
12.15Fonctions de plusieurs variables . . . 570
12.15.1 Graphes de fonctions à plusieurs variables . . . 573
12.15.2 Courbes de niveau . . . 574
12.16Dérivée 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
12.17Dérivée directionnelle de fonctions composées . . . 584
12.18Formes 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
12.19Différentielle . . . 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
14.10Fonctions 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
14.11Formes différentielles exactes et fermées . . . 835
14.12Théorème 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
14.13Rappel sur les intégrales usuelles . . . 847
14.14Intégrales le long de chemins . . . 848
14.14.1 Circulation d'un champ de vecteur . . . 848
14.15Circulation d'un champ conservatif . . . 850
14.16Intégration 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
14.17Les intégrales triples . . . 858
14.17.1 Volume . . . 860
14.18Un 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
14.19Aire et primitive . . . 863
14.19.1 Longueur d'arc de courbe . . . 864
14.19.2 Aire de révolution . . . 864
14.20L'aire en dessous d'une courbe . . . 864
14.21Propriétés des intégrales . . . 866
14.22Techniques 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 16

Trigonométrie, isométries

16.1 Isométries de l'espace euclidien

Nous considérons l'espace affine euclidien $A = \mathcal{E}_n(\mathbb{R})$ modelé sur $\mathbb{R}^n$ avec sa métrique usuelle. Un premier grand résultat sera le théorème 16.4 qui dira que les isométries de cet espace sont des applications linéaires.

16.1.1 Structure du groupe $\mathrm{Isom}(\mathbb{R}^n)$

Exemple 16.1
La forme quadratique $q(x) = x_1^2 + x_2^2$ donne la norme euclidienne. La forme bilinéaire associée est $b(x, y) = x_1y_1 + x_2y_2$, qui est le produit scalaire usuel. △

Il ne faudrait pas déduire trop vite que la formule $\|x\|^2 = q(x)$ donne une norme dès que q est non dégénérée. En effet q peut ne pas être définie positive. La forme $q(x) = x_1^2 - x_2^2$ prend des valeurs positives et négatives. A fortiori $d(x, y) = q(x - y)$ ne donne pas toujours une distance.

Définition 16.2.
*Une **isométrie** pour la forme quadratique q est une application bijective $f\colon V \to V$ telle que $q(x-y) = q\big(f(x)-f(y)\big)$. Dans les cas où q donne une distance, alors c'est une isométrie au sens usuel.*

Lemme 16.3.
Soit q une forme quadratique et b la forme bilinéaire associée par le lemme 9.243. Pour une application bijective $f\colon E \to E$ telle que $f(0) = 0$, les conditions suivantes sont équivalentes :

(1) $b\big(f(x), f(y)\big) = b(x, y)$ pour tout $x, y \in E$;

(2) $q\big(f(x) - f(y)\big) = q(x - y)$ pour tout $x, y \in E$.

Démonstration. Dans le sens direct, en posant $x = y$ nous trouvons tout de suite $q(f(x)) = q(f)$; ensuite en utilisant la distributivité de b,

$$\begin{aligned} q\big(f(x) - f(y)\big) &= b\big(f(x) - f(y), f(x) - f(y)\big) && \text{(16.1a)}\\ &= q\big(f(x)\big) - 2b\big(f(x), f(y)\big) + q\big(f(y)\big) && \text{(16.1b)}\\ &= q(x) + q(y) - 2b(x, y) && \text{(16.1c)}\\ &= q(x - y). && \text{(16.1d)} \end{aligned}$$

Dans l'autre sens, nous commençons par remarquer que l'hypothèse $f(0) = 0$ donne $q(x) =$

$q\big(f(x)\big)$. Ensuite nous utilisons l'identité de polarisation (9.563) :

$$b\big(f(x),f(y)\big) = \frac{1}{2}\big[q\big(f(x)\big) + q\big(f(y)\big) - q\big(f(x-y)\big)\big] \tag{16.2a}$$
$$= \frac{1}{2}\big[q(x) + q(y) - q(x-y)\big] \tag{16.2b}$$
$$= b(x,y). \tag{16.2c}$$

□

Théorème 16.4 ([193])**.**
Soit $f\colon E \to E$ une bijection telle que

$$q(x-y) = q\big(f(x) - f(y)\big) \tag{16.3}$$

pour tout $x, y \in E$. Alors

(1) si $f(0) = 0$, alors f est linéaire ;

(2) si $f(0) \neq 0$ alors f est affine[1]

Démonstration. Si $f(0) = 0$, nous savons par le lemme 16.3 que $b\big(f(x),f(y)\big) = b(x,y)$. Soit $z \in E$; étant donné que f est bijective nous pouvons considérer l'élément $f^{-1}(z) \in E$ et calculer

$$b\big(f(x+y),z\big) = b\big(f(x+y), f(f^{-1}(z))\big) \tag{16.4a}$$
$$= b(x+y, f^{-1}(z)) \tag{16.4b}$$
$$= b(x, f^{-1}(z)) + b(y, f^{-1}(z)) \tag{16.4c}$$
$$= b(f(x),z) + b(f(y),z) \tag{16.4d}$$
$$= b\big(f(x)+f(y),z\big), \tag{16.4e}$$

donc $f(x+y) = f(x) + f(y)$ par le lemme 9.249.

De la même façon on trouve $b\big(f(\lambda x),z\big) = b\big(\lambda f(x),z\big)$ qui prouve que $f(\lambda x) = \lambda f(x)$ et donc que f est linéaire.

Si $f(0) \neq 0$, alors nous posons $g(x) = f(x) - f(0)$ qui vérifie $g(0) = 0$ et

$$q\big(g(x) - g(y)\big) = q\big(f(x) - f(0) - f(y) + f(0)\big) = q(x-y). \tag{16.5}$$

Nous pouvons donc appliquer le premier point à g, déduire que g est linéaire et donc que f est affine. □

Nous pouvons maintenant particulariser tout cela au cas de $\mathbb{R}^n$ muni du produit scalaire usuel et de la norme associée pour voir quel résultat nous avons à peine prouvé.

Lemme 16.5 ([194])**.**
Une isométrie est bijective (nous sommes en dimension finie).

Démonstration. Si $f\colon E \to E$ est une isométrie, elle est linéaire par le théorème 16.4. Elle vérifie également $\|f(x)\| = \|x\|$, et donc $f(x) = 0$ si et seulement si $x = 0$, c'est à dire que f est injective. Elle est alors bijective par le corollaire 7.24 du théorème du rang. □

Nous notons ici $T(n)$ le groupe des translations sur $\mathbb{R}^n$. Un élément de $T(n)$ est une translation τ_v donnée par un vecteur v et agissant sur $\mathbb{R}^n$ par

$$\begin{aligned}\tau_v\colon \mathbb{R}^n &\to \mathbb{R}^n\\ x &\mapsto x + v.\end{aligned} \tag{16.6}$$

Ce groupe est isomorphe au groupe abélien $(\mathbb{R}^n, +)$, et nous allons souvent identifier τ_v à v.

Si vous ne voulez pas savoir ce qu'est un produit semi-direct de groupes, vous pouvez lire seulement le point (1) du théorème suivant, et passer directement à la remarque 16.7.

1. Lemme 11.48.

Théorème 16.6.
Un peu de structure sur $\mathrm{Isom}(\mathbb{R}^n)$.

(1) L'application

$$\begin{aligned}\psi\colon T(n)\times \mathrm{O}(n) &\to \mathrm{Isom}(\mathbb{R}^n)\\ (v,\Lambda) &\mapsto \tau_v\circ\Lambda\end{aligned} \tag{16.7}$$

est une bijection. Ici, $T(n)$ *est le groupe des translations de* $\mathbb{R}^n$.

(2) Un couple $(v,\Lambda)\in T(n)\times \mathrm{SO}(n)$ *agit sur* $x\in\mathbb{R}^n$ *par*

$$(v,\Lambda)x=\Lambda x+v \tag{16.8}$$

au sens où $\psi(v,\Lambda)x=\Lambda x+v$.

(3) En tant que groupes,

$$\mathrm{Isom}(\mathbb{R}^n)\simeq T(n)\times_\rho \mathrm{O}(n) \tag{16.9}$$

où ρ *représente l'action adjointe de* $\mathrm{O}(n)$ *sur* $T(n)$ *et* $\times_\rho$ *dénotes le produit semi-direct de la définition 3.93.*

Démonstration. Point par point.

(1) Prouvons que l'application proposée est injective et surjective. Notons aussi que ce point ne parle pas de structure de groupe, mais seulement d'une bijection en tant qu'ensembles.

Injection Si $\psi(v,\Lambda)=\psi(w,\Lambda')$ alors en appliquant sur $x=0$ nous avons tout de suite $v=w$. Et ensuite $\Lambda=\Lambda'$ est immédiat.

Surjection Une isométrie $g\in\mathrm{Isom}(\mathbb{R}^n)$ est une application $g\colon\mathbb{R}^n\to\mathbb{R}^n$ telle que $d(x,y)=d\big(g(x),g(y)\big)$. Dans le cas de $\mathbb{R}^n$ cela se traduit par

$$\|x-y\|=\big\|g(x)-g(y)\big\|, \tag{16.10}$$

Vu que $x\mapsto\|x\|$ est une forme quadratique, elle tombe sous le coup du théorème 16.4, ce qui nous permet de dire que g est affine. Or par définition une application est affine lorsqu'elle est la composée d'une translation et d'une application linéaire.

(2) C'est seulement le fait que $(\tau_v\circ\Lambda)x=\tau_v\big(\Lambda x\big)=\Lambda(x)+v$.

(3) Nous allons étudier l'application

$$\psi\colon T(n)\times_\rho O(n)\to\mathrm{Isom}(\mathbb{R}^n). \tag{16.11}$$

Le produit semi-direct est bien définit Il faut montrer que

$$\begin{aligned}\rho\colon O(n)&\to \mathrm{Aut}\big(T(n)\big)\\ \Lambda&\mapsto \mathbf{Ad}(\Lambda)\end{aligned} \tag{16.12}$$

est correcte.

D'abord pour $\Lambda\in O(n)$, nous avons bien $\rho_\Lambda(\tau_v)\in T(n)$ parce qu'en appliquant à $x\in\mathbb{R}^n$,

$$(\Lambda\tau_v\Lambda^{-1})(x)=\Lambda\big(\tau_v(\Lambda^{-1}x)\big)=\Lambda\big(\Lambda^{-1}x+v\big)=x+\Lambda(v)=\tau_{\Lambda(v)}(x). \tag{16.13}$$

Donc $\rho_\Lambda(\tau_v)=\tau_{\Lambda(v)}$.

De plus, $\rho_\Lambda\in\mathrm{Aut}\big(T(n)\big)$ parce que

$$\rho_\Lambda\big(\tau_v\circ\tau_w\big)=\rho_\Lambda(\tau_v)\circ\rho_\Lambda(\tau_v), \tag{16.14}$$

comme on peut aisément vérifier que les deux membres sont égaux à $\tau_{\Lambda(v+w)}$.

ψ est une bijection Cela est déjà vérifié.

ψ **est un homomorphisme** Nous avons d'une part

$$\psi\big((v,g)(w,h)\big) = \psi\big(v\rho_g(w), gh\big) = \tau_v \circ g \circ \tau_w \circ g^{-1} \circ g \circ h = \tau_v \circ g \circ \tau_w \circ h. \tag{16.15}$$

Et d'autre part,

$$\psi(v,g) \circ \psi(w,h) = \tau_v \circ g \circ \tau_w \circ h, \tag{16.16}$$

ce qui est la même chose.

□

Remarque 16.7.
Notons au passage la loi de groupe sur les couples qui est donnée, pour tout $v, v' \in \mathbb{R}^n$, $\Lambda, \Lambda' \in \mathrm{SO}(n)$, par

$$(v,\Lambda) \cdot (v',\Lambda') = (\Lambda v' + v, \Lambda\Lambda') \tag{16.17}$$

comme le montre le calcul suivant :

$$\begin{aligned}(v,\Lambda) \cdot (v',\Lambda')x &= (v,\Lambda)(\Lambda' x + v') && (16.18a)\\ &= \Lambda\Lambda' x + \Lambda v' + v && (16.18b)\\ &= (\Lambda v' + v, \Lambda\Lambda')x. && (16.18c)\end{aligned}$$

Proposition 16.8 ([195])**.**
Soit $n \geqslant 1$ et un élément R de $\mathrm{O}(n)$ de déterminant -1 tel que $R^2 = \mathrm{Id}$. En posant $C_2 = \{\mathrm{Id}, R\}$ nous avons

$$\mathrm{O}(n) = \mathrm{SO}(n) \times_\rho C_2 \tag{16.19}$$

Démonstration. Notons que pour R nous pouvons prendre par exemple $(x_1, \dots, x_n) \mapsto (-x_1, x_2, \dots, x_n)$. Ce que nous allons montrer être un isomorphisme est :

$$\begin{aligned}\psi\colon \mathrm{SO}(n) \times C_2 &\to \mathrm{O}(n)\\ (A,h) &\mapsto Ah.\end{aligned} \tag{16.20}$$

Injectif Soient $A, B \in \mathrm{SO}(n)$ et $h, k \in C_2$ tels que $\psi(A,h) = \psi(B,k)$, c'est à dire tels que $Ah = Bk$. Vu que $\det(A) = \det(B) = 1$ nous avons $\det(h) = \det(k)$. Mais comme C_2 contient un élément de déterminant 1 et un élément de déterminant -1, nous avons $h = k$. De là $A = B$.

Surjectif Soit $X \in \mathrm{O}(n)$. Si $\det(X) = 1$ alors $X \in \mathrm{SO}(n)$ et $X = \psi(X, \mathbb{1})$. Si par contre $\det(X) = -1$ alors $XR \in \mathrm{SO}(n)$ parce que $\det(XR) = 1$ et nous avons

$$\psi(XR, R) = XR^2 = X. \tag{16.21}$$

Homomorphisme Nous avons

$$\psi\Big((A,h)(B,k)\Big) = \psi\big(A\rho_h(B), hk\big) = A(hBh^{-1})hk = AhBk, \tag{16.22}$$

tandis que

$$\psi(A,h)\psi(B,k) = AhBk, \tag{16.23}$$

qui est la même chose.

□

Lemme 16.9 ([196])**.**
Si $n \geqslant 3$, alors toute droite est intersection de deux plans non isotropes.

Proposition 16.10 ([195])**.**
Si une isométrie de $\mathbb{R}^n$ fixe un ensemble F de points, alors elle fixe l'espace affine engendrée par F.

Démonstration. Soit $f \in \mathrm{Isom}(\mathbb{R}^n)$ fixant F. Par le théorème 16.4, c'est une application affine et l'ensemble $\mathrm{Fix}(f)$ des points fixés par f est un sous-espace affine de $\mathbb{R}^n$, grâce à la proposition 11.50.

Donc $\mathrm{Fix}(f)$ est un espace affine contenant F. Vu que l'espace affine engendré par F est l'intersection de tous les espace affines contenant F, il est en particulier contenu dans $\mathrm{Fix}(f)$. □

Corollaire 16.11.
Si f et g sont des isométries de $\mathbb{R}^n$ qui coïncident sur F, alors elles coïncident sur l'espace affine engendré par F.

Démonstration. Nous considérons $h = g^{-1} \circ f$ qui est une isométrie de $\mathbb{R}^n$ fixant F. Elle fixe donc, par la proposition 16.10, l'espace affine engendré par F. Or tout point fixé par h est un point sur lequel g et f coïncident. □

16.1.2 Classification des isométries de $\mathbb{R}$

Définition 16.12.
*Soit $x \in \mathbb{R}$; nous notons σ_x la **réflexion** par rapport à x, c'est à dire*

$$\sigma_x(y) = 2x - y. \tag{16.24}$$

Théorème 16.13 ([195])**.**
Toute isométrie de $\mathbb{R}$ est composée d'au plus 2 réflexions. Plus précisément toute isométrie de $\mathbb{R}$ est dans une des trois catégories suivantes :

- *l'identité (0 réflexions),*
- *les réflexions,*
- *les translations (2 réflexions)*

Démonstration. Nous divisions la preuve en fonction du nombre de points fixés par l'isométrie $f \in \mathrm{Isom}(\mathbb{R})$.

f fixe deux points distincts Alors elle fixe l'espace affine engendrée par ces deux points par la proposition 16.10. Donc f fixe tout $\mathbb{R}$ et est l'identité.

f fixe un unique point Soit x l'unique point fixé par f et considérons $y \neq x$. Vu que $x = f(x)$ et que f est une isométrie,

$$d\big(x, f(y)\big) = d\big(f(x), f(y)\big) = d(x, y). \tag{16.25}$$

Donc $f(y)$ est à égale distance de x que y. Autrement dit, $f(y)$ est soit y soit $\sigma_x(y)$. Mais comme x est unique point fixe, $f(y) = \sigma_x(y)$. Ce raisonnement étant valable pour tout $y \neq x$ nous avons $f = \sigma_x$.

f n'a pas de points fixes Soit $x \in \mathbb{R}$ et $y = \frac{x+f(x)}{2}$. Nous posons $g = \sigma_y \circ f$. Alors x est un point fixe de g parce que

$$g(x) = \sigma_y\big(f(x)\big) = 2y - f(x) = x. \tag{16.26}$$

Donc soit g est l'identité soit g est une réflexion (par les points précédents). La possibilité $g = \mathrm{Id}$ est exclue parce que cela ferait $f = \sigma_y$ alors que f n'a pas de points fixes. Donc g est une réflexion ; et comme x est un point fixe de g nous avons $g = \sigma_x$. Au final

$$f = \sigma_y \circ \sigma_x. \tag{16.27}$$

Montrons que cela implique que f est une translation :

$$\sigma_y\sigma_x(z) = \sigma_y(2x - z) = 2y - 2x + z = z + 2(y - x). \tag{16.28}$$

Donc $\sigma_y \circ \sigma_x$ est la translation de vecteur $2(y - x)$.

□

16.1.3 Segment, plan médiateur et équidistance

Lemme 16.14.
Un point M est sur la médiatrice du segment $[A, B]$ si et seulement si $\|M - A\| = \|M - B\|$.

Lemme 16.15.
Soient A et B de points de $\mathbb{R}^3$. Alors le plan médiateur du segment $[A, B]$ est le lieu des points de $\mathbb{R}^3$ à être équidistants de A et B.

Démonstration. Nous nommons σ ce plan.

Soit X un point équidistant de A et B. Alors dans le plan (A, B, X), le triangle ABX est isocèle en X, et la hauteur issue de X coupe perpendiculairement $[A, B]$ en son milieu. Cela prouve que X est dans le plan médiateur du segment $[A, B]$ (lemme 16.14).

Mettons au contraire que X est dans le plan médiateur de $[A, B]$. Nous avons $(X, M) \perp (A, B)$. Donc le triangle A, B, X est isocèle en X et donc X est équidistant de A et B. □

16.1.4 Isométries du tétraèdre régulier

Définition 16.16.
*Un polyèdre **régulier** est un polyèdre dont les faces sont des polygones réguliers identiques dont tous les sommets joignent le même nombre d'arrêtes.*

Le **tétraèdre** est une pyramide à base triangulaire dont toutes les faces sont des triangles équilatéraux.

Proposition 16.17 (Isométries affines du tétraèdre régulier)**.**
Soit un tétraèdre régulier T et son groupe d'isométries affines $\mathrm{Iso}(T)$ *(définition 11.54). Alors*

$$\mathrm{Iso}(T) \simeq S_4 \tag{16.29}$$

où S_4 est le groupe des permutations de quatre objets.

Démonstration. Commençons par prouver qu'une isométrie préserve les sommets : l'image d'un sommet est un sommet. Pour cela nous considérons $g \in \mathrm{Iso}(T)$ et nous supposons que l'image d'un sommet x soit à l'intérieur d'une arrête. Soient $g(y)$ et $g(z)$ deux points distincts de cette arrête situés à égale distance de $g(x)$. Cela est possible parce que g est une bijection de $\mathbb{R}^3$. Aussi : $y \neq z$. Mais une application affine préserve l'alignement (vous ne le croyez pas ? regardez la forme donnée par le lemme (11.48)), donc x, y et z foment un triangle isocèle en x de points alignés et appartenant à T. Cela est impossible si x est un sommet.

Donc l'image d'un sommet est un sommet. Si nous numérotons les sommets $x_1, \ldots, x_4$, nous obtenons un morphisme de groupe $\varphi\colon \mathrm{Isom}(T) \to S_4$ qui envoie g sur la permutation qui envoie 1 sur le numéro du sommet $g(x_1)$, 2 sur le numéro du sommet $g(x_2)$, etc.

Le morphisme φ est injectif Supposons $\varphi(g_1) = \varphi(g_2)$. Alors $g_1^{-1} \circ g_2$ est une isométrie de $(\mathbb{R}^3, d)$ qui fixe les quatre sommets. Une application affine $\mathbb{R}^4 \to \mathbb{R}^3$ fixant 4 point est l'identité par le lemme 11.45. Donc $g_1^{-1} g_2 = \mathrm{Id}$, ce qui prouve que $g_1 = g_2$. Vous noterez que nous utilisons l'unicité de l'inverse dans un groupe.

φ est surjectf Nous savons que S_4 est engendré par les transpositions (proposition 8.38). Or les transpositions sont dans l'image de φ. En effet, notons les sommets de notre tétraèdre par A, B, D et D et considérons la transposition $A \leftrightarrow B$. Elle est l'image par φ de la réflexion selon le plan σ, médiateur du segment $[A, B]$. Pour nous assurer de cela, nous devons nous assurer que C et D appartiennent à σ. Cela est le contenu du lemme 16.15.

Conclusion L'application φ est un morphisme bijectif, c'est à dire un isomorphisme.

□

16.18.
Lorsque le tétraèdre a son barycentre en l'origine de $\mathbb{R}^3$, l'isomorphisme $\varphi\colon \mathrm{Iso}(T) \to S_4$ donne

une représentation de dimension 3 de S_4. Nous allons calculer les caractères de S_4 en la section 17.5 sans avoir besoin de savoir que l'une des représentations de dimension 3 est cella que nous venons de trouver via le groupe des isométries du tétraèdre. Nous allons cependant également y calculer les caractères de la représentations φ, pour le sport.

Plus de détails en 17.5.2.

16.2 Trigonométrie

16.2.1 Définitions, périodicité et quelque valeurs remarquables

Proposition-définition 16.19 (Défintion du cosinus et du sinus)**.**
La série

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \tag{16.30}$$

définit une fonction $\cos\colon \mathbb{R} \to \mathbb{R}$ *de classe* C^∞*. Nous l'appelons* ***cosinus***

La série

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \tag{16.31}$$

définit une fonction $\sin\colon \mathbb{R} \to \mathbb{R}$ *de classe* C^∞*. Nous l'appelons* ***sinus***

Démonstration. La série entière définissant $\cos(x)$ a pour coefficients

$$a_n = \begin{cases} 0 & \text{si } n \text{ est impair} \\ \frac{(-1)^{n/2}}{n!} & \text{si } n \text{ est pair.} \end{cases} \tag{16.32}$$

Nous pouvons la majorer par la série entière donnée par les coefficients

$$b_n = \begin{cases} 1/n! & \text{si } n \text{ est impair} \\ \frac{(-1)^{n/2}}{n!} & \text{si } n \text{ est pair.} \end{cases} \tag{16.33}$$

Quelle que soit la parité de k nous avons toujours

$$\left|\frac{b_{k+1}}{b_k}\right| = \frac{1}{k+1}, \tag{16.34}$$

de telle sorte que la formule d'Hadamard (15.60) nous donne $R = \infty$ pour la série $\sum_{k=0}^{\infty} b_k x^k$. A fortiori [2] le rayon de convergence pour la série du cosinus est infini.

L'assertion concernant le sinus se démontre de même.

En ce qui concerne le fait que les fonctions sin et cos sont de classe C^∞ sur $\mathbb{R}$, il faut invoquer le corollaire 15.43. □

Lemme 16.20.
En ce qui concerne la dérivation, nous avons

$$\sin' = \cos \tag{16.35a}$$
$$\cos' = -\sin. \tag{16.35b}$$

Démonstration. Il s'agit de se permettre de dériver terme à terme (proposition 15.41) les séries qui définissent le sinus et le cosinus. □

Lemme 16.21.
Les fonctions sinus et cosinus vérifient

$$\cos^2(x) + \sin^2(x) = 1 \tag{16.36}$$

pour tout $x \in \mathbb{R}$*.*

2. Remarque 15.25.

Démonstration. Posons $f(x) = \sin^2(x) + \cos^2(x)$ et dérivons :

$$f'(x) = 2\sin(x)\cos(x) + 2\cos(x)(-)\sin(x) = 0. \tag{16.37}$$

La fonction f est donc constante par le corollaire 12.133. Nous avons donc pour tout x :

$$f(x) = f(0) = \sin^2(0) + \cos^2(0) = 1. \tag{16.38}$$

Le dernier calcul s'obtient en substituant directement x par zéro dans les séries : $\sin(0) = 0$ et $\cos(0) = 1$. □

Lemme 16.22.
Nous avons la formule

$$e^{ix} = \cos(x) + i\sin(x) \tag{16.39}$$

pour tout $x \in \mathbb{R}$.

Démonstration. Il faut partir de la définition de l'exponentielle (15.121), et remarquer que i^k vaut 1, i, -1, $-i$. Donc un terme sur deux est imaginaire pur et parmi ceux-là, un sur deux est positif. À bien y regarder, les termes imaginaires purs forment la série du sinus et ceux réels la série du cosinus. □

Lemme 16.23.
Nous avons les formules d'addition d'angles

$$\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b) \tag{16.40a}$$
$$\sin(a+b) = \cos(a)\sin(b) + \sin(a)\cos(b) \tag{16.40b}$$
$$\cos(a-b) = \cos(a)\cos(b) + \sin(a)\sin(b) \tag{16.40c}$$

pour tout a, b réels.

Démonstration. Nous utilisons la formule d'addition dans l'exponentielle, proposition (15.151) et la formule (16.39) avant de séparer les parties réelles et imaginaires :

$$e^{i(a+b)} = e^{ia}e^{ib} = \cos(a)\cos(b) - \sin(a)\sin(b) + i\big(\cos(a)\sin(b) + \sin(a)\cos(b)\big). \tag{16.41}$$

Cela est également égal à

$$\cos(a+b) + i\sin(a+b). \tag{16.42}$$

En identifiant les parties réelle et imaginaires, nous obtenons les formules (16.40a) et (16.40c) annoncées.

Pour la formule (16.40c), il suffit de se souvenir que $\sin(-b) = -\sin(b)$ et $\cos(-b) = \cos(b)$ (ces deux égalités sont immédiatement visibles sur les développements en série : l'un a uniquement des puissances paires et l'autre impaires) et d'écrire (16.40a) avec $-b$ au lieu de b. □

Corollaire 16.24.
Les formules suivantes pour les duplications d'angles s'ensuivent :

$$\cos(2a) = \cos^2(a) - \sin^2(a) \tag{16.43a}$$
$$\sin(2a) = 2\cos(a)\sin(a). \tag{16.43b}$$

Démonstration. Poser $b = a$ dans les relations du lemme 16.23. □

Lemme 16.25.
Un sous-groupe de $(\mathbb{R}, +)$ est soit dense dans $\mathbb{R}$ soit de la forme $p\mathbb{Z}$ pour un certain réel $p \neq 0$.

Démonstration. Soit A, un sous groupe de $(\mathbb{R},+)$ qui ne soit pas dense. Soit un intervalle $]a,b[$ qui n'intersecte pas A (si vous voulez frimer, vous noterez ici que nous utilisons le fait que les intervalles ouverts forment une base de la topologie de $\mathbb{R}$). Si $d=|b-a|$, l'ensemble A ne contient pas deux éléments séparés par strictement moins de d. Soit p, le plus petit élément strictement positif de A ; nous avons $p\geqslant d$ (parce que $0\in A$ de toutes façons).

Vu que A est un groupe nous avons $p\mathbb{Z}\subset A$.

Pour l'inclusion inverse, si $x\in A$ est hors de $p\mathbb{Z}$, il existe un $y\in p\mathbb{Z}$ avec $|x-y|<p$. Et donc le nombre $|x-y|$ est dans A tout en étant plus petit que p. Contradiction. □

Proposition 16.26 ([197])**.**
La fonction cos *est périodique et le nombre $T>0$ est une période si et seulement si* $\cos(T)=1$ *et* $\sin(T)=0$.

Démonstration. Plusieurs étapes.

La fonction cosinus n'est pas toujours positive Supposons d'abord que $\cos(x)>0$ pour tout $x\in\mathbb{R}$. Dans ce cas, la fonction sin est strictement croissante. Mais les deux fonctions sont bornées par 1 du fait de la formule $\cos^2(x)+\sin^2(x)=1$. La fonction sin étant croissante et bornée, elle est convergente vers un réel par la proposition 12.88 :

$$\lim_{x\to\infty}\sin(x)=\ell \tag{16.44}$$

pour un certain $\ell>0$. Avec ça nous avons aussi (pour cause de dérivée) $\lim_{x\to\infty}\sin'(x)=0$, c'est à dire $\lim_{x\to\infty}\cos(x)=0$. Mais vu que $\cos^2(x)+\sin^2(x)=1$ nous avons $\lim_{x\to\infty}\sin(x)=1$. Mézalor $\lim_{x\to\infty}\cos'(x)=-1$, ce qui donne que la fonction cos n'est pas bornée. Cela est impossible. Nous en déduisons que $\cos(x)$ n'est pas toujours positive.

Il existe $T>0$ tel que $\cos(T)=1$ et $\sin(T)=0$ Par ce que nous venons de faire, il existe $r>0$ tel que $\cos(r)=0$. Pour cette valeur, nous avons aussi obligatoirement $\sin(r)=\pm1$. Nous avons aussi, en utilisant les formules (16.40),

$$\cos(2r)=\cos^2(r)-\sin^2(r)=-1 \tag{16.45a}$$
$$\sin(2r)=2\cos(r)\sin(r)=0. \tag{16.45b}$$

et par conséquent

$$\cos(4r)=\cos^2(2r)-\sin^2(2r)=1 \tag{16.46a}$$
$$\sin(4r)=2\cos(2r)\sin(2r)=0. \tag{16.46b}$$

Donc $T=4r$ fonctionne.

Si T est une période Nous entrons dans le vif de la preuve. Soit un $T>0$ tel que $\cos(x+T)=\cos(x)$ pour tout $x\in\mathbb{R}$. Avec la formule d'addition d'angle[3] dans le cosinus nous cherchons un T tel que

$$\cos(x+T)=\cos(x)\cos(T)-\sin(x)\sin(T)=\cos(x) \tag{16.47}$$

et donc tel que

$$\cos(x)\big(\cos(T)-1\big)=\sin(x)\sin(T). \tag{16.48}$$

Nous dérivons cette équation :

$$-\sin(x)\big(\cos(T)-1\big)=\cos(x)\sin(T). \tag{16.49}$$

Nous multiplions chacune des deux équations (16.48) et (16.49) par $\sin(x)$ et $\cos(x)$ pour obtenir les quatre relations suivantes :

$$\cos^2(x)\big(\cos(T)-1\big)-\sin(x)\cos(x)\sin(T)=0 \tag{16.50a}$$
$$-\sin(x)\cos(x)\big(\cos(T)-1\big)-\cos^2(x)\sin(T)=0 \tag{16.50b}$$
$$\sin(x)\cos(x)\big(\cos(T)-1\big)-\sin^2(x)\sin(T)=0 \tag{16.50c}$$
$$\sin^2(x)\big(\cos(T)-1\big)-\sin(x)\cos(x)\sin(T)=0 \tag{16.50d}$$

3. Rien ne nous empêche de donner ce nom à ces formules, mais seriez vous capable de définir précisément le mot «angle» ?

En faisant (16.50a) moins (16.50d) nous trouvons $\cos(T) = 1$. Et en sommant (16.50b) avec (16.50c) nous avons $-\sin(T) = 0$.

Si $T > 0$ est tel que $\sin(T) = 0$ et $\cos(T) = 1$ Alors la formule d'addition d'angle donne tout de suite

$$\cos(x + T) = \cos(x). \tag{16.51}$$

À de niveau nous croyons avoir prouvé que cos était périodique et que la période est donnée par

$$\min\{T > 0 \text{ tel que } \sin(T) = 0, \cos(T) = 1\}. \tag{16.52}$$

Or rien n'est moins sûr parce qu'il pourrait arriver que ce minimum n'existe pas, c'est à dire que l'infimum soit zéro. Autrement dit, il peut arriver que l'ensemble des périodes soit dense. Plus précisément, soit $P \subset \mathbb{R}$ l'ensemble des périodes de cos. C'est un sous-groupe de $(\mathbb{R}, +)$ et le lemme 16.25 nous dit que P est soit dense dans $\mathbb{R}$ soit de la forme $p\mathbb{Z}$ pour un $p > 0$.

Si P est dense, soit $t \in \mathbb{R}$ et une suite (t_n) dans P telle que $t_n \to t$. Pour tout x et tout n nous avons

$$\cos(x + t_n) = \cos(x), \tag{16.53}$$

Vu que la fonction cosinus est continue, nous pouvons passer à la limite et écrire $\cos(x+t) = \cos(x)$. Cela étant valable pour tout x et pour tout t, la fonction cosinus est constante. Or nous savons que ce n'est pas le cas, donc P n'est pas dense. Donc cosinus est périodique. □

Définition 16.27 (Le nombre π)**.**
Le nombre $\pi > 0$ est donné par

$$2\pi = \min\{T > 0 \text{ tel que } \cos(x + T) = \cos(x)\, \forall x\}. \tag{16.54}$$

Par ce qui a été dit dans la démonstration nous avons aussi

$$2\pi = \min\{T > 0 \text{ tel que } \sin(T) = 0, \cos(T) = 1\}. \tag{16.55}$$

Notons que tout ceci ne nous donne pas la plus petite indication d'ordre de grandeur de la valeur de π. Cela peut encore être 0.1 autant que 500.

Proposition 16.28 ([197])**.**
Des propriétés à la chaîne à propos des sinus, cosinus et de leurs périodes.

(1) Le nombre 2π est le plus petit tel que

$$\begin{cases} \cos(2\pi) = 1 & \text{(16.56a)} \\ \sin(2\pi) = 0. & \text{(16.56b)} \end{cases}$$

(2) Le nombre 2π est également la période de la fonction sin*.*

(3) Nous avons $\cos(\pi) = -1$ et $\sin(\pi) = 0$.

(4) Pour tout $a \in \mathbb{R}$ nous avons

$$\cos(a + \pi) = -\cos(\pi) \tag{16.57a}$$
$$\sin(a + \pi) = -\sin(\pi). \tag{16.57b}$$

(5) Le nombre π est le plus petit $T > 0$ tel que $\cos(T) = -1$ et $\sin(T) = 0$.

(6) Nous avons

$$\begin{cases} \cos(\pi/2) = 0 & \text{(16.58a)} \\ \sin(\pi/2) = 1. & \text{(16.58b)} \end{cases}$$

(7) Nous avons les formules

$$\begin{cases} \cos(x + \pi/2) = -\sin(x) & \text{(16.59a)} \\ \sin(x + \pi/2) = \cos(x) & \text{(16.59b)} \end{cases}$$

pour tout $x \in \mathbb{R}$.

(8) *Le nombre $\pi/2$ est le plus petit T vérifiant* $\sin(T) = 1$, $\cos(T) = 0$.

(9) *Nous avons les valeurs*

$$\begin{cases} \cos(\frac{3\pi}{2}) = 0 & (16.60a) \\ \sin(\frac{3\pi}{2}) = -1. & (16.60b) \end{cases}$$

(10) *Le nombre $\pi/2$ est le plus petit $T > 0$ tel que* $\cos(T) = 0$.

Démonstration. C'est parti.

(1) Le fond de la proposition 16.26 est que toutes les périodes $T > 0$ vérifient $\cos(T) = 1$ et $\sin(T) = 0$. La définition de π est que c'est la plus petite période.

(2) En utilisant le fait que l'une est la dérivée de l'autre, si T est une période de cos nous avons

$$\begin{aligned} \sin(x + T) &= -\cos'(x + T) & (16.61a) \\ &= -\lim_{\epsilon \to 0} \frac{\cos(x + T + \epsilon) - \cos(x + T)}{\epsilon} & (16.61b) \\ &= -\lim_{\epsilon \to 0} \frac{\cos(x + \epsilon) - \cos(x)}{\epsilon} & (16.61c) \\ &= -\cos'(x) & (16.61d) \\ &= \sin(x). & (16.61e) \end{aligned}$$

Nous déduisons que toute période de cos est une période de sin. De la même façon, nous pouvons prouver le contraire : toute période de sin est une période de cos.

(3) D'un côté nous avons

$$\cos(2\pi) = \cos^2(\pi) - \sin^2(\pi) = 1 \tag{16.62}$$

parce que $\cos(2\pi) = \cos(0) = 1$. Vu que $\cos(\pi)$ et $\sin(\pi)$ sont bornés par -1 et 1, nous devons avoir $\sin(\pi) = 0$ et $\cos(\pi) = \pm 1$.

Mais d'un autre côté, le nombre 2π est le plus petit T vérifiant $\cos(T) = 1$, $\sin(T) = 0$. Donc avoir $\cos(\pi) = 1$ n'est pas possible. Nous concluons

$$\begin{cases} \cos(\pi) = -1 & (16.63a) \\ \sin(\pi) = 0. & (16.63b) \end{cases}$$

(4) Il s'agit d'utiliser les formules d'addition d'angles du lemme 16.23 pour calculer $\cos(a + \pi)$ et $\sin(a + \pi)$ en tenant compte du fait que $\cos(\pi) = -1$ et $\sin(\pi) = 0$.

(5) Soit $a \in]0, \pi[$ tel que $\cos(a) = -1$ et $\sin(a) = 0$. Alors nous avons

$$\begin{aligned} \cos(a + \pi) &= -\cos(\pi) = 1 & (16.64a) \\ \sin(a + \pi) &= -\sin(\pi) = 0, & (16.64b) \end{aligned}$$

ce qui donnerait $a + \pi \in]\pi, 2\pi[$ dont le cosinus est 1 et le sinus est zéro. Mais nous savons déjà que 2π est le minimum pour cette propriété.

(6) Nous avons

$$-1 = \cos(\pi) = \cos^2(\pi/2) - \sin^2(\pi/2), \tag{16.65}$$

donc $\cos(\pi/2) = 0$ et $\sin^2(\pi/2) = 1$, ce qui donne $\sin(\pi/2) = \pm 1$.

Nous devons départager le $\pm$. Pour cela nous savons que $\sin'(0) = \cos(0) = 1$, donc il existe $\epsilon > \epsilon$ tel que pour tout $x \in]0, \epsilon[$ nous avons $0 < \cos(x) < 1$ et $0 \sin(x) < 1$. Nous choisissons ϵ plus petit que $\pi/2$.

Supposons que $\sin(\pi/2) = -1$. Le théorème des valeurs intermédiaires 12.21 dit qu'il existe $x_0 \in]\epsilon, \pi/2[$ tel que $\sin(x_0) = 0$. Pour cette valeur de x_0 nous devons aussi avoir $\cos(x_0) =$

± 1. Mais vu que 2π est minium pour avoir cos = 1 et sin = 0 nous devons avoir $\cos(x_0) = -1$. Alors nous avons aussi

$$\cos(x_0 + \pi) = \cos(x_0)\cos(\pi) - \sin(x_0)\sin(\pi) = -\cos(x_0) = 1 \tag{16.66a}$$
$$\sin(x_0 + \pi) = \cos(x_0)\sin(\pi) + \sin(x_0)\cos(\pi) = \sin(x_0) = 0. \tag{16.66b}$$

Encore une fois par minimalité de 2π, cela ne va pas. Conclusion : $\sin(\pi/2) = 1$.

(7) Il s'agit encore d'utiliser les formules d'addition d'angle en tenant compte du fait que $\cos(\pi/2) = 0$ et $\sin(\pi/2) = 1$.

(8) Supposons $x_0 \in]0, \pi/2[$ tel que $\sin(x_0) = 1$ et $\cos(x_0) = 0$. En utilisant les formules (16.59) nous avons

$$\cos(x_0 + \pi/2) = -1 \tag{16.67a}$$
$$\sin(x_0 + \pi/2) = 0, \tag{16.67b}$$

avec $x_0 + \pi/2 < \pi$. Cela contredirait la minimalité de π.

(9) Il s'agit d'utiliser les formules (16.59) :

$$\cos(\frac{3\pi}{2}) = \cos(\pi + \pi/2) = -\sin(\pi) = 0 \tag{16.68a}$$
$$\sin(\frac{3\pi}{2}) = \sin(\pi + \pi/2) = \cos(\pi) = -1. \tag{16.68b}$$

(10) Si $\cos(x_0) = 0$ alors $\sin(x_0) = -1$ (parce que $\sin(x_0) = 1$ est déjà exclu). Alors $\cos(x_0 + \pi/2) = 1$ et $\sin(x_0 + \pi/2) = 0$, ce qui est également impossible.

□

Tout cela nous permet d'écrire le tableau de variations de sinus et cosinus.

Proposition 16.29.
L'application

$$\begin{aligned} \gamma\colon [0, 2\pi[&\to S^1 \subset \mathbb{R}^2 \\ t &\mapsto \big(\cos(t), \sin(t)\big) \end{aligned} \tag{16.69}$$

est une bijection continue.

Démonstration. La continuité découle de la continuité des composantes. Le fait que l'image de γ soit dans S^1 découle immédiatement du fait que $\sin^2 + \cos^2 = 1$.

Pour la bijection, il faut injectif et surjectif.

Injectif Soient $x_1 < x_2$ tels que $\sin(x_1) = \sin(x_2)$ et $\cos(x_1) = \cos(x_2)$. Supposons pour fixer les idées que $\sin(x_1) > 0$ et $\cos(x_1) > 0$: si ce n'est pas le cas, il faut traiter séparément les 4 possibilités de combinaisons de signes.

Nous avons obligatoirement $x_1, x_2 \in [0, \frac{\pi}{2}[$. Vu que $\sin(x_1) = \sin(x_2)$, il existe par le théorème de Rolle 12.127 un élément $c \in]x_1, x_2[$ tel que $\sin'(c) = 0$, c'est à dire $\cos(c) = 0$. Cela contredirait la proposition 16.28(10) à moins que $x_1 = x_2$.

Surjectif Soient x, y tels que $x^2 + y^2 = 1$. Supposons pour varier les plaisirs que $x < 0$ et $y > 0$. Vu que la fonction cos va de 0 à -1 lorsque x va de $\pi/2$ à π, le théorème des valeurs intermédiaires donne $t \in [\pi/2, \pi]$ tel que $\cos(t) = x$. Pour cette valeur de x nous avons

$$\cos^2(x) + \sin^2(x) = 1, \tag{16.70}$$

et donc $\sin^2(x) = y^2$, ce qui donne $\sin(x) = \pm y$. Mais pour $x \in [\pi/2, \pi]$ nous avons $\sin(t) > 0$. Par conséquent $\sin(t) = y$.

□

Lemme 16.30.
Nous avons les valeurs remarquables

$$\sin(\frac{\pi}{4}) = \cos(\frac{\pi}{4}) = \frac{\sqrt{2}}{2}. \tag{16.71}$$

Je crois que vous devriez pouvoir faire la preuve tout seul. Si ça ne va pas, contactez-moi.

16.2.2 Exemples

Nous mettons ici quelque exemples concernant les fonctions trigonométriques, qui n'ont pas pu être mis dans les chapitres le plus adapté, parce que ces derniers sont plus haut dans la table des matière.

Exemple 16.31(Taylor)
Le développement du cosinus est donné par

$$\cos(x) = 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} \cdots \tag{16.72}$$

Nous avons donc l'existence d'une fonction $h_1 \in o(x^2)$ telle que $\cos(x) = 1 - \frac{x^2}{2} + h_1(x)$. Il existe aussi une autre fonction $h_2 \in o(x^4)$ telle que $\cos(x) = 1 - \frac{x^2}{2} + \frac{x^4}{4!} + h_2(x)$. △

Exemple 16.32(Limite et prolongement par continuité)
La fonction

$$f(x) = \frac{\cos(x) - 1}{x} \tag{16.73}$$

n'est pas définie en $x = 0$, mais en la limite

$$\lim_{x \to 0} \frac{\cos(x) - 1}{x} \tag{16.74}$$

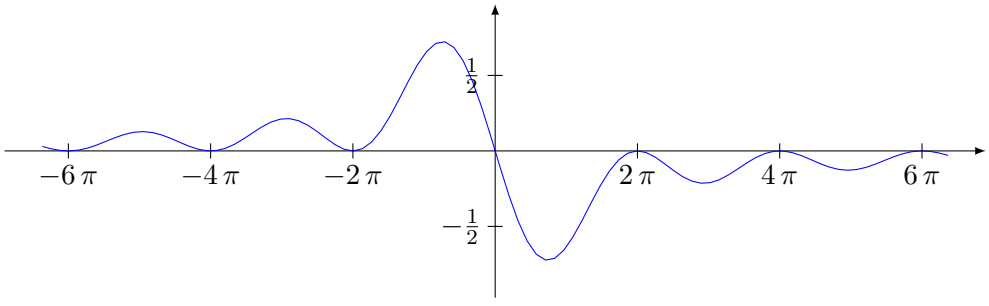
nous reconnaissons la limite définissant la dérivée du cosinus en 0, c'est à dire que

$$\lim_{x \to 0} \frac{\cos(x) - 1}{x} = \sin(0) = 0. \tag{16.75}$$

Nous avons donc le prolongement par continuité

$$\tilde{f}(x) = \begin{cases} \frac{\cos(x)-1}{x} & \text{si } x \neq 0 \\ 0 & \text{sinon.} \end{cases} \tag{16.76}$$

Encore une fois, le graphe de la fonction $\tilde{f}$ ne présente aucune particularité autour de $x = 0$.



△

Exemple 16.33(Un calcul heuristique de limite)
Soit à calculer la limite suivante :

$$\lim_{x\to 0} \frac{e^{-2\cos(x)+2}\sin(x)}{\sqrt{e^{2\cos(x)+2}-1}}. \tag{16.77}$$

La stratégie que nous allons suivre pour calculer cette limite est de développer certaines parties de l'expression en série de Taylor, afin de simplifier l'expression. La première chose à faire est de remplacer $e^{y(x)}$ par $1+y(x)$ lorsque $y(x)\to 0$. La limite devient

$$\lim_{x\to 0} \frac{\big(-2\cos(x)+3\big)\sin(x)}{\sqrt{-2\cos(x)+2}}. \tag{16.78}$$

Nous allons maintenant remplacer $\cos(x)$ par 1 au numérateur et par $1-x^2/2$ au dénominateur. Pourquoi ? Parce que le cosinus du dénominateur est dans une racine, donc nous nous attendons à ce que le terme de degré deux du cosinus donne un degré un en dehors de la racine, alors que du degré un est exactement ce que nous avons au numérateur : le développement du sinus commence par x.

Nous calculons donc

$$\lim_{x\to 0} \frac{\sin(x)}{\sqrt{-2\left(1-\frac{x^2}{2}\right)+2}} = \lim_{x\to 0}\frac{\sin(x)}{x} = 1. \tag{16.79}$$

Tout ceci n'est évidement pas très rigoureux, mais en principe vous avez tous les éléments en main pour justifier les étapes. △

Proposition 16.34 ([198])**.**
Soient des fonction $f,g\colon I\to\mathbb{R}$ de classe C^1 sur l'ouvert I de $\mathbb{R}$ telles que $f^2+g^2=1$. Soit $t_0\in I$ et θ_0 tel que $f(t_0)=\cos(\theta_0)$ et $g(t_0)=\sin(\theta_0)$.

Alors il existe une unique fonction continue $\theta\colon I\to\mathbb{R}$ telle que

$$\begin{cases} \theta(t_0)=\theta_0 & \text{(16.80a)}\\ f=\cos\circ\theta & \text{(16.80b)}\\ g=\sin\circ\theta. & \text{(16.80c)}\end{cases}$$

Démonstration. Nous commençons par l'existence, en passant par les nombres complexes. Soit $h\colon I\to\mathbb{C}$ définie par $h=f+ig$. Nous avons $h\bar{h}=1$ et nous définissons

$$\theta(t)=\theta_0-i\int_{t_0}^{t} h'(s)\overline{h(s)}ds. \tag{16.81}$$

Cette intégrale existe pour tout t parce que les fonctions f et g étant de classe C^∞, elles sont bornées sur le compact $[t_0,t]$. De plus θ est une fonction continue parce que c'est une primitive (proposition 13.204) [4].

La dérivée de θ est la fonction $s\mapsto -ih'(s)\overline{h(s)}$.

Utilisant la formule du lemme 16.22 sur la forme trigonométrique des nombres complexes, nous calculons :

$$\frac{d}{dt}\Big[he^{-i\theta}\Big]_{t=0} = e^{-i\theta}(h'-h\theta') = e^{-i\theta}(h'-ih(-i)h'\bar{h}) = 0. \tag{16.82}$$

Par conséquent il existe $c\in\mathbb{C}$ tel que $he^{-i\theta}=c$. Mais $h(t_0)=f(t_0)+ig(t_0)=\cos(\theta_0)+i\sin(\theta_0)=e^{i\theta_0}$, du coup

$$h(t_0)e^{-i\theta(t_0)}=c \tag{16.83}$$

4. En réalité nous appliquons la proposition 9.29 à chacune des parties réelles et imaginaires de la fonction $s\mapsto h'(s)\overline{h(s)}$.

donne immédiatement $c = 1$, ou encore $e^{i\theta(t)} = h(t)$, c'est à dire que

$$f + ig = \cos \circ\theta + i \sin \circ\theta, \tag{16.84}$$

ce qu'il fallait pour l'existence.

Pour l'unicité nous supposons avoir une autre fonction, α qui satisfait aux exigences. Pour tout $t \in I$ nous avons

$$e^{i\theta(t)} = e^{i\alpha(t)}. \tag{16.85}$$

Il existe donc une fonction $n\colon I \to \mathbb{N}$ telle que $\theta(t) = \alpha(t) + 2n(t)\pi$. Par continuité de θ et α, la fonction n doit être constante, mais vu que $\theta(t_0) = \alpha(t_0)$ nous avons $n = 1$. □

16.2.3 Très modeste approximation de π

Nous sommes en droit de vouloir une valeur approchée de π.

Lemme 16.35.
Nous avons l'approximation numérique

$$2\sqrt{2} < \pi < 4. \tag{16.86}$$

Démonstration. Grace au lemme 16.30 nous savons que la fonction sin passe de 0 à $\sqrt{2}/2$ sur un intervalle de taille $\pi/4$ avec une dérivé majorée par 1. Par conséquent

$$\frac{\pi}{4} > \frac{\sqrt{2}}{2} \tag{16.87}$$

et donc [5]

$$\pi > 2\sqrt{2} \simeq 2.82 \tag{16.88}$$

De plus la fonction sin passe de 0 à $\sqrt{2}/2$ sur un intervalle de taille $\pi/4$ avec une dérivée majorée par $\sqrt{2}/2$, donc

$$\frac{\pi}{4} < \frac{\sqrt{2}/2}{\sqrt{2}/2}, \tag{16.89}$$

ce qui donne

$$\pi < 4. \tag{16.90}$$

□

Pour avoir une meilleur approximation de π, nous pouvons remarquer que $\pi \in]2.82, 4[$, et que cet intervalle est suffisamment petit pour ne pas recouvrir l'intervalle correspondant pour 2π. L'équation $\cos(x) = -1$ possède donc une unique solution dans cet intervalle (et cette solution est π). Nous pouvons donc faire une dichotomie pour trouver la valeur de π, pourvu que nous ayons une façon d'évaluer des valeurs de $\cos(x)$ de façon pas trop ridicule.

16.2.4 Cercle trigonométrique

Le cercle trigonométrique est le cercle de rayon 1 représenté à la figure 16.1. Sa longueur est 2π.

Nous verrons plus tard que la longueur de l'arc de cercle intercepté par un angle θ est égal à θ. Les radians sont donc l'unité d'angle les plus adaptés au calcul de longueurs sur le cercle.

5. Sérieusement, êtes vous capables de trouver une approximation de $\sqrt{2}$ en ne vous basant que sur des choses vues jusqu'ici ?

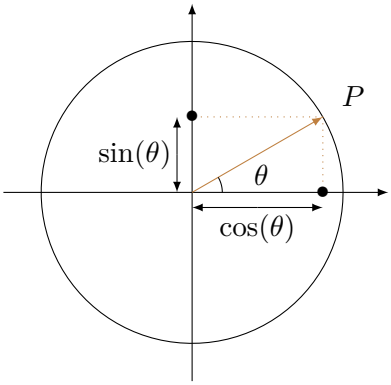


FIGURE 16.1 – Le cercle trigonométrique.

16.2.5 Les fonctions tangente et arc tangente

Définition 16.36.
La fonction ***tangente*** *est :*

$$\tan(x) = \frac{\sin(x)}{\cos(x)} \tag{16.91}$$

où sin *et* cos *sont de la définition 16.19.*

La fonction tangente n'est pas définie sur les points de la forme $x = \frac{\pi}{2} + k\pi$, $k \in \mathbb{Z}$. Une interprétation géométrique, qui justifie le nom, est donnée sur la figure 16.2.

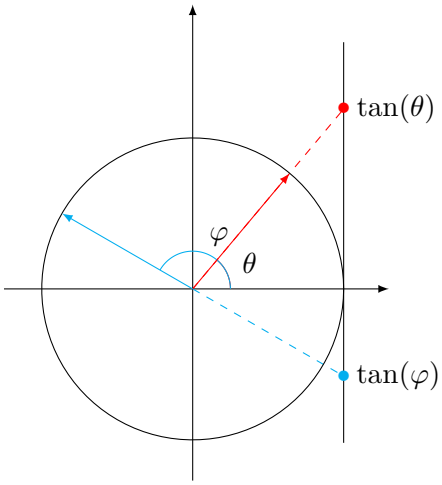


FIGURE 16.2 – Interprétation géométrique de la fonction tangente. La tangente de l'angle θ est positive (et un peu plus grande que 1) tandis que celle de la tangente de l'angle φ est négative.

Proposition 16.37.
La fonction

$$\begin{aligned} \tan\colon]-\frac{\pi}{2}, \frac{\pi}{2}[&\to \mathbb{R} \\ x &\mapsto \tan(x) \end{aligned} \tag{16.92}$$

est une bijection.

Démonstration. Le cosinus ne s'annulant pas sur l'intervalle donné, la fonction est bien définie. Nous avons

$$\lim_{x \to \pi/2^-} \tan(x) = +\infty \tag{16.93}$$

parce que la limite du sinus est 1 est celle du cosinus est zéro par les valeurs positives. Le même raisonnement donne la limite en $-\pi/2$ qui vaut $-\infty$. Le théorème des valeurs intermédiaires[6] dit que la fonction tangente est alors surjective sur $\mathbb{R}$.

Par ailleurs en utilisant les règles de calcul comme la dérivation du quotient 12.115(3) nous trouvons

$$\tan'(x) = \tan^2(x) + 1, \tag{16.94}$$

ce qui nous donne une dérivée partout strictement positive, et donc une fonction strictement croissante et donc injective. □

Le graphe de la fonction tangente est sur la figure 16.3.

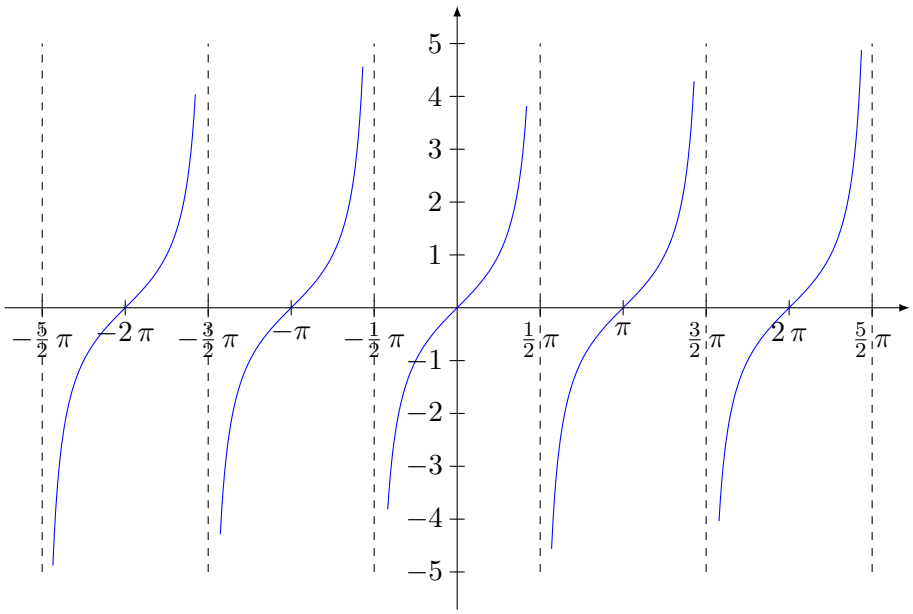


FIGURE 16.3 – Le graphe de la fonction tangente.

En ce qui concerne la bijection réciproque nous avons le théorème suivant.

Théorème 16.38.
La fonction

$$\begin{aligned} \arctan\colon \mathbb{R} &\to \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\\ x &\mapsto \arctan(x) \end{aligned} \tag{16.95}$$

nommée ***arc tangente*** *est*

(1) impaire et strictement croissante sur $\mathbb{R}$*.*

(2) dérivable sur $\mathbb{R}$ *de dérivée*

$$\arctan'(x) = \frac{1}{1+x^2}. \tag{16.96}$$

Démonstration. La fonction sinus étant impaire (visible sur son développement de définition (16.31)), et la fonction cosinus étant paire, la fonction tangente est impaire et sa réciproque l'est tout autant.

La fonction arc tangente est également dérivable (donc continue) par la proposition 12.42 parce que la fonction tangente l'est. Notons qu'ici nous nous sommes restreint à $]-\pi/2, \pi/2[$. Sinon, le résultat est faux.

La formule proposée pour la dérivée provient également de la proposition 12.42 et de la dérivée de la tangente : □

6. Théorème 12.21.

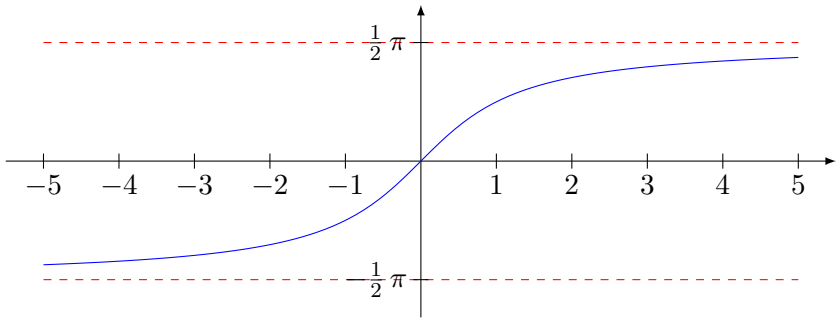
Lemme 16.39.
Nous avons la valeur remarquable

$$\arctan(1/\sqrt{3}) = \frac{\pi}{6}. \tag{16.97}$$

Le nombre $\arctan(x_0)$ se calcule en cherchant l'angle $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ dont la tangente vaut x_0. Nous obtenons le tableau de valeurs suivant :

x	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$
$\arctan(x)$	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$

En ce qui concerne la représentation graphique de la fonction $x \mapsto \arctan(x)$, elle s'obtient «en retournant» la partie entre $-\frac{\pi}{2}$ et $\frac{\pi}{2}$ du graphique de la fonction tangente :



16.2.6 La fonction arc sinus

Nous voulons étudier la fonction

$$\begin{aligned} \sin\colon \mathbb{R} &\to [-1,1] \\ x &\mapsto \sin(x) \end{aligned} \tag{16.98}$$

et sa réciproque éventuelle.

La fonction sinus est continue sur $\mathbb{R}$ mais n'est pas bijective : elle prend une infinité de fois chaque valeur de $J = [-1, 1]$. Pour définir une bijection réciproque de la fonction sinus en utilisant le théorème 12.41, nous devons donc choisir un intervalle à partir duquel la fonction sinus est monotone. Nous choisissons l'intervalle

$$I = [-\frac{\pi}{2}, \frac{\pi}{2}]. \tag{16.99}$$

La fonction

$$\begin{aligned} \sin\colon [-\frac{\pi}{2}, \frac{\pi}{2}] &\to [-1,1] \\ x &\mapsto \sin(x) \end{aligned} \tag{16.100}$$

est une bijection croissante et continue. Nous avons donc le résultat suivant.

Théorème 16.40 (Définition et propriétés de arc sinus)**.**
Nous nommons ***arc sinus*** *la bijection inverse de la fonction* $\sin\colon I \to J$*. La fonction*

$$\begin{aligned} \arcsin\colon [-1,1] &\to [-\frac{\pi}{2}, \frac{\pi}{2}] \\ x &\mapsto \arcsin(x) \end{aligned} \tag{16.101}$$

ainsi définie est

(1) continue et strictement croissante ;

(2) impaire : pour tout $x \in [-1, 1]$ *nous avons* $\arcsin(-x) = -\arcsin(x)$.

Démonstration. Nous prouvons le fait que arcsin est impaire. Un élément de l'ensemble de définition de arcsin est de la forme $y = \sin(x)$ avec $x \in [-\pi/2, \pi/2]$. La relation (12.45) s'écrit dans notre cas

$$x = \arcsin\big(\sin(x)\big). \tag{16.102}$$

Nous écrivons d'une part cette équation avec $-x$ au lieu de x :

$$-x = \arcsin\big(\sin(-x)\big) = \arcsin\big(-\sin(x)\big) = \arcsin(-y); \tag{16.103}$$

et d'autre part nous multiplions (16.102) par -1 :

$$-x = -\arcsin\big(\sin(x)\big) = -\arcsin(y). \tag{16.104}$$

En égalisant les valeurs (16.103) et (16.104) nous trouvons

$$\arcsin(-y) = -\arcsin(y), \tag{16.105}$$

ce qui signifie que arcsin est une fonction impaire. □

Notons que cette preuve repose sur le fait que tout élément de l'ensemble de définition de la fonction arc sinus peut être écrit sous la forme $\sin(x)$ pour un certain x.

Si $x_0 \in [-1, 1]$ est donné, calculer $\arcsin(x_0)$ revient à trouver un angle θ_0 dans $[-\frac{\pi}{2}, \frac{\pi}{2}]$ pour lequel $\sin(\theta_0) = x_0$. Un tel angle sera forcément unique.

Remarque 16.41.
La définition de arc sinus découle du choix de l'intervalle I, qui est une convention. Il aurait été possible de faire un choix différent : pourriez vous trouver la réciproque de la fonction sinus sur l'intervalle $[\pi/2, 3\pi/2]$? Le mieux est de l'écrire comme une translatée de arc sinus, en utilisant le fait que sinus est une fonction périodique.

Exemple 16.42
Pour calculer $\arcsin(1)$, il faut chercher un angle entre $-\frac{\pi}{2}$ et $\frac{\pi}{2}$ ayant 1 pour sinus : résoudre $\sin(\theta) = 1$. La solution est $\theta = \frac{\pi}{2}$ et nous avons donc $\arcsin(1) = \frac{\pi}{2}$. △

À l'aide des valeurs remarquables de la fonction sinus nous obtenons le tableau suivant de valeurs remarquables pour l'arc sinus.

x	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1
$\arcsin(x)$	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$

Les autres valeurs remarquables peuvent être déduites du fait que l'arc sinus est une fonction impaire.

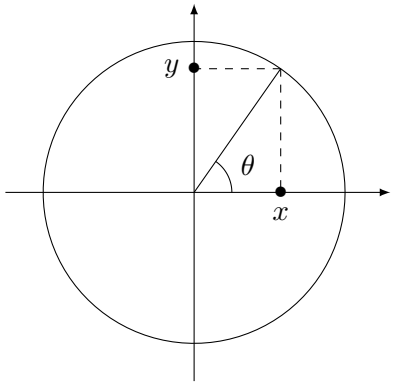
En ce qui concerne la dérivabilité de la fonction arc sinus, en application de la proposition 12.42 elle est dérivable en tout $y = \sin(x)$ tel que $\sin'(x) \neq 0$, c'est à dire tel que $\cos(x) \neq 0$. Or $\cos(x) = 0$ pour $x = \pm\frac{\pi}{2}$, ce qui correspond à $y = \sin(\pm\frac{\pi}{2}) = \pm 1$. La fonction arc sinus est donc dérivable sur $]-1, 1[$. Nous avons donc la propriété suivante pour la dérivabilité.

Proposition 16.43.
La fonction arc sinus est continue sur $[-1, 1]$ *et dérivable sur* $]-1, 1[$. *Pour tout* $y \in]-1, 1[$, *la dérivée est donnée par la formule* (12.54), *qui dans ce cas s'écrit*

$$\arcsin'(y) = \frac{1}{\cos\big(\arcsin(y)\big)} = \frac{1}{\sqrt{1-y^2}}. \tag{16.106}$$

La dernière égalité viens du fait que si $x = \arcsin(y)$ alors $y = \sin(x)$ et $\cos(x) = \sqrt{1-\sin^2(x)} = \sqrt{1-y^2}$.

Pour comprendre la dernière égalité, remarquer que dans le dessin suivant, $\theta = \arcsin(y)$, donc $y = \sin(\theta)$, et $x = \cos(\theta)$.



Notons enfin que le graphe de la fonction arc sinus est donné à la figure 16.4.

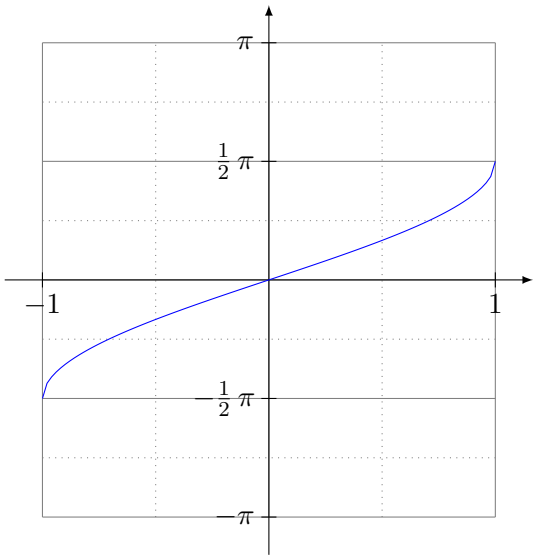


FIGURE 16.4 – Le graphe de la fonction $x \mapsto \arcsin(x)$

16.2.7 La fonction arc cosinus

Nous voulons étudier la fonction

$$\cos\colon \mathbb{R} \to [-1,1] \tag{16.107}$$

et son éventuelle réciproque. Encore une fois il n'est pas possible d'en prendre la réciproque globale parce que ce n'est pas une bijection ; ne fut-ce que parce qu'elle est périodique (proposition 16.26). Nous choisissons de considérer l'intervalle $[0,\pi]$ sur lequel la fonction cosinus est continue et strictement monotone décroissante.

Nous avons alors le résultat suivant :

Proposition-définition 16.44.
Pour définir la fonction arcsinus.

(1) La fonction

$$\cos\colon [0,\pi] \to [-1,1] \tag{16.108}$$

est une bijection continue strictement décroissante.

(2) Sa bijection réciproque est la fonction

$$\arccos\colon [-1,1] \to [0,\pi] \tag{16.109}$$

nommée ***arc cosinus****.*

(3) La fonction arc cosinus est continue, strictement décroissante.

(4) Elle est dérivable et pour tout $y \in]-1,1[$, sa dérivée est donnée par

$$\arccos'(y) = \frac{1}{-\sin\big(\arccos(y)\big)} = \frac{-1}{\sqrt{1-y^2}}. \tag{16.110}$$

Démonstration. La fonction cosinus est continue et même de classe C^∞ par la proposition 16.19. Elle est strictement décroissant parce que sa dérivée $(-\sin)$ y est strictement positive (strictement à dans l'intérieur du domaine).

Le fait que arc cosinus soit une bijection continue strictement monotone est dans le théorème de la bijection 12.41. La dérivabilité et la formule sont de la proposition 12.42. □

Pour $y_0 \in [-1,1]$, trouver la valeur de $\arccos(y_0)$ revient à résoudre l'équation $\cos(x_0) = y_0$. Cela nous permet de construire une tableau de valeurs :

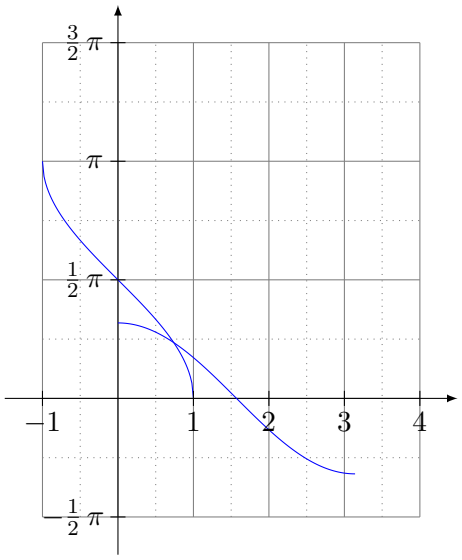
x	-1	$-\frac{\sqrt{3}}{2}$	$-\frac{\sqrt{2}}{2}$	$-\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1
$\arccos(x)$	π	$\frac{5\pi}{6}$	$\frac{3}{4}\pi$	$\frac{2}{3}\pi$	$\frac{1}{2}\pi$	$\frac{\pi}{3}$	$\frac{1}{4}\pi$	$\frac{1}{6}\pi$	0

Remarque 16.45.
Certes la fonction cosinus est paire (vue sur $\mathbb{R}$), mais la fonction arc cosinus ne l'est pas car elle est une bijection entre $[-1,1]$ et $[0,\pi]$.

Exemple 16.46
Cherchons $\arccos(\frac{1}{2})$. Il faut trouver un angle $\theta \in [0,\pi]$ tel que $\cos(\theta) = \frac{1}{2}$. La solution est $\theta = \frac{\pi}{3}$. Donc $\arccos(\frac{1}{2}) = \frac{\pi}{3}$.

Il n'est cependant pas immédiat d'en déduire la valeur de $\arccos(-\frac{1}{2})$. En effet $\theta = \arccos(-\frac{1}{2})$ si et seulement si $\cos(\theta) = -\frac{1}{2}$ avec $\theta \in [0,\pi]$. La solution est $\theta = \frac{2\pi}{3}$. △

En ce qui concerne la représentation graphique, il suffit de tracer la fonction cosinus entre 0 et π puis de prendre le symétrique par rapport à la droite $y = x$.



16.2.8 Une meilleure approximation de π

Nous avions laissé le nombre π avec l'approximation assez minable de $2\sqrt{2} < \pi < 4$ en le lemme 16.35. Nous pouvons maintenant faire nettement mieux.

Le lemme 16.39 donne

$$\arctan(1/\sqrt{3}) = \pi/6 \tag{16.111}$$

et l'idée est de donner un développement de arctan autour de zéro, de l'évaluer en $1/\sqrt{3}$ et d'égaliser le résultat à $\pi/6$. Tout cela donne lieu à des calcules peut-être fastidieux, mais comme un gars l'a fait dès l'an 1424[199] pour trouver 16 décimales corrects, nous faisons comme si c'était facile.

Le développement en série de Taylor 12.677 d'arc tangente autour de $x = 0$ est donné par

$$\arctan(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{2k+1}, \tag{16.112}$$

valable pour $x \in]-1, 1[$. Avec cela nous avons

$$\arctan(\frac{1}{\sqrt{3}}) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)3^k} \times \frac{1}{\sqrt{3}} = \frac{\pi}{6}, \tag{16.113}$$

et donc

$$\pi = \frac{6}{\sqrt{3}} \sum_{k=0}^{\infty} \frac{(-1)k}{(2k+1)3^k}. \tag{16.114}$$

Pour donner une idée du fait que ça fonctionne pas mal, voici le calcul pour quelque termes :

```
1
2  SageMath version 7.3, Release Date: 2016-08-04
3  Type "notebook()" for the browser-based notebook interface.
4  Type "help()" for help.
5
6 sage: n(taylor(arctan(x),x,0,5)(1/sqrt(3)))*6
7 3.15618147156995
8 sage: n(taylor(arctan(x),x,0,10)(1/sqrt(3)))*6
9 3.14260474566308
10 sage: n(taylor(arctan(x),x,0,20)(1/sqrt(3))*6-pi)
11 -2.14265171338823e-6
12 sage: n(taylor(arctan(x),x,0,58)(1/sqrt(3))*6-pi)
13 8.88178419700125e-16
```

tex/sage/sageSnip012.sage

Calculer 5 termes donne déjà 3.15. Et on est à 10^{-6} de la bonne réponse avec 20 termes. Et avec 58 termes, on n'est à 10^{-16}.

Problèmes et choses à faire

Pour bien faire, il faudrait étudier le reste et donner un encadrement.

16.2.9 Forme polaire ou trigonométrique des nombres complexes

Un nombre complexe étant représenté par deux nombres, on peut le représenter dans un plan appelé « plan de Gauss ». La plupart des opérations sur les nombres complexes ont leur interprétation géométrique dans ce plan.

Dans le plan de Gauss, le module d'un complexe z représente la distance entre 0 et z. On appelle *argument* de z (noté $\arg z$) l'angle (déterminé à 2π près) entre le demi-axe des réels positifs et la demi-droite qui part de 0 et passe par z. Le module et l'argument d'un complexe permettent de déterminer univoquement ce complexe puisqu'on a la formule

$$z = a + bi = |z| \left(\cos(\arg(z)) + i \sin(\arg(z))\right)$$

L'argument de z se détermine via les formules

$$\frac{a}{|z|} = \cos(\arg(z)) \quad \frac{b}{|z|} = \sin(\arg(z))$$

ou encore par la formule

$$\frac{b}{a} = \tan(\arg(z)) \quad \text{en vérifiant le quadrant.}$$

La vérification du quadrant vient de ce que la tangente ne détermine l'angle qu'à π près.

16.2.10 Angle entre deux vecteurs

Si a et b sont des réels, l'inégalité $|a| \leqslant b$ peut se développer en une double inégalité

$$-b \leqslant a \leqslant b. \tag{16.115}$$

L'inégalité de Cauchy-Schwarz (9.94) devient alors

$$-\|X\|\|Y\| \leqslant X \cdot Y \leqslant \|X\|\|Y\|. \tag{16.116}$$

Si $X \neq 0$ et $Y \neq 0$, nous en déduisons

$$-1 \leqslant \frac{X \cdot Y}{\|X\|\|Y\|} \leqslant 1. \tag{16.117}$$

Il existe donc par la proposition 16.44 un angle $\theta \in [0, \pi]$ tel que

$$\cos(\theta) = \frac{X \cdot Y}{\|X\|\|Y\|}. \tag{16.118}$$

Définition 16.47.
*L'angle ainsi défini est l'**angle** entre X et Y. La définition* (16.118) *est souvent utilisée sous la forme*

$$X \cdot Y = \|X\|\|Y\|\cos(\theta). \tag{16.119}$$

Notez que les angles sont toujours des angles plus petits ou égaux à 180°.

La longueur de la projection du point P sur la droite horizontale va naturellement être égale à $\cos(\theta)$. En effet, si nous notons X un vecteur horizontal de norme 1, cette projection est donné par $P \cdot X$. Mais en reprenant l'équation (16.119), nous voyons que

$$P \cdot X = \|P\|\|X\|\cos(\theta), \tag{16.120}$$

tandis qu'ici nous avons $\|P\| = \|X\| = 1$.

Nous appelons $\sin(\theta)$ la longueur de la projection sur l'axe vertical.

Quelques dessins nous convainquent que

$$\begin{aligned} \sin(\theta + 2\pi) &= \sin(\theta) \quad \cos(\theta + 2\pi) = \sin(\theta), \\ \sin(\theta + \frac{\pi}{2}) &= \cos(\theta) \quad \cos(\theta + \frac{\pi}{2}) = -\sin(\theta), \\ \sin(\pi - \theta) &= \sin(\theta) \quad \cos(\pi - \theta) = -\cos(\theta). \end{aligned} \tag{16.121}$$

Le théorème de Pythagore nous montre aussi l'importante relation

$$\sin^2(\theta) + \cos^2(\theta) = 1. \tag{16.122}$$

Quelques valeurs remarquables pour les sinus et cosinus :

$$\begin{aligned} &\sin 0 = 0, \quad \sin\frac{\pi}{6} = \frac{1}{2}, \quad \sin\frac{\pi}{4} = \frac{\sqrt{2}}{2}, \quad \sin\frac{\pi}{3} = \frac{\sqrt{3}}{2}, \quad \sin\frac{\pi}{2} = 1, \quad \sin\pi = 0 \\ &\cos 0 = 1, \quad \cos\frac{\pi}{6} = \frac{\sqrt{3}}{2}, \quad \cos\frac{\pi}{4} = \frac{\sqrt{2}}{2}, \quad \cos\frac{\pi}{3} = \frac{1}{2}, \quad \cos\frac{\pi}{2} = 0, \quad \cos\pi = -1 \end{aligned} \tag{16.123}$$

Nous pouvons prouver simplement que $\sin(30°) = \frac{1}{2}$ et $\cos(30°) = \frac{\sqrt{3}}{2}$ en s'inspirant de la figure 16.5.

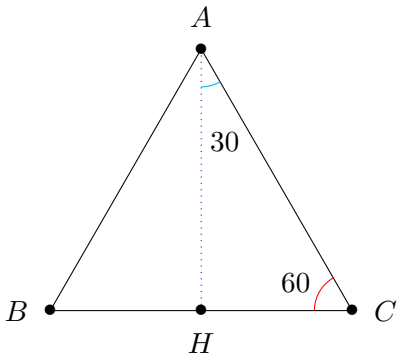


FIGURE 16.5 – Un triangle équilatéral de côté 1.

16.2.11 Aire du parallélogramme

Remarque 16.48.
Le nombre $\|a\|\|b\|\sin(\theta)$ est l'aire du parallélogramme formé par les vecteurs a et b, comme cela se voit sur la figure 16.6.

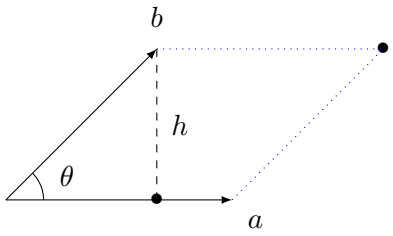


FIGURE 16.6 – Calculer l'aire d'un parallélogramme.

Proposition 16.49.
Nous avons

$$\|a \times b\| = \|a\|\|b\|\sin(\theta) \tag{16.124}$$

où $\theta \in [0.\pi]$ est l'angle formé par a et b.

Démonstration. En utilisant la décomposition du produit vectoriel, nous avons

$$\begin{aligned}
\|a \times b\|^2 &= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}^2 + \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}^2 + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}^2 \\
&= (a_2b_3 - b_2a_3)^2 + (a_1b_3 - a_3b_1)^2 + (a_1b_2 - a_2b_1)^2 \\
&= (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) - (a_1b_1 + a_2b_2 + a_3b_3)^2 \\
&= \|a\|^2\|b\|^2 - (a \cdot b)^2 \\
&= \|a\|^2\|b\|^2 - \|a\|^2\|b\|^2\cos^2(\theta) \\
&= \|a\|^2\|b\|^2\big(1 - \cos^2(\theta)\big) \\
&= \|a\|^2\|b\|^2\sin^2(\theta).
\end{aligned} \tag{16.125}$$

D'où le résultat. Nous avons utilisé la formule de la définition (16.47) donnant l'angle en fonction du produit scalaire. □

16.50.
Si les vecteurs a, b et c ne sont pas coplanaires, alors la valeur absolue du produit mixte (voir équation (9.173)) $a \cdot (b \times c)$ donne le volume du parallélépipède construit sur les vecteurs a, b et c.

En effet si φ est l'angle entre $b\times c$ et a, alors la hauteur du parallélépipède vaut $\|a\|\cos(\varphi)$ parce que la direction verticale est donnée par $b\times c$, et la hauteur est alors la «composante verticale» de a. Par conséquent, étant donné que $\|b\times c\|$ est l'aire de la base, le volume du parallélépipède vaut [7]

$$V = \|b\times c\|\|a\|\cos(\varphi). \tag{16.126}$$

Or cette formule est le produit scalaire de a par $b\times c$; ce dernier étant donné par le déterminant de la matrice formée des composantes de a, b et c grâce à la formule (9.173).

La valeur absolue du déterminant

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \tag{16.127}$$

est l'aire du parallélogramme déterminé par les vecteurs $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ et $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$. En effet, d'après la remarque 16.48, l'aire de ce parallélogramme est donnée par la norme du produit vectoriel

$$\begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix} = \begin{vmatrix} e_x & e_y & e_z \\ a_1 & a_2 & 0 \\ b_1 & b_2 & 0 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} e_z, \tag{16.128}$$

donc la norme $\|a\times b\|$ est bien donnée par la valeur absolue du déterminant (16.127).

16.3 Classification des isométries dans $\mathbb{R}^2$

16.3.1 Réflexions

Soit un espace vectoriel E de dimension 2 muni d'un produit scalaire [8]. Cela pourrait très bien être $\mathbb{R}^2$, mais nous allons nous efforcer de l'appeler E pour rester un peu général.

Lemme-définition 16.51 (Caractérisation des réflexions)**.**
Soit une droite ℓ de $\mathbb{R}^2$. Il existe une unique application $f\colon \mathbb{R}^2\to\mathbb{R}^2$ telle que

(1) $f(x)=x$ pour tout $x\in\ell$.

(2) f échange les côtés de ℓ.

(3) f laisse invariants les droites perpendiculaires à ℓ et les cercles dont le centre est sur ℓ.

*Cette application est la **réflexion** d'axe ℓ.*

Démonstration. Soit x hors de ℓ et p la droite perpendiculaire à ℓ et passant par x. Nous avons $f(x)\in p$. En nommant P l'intersection entre ℓ et p, nous considérons le cercle $S(P,\|Px\|)$ qui est un cercle dont le centre est sur ℓ. Il contient x et donc $f(x)\in S(P,\|Px\|)$.

Donc $f(x)\in p\cap S(P,\|Px\|)$. L'intersection entre un cercle et une droite contient de façon générique deux point. L'un est x, mais $f(x)=x$ n'est pas possible parce que x est hors de ℓ et f doit inverser les côtés de ℓ. Donc $f(x)$ est l'autre.

Cela prouve l'unicité. En ce qui concerne l'existence, il suffit de noter que la réflexion σ_ℓ satisfait les contraintes. □

Lemme 16.52.
Soit une droite ℓ est $A\in E$. Alors

$$\sigma_\ell(A) = 2\,\texttt{proj}_\ell(A) - A \tag{16.129}$$

où $\texttt{proj}_\ell$ est l'opération de projection orthogonale sur la droite ℓ.

7. Le calcul de ce volume mériterait une certaine réflexion, surtout à partir du moment où nous avons décidé de définir les fonctions trigonométriques à partir de son développement (définition 16.19).

8. Définition 9.24.

Démonstration. Nous posons $f(X) = 2\texttt{proj}_\ell(X) - X$ et nous allons montrer que $f = \sigma_\ell$ en vérifiant les conditions de la définition 16.51. Nous nous gardons bien de faire un raisonnement du type «nous allons montrer que f et σ_ℓ coïncident sur deux points, et sont donc égales par le corollaire 16.11» parce que nous ne savons pas encore que σ_ℓ est une application affine, ni même que c'est une isométrie.

Si $X \in \ell$ alors $\texttt{proj}_\ell(X) = X$ et nous avons $f(X) = 2X - X = X$. Donc ℓ est conservée.

En ce qui concerne les deux côtés de ℓ, il existe une application linéaire $s\colon E \to \mathbb{R}$ et une constante $c \in \mathbb{R}$ telles qu'en posant $\ell(X) = s(X) + c$, la droite ℓ soit le lieux des points X tels que $\ell(X) = 0$. Un côté de la droite est $\ell < 0$ et l'autre côté est $\ell > 0$. Nous avons :

$$\begin{aligned}
\ell\big(f(A)\big) &= \ell\big(2\texttt{proj}_\ell(A) - A\big) && (16.130a)\\
&= s(2\texttt{proj}_\ell(A) - A) + c && (16.130b)\\
&= 2s\big(\texttt{proj}_\ell(A)\big) - s(A) + c && (16.130c)\\
&= s\big(\texttt{proj}_\ell(A)\big) - s(A) && (16.130d)\\
&= -c - s(A) && (16.130e)\\
&= -\ell(A) && (16.130f)
\end{aligned}$$

où nous avons utilisé le fait que, $\texttt{proj}_\ell(A)$ étant sur ℓ, $s\big(\texttt{proj}_\ell(A)\big) + c = 0$. Nous avons donc $\ell\big(f(A)\big) = -\ell(A)$, ce qui indique que A et $f(A)$ sont de part et d'autre de ℓ.

Si d est une droite perpendiculaire à ℓ et si $A \in d$ alors $f(A) = 2\texttt{proj}_\ell(A) - A = \big(\texttt{proj}_\ell(A) - A\big) + A \in d$ parce que $\texttt{proj}_\ell(A) \in d$ du fait que d soit précisément perpendiculaire à ℓ. Nous avons aussi utilisé le fait que si $A, B, C \in d$ alors $(B - A) + C \in d$; pensez que $B - A$ est un vecteur directeur et que C est un point de d.

Enfin soit $K \in \ell$ et un cercle $S(K, r)$ centré en K. Soit $A \in S(K, r)$; nous devons vérifier que $f(A) = S(K, r)$. Le segment $[A, f(A)]$ est par définition perpendiculaire à ℓ. Soit M, le milieu, qui est sur la droite ℓ. Les triangles AMK et $f(A)MK$ sont rectangles en M, et $\|AM\| = \|Mf(A)\|$. Le théorème de Pythagore donne $\|AK\| = \|f(A)K\|$. Donc le cercle centré en K est donc préservé par f.

Nous en déduisons que $f = \sigma_\ell$. □

Proposition 16.53 ([200])**.**
Une réflexion est une isométrie de (E, d) où $d(A, B) = \|A - B\|$.

Démonstration. Soient $A, B \in E$; il faut vérifier que $\|A - B\| = \|\sigma_\ell(A) - \sigma_\ell(B)\|$. Pour cela nous écrivons

$$B - A = \underbrace{B - \texttt{proj}_\ell(B)}_{=a} + \underbrace{\texttt{proj}_\ell(B) - \texttt{proj}_\ell(A)}_{=b} + \underbrace{\texttt{proj}_\ell(A) - A}_{=c}. \tag{16.131}$$

Vu que $b \perp a$ et $b \perp c$ nous avons

$$\|B - A\| = \langle B - A, B - A\rangle = \|a\|^2 + 2\langle a, c\rangle + \|b\|^2 + \|c\|^2. \tag{16.132}$$

Nous pouvons faire le même jeu avec $\sigma_\ell(B) - \sigma_\ell(A)$ en tenant compte du fait que $\texttt{proj}_\ell\big(\sigma_\ell(X)\big) = \texttt{proj}_\ell(X)$ et que

$$\sigma_\ell(A) - \texttt{proj}_\ell(A) = 2\texttt{proj}_\ell(A) - A - \texttt{proj}_\ell(A) = -\big(A - \texttt{proj}_\ell(A)\big). \tag{16.133}$$

Là nous avons utilisé le lemme 16.52. Ce que nous trouvons est que

$$\sigma_\ell(B) - \sigma_\ell(A) = -a + b - c, \tag{16.134}$$

et donc encore une fois

$$\|\sigma_\ell(B) - \sigma_\ell(A)\| = \|a\|^2 - 2\langle a, c\rangle + \|b\|^2 + \|c\|^2. \tag{16.135}$$

□

Remarque 16.54.
Il faut bien comprendre que si l'axe de la réflexion ne passe par par 0 (le zéro de l'espace vectoriel normé $(E, \|.\|)$), la réflexion n'est pas une isométrie de $(E, \|.\|)$ au sens où nous n'avons pas $\|\sigma_\ell(x)\| = \|x\|$.

Lemme 16.55.
Si A' est l'image de A par σ_ℓ alors ℓ est la médiatrice du segment $[A, A']$.

Démonstration. Soit $M \in \ell$. Nous avons

$$\|A - M\|^2 = \|\,\mathtt{proj}_\ell(A) - A\|^2 + \|\,\mathtt{proj}_\ell(A) - M\|^2 \tag{16.136}$$

parce que $A - \mathtt{proj}_{ell}(A) \perp M - \mathtt{proj}_\ell(A)$. Par ailleurs, vu que $\sigma_\ell(A) = 2\,\mathtt{proj}_\ell(A) - A$ et que $\mathtt{proj}_\ell(A) = \mathtt{proj}_\ell(A')$,

$$\|\,\mathtt{proj}_\ell(A) - A\| = \|\,\mathtt{proj}_\ell(A') - A'\|. \tag{16.137}$$

Nous avons donc

$$\|\sigma_\ell(A) - M\|^2 = \|A - M\|^2, \tag{16.138}$$

ce qui prouve que M est sur la médiatrice de $[A', A]$ par le lemme 16.14. □

16.56.
Si l est une droite dans $\mathbb{R}^2$, nous avons la réflexion $\sigma_l \in \mathrm{Isom}(\mathbb{R}^2)$ d'axe l. Cela est une isométrie et donc une application affine par le théorème 16.4. Le lemme suivant détermine comment la réflexion σ_ℓ se décompose en une translation et une application linéaire.

Lemme 16.57.
Soit une droite ℓ. Alors

$$\sigma_\ell = \tau_{2w} \circ \sigma_{\ell_0} \tag{16.139}$$

où ℓ_0 est la droite parallèle à ℓ passant par l'origine, et w est le vecteur perpendiculaire à ℓ tel que $\ell_0 = \ell + v$.

Démonstration. Il faut trouver trois points non alignés sur lesquels les deux applications coïncident ; cela suffira par le corollaire 16.11.

Pour tous les points de ℓ_0, l'égalité fonctionne parce que si $x \in \ell_0$,

$$\sigma_\ell(x) = x + 2w, \tag{16.140}$$

tandis que

$$\sigma_{\ell_0}(x) + 2w = x + 2w \tag{16.141}$$

du fait que $\sigma_{\ell_0}(x) = x$.

Si $x \in \ell$, alors

$$\sigma_\ell(x) = x \tag{16.142}$$

tandis que

$$\sigma_{\ell_0}(x) + 2w = x - 2w + 2w = x. \tag{16.143}$$

Donc les applications affines σ_ℓ et $x \mapsto \sigma_{\ell_0}(x) + 2w$ coïncident sur ℓ et ℓ_0. Elles coïncident donc partout. □

16.3.2 Rotations

Définition 16.58 (Rotation en dimension 2)**.**
*Une **rotation** d'un espace euclidien de dimension 2 est une composée de deux réflexions d'axes non parallèles. L'identité est une rotation.*

16.59.
Quelque remarques à propos de cette définition.

(1) Attention : nous ne parlons pas encore de rotations «vectorielles» : ici le centre de la rotation (que nous n'avons pas encore défini) peut ne pas être 0.

(2) Dans la même veine : plus tard, lorsque nous saurons que les rotations sont des isométries de (E, d) où $d(X, Y) = \|X - Y\|$, nous allons en réalité beaucoup plus souvent parler de rotations centrées en l'origine qu'en un point quelconque. C'est pourquoi à partir de 16.69 nous dirons le plus souvent «rotation» pour «rotation centrée en 0». D'où les énoncés comme «les rotations sont les matrice orthogonales» (corollaire 16.81) , qui *stricto senus* de la définition 16.58 sont faux.

(3) Une rotation est composée de deux réflexions d'axes non parallèles. Il est cependant trop tôt pour décréter que l'intersection de ces axes est le centre de la rotation. Rien ne dit en effet pour l'instant que deux décompositions différentes de la même rotation, avec des axes différents donnent le même point d'intersection.

(4) Pourquoi ajouter l'identité ? Pour avoir un groupe. Dans le cas vectoriel, il est suffisant de demander d'être une composée de deux réflexions, parce que toutes les réflexions vectorielles ont des axes qui s'intersectent en 0. Le cas des axes parallèles est seulement le cas des axes confondus et revient à l'identité.

Si nous voulons avoir un groupe même pour les rotations centrées ailleurs qu'en zéro, nous devons ajouter «à la main» l'identité.

Toutes ces remarques se résument par : «tout devient compliqué du fait que nous voulons considérer également les rotations centrées ailleurs qu'en zéro». En se contentent du cas vectoriel, de nombreuses choses sont plus simples.

Corollaire 16.60.
Si $A \neq B$ dans E alors il existe une unique réflexion envoyant A sur B.

Démonstration. En ce qui concerne l'existence, la réflexion dont l'axe est la médiatrice de $[A, B]$ fait l'affaire. En ce qui concerne l'unicité, le lemme 16.55 nous dit que si A est envoyé sur B, l'axe est forcément la médiatrice de $[A, B]$. □

Lemme 16.61 ([194])**.**
La rotation $r = \sigma_1 \circ \sigma_2$, si elle est différente de l'identité, ne fixe que le point d'intersection $\ell_1 \cap \ell_2$.

Démonstration. Nous nommons $O = \ell_1 \cap \ell_2$. Soit $A \in E$, et supposons que $r(A) = A$. Nous avons $\sigma_1 \circ r = s_2$ et donc

$$\sigma_1(A) = (\sigma_1 \circ r)(A) = s_2(A). \tag{16.144}$$

On pose $B = \sigma_1(A)$. Alors σ_1 et σ_2 envoient tout deux A sur B.

Si $A = B$ alors A est fixé par σ_1 et donc appartient à ℓ_1. Même chose pour A est fixé par σ_2 et donc $A \in \ell_2$. Cela donne $A = B = O$, et donc le point fixé par r est O.

Si $A \neq B$ alors il existe une unique réflexion envoyant A sur B (corollaire 16.60). L'unicité signifie que $\sigma_1 = \sigma_2$. Dans ce cas, $r = \sigma_1 \circ \sigma_2 = \mathrm{Id}$. □

16.62.
La rotation $\sigma_1 \circ \sigma_2$ laisse évidemment fixé le point $\ell_1 \cap \ell_2$. Si $\sigma_1 \circ \sigma_2 = \sigma_a \circ \sigma_b$ alors rien n'oblige les axes de σ_1 et σ_2 d'être identiques à ceux que σ_a et σ_b. Mais l'intersection $\ell_1 \cap \ell_2$ doit être la même que l'intersection $\ell_a \cap \ell_b$ parce que c'est l'unique point fixé par la composée. Cela nous permet de poser les définition suivante.

Définition 16.63.
*Le **centre** d'une rotation (autre que l'identité) est l'unique point fixé par la rotation.*

Comme expliqué dans 16.62, le centre de la rotation s'obtient en l'écrivant comme composée de deux réflexions et en considérant le point d'intersection des axes.

Lemme 16.64.
Les rotation sont des isométries pour la distance : $\|X - Y\| = \|r(X) - r(Y)\|$.

Démonstration. Si $r = \sigma_1 \circ \sigma_2$, en utilisant le fait que σ_1 et σ_2 sont des isométries de (E, d) (16.53) nous avons :

$$d(X, Y) = d\big(\sigma_2(X), \sigma_2(Y)\big) = d\big(\sigma_1\sigma_2(X), \sigma_1\sigma_2(Y)\big) = d\big(r(X), r(Y)\big). \tag{16.145}$$

□

Ce lemme nous dit qu'une rotation de centre O vérifie $\|OX\| = \|Or(X)\|$ pour tout X.

Proposition 16.65 ([194])**.**
Soient $A, B, O \in E$ tels que $\|AO\| = \|BO\| \neq 0$. Alors il existe une unique rotation r centrée en O telle que $r(A) = B$.

Problèmes et choses à faire
Attention : la preuve qui suit contient de nombreuses galipettes et improvisations personnelles. Relisez-la attentivement avant de la prendre pour argent comptant.

La difficulté tient essentiellement à ce que cette preuve traite de façon vectorielle (tous les points sont des éléments de E) un énoncé qui est essentiellement affine : tous les points doivent être vus comme vecteurs partant de O.

Si vous comparez la preuve donnée ici avec celle de [194], vous remarquerez que dans ce dernier, seule la partie «A et O son alignés» est présente. C'est parce que lui, il se met directement dans le cas vectoriel et $O = 0$. Il a donc une preuve un tout petit peu moins générale, mais au moins ses isométries sont linéaires et non affines.

Démonstration. Existence et unicité séparément.

Existence Si $A = B$, l'identité fait l'affaire. Sinon, $\|A-O\| = \|B-O\|$ implique que la médiatrice de $[A, B]$ contient O. Soit σ_m la réflexion selon cette médiatrice. La rotation $\sigma_m \circ \sigma_{(AO)}$ convient.

Unicité Soit r une rotation de centre O et telle que $r(A) = B$. Si $A = B$ alors $r = \mathrm{Id}$ parce qu'une rotation autre que l'identité ne fixe que son centre par le lemme 16.61. Nous supposons que $A \neq B$.

Nous posons $g = \sigma_m \circ r$. Alors $g(A) = \sigma_m(B) = A$ parce que $\sigma_m(B) = A$ et $r(A) = B$. Cela signifie que g est une isométrie qui fixe A.

Si A et O ne sont pas alignés Attention : ici O est un point de E, pas le zéro de l'espace vectoriel E. Lorsqu'on dit que A et O ne sont pas alignés, nous parlons bien d'alignement avec le zéro de E.

Nous avons $g(A) = A$ et $g(O) = O$. Donc g coïncide avec $\sigma_{(AO)}$ en deux points non alignés, c'est à dire en deux points pour lesquels l'espace engendré est tout E. Nous en déduisons que $g = \sigma_{(AO)}$.

Si A et O sont alignés Soit maintenant un point C tel que $A - O \perp C - O$ et

$$\|OC\| = \|OA\| = \|OB\|. \tag{16.146}$$

Vu que g est une isométrie pour la distance sur E, pas pour la norme, nous ne pouvons pas écrire $g(C - O) \perp g(A - O)$ à partir de $C - O \perp A - O$. Nous décomposons $g(X) = s(X) + G$ où s est linéaire sur E. Il est vite vu que s est une isométrie de $(E, \|.\|)$:

$$\|X - Y\| = \|g(X) - g(Y)\| = \|s(X) + G - s(y) - G\| = \|s(X) - s(Y)\| = \|s(X - Y)\| \tag{16.147}$$

pour tout $X, Y \in E$. Nous avons de plus $g(A) = A$ et $g(O) = O$, ce qui donne $O = s(O) + G$ et $A = s(A) + G$. En égalisant les valeurs de G nous avons

$$O - s(O) = A - s(A). \tag{16.148}$$

Vu que s est une isométrie (une vraie) nous avons

$$s(A - O) \perp s(C - O), \tag{16.149}$$

mais $s(A-O)=s(A)-s(O)=A-O$ par (16.148). Donc

$$A-O\perp s(C-O). \tag{16.150}$$

Nous en concluons que $s(C-O)=\pm(C-O)$. Parce que les vecteurs $\pm(C-O)$ sont les deux seuls de norme $\|AO\|=\|CO\|$ à être perpendiculaire à $A-O$. Rappel : la définition de C et le fait que nous soyons en dimension 2.o

Est-il possible d'avoir $s(C-O)=C-O$? Cela donnerait

$$g(A-O)=s(A)-s(O)+G=s(A)-O+O-s(O)+G=A-O+G \tag{16.151a}$$
$$g(C-O)=C-O+G, \tag{16.151b}$$

ce qui signifierait que g et τ_G coïncideraient sur les points $A-O$ et $C-O$, et donc seraient égaux par le corollaire 16.11. Cela est cependant impossible parce que g fixe au moins les points A et O alors que la translation ne fixe aucun point. Nous en déduisons $s(C-O)=-(C-O)$.

Nous avons aussi, parce que (AO) est une droite passant par l'origine que

$$\sigma_{(AO)}(A-O)=A-O \tag{16.152}$$

et parce que $C-O$ est perpendiculaire à cette droite :

$$\sigma_{(AO)}(C-O)=-(C-O). \tag{16.153}$$

Nous avons donc quand même que g et $\sigma_{(AO)}$ coïncident sur deux points non alignés : $A-O$ et $C-O$.

Dans tous le cas, $g=\sigma_{(AO)}$. Nous avons donc

$$\sigma_{(OA)}=\sigma_m\circ r, \tag{16.154}$$

et donc r est fixé à

$$r=\sigma_m\circ\sigma_{(OA)}. \tag{16.155}$$

□

16.66.
Anticipons un peu et faisons semblant de déjà connaitre les matrices et les fonctions trigonométriques. La proposition 16.65 nous dit qu'il existe une seule rotation amenant A sur B. Vous pourriez objecter que le point $(1,0)$ peut être amené sur $(0,-1)$ soit par la rotation d'angle $3\pi/2$, soit par celle d'angle $-\pi/2$. Il n'en est rien parce que ces deux rotations sont les mêmes ! Pensez-y. En tant qu'application $\mathbb{R}^2\to\mathbb{R}^2$, la rotation $R_{3\pi/2}$ est égale à $R_{-\pi/2}$.

Une rotation donnée peut être écrite de beaucoup de façons comme composée de deux réflexions. En fait d'autant de façons qu'il y a de réflexions.

Proposition 16.67 ([194])**.**
Soit une rotation r de E centrée en O. Pour toute réflexion σ_ℓ telle que le centre de r soit sur ℓ, il existe une réflexion σ_1 tells que $r=\sigma_1\circ\sigma_\ell$. Il existe aussi une réflexion σ_2 telle que $r=\sigma_\ell\circ s_2$.

Démonstration. Si $r=\mathrm{Id}$ c'est bon avec $s_1=s_2=\sigma_\ell$. Sinon nous considérons $A\neq O$ sur ℓ, et $B=r(A)$. Nous savons que $B\neq A$ parce que O est le seul point de E fixé par r (proposition 16.61). Il existe une réflexion (unique) σ_1 faisant $\sigma_1(A)=B$, et c'est le réflexion dont l'axe est la médiatrice de $[A,B]$. Le point O est sur cette médiatrice parce que les rotations sont des isométries de (E,d) (lemme 16.64).

La rotation $\sigma_1\circ\sigma_\ell$ vérifie

$$(\sigma_1\circ\sigma_\ell)(A)=\sigma_1(A)=B. \tag{16.156}$$

Or $\|OA\| = \|OB\|$, donc il y a unicité de la rotation centrée en O portant A sur B (proposition 16.65) ; nous avons donc $r = \sigma_1 \circ \sigma_\ell$.

En ce qui concerne $r = \sigma_\ell \circ \sigma_2$, il faut appliquer ce que nous venons de faire à la rotation r^{-1} : il existe σ_2 tel que $r^{-1} = \sigma_2 \circ \sigma_\ell$, ce qui donne

$$r = \sigma_\ell \circ \sigma_2. \tag{16.157}$$

□

Proposition 16.68 ([194])**.**
Les rotations basées en O forment un groupe abélien.

Démonstration. L'identité est une rotation par définition. En ce qui concerne l'inverse, si $r = \sigma_1\sigma_2$ alors $r^{-1} = \sigma_2\sigma_1$. Nous commençons maintenant les choses pas tout à fait évidentes.

Composition Soient des rotations r, r' centrées en O. Soit également une réflexion σ dont l'axe contient O. Alors la proposition 16.67 nous donne l'existence de σ_1 et σ_2 tels que $r = \sigma_1\sigma$ et $r' = \sigma\sigma_2$. Avec ça, la composition donne

$$rr' = \sigma_1\sigma\sigma\sigma_2 = \sigma_1\sigma_2, \tag{16.158}$$

qui est encore une rotation.

Commutativité Soient deux rotations r et r' ainsi que des décompositions $r = \sigma_1\sigma$, $r' = \sigma\sigma_2$. Nous avons

$$rr' = \sigma_1\sigma_2 \tag{16.159a}$$
$$r'r = \sigma\sigma_2\sigma_2\sigma. \tag{16.159b}$$

Vu que $t = \sigma_2\sigma_1$ est une rotation nous pouvons encore appliquer la proposition 16.67 pour avoir $t = \sigma_2\sigma_1 = \sigma\sigma_3$. Avec ça,

$$r'r = \sigma\sigma\sigma_3\sigma = \sigma_3\sigma. \tag{16.160}$$

Mais aussi $rr' = \sigma_1\sigma_2 = t^{-1} = \sigma_3\sigma$. Nous avons donc bien $rr' = r'r$, et le groupe est commutatif.

□

16.3.3 Rotations vectorielles

16.69.
Jusqu'à présent nous avons parlé de rotations «affines». Parmi elles, Les rotations centrées en 0 (zéro, l'origine de E comme espace vectoriel) sont de particulière importance. Ce sont des applications linéaires, et même des isométries. Dans la suite, nous allons souvent dire simplement «rotation» pour dire «rotation centrée en 0».

Vu que nous allons maintenant prendre un point de vue plus vectoriel, nous allons noter les points de E avec des lettres comme x, y, u, v et plus avec des majuscules, comme quand on avait un point de vue affin. En même temps, nous allons noter les applications $E \to E$ par des lettres comme A et ne plus écrire les parenthèses. Bref, nous écrivons Au au lieu de $r(A)$.

Lemme 16.70.
En dimension 2, les réflexions (vectorielles, c'est à dire dont l'axe passe par 0) ont un déterminant -1.

Démonstration. Soit une réflexion d'axe ℓ. Prenons une base orthonormale de E constituée de e_1 sur ℓ et de $e_2 \perp \ell$. Alors $\sigma_\ell(e_1) = e_1$ et $\sigma_\ell(e_2) = -e_2$. la formule du déterminant donne

$$\det(\sigma_\ell) = e_1^*\big(\sigma_\ell(e_1)\big)e_1^*\big(\sigma_\ell(e_2)\big) - e_2^*\big(\sigma_\ell(e_1)\big)e_1^*\big(\sigma_\ell(e_2)\big) = 1 \times (-1) - 0 \times 0 = -1. \tag{16.161}$$

Nous utilisons de façon cruciale le fait que le calcul du déterminant ne dépende pas de la base choisie, lemme 9.8. □

Proposition 16.71.
Les rotations[9] *sont*

(1) des applications orthogonales au sens de la définition 9.47,

(2) des applications de déterminant 1*,*

Démonstration. Nous avons, pour tout $u \in E$ l'égalité de la norme $\|u\|$ et $\|Au\|$ par le lemme 16.64 appliqué à $Y = 0$. En terme de produits scalaires nous avons alors $\langle Au, Au\rangle = \langle u, u\rangle$, et donc

$$\langle A^*Au, u\rangle = \|u\|^2. \tag{16.162}$$

En particulier si $\{e_i\}_{i=1,\dots,n}$ est une base orthonormée de E nous avons

$$(A^*Ae_i)_i = \|e_i\|^2 = 1, \tag{16.163}$$

ce qui donne $\|A^*Ae_i\| \geqslant 1$, avec égalité si et seulement si $A^*Ae_i = e_i$. Ici nous avons utilisé le fait que $\langle x, e_i\rangle = x_i$, et le fait que pour tout i nous ayons $\|x\| \geqslant |x_i|$, avec égalité seulement si x est un multiple de e_i.

Par ailleurs l'inégalité de Cauchy-Schwarz 9.25 nous donne

$$\|u\|^2 = |\langle A^*Au, u\rangle| \leqslant \|A^*Au\|\|u\| \tag{16.164}$$

et donc

$$\|u\| \leqslant \|A^*Au\|. \tag{16.165}$$

Encore une fois, en appliquant cela à $u = e_i$ nous trouvons $1 \leqslant \|A^*Ae_i\|$. Vu que nous avions déjà l'inégalité dans l'autre sens, $\|A^*Ae_i\| = 1$. Et le cas d'égalité est uniquement possible avec $A^*Ae_i = e_i$.

Donc pour tout i de la base nous avons $A^*Ae_i = e_i$. Nous avons donc $A^*A = \mathrm{Id}$ et l'application A est orthogonale.

En ce qui concerne le déterminant, les réflexions sont de déterminant -1 par le lemme 16.70, donc $A = \sigma_1 \circ \sigma_2$ est de déterminant 1. Nous avons utilisé le fait que le déterminant était un morphisme : proposition 9.10(1). □

Remarque 16.72.
Nous ne savons pas encore que les rotations forment tout le groupe $\mathrm{SO}(2)$ des endomorphismes orthogonaux de déterminant 1. Il faudra attendre le corollaire 16.81 pour le savoir.

Lemme 16.73.
L'application $-\mathrm{Id}$ *est une rotation.*

Démonstration. Soit une base orthonormée $\{e_1, e_2\}$ de E et la rotation $r = \sigma_1\sigma_2$ où σ_i est la réflexion le long de l'axe $\ell_i = \{te_i\}_{t\in\mathbb{R}}$. Faut-il vous prouver que $r = -\mathrm{Id}$? La réflexion σ_2 retourne la composante y d'un vecteur écrit dans la base $\{e_1, e_2\}$ sans toucher à la composante x. La réflexion σ_1 fait le contraire. □

16.3.4 Angle

Avant d'aborder la classification des isométries, nous devons parler de l'angle entre deux droites. Si ℓ_1 et ℓ_2 sont deux droites, alors il est bien clair deux angles peuvent prétendre être «l'angle entre ℓ_1 et ℓ_2». De plus chacun de ces deux angles sont doubles parce que si α peut prétendre être l'angle entre ℓ_1 et ℓ_2, alors $-\alpha$ peut également prétendre.

Lemme-définition 16.74.
Si ℓ_1 et ℓ_2 sont deux droites sécantes au point O et si $x \in \ell_1$ n'est pas O, alors il existe un unique $\alpha \in [0, \pi[$ tel que $R_O(\alpha)x \in \ell_2$. La valeur de α ne dépend pas du choix du point $x \in \ell_1$.

*Cet angle α est l'**angle** de ℓ_1 à ℓ_2.*

9. Centrées en 0, nous ne le répéterons pas !

Remarque 16.75.
Nous ne parlons pas de l'angle entre ℓ_1 et ℓ_2 mais bien de l'angle *de* ℓ_1 *à* ℓ_2. L'ordre des droites est important.

16.76.
Pour la suite, $R_O(\alpha)$ est la rotation d'angle α autour du point O tandis que $R(\alpha)$ est la rotation d'angle α autour de l'origine. En termes matriciels, la rotation d'angle α est donnée par

$$R(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}, \tag{16.166}$$

Lemme 16.77 ([1])**.**
Soit $A \in \mathbb{R}^2$ et une droite ℓ_1. Soit ℓ_2 une droite passant par A et intersectant ℓ_1 en O. Alors

$$\sigma_{\ell_1}(A) = R_O(-2\alpha)A \tag{16.167}$$

où α est l'angle de ℓ_1 à ℓ_2.

Démonstration. Nous allons utiliser des coordonnées autour de O. Il existe un vecteur v tel que

$$A = O + v \tag{16.168}$$

Par définition de l'angle α, la droite ℓ_2 s'obtient par rotation d'angle α depuis la droite ℓ_1. Donc le point

$$B = R_O(-\alpha)A \tag{16.169}$$

est sur ℓ_1.

Nous allons prouver que le point

$$D = R_O(-2\alpha)A \tag{16.170}$$

est $D = \sigma_{\ell_1} A$.

Nous commençons par montrer que la droite (DA) est perpendiculaire à ℓ_1, c'est à dire que

$$(D - A) \cdot (B - O) = 0. \tag{16.171}$$

En utilisant le fait que

$$R_O(\alpha)(O + X) = O + R(\alpha)X, \tag{16.172}$$

nous avons

$$D - A = R_O(-2\alpha)(O + v) - (O + v) = O + R(-2\alpha)v - O - v = R(-2\alpha)v - v \tag{16.173}$$

et de la même façon,

$$B - O = R(-\alpha)v. \tag{16.174}$$

Notons que tous les O se sont simplifiés et qu'il ne reste que des rotation usuelles. En utilisant le fait que $R(\alpha)$ est une isométrie, nous pouvons alors calculer

$$(D - A) \cdot (B - O) = \langle R(-2\alpha)v - v, R(-\alpha)v \rangle \tag{16.175a}$$
$$= \langle R(-\alpha)v - R(\alpha)v, v \rangle. \tag{16.175b}$$

En utilisant la matrice de rotation (16.166) nous trouvons

$$\big(R(-\alpha) - R(\alpha)\big)v = \begin{pmatrix} 2\sin(\alpha)v_2 \\ -2\sin(\alpha)v_1 \end{pmatrix} \tag{16.176}$$

et donc

$$\big\langle \big(R(-\alpha) - R(\alpha)\big)v, v \big\rangle = 0. \tag{16.177}$$

Le point D est bien sûr la droite perpendiculaire à ℓ_1 et passant par A. Mais vu que D est obtenu à partir de A par une rotation, le point D est également sur le cercle de rayon $\|OA\|$ et centré en O. Ce cercle possède exactement deux intersections avec cette droite. Le premier est A et le second est $\sigma_{\ell_1}(A)$. Vu que D n'est pas A, nous avons $D = \sigma_\ell(A)$. □

16.3.5 Classification

Théorème 16.78 ([195]).
Toute isométrie du plan est une composition d'au plus 3 *réflexions.*

Démonstration. Encore une fois nous décomposons la preuve en fonction du nombre de points fixes.

Si f n'a pas de points fixes Soit $x \in \mathbb{R}^2$ et l, la médiatrice du segment $[x, f(x)]$. Par construction, $f(x) = \sigma_l(x)$. Nous posons $g = \sigma_l \circ f$, et nous avons

$$g(x) = x. \tag{16.178}$$

Donc nous avons $f = \sigma_l \circ g$ avec $x \in \mathrm{Fix}(g)$.

Si f a un unique point fixe Soit x cet unique point fixe. Soit $y \neq x$ et l la médiatrice de $[y, f(y)]$. En posant $g = \sigma_l \circ f$ nous avons

$$g(y) = y \tag{16.179}$$

et $g(x) = x$ parce que

$$d\big(x, f(y)\big) = d\big(f(x), f(y)\big) = d(x, y), \tag{16.180}$$

ce qui donne que x est à égale distance de y et de $f(y)$, c'est à dire que $x \in l$ et par conséquent $g(x) = (\sigma_l \circ f)(x) = \sigma_l(x) = x$.

Donc g fixe x et y et donc toute la droite (xy).

Si f fixe une droite Soit l une droite fixée par f, et soient $x, y \in l$ et $z \notin l$ (avec $x \neq y$). Le fait que x et y soient des points fixes de f implique

$$\begin{cases} d\big(x, f(z)\big) = d(x, z) & \text{(16.181a)} \\ d\big(y, f(z)\big) = d(y, z) & \text{(16.181b)} \end{cases}$$

ce qui signifie que $f(z)$ est sur l'intersection des deux cercles [10] $S\big(x, d(x, z)\big)$ et $S\big(y, d(y, z)\big)$, et comme ce sont deux cercles centrés sur la droite l, les intersections sont liées par σ_l. Autrement dit, les intersections sont z et $\sigma_l(z)$.

Si $f(z) = z$ alors f fixe trois points non alignés et fixe dont $\mathbb{R}^2$, c'est à dire $f = \mathrm{Id}$.

Si par contre $f(z) = \sigma_l(z)$ alors les isométries f et σ_l coïncident sur trois points et coïncident donc partout par le corollaire 16.11 : $f = \sigma_l$.

Conclusion Nous avons montré que si $\mathrm{Fix}(f)$ a dimension m, alors il existe une droite pour laquelle $f = \sigma_l \circ g$ avec $\dim\big(\mathrm{Fix}(g)\big) > m$. Donc il faux au maximum trois pas pour avoir $\dim\big(\mathrm{Fix}(g)\big) = 2$ c'est à dire pour avoir $g = \mathrm{Id}$.

□

Définition 16.79.
Une ***réflexion glissée*** *est une transformation du plan de la forme $\tau_v \circ \sigma_\ell$ où le vecteur v est parallèle à la droite ℓ.*

Théorème 16.80 ([195]).
Les isométries du plan sont exactement

(1) l'identité (composée de 0 *réflexions),*

(2) les réflexions,

(3) les translations (composées de 2 *translations d'axes parallèles),*

(4) les rotations (composées de 2 *réflexions d'axes non parallèles),*

(5) les réflexions glissées (composées de 3 *réflexions)*

10. L'intersection existe pare que $d(x, z) + d(y, z) > d(x, y)$.

Démonstration. Nous savons déjà que $f \in \mathrm{Isom}(\mathbb{R}^2)$ est une composée de 0, 1, 2 ou 3 réflexions.

Zéro réflexions Alors c'est l'identité. Ce n'est pas très profond.

Une réflexion Alors f est une réflexion. Toujours pas très profond.

Deux réflexions Soit $f = \sigma_{\ell_1} \circ \sigma_{\ell_2}$. Maintenant ça s'approfondit un bon coup.

Nous supposons d'abord que $\ell_1 \parallel \ell_2$. Dans ce cas nous allons prouver que $f = \tau_{2v}$ où v est le vecteur perpendiculaire à ℓ_1 tel que $\ell_1 + v = \ell_2$. Nous allons utiliser le lemme 16.57 pour montrer que $\sigma_{\ell_1} \circ \sigma_{\ell_2} = \tau_{2v}$. Nous avons

$$\ell_1 = \ell_0 + w \tag{16.182a}$$
$$\ell_2 = \ell_0 + w + v \tag{16.182b}$$

où w est un vecteur perpendiculaire à ℓ_1 et ℓ_0 est la droite passant par l'origine et parallèle à ℓ_1 et ℓ_2. Avec cela,

$$\begin{aligned}(\sigma_{\ell_1} \circ \sigma_{\ell_2})(x) &= \sigma_{\ell_1}\big(\sigma_{\ell_0}(x) + 2w\big) && (16.183a)\\ &= \sigma_{\ell_0}\big(\sigma_{\ell_0}(x) + 2w\big) + 2(v + w) && (16.183b)\\ &= x + \underbrace{\sigma_{\ell_0}(2w)}_{-2w} + 2v + 2w && (16.183c)\\ &= x + 2v. && (16.183d)\end{aligned}$$

Donc si f est composée de deux réflexions d'axes parallèles, alors f est une translation.

Toujours dans le cas où f est composée de deux réflexions, nous supposons que $f = \sigma_{\ell_2} \circ \sigma_{\ell_1}$ avec ℓ_1 et ℓ_2 non parallèles. Nous notons O le point d'intersection, et nous allons voir que $f = R_O(2\alpha)$ où α est l'angle de ℓ_1 à ℓ_2 donné par le lemme 16.74.

Soit $x \in \ell_1$. Alors

$$f(x) = \sigma_{\ell_2}(x), \tag{16.184}$$

et le lemme 16.77 nous donne un moyen de calculer $\sigma_{\ell_2}(x)$ parce que ℓ_1 est une droite passant par x et coupant ℓ_1 au point O. Le lemme dit que $\sigma_{\ell_2}(x) = R_O(2\alpha)$. Remarque : c'est bien 2α et non -2α parce qu'il s'agit de l'angle de ℓ_2 à ℓ_2 ; il y a inversion des numéros entre ici et l'énoncé du lemme.

Nous avons donc bien $f(x) = R_O(2\alpha)x$ pour $x \in \ell_1$.

Si $y \in \ell_2$ alors

$$f(y) = \sigma_{\ell_2}\big(R_O(-2\alpha)y\big) \tag{16.185}$$

Nous posons $z = \sigma_{\ell_1}(y) = R_O(-2\alpha)y$. Soit la droite ℓ_3 passant par O et z. Vu que $R_O(2\alpha)z = y \in \ell_2$, l'angle de ℓ_3 à ℓ_2 est 2α. Par conséquent

$$\sigma_{\ell_2}(z) = R_O\big(-2 \times (-2\alpha)\big)z = R_O(4\alpha)z = R_O(4\alpha)R_O(-2\alpha)y = R_O(2\alpha)y. \tag{16.186}$$

Donc les transformations f et $R_O(2\alpha)$ coïncident pour tous les points des droites ℓ_1 et ℓ_2, qui ne sont pas parallèles. Cela prouve que $f = R_O(2\alpha)$.

Trois réflexions Nous écrivons $f = \sigma_{\ell_3} \circ \sigma_{\ell_2} \circ \sigma_{\ell_1}$. Nous allons transformer cela progressivement en une symétrie glissée en passant par plusieurs étapes :

(1) $f = \sigma_\ell \circ \tau_v$,

(2) $f = \tau_v \circ \sigma_\ell$,

(3) $f = \tau_v \circ \sigma_\ell$ avec $v \parallel \ell$.

À chacune de ces étapes, v et ℓ vont changer. La dernière est une réflexion glissée.

Nous commençons par supposer $\ell_2 \parallel \ell_3$. Dans ce cas, $\sigma_{\ell_3} \circ \sigma_{\ell_2}$ est une translation, comme nous l'avons déjà vu. Alors $f = \tau_v \circ \sigma_{\ell_1}$ et nous sommes déjà dans le cas (2).

Nous supposons que ℓ_2 n'est pas parallèle à ℓ_3. Dans ce cas, si $O = \ell_2 \cap \ell_3$ nous avons

$$\sigma_{\ell_3} \circ \sigma_{\ell_2} = R_O(2\alpha) \tag{16.187}$$

où α est l'angle de ℓ_2 à ℓ_3. En réalité tant que l'angle de ℓ_3' à ℓ_2' est α nous avons

$$\sigma_{\ell_3'} \circ \sigma_{\ell_2'} = \sigma_{\ell_3} \circ \sigma_{\ell_2} = R_O(2\alpha). \tag{16.188}$$

Nous choisissons ℓ_2' parallèle à ℓ_1, de telle sorte à ce que $\sigma_{\ell_2'} \circ \sigma_{\ell_1}$ soit une translation. Alors nous avons

$$f = \sigma_{\ell_3} \circ \sigma_{\ell_2} \circ \sigma_{\ell_1} = \sigma_{\ell_3} \circ \sigma_{\ell_2'} \circ \sigma_{\ell_1'} = \sigma_{\ell_3} \circ \tau_v. \tag{16.189}$$

où v est le vecteur de la translation en question.

Nous avons donc prouvé que toute composition de trois réflexions peut être écrite soit sous la forme (1) soit sous la forme (2).

Nous prouvons à présent que toute transformation de la forme (1) peut être écrite sous la forme (2). Plus précisément nous allons prouver que si ℓ est une droite, v un vecteur et ℓ_0 la droite parallèle à ℓ passant par l'origine, alors

$$\sigma_\ell \circ \tau_v = \tau_{\sigma_{\ell_0}(v)} \circ \sigma_l \tag{16.190}$$

D'abord nous savons que $\sigma_\ell(x) = \sigma_{\ell_0}(x) + 2w$ où w est le vecteur tel que $\ell = \ell_0 + w$. Ensuite c'est un simple calcul utilisant le fait que σ_{ℓ_0} est linéaire :

$$(\sigma_\ell \circ \tau_v)(x) = \sigma_l(x + v) = \sigma_{\ell_0}(x) + \sigma_{\ell_0}(v) + 2w, \tag{16.191}$$

et

$$(\tau_{\sigma_{\ell_0}(v)} \circ \sigma_\ell)(x) = \sigma_{\ell_0}(v) + \sigma_\ell(x) = \sigma_{\ell_0}(v) + \sigma_{\ell_0}(x) + 2w. \tag{16.192}$$

L'égalité est faite.

Nous montrons maintenant que toute transformation de la forme (2) peut être mise sous la forme (3). Soit donc $f = \tau_v \circ \sigma_\ell$ où v et ℓ ne sont pas spécialement parallèles.

Pour cela nous décomposons $v = v_1 + v_2$ avec $v_1 \perp \ell$ et $v_2 \parallel \ell$ et nous posons $\ell' = \ell + \frac{1}{2}v_1$. Nous montrons que

— $\tau_v \circ \sigma_\ell = \tau_{v_2} \circ \sigma_{\ell'}$

— $v_2 \parallel \ell'$.

Pour le deuxième point, $v_2 \parallel \ell$ et bien entendu $\ell' \parallel \ell$. Donc $v_2 \parallel \ell'$.

Soit ℓ_0 la droite parallèle à ℓ et ℓ' et passant par l'origine. Soit aussi le vecteur w tel que $\ell = \ell_0 + w$. Alors nous avons

$$\begin{cases} \sigma_\ell = \sigma_{\ell_0} + 2w & \text{(16.193a)} \\ \sigma_{\ell'} = \sigma_{\ell_0} + 2w + v_1 & \text{(16.193b)} \end{cases}$$

Nous avons

$$(\tau_v \circ \sigma_\ell)(x) = v + \sigma_{\ell_0}(x) + 2w \tag{16.194}$$

et

$$\begin{aligned} (\tau_{v_2} \circ \sigma_{\ell'})(x) &= v_2 + \sigma_{\ell_0}(x) + 2w + v_1 && \text{(16.195a)} \\ &= \sigma_{\ell_0}(x) + v + 2w && \text{(16.195b)} \end{aligned}$$

où dans la dernière ligne, nous avons regroupé $v_1 + v_2 = v$. Et voila.

□

Corollaire 16.81 ([1])**.**
Les rotations vectorielles forment exactement le groupe SO(2)*.*

Démonstration. Nous savons déjà de la proposition 16.71 que les rotations (composées de 2 réflexions, par la définition 16.58) sont toutes des applications de SO(2).

Pour l'autre sens, une application de SO(2) est une isométrie et est donc, par le théorème 16.80 une des choses suivantes :

— identité
— réflexion
— translation
— réflexion glissée
— rotation (composée de réflexions d'axes non parallèles).

L'identité est possible et c'est une rotation. La réflexion est impossible parce qu'une réflexion linéaire (non affine) est de déterminant -1 (lemme 16.70). Une translation c'est impossible parce que les éléments de SO(2) laissent invariant un point alors que les translations ne laissent aucun point invariant.

Quid des réflexions glissées ? Vue la forme (16.129) de la réflexion, la réflexion glissée $\tau_v \circ \sigma_\ell$ a pour formule

$$\alpha(x) = (\tau_v \circ \sigma_\ell)(x) = v + 2\,\mathtt{proj}_\ell(x) - x. \tag{16.196}$$

Nous allons étudier les conditions nécessaires pour que cela décrive un élément de SO(2), et nous allons voir que c'est impossible.

D'abord si ℓ passe par l'origine, alors $\alpha(0) = v$ et nous devons avoir $v = 0$. Dans ce cas $\alpha = \sigma_\ell$, ce qui n'est pas possible parce que α doit avoir un déterminant 1 alors que la réflexion a un déterminant -1.

Donc ℓ ne passe pas par l'origine. La condition $\alpha(0) = 0$ impose $v = -2\,\mathtt{proj}_\ell(0)$. Alors

$$\alpha(x) = -2\,\mathtt{proj}_\ell(0) + 2\,\mathtt{proj}_\ell(x) - x. \tag{16.197}$$

Le calcul de $\alpha\big(\mathtt{proj}_\ell(0)\big)$ donne :

$$\alpha\big(\mathtt{proj}_\ell(0)\big) = -2\,\mathtt{proj}_\ell(0) + 2\,\mathtt{proj}_\ell\big(\mathtt{proj}_\ell(0)\big) - \mathtt{proj}_\ell(0). \tag{16.198}$$

Vu que la projection de la projection est la projection, cela se réduit à

$$\alpha\big(\mathtt{proj}_\ell(0)\big) = -\,\mathtt{proj}_\ell(0). \tag{16.199}$$

Le vecteur $\mathtt{proj}_\ell(0)$ est donc renversé par l'action de α.

Considérons une base orthonormale $\{e_1, e_2\}$ de E où e_1 est un multiple de norme 1 de $\mathtt{proj}_\ell(0)$. Le déterminant de α pour cette base est [11] :

$$\det(\alpha) = \langle e_1, \alpha(e_1)\rangle\langle e_2, \alpha(e_2)\rangle - \langle e_1, \alpha(e_2)\rangle\langle e_2, \alpha(e_1)\rangle. \tag{16.200}$$

Là dedans nous avons $\langle e_1, \alpha(e_1)\rangle = -1$ et $\langle e_2, \alpha(e_1)\rangle = 0$, et le tout doit faire 1. Donc

$$\langle e_2, \alpha(e_2)\rangle = -1, \tag{16.201}$$

ce qui implique que $\alpha(e_2) = -e_2$. En particulier $\alpha = -\mathrm{Id}$. Nous avons prouvé jusqu'ici que si α est une réflexion glissée dans SO(2) alors $\alpha = -\mathrm{Id}$.

Même pas besoin de se poser de grandes questions sur savoir si il est possible de construire une symétrie glissée égale à $-\mathrm{Id}$ (c'est pas possible), il suffit de dire que de toutes façons, $-\mathrm{Id}$ est tout de même une rotation (lemme 16.73).

Il reste les rotations. Tout élément de SO(2) est une rotation. □

16.4 Isométries dans $\mathbb{R}^n$

Définition 16.82.
*Un **hyperplan** de $\mathbb{R}^n$ est un sous-espace affine de dimension $n-1$.*

Lemme-définition 16.83.
Si un hyperplan H de $\mathbb{R}^n$ est donné, et si $x \in \mathbb{R}^n$, il existe un unique point $y \in \mathbb{R}^n$ tel que

(1) $x - y \perp H$,

11. La définition 9.6 avec la formule (9.17).

(2) Le segment $[x, y]$ coupe H en son milieu.

*La **réflexion** σ_H est l'application $\sigma_H : \mathbb{R}^n \to \mathbb{R}^n$ qui à x fait correspondre ce y.*

Démonstration. Il faut vérifier que les conditions données définissent effectivement un unique point de $\mathbb{R}^n$. Soit H_0 le sous-espace vectoriel parallèle à H et une base orthonormée $\{e_1, \dots, e_{n-1}\}$ de H_0. Nous complétons cela en une base orthonormée de $\mathbb{R}^n$ avec un vecteur e_n. Si $H = H_0 + v$, quitte à décomposer v en une partie parallèle et une partie perpendiculaire à H, nous avons

$$H = H_0 + \lambda e_n \tag{16.202}$$

pour un certain λ.

Une droite passant par x et perpendiculaire à H est de la forme $t \mapsto x + te_n$. Si $x = \sum_{i=1}^n x_i e_i$ alors l'unique point de cette droit à être dans H est le point tel que $x_n e_n + te_n = \lambda e_n$, c'est à dire $t = -x_n$. L'unique point y sur cette droite à être tel que $[x, y]$ coupe H en son milieu est celui qui correspond à $t = -2x_n$. □

Notons au passage que cette preuve donne une formule pour σ_H :

$$\sigma_H(x) = \sum_{i=1}^{n-1} x_i e_i - x_n e_n. \tag{16.203}$$

Il s'agit donc de changer le signe de la composante perpendiculaire à H.

Lemme 16.84.
Dans cette même base si H_0 est l'hyperplan parallèle à H et passant par l'origine, nous écrivons $H = H_0 + \lambda e_n$ pour un certain λ. Alors

$$\sigma_H = \sigma_{H_0} + 2\lambda e_n. \tag{16.204}$$

Démonstration. Un élément $x \in \mathbb{R}^n$ peut être décomposé dans la base adéquate en $x = x_H + x_n e_n$. Nous savons de la formule (16.203) que

$$\sigma_H(x) = x_H - x_n e_n. \tag{16.205}$$

Mais vu que $\sigma_{H_0}(x_H) = x_H - 2\lambda e_n$ nous avons

$$\sigma_{H_0}(x) + 2\lambda e_n = \sigma_{H_0}(x_H + x_n e_n) + 2\lambda e_N = x_H - 2\lambda e_n - x_n e_n + 2\lambda e_n = x_H - x_n e_n. \tag{16.206}$$

□

Le lemme suivant est une généralisation du fait que tous les points de la médiatrice d'un segment sont à égale distance des deux extrémités du segment (très utile lorsqu'on étudie les triangles isocèles).

Lemme 16.85 ([195])**.**
Soient deux points distincts $x_0, y_0 \in \mathbb{R}^n$ l'ensemble $H \subset \mathbb{R}^n$ donné par

$$H = \{x \in \mathbb{R}^n \text{ tel que } d(x, x_0) = d(x, y_0)\}. \tag{16.207}$$

Alors H est l'hyperplan orthogonal au vecteur $v = y_0 - x_0$ et H passe par le milieu du segment $[x_0, y_0]$.

Démonstration. Nous savons que

$$d(x, x_0)^2 = \langle x - x_0, x - x_0 \rangle = \|x\|^2 + \|x_0\|^2 - 2\langle x, x_0 \rangle, \tag{16.208}$$

ou encore

$$\|x_0\|^2 - \|y_0\|^2 = 2\langle x, x_0 - y_0 \rangle. \tag{16.209}$$

En posant $v = y_0 - x_0$ et en considérant la forme linéaire

$$\begin{aligned}\beta\colon \mathbb{R}^n &\to \mathbb{R}\\ x &\mapsto \langle x, v\rangle,\end{aligned} \tag{16.210}$$

Nous avons $x \in H$ si et seulement si $\beta(x) = \frac{1}{2}\big(\|y_0\|^2 - \|x_0\|^2\big) = \lambda$. En d'autres termes, $H = \beta^{-1}(\lambda)$. Par la proposition 11.22 la partie H est un sous-espace affine. C'est même un translaté de $\ker(\beta)$, et comme $\ker(\beta)$ est l'espace vectoriel des vecteurs perpendiculaires à v, nous avons $\dim(H) = \dim\big(\ker(\beta)\big) = n - 1$.

Le fait que H contienne le milieu du segment $[x_0, y_0]$ est par définition. □

Pour le lemme suivant, et pour que la récurrence se passe bien nous disons que l'ensemble vide est un espace vectoriel de dimension -1.

Lemme 16.86.
Si $f \in \mathrm{Isom}(\mathbb{R}^n)$ satisfait

$$\dim\big(\mathrm{Fix}(f)\big) = n - k \tag{16.211}$$

alors f peut être écrit comme composition d'au plus k réflexions hyperplanes.

Démonstration. Nous faisons une récurrence sur $k \geqslant 0$.

Pour l'initialisation, si $k = 0$ alors $\dim\big(\mathrm{Fix}(f)\big) = n$, c'est à dire que f fixe tout $\mathbb{R}^n$, autant dire que f est l'identité, une composition de zéro réflexions.

Pour la récurrence, nous supposons que le lemme est démontré jusqu'à $k \geqslant 0$. Soit donc $f \in \mathrm{Isom}(\mathbb{R}^n)$ tel que

$$\dim\big(\mathrm{Fix}(f)\big) = n - (k+1). \tag{16.212}$$

Vu que $k \geqslant 0$, la dimension de $\mathrm{Fix}(f)$ est strictement plus petite que n, donc il existe un $x_0 \in \mathbb{R}^n$ tel que $f(x_0) \neq x_0$. Nous posons

$$H = \{x \in \mathbb{R}^n \text{ tel que } d(x, x_0) = d\big(x, f(x_0)\big)\}. \tag{16.213}$$

Par le lemme 16.85, ce H est l'hyperplan orthogonal à $v = f(x_0) - x_0$ et passant par le milieu du segment $[x_0, f(x_0)]$.

Nous posons $g = \sigma_H \circ f$. Vu que $g(x_0) = \sigma_H(f(x_0)) = x_0$, ce x_0 est un point fixe de g. Le fait que $\sigma_H\big(f(x_0)\big) = x_0$ est vraiment la définition de l'hyperplan H.

Nous avons donc

$$x_0 \in \mathrm{Fix}(g) \backslash \mathrm{Fix}(f). \tag{16.214}$$

Mais nous prouvons de plus que $\mathrm{Fix}(f) \subset \mathrm{Fix}(g)$. En effet si $y \in Fix(f)$ alors $y \in H$ parce que

$$d(y, x_0) = d\big(f(y), f(x_0)\big) = d\big(y, f(x_0)\big). \tag{16.215}$$

Vu que $y \in H$ nous avons $y \in \mathrm{Fix}(g)$ parce que

$$g(y) = \sigma_H\big(f(y)\big) = \sigma_H(y) = y. \tag{16.216}$$

Tout cela pour dire que l'ensemble $\mathrm{Fix}(g)$ est *strictement* plus grand que $\mathrm{Fix}(f)$. Et comme ce sont des espaces affines nous pouvons parler de dimension :

$$\dim\big(\mathrm{Fix}(g)\big) > \dim\big(\mathrm{Fix}(f)\big). \tag{16.217}$$

Par hypothèse de récurrence, l'application g peut être écrite comme composition de k réflexions. Donc l'application

$$f = \sigma_H \circ g \tag{16.218}$$

est une composition de $k + 1$ réflexions. □

Lemme 16.87.
Soit un hyperplan H et un vecteur v de $\mathbb{R}^n$. Nous avons

$$\tau_v \circ \sigma_H \circ \tau_v^{-1} = \sigma_{\tau_v(H)}. \tag{16.219}$$

Démonstration. Pour ce faire nous considérons une base adaptée. Les vecteurs $\{e_1, \ldots, e_{n-1}\}$ forment une base orthonormée de H_0 et e_n complète en une base orthonormée de $\mathbb{R}^n$. Soit H_0 l'hyperplan parallèle à H et passant par l'origine ; nous avons, pour un certain $\lambda \in \mathbb{R}$,

$$H = H_0 + \lambda e_n \tag{16.220}$$

D'un autre côté, le vecteur v peut être décomposé en $v = v_1 + v_2$ où $v_1 \perp H$ et $v_2 \parallel H$. Alors

$$\tau_v(H) = H + v = H + v_2 = H_0 + \lambda e_n + v_2. \tag{16.221}$$

Nous pouvons maintenant utiliser le lemme 16.84 pour exprimer la transformation $\sigma_{\tau_v(H)}$:

$$\sigma_{\tau_v(H)}(x) = \sigma_{H_0}(x) + 2\lambda e_n + 2v_2 \tag{16.222}$$

Mais d'autre part,

$$(\tau_v \circ \sigma_H \circ \tau_v^{-1})(x) = v + \sigma_H(x - v) = v + \sigma_{H_0}(x - v) + 2\lambda e_n. \tag{16.223}$$

Vue la décomposition de $v = v_1 + v_2$ nous avons $\sigma_{H_0}(v) = -v_1 + v_2$ et donc

$$(\tau_v \circ \sigma_H \circ \tau_v^{-1})(x) = v + \sigma_{H_0}(x) + v_1 - v_2 + 2\lambda e_n = \sigma_{H_0} + 2v_1 + 2\lambda e_n. \tag{16.224}$$

Les expressions (16.222) et (16.224) coïncident, d'où l'égalité recherchée. □

Théorème 16.88 ([195])**.**
Toute isométrie de $\mathbb{R}^n$ peut être écrite comme composition d'au plus $n + 1$ réflexions.

Une isométrie de $\mathbb{R}^n$ préserve l'orientation si et seulement si est elle composition d'un nombre pair de réflexions.

Démonstration. La première partie de ce théorème n'est rien d'autre que le lemme 16.86 parce que le pire cas est celui où le fixateur de f est réduit à l'ensemble vide, et dans ce cas l'application f est une composition de $n + 1$ réflexions.

Pour la seconde partie nous définissons

$$\begin{aligned} \epsilon\colon \mathrm{Isom}(\mathbb{R}^n) &\to \{\pm 1\} \\ \tau_v \circ \alpha &\mapsto \det(\alpha) \end{aligned} \tag{16.225}$$

où nous nous référons à la décomposition unique d'un élément de $\mathrm{Isom}(\mathbb{R}^n)$ sous la forme $\tau_v \circ \alpha$ avec $\alpha \in O(n)$ donnée par le théorème 16.6(3).

Le noyau de ϵ est alors la partie

$$\ker(\epsilon) = \mathbb{R}^n \times_{\mathbf{Ad}} \mathrm{SO}(n). \tag{16.226}$$

Une isométrie f préserve l'orientation si et seulement si $\epsilon(f) = 1$. Vu que toutes les isométries sont des composition de réflexions (première partie), il nous suffit de montrer que $\epsilon(\epsilon_H) = -1$ pour qu'une isométrie préserve l'orientation si et seulement si elle est composition d'un nombre pair de réflexions.

Nous commençons par prouver que pour tout vecteur v, $\epsilon\big(\sigma_H\big) = \epsilon\big(\sigma_{\tau_v(H)}\big)$. Pour cela nous utilisons le lemme 16.87 et le fait que ϵ est un homomorphisme :

$$\epsilon(\sigma_{\tau_v(H)}) = \epsilon(\tau_v)\epsilon(\sigma_H)\epsilon(\tau_v^{-1}) = \epsilon(\sigma_H) \tag{16.227}$$

parce que la partie linéaire d'une translation est l'identité (et donc $\epsilon(\tau_v) = 1$ pour tout v).

Nous avons donc $\epsilon(\sigma_H) = \epsilon(\sigma_{H_0})$. En ce qui concerne σ_{H_0}, dans la base adaptée la matrice est

$$\sigma_{H_0} = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & -1 \end{pmatrix}, \tag{16.228}$$

dont le déterminant est -1. □

Pour en savoir plus sur le groupe des isométries, il faut lire le théorème de Cartan-Dieudonné dans [196].

16.4.1 Groupes finis d'isomorphismes

Définition 16.89.
*Si X est une partie finie de $\mathbb{R}^n$, le **barycentre** de X est le point*

$$B_X = \frac{1}{|X|} \sum_{x \in X} x \tag{16.229}$$

où $|X|$ est le cardinal de X.

Cela est à mettre en relation avec la définition dans le cadre affine 11.24.

Lemme 16.90 ([195])**.**
Soit une partie finie X de $\mathbb{R}^n$ et une application affine $f \in \mathrm{Aff}(\mathbb{R}^n)$. Alors

$$f(B_X) = B_{f(X)}. \tag{16.230}$$

Démonstration. Nous savons que toute application affine est une composée de translation et d'une application linéaire : $f = \tau_v \circ g$ avec $v \in \mathbb{R}^n$ et $g \in \mathrm{GL}(n, \mathbb{R})$. Nous vérifions le résultat séparément pour τ_v et pour g.

D'une part,

$$B_{\tau_v(X)} = \frac{1}{|\tau_v(X)|} \sum_{y \in \tau_v(X)} y = \frac{1}{|X|} \sum_{x \in X} (x + v) = B_x + \frac{1}{|X|} \sum_{x \in X} v = B_x + v = \tau_v(B_X). \tag{16.231}$$

Nous avons utilisé le fait que X et $\tau_v(X)$ possèdent le même nombre d'éléments, ainsi que le fait d'avoir une somme de $|X|$ termes tous égaux à v.

D'autre part,

$$B_{g(X)} = \frac{1}{|X|} \sum_{x \in X} g(x) = g\Big(\frac{1}{|X|} \sum_{x \in X} x\Big) = g(B_X) \tag{16.232}$$

où nous avons utilisé la linéarité de g dans tous ses retranchements. □

Proposition 16.91.
Points fixes d'un sous-groupe.

(1) Soit H un sous-groupe finie de $\mathrm{Isom}(\mathbb{R}^n)$. Alors il existe $v \in \mathbb{R}^n$ tel que $f(v) = v$ pour tout $f \in H$.

(2) Si H est un sous-groupe de $\mathrm{Isom}(\mathbb{R}^n)$ n'acceptant pas de points fixes, alors il est infini.

Démonstration. Le groupe H agit sur $\mathbb{R}^n$, et si $x \in \mathbb{R}^n$ nous pouvons considérer son orbite Hx, qui est une partie finie de $\mathbb{R}^n$. Considérons son barycentre

$$v = B_{Hx} \tag{16.233}$$

Soit $f \in H$. Alors $f(v) = f(B_{Hx}) = B_{f(Hx)} = B_{Hx} = v$, donc v est fixé par H.

La seconde affirmation n'est rien d'autre que la contraposée de la première. □

Proposition 16.92.
À propos de groupes finis d'isométries.

(1) Tout sous groupe finie de $\mathrm{Isom}(\mathbb{R}^n)$ *est isomorphe à un sous-groupe fini de* $\mathrm{O}(n)$.

(2) Tout sous-groupe fini de $\mathrm{Isom}^+(\mathbb{R}^n)$ *est isomorphe à un sous-groupe fini de* $\mathrm{SO}(n)$.

Démonstration. Soit H un sous-groupe fini de $\mathrm{Isom}(\mathbb{R}^n)$ et $v \in \mathbb{R}^n$ un élément fixé par H (comme garantit par la proposition 16.91). Nous posons

$$\begin{aligned}\phi\colon H &\to \mathrm{Isom}(\mathbb{R}^n)\\ f &\mapsto \tau_v^{-1} \circ f \circ \tau_v.\end{aligned} \tag{16.234}$$

ϕ **est un homomorphisme** Les opération du type $\phi = \mathbf{Ad}(\tau_v)$ sont toujours des homomorphismes.

ϕ **consiste à extraire la partie linéaire** Si $f = \tau_w \circ g$ alors

$$\phi(f)(x) = (\tau_{-v} \circ \tau_w \circ g \circ \tau_v)(x) \tag{16.235a}$$
$$= \tau_{w-v}(g(x) + g(v)) \tag{16.235b}$$
$$= g(x) + g(v) - v + w \tag{16.235c}$$

Mais $g(v) + w = f(v)$ et nous savons que $f(v) = v$. Donc il ne reste que $\phi(f)(x) = g(x)$.

ϕ **est injective** Si $f = \tau_w \circ g$ vérifie $\phi(f) = \mathrm{Id}$, il faut en particulier que $g = \mathrm{Id}$. Mais H est fini et ne peut donc pas contenir de translations non triviales. Donc $w = 0$ et $f = \mathrm{Id}$.

Donc ϕ est une injection à valeur dans les transformation linéaires de $\mathrm{Isom}(\mathbb{R}^n)$. Autrement dit, ϕ est un isomorphisme entre H et son image, laquelle image est dans $\mathrm{O}(n)$.

En ce qui concerne la seconde partie, si $f \in \mathrm{Isom}^+(\mathbb{R}^n)$, alors $\phi(f)$ y est aussi, tout en étant linéaire. Donc $\phi(f) \in \mathrm{SO}(n)$. □

L'extraction de la partie linéaire est injective ? Certe c'est prouvé, mais on peut se demander ce qu'il se passe si H contient deux éléments qui ont la même partie linéaire. Cela n'est pas possible parce si $f_1 = \tau_{w_1} \circ g$ et $f_2 = \tau_{w_2} \circ g$ sont dans H alors $f_1 f_2^{-1} = \tau_{w_1+w_2}$ est également dans H, ce qui n'est pas possible si H est fini.

Définition 16.93 (Groupe de symétrie d'une partie de $\mathbb{R}^n$[195])**.**
Si Y *est une partie de* $\mathbb{R}^n$*, nous définissons le* ***groupe des symétries*** *de* Y *par*

$$\mathrm{Sym}(Y) = \{f \in \mathrm{Isom}(\mathbb{R}^n) \text{ tel que } f(Y) = Y\}. \tag{16.236}$$

Nous définissons aussi le groupe des symétries propres de Y *par*

$$\mathrm{Sym}^+(Y) = \{f \in \mathrm{Isom}^+(\mathbb{R}^n) \text{ tel que } f(Y) = Y\}. \tag{16.237}$$

Théorème 16.94 ([195])**.**
Soit $Y \subset \mathbb{R}^2$ *tel que le groupe* $\mathrm{Sym}^+(Y)$ *soit fini d'ordre* n*. Alors c'est un groupe cyclique d'ordre* n*.*

Si $\mathrm{Sym}^+(Y)$ *est fini, alors* $\mathrm{Sym}(Y)$ *est soit cyclique d'ordre* n *soit isomorphe au groupe diédral d'ordre* $2n$*.*

Démonstration. Nous savons déjà par la proposition 16.92 que $\mathrm{Sym}^+(Y)$ est isomorphe à un sous-groupe H^+ d'ordre n de $\mathrm{SO}(2)$. Vérifions que ce groupe est cyclique. Si $n = 1$, c'est évident. Si $n \geqslant 2$ alors nous savons que H^+ est constitué de rotations d'angles dans $[0, 2\pi[$ et vu que c'est un ensemble fini, il possède une rotation d'angle minimal (à part zéro). Notons α_0 cet angle.

Nous montrons que H^+ est engendré par la rotation d'angle α_0. Soit une rotation d'angle α. Étant donné que $\alpha_0 < \alpha$ nous pouvons effectuer la division euclidienne[12] de α par α_0 et obtenir

$$\alpha = k\alpha_0 + \beta \tag{16.238}$$

12. Théorème 3.25.

avec $\beta < \alpha_0$. Mézalors $R(\beta) = R(\alpha)R(\alpha_0)^{-k}$ est également un élément du groupe. Cela contredit la minimalité dès que $\beta \neq 0$. Avoir $\beta = 0$ revient à dire que α est un multiple de α_0, ce qui signifie que le groupe H^+ est cyclique engendré par α_0.

Notons au passage que nous avons automatiquement $\alpha_0 = \frac{2\pi}{n}$ parce qu'il faut $R(\alpha_0)^n = \text{Id}$. Nous avons prouvé que $\text{Sym}^+(Y)$ est cyclique d'ordre n.

Nous étudions maintenant le groupe $\text{Sym}(Y)$. Par la proposition 16.92 nous avons un homomorphisme injectif

$$\phi\colon \text{Sym}(Y) \to \text{O}(2), \tag{16.239}$$

et en posant $H = \phi\big(\text{Sym}(Y)\big)$ nous avons un isomorphisme de groupes $\phi\colon \text{Sym}(Y) \to H$. Nous savons aussi que ce ϕ se restreint en

$$\phi\colon \text{Sym}^+(Y) \to H^+ \subset \text{SO}(2) \tag{16.240}$$

où $H^+ = \phi\big(\text{Sym}^+(Y)\big) = H \cap \text{SO}(2)$. Le groupe H^+ est cyclique et est engendré par la rotation $R(2\pi/n)$.

Supposons un instant que $H \subset \text{SO}(2)$. Alors nous avons $H = H^+$ et ϕ est un isomorphisme entre $\text{Sym}(Y)$ et le groupe cyclique engendré par $R(2\pi/n)$.

Nous supposons à présent que H n'est pas un sous-ensemble de $\text{SO}(2)$. Quelles sont les isométries de $\mathbb{R}^2$ qui ne sont pas de déterminant 1 ? Il faut regarder dans le théorème 16.80 quelles sont les isométries contenant un nombre impair de réflexions. Ce sont les réflexions et les réflexions glissées. Or il ne peut pas y avoir de réflexions glissées dans un groupe fini parce que si f est une réflexion glissée, tous les f^k sont différents.

Nous en déduisons que si H n'est pas inclus à $\text{SO}(2)$, il contient une réflexion que nous nommons σ. Nous allons en déduire que $H \simeq H^+ \times_{\mathbf{Ad}} C_2$ où $C_2 = \{\text{Id}, \sigma\}$. Si $h \in H$ nous pouvons écrire

$$h = (h\sigma^\epsilon)\sigma^\epsilon \tag{16.241}$$

pour n'importe quelle valeur de ϵ, et en particulier pour $\epsilon = \pm 1$.

Si $h \in \text{SO}(2)$ alors nous écrivons $h = h\epsilon^0$ et si $h \notin \text{SO}(2)$ nous écrivons $h = (h\sigma)\sigma$. Vu que $h\sigma \in \text{SO}(2)$, cette dernière écriture est encore de la forme $\text{SO}(2) \times C_2$. Quoi qu'il en soit tout élément de H s'écrit comme un produit

$$H = H^+C_2. \tag{16.242}$$

Cette décomposition est unique parce que si $h_1c_1 = h_2c_2$ alors $h_2^{-1}h_1 = c_2c_1^{-1}$, et comme $h_2^{-1}h_1 \in H^+$ nous avons $c_2c_1^{-1} \in H^+$ et donc $c_1 = c_2$. Partant nous avons aussi $h_1 = h_2$. Pour avoir le produit semi-direct il faut encore montrer que $\mathbf{Ad}(C_2)H^+ \subset H^+$. Le seul cas à vérifier est $\mathbf{Ad}(\sigma)H^+ \subset H^+$. Vu que les éléments de H^+ sont caractérisés par le fait d'avoir un déterminant positif, nous avons

$$\mathbf{Ad}(\sigma)R(\alpha) = \sigma R(\alpha)\sigma^{-1} \in H^+. \tag{16.243}$$

□

Remarque 16.95.
Tout ceci est cohérent avec le théorème de Burnside 9.238 parce que le sous-groupe fini de $\text{SO}(n)$ engendré par la rotation $R(2\pi/n)$ est un groupe d'exposant fini, à savoir que si h est dans ce groupe, $h^n = \text{Id}$.

16.4.2 Théorème de Sylvester

Théorème 16.96 (de Sylvester)**.**
Soit Q une forme quadratique réelle de signature (p, q). Alors pour toute base orthonormée on a

$$p = \text{Card}\{i \text{ tel que } Q(e_i) > 0\} \tag{16.244a}$$
$$q = \text{Card}\{i \text{ tel que } Q(e_i) < 0\}. \tag{16.244b}$$

Le rang de Q est $p+q$.

Si A est la matrice de Q dans une base, alors il existe une matrice inversible P telle que

$$P^tAP = \begin{pmatrix} -\mathbb{1}_q & & \\ & \mathbb{1}_p & \\ & & 0 \end{pmatrix}. \tag{16.245}$$

16.4.3 Groupe diédral

16.4.3.1 Définition et générateurs : vue géométrique

Définition 16.97.
*Le **groupe diédral** D_n est le groupe des isométries de $\mathbb{R}^2$ laissant invariant un polygone régulier à n côtés.*

Le groupe diédral peut être vu comme le stabilisateur de l'ensemble

$$\{e^{2ik\pi/n}, k = 0, \dots, n-1\} \tag{16.246}$$

dans le groupe des isométries affines de $\mathbb{C}^*$.

Si $f \in D_n$, alors $f(e^{2ik\pi/n})$ doit être l'un des $e^{2ik'\pi/n}$, et vu que f conserve les longueurs dans $\mathbb{C}$, nous devons avoir

$$1 = d(0, e^{2ik\pi/n}) = d\big(f(0), e^{2ik'\pi/n}\big). \tag{16.247}$$

Donc $f(0)$ est à l'intersection de tous les cercles de rayon 1 centrés en les $e^{2ik\pi/n}$, ce qui montre que $f(0))0$ (dès que $n \geqslant 3$). Par conséquent notre étude du groupe diédral ne doit prendre en compte que les isométries vectorielles de $\mathbb{R}^2$. En d'autres termes

$$D_n \subset O(2, \mathbb{R}). \tag{16.248}$$

Proposition 16.98 ([201])**.**
Le groupe D_n contient un sous groupe cyclique d'ordre 2 et un sous groupe cyclique d'ordre n.

Démonstration. Si s est la réflexion d'axe $\mathbb{R}$, alors s est d'ordre 2. De plus s est bien dans D_n parce que

$$s\big(e^{2ki\pi/n}\big) = e^{2(n-k)i\pi/n}. \tag{16.249}$$

De la même façon, la rotations d'angle $2\pi/n$, que l'on note r, agit sur les racines de l'unité et engendre un le groupe d'ordre n des rotations d'angle $2k\pi/n$. □

Notons que la conjugaison complexe ne fait pas spécialement partie du groupe D_n. En effet pour $n = 3$ par exemple les points fixes sont $A_1 = (1, 0)$, $A_2 = (-\frac{1}{2}, \frac{\sqrt{3}}{2})$ et $A_3 = (\frac{1}{2}, -\frac{\sqrt{3}}{2})$. La conjugaison complexe envoie évidemment A_1 sur A_1, mais pas du tout A_2 sur A_3.

Proposition 16.99 ([201])**.**
Nous avons $(sr)^2 = \mathrm{Id}$.

Démonstration. Si $z^n = 1$, alors

$$(srsr)z = srse^{2i\pi/n}z = sr\big(e^{-2\pi i/n}\bar{z}\big) = s\bar{z} = z. \tag{16.250}$$

□

Proposition 16.100 ([201])**.**
Le groupe diédral D_n est engendré par s et r. De plus tous les éléments de D_n s'écrivent sous la forme $s \circ r^m$.

Démonstration. Nous considérons les points $A_0 = 1$ et $A_k = e^{2ki\pi/n}$ avec $k \in \{1, \dots, n-1\}$. Par convention, $A_n = A_0$. L'action des éléments s et r sur ces points est

$$r(A_k) = A_{k+1} \tag{16.251a}$$
$$s(A_k) = A_{n-k}. \tag{16.251b}$$

Cette dernière est l'équation (16.249).

Soit $f \in D_n$. Étant donné que c'est une isométrie de $\mathbb{R}^2$ avec un point fixe (le point 0), f est soit une rotation soit une réflexion.

Supposons pour commencer que un des A_k est fixé par f. Dans ce cas f a deux points fixes : O et A_k et est donc la réflexion d'axe (OA_k). Dans ce cas, nous avons $f = s \circ r^{n-2k}$. En effet

$$s \circ r^{n-2k}(A_k) = s(A_{k+n-2k}) = s(A_{n-k}) = A_k. \tag{16.252}$$

Donc O et A_k sont deux points fixes de l'isométrie f ; donc f est bien la réflexion sur le bon axe.

Nous passons à présent au cas où f ne fixe aucun des A_k.

(1) Supposons que f soit une rotation. Si $f(A_k) = A_m$, alors l'angle de la rotation est

$$\frac{2(m-k)\pi}{n}, \tag{16.253}$$

et donc $f = r^{m-k}$, qui est de la forme demandée.

(2) Supposons à présent que f soit une réflexion d'axe Δ. Cette fois, Δ ne passe par aucun des points A_k, par contre Δ passe par 0. Nous commençons par montrer que Δ doit être la médiatrice d'un des côtés $[A_p, A_{p+1}]$ du polygone. Vu que Δ passe par O et n'est aucune des droites (OA_k), cette droite passe par l'intérieur d'un des triangles OA_pA_{p+1} et intersecte donc le côté correspondant.

Notre tâche est de montrer que Δ coupe $[A_p, A_{p+1}]$ en son milieu. Dans ce cas, Δ sera automatiquement perpendiculaire parce que le triangle OA_pA_{p+1} est isocèle en O. Nommons l la longueur des côtés du polygone, $P = \Delta \cap [A_p, A_{p+1}]$, $x = d(A_p, P)$ et $\delta = d(A_p, \Delta)$. Vu que f est la symétrie d'axe Δ, nous avons aussi $d\big(f(A_p), \Delta\big) = \delta$ et $d\big(A_p, f(A_p)\big) = 2\delta$. D'autre part, par la définition de la distance, $\delta < x$. Si $x < \frac{l}{2}$, alors $\delta < \frac{\delta}{2}$ et donc $d\big(A_p, f(A_p)\big) < l$. Or cela est impossible parce que le polygone ne possède aucun sommet à distance plus courte que l de A_p.

De la même manière si $x > \frac{l}{2}$, nous raisonnons avec A_{p+1} pour obtenir une contradiction. Nous en concluons que la seule possibilité est $x = \frac{l}{2}$, et donc $f(A_p) = A_{p+1}$. Montrons alors que $f = s \circ r^{n-2p-1}$. Il faut montrer que c'est une réflexion qui envoie A_p sur A_{p+1}. D'abord c'est une réflexion parce que

$$\det(sr^{n-2p-1}) = \det(s)\det(r^{n-2p-1}) = -1 \tag{16.254}$$

parce que $\det(s) = -1$ alors que $\det(r^k) = 1$ parce que r est une rotation dans $\mathrm{SO}(2)$. Ensuite nous avons

$$s \circ r^{n-2p-1}(A_p) = s(A_{p+n-2p-1}) = s(A_{n-p-1}) = A_{n-(n-p-1)} = A_{p+1}. \tag{16.255}$$

Donc $s \circ r^{n-2p-1}$ est bien une réflexion qui envoie A_p sur A_{p+1}.

□

Corollaire 16.101.
La liste des éléments de D_n est

$$D_n = \{1, r, \dots, r^{n-1}, s, sr, \dots, sr^{n-1}\} \tag{16.256}$$

et $|D_n| = 2n$.

Démonstration. Nous savons par la proposition 16.100 que tous les élément de D_n s'écrivent sous la forme r^k ou sr^k. Vu que r est d'ordre n, il ne faut considérer que $k \in \{1, \ldots, n-1\}$. Les éléments 1, r,..., r^{n-1} sont tous différents, et sont (pour des raisons de déterminant) tous différents des sr^k. Les isométries sr^k sont toutes différentes entre elles pour essentiellement la même raison :

$$sr^k(A_p) = s(A_{p+k}) = A_{n-p+k} \tag{16.257}$$

donc si $k \neq k'$, $sr^k(A_p) \neq sr^{k'}(A_p)$. La liste des éléments de D_n est donc

$$D_n = \{1, r, \ldots, r^{n-1}, s, sr, \ldots, sr^{n-1}\} \tag{16.258}$$

et donc $|D_n| = 2n$. □

Exemple 16.102
Nous considérons le carré $ABCD$ dans $\mathbb{R}^2$ et nous cherchons les isométries de $\mathbb{R}^2$ qui laissent le carré invariant. Nous nommons les points comme sur la figure 16.7. La symétrie d'axe vertical est nommée s et la rotation de 90 degrés est notée r.

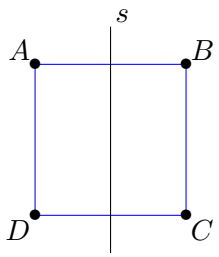


FIGURE 16.7 – Le carré dont nous étudions le groupe diédral.

Il est facile de vérifier que toutes les symétries axiales peuvent être écrites sous la forme $r^i s$. De plus le groupe engendré par s agit sur le groupe engendré par r parce que

$$(srs^{-1})(A, B, C, D) = sr(B, A, D, C) = s(A, D, C, B) = (B, C, D, A), \tag{16.259}$$

c'est à dire $srs^{-1} = r^{-1}$. Nous sommes alors dans le cadre du corollaire 3.95 et nous pouvons écrire que

$$D_4 = \mathrm{gr}(r) \times_\sigma \mathrm{gr}(s). \tag{16.260}$$

△

16.4.3.2 Générateurs : vue abstraite

16.103.
Nous allons montrer que D_n peut être décrit de façon abstraite en ne parlant que de ses générateurs. Nous considérons un groupe G engendré par des éléments a et b tels que

(1) a est d'ordre 2,
(2) b est d'ordre n avec $n \geqslant 3$,
(3) $abab = e$.

Nous allons prouver que ce groupe doit avoir la même liste d'éléments que celle du corollaire 16.101.

Proposition 16.104 ([201])**.**
Le groupe G n'est pas abélien.

Démonstration. Nous savons que $abab = e$, donc $aba = b^{-2}$, mais $b^{-2} \neq e$ parce que b est d'ordre $n > 2$. Donc $aba \neq e$. En manipulant un peu :

$$e \neq aba = (ab)(ba^{-1})^{-1} = (ab)(ba)^{-1} \tag{16.261}$$

parce que $a^{-1} = a$. Donc $ab \neq ba$. □

Lemme 16.105 ([201]).
Pour tout k entre 1 et $n-1$ nous avons

$$\boldsymbol{Ad}(a)b^k = ab^k a^{-1} = ab^k a = b^{-k}. \tag{16.262}$$

Démonstration. Nous faisons la démonstration par récurrence. D'abord pour $k = 1$, nous devons avoir $aba = b^{-1}$, ce qui est correct parce que par construction de G nous avons $abab = e$. Ensuite nous supposons que le lemme tient pour k et nous regardons ce qu'il se passe avec $k + 1$:

$$ab^{k+1}ba = ab^k ba = \underbrace{ab^k a}_{b^{-k}}\underbrace{aba}_{b^{-1}} = b^{-k}b^{-1} = b^{-(k+1)}. \tag{16.263}$$

□

Proposition 16.106.
L'élément a n'est pas une puissance de b.

Démonstration. Supposons le contraire : $a = b^k$. Dans ce cas nous aurions

$$e = (ab)(ab) = b^{k+1}b^{k+1} = b^{2k+2} = b^{2k}b^2 = a^2b^2 = b^2, \tag{16.264}$$

ce qui signifierait que b est d'ordre 2, ce qui est exclu par construction. □

Proposition 16.107 ([201]).
La liste des éléments de G est donnée par

$$G = \{1, b, \cdots, b^{n-1}, a, ab, \ldots, ab^{n-1}\} = \{a^\epsilon b^k\}_{\substack{\epsilon=0,1\\ k=0,\ldots,n-1}} \tag{16.265}$$

Les éléments de ces listes sont distincts.

Démonstration. Étant donné que a n'est pas une puissance de b, les éléments 1, a, b,..., b^{n-1} sont distincts. De plus si k et $m = k + p$ sont deux éléments distincts de $\{1, \ldots, n-1\}$, nous avons $ab^k \neq ab^m$ parce que si $ab^k = ab^{k+p}$, alors $a = ab^p$ avec $p < n$, ce qui est impossible. Pour la même raison, $ab^k \neq e$, et $ab^k \neq b^m$.

Au final les éléments $1, a, b, \ldots, b^{n-1}, ab, \ldots, ab^{n-1}$ sont tous différents. Nous devons encore voir qu'il n'y en a pas d'autres.

Par définition le groupe G est engendré par a et b, donc tout élément $x \in G$ s'écrit $x = a^{m_1}b^{k_1}\ldots a^{m_r}b^{k_r}$ pour un certain r et avec pour tout i, $k_i \in \{1, \ldots, n-1\}$ (sauf k_r qui peut être égal à zéro) et $m_i = 1$, sauf m_1 qui peut être égal à zéro. Donc

$$x = a^m b^{k_1} ab^{k_2} a \ldots b^{k_{r-1}} ab^{k_r} \tag{16.266}$$

où m et k_r peuvent éventuellement être zéro. En utilisant le lemme 16.105 sous la forme $b^{k_i}a = ab^{-k_i}$, quitte à changer les valeurs des exposants, nous pouvons passer tous les a à gauche et tous les b à droite pour finir sous la forme $x = a^k b^m$.

Donc non, il n'existe pas d'autres éléments dans G que ceux déjà listés. □

Lemme 16.108 ([1]).
Tout élément de G s'écrit de façon unique sous la forme $a^\epsilon b^k$ ou $b^k a^\epsilon$ avec $\epsilon = 0, 1$ et $k = 0, \ldots, n-1$.

Démonstration. Nous commençons par la forme $a^\epsilon b^k$. L'existence est la proposition 16.107. Pour l'unicité nous supposons $a^\epsilon b^k = a^\sigma b^l$ et nous décomposons en 4.

$\underline{\epsilon = 0, \sigma = 0}$ Alors $b^k = b^l$. Mais b étant d'ordre n et k, l étant égaux au maximum à $n-1$, cette égalité implique $k = l$.

$\underline{\epsilon = 0, \sigma = 1}$ Alors $b^k = ab^l$, ce qui donne $a = b^{k-l}$, ce qui est interdit par la proposition 16.106.

$\underline{\epsilon = 1, \sigma = 0}$ Même problème.

$\underline{\epsilon = 1, \sigma = 1}$ Encore une fois $b^k = b^l$ implique $k = l$.

En ce qui concerne la forme $b^k a^\epsilon$, l'existence est à montrer. Soit l'élément $g = a^\epsilon b^k$ et cherchons à le mettre sous la forme $b^l a^\sigma$. Si $\epsilon = 0$ c'est évident. Sinon $\epsilon = 1$ et nous avons par le lemme 16.105

$$ab^k = b^{-k}a^{-1} = b^{-k}b^n a = b^{-k}a. \tag{16.267}$$

En ce qui concerne l'unicité, nous refaisons 4 cas pour $b^k a^\epsilon = b^l a^\sigma$ comme précédemment et ils se traitement exactement comme précédemment. □

Théorème 16.109.
Les groupes G et D_n sont isomorphes.

Démonstration. Nous utilisons l'application

$$\begin{aligned} \psi\colon G &\to D_n \\ a^k b^m &\mapsto s^k r^m. \end{aligned} \tag{16.268}$$

C'est évidemment bien défini et bijectif, mais c'est également un homomorphisme parce que si nous calculons ψ sur un produit, nous devons comparer

$$\psi\big(a^{k_1}b^{m_1}a^{k_2}b^{m_2}\big) \tag{16.269}$$

avec

$$\psi\big(a^{k_1}b^{m_1}\big)\psi\big(a^{k_2}b^{m_2}\big) = s^{k_1}r^{m_1}s^{k_2}r^{m_2}. \tag{16.270}$$

Vu que D_n et G ont les mêmes propriétés qui permettent de permuter a et b ou s et r, l'expression à l'intérieur du ψ dans (16.269) se simplifie en $a^k b^m$ avec les même k et n que l'expression à droite dans (16.270) ne se simplifie en $s^k r^m$. □

Corollaire 16.110.
Toutes les propriétés démontrées pour G sont vraies pour D_n. En particulier, avec quelques redites :

1. *Le groupe D_n peut être défini comme étant le groupe engendré par un élément s d'ordre 2 et un élément r d'ordre $n-1$ assujettis à la relation $srsr = e$.*
2. *Le groupe D_n n'est pas abélien.*
3. *Pour tout $k \in \{1, \dots, n-1\}$ nous avons $sr^k s = r^{-k}$.*
4. *L'élément s ne peut pas être obtenu comme une puissance de r.*
5. *La liste des éléments de D_n est*
$$D_n = \{1, r, \dots, r^{n-1}, s, sr, \dots, sr^{n-1}\} \tag{16.271}$$
6. *Le groupe diédral D_n est d'ordre $2n$.*

Proposition 16.111.
En posant $C_n = \{r^k\}_{k=0,\dots,n-1}$ et $C_2 = \{a^\epsilon\}_{\epsilon=0,1}$, nous pouvons exprimer D_n comme le produit semi-direct

$$D_n = C_n \times_\rho C_2 \tag{16.272}$$

où ρ désigne l'action adjointe.

Démonstration. L'isomorphisme est :

$$\begin{aligned} \psi\colon C_n \times_\rho C_2 &\to D_n \\ (b^k, a^\epsilon) &\mapsto b^k a^\epsilon. \end{aligned} \tag{16.273}$$

Action adjointe L'application $\rho_{a^\epsilon} = \mathbf{Ad}(a^\epsilon)$ est toujours un homomorphisme. Vu que a^ϵ est soit e soit a, nous allons nous restreindre à a et oublier l'exposant ϵ. Il faut montrer que$\mathbf{Ad}(a) \in \mathrm{Aut}(C_n)$. En utilisant le lemme 16.105,

$$\mathbf{Ad}(a)b^k = ab^k a^{-1} = b^{-k} = b^{n-k}. \tag{16.274}$$

L'application $\mathbf{Ad}(a)\colon C_n \to C_n$ est donc bijective et homomorphique. Ergo isomorphisme.

Injectif Si $\psi(b^k, a^\epsilon) = \psi(b^l, a^\sigma)$, alors par unicité du lemme 16.108 nous avons $k = l$ et $\epsilon = \sigma$.

Surjectif Par la partie «existence» du lemme 16.108.

Homomorphisme L'homomorphisme est toujours de mise lorsque l'on prend deux sous-groupes d'un même groupe (ici le groupe des isométries de $\mathbb{R}^2$) et que l'on tente de faire un produit semi-direct en utilisant l'action adjointe. Dans notre cas, le calcul est :

$$\psi\big((b^k, a^\epsilon)(b^l, a^\sigma)\big) = b^k \rho_{a^\epsilon}(b^l) a^{\epsilon+\sigma} = b^k a^\epsilon b^l a^{-\epsilon} a^{\epsilon+\sigma} = b^k a^\epsilon b^l a^\sigma = \psi(b^k, a^\epsilon)\psi(b^l, a^\sigma). \tag{16.275}$$

□

16.4.3.3 Classes de conjugaison

Pour les classes de conjugaison du groupe diédral nous suivons [202].

D'abord pour des raisons de déterminants[13], les classes des éléments de la forme r^k et de la forme sr^k ne se mélangent pas. Nous notons $C(x)$ la classe de conjugaison de x, et $y \cdot x = yxy^{-1}$.

Les relations que nous allons utiliser sont

$$sr^k s = r^{-k} \tag{16.276a}$$

$$rs = sr^{-1} = sr^{n-1}. \tag{16.276b}$$

La classe de conjugaison qui ne rate jamais est bien entendu $C(1) = 1$. Nous commençons les vraies festivités $C(r^m)$. D'abord $r^k \cdot r^m = r^m$, ensuite

$$(sr^k) \cdot r^m = sr^k r^m r^{-k} s^{-1} = sr^m s^{-1} = r^{-m}. \tag{16.277}$$

Donc

$$C(r^m) = \{r^m, r^{-m}\}. \tag{16.278}$$

À ce niveau il faut faire deux remarques. D'abord si $m > \frac{n}{2}$, alors $C(r^m)$ est la classe de C^{n-m} avec $n - m < \frac{n}{2}$. Donc les classes que nous avons trouvées sont uniquement à lister avec $m < \frac{n}{2}$. Ensuite si $m = \frac{n}{2}$ alors $r^m = r^{-m}$ et la classe est un singleton. Cela n'arrive que si n est pair.

Nous passons ensuite à $C(s)$. Nous avons

$$r^k \cdot s = r^k s r^{-k} = s s r^k s r^{-k} = s r^{-k} r^{-k} = s r^{n-2k}, \tag{16.279}$$

et

$$(sr^k) \cdot s = \underbrace{sr^k s}_{r^{-k}} r^{-k} s^{-1} = r^{-2k} s = r^{n-2k} s = s r^{(n-1)(n-2k)} = s r^{n^2-2kn-n+2k} = s r^{2k}. \tag{16.280}$$

donc

$$C(s) = \{sr^{n-2k}, sr^{2k}\}_{k=0,\dots,n-1}. \tag{16.281}$$

Ici aussi l'écriture n'est pas optimale : peut-être que pour certains k il y a des doublons. Nous reportons l'écriture exacte à la discussion plus bas qui distinguera n pair de n impair. Notons juste que si n est pair, l'élément sr n'est pas dans la classe $C(s)$.

13. Vous notez qu'ici nous utilisons un argument qui utilise la définition de D_n comme isométries de $\mathbb{R}^2$. Si nous avions voulu à tout prix nous limiter à la définition «abstraite» en termes de générateurs, il aurait fallu trouver autre chose.

Nous en faisons donc à présent le calcul en gardant en tête le fait qu'il n'a de sens que si n est pair. D'abord

$$s \cdot (sr) = ssrs = rs = sr^{n-1}. \tag{16.282}$$

Ensuite

$$(sr^k) \cdot (sr) = sr^k srr^{-k}s = r^{-2k+1}s = sr^{2k-1}. \tag{16.283}$$

Avec $k = \frac{n}{2}$, cela rend $s \cdot (sr)$, donc pas besoin de le recopier. Nous avons

$$C(sr) = \{sr^{2k-1}\}_{k=1,\dots,n-1}. \tag{16.284}$$

16.4.3.4 Le compte pour n pair

Si n est pair, nous avons les classes

$$C(1) = \{1\} \qquad \text{1 élément} \tag{16.285a}$$

$$C(r^m) = \{r^m, r^{m-1}\} \qquad \text{pour } 0 < m < \frac{n}{2} \qquad \frac{n}{2} - 1 \text{ fois 2 éléments} \tag{16.285b}$$

$$C(r^{n/2}) = \{r^{n/2}\} \qquad \text{1 élément} \tag{16.285c}$$

$$C(s) = \{sr^{2k}\}_{k=0,\dots,\frac{n}{2}-1} \qquad \frac{n}{2} \text{ éléments} \tag{16.285d}$$

$$C(sr) = \{sr^{2k+1}\}_{k=0,\dots,\frac{n}{2}-1} \qquad \frac{n}{2} \text{ éléments.} \tag{16.285e}$$

Au total nous avons bien listé $2n$ éléments comme il se doit, dans $\frac{n}{2} + 3$ classes différentes.

16.4.3.5 Le compte pour n impair

Si n est impair, nous avons les classes

$$C(1) = \{1\} \qquad \text{1 élément} \tag{16.286a}$$

$$C(r^m) = \{r^m, r^{m-1}\} \qquad \text{pour } 0 < m < \frac{n-1}{2} \qquad \frac{n-1}{2} \text{ fois 2 éléments} \tag{16.286b}$$

$$C(s) = \{sr^k\}_{k=0,\dots,n-1} \qquad n \text{ éléments} \tag{16.286c}$$

Au total nous avons bien listé $2n$ éléments comme il se doit, dans $\frac{n+3}{2}$ classes différentes.

16.4.4 Applications : du dénombrement

16.4.4.1 Le jeu de la roulette

Soit une roulette à n secteurs que nous voulons colorier en q couleurs[203]. Nous voulons savoir le nombre de possibilités à rotations près. Soit d'abord E l'ensemble des coloriages possibles sans contraintes ; il y a naturellement q^n possibilités. Sur l'ensemble E, le groupe cyclique G des rotations d'angle $2\pi/n$ agit. Deux coloriages étant identiques si ils sont reliés par une rotation, la réponse à notre problème est donné par le nombre d'orbites de l'action de G sur E qui sera donnée par la formule du théorème de Burnside 3.85.

Nous devons calculer $\text{Card}\big(\text{Fix}(g)\big)$ pour tout $g \in G$. Soit g, un élément d'ordre d dans G. Si g agit sur la roulette, chaque secteur a une orbite contenant d éléments. Autrement dit, g divise la roulette en n/d secteurs. Un élément de E appartenant à $\text{Fix}(g)$ doit colorier ces n/d secteurs de façon uniforme ; il y a $q^{n/d}$ possibilités.

Il reste à déterminer le nombre d'éléments d'ordre d dans G. Un élément de G est donné par un nombre complexe de la forme $e^{2ik\pi/n}$. Les éléments d'ordre d sont les racines primitives[14] $d$ièmes de l'unité. Nous savons que –par définition– il y a $\varphi(d)$ telles racines primitives de l'unité. Bref il y a $\varphi(d)$ éléments d'ordre d dans G.

14. Une racine non primitive 8ième de l'unité est par exemple i. Certes $i^8 = 1$, mais $i^4 = 1$ aussi. Le nombre i est d'ordre 4.

La formule de Burnside nous donne maintenant le nombre d'orbites :

$$\frac{1}{n}\sum_{d|n}\varphi(d)q^{n/d}. \tag{16.287}$$

Cela est le nombre de coloriage possibles de la roulette à n secteurs avec q couleurs.

16.4.4.2 L'affaire du collier

Nous avons maintenant des perles de q couleurs différentes et nous voulons en faire un collier à n perles. Cette fois non seulement les rotations donnent des colliers équivalents, mais en outre les symétries axiales (il est possible de retourner un collier, mais pas une roulette). Le groupe agissant sur E est maintenant le groupe diédral [15] D_n conservant un polygone a n sommets.

Nous devons séparer le cas n impair du cas n pair.

Si n est impair, alors les axes de symétries passent par un sommet par le milieu du côté opposé. Le groupe D_n contient n symétries axiales. Nous avons donc maintenant

$$|G| = 2n. \tag{16.288}$$

Nous écrivons la formule de Burnside

$$\mathrm{Card}(\Omega) = \frac{1}{2n}\sum_{g\in G}\mathrm{Card}\big(\mathrm{Fix}(g)\big). \tag{16.289}$$

Si g est une rotation, le travail est déjà fait. Si g est une symétrie, nous avons le choix de la couleur du sommet par lequel passe l'axe et le choix de la couleur des $(n-1)/2$ paires de sommets. Cela fait

$$qq^{(n-1)/2} = q^{\frac{n+1}{2}} \tag{16.290}$$

possibilités. Nous avons donc

$$\mathrm{Card}(\Omega) = \frac{1}{2n}\left(\sum_{d|n}q^{n/d}\varphi(d) + nq^{\frac{n+1}{2}}\right). \tag{16.291}$$

Si n est pair, le choses se compliquent un tout petit peu. En plus de symétries axiales passant par un sommet et le milieu du côté opposé, il y a les axes passant par deux sommets opposés. Pour colorier un collier en tenant compte d'une telle symétrie, nous pouvons choisir la couleur des deux perles par lesquelles passe l'axe ainsi que la couleur des $(n-2)/2$ paires de perles. Cela fait en tout

$$q^2q^{\frac{n-2}{2}} = q^{\frac{n+2}{2}}. \tag{16.292}$$

Le groupe G contient $n/2$ tels axes.

Notons que cette fois G ne contient plus que $n/2$ symétries passant par un sommet et un côté. L'ordre de G est donc encore $2n$. La formule de Burnside donne

$$\mathrm{Card}(\Omega) = \frac{1}{2n}\left(\sum_{d\,|\,n}\varphi(d)q^{n/d} + \frac{n}{2}q^{(n+2)/2} + \frac{n}{2}q^{n/2}\right). \tag{16.293}$$

16.5 Angles et rotations dans $\mathbb{R}^2$

16.5.1 Angles orientés, rotations

Nous avons défini les rotations planes en 16.58, et montré que pour les rotations vectorielles, nous avions en réalité le groupe SO(2) en le corollaire 16.81. Il est temps de donner une forme matricielle aux rotations et de lier cela aux fonctions trigonométriques.

15. Définition 16.97.

Lemme 16.112.
Si $A \in \mathrm{O}(2)$ alors il existe un unique $\theta \in [0, 2\pi[$ et un unique $\epsilon = \pm 1$ tels que

$$A = \begin{pmatrix} \cos(\theta) & -\epsilon \sin(\theta) \\ \sin(\theta) & \epsilon \cos(\theta) \end{pmatrix} \tag{16.294}$$

Démonstration. Soit une matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ et imposons qu'elle soit dans $\mathrm{O}(2)$. Le fait que A soit orthogonale impose

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{16.295}$$

Nous avons alors le système

$$\begin{cases} a^2 + b^2 = 1 & \text{(16.296a)} \\ b^2 + d^2 = 1 & \text{(16.296b)} \\ ab + cd = 0 & \text{(16.296c)} \end{cases}$$

La proposition 16.29 nous permet de déduire qu'il existe un unique $\theta \in [0, 2\pi[$ tel que $a = \cos(\theta)$, $c = \sin(\theta)$, ainsi que plusieurs $\alpha \in \mathbb{R}$ tel que $b = \cos(\alpha)$, $d = \sin(\alpha)$.

Note : si nous voulons $\alpha \in [0, 2\pi[$, alors il y a unicité. Ici nous ne nous attachons pas à cette contrainte ; nous savons qu'il en existe plusieurs, et nous allons en fixer un en fonction de θ. Le α ainsi fixé ne sera peut-être pas dans $[0, 2\pi[$, mais ce ne sera pas grave.

Les angles θ et α sont alors liés par la contrainte

$$\cos(\theta)\cos(\alpha) + \sin(\theta)\sin(\alpha) = 0. \tag{16.297}$$

Utilisant l'identité (16.40c) cela signifie que $\cos(\theta - \alpha) = 0$. Donc

$$\alpha \in \{\theta + \frac{\pi}{2} + k\pi\}_{k \in \mathbb{Z}}. \tag{16.298}$$

Si k est pair, ça donne

$$\cos(\alpha) = -\sin(\theta) \tag{16.299a}$$
$$\sin(\alpha) = \cos(\theta) \tag{16.299b}$$

et alors

$$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{16.300}$$

Si au contraire k est impair, alors

$$\cos(\alpha) = \sin(\theta) \tag{16.301a}$$
$$\sin(\alpha) = -\cos(\theta), \tag{16.301b}$$

et

$$A = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}. \tag{16.302}$$

Nous avons démontré qu'une matrice de $\mathrm{O}(2)$ était forcément d'une des deux formes (16.300) ou (16.302). Il est maintenant facile de vérifier que ces deux matrices sont effectivement dans $\mathrm{O}(2)$. □

Lemme 16.113.
Toute rotation s'écrit (dans la base canonique) de façon unique sous la forme

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \tag{16.303}$$

avec $\theta \in [0, 2\pi[$.

Démonstration. Vu que SO(2) est la partie de O(2) constitué des matrices de déterminant 1, nous pouvons reprendre la forme donnée par le lemme 16.112 et fixer ϵ par la contrainte sur le déterminant.

Nous avons, en utilisant la relation du lemme 16.21,

$$\det\begin{pmatrix}\cos(\theta) & -\epsilon\sin(\theta)\\ \sin(\theta) & \epsilon\cos(\theta)\end{pmatrix}=\epsilon, \tag{16.304}$$

et donc il faut et suffit de fixer $\epsilon = 1$. □

Corollaire 16.114 ([1]).
Nous avons une bijection

$$\begin{aligned}\psi\colon\ \mathrm{SO}(2)&\to[0,2\pi[\\ \begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}&\mapsto\theta,\end{aligned} \tag{16.305}$$

et un isomorphisme de groupe

$$\begin{aligned}\varphi\colon\ \mathrm{SO}(2)&\to\mathrm{U}(1)=\{e^{i\theta}\}_{\theta\in\mathbb{R}}\\ \begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}&\mapsto e^{i\theta}.\end{aligned} \tag{16.306}$$

Démonstration. La première assertion est une paraphrase du lemme 16.113. Pour la seconde, il faut vérifier que c'est bien un morphisme et une bijection.

Pour le morphisme, ce sont les formules d'addition d'angle du lemme 16.23 qui jouent. En ce qui concerne la bijection...

Surjection Vu que $e^{i\theta+2ki\pi}=e^{i\theta}$, tout élément de U(1) est exponentielle de $i\theta$ pour un $\theta\in[0,2\pi[$.

Injection Nous avons $\varphi(A)=\varphi(B)$ lorsque les formes matricielles de A et B sous forme trigonométrique sont avec des angles différents d'un multiple de 2π. Vu que les fonctions trigonométriques sont périodiques, nous avons $A=B$ parce que leurs matrices sont égales.

□

Proposition 16.115 ([204]).
Si u et v sont des vecteurs unitaires[16] *de $\mathbb{R}^2$ alors il existe une unique rotation*[17] *f telle que $f(u)=v$.*

Démonstration. C'est la proposition 16.65 appliquée à $O=(0,0)$. □

Notons l'unicité. Nous ne faisons pas de différences entre R_θ et $R_{\theta+2\pi}$ et les autres $R_{\theta+2k\pi}$. En particulier si une rotation T est donnée, dire «$T=R_\theta$» ne définit pas un nombre θ de façon univoque. Par contre ça définit une classe modulo 2π, c'est à dire un élément $\theta\in\mathbb{R}/2\pi$.

Nous avons déjà défini le groupe SO(2) en la définition 9.51 et nous avons déterminé ses matrices dans $\mathbb{R}^2$ en le lemme 16.113.

La proposition 16.115 donne une application

$$T\colon S^1\times S^1\to\mathrm{SO}(2). \tag{16.307}$$

Et nous avons une relation d'équivalence sur $S^1\times S^1$ donnée par $(u,v)\sim(u',v')$ si et seulement si il existe $g\in\mathrm{SO}(2)$ telle que $g(u)=u'$ et $g(v)=v'$.

Définition 16.116 (Angle orienté[204]).
*Les classes de $S^1\times S^1$ pour cette relation d'équivalence sont les **angles orientés de vecteurs**. Nous notons $[u,v]$ la classe de (u,v).*

16. De norme 1.
17. Définition 16.58.

Proposition 16.117.
Nous avons $T(u,v)=T(u',v')$ si et seulement si $(u,v)\sim(u',v')$.

Démonstration. En utilisant la commutativité du groupe SO(2) nous avons équivalence entre les affirmations suivantes :

— $(u,v)\sim(u',v')$
— $T(u,u')=T(v',v')$
— $T(u,u')\circ T(u',v)=T(v,v')\circ T(u',v)$
— $T(u,v)=T(u',v')$.

□

Proposition 16.118.
Nous avons une bijection

$$\begin{aligned} S\colon \frac{S^1\times S^1}{\sim} &\to \mathrm{SO}(2)\\ [u,v] &\mapsto T(u,v). \end{aligned} \tag{16.308}$$

Démonstration. En plusieurs points.

S est bien définie En effet si $[u,v]=[z,t]$ alors $T(u,v)=T(z,t)$.

Injectif Si $S[u,v]=S[z,t]$ alors $T(u,v)=T(z,t)$, qui implique $(u,v)\sim(z,t)$ par la proposition 16.117.

Surjectif Nous avons $R_\theta=T(u,R_\theta u)$.

□

Définition 16.119 (Somme d'angles orientés[204])**.**
Si $[u,v]$ et $[z,t]$ sont des angles orientés, nous définissons la somme par

$$[u,v]+[z,t]=S^{-1}\Big(S[u,v]\circ S[z,t]\Big). \tag{16.309}$$

Lemme 16.120.
Quelque propriétés des angles plats liées à la somme.

(1) $(S^1\times S^1)/\sim$ est un groupe commutatif.

(2) Relations de Chasles :

$$[u,v]+[v,w]=[u,w]. \tag{16.310}$$

(3) $-[u,v]=[v,u]$.

Démonstration. Pour la relation de Chasles, ça se base sur la propriété correspondante sur T :

$$\begin{aligned} [u,v]+[v,w]&=S^{-1}\Big(T(u,v)\circ T(v,w)\Big) && \text{(16.311a)}\\ &=S^{-1}\big(T(u,w)\big) && \text{(16.311b)}\\ &=[u,w]. && \text{(16.311c)} \end{aligned}$$

Pour l'inverse, la vérification est que

$$[u,v]+[v,u]=[u,u]=0. \tag{16.312}$$

□

Définition 16.121.
*La **mesure** de l'angle orienté $[u,v]$ est $[\theta]_{2\pi}$ si $T[u,v]=R_\theta$.*

Notons dans cette définition qu'écrire $T[u,v] = R_\theta$ dans $\mathrm{SO}(2)$ ne définit pas θ, mais seulement sa classe modulo 2π. C'est pour cela que la mesure de l'angle orienté n'est également définie que modulo 2π.

Pour la suite nous allons nous intéresser à des vecteurs qui ont, dans l'idée, un point de départ et un point d'arrivée. Si $A, B \in \mathbb{R}^2$ nous notons

$$\overrightarrow{AB} = \frac{B-A}{\|B-A\|}. \tag{16.313}$$

C'est le vecteur unitaire dans la direction «de B vers A».

Théorème 16.122 (Théorème de l'angle inscrit[205])**.**
Soit un cercle Γ de centre O et trois points distincts $A, B, M \in \Gamma$. Alors

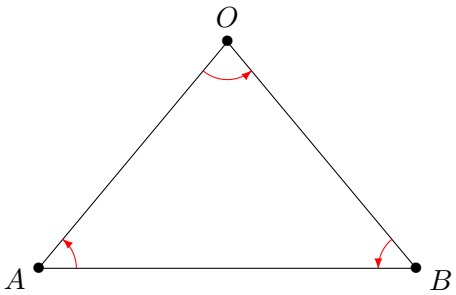
$$2(\overrightarrow{MA}, \overrightarrow{MB}) \in (\overrightarrow{OA}, \overrightarrow{OB})_{2\pi} \tag{16.314}$$

où l'indice 2π indique la classe modulo 2π.

Démonstration. Le triangle MOA est isocèle en O, donc les angles à la base sont égaux. Et de plus la somme des angles est dans $[\pi]_{2\pi}$. Bon, entre nous, nous savons que la somme des angles est exactement π, mais comme nous n'avons pas défini les angles autrement que modulo π, nous ne pouvons pas dire mieux. Donc

$$2(\overrightarrow{AB}, \overrightarrow{AO}) + (\overrightarrow{OB}, \overrightarrow{OA}) \in [\pi]_{2\pi}. \tag{16.315}$$

Il faut être sûr de l'orientation de tout cela. Le nombre $(\overrightarrow{AB}, \overrightarrow{AO})$ est l'angle qui sert à amener $\overrightarrow{AB}$ sur $\overrightarrow{AO}$. Vu que nous l'avons choisit dans le sens trigonométrique, il faut bien prendre les autres dans le sens trigonométrique et utiliser $(\overrightarrow{OA}, \overrightarrow{OB})$ et non $(\overrightarrow{OB}, \overrightarrow{OA})$.



De la même manière sur le triangle MOB nous écrivons

$$2(\overrightarrow{MB}, \overrightarrow{MO}) + (\overrightarrow{OM}, \overrightarrow{OB}) \in [\pi]_{2\pi}. \tag{16.316}$$

Nous faisons la différence entre les deux équations en nous souvenant que $-(\overrightarrow{MB}, \overrightarrow{MO}) = (\overrightarrow{MO}, \overrightarrow{MB})$ et les relations de Chasles du lemme 16.120(2) nous avons :

$$2(\overrightarrow{MA}, \overrightarrow{MB}) + (\overrightarrow{OB}, \overrightarrow{OA}) \in [0]_{2\pi}. \tag{16.317}$$

□

16.123.
Comment exprimer le fait qu'un angle orienté soit égal à θ modulo π alors que les angles orientés sont des classes modulo 2π ? Nous ne pouvons certainement pas écrire

$$(u,v) = [\theta]_\pi \tag{16.318}$$

parce que (u,v) est un élément de $S^1 \times S^1$ alors que $[\theta]_\pi$ est un ensemble de nombres. Nous pouvons écrire

$$[u,v] \subset [\theta]_\pi. \tag{16.319}$$

C'est cohérent parce que nous avons des deux côtés des ensembles de nombres. Les opérations permises sont l'égalité ou l'inclusion. L'égalité entre les deux ensembles n'est pas possible parce que la différence minimale ente deux éléments dans $[u,v]$ est 2π alors que celle dans $[\theta]_\pi$ est π.

Si u et v forment un angle droit, nous avons

$$[u,v] = \{\frac{\pi}{2} + 2k\pi\}_{k\in\mathbb{Z}}. \tag{16.320}$$

Et cela est bien un sous-ensemble de $[\pi/2]_\pi$.

Pour exprimer que la différence entre deux angles orientés diffèrent de π nous devrions écrire :

$$[u,v] \subset [a,b]_\pi \tag{16.321}$$

où le membre de droite signifie la classe modulo π d'un représentant de $[a,b]$.

Nous allons cependant nous permettre d'écrire

$$[u,v] = [a,b]_\pi \tag{16.322}$$

voire carrément

$$(u,v) = (a,b)_\pi. \tag{16.323}$$

Cette dernière égalité devant être comprise comme voulant dire que l'angle pour passer de u à v est soit le même que celui pour alle de a à b soit ce dernier plus π.

Théorème 16.124 ([205])**.**
Soient 4 points distincts du plan A,B,C,D. Ils sont alignés ou cocycliques[18] *si et seulement si*

$$(\overrightarrow{CA},\overrightarrow{CB}) = (\overrightarrow{DA},\overrightarrow{DB})_\pi. \tag{16.324}$$

Nous allons seulement démontrer l'implication directe.

Démonstration. Si les quatre points sont alignés nous avons $[\overrightarrow{CA},\overrightarrow{CB}] = [0]_{2\pi}$ et $[\overrightarrow{DA},\overrightarrow{DB}] = [0]_{2\pi}$. En particulier nous avons

$$[\overrightarrow{CA},\overrightarrow{CB}] = [\overrightarrow{DA},\overrightarrow{DB}] \tag{16.325}$$

et a fortiori l'égalité modulo π au lieu de 2π.

Nous nous relâchons en termes de notations. Si les quatre points sont cocycliques, nous pouvons utiliser le théorème de l'angle inscrit 16.122 dans les triangles ABC et ADB :

$$2(\overrightarrow{CA},\overrightarrow{CB}) = (\overrightarrow{OA},\overrightarrow{OB})_{2\pi} \tag{16.326a}$$
$$2(\overrightarrow{DA},\overrightarrow{DB}) = (\overrightarrow{OA},\overrightarrow{OB})_{2\pi}, \tag{16.326b}$$

ce qui donne $2(\overrightarrow{CA},\overrightarrow{CB}) = 2(\overrightarrow{DA},\overrightarrow{DB})_{2\pi}$ et donc

$$(\overrightarrow{CA},\overrightarrow{CB}) = (\overrightarrow{DA},\overrightarrow{DB})_\pi. \tag{16.327}$$

Comme annoncé, nous ne faisons pas la preuve dans l'autre sens ; elle peut être trouvée dans [205]. □

16.5.2 Angles et nombres complexes

Les nombres complexes peuvent être repérés par une norme et un angle, ce qui en fait un terrain propice à l'utilisation des angles orientés. Nous en ferons d'ailleurs usage dans $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ pour parler d'alignement, de cocyclicité et de birapport dans la proposition 21.85.

Soient deux éléments $z_1, z_2 \in \mathbb{C}$. Nous les écrivons sous la forme $z_1 = r_1 e^{i\theta_1}$ et $z_2 = r_2 e^{i\theta_2}$; remarquons que cela ne définit θ_i qu'à 2π près. Nous avons

$$[z_1,z_2] = [\theta_2 - \theta_1]_{2\pi}. \tag{16.328}$$

18. C'est à dire sur un même cercle.

Soient maintenant $a, b, c, d \in \mathbb{C}$. Nous écrivons $\overrightarrow{ab}$ le vecteur unitaire dans le sens «de a vers b», c'est à dire un multiple positif bien choisi du nombre $b-a$. Nous notons θ_{ab} l'argument du nombre complexe $b-a$, et nous avons encore

$$[\overrightarrow{ab}, \overrightarrow{cd}] = [\theta_{ab} - \theta_{cd}]. \tag{16.329}$$

Avec toutes ces notations, ce qui est bien est que les produits et quotients de nombres complexes se comportent très bien par rapport aux angles : l'argument de a/b est $\theta_a - \theta_b$ et en particulier l'argument de

$$\frac{a-b}{c-d} \tag{16.330}$$

est dans la classe de l'angle orienté

$$[\overrightarrow{ba}, \overrightarrow{dc}]. \tag{16.331}$$

16.5.3 Sous-groupes finis de SO(2)

Lemme 16.125 ([1])**.**
Tout sous groupe fini de SO(2) *est cyclique.*

Démonstration. Soit uns sous-groupe fini G de SO(2). Nous savons que SO(2) est isomorphe à U(1) par le corollaire 16.114, et en bijection avec $[0, 2\pi[$. Vu que G est fini, l'ensemble $G\backslash\{e\}$ il possède, dans $[0, 2\pi[$ un élément minimum non nul. Soit g_0 ce minimum.

Soit un élément g_1 de G qui ne serait ni l'identité ni un multiple de g_0. En particulier tous les nombres du type $g_1 - kg_0$ sont dans G (l'image de G dans $[0, 2\pi[$ en fait). Si g_1 n'est pas un multiple de g_0, il n'en reste pas moins que $g_1 = \lambda g_0$; alors en prenant pour k la partie entière de λ, l'élément $g_1 - kg_0$ est plus petit que g_0. Contradiction. □

16.6 Un peu de structure de O(n)

16.6.1 Valeurs propres dans O(n)

Proposition 16.126 ([206])**.**
Soit une matrice $A \in O(n)$. Si $\lambda \in \mathbb{C}$ est une valeur propre de A, alors $\bar{\lambda}$ est également une valeur propre de A, et de plus $|\lambda| = 1$.

Démonstration. Dire que $\lambda \in \mathbb{C}$ est une valeur propre de A signifie qu'il exist $x \in \mathbb{C}^n$ (non nul) tel que $Ax = \lambda x$. Vu que les éléments de la matrice A sont réels,

$$A\bar{x} = \bar{A}\bar{x} = \overline{Ax} = \overline{\lambda x} = \bar{\lambda}\bar{x}. \tag{16.332}$$

Donc $\bar{\lambda}$ est une valeur propre de A pour le vecteur propre $\bar{x}$.

Soit λ une valeur propre de A de vecteur propre x. Alors nous avons d'une part

$$\langle \overline{Ax}, Ax \rangle = \langle A^t A\bar{x}, x \rangle = \langle x, \bar{x} \rangle, \tag{16.333}$$

et d'autre part

$$\langle \overline{Ax}, Ax \rangle = \langle \bar{\lambda}\bar{x}, \lambda x \rangle = |\lambda|^2 \langle \bar{x}, x \rangle. \tag{16.334}$$

Vu que $x \neq 0$ nous avons aussi $\langle \bar{x}, x \rangle \neq 0$. Par conséquent $|\lambda|^2 = 1$ et $|\lambda| = 1$. □

Lemme 16.127 ([206])**.**
Soit un espace vectoriel euclidien E de dimension finie et une isométrie f de E. Soit F un sous-espace de E stable par f. Alors $F^\perp$ est stable par f.

Démonstration. La restriction $f_F \colon F \to F$ est encore une isométrie ; elle est donc inversible : pour tout $y \in F$, il existe $x \in F$ tel que $y = f(x)$. Soit $a \in F^\perp$; nous montrons que $f(a) \in F^\perp$. Soit donc $y \in F$ et calculons :

$$\langle y, f(a) \rangle = \langle f(x), f(a) \rangle = \langle x, a, \rangle = 0 \tag{16.335}$$

parce que $x \in F^\perp$. □

Proposition 16.128 ([206]).
Soit une isométrie $f\colon \mathbb{R}^3 \to \mathbb{R}^3$.

(1) L'application linéaire f possède au moins une valeur propre réelle qui vaut ± 1.

(2) Il existe une base orthonormée de $\mathbb{R}^3$ dans laquelle la matrice de f est de la forme

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \cos(\theta) & -\epsilon\cos(\theta) \\ 0 & \sin(\theta) & \epsilon\cos(\theta) \end{pmatrix} \tag{16.336}$$

avec $\epsilon, \lambda = \pm 1$ et $\theta \in [0, 2\pi[$.

Démonstration. Le polynôme caractéristique de f, donné par $\det(f - \lambda\,\mathrm{Id})$, est à coefficients réels et de degré 3. Il possède dont au moins une solution réelle par le corollaire 5.191. Soit donc une valeur propre réelle λ de χ_f ; par le lemme 16.126 nous avons $\lambda = \pm 1$. Soit u_1 le vecteur propre correspondant. Nous notons F l'espace engendré par u_1.

Nous avons $f(F) = F$ et donc $f(F^\perp) = F^\perp$ par le lemme 16.127. Soit une base orthonormée $\{u_2, u_3\}$ de $F^\perp$ et la matrice B de la restriction f_p à $F^\perp$. Vu que l'application f_p est une isométrie de $F^\perp$, la matrice B est, par le lemme 16.112, de la forme

$$B = \begin{pmatrix} \cos(\theta) & -\epsilon\sin(\theta) \\ \sin(\theta) & \epsilon\cos(\theta) \end{pmatrix} \tag{16.337}$$

pour un certain $\theta \in [0, 2\pi[$ et $\epsilon = \pm 1$.

Dans la base $\{u_1, u_2, u_3\}$ de $\mathbb{R}^3$, la matrice de f est alors

$$\begin{pmatrix} \lambda & 0 \\ 0 & B \end{pmatrix}, \tag{16.338}$$

comme annoncé. □

Pour classifier les isométries de $\mathbb{R}^3$, nous pouvons nous baser sur les possibilités de la matrice donne dans le lemme 16.128. Il y a essentiellement quatre possibilités suivant les valeurs de $\lambda = \pm 1$ et $\epsilon = \pm 1$.

Si $\epsilon = \lambda = 1$ Alors la matrice est

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix} \tag{16.339}$$

et l'isométrie correspondante est la rotation d'angle $-\theta$ autour de la droite de u_1.

Si $\epsilon = \lambda = -1$ Alors la matrice est

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & \sin(\theta) & -\cos(\theta) \end{pmatrix} \tag{16.340}$$

Cette application est plus subtile, parce que même dans le plan $\mathrm{Span}(u_2, u_3)$, ce n'est pas une rotation. Nous allons montrer qu'il s'agit d'une réflexion autour de la droite d'angle $\theta/2$ dans le plan $\mathrm{Span}(u_2, u_3)$. Nous nommons D cette droite. Dans la base $\{u_1, u_2, u_3\}$, les points de cette droite sont de la forme [19]

$$\big(0, \cos(\theta/2), \sin(\theta/2)\big). \tag{16.341}$$

L'image de u_1 par cette réflexion est $-u_1$, c'est clair.

19. Les plus acharnés remarquerons que $\{u_1, u_2, u_3\}$ est un ensemble, qui est une base. Mais un ensemble n'est pas ordonné, alors que pour écrire l'équation de droite qui suit, nous supposons un ordre. Je laisse au tel lecteur le soin de trouver une bonne notation.

Faisons en détail l'image de u_3. Nous devons démontrer que la droite D coupe le segment $[u_3, A(u_3)]$ perpendiculairement en son milieu.

Dans le plan $\text{Span}(u_2, u_3)$ nous avons $u_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ et $A(u_3) = \begin{pmatrix} \sin(\theta) \\ \cos(\theta) \end{pmatrix}$. Le milieu du segment $[u_3, A(u_3)]$ est le point

$$M = \left(\frac{\sin(\theta)}{2}, \frac{1 - \cos(\theta)}{2} \right). \tag{16.342}$$

Les formules de duplication d'angle du corollaire 16.24 nous permettent d'écrire $\sin(\theta)$ et $\cos(\theta)$ en fonction de $\sin(\theta/2)$ et $\cos(\theta/2)$, et donc d'exprimer le point M de la façon suivante :

$$\begin{aligned} M &= \left(\cos(\theta/2)\sin(\theta/2), \frac{1 - \big(\cos^2(\theta/2) - \sin^2(\theta/2)\big)}{2} \right) && \text{(16.343a)} \\ &= \big(\cos(\theta/2)\sin(\theta/2), \sin^2(\theta/2) \big) && \text{(16.343b)} \\ &= \sin(\theta/2)\big(\cos(\theta/2), \sin(\theta/2) \big). && \text{(16.343c)} \end{aligned}$$

Ce point fait donc partie de la droite D. La droite D coupe le segment $[u_3, A(u_3)]$ en son milieu.

En ce qui concerne l'orthogonalité, nous calculons le produit scalaire

$$\big(A(u_3) - u_3\big) \cdot \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix} = \sin(\theta)\cos(\theta/2) - \big(1 + \cos(\theta)\big)\sin(\theta/2) = 0 \tag{16.344}$$

où nous avons encore utilisé les duplications d'angles et le fait que $1 = \cos^2(\theta/2) + \sin^2(\theta/2)$ (lemme 16.21).

Si $\epsilon = -1$ et $\lambda = 1$ Alors la matrice est

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & \sin(\theta) & -\cos(\theta) \end{pmatrix}. \tag{16.345}$$

C'est la symétrie orthogonale par le plan engendré par u_1 et $v = \cos(\theta/2)u_2 + \sin(\theta/2)u_3$. Le vecteur u_1 est bien évidemment préservé par A. En ce qui concerne le vecteur v,

$$A(v) = \cos(\theta/2) \begin{pmatrix} 0 \\ \cos(\theta) \\ \sin(\theta) \end{pmatrix} + \sin(\theta/2) \begin{pmatrix} 0 \\ -\sin(\theta) \\ \cos(\theta) \end{pmatrix} = \begin{pmatrix} 0 \\ \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix} = v. \tag{16.346}$$

Nous avons sauté quelque étapes de calcul mettant en scène les formules de duplication d'angle : exprimer $\cos(\theta) = \cos^2(\theta/2) - \sin^2(\theta/2)$ et $\sin(\theta) = 2\cos(\theta/2)\sin(\theta/2)$.

Pour achever, nous devons trouver un vecteur w perpendiculaire au plan, et montrer qu'il est envoyé par A sur $-w$. Un vecteur $w = xu_1 + yu_2 + zu_3$ est perpendiculaire au plan si les deux égalités suivantes sont satisfaites :

$$\begin{aligned} \big(\cos(\theta/2)u_2 + \sin(\theta/2)u_3\big) \cdot (xu_1 + yu_2 + zu_3) = 0 && \text{(16.347a)} \\ u_1 \cdot (xu_1 + yu_2 + zu_3) = 0. && \text{(16.347b)} \end{aligned}$$

Nous avons immédiatement $x = 0$ et ensuite la relation

$$y\cos(\theta/2) + z\sin(\theta/2) = 0. \tag{16.348}$$

En ne regardant que les deux dernières composantes pour alléger l'écriture,

$$A(w) = y \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} + z \begin{pmatrix} \sin(\theta) \\ -\cos(\theta) \end{pmatrix}. \tag{16.349}$$

Le but est de montrer que cela est égal à $-y\cos(\theta/2) - z\sin(\theta/2)$.
Notons $c = \cos(\theta/2)$ et $s = \sin(\theta/2)$. Alors $A(w)_2 = y(c^2 - s^2) + 2zcs$. Évacuons tout de suite les deux cas limite : si $c = 0$ alors $A(w)_2 = -y$ (parce que $s = \pm 1$) et c'est bon. Si $s = 0$, alors $A(w)_2 = y$, mais la relation (16.348) donne $y = 0$, donc c'est bon aussi. Dans la cas générique, $z = -yc/2$ et

$$A(w)_2 = y(c^2 - s^2) - 2cs\frac{yc}{s} = -y(c^2 + s^2) = -y. \tag{16.350}$$

En ce qui concerne $A(w)_3$, c'est très similaire :

$$A(w)_3 = 2ysc - z(c^2 - s^2). \tag{16.351}$$

Avec $z = 0$ c'est $-z$, donc c'est bon. Avec $c = 0$ c'est z mais $z = 0$. Et pour le cas générique, la substitution $y = -zs/c$ donne le résultat.

Si $\epsilon = 1$ et $\lambda = -1$ Alors la matrice est

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{16.352}$$

Cela est la composition entre la symétrie de plan $\text{Span}(u_2, u_3)$ et la rotation d'angle θ dans ce plan.

16.6.2 Sous-groupes finis de SO(3)

Lemme 16.129 ([1]).
Points fixes pour SO(3).

(1) Tout élément de SO(3) *possède une droite de point fixes.*

(2) Tout élément non trivial de SO(3) *possède une seule droite de points fixes.*

Démonstration. Le polynôme caractéristique d'un élément de SO(3) est de degré trois et possède dont (en comptant les multiplicités) trois racines dont une réelle par le corollaire 5.191. Vu que nous sommes en dimension impaire, le coefficient du terme de degré 3 est -1 et le polynôme caractéristique de $g \in \text{SO}(3)$ s'écrit

$$\chi_g(X) = -(X - \lambda_1)(X - \bar{\lambda}_1)(X - s) \tag{16.353}$$

avec $s = \pm 1$ que nous allons tout de suite fixer. Nous savons que $\det(g) = \chi_g(0)$ mais aussi que $\det(g) = 1$. Donc

$$1 = \det(g) = \chi_g(0) = \lambda_1\bar{\lambda}_1 s = s. \tag{16.354}$$

Tout cela pour dire que tout élément de SO(3) possède une valeur propre égale à 1, et donc une droite de points fixes.

Pour continuer, supposons que g possède deux droites distinctes de points fixes. En particulier g fixe un plan. Une base orthonormée de $\mathbb{R}^3$ peut être choisie en prenant deux vecteurs e_1, e_2 dans ce plan et un vecteur e_3 perpendiculaire au plan.

Vu que g est une isométrie, la base reste orthonormée sous l'action de g. Donc g a pour matrice

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \pm 1 \end{pmatrix}. \tag{16.355}$$

Pour que le déterminant soit 1, il faut que la matrice soit l'identité. □

Proposition 16.130 ([1, 207, 34, 208, 209]).
Les sous-groupes finis de SO(3) *sont :*

(1) les groupes cycliques $\mathbb{Z}/n\mathbb{Z}$,
(2) les groupes diédraux D_n,
(3) le groupe alterné A_4,
(4) le groupe alterné A_5
(5) le groups symétrique S_4.

Démonstration. Soit G, un sous-groupe fini de SO(3). Par la proposition 9.50, les éléments de G sont des isométries de $\mathbb{R}^3$, et le lemme 16.129 dit que tout élément de G possède une droite de points fixes.

Un point de la sphère unité fixé par $g \in G$ est un **pôle** de g. Nous nommons Ω l'ensemble des pôles des éléments non triviaux de G.

Une action Le groupe G agit sur Ω. En effet si $x \in \Omega$, alors x est fixé par un élément g. Montrons que $h(x)$ est également fixé par un élément de G. Par dur : $(h^{-1}gh)h(x) = h(x)$; donc $h(x)$ est un pôle de $h^{-1}gh$.

Les fixateurs sont cycliques Nous montrons à présent que pour tout $x \in \Omega$, le sous-groupe $\mathrm{Fix}(x)$ est cyclique. Soit donc $x \in \Omega$, le plan orthogonal $\sigma = \mathrm{Span}(x)^\perp$ et $h \in \mathrm{Fix}(x)$. Nous avons $h(\sigma) = \sigma$. En effet si $y \in \sigma$ nous avons

$$0 = y \cdot x = h(y) \cdot h(x) = h(y) \cdot x, \tag{16.356}$$

donc $h(y)$ est perpendiculaire à x. L'inclusion inverse se démontre de même : si $y \in \sigma$ alors $y = h\big(h^{-1}(y)\big)$ alors que $h^{-1}(y) \in \sigma$.

La restriction de h à σ est une isométrie de σ. Prenant une isométrie $f\colon \sigma \to \mathbb{R}^2$, l'application

$$\begin{aligned} \varphi\colon \mathrm{Fix}(x) &\to \mathrm{SO}(2) \\ h &\mapsto f \circ h \circ f^{-1}. \end{aligned} \tag{16.357}$$

est un morphisme injectif de groupes. En effet nous avons d'une part

$$\varphi(hh') = f \circ h \circ h' \circ f^{-1} = fhf^{-1}fh'f^{-1} = \varphi(h)\varphi(h'), \tag{16.358}$$

d'où le morphisme. Et d'autre part, si $\varphi(h) = \varphi(h')$ alors $f \circ h \circ f^{-1} = g \circ h' \circ f^{-1}$, qui donne immédiatement $h = h'$.

Nous en déduisons que $\mathrm{Fix}(x)$ est isomorphe à un sous-groupe de $SO(2)$ (l'image de φ). Le lemme 16.125 en fait un groupe cyclique.

Taille des fixateurs Soient Ω_i les orbites. Si $x, y \in \Omega_i$ alors nous montrons que $|\mathrm{Fix}(x)| = |\mathrm{Fix}(y)|$ avec la bijection

$$\begin{aligned} \varphi\colon \mathrm{Fix}(x) &\to \mathrm{Fix}(y) \\ h &\mapsto g^{-1}hg \end{aligned} \tag{16.359}$$

où g est choisit de façon à avoir $y = g(x)$ (possible parce que x et y sont dans la même orbite). Cela est surjectif parce que si $k \in \mathrm{Fix}(x)$ alors $k = \varphi(gkg^{-1})$ et l'on vérifie que $gkg^{-1} \in \mathrm{Fix}(y)$. L'application φ est également injective parce que si $ghg^{-1} = gh'g^{-1}$ alors $h = h'$.

Un peu de notations Vu que tous les fixateurs des éléments d'une orbite ont la même taille (finie), nous pouvons noter

$$n_i = |\mathrm{Fix}(x_i)| \tag{16.360}$$

pour $x_i \in \Omega_i$. Nous notons également r le nombre d'orbites de G.

La formule de Burnside du théorème 3.85, avec les notations d'ici, donne

$$r = \frac{1}{|G|} \sum_{g \in G} |\mathrm{Fix}(g)|. \tag{16.361}$$

Une belle formule Soit l'ensemble

$$A = \{(g, x) \text{ tel que } g \in G\backslash\{e\}, x \in \mathrm{Fix}(g)\} \tag{16.362}$$

où par $\mathrm{Fix}(g)$ nous entendons les pôles de G fixés par g.

Il y a $|G|-1$ possibilités pour la composante g, mais chaque élément $g\neq e$ possède exactement deux pôles, donc l'ensemble A contient exactement $2(|G|-1)$ éléments.

Nous pouvons calculer le nombre d'éléments dans A d'une autre façon : pour chaque $x\in\Omega$ nous avons $|\mathrm{Fix}(x)|-1$ éléments de $G\backslash\{e\}$ qui fixent x. Donc

$$|A|=\sum_{x\in\Omega}\big(|\mathrm{Fix}(x)|-1\big). \tag{16.363}$$

Mais $|\mathrm{Fix}(x)|$ est constant sur les orbites. Nous coupons donc la somme sur Ω en plusieurs sommes sur les orbites Ω_i :

$$|A|=\sum_{i=1}^{r}\sum_{x\in\Omega_i}\big(|\mathrm{Fix}(x)|-1\big)=\sum_i|\Omega_i|(n_i-1). \tag{16.364}$$

En égalisant les deux façons de calculer $|A|$, nous déduisons la formule

$$2\big(|G|-1\big)=\sum_i|\Omega_i|(n_i-1). \tag{16.365}$$

Nous utilisons ensuite la relation orbite-stabilisateur, proposition 3.78 : $|\mathrm{Fix}(x_i)||\Omega_i|=|G|$; la formule (16.365) devient

$$2\big(|G|-1\big)=\sum_i|G|-\sum_i\frac{|G|}{n_i}=r|G|+|G|\sum_i\frac{1}{n_i}, \tag{16.366}$$

ou encore, en simplifiant par $|G|$:

$$2-\frac{2}{|G|}=r-\sum_i\frac{1}{n_i}=\sum_{i=1}^{r}\left(1-\frac{1}{n_i}\right). \tag{16.367}$$

Nous pouvons aussi repartir de (16.365) et sommer de façon plus simple $\sum_i|\Omega_i|=|\Omega|$ pour obtenir

$$2\big(|G|-1\big)=r|G|-|\Omega| \tag{16.368}$$

où Ω est l'ensemble des pôles de $G\backslash\{e\}$.

Quelles sont les possibilités ? Les nombres $|G|$, r et n_i sont des entiers. Nous allons vois qu'il n'y a pas des centaines de possibilités pour satisfaire la relation (16.367). D'abord, pour toute valeur de $|G|$ (strictement plus grande que 1),

$$1\leqslant 2-\frac{2}{|G|}<2. \tag{16.369}$$

Ensuite, si g fixe x alors g^{-1} fixe également x, de sorte que $n_i=|\mathrm{Fix}(x_i)|\geqslant 2$ pour tout i. Donc tous les termes dans la somme à droite de (16.367) sont dans $[\frac{1}{2},1[$. Nous avons donc au minimum deux termes, et au maximum trois. Autrement dit : $r=2$ ou $r=3$.

Si $r=2$ Le plus simple est de repartir de (16.368). En posant $r=2$ nous trouvons tout de suite $|\Omega|=2$. Il y a donc exactement deux pôles pour l'action de G sur la sphère unité.

Tous les éléments de G laissent donc le même axe invariant et G est un sous groupe des isométries du plan qui lui est perpendiculaire. Autrement dit, G est un sous-groupe fini de $\mathrm{SO}(2)$ et donc cyclique par le lemme 16.125.

Les possibilités pour $r=3$ Nous devons voir les solutions entières $(n_1,n_2,n_3,|G|)$ de

$$2-\frac{2}{|G|}=3-\frac{1}{n_1}-\frac{1}{n_2}-\frac{1}{n_3}<2. \tag{16.370}$$

Il faut en particulier que

$$\frac{1}{n_1}+\frac{1}{n_2}+\frac{1}{n_3}>1, \tag{16.371}$$

ce qui signifie qu'au moins un des n_i doit être 1 ou 2, mais qu'il n'est pas possible que tous les n_i soient plus grands ou égaux à 3. Vu que $n_i=|\operatorname{Fix}(x_i)|\geqslant 2$, nous en déduisons qu'au moins un des n_i doit valoir 2. Nous posons donc $n_1=2$.

De plus, nous savons que les n_i doivent diviser $|G|$. Donc $|G|$ est pair.

Si $n_2=2$ Nous sommes dans le cas $r=3$, $n_1=2$, $n_2=2$. Nous avons

$$\frac{1}{n_3}=\frac{2}{|G|}, \tag{16.372}$$

mais aussi $n_3=|G|/|\Omega_3|$ d'où nous déduisons que $|\Omega_3|=2$. Nous avons donc une orbite à deux éléments. Soit $\Omega_3=\{x,y\}$ avec $x\neq y$.

Le groupe $\operatorname{Fix}(x)$ est un groupe à $|G|/2$ éléments. Il est donc normal par le lemme 3.59. Si $g\in G$ est tel que $g(x)=y$ alors nous avons $\operatorname{Fix}(y)=g\operatorname{Fix}(x)g^{-1}$, mais comme $\operatorname{Fix}(x)$ est normal nous avons $\operatorname{Fix}(x)=\operatorname{Fix}(y)$. Donc tous les éléments de $\operatorname{Fix}(x)$ fixent x et y. Le groupe $\operatorname{Fix}(x)$ est donc un sous-groupe de $\mathrm{SO}(2)$ est est cyclique comme vu plus haut.

Mais de plus nous avons forcément $y=-x$ parce qu'un élément de G qui fixe un point fixe également l'opposé. Vu que $\Omega_3=\{x,-x\}$, il existe $s\in G$ tel que $s(x)=-x$. Évidemment, s n'est pas dans $\operatorname{Fix}(x)$ et les points fixes de s ne sont pas parmi x et $-x$. Donc l'élément s^2 a au moins 4 points fixes : les deux de s ainsi que x et $-x$. Il a donc au moins deux droites de points fixes et est donc l'identité : $s^2=e$.

De plus, vu que $s(y)$ doit être égal soit à x soit à y, et vu que $s(x)=y$, l'injectivité de s donne $s(y)=x$.

Soit a, un générateur de $\operatorname{Fix}(x)$. Nous allons montrer que $G=\operatorname{gr}(s,sa)$. Nous avons déjà

$$(sa)(x)=s(x)=y \tag{16.373a}$$

$$(sa)(y)=s(y)=x. \tag{16.373b}$$

Donc sa inverse x et y. Mais sa a ses propres deux points fixes (qui ne sont ni x ni y). L'élément $(sa)^2$ a deux quatre points fixes sur la sphère unitié : x, y et les deux de sa. Nous en déduisons que $(sa)^2=e$.

Nous nous souvenons que a est un générateur $\operatorname{Fix}(x)$. Mais $a=s\cdot sa$, donc $a^k=(ssa)^k$. Nous en déduisons que $\operatorname{gr}(s,sa)$ contient au moins $\operatorname{Fix}(x)$.

D'autre part si h et h' sont des éléments distincts dans $\operatorname{Fix}(x)$, alors sh et sh' sont des éléments distincts de $\operatorname{gr}(s,sa)$ qui ne sont pas dans $\operatorname{Fix}(x)$. Autrement dit, la partie

$$A=\{sh \text{ tel que } h\in Fix(x)\} \tag{16.374}$$

est une partie de même cardinal que $\operatorname{Fix}(x)$ tout en n'ayant aucune intersection avec $\operatorname{Fix}(x)$ (note : l'identité n'est pas dans A). Mais $|\operatorname{Fix}(x)|=|G|/2$, donc $A\cup\operatorname{Fix}(x)=G$. Et justement $A\cup G\subset\operatorname{gr}(s,sa)$. Nous en déduisons que $\operatorname{gr}(s,sa)=G$.

Le théorème 16.109 nous assure que le groupe G est alors le groupe diédral parce que les éléments s et sa vérifient les relations données en 16.103.

Si $r=3$, les autres cas possibles Nous repartons de (16.367) en posant $r=3$. Nous obtenons ceci :

$$1+\frac{2}{|G|}=\frac{1}{n_1}+\frac{1}{n_2}+\frac{1}{n_3}. \tag{16.375}$$

Nous avons déjà vu que $n_1=2$ était obligatoire, et que tous les cas où deux des n_i sont égaux à 2 sont déjà couverts. Donc n_2 et n_3 vallent 3 ou plus.

Nous trions les n_i dans l'ordre croissant. Si $n_2=4$ ou plus, alors n_3 vaut 4 ou plus. Mais

$$\frac{1}{2}+\frac{1}{4}+\frac{1}{4}=1<1+\frac{2}{|G|}. \tag{16.376}$$

Donc $n_3 = 3$ est obligatoire. Nous avons alors l'inégalité suivante qui restreint n_3 :

$$\frac{1}{n_3} = \frac{1}{6} + \frac{3}{|G|} > \frac{1}{6}. \tag{16.377}$$

Donc n_3 est 3, 4 ou 5.

Les derniers cas à couvrir sont :

— $(n_1, n_2, n_3) = (2, 3, 3)$. Dans ce cas, $\frac{7}{6} = 1 + \frac{2}{|G|}$, donc $|G| = 12$.
— $(n_1, n_2, n_3) = (2, 3, 4)$. Dans ce cas, $|G| = 24$.
— $(n_1, n_2, n_3) = (2, 3, 5)$. Dans ce cas, $|G| = 60$.

Le cas $(2, 3, 3)$ Nous utilisons les relations $n_i|\Omega_i| = |G|$ pour savoir la taille des orites. Nous avons :

(1) $2|\Omega_1| = 12$, donc $|\Omega_1| = 6$, $3|\Omega_2| = 12$, donc $|\Omega_2| = 4$, $3|\Omega_3| = 12$, donc $|\Omega_3| = 4$.

Nous avons $G \cdot \Omega_2 = \Omega_2$. D'une part parce que, par définition d'une orbite, $G \cdot \Omega_2 \subset \Omega_2$, et d'autre part parce que si $x \in \Omega_2$, alors $g^{-1}(x) \in \Omega_2$ et $g\big(g^{-1}(x)\big) = x$; donc Ω_2 est bien dans l'image de Ω_2 par G. Nous avons donc un morphisme $s\colon G \to S_{\Omega_2}$ que nous allons immédiatement voir comme

$$s\colon G \to S_4 \tag{16.378}$$

où S_4 est le groupe des permutations de $\{1, 2, 3, 4\}$.

Voyons que s est injective. Si $s(g) = s(h)$, alors $s(gh^{-1}) = \mathrm{Id}$. Autrement dit, l'élément sh^{-1} de G est l'identité sur Ω_2 qui contient 4 éléments. Fixant 4 points (au moins), l'élément sh^{-1} est l'identité. Par conséquent

$$s\colon G \to s(G) \subset S_4 \tag{16.379}$$

est un isomorphisme entre G et un sous-groupe de S_4. Mais $|G| = 12$ et $|S_4| = 24$, donc G est d'indice deux dans S_4 et est donc le groupe alterné A_4 par la proposition 8.41(3).

Le cas $(2, 3, 4)$ Nous avons $|G = 24|$ et les orbites ont pour taille :

— $2|\Omega_1| = 24$, donc $|\Omega_1| = 12$,
— $3|\Omega_2| = 24$, donc $|\Omega_2| = 8$,
— $4|\Omega_3| = 24$, donc $|\Omega_3| = 6$.

Ω_2 vient par paires Soit $x \in \Omega$ tel que $|\operatorname{Fix}(x)| = 3$. Alors $x \in \Omega_2$ parce que x est forcément dans un des Ω_i et tout élément x_i de Ω_i vérifie $|\operatorname{Fix}(x_i)| = n_i$. Mais comme les éléments de $\mathrm{SO}(3)$ sont des applications linéaires, ceux qui fixent x fixent également $-x$. Cela pour dire que si $x \in \Omega_2$, alors $-x \in \Omega_2$. Nous avons donc quatre éléments distincts a_1, a_2, a_3 et a_4 tels que

$$\Omega_2 = \{\pm a_1, \pm a_2, \pm a_3, \pm a_4\}. \tag{16.380}$$

Action sur les couples Nous prétendons que G agit sur l'ensemble des couples $\{\pm a_i\}$. C'est encore la linéarité qui joue : l'élément $g(a_i)$ est forcément un des $\pm a_k$ (éventuellement $k = i$). Si $g(a_i) = a_k$, alors $g(-a_i) = -a_k$. Autrement dit, pour tout i, il existe un k tel que $g\big(\{a_i, -a_i\}\big) = \{a_k, -a_k\}$. Cette association $i \mapsto k$ est bijective (sinon g ne serait pas bijective), et fournit donc un morphisme de groupes

$$s\colon G \to S_4. \tag{16.381}$$

s est injective Nous prouvons à présent que $s(g) = \mathrm{Id}$ si et seulement si $g = e$. Dans un sens c'est évident : $g(e) = \mathrm{Id}$. Dans l'autre sens, nous devons prouver que si $g(a_i) \in \pm a_i$ pour tout i alors $g = e$.

Si $g(a_i) = a_i$ pour tout i, alors g stabilise 4 points et l'affaire est pliée. Nous supposons qu'au moins un des a_i n'est pas stabilisé par g. Pour fixer les idées nous disons que c'est a_1. Nous avons donc $g(a_1) = -a_1$. (oui : $g(a_1) = -a_1$ et non $\pm a_k$ pour un autre k parce que nous sommes sous l'hypothèse que g stabilise les couples)

L'élément g^2 fixe tout Ω_2 ; donc $g^2 = e$. Nommons $\pm b$ les points fixes de g. Si $b \in \Omega_2$ alors $|\operatorname{Fix}(b)| = 3$, c'est à dire que les éléments de G qui fixent b sont dans un groupe d'ordre 3, et le corollaire 3.55 nous indique que ces éléments ne peuvent être que d'ordre 1 ou 3, pas deux. Nous en déduisons que b n'est pas dans Ω_2 et donc que $g(a_i) = -a_i$ pour tout i.

Jusqu'à présent nous avons prouvé que si $g \in \ker(s)$ est non trivial, alors $g(a_i) = -a_i$ pour tout i.

Soit maintenant $h \in G$. Vu que Ω_2 est une orbite, $h(a_i) \in \Omega_2$ et nous notons $h(a_i) = \epsilon a_k$ avec $\epsilon = \pm 1$ et éventiellement $k = i$ ou éventiellement pas. Nous avons :

$$(h^{-1}gh)(a_i) = \epsilon(h^{-1}g)(a_k) = -\epsilon h^{-1}(a_k) = -\epsilon^2 a_i = -a_i. \tag{16.382}$$

Donc g et $h^{-1}gh$ ont même restriction à Ω_2. En particulier $h^{-1}ghg^{-1}$ est l'identité sur Ω_2 et est donc l'identité.

Pour tout h nous avons $g = h^{-1}gh$. Les points fixes de $h^{-1}gh$ sont $\pm h^{-1}(b)$, mais aussi $\pm b$. Nous avons donc égalité d'ensemble $\{h(b), -h(b)\} = \{b, -b\}$ pour tout $h \in G$ (notez le changement de notation $h \to h^{-1}$). Cela signifie que $\{b, -b\}$ est une orbite de G. Maizon'a pas d'orbites de cardinal deux ; contradition. Nous en déduisons que e est l'unique élément de $\ker(s)$.

Conclusion La partie $s(G)$ est un sous-groupe de S_4 isomorphe à G. Mais au niveau des cardinaux, $|G| = 24$ en même temps que $|S_4| = 24$. Donc $G \simeq s(G) \simeq S_4$.

Nous passons au cas $(2, 3, 5)$, et comme ça va être long et douloureux [20], nous sautons un niveau d'indentation.

Au niveau du cardinal de G,

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{5} = 1 + \frac{2}{|G|}, \tag{16.383}$$

donc $|G| = 60$. Et pour les orbites, $|\Omega_1| = 30$, $|\Omega_2| = 20$, $|\Omega_3| = 12$.

La proposition 8.49 nous indique que le seul groupe simple d'ordre 60 est le groupe A_5. Nous allons donc nous atteler à prouver que G est simple. Vous êtes prêts ?

Fixateurs et ordres Tous les éléments de G sont dans un fixateur de type $\operatorname{Fix}(x)$, et comme l'ordre d'un élément divise l'ordre du groupe (corollaire 3.55), tous les éléments de G ont un ordre 2, 3 ou 5. Nous sommes dans un cas très particuliers parce que

— Les trois nombres 2, 3 et 5 sont des nombres premiers distincts. Donc «diviser n_i» signifie pratiquement «être égal à n_i», surtout lorsqu'on parle de l'ordre d'un élément, qui ne peut pas être 1.

— Il existe une seule orbite de chaque taille.

Nous notons $G(n_i)$ l'ensemble des éléments de G d'ordre n_i. Les parties $G(n_i)$ ne contiennent pas l'identité.

$g \in G(n_i)$ **implique** $\operatorname{Fix}(g) \subset \Omega_i$ Si $g \in G(n_i)$ et $x \in \operatorname{Fix}(g)$ alors $x \in \Omega_i$. En effet $x \in Fix(g)$ signifie $g(x) = x$ et donc aussi $g \in \operatorname{Fix}(x)$. Donc l'ordre de g divise $|Fix(x)|$, alors que l'ordre de g est n_i et que les possibilités pour $|\operatorname{Fix}(x)|$ sont exactement les n_i, lesquels sont premiers entre eux. Donc $|\operatorname{Fix}(x)| = n_i$ et $x \in \Omega_i$.

$|\operatorname{Fix}(x)| = n_i$ **implique** $x \in \Omega_i$ Comme plus haut, $g \in \operatorname{Fix}(x)$ implique que l'ordre de g divise n_i et est donc égal à n_i. Autrement dit, $g \in G(n_i)$. De plus $g \in \operatorname{Fix}(x)$ implique $x \in \operatorname{Fix}(g)$. Par le cas juste au-dessus nous déduisons $x \in \Omega_i$.

a **et** $-a$ **dans la même orbite** Nous avons évidemment $\operatorname{Fix}(a) = \operatorname{Fix}(-a)$ du fait que les éléments de G sont des applications linéaires. Si $|\operatorname{Fix}(a)| = n_i$ alors $a \in \Omega_i$ et aussi $|\operatorname{Fix}(-a)| = |\operatorname{Fix}(a)| = n_i$ et aussi $-a \in \Omega_i$.

20. Mais pas autant que le théorème 30.28, cependant.

Nombre de $\operatorname{Fix}(x_i)$ Soient $a, b \in \Omega_i$. Nous avons $|\operatorname{Fix}(a)| = |\operatorname{Fix}(b)| = n_i$ et $\operatorname{Fix}(a) = \operatorname{Fix}(b)$ si et seulement si $b = -a$ parce qu'un élément qui fixe a et b fixe automatiquement a, b, $-a$, et $-b$. Aucun élément non trivial ne peut fixer quatre points distincts. Autrement dit,

$$\operatorname{Fix}(a) \cap \operatorname{Fix}(b) = \begin{cases} \operatorname{Fix}(a) & \text{si } a = \pm b \\ \{e\} & \text{sinon.} \end{cases} \tag{16.384}$$

Chaque élément $x_i \in \Omega_i$ a son fixateur (il y en aurait $|\Omega_i| = 60/n_i$), mais ces fixateurs sont égaux deux à deux, donc il y a seulement $\frac{60}{2n_i}$ groupes distincts de la forme $|\operatorname{Fix}(x_i)|$ avec $x_i \in \Omega_i$.

Récapitulatif En reprenant ce que nous venons de dire avec $i = 1, 2, 3$ nous trouvons :

(1) $n_1 = 2$, avec $|\Omega_1| = 30$ et 15 groupes du type $\operatorname{Fix}(x_1)$ avec x_1 parcourant Ω_1.

(2) $n_2 = 3$, avec $|\Omega_2| = 20$ et 10 groupes du type $\operatorname{Fix}(x_2)$ avec x_2 parcourant Ω_2.

(3) $n_3 = 5$, avec $|\Omega_3| = 12$ et 6 groupes du type $\operatorname{Fix}(x_3)$ avec x_3 parcourant Ω_3.

Un élément non trivial de G se trouve forcément dans un et un seul de ces sous-groupes. Plus précisément, si $g \in G(n_i)$ alors g est dans un des $\operatorname{Fix}(x_i)$ avec $x_i \in \Omega_i$.

Comptons pour être sûr de ne pas s'être trompé. Chacune des lignes décrit 30 éléments de G ; par exemple pour la seconde ligne donn 10 groupes de taille $|\operatorname{Fix}(x_2)| = n_2 = 3$. Mais tout ces groupes ont pour intersection exactement $\{e\}$. Donc le comptage des éléments se fait comme suit :

$$3 \times 30 - 15 - 10 - 6 + 1. \tag{16.385}$$

Le dernier $+1$ est parce que nous aurions décompté l'identité une fois de trop. Bref, on a bien 60 éléments comme il se doit.

Un ensemble à calculer deux fois Soient les ensembles A_2, A_3 et A_5 définis par

$$A_i = \{(g, a) \in G(n_i) \times \Omega_i \text{ tel que } g(a) = a\} \tag{16.386}$$

où $G(n_i)$ est la partie de G des éléments d'ordre n_i.

Nous avons

$$|A_i| = \sum_{g \in G(n_i)} |\operatorname{Fix}(g) \cap \Omega_i|. \tag{16.387}$$

Mais les éléments de $G(n_i)$ sont d'ordre n_i, et par ce que nous avons dit plus haut, tous les éléments de $\operatorname{Fix}(g)$ sont dans Ω_i. Donc $\operatorname{Fix}(g) \cap \Omega_i = \operatorname{Fix}(g)$. Nous avons alors

$$|A_i| = \sum_{g \in G(n_i)} |\operatorname{Fix}(g)| = 2|G(n_i)| \tag{16.388}$$

parce que $|\operatorname{Fix}(g)| = 2$ pour tout g.

En compatant $|A_i|$ dans l'autre sens, nous avons

$$|A_i| = \sum_{x \in \Omega_i} |\operatorname{Fix}(x) \cap G(n_i)| \tag{16.389}$$

Vu que $x \in \Omega_i$, les éléments de $\operatorname{Fix}(x)$ sont d'ordre n_i [21] (sauf e), et comme $G(n_i)$ est justement l'ensemble des éléments d'ordre n_i dans G nous avons $\operatorname{Fix}(x) \cap G(n_i) : \operatorname{Fix}(x) \backslash \{e\}$. Cela pour dire que

$$|A_i| = \sum_{x \in \Omega_i} \Big(|\operatorname{Fix}(x)| - 1 \Big) \tag{16.390a}$$

$$= \sum_{x \in \Omega_i} |\operatorname{Fix}(x)| - \sum_{x \in \Omega_i} 1 \tag{16.390b}$$

$$= \sum_{x \in \Omega_i} n_i - |\Omega_i| \qquad |\operatorname{Fix}(x)| = n_i \text{ pcq } x \in \Omega_i \tag{16.390c}$$

$$= |\Omega_i| n_i - |\Omega_i| = |G| - |\Omega_i|. \tag{16.390d}$$

21. Encore et toujoure parce que les éléments de $\operatorname{Fix}(x)$ ont un ordre qui divise $|\operatorname{Fix}(x)| = n_i$ et que n_i est premier, et que nous avons exclu l'identité.

En égalisant cela à la valeur $2|G(n_i)|$ déjà trouvée, nous déduisons les valeurs des $|G(n_i)|$:

$$|G(n_i)| = \frac{|G| - |\Omega_i|}{2}. \tag{16.391}$$

Nous avons alors

(1) $|G(2)| = 15$
(2) $|G(3)| = 20$
(3) $|G(5)| = 24$

Les Sylow de G Les p-Sylow sont définis en 8.6, et le super théorème qui répond à toutes les questions est le théorème 8.11. Dans notre cas, les diviseurs premiers de $|G| = 60$ sont 2, 3 et 5. Il faut faire attention au 2 parce que sa plus haute puissance dans la décomposition de 60 est 4 et non 2. Nous avons :

(1) Un 2-Sylow est un sous-groupe d'ordre 4.
(2) Un 3-Sylow est un sous-groupe d'ordre 3.
(3) Un 5-Sylow est un sous-groupe d'ordre 5.

Entre autres :

(1) Les 10 sous-groupes $\text{Fix}(x_2)$ avec $x_2 \in \Omega_2$ sont des 3-Sylow de G.
(2) Les 6 sous-groupes $\text{Fix}(x_3)$ avec $x_3 \in \Omega_3$ sont des 5-Sylow de G.
(3) Les 15 sous-groupes $\text{Fix}(x_1)$ avec $x_1 \in \Omega_1$ sont d'ordre 2 et ne sont donc pas des 2-Sylow de G.

Tous les 3-Sylow et les 5-Sylow Nous avons déjà trouvé 10 3-Sylow et 6 5-Sylow. Nous montrons à présent qu'il n'y en a pas d'autres. Le théorème de Sylow 8.11(4) nous indique que le nombre n_3 de 3-Sylow est :
— diviseur de 60,
— dans $[1]_3$
— au moins 10.
Les diviseurs de 60 sont :

$$1, 5, 3, 15, 2, 10, 6, 30, 4, 20, 12, 60. \tag{16.392}$$

Le seul qui vérifie toutes les conditions est 10. Donc G possède seulement 10 3-Sylow et ils sont tous de la forme $\text{Fix}(x_2)$ avec $x_2 \in \Omega_2$.

Même raisonnement pour les 5-Sylow : il faut
— diviseur de 60,
— dans $[1]_5$
— au moins 6.
La seule possibilité est 6.

Sous-groupe normal Soit H, un sous-groupe normal de G. Notre but étant de prouver que G est simple, nous voulons prouver que H est soit $\{e\}$ soit G. Nous supposons que H est non trivial, et nous allons prouver que $H = G$.

Le théorème de Lagrange 3.54(1) nous dit que $|H|$ divise $|G|$. Le nombre $|H|$ ne peut donc avoir que 2, 3 et 5 comme facteurs premiers. Avec une mention spéciale pour le 2 : $|H|$ pourrait être divisible aussi par 4.

Diviseurs de $|H|$ Soit un sous-groupe normal H de G. Vu que c'est un sous-groupe sont ordre divise celui de G (encore et toujours le théorème de Lagrange 3.54), et donc les facteurs premiers de $|H|$ ne peuvent être que 2, 3 et 5.

Si $|H|$ est divisible en 3 Alors H contient au moins un 3-Sylow. Mais nous avons vu que les 3-Sylow de H sont les 3-Sylow de G. Donc H contient tous les 3-Sylow de G, parce que les 3-Syow sont conjugués et H est normal.

Soit E l'ensemble des sous-groupes de H. Vu qu'il est normal, H agit sur E par conjugaison, et les 3-Sylow forment une orbite. Si α est un 3-Sylow, la formule des classes (proposition 3.78(2)) nous donne

$$|H| = |\operatorname{Fix}(\alpha)||\mathcal{O}_\alpha|. \tag{16.393}$$

Mais l'orbite $\mathcal{O}_\alpha$ de α est l'ensemble des 3-Sylow, de sorte que $|\mathcal{O}_\alpha| = 10$. Donc $|H|$ est divisible en 10.

Mais il y a pire : H contient au moins les 10 sous-groupes $\operatorname{Fix}(x_2)$ pour $x_2 \in \Omega_2$. Ce sont 10 groupes de $|\operatorname{Fix}(x_2)| = 3$ éléments. En décomptant e qui est dans l'intersection, cela fait

$$10 \times |\operatorname{Fix}(x_2)| - 10 + 1 = 21 \tag{16.394}$$

éléments. Donc H contient au moins 21 éléments. Le nombre $|H|$ est donc :

— diviseur de 60
— multiple de 10
— au moins 21.

Donc c'est 30 ou 60.

Si $|H|$ est divisible en 5 Le même raisonnement tient et $|H|$ est 30 ou 60.

Nous restons avec les possibilités $|H|$ égal à 2, 4, 30 ou 60.

Si $|H| = 4$ Alors H contient au moins un 2-Sylow. Un 2-Sylow de H est un sous-groupe contenant 4 éléments qui sont d'ordre 2^m. Le seul m possible dans G est $m = 1$. Vu qu'un 2-Sylow de H contient 4 éléments, nous sommes dans le cas où H est un 2-Sylow. Il est donc le seul 2-Sylow de H parce que H est normal et que tous les 2-Sylow sont conjugués.

Mais tous les sous-groupes d'ordre 2 sont contenus dans un 2-Sylow. En particulier tous les 15 groupes $\operatorname{Fix}(x_1)$ sont dans l'unique 2-Sylow H qui est soit-disant d'ordre 4. IL y a là une belle impossibilité.

Donc le cas $|H| = 4$ est hors-concours.

Si $|H| = 2$ Alors $H = \{e, g\}$ avec $g^2 = e$. Si $h \in G$, l'élément hgh^{-1} ne peut être que e ou g (parce que H est normal). Le premier cas est $g = e$, et le second donne $gh = hg$. Donc g est dans le centre de G : il commute avec tous les éléments de G.

Vu que $g \in G(2)$, nous avons que les éléments $a \in \operatorname{Fix}(g)$ sont forcément dans Ω_1 parce que les points dont les fixateurs sont formés d'éléments d'ordre 2 sont dans Ω_1. Soit $h \in G$. Nous avons $g = hgh^{-1}$ et donc aussi

$$\big(hgh^{-1}\big)\big(h(a)\big) = hg(a) = h(a), \tag{16.395}$$

donc $h(a)$ et $-h(a)$ sont des points fixes de hgh^{-1}. Ce sont donc également de points fixes de g. Nous en déduisons que g a pour points fixes les points a, $-a$, $h(a)$ et $-h(a)$. Vu que g n'est pas e, ces quatre points ne peuvent pas être distincts. Vu que $h(a)$ ne peut pas être $-h(a)$, nous avons forcément $h(a) = \pm a$.

Donc l'orbite de a ne contiendrait que 2 éléments. Pas possible.

Si $|H| = 30$ À part $|H| = 60$, le dernier cas à traiter est $|H| = 30$. Nous rappellons obligeament que

(1) $|G(2)| = 15$
(2) $|G(3)| = 20$
(3) $|G(5)| = 24$.

Si H possède 30 éléments, le théorème de Sylow dit que H contient au moins un 3-Sylow et un 5-Sylow, et donc tous. Vu que pour 3 et 5, les Sylow de H et de G sont les mêmes et bien idéntifiés, nous allons nous baser dessus. Le sous-groupe H contient tous les 3 et 5-Sylow, donc le comptage des éléments est :

$$10 \times |\operatorname{Fix}(x_2)| + 6 \times |\operatorname{Fix}(x_3)| - 15 = 45. \tag{16.396}$$

Nous aurions aussi pu ajouter $+4 - 1$ pour compter au moins un 2-Sylow.

Donc dès que H compte 30 éléments, il en compte 45 et donc 60 parce qu'il n'y a pas de diviseurs de 60 entre 45 et 60.

□

Problèmes et choses à faire

La démonstration des groupes finis de SO(3) est longue. Je me demande si il n'y a pas moyen de faire plus court. Par exemple [207] utilise le théorème de Cauchy 3.56 que je n'utilise pas. D'autre part, toutes les références me semble utiliser plus ou moins implicitement le fait que si le sous-groupe normal H contient un élément de $G(n_i)$, alors il les continent tous. J'avoue ne pas trop comprendre pourquoi.

16.7 Coordonnées polaires

16.7.1 Les coordonnées polaires

On a vu qu'un point M dans $\mathbb{R}^2$ peut être représenté par ses abscisses x et ses ordonnées y. Nous pouvons également déterminer le même point M en donnant un angle et une distance comme montré sur la figure 16.8.

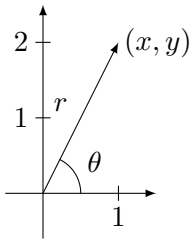


FIGURE 16.8 – Un point en coordonnées polaires est donné par sa distance à l'origine et par l'angle qu'il faut avec l'horizontale.

Le même point M peut être décrit indifféremment avec les coordonnées (x, y) ou bien avec (r, θ).

Remarque 16.131.
L'angle θ d'un point n'étant a priori défini qu'à un multiple de 2π près, nous convenons de toujours choisir un angle $0 \leqslant \theta < 2\pi$. Par ailleurs l'angle θ n'est pas défini si $(x, y) = (0, 0)$.

La coordonnée r est toujours positive.

En utilisant la trigonométrie, il est facile de trouver le changement de variable qui donne (x, y) en fonction de (r, θ) :

$$\begin{cases} x = r\cos(\theta) & \text{(16.397a)} \\ y = r\sin(\theta). & \text{(16.397b)} \end{cases}$$

16.7.1.1 Transformation inverse : théorie

Voyons la question inverse : comment retrouver r et θ si on connait x et y ? Tout d'abord,

$$r = \sqrt{x^2 + y^2} \tag{16.398}$$

parce que la coordonnée r est la distance entre l'origine et (x, y). Comment trouver l'angle ? Nous supposons $(x, y) \neq (0, 0)$. Si $x = 0$, alors le point est sur l'axe vertical et nous avons

$$\theta = \begin{cases} \pi/2 & \text{si } y > 0 \\ 3\pi/2 & \text{si } y < 0 \end{cases} \tag{16.399}$$

Notez que si $y < 0$, conformément à notre convention $\theta \geqslant 0$, nous avons noté $\frac{3\pi}{2}$ et non $-\frac{\pi}{2}$.

Supposons maintenant le cas général avec $x \neq 0$. Les équations (16.397) montrent que

$$\tan(\theta) = \frac{y}{x}. \tag{16.400}$$

Nous avons donc

$$\theta = \tan^{-1}\left(\frac{y}{x}\right). \tag{16.401}$$

La fonction inverse de la fonction tangente est celle définie plus haut.

16.7.1.2 Transformation inverse : pratique

Le code suivant utilise Sage.

```
#! /usr/bin/sage -python
# -*- coding: utf8 -*-

from sage.all import *

def PointToPolaire(x,y):
  x=SR(x)
  y=SR(y)
  r = sqrt(x**2+y**2)
  if x == 0:
    if y > 0:
      alpha = pi/2
    if y < 0:
      alpha = 3*pi/2
    if y == 0 :
      raise ValueError,"Pas d'angle pour le point (0,0) !!"
  else :
    alpha = atan(y/x)
  if (x < 0) and (y == 0) :
    alpha = pi
  if (x < 0) and (y > 0) :
    alpha = alpha + pi
  if (x < 0) and (y < 0 ) :
    alpha = alpha + pi
  alpha=alpha.simplify_trig()
  return (r,alpha)

print PointToPolaire(1,1)
print PointToPolaire(-2,1)
print PointToPolaire(6*sqrt(3)/2,3)
```

tex/frido/calculAngle.py

Son exécution retourne :

```
(sqrt(2), 1/4*pi)
(sqrt(5), pi - arctan(1/2))
(6, 1/6*pi)
```

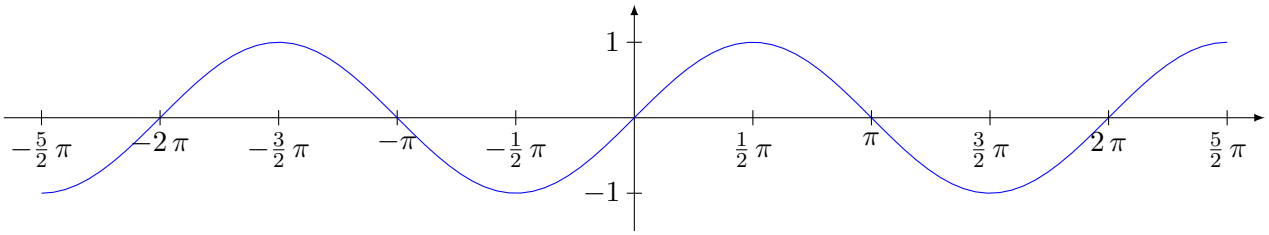
Notez que ce sont des valeurs *exactes*. Ce ne sont pas des approximations, Sage travaille de façon symbolique.

Voici un tableau qui rappelle les valeurs à retenir pour les fonctions sinus, cosinus et tangente.

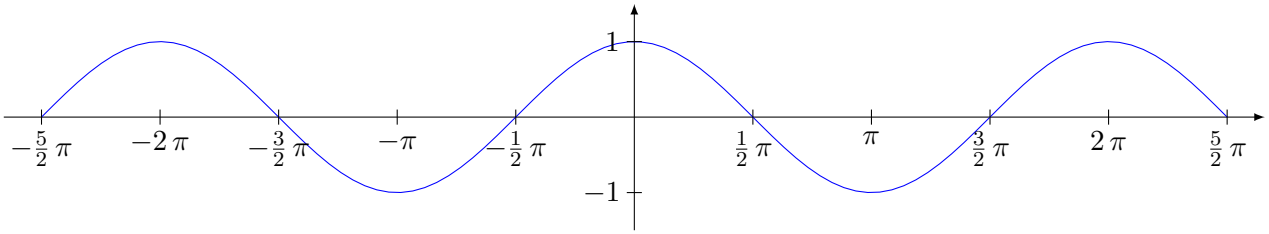
x	$\sin(x)$	$\cos(x)$	$\tan(x)$
0	0	1	0
$\pi/6$	$1/2$	$\sqrt{3}/2$	$\sqrt{3}/3$
$\pi/6$	$1/2$	$\sqrt{3}/2$	$\sqrt{3}/3$
$\pi/4$	$\sqrt{2}/2$	$\sqrt{2}/2$	1
$\pi/3$	$\sqrt{3}/2$	$1/2$	$\sqrt{3}$
$\pi/2$	1	0	N.D.

où «N.D.» signifie «non défini».

Rappelons le graphe de la fonction sinus :



celui de la fonction cosinus :



Lemme 16.132.
Pour toute valeur de $x \in \mathbb{R}$ on a $|\sin(x)| \leqslant |x|$.

Démonstration. Nous séparons des cas en fonction des valeurs.

- Si $0 \leqslant x \leqslant \pi/2$ alors le sinus de x est la longueur du cathète verticale du triangle rectangle de sommets $O = (0,0)$, $A = (\cos(x), \sin(x))$ et $B = (\cos(x), 0)$. Le triangle de sommets A, B et $C = (1,0)$ est aussi rectangle et nous savons que chacun des cathètes ne peut pas être plus long que l'hypoténuse. Donc $\sin(x)$ est inférieur à la longueur du segment AC. À son tour le segment AC ne peut pas être plus long que l'arc de cercle $\widehat{A0C}$, car le chemin le plus court entre deux points du plan est toujours donné par un morceau de droite. La longueur de l'arc de cercle $\frown AC$ est *par définition* la mesure en radiants de l'angle $\widehat{AOC}$, qui est x et on a l'inégalité $\sin(x) \leqslant x$.
- Si $-\pi/2 \leqslant x \leqslant 0$ le même raisonnement que au point précédent permet de conclure que $\sin(x) \leqslant |x|$.
- Nous savons par ailleurs que la fonction sinus prend ses valeurs dans l'intervalle $[-1,1]$ et donc pour tout x tel que $|x| \geqslant \pi/2 \equiv 1,57\ldots$ on a forcement $|\sin(x)| \leqslant |x|$.

□

16.7.2 Coordonnées polaires comme changement de variables pour intégrales

Le théorème 13.224 manque un peu d'exemples. Nous allons en voir quelque uns maintenant.

Exemple 16.133
Coordonnées polaires : On veut évaluer l'intégrale de la fonction $f(x,y) = x^2 + y^2$ sur la région V suivante :

$$V = \{(x,y) \in \mathbb{R}^2 \,|\, x^2 + y^2 \leqslant 1,\ x > 0,\ y > 0\}.$$

On peut faire le calcul directement,

$$\int_V f(x,y)\,dV = \int_0^1 \int_0^{\sqrt{1-x^2}} x^2 + y^2\,dy\,dx = \int_0^1 x^2\sqrt{1-x^2} + \frac{(1-x^2)^{3/2}}{3}\,dx$$

mais c'est un peu ennuyeux. On peut simplifier beaucoup les calculs avec un changement de variables vers les coordonnées polaires. Dans ce cas, on sait bien que le difféomorphisme à utiliser est $\phi(r,\theta) = (r\cos\theta, r\sin\theta)$. Le jacobien J_ϕ est

$$J_\phi(r,\theta) = \begin{vmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{vmatrix} = r, \tag{16.402}$$

qui est toujours positif. La fonction f peut s'écrire comme $f(\phi(r,\theta)) = r^2$ et $\phi^{-1}(V) =]0,1]\times]0,\pi/2[$. La formule du changement de variables nous donne

$$\int_V f(x,y)\,dV = \int_0^{\pi/2}\int_0^1 r^3 dr\,d\theta = \int_0^{\pi/2}\frac{1}{4}\,d\theta = \frac{\pi}{8}.$$

△

Exemple 16.134
Coordonnées cylindriques : On veut calculer le volume de la région A définie par l'intersection entre la boule unité et le cylindre qui a pour base un disque de rayon $1/2$ centré en $(0,1/2)$

$$A = \{(x,y,z)\in\mathbb{R}^3 \,|\, x^2+y^2+z^1 \leqslant 1\} \cap \{(x,y,z)\in\mathbb{R}^3 \,|\, x^2+(y-1/2)^2 \leqslant 1/4\}.$$

On peut décrire A en coordonnées cylindriques

$$\begin{aligned} A = \Big\{(r,\theta,z)\in]0,+\infty[\times[-\pi,\pi[\times\mathbb{R}\,| \\ -\pi/2<\theta<\pi,\ 0<r\leqslant\sin\theta,\ -\sqrt{1-r^2}\leqslant z\leqslant\sqrt{1-r^2}\Big\}. \end{aligned} \tag{16.403}$$

Le jacobien de ce changement de variables, J_{cyl}, est

$$J_{cyl}(r,\theta),z = \begin{vmatrix} \cos\theta & \sin\theta & 0 \\ -r\sin\theta & r\cos\theta & 0 \\ 0 & 0 & \end{vmatrix} = r, \tag{16.404}$$

qui est toujours positif. Le volume de A est donc

$$\int_{\mathbb{R}^3}\chi_A(x,y,z)\,dV = \int_{-\pi/2}^{\pi/2}\int_0^{\sin\theta}\int_{-\sqrt{1-r^2}}^{\sqrt{1-r^2}} r\,dz\,dr\,d\theta = \frac{2\pi}{8}+\frac{8}{9}.$$

△

Exemple 16.135
Volume d'un solide de révolution : Soit $g:[a,b]\to\mathbb{R}_+$ une fonction continue et positive. On dit que le solide A décrit par

$$A = \Big\{(x,y,z)\in\mathbb{R}^3 \,|\, z\in[a,b],\ \sqrt{x^2+y^2}\leqslant g^2(z)\Big\}$$

est un solide de révolution. Afin de calculer son volume, on peut décrire A en coordonnées cylindriques,

$$A = \big\{(r,\theta,z)\in]0,+\infty[\times[-\pi,\pi[\times\mathbb{R}\,|\,a\leqslant z\leqslant b,\ 0<r^2\leqslant g^2(z)\big\}.$$

Le jacobien de ce changement de variables est $J_{cyl} = r$, comme dans l'exemple précédent. Le volume de A est donc

$$\int_{\mathbb{R}^3}\chi_A(x,y,z)\,dV = \int_a^b\int_{-\pi}^{\pi}\int_0^{g(z)} r\,dr\,d\theta\,dz = \int_a^b \pi g^2(z)\,dz.$$

Cette formule peut être utilisée pour tout solide de révolution. △

Exemple 16.136
Coordonnées sphériques : On veut calculer le volume du cornet de glace A

$$A = \Big\{(x,y,z)\in\mathbb{R}^3 \,|\, (x,y)\in\mathbb{S}^2,\ \sqrt{x^2+y^2}\leqslant z\leqslant\sqrt{1-x^2-y^2}\Big\}.$$

On peut décrire A en coordonnées sphériques.

$$A = \{(\rho, \theta, \phi) \in]0, +\infty[\times[-\pi, \pi[\times[0, \pi[\,|\, 0 < \phi \leqslant \pi/4,\ 0 < \rho \leqslant 1\}.$$

Le jacobien de ce changement de variables J_{sph} est

$$J_{sph}(\rho, \theta, \phi) = \begin{vmatrix} \cos\theta\sin\phi & \sin\theta\sin\phi & \cos\phi \\ -\rho\sin\theta\sin\phi & \rho\cos\theta\sin\phi & 0 \\ \rho\cos\theta\cos\phi & \rho\sin\theta\cos\phi & -\rho\sin\phi \end{vmatrix} = \rho^2\sin\phi, \tag{16.405}$$

Le volume de A est donc

$$\int_{\mathbb{R}^3} \chi_A(x, y, z)\, dV = \int_{-\pi}^{\pi}\int_0^{\pi/4}\int_0^1 \rho^2\sin\phi\, d\rho\, d\phi\, d\theta = \frac{2\pi}{3}\left(1 - \frac{1}{\sqrt{2}}\right).$$

△

16.7.3 Changement de variables

Le domaine $E = \{(x, y) \in \mathbb{R}^2$ tel que $x^2+y^2 < 1\}$ s'écrit plus facilement $E = \{(r, \theta)$ 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

$$\int_E f = \int_0^{2\pi}\int_0^1 f(r\cos\theta, r\sin\theta) dr d\theta. \tag{16.406}$$

Hélas, non ; la vie n'est pas aussi simple.

Théorème 16.137.
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

$$\int_F f(x)dy = \int_{g^{-1}(F)f\left(g(x)\right)} |Jg(x)|dx \tag{16.407}$$

Pour rappel, Jg est le déterminant de la matrice jacobienne (aucun lien de parenté) donnée par

$$Jg = \det\begin{pmatrix} \partial_x g_1 & \partial_y g_1 \\ \partial_x g_2 & \partial_t g_2 \end{pmatrix}. \tag{16.408}$$

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.

16.7.3.1 Coordonnées polaires

Les coordonnées polaires sont données par le difféomorphisme

$$\begin{aligned} g\colon]0, \infty[\times]0, 2\pi[&\to \mathbb{R}^2\backslash D \\ (r, \theta) &\mapsto \big(r\cos(\theta), r\sin(\theta)\big) \end{aligned} \tag{16.409}$$

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

$$Jg = \det\begin{pmatrix} \partial_r x & \partial_\theta x \\ \partial_r y & \partial_\theta y \end{pmatrix} = \det\begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix} = r. \tag{16.410}$$

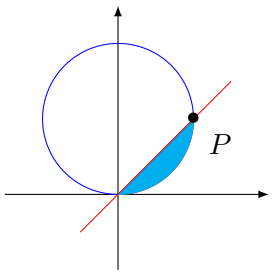


FIGURE 16.9 – Passage en polaire pour intégrer sur un morceau de cercle.

Exemple 16.138
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 16.9. 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 : $y^2 - y = (y - \frac{1}{2})^2 - \frac{1}{4}$.

Le passage en polaire transforme les équations du bord du domaine en

$$\begin{aligned}\cos(\theta) &= \sin(\theta)\\ r^2 &= r\sin(\theta).\end{aligned} \tag{16.411}$$

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

$$\int_0^{\pi/4} \left(\int_0^{\sin(\theta)} \sqrt{1-r^2} r \, rd \right). \tag{16.412}$$

Remarquez la présence d'un r supplémentaire pour le jacobien.

Notez que les coordonnées du point P sont $(1,1)$. △

16.7.3.2 Coordonnées sphériques

Les coordonnées sphériques sont données par

$$\left\{\begin{array}{lll} x = r\cos\theta\sin\varphi & & r \in]0,\infty[\\ y = r\sin\theta\sin\varphi & \text{avec} & \theta \in]0,2\pi[\\ z = r\cos\varphi & & \phi \in]0,\pi[. \end{array}\right. \tag{16.413}$$

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

$$\int_0^R dr \int_0^{2\pi} d\theta \int_0^{\pi} r^2\sin(\phi)d\phi = 4\pi R = \frac{4}{3}\pi R^3. \tag{16.414}$$

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 (16.413) est encore une paramétrisation de $\mathbb{R}^3$ si on intervertit le domaine des angles :

$$\begin{aligned}\theta &: 0 \to \pi\\ \phi &: 0 \to 2\pi,\end{aligned} \tag{16.415}$$

alors nous paramétrons encore parfaitement bien la sphère, mais hélas

$$\int_0^R dr \int_0^{\pi} d\theta \int_0^{2\pi} r^2\sin(\phi)d\phi = 0. \tag{16.416}$$

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 (16.416) n'est pas correcte. Il faut la remplacer par

$$\int_0^R dr\int_0^\pi d\theta\int_0^\pi r^2\sin(\phi)d\phi - \int_0^R dr\int_0^\pi d\theta\int_\pi^{2\pi} r^2\sin(\phi)d\phi = \frac{4}{3}\pi R^3 \tag{16.417}$$

16.7.4 Coordonnées polaires

Soit T la fonction de $]0,+\infty[\times\mathbb{R}$ dans $\mathbb{R}^2\backslash\{(0,0)\}$ définie par

$$\begin{array}{rccc} T: &]0,+\infty[\times\mathbb{R} & \to & \mathbb{R}^2\backslash\{(0,0)\} \\ & (r,\theta) & \mapsto & (r\cos\theta, r\sin\theta), \end{array} \tag{16.418}$$

Cette fonction est surjective. Elle est bijective sur chaque bande de la forme $]0,+\infty[\times[a-\pi, a+\pi[$. Si $a=0$ l'inverse de T est la fonction $T^{-1}(x,y) = (\sqrt{x^2+y^2}, \text{arctg}(y/x))$. Soit $P=(x,y)$ un élément dans $\mathbb{R}^2$, on dit que $r=\sqrt{x^2+y^2}$ est le rayon de P et que $\theta = \text{arctg}(y/x)$ est son argument principal. L'origine ne peut pas être décrite en coordonnées polaires parce que si son rayon est manifestement zéro, on ne peut pas lui associer une valeur univoque de l'angle θ.

Exemple 16.139
L'équation du cercle de rayon a et centre $(0,0)$ en coordonnées polaires est $r=a$. △

Exemple 16.140
Une équation possible pour la demi-droite $x=y$, $x>0$, est $\theta=\pi/4$. △

16.7.5 Coordonnées cylindriques

Soit T la fonction de $]0,+\infty[\times\mathbb{R}^2$ dans $\mathbb{R}^3\backslash\{(0,0,0)\}$ définie par

$$\begin{array}{rccc} T: &]0,+\infty[\times\mathbb{R}\times\mathbb{R} & \to & \mathbb{R}^3\backslash\{(0,0,0)\} \\ & (r,\theta,z) & \mapsto & (r\cos\theta, r\sin\theta, z), \end{array} \tag{16.419}$$

Cette fonction est surjective. Elle est bijective sur chaque bande de la forme $]0,+\infty[\times[a-\pi, a+\pi[\times\mathbb{R}$, a dans $\mathbb{R}$. Il n'y a presque rien de nouveau par rapport aux coordonnées polaires. Les coordonnées cylindriques sont intéressantes si on décrit un objet invariant par rapport aux rotations autour de l'axe des z.

Exemple 16.141
Il faut savoir ce que décrivent les équations les plus simples en coordonnées cylindriques,

- $r \leqslant a$, pour a constant dans $]0,+\infty[$, est le cylindre de hauteur infinie qui a pour axe l'axe des z et pour base le disque de rayon a centré à l'origine,
- $r=a$ est la surface du cylindre,
- $\theta=b$ est un demi-plan ouvert et sa fermeture contient l'axe des z,
- $z=c$ est un plan parallèle au plan x-y.

△

Exemple 16.142
Un demi-cône qui a son sommet en l'origine et pour axe l'axe des z est décrit par $z=dr$. Si d est positif il s'agit de la moitié supérieure du cône, si $d<0$ de la moitié inférieure. △

Exemple 16.143
De même, la sphère de rayon a et centrée à l'origine est l'assemblage des calottes $z=\sqrt{a^2-r^2}$ et $z=-\sqrt{a^2-r^2}$. △

16.7.6 Coordonnées sphériques

Soit T la fonction de $]0,+\infty[\times\mathbb{R}^2$ dans $\mathbb{R}^3\backslash\{(0,0,0)\}$ définie par

$$\begin{array}{rccc} T: &]0,+\infty[\times\mathbb{R}\times\mathbb{R} & \to & \mathbb{R}^3\backslash\{(0,0,0)\} \\ & (\rho,\theta,\phi) & \mapsto & (\rho\cos\theta\sin\phi, \rho\sin\theta\sin\phi, \rho\cos\phi), \end{array} \tag{16.420}$$

Cette fonction est surjective. Elle est bijective sur chaque bande de la forme $]0,+\infty[\times[a-\pi,a+\pi[\times[b-\pi/2,b+\pi/2[$, a et b dans $\mathbb{R}$. Si $a=0$ et $b=-\pi/2$ la fonction inverse T^{-1} est donnée donnée

$$\begin{array}{rccc} T: & \mathbb{R}^3\backslash\{(0,0,0)\} & \to &]0,+\infty[\times[-\pi,\pi[\times[0,\pi[\\ & (x,y,z) & \mapsto & \left(\sqrt{x^2+y^2+z^2}, \operatorname{arctg}\frac{y}{x}, \arccos\left(\frac{z}{\sqrt{x^2+y^2+z^2}}\right)\right). \end{array} \tag{16.421}$$

Soit P un point dans $\mathbb{R}^3$. L'angle ϕ est l'angle entre le demi-axe positif des z et le vecteur $\overrightarrow{OP}$, ρ est la norme de $\overrightarrow{OP}$ et θ est l'argument en coordonnées polaires de la projection de $\overrightarrow{OP}$ sur le plan x-y.

Remarque 16.144.
Dans la littérature, les angles θ et ϕ sont parfois inversés (voire, changent de nom, par exemple φ au lieu de ϕ). Il faut donc être très prudent lorsqu'on veut utiliser dans un cours des formules données dans un autre cours.

Exemple 16.145
Il faut connaître le sens des équations plus simples,

- $\rho \leqslant a$, pour a constant dans $]0,+\infty[$, est la boule fermée de rayon a centrée à l'origine,
- $\rho = a$ est la sphère de rayon a centrée à l'origine,
- $\theta = b$ est un demi-plan ouvert et sa fermeture contient l'axe des z,
- $\phi = c$ est un demi-cône qui a son sommet à l'origine et pour axe l'axe des z. Si c est positif il s'agit de la moitié supérieure du cône, si $d<0$ de la moitié inférieure.

△

16.7.7 Coordonnées cylindriques et sphériques

Les **coordonnées cylindriques** sont un perfectionnement des coordonnées polaires. Il s'agit simplement de donner le point (x,y,z) en faisant la conversion $(x,y)\mapsto(r,\theta)$ et en gardant le z. Les formules de passage sont

$$\begin{cases} x = r\cos(\theta) & \text{(16.422a)} \\ y = r\sin(\theta) & \text{(16.422b)} \\ z = z. & \text{(16.422c)} \end{cases}$$

Les **coordonnées sphériques** sont ce qu'on appelle les «méridiens» et «longitudes» en géographie. Les formules de transformation sont

$$\begin{cases} x = \rho\sin(\theta)\cos(\varphi) & \text{(16.423a)} \\ y = \rho\sin(\theta)\sin(\varphi) & \text{(16.423b)} \\ z = \rho\cos(\theta) & \text{(16.423c)} \end{cases}$$

avec $0\leqslant\theta\leqslant\pi$ et $0\leqslant\varphi<2\pi$.

Remarque 16.146.
Attention : d'un livre à l'autre les conventions sur les noms des angles changent. N'essayez donc pas d'étudier par cœur des formules concernant les coordonnées sphériques trouvées autre part.

Par exemple sur le premier dessin de wikipédia, l'angle φ est noté θ et l'angle θ est noté Φ. Mais vous noterez que sur cette même page, les convention de noms de ces angles changent plusieurs fois.

Trouvons le changement inverse, c'est à dire trouvons ρ, θ et φ en termes de x, y et z. D'abord nous avons

$$\rho = \sqrt{x^2 + y^2 + z^2}. \tag{16.424}$$

Ensuite nous savons que

$$\cos(\theta) = \frac{z}{\rho} \tag{16.425}$$

détermine de façon unique[22] un angle $\theta \in [0, \pi]$. Dès que ρ et θ sont connus, nous pouvons poser $r = \rho \sin\theta$ et alors nous nous trouvons avec les équations

$$\begin{cases} x = r\cos(\varphi) & \text{(16.426a)} \\ y = r\sin(\varphi), & \text{(16.426b)} \end{cases}$$

qui sont similaires à celles déjà étudiées dans le cas des coordonnées polaires.

16.7.8 Récapitulatif des changements de variables

En pratique, nous retiendrons les formules suivantes :

16.7.8.1 Coordonnées polaires

$$\begin{cases} x = r\cos\theta & \text{(16.427a)} \\ y = r\sin\theta & \text{(16.427b)} \end{cases}$$

avec $r \in]0, \infty[$ et $\theta \in [0, 2\pi[$. Le jacobien vaut r.

16.7.8.2 Coordonnées cylindriques

$$\begin{cases} x = r\cos\theta & \text{(16.428a)} \\ y = r\sin\theta & \text{(16.428b)} \\ z = z & \text{(16.428c)} \end{cases}$$

avec $r \in]0, \infty[$, $\theta \in [0, 2\pi[$ et $z \in \mathbb{R}$. Le jacobien vaut r.

16.7.8.3 Coordonnées sphériques

$$\begin{cases} x = \rho\cos\theta\sin\phi & \text{(16.429a)} \\ y = \rho\sin\theta\sin\phi & \text{(16.429b)} \\ z = \rho\cos\phi & \text{(16.429c)} \end{cases}$$

avec $\rho \in]0, \infty[$, $\theta \in [0, 2\pi[$ et $\phi \in [0, \pi[$. Le jacobien vaut $-\rho^2\sin(\phi)$.

N'oubliez pas que lorsqu'on effectue un changement de variables dans une intégrale, la *valeur absolue* du jacobien apparaît.

Cependant notre convention de coordonnées sphériques fait venir $\sin(\phi)$ avec $\phi \in [0, \pi[$; vu que le signe de $\sin(\phi)$ y est toujours positif, cette histoire de valeur absolue est sans grandes conséquent. Ce n'est pas le cas de toutes les conventions possibles.

22. Le problème $\rho = 0$ ne se pose pas ; pourquoi ?

16.7.9 une limite

Proposition 16.147.
Nous avons la limite

$$\lim_{x\to 0} \frac{\sin(x)}{x} = 1. \tag{16.430}$$

Démonstration. On commence par observer que la fonction $g(x) = \frac{\sin(x)}{x}$ est un rapport entre deux fonction impaires et est donc une fonction paire. Nous pouvons alors nous réduire à considérer le cas où x est positif. La première étape de cette preuve nous dit que $g(x) \leqslant 1$ pour tout $x \in \mathbb{R}^{+,*}$.

Nous voulons étudier le comportement de g dans un voisinage de 0. Nous pouvons alors supposer que x soit inférieur à $\pi/2$. Soit $D = (1, \tan(x))$. On voit très bien dans le dessin que l'aire du triangle de sommets O, D et C est supérieure à l'aire du secteur circulaire de sommets O, A et C. Ces deux aires peuvent être calculées très facilement et nous obtenons

$$\frac{\sin(x)}{2\cos(x)} \geqslant \frac{x}{2}.$$

À partir de cette dernière inégalité nous pouvons écrire

$$1 \geqslant \frac{\sin(x)}{x} \geqslant \cos(x).$$

En prenant la limite lorsque x tend vers 0 dans les trois membres de l'inégalité la règle de l'étau nous permet d'obtenir la limite remarquable (16.430). □

Chapitre 17

Représentations et caractères

17.1 Représentations et caractères

Définition 17.1.
Une représentation est ***fidèle*** *si elle est injective en tant que application $G \to \mathrm{GL}(V)$. Ce ne sont pas chacun des $\rho(g)$ qui doivent être injectifs. La dimension de V est le* ***degré*** *de la représentation (V, ρ).*

Définition 17.2.
Si G est un groupe, l'ensemble des homomorphismes $\mathrm{Hom}(G, \mathbb{C}^*)$ est un groupe pour la multiplication. Un élément de $\mathrm{Hom}(G, \mathbb{C}^*)$ est un ***caractère abélien****. Le nom «abélien» vient du fait que le caractère prenne ses valeurs dans $\mathbb{C}^*$. Nous notons $\hat{G} = \mathrm{Hom}(G, \mathbb{C}^*)$.*

Théorème 17.3.
Soit G un groupe abélien fini. Alors G est isomorphe à $\hat{G}$.

L'isomorphisme n'est pas canonique.

Démonstration. Étant donné la structure des groupes abéliens finis donnée par le théorème 8.24, nous commençons par nous concentrer sur $G = \mathbb{Z}/n\mathbb{Z}$. Nous allons montrer que

$$\mathrm{Hom}(\mathbb{Z}/n\mathbb{Z}) \simeq \mathrm{U}_n = \{\xi \in \mathbb{C} \text{ tel que } \xi^n = 1\}. \tag{17.1}$$

Pour cela nous avons l'isomorphisme

$$\begin{aligned} \psi\colon \mathrm{Hom}(\mathbb{Z}, \mathbb{C}^*) &\to \mathbb{C}^* \\ f &\mapsto f(1). \end{aligned} \tag{17.2}$$

Notons que si $f \in \mathrm{Hom}(\mathbb{Z}, \mathbb{C}^*)$, alors $f(k) = f(1)^k$, donc ψ est bien un isomorphisme. Cela nous amène à définir

$$\begin{aligned} \varphi\colon \mathrm{Hom}\Big((\mathbb{Z}/n\mathbb{Z}, +), (\mathbb{C}^* n\ \cdot\ )\Big) &\to \mathrm{U}_n \\ g &\mapsto f(1). \end{aligned} \tag{17.3}$$

Remarquons que pour tout $f \in \mathrm{Hom}(\mathbb{Z}/n\mathbb{Z}, \mathbb{C}^*)$ on a bien $f(1)^n = 1$. En effet si $[k] \in \mathbb{Z}/n\mathbb{Z}$, alors $f\big([k]\big) = f(1)^k$ et en particulier

$$f(1)^n = f([n]) = f(0) = 1. \tag{17.4}$$

Donc $f(1) \in \mathrm{U}_n$. Le φ est injective parce que si $f(1) = g(1)$ alors $f = g$ du fait que $f(k) = f(1)^k = g(1)^k = g(k)$.

Nous en sommes à avoir prouvé que $\widehat{\mathbb{Z}/n\mathbb{Z}} \simeq \mathrm{U}_n$. Il faudrait encore montrer que $\mathrm{U}_n \simeq \mathbb{Z}/n\mathbb{Z}$. Pour cela nous nous rappelons du lemme 3.98 nous ayant raconté que le groupe U_n des racines de l'unité était cyclique et d'ordre n. Il est donc bien isomorphe à $\mathbb{Z}/n\mathbb{Z}$.

Passons au cas où

$$G \simeq \mathbb{Z}/d_1\mathbb{Z} \times \mathbb{Z}/d_2\mathbb{Z} \times \ldots \times \mathbb{Z}/n_k\mathbb{Z}. \tag{17.5}$$

Dans ce cas nous montrons que

$$\begin{aligned} \alpha\colon & \bigtimes_{i=1}^{k} \mathrm{Hom}(\mathbb{Z}/d_i\mathbb{Z}, \mathbb{C}^*) \to \mathrm{Hom}(G, \mathbb{C}^*) \\ & \alpha(\chi_1, \ldots, \chi_k)(g_1, \ldots, g_k) = \chi_1(g_1) \ldots \chi_k(g_k). \end{aligned} \tag{17.6}$$

Ce α est injectif parce qu'en appliquant l'égalité

$$\alpha(\chi_1, \ldots, \chi_k) = \alpha(\chi'_1, \ldots, \chi'_k) \tag{17.7}$$

à l'élément $g = (9, \ldots, 1, \ldots, 0)$ alors nous trouvons $\chi_i(1) = \chi'_i(1)$ parce que $\chi_j(0) = 1$. Du coup $\chi_i = \chi'_i$.

L'application α est en plus surjective. En effet si $\chi \in \mathrm{Hom}(G, \mathbb{C}^*)$, alors nous définissons

$$\chi_i(g_i) = \chi(0, \ldots, g_i, \ldots, 0), \tag{17.8}$$

et nous avons alors $\alpha(\chi_1, \ldots, \chi_k) = \chi$.

Nous devons encore montrer que α est un homomorphisme. Si $\chi, \chi' \in \bigtimes_{i=1}^{k} \mathrm{Hom}(\mathbb{F}_{d_i}, \mathbb{C}^*)$, alors

$$\begin{aligned} \alpha(\chi\chi')(g_1, \ldots, g_k) &= (\chi_1\chi'_1)(g_1) \ldots (\chi_k\chi'_k)(g_k) && \text{(17.9a)} \\ &= \chi_1(g_1) \ldots \chi_k(g_k)\chi'_1(g_1) \ldots \chi'_k(g_k) && \text{(17.9b)} \\ &= \alpha(\chi)(g_1, \ldots, g_k)\alpha(\chi')(g_1, \ldots, g_k) && \text{(17.9c)} \\ &= \big(\alpha(\chi)\alpha(\chi')\big)(g_1, \ldots, g_k). && \text{(17.9d)} \end{aligned}$$

Donc $\alpha(\chi\chi') = \alpha(\chi)\alpha(\chi')$. □

Théorème 17.4.
Soit G un groupe abélien fini. Les groupes G et $\hat{\hat{G}}$ sont isomorphes et un isomorphisme canonique est donné par $\alpha\colon g \mapsto f_g$ donné par

$$f_g(\chi) = \chi(g). \tag{17.10}$$

Démonstration. D'abord f_g est bien un caractère de $\hat{G}$ parce que

$$f_g(\chi\chi') = (\chi\chi')(g) = \chi(g)\chi'(g) = f_g(\chi)f_g(\chi'). \tag{17.11}$$

Le fait que α soit un homomorphisme de groupes est direct :

$$f_{gg'}(\chi) = \chi(gg') = \chi(g)\chi(g') = f_g(\chi)f_{g'}(\chi) = (f_gf_{g'})(\chi). \tag{17.12}$$

D'autre part nous savon que G et $\hat{\hat{G}}$ ont le même cardinal. Il suffit donc de prouver l'injectivité de α pour être sûr de la bijectivité. Pour cela nous devons prouver que si $g \neq e$ alors $f_g \neq f_e$. Nous savons que pour tout caractère $\chi \in \hat{G}$, $f_e(\chi) = \chi(e) = 1$. Donc pour tout $g \in G\backslash\{e\}$, nous devons trouver $\chi \in \hat{G}$ tel que $\chi(g) \neq 1$.

En vertu de ce que nous connaissons sur la structure des groupes abéliens finis (théorème 8.24), nous commençons $G = \mathbb{Z}/n\mathbb{Z}$ et considérons le caractère donné par $\chi([1]) = e^{2i\pi/n}$. Ce χ est un isomorphisme entre G et $\mathrm{U}(n)$; nous n'avons $\chi([k]) = 0$ que si $[k] = [n] = [0]$. Pour rappel dans $\mathbb{Z}/n\mathbb{Z}$, le neutre est $e = 0$ et non $e = 1$.

Passons au cas général :

$$G \simeq \mathbb{Z}/n_1\mathbb{Z} \times \ldots \times \mathbb{Z}/n_k\mathbb{Z} \tag{17.13}$$

Si $g = (g_1, \ldots, g_k)$ est non nul dans G, alors il existe i tel que $g_i \neq 0$ et on prend

$$\chi(g_1, \ldots, g_k) = \chi_i(g_i) \tag{17.14}$$

où χ_i est le caractère $\chi_i([1]) = e^{2\pi i/n_i}$. Ce χ est alors un caractère non trivial de G.

□

17.1.1 Crochet de dualité et transformée de Fourier

Si G est un groupe abélien, nous définissons le crochet de dualité entre G et $\hat{G}$ par

$$\begin{aligned}\langle ., . \rangle \colon G \times \hat{G} &\to \mathbb{C}^* \\ \langle g, \chi \rangle &= \chi(g).\end{aligned} \tag{17.15}$$

Notons que l'image de ce crochet n'est pas $\mathbb{C}^*$ entier, mais seulement le groupe unitaire $\mathrm{U}(n)$ où n est l'exposant [1] de G.

Si f, g sont des applications de G dans $\mathbb{C}$, alors on leur associe le produit scalaire

$$\langle f, g \rangle = \frac{1}{|G|} \sum_{s \in G} f(s)\overline{g(s)}. \tag{17.16}$$

Lemme 17.5.
Les caractères de G forment une base orthonormée de $\mathbb{C}^G$ pour ce produit scalaire.

Démonstration. Étant donné que les $\chi(s)$ sont des nombres complexe de module 1, nous avons $\chi(s)\overline{\chi(s)} = 1$ et par conséquent $\langle \chi, \chi \rangle = 1$.

Si par contre $\chi \neq \chi'$, alors il existe $s_\in G$ tel que $\chi(s_0) \neq \chi'(s_0)$. Dans ce cas en effectuant un changement de variable $s \to s_0 s$ dans la sommation,

$$\langle \chi, \chi' \rangle = \frac{1}{|G|} \sum_{s \in G} \chi(s)\overline{\chi'(s)} \tag{17.17a}$$

$$= \frac{1}{|G|} \sum_{s \in G} \chi(s_0 s)\overline{\chi'(s_0 s)} \tag{17.17b}$$

$$= \frac{1}{|G|} \chi(s_0)\chi'(s_0) \sum_{s \in G} \chi(s)\overline{\chi'(s)}. \tag{17.17c}$$

Donc nous avons trouvé

$$\langle \chi, \chi' \rangle \big(1 - \chi(s_0)\overline{\chi'(s_0)}\big) = 0. \tag{17.18}$$

Mais vu que $\chi(s_0) \neq \chi(s_0')$, la parenthèse est non nulle (pour rappel $\chi(s_0)$ est un complexe de module 1) et par conséquent $\langle \chi, \chi' \rangle = 0$.

Nous déduisons immédiatement que les caractères forment une famille libre parce que si $\sum_i \chi_i = 0$ (la somme est sur tous les caractères), alors en prenant le produit scalaire avec χ_k,

$$\sum_i a_i \langle \chi_k, \chi_i \rangle = 0, \tag{17.19}$$

et donc $a_k = 0$.

Les caractères forment donc un système libre orthonormé. De plus l'espace engendré à la bonne dimension parce que le cardinal de l'ensemble des caractères est la dimension (complexe) de l'espace des fonction de G dans $\mathbb{C}$ parce que, en utilisant l'isomorphisme entre G et $\hat{G}$,

$$\operatorname{Card} \hat{G} = \operatorname{Card}(G) = \dim_{\mathbb{C}} \mathbb{C}^G. \tag{17.20}$$

La première □

Du fait que les caractères forment une base orthonormée, nous pouvons écrire, pour toute application $f \colon G \to \mathbb{C}$,

$$f = \sum_{\chi \in \hat{G}} \langle \chi, f \rangle \chi. \tag{17.21}$$

À une fonction $f \colon G \to \mathbb{C}$ nous associons la **transformée de Fourier**

$$\begin{aligned}\hat{f} \colon \hat{G} &\to \mathbb{C} \\ \chi &\mapsto \langle \chi, f \rangle.\end{aligned} \tag{17.22}$$

1. Définition 3.15.

Nous avons donc aussi une espèce de formule d'inversion

$$f = \sum_{\chi \in \hat{G}} \hat{f}(\chi)\chi \tag{17.23}$$

qui n'est qu'une réécriture de 17.21.

17.1.2 Groupes non abéliens

Nous avons vu que le groupe des caractères $\hat{G}$ contenait toute l'information sur un groupe abélien. Malheureusement, pour les groupes non abéliens, ça ne va pas suffire, et nous allons introduire la notion de représentations, dont les caractères seront un cas particulier de dimension un.

Proposition 17.6.
Soit G un groupe (pas spécialement abélien). Nous avons

$$\hat{G} \simeq \text{Hom}\big(G/D(G), \mathbb{C}^*\big). \tag{17.24}$$

Démonstration. Ce qui fait fonctionner la preuve est le fait que si $f\colon G \to \mathbb{C}^*$ est un homomorphisme, alors f s'annule sur $D(G)$. L'isomorphisme est

$$\begin{aligned} \psi\colon \hat{G} &\to \text{Hom}\big(G/D(G), \mathbb{C}^*\big) \\ \psi(f)[g] &= f(g). \end{aligned} \tag{17.25}$$

Cette application est bien définie parce que si f est un homomorphisme,

$$f(gklk^{-1}l^{-1}) = f(g). \tag{17.26}$$

D'autre part ψ est un homomorphisme de groupe parce que

$$\psi(f_1f_2)[g] = (f_1f_2)(g) = f_1(g)f_2(g) = \psi(f_1)[g]\psi(f_2)[g] = \big(\psi(f_1)\psi(f_2)\big)[g]. \tag{17.27}$$

Pour l'injectivité de ψ, soit f_1 et f_2 telles que $\psi(f_1) = \psi(f_2)$. Alors pour tout $g \in G$ nous avons

$$\psi(f_1)[g] = \psi(f_2)[g] \tag{17.28}$$

et donc $f_1(g) = f_2(g)$.

Enfin ψ est surjective. En effet, soit $\bar{f} \in \text{Hom}\big(G/D(G), \mathbb{C}^*\big)$. Alors nous obtenons $\psi(f) = \bar{f}$ en posant

$$f(g) = \bar{f}[g]. \tag{17.29}$$

Il faut juste vérifier que le f ainsi défini est dans $\hat{G}$, c'est à dire que $f(g_1g_2) = f(g_1)f(g_2)$. □

Cette proposition nous montre que

$$\hat{G} = \widehat{G/D(G)}, \tag{17.30}$$

alors que $G/D(G)$ est abélien ; il n'est donc pas tellement possible que $\hat{G}$ contienne beaucoup d'informations intéressantes sur G.

17.1.3 Représentations linéaires des groupes finis

Si $\dim V = 1$, alors $\text{GL}(V) = \mathbb{C}^*$ et les représentation sont les caractères abéliens.

Exemple 17.7
Considérons le triangle équilatéral A, B, C, par exemple donné par les points

$$\begin{cases} A = 1 & \text{(17.31a)} \\ B = (-\frac{1}{2}, \frac{\sqrt{3}}{2}) & \text{(17.31b)} \\ C = (-\frac{1}{2}, -\frac{\sqrt{3}}{2}) & \text{(17.31c)} \\ & \text{(17.31d)} \end{cases}$$

Dans la base (pas orthonormée) $\{A, B\}$ de $\mathbb{R}^2$, ces trois points sont donnés par

$$A = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad C = \begin{pmatrix} -1 \\ -1 \end{pmatrix}. \tag{17.32}$$

Le groupe symétrique S_3 agit sur le triangle par permutation des sommets. Vues dans la base $\{A, B\}$, les transpositions correspondent aux matrices

$$(A, B) \to \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{17.33a}$$

$$(A, C) \to \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} \tag{17.33b}$$

$$(B, C) \to \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}. \tag{17.33c}$$

La permutation (A, B, C) s'écrit comme $(A, B, C) = (A, C)(A, B)$ et on lui associe la matrice

$$(A, B, C) \to \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}. \tag{17.34}$$

C'est bien le produit des matrices de (A, C) et de (A, B). De la même façon nous avons

$$(BAC) < ++ > \tag{17.35}$$

<++> △

Si (V, ρ) et (V', ρ') sont deux représentations du groupe G, alors nous définissons la **somme directe** par $\big(V \oplus V', \rho \oplus \rho'\big)$ donné par

$$(\rho \oplus \rho')(g) = \begin{pmatrix} \rho(g) & 0 \\ 0 & \rho'(g) \end{pmatrix} \in \mathrm{GL}(V \oplus V'). \tag{17.36}$$

Nous noterons souvent $2V$ pour la représentations $(V, \rho) \oplus (V, \rho)$ et plus généralement l'écriture

$$V = \bigoplus_i k_i W_i \tag{17.37}$$

signifiera la représentation somme de k_i termes de la représentation W_i. Ici encore un abus est commis entre la représentation (ρ_i, W_i) et l'espace W_i.

17.1.4 Module

Nous considérons la $\mathbb{C}$-algèbre $G[\mathbb{C}]$ des combinaisons (formelles) d'éléments de G à coefficients dans G, c'est à dire l'ensemble

$$\mathbb{C}[G] = \{\sum_{s \in G} a_s s\} \tag{17.38}$$

avec le produit hérité de la bilinéarité :

$$\sum_{s \in G} \sum_{t \in G} a_s b_t st = \sum_s \sum_t a_s b_{s^{-1}t} t, \tag{17.39}$$

et la somme

$$(\sum_s a_s s) + \sum_t b_t t = \sum_{s \in G} (a_s + b_s) s. \tag{17.40}$$

Le tout est une $\mathbb{C}$-algèbre agissant sur V par

$$\left(\sum_s a_s s\right) v = \sum_{s \in G} a_s \rho(s) v \in V \tag{17.41}$$

Les sous-modules indécomposables seront les représentations irréductibles.

Définition 17.8.
La représentation (V, ρ) du groupe G est ***irréductible*** *si les seuls sous-espaces invariants de V sous $\rho(G)$ sont V et $\{0\}$.*

Exemple 17.9
La représentation de S_3 sur $\mathbb{R}^2$ donnée par les permutations des sommets d'un triangle équilatéral donnée dans l'exemple 17.7 est irréductible. △

La question qui vient est de savoir si une représentation possédant des sous-espaces invariants peut être écrite comme la somme de représentations irréductibles.

Proposition 17.10.
Soit (V, ρ) une représentation linéaire de dimension finie d'un groupe fini[2]*. Si W_1 est un sous-espace stable*[3]*, alors il existe un sous-espace W_2 également stable et tel que $V = W_1 \oplus W_2$.*

Toute représentation linéaire est décomposable en représentations irréductibles.

Démonstration. Soit $P\colon V \to V$ un projecteur sur W_1, c'est à dire que $P^2 = P$ et $P(V) = W_1$. Pour construire un tel projecteur, on peut par exemple prendre un supplémentaire de W_1 dans V puis utiliser la décomposition[4]. Nous considérons l'opérateur

$$P_G = \frac{1}{|G|} \sum_{g \in G} \rho(g) \circ P \circ \rho(g)^{-1}. \tag{17.42}$$

Prouvons que ce P_G est encore un projecteur. D'abord pour tout $g \in G$ nous avons

$$\rho(g) P_G \rho(g)^{-1} = \frac{1}{|G|} \sum_{s \in G} \rho(gs) P \rho(gs)^{-1} = P_G. \tag{17.43}$$

La dernière égalité est un changement de variables dans la somme[5]. Cela signifie que $P_G \rho = \rho P_G$. Nous avons même $P_G P = P$ parce que si $v \in W_1$, alors

$$P_G(v) = \frac{1}{|G|} \sum_{s \in G} \rho(s) P \underbrace{\rho(s)^{-1} v}_{\in W_1} \tag{17.44a}$$

$$= \frac{1}{|G|} \sum_{s} \rho(s) \rho(s)^{-1} v \tag{17.44b}$$

$$= v. \tag{17.44c}$$

Avec cela nous pouvons conclure que $P_G^2 = P_G$ parce que

$$P_G \circ P_G = \frac{1}{|G|} \sum_{g} P_G \rho(g) P \rho(g)^{-1} \tag{17.45a}$$

$$= \frac{1}{|G|} \sum_{g} \rho(g) P_G P \rho(g)^{-1} \tag{17.45b}$$

$$= \frac{1}{|G|} \sum_{g} \rho(g) P \rho(g)^{-1} \tag{17.45c}$$

$$= P_G. \tag{17.45d}$$

Donc P_G est un projecteur, est stable sous les conjugaisons par $\rho(g)$ et commute avec $\rho(g)$. Nous décomposant Id de façon évidente en

$$\mathrm{Id} = P_G + (\mathrm{Id} - P_G). \tag{17.46}$$

2. La démonstration marche aussi pour les groupes compacts, mais il faudrait des intégrales.

3. c'est à dire si ρ n'est pas irréductible.

4. Ou encore prendre une base de W_1, l'étendre en une base de V et définir P comme l'annulation des coefficients des vecteurs «complétant» la base.

5. Et c'est ça qui demande un peu de technique pour écrire la preuve dans le cas d'un groupe compact : il faut une mesure de Haar.

Étant donné que l'opérateur P_G commute avec tous les $\rho(g)$, les noyaux de P_G et $\mathrm{Id} - P_G$ sont des sous-espaces invariants. Vu que P_G est un projecteur, nous avons $q(P_G) = 0$ avec $q(X) = X^2 - X$. Pour appliquer le lemme des noyaux (théorème 9.102), nous remarquons que $q(X) = X(X-1)$ et donc

$$V = \ker P_G \oplus \ker(P_G - \mathbb{1}). \tag{17.47}$$

Si nous posons $W_2 = \ker P_G$, il reste à voir que $\ker(P_G\mathbb{1}) = W_1$. D'abord $W_1 \subset \ker(P_G - \mathrm{Id})$ parce que si $w \in W_1$, ce dernier étant stable,

$$P_G w = \frac{1}{|G|} \sum_{g\in G} \rho(g) P \underbrace{\rho(g)^{-1} w}_{\in W_1} \tag{17.48a}$$

$$= \frac{1}{|G|} \sum_{g\in G} w \tag{17.48b}$$

$$= w. \tag{17.48c}$$

Pour prouver l'inclusion inverse, nous savons que P_G et P sont des projecteurs tels que $P_G P = P$, ce qui signifie que l'image de P_G est inclue à celle de P, c'est à dire à W_1. Mais $\mathrm{Image}(P_G) = \ker(\mathbb{1} - P_G)$, donc

$$\ker(\mathbb{1} - P_G) = \mathrm{Image}(P_G) \subset \mathrm{Image}(P) = W_1. \tag{17.49}$$

La représentation ρ se décompose donc en deux sous-représentations (ρ, W_1) et ρ, W_2. Si l'une des deux n'est pas irréductible, le processus peut recommencer. Vu que la dimension de V est finie, toute représentation se décompose en une somme finie de représentation irréductibles. □

17.1.5 Structure hermitienne

Soit (ρ, V) une représentation de G sur un espace vectoriel complexe V. Nous voulons munir V d'un produit scalaire hermitien (définition 9.41) tel que les opérateurs $\rho(g)$ soient tous des isométries. C'est à dire que nous voudrions définir $\langle u, v\rangle_G$ de telle sorte à avoir

$$\langle \rho(g)u, \rho(g)v\rangle_G = \langle u, v\rangle_G \tag{17.50}$$

pour tout $g \in G$. Nous commençons par considérer un produit hermitien $\langle ., .\rangle$ quelconque et puis nous définissons

$$\langle u, v\rangle_G = \frac{1}{|G|} \sum_{g\in G} \langle \rho(g)u, \rho(g)v\rangle. \tag{17.51}$$

Nous devons vérifier que c'est un produit. La seule des conditions dont la vérification n'est pas immédiate est celle de positivité. Pour tout $g \in G$ et tout $v \in V$, nous avons $\langle \rho(g)v, \rho(g)v\rangle$ est positif et nul si et seulement si $\rho(g)v = 0$. Étant donné que $\rho(e)v = v$, parmi les termes de la somme

$$\langle u, u\rangle_G = \frac{1}{|G|} \sum_{g\in G} \langle \rho(g)v, \rho(g)v\rangle, \tag{17.52}$$

au moins un est strictement positif (pourvu que $v \neq 0$) ; les autres sont positifs ou nuls. Par conséquent $\langle v, v\rangle_G = 0$ si et seulement si $v = 0$.

Donc les groupes finis peuvent être vus comme des parties de groupes d'isométrie. De la même façon, en utilisant une mesure de Haar pour faire la moyenne, nous pouvons plonger les groupes compacts dans des groupes unitaires.

17.1.6 Caractères

Définition 17.11.
Soit (V, ρ) une représentation linéaire du groupe G. Le ***caractère*** *de ρ est la fonction*

$$\begin{aligned} \chi_\rho \colon G &\to \mathbb{C} \\ s &\mapsto \mathrm{Tr}\big(\rho(s)\big). \end{aligned} \tag{17.53}$$

Par invariance cyclique de la trace, nous avons

$$\chi_\rho(sts^{-1}) = \chi_\rho(t), \tag{17.54}$$

ce qui fait que le caractère est une fonction constante sur les classes de conjugaison.

D'après sa fiche wikipédia, le marquis de Sade, passionné de théâtre, faisait des représentations qui avaient du caractère.

Un **caractère irréductible** est un caractère d'une représentation irréductible.

Définition 17.12.
*Une application $f\colon G \to \mathbb{C}$ est **centrale** si elle est constante sur les classes de conjugaison.*

Les traces sont des applications centrales.

L'ensemble des fonctions centrales sur un groupe fini (ou tout au moins ayant un nombre fini de classes de conjugaison) est un $\mathbb{C}$-espace vectoriel de dimension égale au nombre de classes, et nous pouvons mettre le produit scalaire

$$\langle f, g\rangle = \frac{1}{|G|}\sum_{s\in G} f(s)\overline{g(s)}. \tag{17.55}$$

C'est une forme hermitienne sur l'espace des fonction centrales.

17.2 Équivalence de représentations et caractères

Cette section prend des éléments des articles lemme de Schur, caractère d'une représentation, fonction centrale et trace de wikipédia.

Nous disons que les deux représentations (V, ρ) et (V', ρ') sont **équivalentes** s'il existe une bijection linéaire $f\colon V \to V'$ telle que

$$f \circ \rho = \rho' \circ f. \tag{17.56}$$

Nous disons alors que f **entrelace** ρ et ρ'.

Théorème 17.13 (Théorème de Schur)**.**
Si (V, ρ) et (V', ρ') sont des représentations irréductibles non équivalentes alors la seule application linéaire $f\colon V \to V'$ entrelaçant ρ et ρ' est la fonction nulle.

En d'autres termes, soit les représentations sont équivalentes (et il y a un isomorphisme), soit il n'y a même pas un homomorphisme.

Démonstration. Soit $f \in \mathcal{L}(V, V')$ telle que $f \circ \rho = \rho' \circ f$. Alors $\ker f$ est un sous-espace stable sous $\rho(G)$, et $\text{Image}(f)$ est un sous-espace de V' stable par $\rho'(G)$. Par irréductibilité, nous avons que $\ker(f) = \{0\}$ ou V. Même chose pour $\text{Image}(f)$. Il y a deux possibilités.

(1) Si $\ker(f) = \{0\}$, alors $\text{Image}(f) \neq \{0\}$ et alors $\text{Image}(f) = V'$. Du coup f est injective et surjective, c'est à dire est un isomorphisme.
(2) Si $\ker(f) = V$, alors $f = 0$.

□

Corollaire 17.14 (Schur pour les représentations sur $\mathbb{C}$)**.**
Soit (V, ρ) une représentation irréductible, alors l'ensemble

$$\text{End}_G(V, \rho) = \{f \in \text{End}(V) \text{ tel que } \rho \circ f = f \circ \rho\} \tag{17.57}$$

est l'ensemble des homothéties.

Démonstration. Soit $f \in \text{End}_G(V, \rho)$. Vu que l'espace est sur $\mathbb{C}$, l'endomorphisme f a une valeur propre λ. L'opérateur $g = f - \lambda\mathbb{1}$ est aussi un opérateur d'entrelacement de ρ alors que $\ker(g) \neq \{0\}$ par définition de valeur propre. Du coup $\ker(g) = V$, ce qui signifie que f est l'isométrie de rapport λ : $f = \lambda\,\text{Id}$. □

Lemme 17.15.
Si (ρ, V) et (ρ', V') sont des représentations équivalentes de caractères χ et χ', alors $\chi = \chi'$.

Démonstration. Si $A\colon V \to V'$ est un isomorphisme d'espace vectoriel entrelaçant ρ et ρ', c'est à dire si pour tout g, $\rho'(g)A = A\rho(g)$, alors $\rho'(g) = A\rho(g)A^{-1}$ et

$$\chi'(g) = \mathrm{Tr}\big(\rho'(g)\big) = \mathrm{Tr}\big(A\rho(g)A^{-1}\big) = \mathrm{Tr}\big(\rho(g)\big) \tag{17.58}$$

parce que la trace est un invariant de similitude (lemme 9.135). □

Lemme 17.16.
Si χ est le caractère de la représentation complexe (V, ρ) du groupe fini G, alors pour tout $g \in G$ nous avons $\chi(g^{-1}) = \overline{\chi(g)}$.

Démonstration. Par le corollaire 3.55 au théorème de Lagrange, nous avons $g^{|G|} = e$ et donc en tant qu'opérateur, $\rho(g)^{|G|} = \mathbb{1}$. Les valeurs propres de $\rho(g)$ sont donc des racines de l'unité. Si nous notons λ_i ces valeurs propres, alors $\chi(g) = \sum_i \lambda_i$, et en considérant la matrice dans sa base de diagonalisation (lemme de Schur complexe, 9.152), nous voyons que

$$\chi(g^{-1}) = \mathrm{Tr}\big(\rho(g)^{-1}\big) = \sum_i \frac{1}{\lambda_i}. \tag{17.59}$$

Mais λ_i étant une racine de l'unité nous avons $\frac{1}{\lambda_i} = \bar{\lambda}_i$, ce qui fait que

$$\chi(g^{-1}) = \sum_i \bar{\lambda}_i = \overline{\chi(g)}. \tag{17.60}$$

□

Proposition 17.17.
Soient deux représentations irréductibles complexes (V, ρ) et (V', ρ') du même groupe fini G, et χ et χ' leurs caractères respectifs. Nous avons

(1) $\langle \chi, \chi' \rangle = 0$ si ρ et ρ' ne sont pas équivalentes.

(2) $\langle \chi, \chi' \rangle = 1$ si les représentations sont équivalentes.

Démonstration. Nous considérons les bases $\{e_1, \dots, e_n\}$ de V et $\{f_1, \dots, f_m\}$ de V'. Puis nous considérons la matrice $F(k,l) = E_{kl} \in \mathbb{M}_{m,n}(\mathbb{C})$ où pour rappel, E_{kl} est la matrice de composantes $(E_{kl})_{ij} = \delta_{ki}\delta_{lj}$. Nous posons

$$F_G(k,l) = \frac{1}{|G|} \sum_{g \in G} \rho(g) \circ F(k,l) \circ \rho'(g)^{-1}. \tag{17.61}$$

En nous permettant de ne pas réécrire les indices k et l de F et F_G, nous montrons que F_G entrelace ρ et ρ' :

$$F_G \circ \rho'(t) = \frac{1}{|G|} \sum_{s \in G} \rho(s) \circ F \circ \rho'(s^{-1}) \circ \rho(t) \tag{17.62a}$$

$$= \frac{1}{|G|} \sum_{s} \rho(s) F \rho'(s^{-1}t) \tag{17.62b}$$

$$= \frac{1}{|G|} \sum_{k} \rho(tk) F \rho'(k^{-1}) \tag{17.62c}$$

$$= \frac{1}{|G|} \rho(t) \sum_{k} \rho(k) F \rho'(k^{-1}) \tag{17.62d}$$

$$= \rho(t) \circ F_G. \tag{17.62e}$$

Dans ce calcul nous avons effectué le changement de variables $k = (s^{-1}t)^{-1}$ qui donne $s = tk$.

Par ailleurs nous avons

$$\Big(\rho(g)F(k,l)\rho'(g^{-1})\Big)_{ij} = \sum_{r=1}^{n}\sum_{s=1}^{m}\rho(g)_{ir}F(k,l)_{rs}\rho'(g^{-1})_{sj} \tag{17.63a}$$

$$= \sum_{rs}\rho(g)_{ir}\delta_{kr}\delta_{ls}\rho'(g^{-1})_{sj} \tag{17.63b}$$

$$= \rho(g)_{ik}\rho'(g^{-1})_{lj}, \tag{17.63c}$$

et par conséquent

$$F_G(k,l)_{ij} = \frac{1}{|G|}\sum_{g\in G}\rho(g)_{ik}\rho'(g^{-1})_{lj}. \tag{17.64}$$

Si χ et χ' sont les caractères de ρ et ρ', alors nous avons le produit (17.55) qui donne

$$\langle\chi,\chi'\rangle = \frac{1}{|G|}\sum_{g\in G}\chi(g)\overline{\chi'(g)} \tag{17.65a}$$

$$= \frac{1}{|G|}\sum_{g}\chi(g)\chi'(g^{-1}) \qquad \text{lemme 17.16} \tag{17.65b}$$

$$= \frac{1}{|G|}\sum_{g}\sum_{i=1}^{n}\sum_{j=1}^{m}\rho(g)_{ii}\rho'(g^{-1})_{jj} \tag{17.65c}$$

$$= \sum_{ij}F_G(i,j)_{ij} \qquad \text{par (17.64).} \tag{17.65d}$$

Si les représentations ρ et ρ' ne sont pas équivalentes, le fait que F_G en soit un opérateur d'entrelacement implique par le théorème de Schur 17.13 que $F_G = 0$ et donc $\langle\chi,\chi'\rangle = 0$.

Si au contraire les représentation sont équivalentes, alors le lemme 17.15 nous dit que $\chi = \chi'$ et nous reprenons la définition :

$$\langle\chi,\chi\rangle = \frac{1}{|G|}\sum_{g}\chi(g)\overline{\chi(g)} = \frac{1}{|G|}\sum_{g\in G}1 = 1 \tag{17.66}$$

parce que les nombres $\chi(g)$ sont des racines de l'unité. □

17.2.1 Représentation régulière

Nous notons λ la **représentation régulière gauche**, agissant sur le $\mathbb{K}$-espace vectoriel des fonctions $G \to \mathbb{K}$ par

$$\Big(\lambda(g)f\Big)(g) = f(g^{-1}h). \tag{17.67}$$

D'autre part nous considérons les fonction $\delta_g\colon G \to \mathbb{K}$ (ici $\mathbb{K}$ est $\mathbb{R}$ ou $\mathbb{C}$ ou pire) définie par

$$\delta_g(h) = \begin{cases} 1 & \text{si } g = h \\ 0 & \text{sinon.} \end{cases} \tag{17.68}$$

La représentation régulière agit sur les fonctions δ_s de la façon suivante :

$$\lambda(g)\delta_s = \delta_{gs} \tag{17.69}$$

parce que $\big(\lambda(g)\delta_s\big)(h) = \delta_s(g^{-1}h) = \delta_{gs}(h)$.

Lemme 17.18.
Le caractère de la représentation régulière gauche est donné par

$$\chi_\lambda = |G|\delta_e. \tag{17.70}$$

Démonstration. Appliquer l'équation (17.70) fonctionne parce que $\chi_\lambda(e)$ est la dimension de l'espace des fonctions sur G, c'est à dire $|G|$. Si par contre $g \neq e$, alors $\lambda(g)$ est une matrice de permutation (dans la base des δ_h) et a donc tous ses éléments diagonaux nuls. □

Si ρ est une représentation et si f est une fonction sur le groupe, alors nous considérons l'opérateur

$$\rho_f = \sum_{g\in G} f(g)\rho(g). \tag{17.71}$$

Proposition 17.19 ([210])**.**
Si (ρ, V) est une représentation irréductible et si f est une fonction centrale sur G, alors l'opérateur ρ_f est une homothétie de V de rapport

$$\frac{1}{\dim V}\sum_{g\in G} f(g)\chi(g) \tag{17.72}$$

où χ est le caractère de ρ.

Démonstration. Nous commençons par voir que ρ_f entrelace ρ. En effet,

$$\rho(t)^{-1}\circ\rho_f\circ\rho(t) = \sum_g f(g)\rho(t^{-1}gt) \tag{17.73a}$$
$$= \sum_h f(tht^{-1})\rho(g) \qquad h = t^{-1}gt \tag{17.73b}$$
$$= \sum_h f(h)\rho(h) \tag{17.73c}$$
$$= \rho_f \tag{17.73d}$$

où en écrivant $f(tht^{-1}) = f(h)$, nous avons utilisé le fait que f était centrale. Étant donné que ρ_f entrelace une représentation irréductible, le lemme de Schur (17.13) nous indique que ρ_f est une homothétie. Soit k le facteur d'homothétie. Alors d'une part $\mathrm{Tr}(\rho_f) = nk$. D'autre part,

$$\mathrm{Tr}(\rho_f) = \mathrm{Tr}\big(\sum_g f(g)\rho(g)\big) \tag{17.74a}$$
$$= \sum_g f(g)\,\mathrm{Tr}\big(\rho(g)\big) \tag{17.74b}$$
$$= \sum_g f(g)\chi(g). \tag{17.74c}$$

Du coup effectivement

$$k = \frac{1}{n}\sum_{g\in G} f(g)\chi(g). \tag{17.75}$$

□

17.2.2 Caractères et représentations : suite et fin

Lemme 17.20.
Un groupe fini n'a (à équivalence près) qu'un nombre fini de représentations irréductibles.

Démonstration. Les caractères irréductibles forment un système orthonormé (proposition 17.17) et donc libre parmi les fonctions centrales. Donc il y a au plus autant de caractères irréductibles que la dimension de l'espace des fonctions centrales ; et ce dernier est de dimension finie donnée par le nombre de classes de conjugaison de G. □

Nous savons que les caractères de deux représentations irréductibles sont égaux. Étant donné qu'il n'existe qu'un nombre fini de représentations irréductibles, il existe un nombre fini de caractères irréductibles. Nous pouvons donc fixer les notations suivantes. Les caractères irréductibles seront notés $\{\varphi_i\}_{i=1,\dots,h}$ et nous noterons (σ_i, W_i) une représentation ayant le caractère φ_i.

Théorème 17.21 ([210])**.**
Soit (ρ, V) une représentation de G de caractère χ. Alors sa décomposition en représentations irréductibles est donnée par

$$(V, \rho) = \bigoplus_{i=1}^{h} k_i(W_i, \sigma_i) \tag{17.76}$$

avec $k_i = \langle \chi, \varphi_i \rangle$. En particulier, à permutation près des facteurs, la décomposition d'une représentation en représentations irréductibles est unique.

Démonstration. La décomposition de χ en caractères irréductibles est donnée par $\chi = \sum_i k_i \varphi_i$; en prenant le produit de cette égalité avec φ_j et en tenant compte de l'othonormalité des caractères irréductibles,

$$\langle \chi, \varphi_j \rangle = \sum_i k_i \langle \varphi_i, \varphi_j \rangle = k_j. \tag{17.77}$$

□

Le théorème suivant est ce qui nous permet de dire que l'étude des caractères et l'étude des représentations, c'est la même chose.

Théorème 17.22.
Soit G un groupe fini[6]*.*

(1) Deux représentations sont équivalentes si et seulement si elles ont même caractères.

(2) Si χ est le caractère d'une représentation, alors

(a) $\langle \chi, \chi \rangle \in \mathbb{N}$

(b) $\langle \chi, \chi \rangle = 1$ si et seulement si la représentation est irréductible.

Démonstration. Nous démontrons chaque point séparément.

(1) Le fait que deux représentations équivalentes aient même caractère est le lemme 17.15. Nous montrons l'autre sens. Si (ρ, V) et (ρ', V') sont deux représentations irréductibles de décompositions

$$V = \bigoplus_i k_i W_i \tag{17.78a}$$

$$V' = \bigoplus_i k'_i W_i, \tag{17.78b}$$

alors si $\chi = \chi'$, nous avons $k_i = k'_i$ et les représentations sont identiques.

(2) Soit (ρ, V) une représentation ayant χ comme caractère. En posant $k_i = \langle \chi, \varphi_i \rangle$ nous avons la décomposition en représentations irréductibles

$$V = \bigoplus_i k_i W_i, \tag{17.79}$$

et aussi

$$\langle \chi, \chi \rangle = \langle \sum_i k_i \varphi_i, \sum_j k_j \varphi_j \rangle = \sum_i k_i^2 \in \mathbb{N}. \tag{17.80}$$

Ce nombre est de plus égal à 1 si et seulement si tous les termes de la somme sont nuls sauf un qui vaudrait 1. Ce cas donne une représentation irréductible.

□

Proposition 17.23.
Si (λ, R) est la représentation régulière gauche de décomposition en représentations irréductibles

$$R = \bigoplus_i k_i W_i, \tag{17.81}$$

alors

6. Nous sommes depuis longtemps dans l'étude des représentations des groupes finis.

(1) $k_i = \dim W_i$,

(2) $\sum_i (\dim W_i)^2 = |G|$,

(3) pour tout $g \in G$, $\sum_i (\dim W_i)\varphi_i(g) = 0$ [7].

(4) Si $\{(n_i, \varphi_i)\}$ *est la liste des couples dimension,caractère des représentations irréductibles non équivalentes, alors pour tout* $s \in G\backslash\{e\}$ *nous avons* $\sum_{i=1}^{p} n_i\varphi_i(s) = 0$ *où la somme porte sur les représentations irréductibles non équivalentes.*

Démonstration. Nous notons r le caractère de la représentation régulière gauche. Nous avons

$$k_i = \langle r, \varphi_i \rangle = \frac{1}{|G|} \sum_{s\in G} r(s)\overline{\varphi_i(s)} = \overline{\varphi_i(e)}. \tag{17.82}$$

Mais $\varphi_i(e) = \dim W_i \in \mathbb{R}$, donc nous avons bien $k_i = \dim W_i$. Le caractère de la représentation régulière peut alors s'exprimer de deux façons :

$$|G|\delta_e = \sum_i (\dim W_i)\varphi_i. \tag{17.83}$$

En évaluant cette égalité en e nous trouvons directement

$$|G| = \sum_i (\dim W_i)^2, \tag{17.84}$$

et en l'évaluant en $s \neq e$, nous trouvons

$$0 = \sum_i (\dim W_i)\varphi_i(s). \tag{17.85}$$

□

Le théorème suivant est valable pour les groupes finis (comme toute cette section).

Théorème 17.24 ([210])**.**
Les caractères irréductibles $\chi_1, \dots, \chi_h$ *forment une base orthonormé des fonctions centrales sur* G.

Démonstration. Nous savons déjà qu'ils forment un système orthonormé. Considérons le sous-espace $H = \text{Span}\{\varphi_i\}_{i=1,\dots,h}$ de l'espace des fonctions centrales sur G. En vertu de la proposition 7.41, il nous suffit de prouver que $H^\perp = 0$. Soit donc f, une fonction centrale appartenant à $H^\perp$. Pour tout i, nous avons $\langle f, \varphi_i \rangle = 0$ et donc aussi $\langle \bar{f}, \bar{\varphi}_i \rangle = 0$.

Considérant une représentation irréductible (σ, W) de caractère φ, nous savons par la proposition 17.19 que l'opérateur

$$\sigma_{\bar{f}} = \sum_g \bar{f}(g)\varphi(g) \tag{17.86}$$

est une homothétie de rapport $\langle \bar{f}, \bar{\varphi} \rangle / \dim W = 0$. Étant donné que toute les représentations sont des sommes directes de représentations irréductibles, en réalité l'opérateur $\rho_{\bar{f}}$ est nul pour toute représentation ρ. En particulier pour la représentation régulière,

$$0 = \lambda_{\bar{f}}(\delta_t) = \sum_{g\in G} \bar{f}(g)\lambda(g)(\delta_t) = \sum_g \bar{f}(g)\delta_{ft}. \tag{17.87}$$

En écrivant cette égalité avec $t = e$ et puis en appliquant à $k \in G$ nous trouvons

$$0 = \sum_g \bar{f}(g)\delta_g(k) = \bar{f}(k). \tag{17.88}$$

Donc $\bar{f} = 0$ et f est nulle. □

7. Cette propriété est appelée «orthogonalité des colonnes» pour une raison qui apparaîtra au moment de compléter le tableau (17.110).

Corollaire 17.25.
Le nombre de représentations irréductibles non équivalentes d'un groupe fini est égal à son nombre de classes de conjugaison.

Démonstration. Le nombre de classes de conjugaison est la dimension de l'espace des fonctions centrales qui elle-même est égale au nombre de caractères irréductibles par le théorème 17.24. Enfin deux caractères irréductibles sont égaux si et seulement si les représentations sous-jacentes sont équivalentes. □

Corollaire 17.26.
Toutes les représentations irréductibles d'un groupe abélien sont de dimension 1.

Démonstration. Le corollaire 17.25 nous dit qu'il y a autant de représentations unitaires qu'il n'y a de représenations irréductibles (non équivalentes). Mais les classes de conjugaisons sont des singletons (lemme 3.18). Nous avons donc exactement $|G|$ représentations irréductibles lorsque G est abélien.

Mais d'autre part la proposition 17.23(2) donne $\sum_i (\dim W_i)^2 = |G|$ lorsque la somme parcours les représentations irréductibles. Il y a $|G|$ termes à la somme, donc tous les termes doivent être 1. □

17.3 Représentation produit tensoriel

Soient ρ et ϕ, deux représentations d'un groupe G sur des espaces vectoriels V et W. La représentation **produit tensoriel** est la représentation

$$\begin{aligned} \rho\otimes\phi\colon G &\to \mathrm{GL}(V\otimes W)\\ (\rho\otimes\phi)(g)(v\otimes w) &= \rho(g)v\otimes\phi(g)w. \end{aligned} \tag{17.89}$$

Pour trouver son caractère, nous considérons une base $\{e_i\}$ de V et une base $\{e_\alpha\}$ de W, et la base $\{e_i\otimes e_\alpha\}$ de $V\otimes W$. Donc

$$(\rho\otimes\phi)(g)(e_i\otimes e_\alpha) = \rho(g)e_i\otimes\phi(g)e_\alpha. \tag{17.90}$$

Nous devons savoir quelle est la composante «$e_i\otimes e_\alpha$» de cette dernière expression, et c'est évidemment

$$\rho(g)_{ii}\rho_{\alpha\alpha}, \tag{17.91}$$

ce qui nous amène à dire que

$$\mathrm{Tr}(\rho\otimes\phi)(g) = \sum_i\sum_\alpha \rho(g)_{ii}\phi(g)_{\alpha\alpha} = \mathrm{Tr}\big(\rho(g)\big)\,\mathrm{Tr}\big(\phi(g)\big), \tag{17.92}$$

c'est à dire au final que

$$\chi_{\rho\otimes\phi} = \chi_\rho\chi_\phi. \tag{17.93}$$

17.4 Exemple sur le groupe symétrique

Soit $G = S_3$, un des premiers groupes finis non abéliens. On en a une représentation de dimension deux en tant que permutation des sommets d'un triangle équilatéral, donnée dans l'exemple 17.7 ; nous notons ρ cette représentation.

Nous y avons aussi la représentation de signature donnée par

$$\begin{aligned} \epsilon\colon S_3 &\to \mathrm{GL}(\mathbb{C})\\ \sigma &\mapsto \epsilon(\sigma)\,\mathrm{Id}\,. \end{aligned} \tag{17.94}$$

Et enfin il y a la représentation triviale. Ce sont les trois représentations irréductibles ; pour rappel il y a autant de représentations irréductibles que de classes de conjugaison (corollaire 17.25).

Classe de conjugaison	taille	χ_1	χ_ϵ	χ_ρ
Id	1	1	1	2
(A, B)	3	1	-1	0
(A, B, C)	2	1	1	-1

Nous calculons par exemple le produit scalaire

$$\langle \chi_1, \chi_\epsilon \rangle = \frac{1}{6}\big(1 \cdot \chi_1(\mathrm{Id})\overline{\chi_\epsilon(\mathrm{Id})} + 3 \cdot \chi_1(A,B)\overline{\chi_\epsilon(A,B)} + 2 \cdot \chi_1(A,B,C)\overline{\chi_\epsilon(A,B,C)}\big) \tag{17.95a}$$

$$= 0. \tag{17.95b}$$

D'autre part nous avons aussi

$$\langle \chi_\rho, \chi_\rho \rangle = \frac{1}{6}(1 \cdot 2 \cdot 2 + 3 \cdot 0 + 2 \cdot 1) = 1. \tag{17.96}$$

17.5 Table des caractères du groupe symétrique S_4

Pour la table des caractères de S_4, voir [2].

17.5.1 Calculs à partir de rien ou presque

Nous savons que les classes de conjugaison dans S_4 sont caractérisées par la structure des décompositions en cycles (proposition 8.34). Elles sont données dans l'exemple 8.36.

Nous avons donc 5 classes de conjugaison, et il nous faut donc 5 représentations irréductibles non équivalentes (corollaire 17.25) dont nous allons chercher les caractères.

La première est la représentation triviale de dimension 1 ; nous notons χ_1 son caractère et nous avons la ligne

	dimension	Id	(12)	(123)	(1234)	(12)(34)
χ_1	1	1	1	1	1	1

(17.97)

Ensuite nous avons la signature qui est un morphisme non trivial $\epsilon\colon S_n \to \{-1, 1\}$. Nous avons alors la ligne

	dimension	Id	(12)	(123)	(1234)	(12)(34)
χ_ϵ	1	1	-1	1	-1	1

(17.98)

Une troisième représentation pas trop compliquée à trouver est celle

$$\begin{aligned} \rho_p &\colon S_4 \to \mathrm{GL}(4, \mathbb{C}) \\ \rho_p(\sigma)e_i &= e_{\sigma(i)}. \end{aligned} \tag{17.99}$$

Cela n'est pas une représentation irréductible parce que $\mathbb{C}^4$ se décompose en deux sous-espaces stables :

$$D = \mathrm{Span}(1, 1, 1, 1) \tag{17.100a}$$

$$H = \{x \in \mathbb{C}^4 \text{ tel que } x_1 + x_2 + x_3 + x_4 = 0\}. \tag{17.100b}$$

La représentation induite sur D est la représentation triviale. Puis sur H, elle induit une autre représentations que nous allons noter ρ_s.

Nous allons à présent déduire le caractère de la représentation ρ_s et prouver qu'elle est irréductible. Il est cependant possible de sauter cette étape en échange d'un certain travail sur les isométries du tétraèdre. Voir la proposition 16.17 et ensuite

Nous avons la décomposition $\rho_p = \rho_1 \oplus \rho_s$ et donc

$$\chi_p = \chi_1 + \chi_s. \tag{17.101}$$

Nous savons déjà χ_1. Le caractère χ_p n'est pas très compliqué parce que $\chi_p(\sigma)$ est une matrice de permutation des vecteurs de base. Donc la matrice $\rho_p(\sigma)$ a un 1 sur la diagonale pour les i tels que $\sigma(i) = i$. Nous avons donc

$$\chi_p(\mathrm{Id}) = 4 \qquad \chi_p(12) = 2 \tag{17.102a}$$
$$\chi_p\big((12)(34)\big) = 0 \qquad \chi_p(123) = 1 \tag{17.102b}$$
$$\chi_p(1234) = 0. \tag{17.102c}$$

Le caractère χ_s peut être calculé par simple soustraction :

	dimension	Id	(12)	(123)	(1234)	(12)(34)
χ_s	3	3	1	0	-1	-1

(17.103)

Avant d'ajouter cette ligne au tableau des représentations irréductibles nous devons savoir si ρ_s en est une. Pour cela, tant que nous avons son caractère nous pouvons utiliser le critère du théorème 17.22 :

$$\langle \chi_s, \chi_s \rangle = \frac{1}{|S_4|} \sum_{\sigma \in S_4} \chi_s(\sigma)^2. \tag{17.104}$$

Nous avons tout de suite $|S_4| = 4 \cdot 3 \cdot 2 = 24$ et puis

$$24\langle \chi_s, \chi_s \rangle = 3^2 + 6 \cdot 1^2 + 8 \cdot 0^2 + 6 \cdot (-1)^2 + 3 \cdot (-1)^2 = 24, \tag{17.105}$$

donc oui, le caractère est irréductible parce que $\langle \chi_s, \chi_s \rangle = 1$. Et nous pouvons donc ajouter la ligne (17.103) à notre tableau. Par ailleurs, nous notons qu'elle est de dimension 3.

Pour le reste nous savons qu'il y a autant de représentations irréductibles que de classes de conjugaison, de telle sorte qu'il ne manque que deux représentations irréductibles. De plus la proposition 17.23 nous dit que si n_i est la dimension de la i^{e} représentation irréductible, alors

$$|S_4| = \sum_i n_i^2. \tag{17.106}$$

Dans notre situation, si nous nommons n_1 et n_2 les dimensions des deux représentations qui nous manquent, nous avons $24 = n_1^2 + n_2^2 + (1^2 + 1^2 + 3^2)$, c'est à dire $n_1^2 + n_2^2 = 13$. Il n'y a pas des tonnes de sommes de deux carrés qui font 13. Il y a $n_1 = 2$ et $n_2 = 3$, et c'est tout.

Nous recherchons donc encore une représentation de dimension 2 et une de dimension 3. Pour cela nous allons un peu regarder les produits tensoriels qui s'offrent à nous. Pour faire une dimension 3, il faut faire le produit d'une de dimension 1 par une de dimension 3. Là encore le choix est très limité et nous demande d'essayer

$$\rho_W = \rho_s \otimes \rho_\epsilon \tag{17.107}$$

qui agit sur l'espace $V_2 \otimes V_\epsilon$ par

$$\rho_W(g)(v \otimes x) = \rho_s(g)v \otimes \rho_\epsilon(g)x. \tag{17.108}$$

Pour savoir son caractère nous utilisons la petite formule toute simple (17.93) : nous multiplions case par case les tableaux (17.103) et (17.98) :

	dimension	Id	(12)	(123)	(1234)	(12)(34)
χ_W	3	3	-1	0	1	-1

(17.109)

Avant de réellement ajouter cette ligne au tableau, nous devons nous assurer qu'elle est bien irréductible. Nous utilisons le même critère : $\langle \chi_W, \chi_W \rangle = 1$, donc c'est bon.

Pour trouver le dernier caractère, que nous nommerons χ_u, il ne faut pas beaucoup d'imagination. Il suffit d'utiliser les relations d'orthogonalité du théorème 17.24, en sachant que la dimension est 2 et qu'alors $\chi_W(\mathrm{Id}) = 2$, c'est pas trop compliqué :

	dimension	Id	(12)	(123)	(1234)	(12)(34)
χ_1	1	1	1	1	1	1
χ_ϵ	1	1	-1	1	-1	1
χ_s	3	3	1	0	-1	-1
χ_W	3	3	-1	0	1	-1
χ_u	2	2	b	c	d	e

(17.110)

Les relations d'orthogonalité des colonnes de la propriété 17.23 nous permettent de calculer les coefficients manquants. En pratique, il suffit de prendre le produit scalaire de chaque ligne avec la première et d'égaler avec zéro. Nous trouvons $b = 0$, $c = 1$, $d = 0$, et $e = 2$. Le tableau final est :

	dimension	Id	(12)	(123)	(1234)	(12)(34)
χ_1	1	1	1	1	1	1
χ_ϵ	1	1	-1	1	-1	1
χ_s	3	3	1	0	-1	-1
χ_W	3	3	-1	0	1	-1
χ_u	2	2	0	-1	0	2

(17.111)

Notons que nous sommes parvenus à remplir la dernière ligne sans rien savoir de la représentation qui va avec. Nous allons cependant donner une interprétation géométrique et fixer cette représentation comme agissant sur le triangle équilatéral en 17.30.

17.5.2 Représentation de S_4 via les isométries du tétraèdre

Une des représentations trouvées (la représentation ρ_s) peut être vue comme le groupe $\mathrm{Iso}(T)$ des isométries affine du tétraèdre. Nous avons vu en la proposition 16.17 qu'il existe un isomorphisme de groupe $S_4 \simeq \mathrm{Iso}(T)$ lorsque T est un tétraèdre régulier de $\mathbb{R}^3$.

Si le barycentre de T est situé à l'origine de $\mathbb{R}^3$, alors les éléments de $\mathrm{Iso}(T)$ sont des applications linéaires parce que

— les affinités laissent invariantes les barycentres (proposition 11.37),
— les affinités qui laissent l'origine invariante sont linéaires (corollaire 11.49).

Nous allons à présent calculer la trace de cette représentation, en utilisant le fait que nous la connaissions explicitement. Nous savons que les caractères sont constants sur les classes de conjugaison ; nous allons donc écrire une matrice par classe de conjugaison (qui sont données dans l'exemple 8.35).

Pour tout cela nous allons considérer un tétraèdre dont le centre (isobary) est en $(0, 0, 0)$ et une base de $\mathbb{R}^3$ formée de trois sommets e_1, e_2 et e_3. Vu que l'isobarycentre des quatre sommes est en $(0, 0, 0)$, le quatrième somme est forcément le point de coordonnées $e_4(-1, -1, -1)$, de telle sorte que $e_1 + e_2 + e_3 + e_4 = 0$.

Les transpositions Quelle isométrie de $\mathbb{R}^3$ permute deux sommets du tétraèdre sans bouger les autres ? Pour permuter les sommets e_1 et e_2 en laissant e_3 et e_4, c'est le symétrie par rapport au plan médiateur de $[e_1, e_2]$. Ce plan passe par les sommets e_3 et e_4, parce que le tétraèdre étant régulier, les points e_3 e_4 sont équidistants de e_1 et e_2. Le lemme 16.15 dit qu'alors ces points dont partie du plan médiateur.

Dans notre base, la matrice de la transposition précédemment nommée (12) est

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{17.112}$$

dont la trace est 1. Donc $\chi_s(12) = 1$.

Les bitranspositions La bitransposition (12)(34) est le produit des transpositions selon les plans médiateur de $[e_1, e_2]$ et $[e_3, e_4]$. Ces deux plans sont perpendiculaires, et l'intersection est la droite qui passe par les milieux. Cette droite est perpendiculaire aux deux segments en même temps. La matrice est :

$$\begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ 0 & 0 & -1 \end{pmatrix} \tag{17.113}$$

parce que $e_1 \mapsto e_2$, $e_2 \mapsto e_1$ et $e_3 \mapsto e_4$. Pour rappel, la matrice est formée des images des vecteurs de base. Cela donne

$$\chi_s\big((12)(34)\big) = -1. \tag{17.114}$$

Les 3-cycles La symétrie qui permute cycliquement les point e_1, e_2 et e_3 est la rotation d'angle $2\pi/3$ dans le plan formé par les extrémités de ces trois vecteurs. Heureusement, la trace est invariante par changement de base ; donc nous pouvons calculer la trace d'une rotation d'angle $2\pi/3$ dans n'importe quelle base. Par exemple :

$$\chi_s\big((12)(34)\big) = \mathrm{Tr}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(2\pi/3) & \sin(2\pi/3) \\ 0 & -\sin(2\pi/3) & \cos(2\pi/3) \end{pmatrix} = 1 + 2\cos(2\pi/3) = 0. \tag{17.115}$$

Notons que, sans cette interprétation géométrique, nous y arrivons aussi facilement : dans notre base le 3-cycle est $e_1 \mapsto e_2 \mapsto e_3 \mapsto e_1$, donc la matrice est :

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \tag{17.116}$$

dont la trace est manifestement nulle : $\chi_s\big((123)\big) = 0$.

Le 4-cycle Il fait $e_1 \mapsto e_2 \mapsto e_3 \mapsto e_4 \mapsto e_1$, dont la matrice est

$$\begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}, \tag{17.117}$$

et la trace est $\chi_s\big((1,2,3,4)\big) = -1$.

Nous avons retrouvé les caractères de la représentation ρ_s, et nous pouvons vérifier qu'elle est irréductible.

17.5.3 À propos de la représentation ρ_u

Nous nous penchons à présent sur la représentations ρ_u dont nous ne savons rien à part qu'elle est de dimension 2 et son caractère.

Lemme 17.27.
Nous avons $\rho_u(s) = \mathrm{Id}$ pour tout $s \in V_4$.

Démonstration. Tous les éléments de V_4 sont conjugués (à part l'identité, mais pour elle le résultat est clair), donc il suffit de prouver le résultat pour un élément quelconque.

L'endomorphisme $\rho_u\big((12)(34)\big)$ est un endomorphisme d'ordre 2 sur $\mathbb{R}^2$ dont la trace est 2. Imposons donc

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{17.118}$$

sous la contrainte $a + d = 2$. La résolution est assez rapide et donne $b = c = 0$, $a = d = 1$.

Vous voulez une démonstration plus technologique ? Remarquez que l'opérateur $A = \rho_u\big((12)(34)\big)$ vérifie $A^2 = 1$, donc le polynôme $X^2 - 1$ est un polynôme annulateur de A. Il est peut-être minimal ou peut être pas, mais en tout cas le polynôme minimal divise celui-là et donc est soit $X - 1$ soit $X + 1$ soit $X^2 - 1$. Dans les trois cas il est scindé à racines simples, et l'endomorphisme A est diagonalisable par le théorème 9.143(3).

Mais comme $A^2 = 1$, les valeurs propres (ce qui est sur la diagonale) de A ne peuvent être que ± 1. La trace étant 2, les éléments diagonaux ne peuvent être que 1. Et $A = \mathrm{Id}$. □

Le groupe V_4 définit en 8.50 est normal dans S_4, donc le quotient S_4/V_4 est un groupe par le lemme 3.51.

Lemme 17.28.
L'application

$$\begin{aligned} \tilde{\rho}_u \colon S_4/V_4 &\to \mathrm{GL}(2,\mathbb{R}) \\ [g] &\mapsto \rho_u(g) \end{aligned} \tag{17.119}$$

est bien définie et donne une représentation irréductible de S_4/V_4.

Démonstration. Montrons que c'est bien définit. Si $s \in V_4$ nous devons prouver que $\rho_u(gs) = \rho_u(g)$. Vu que ρ_u est un homomorphisme (c'est une représentation), et que $\rho_u(s) = \mathrm{Id}$ nous avons directement

$$\rho_u(gs) = \rho_u(g)\rho_u(s) = \rho_u(g). \tag{17.120}$$

Nous devons prouver que la représentation $\tilde{\rho}_u$ est irréductible. Si un sous espace non trivial $\mathrm{Span}(x)$ était stabilisé par $\tilde{\rho}_u$, il serait également stabilisé par ρ_u. Mais comme ρ_u est irréductible, elle ne stabilise personne. □

Lemme 17.29.
Le groupe S_4/V_4 est un groupe non-abélien, isomorphe à S_3.

Démonstration. Le groupe S_4/V_4 a une représentation irréductible de dimension 2, et n'est donc pas abélien par le corollaire 17.26.

Il contient $|S_4|/|V_4| = 24/4 = 6$ éléments (théorème de Lagrange 3.54). Or $6 = 3 \times 2$, donc le groupe S_4/V_4 est dans le cas non-abélien du théorème 8.26(2). Cette partie parle d'unicité du groupe non-abélien d'ordre 6. Or S_3 est un groupe non-abélien d'ordre 6, donc S_4/V_4 est isomorphe à S_3. □

Attention : il n'est pas correct de dire que S_4/V_4 est un sous-groupe de S_4 juste parce que c'est un quotient de S_4 ; ce n'est en général pas vrai (exemple 3.52).

17.30.
Nous sommes maintenant aptes à identifier la représentation ρ_u. D'abord nous nous rappelons de la représentation $\rho_s : S_4 \to \mathrm{Iso}(T)$ de S_4 sur le tétraèdre. Ensuite si A est un sommet dudit tétraèdre et que $S_3 \subset S_4$ est la partie qui fixe A alors nous avons une représentation

$$\rho_s : S_3 \to \mathrm{Iso}(T) \tag{17.121}$$

qui agit en réalité sur le triangle équilatéral T' opposé au sommet A.

Nous avons finalement la chaine d'homomorphismes de groupes

$$S_4 \xrightarrow{\texttt{proj}} S_4/V_4 \xrightarrow{\simeq} S_3 \xrightarrow{\rho_s} \mathrm{Iso}(T') \tag{17.122}$$

Cela est donc une représentation $S_4 \to \mathrm{Iso}(T')$. Elle est de dimension 2 et est irréductible (elle contient les rotations d'angle $2\pi/3$ qui ne fixent aucune direction). Elle est donc la représentation ρ_u qui est la seule irréductible de dimension 2.

Nous avons donc montré que la représentation ρ_u dont nous ne savions rien est la représentation de S_4 sur un triangle équilatéral obtenue à partir de celle de S_4 sur le tétraèdre, en fixant un point.

<++>

17.6 Table de caractères du groupe diédral

Cette section vient de [2] ; nous avons comme but d'établir la table des caractères des représentations complexes du groupe diédral D_n.

17.6.1 Représentations de dimension un

Nous nous occupons des représentations de D_n sur $\mathbb{C}$. Les applications linéaires $\mathbb{C} \to \mathbb{C}$ sont seulement les multiplications par des nombres complexes. Nous cherchons donc $\psi : D_n \to \mathbb{C}^*$.

Nous savons que D_n est généré[8] par s et r. Vu que $s^2 = 1$, nous avons

$$\psi(s)^2 = \psi(s^2) = \psi(1) = 1, \tag{17.123}$$

8. Voir proposition 16.100 et tout ce qui suit.

donc $\psi(s) \in \{-1,1\}$. Nous savons aussi que $srsr = 1$, donc

$$\psi(s)^2\psi(r)^2 = 1, \tag{17.124}$$

ce qui donne $\psi(r) \in \{-1,1\}$.

Nous avons donc quatre représentations de dimension un données par

	$\psi(r) = 1$	$\psi(r) = -1$
$\psi(s) = 1$	ρ^{++}	ρ^{+-}
$\psi(s) = -1$	ρ^{-+}	ρ^{--}

Attention au fait que nous devons aussi avoir la relation $\psi(r)^n = \psi(r^n) = 1$. Donc $\psi(r)$ doit être une racine n^{e} de l'unité. Nous allons donc devoir avoir un compte différent selon la parité de n. Nous en reparlerons à la fin, au moment de faire les comptes. En ce qui concerne les caractères correspondants,

	r^k	sr^k
χ^{++}	1	1
χ^{+-}	$(-1)^k$	$(-1)^k$
χ^{-+}	1	-1
χ^{--}	$(-1)^k$	$(-1)^{k+1}$

Étant donné qu'ils sont tous différents, ce sont des représentations deux à deux non équivalentes, lemme 17.15.

17.6.2 Représentations de dimension deux

Nous cherchons maintenant les représentations $\rho\colon D_n \to \text{End}(\mathbb{C}^2)$. Ici nous supposons connue la liste des éléments de D_n donnée par le corollaire 16.101. Soit $\omega = e^{2i\pi/n}$ et $h \in \mathbb{Z}$; nous considérons la représentation $\rho^{(h)}$ de D_n définie par

$$\rho^{(h)}(r^k) = \begin{pmatrix} \omega^{hk} & 0 \\ 0 & \omega^{-hk} \end{pmatrix} \tag{17.125a}$$

$$\rho^{(h)}(st^k) = \begin{pmatrix} 0 & \omega^{-hk} \\ \omega^{hk} & 0 \end{pmatrix}. \tag{17.125b}$$

Cela donne bien $\rho^{(h)}$ sur tous les éléments de D_n par la proposition 16.100. Nous pouvons restreindre le domaine de h en remarquant d'abord que $\rho^{(h)} = \rho^{(h+n)}$, et ensuite que les représentations $\rho^{(h)}$ et $\rho^{(-h)}$ sont équivalentes. Un opérateur d'entrelacement est donné par $T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, et il est facile de vérifier que $T\rho^{(h)}(x) = \rho^{-h}(x)T$ avec $x = r^k$ puis avec $x = sr^k$.

Donc $\rho^{(h)} \simeq \rho^{(-h)} \simeq \rho^{(n-h)}$ et nous pouvons restreindre notre étude à $0 \leqslant h \leqslant \frac{n}{2}$.

Nous allons séparer les cas $n = 0$, $h = n/2$ et les autres. En effet si nous notons par commodité $a = \omega^h$, alors un vecteur (x, y) est vecteur propre de $\rho^{(h)}(s)$ et de $\rho^{(h)}(r)$ si et seulement s'il vérifie les systèmes d'équations

$$\begin{cases} ax = \lambda x & \text{(17.126a)} \\ \dfrac{1}{a}y = \lambda y & \text{(17.126b)} \end{cases}$$

et

$$\begin{cases} \dfrac{1}{a}y = \mu x & \text{(17.127a)} \\ ax = \mu y & \text{(17.127b)} \end{cases}$$

avec λ et μ des nombres non nuls. Une représentation sera réductible si et seulement si ces deux systèmes acceptent une solution non nulle commune. Il est vite vu que si $x \neq 0$ et $y \neq 0$, alors $a^2 = 1$, ce qui signifie $h = 0$ ou $h = n/2$. Sinon, il n'y a pas de solutions, et la représentation associée est irréductible.

(1) $h = 0$. Nous avons

$$\rho^{(0)}(r^k) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \rho^{(0)}(sr^k) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{17.128}$$

donc le caractère de cette représentation est $\chi^{(0)}(r^k) = 2$ et $\chi^{(0)}(sr^k) = 0$. Donc nous avons

$$\chi^{(0)} = \chi^{++} + \chi^{-+}. \tag{17.129}$$

Il y a maintenant (au moins) quatre façons de voir que la représentation $\rho^{(0)}$ est réductible.

Première méthode Trouver un opérateur d'entrelacement. Pour cela nous calculons les matrices :

$$S(r) = (\rho^{++} \oplus \rho^{-+})(r^k) = \begin{pmatrix} \rho^{++}(r^k) & 0 \\ 0 & \rho^{-+}(r^k) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{17.130a}$$

$$S(sr^k) = (\rho^{++} \oplus \rho^{-+})(sr^k) = \begin{pmatrix} \rho^{++}(sr^k) & 0 \\ 0 & \rho^{-+}(sr^k) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{17.130b}$$

$$\tag{17.130c}$$

Nous cherchons une matrice T telle que $TS(r^k) = \rho^{(0)}(r^k)T$ et $TS(sr^k) = \rho^{(0)}(sr^k)T$. Étant donné que $S(r^k) = \mathbb{1} = \rho^{(0)}(r^k)$, la première contrainte n'en est pas une. Nous pouvons vérifier qu'avec $T = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, nous avons bien

$$T \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{17.131}$$

Donc ce T entrelace $\rho^{++} \oplus \rho^{-+}$ avec $\rho^{(0)}$ qui sont donc deux représentations équivalentes. Donc $\rho^{(0)}$ est réductible et ça ne nous intéresse pas de la lister.

Seconde méthode Invoquer le théorème 17.22(1) pour dire que, les caractères étant égaux, les représentations sont équivalentes.

Troisième méthode Utiliser le théorème 17.22(2) et calculer $\langle \chi^{(0)}, \chi^{(0)} \rangle$:

$$\langle \chi^{(0)}, \chi^{(0)} \rangle = \frac{1}{|D_n|} \sum_{g \in D_n} |\chi^{(0)}(g)|^2 \tag{17.132a}$$

$$= \frac{1}{2n}\big(4 + 0 + 4(n-1)\big) \tag{17.132b}$$

$$= 2. \tag{17.132c}$$

Ici le 4 est pour le 1, le zéro est pour les termes sr^k et $4(n-1)$ est pour les $n-1$ termes r^k. Vu que le résultat n'est pas 1, la représentation $\rho^{(0)}$ n'est pas irréductible.

Quatrième méthode Regarder les solutions des systèmes (17.126) et (17.127) dont nous avons parlé plus haut.

La première méthode a l'avantage d'être simple et ne demander aucune théorie particulière à part les définitions. La seconde méthode est la plus rapide, mais demande un théorème très puissant. La troisième utilise également un théorème assez avancé, mais a l'avantage sur les deux autres méthodes de ne pas avoir besoin de savoir a priori un candidat décomposition de $\rho^{0)}$; cette méthode est applicable même sans faire la remarque que $\chi^{(0)} = \chi^{++} + \chi^{-+}$. Quoi qu'il en soit, nous ne listons pas $\chi^{(0)}$ dans notre table de caractères.

(2) $h = n/2$. Vu que $\omega^{n/2} = e^{i\pi} = -1$, nous avons

$$\rho^{(n/2)}(r^k) = \begin{pmatrix} (-1)^k & 0 \\ 0 & (-1)^k \end{pmatrix} \quad \rho^{(n/2)}(sr^k) = \begin{pmatrix} 0 & (-1)^k \\ (-1)^k & 0 \end{pmatrix} \quad , \tag{17.133}$$

et donc

$$\chi^{(n/2)}(r^k) = 2(-1)^k \tag{17.134a}$$
$$\chi^{(n/2)}(sr^k) = 0. \tag{17.134b}$$

Il est vite vu que $\chi^{(n/2)} = \chi^{+-} + \chi^{-+}$. Ergo la représentation $\rho^{(n/2)}$ n'est pas irréductible.

(3) $0 < h < \frac{n}{2}$. Dans ce cas nous avons $\omega^h \neq \omega^{-h}$, et en regardant les systèmes d'équations donnés plus haut, nous voyons que $\rho^{(h)}(s)$ et $\rho^{(h)}(r)$ n'ont pas de vecteurs propres communs. Donc ces représentations sont irréductibles.

Nous devons cependant encore vérifier si elles sont deux à deux non équivalentes. Supposons que pour $h \neq h'$ nous ayons une matrice $T \in \mathrm{GL}(2, \mathbb{C})$ telle que $T\rho^{(h)}(r)T^{-1} = \rho^{(h')}(r)$. Cela impliquerait en particulier que les matrices $\rho^{(h)}(r)$ et $\rho^{(h')}(r)$ aient même valeurs propres. Nous aurions donc $\{\omega^h, \omega^{-h}\} = \{\omega^{h'}, \omega^{-h'}\}$. Mais cela est impossible avec $0 < h < h' < \frac{n}{2}$. Donc toutes ces représentations sont distinctes.

Le caractère de la représentation $\rho^{(h)}$ est $\chi^{(h)}(r^k) = \omega^{hk} + \omega^{-hk} = 2\cos\left(\frac{2\pi hk}{n}\right)$.
Nous ajoutons donc la ligne suivante à notre liste :

	r^k	sr^k
$\chi^{(h)}$	$2\cos\left(\frac{2\pi hk}{n}\right)$	0

17.6.3 Le compte pour n pair

Nous avons 4 représentations de dimension 1 puis $\frac{n}{2} - 1$ représentations de dimension 2. En tout nous avons

$$\frac{n}{2} + 3 \tag{17.135}$$

représentations irréductibles modulo équivalence. Cela fait le compte en vertu des classes de conjugaisons listées en 16.4.3.4. Pour rappel, le nombre de représentations non équivalentes est égal au nombre de classes de conjugaison par le corollaire 17.25. Notons que c'est cela qui justifie le fait que nous ne devons pas chercher d'autres représentations. Nous sommes sûrs de les avoir toutes trouvées.

17.6.4 Le compte pour n impair

Nous avions fait mention plus haut du fait que si ψ est une représentation de dimension 1, le nombre $\psi(r)$ devait être une racine n^{e} de l'unité. Donc en dimension 1 nous avons seulement les représentations ρ^{++} et ρ^{-+}. Pour celles de dimension 2, nous en avons $\frac{n-1}{2}$. En tout nous avons donc

$$\frac{n+3}{2} \tag{17.136}$$

représentations irréductibles modulo équivalence. Cela fait le compte en vertu des classes de conjugaisons listées en 16.4.3.5.

Chapitre 18

Suite de l'analyse

18.1 Densité des polynômes

18.1.1 Théorème de Stone-Weierstrass

Voir le thème 11.

Note : le lemme 18.1 est utilisé dans la démonstration du théorème 18.4 ; c'est pour cela que nous l'avons isolé.

Lemme 18.1.
Il existe une suite de polynômes sur $[0,1]$ *convergeant uniformément vers la fonction racine carré.*

Démonstration. Nous donnons cette suite par récurrence :

$$P_0(t) = 0 \tag{18.1a}$$

$$P_{n+1}(t) = P_n(t) + \frac{1}{2}\big(t - P_n(t)^2\big). \tag{18.1b}$$

Nous commençons par montrer que pour tout $t \in [0,1]$, $P_n(t) \in [0,\sqrt{t}]$. Pour P_0, c'est évident. Ensuite nous avons

$$P_{n+1}(t) - \sqrt{t} = P_n(t) - \sqrt{t} + \frac{1}{2}(t - P_n(t)^2) \tag{18.2a}$$

$$= \big(P_n(t) - \sqrt{t}\big)\left(1 - \frac{1}{2}\frac{t - P_n(t)^2}{P_n(t) - \sqrt{t}}\right) \tag{18.2b}$$

$$= \big(P_n(t) - \sqrt{t}\big)\left(1 - \frac{\sqrt{t} + P_n(t)}{2}\right) \tag{18.2c}$$

$$\leqslant 0 \tag{18.2d}$$

parce que $\sqrt{t} \leqslant 1$ et $P_n(t) \leqslant 1$ par hypothèse de récurrence.

Nous savons au passage que $P_n(t)$ est une suite réelle croissante parce que $t - P_n(t)^2 \geqslant t - (\sqrt{t})^2 = 0$. La suite $P_n(t)$ est donc croissante et majorée par $\sqrt{t}$; elle converge donc. Les candidats limites sont déterminés par l'équation

$$\ell = \ell + \frac{1}{2}(t - \ell^2), \tag{18.3}$$

dont les solutions sont $\ell = \pm\sqrt{t}$. La suite étant positive, nous avons une convergence ponctuelle de P_n vers la racine carré. Cette suite étant une suite croissante de fonctions continues sur un compact, convergeant ponctuellement vers une fonction continue, la convergence est uniforme par le théorème de Dini 12.302. □

Lemme 18.2.
Soit K*, un compact de* $\mathbb{R}$ *et* f_n *une suite de fonctions sur* K *convergeant uniformément vers* f*. Soit* $g\colon X \to K$ *une fonction depuis un espace topologique* K*. Alors* $f_n \circ g$ *converge uniformément vers* $f \circ g$*.*

Démonstration. En effet, pour tout $x \in X$ nous avons

$$\|(f_n \circ g) - (f \circ g)\|_\infty = \sup_{x\in X} \|f_n\big(g(x)\big) - f\big(g(x)\big)\| \leqslant \|f_n - f\|_\infty. \tag{18.4}$$

Par conséquent, si $\epsilon\, 0$ est donné, il suffit de choisir n de telle sorte à avoir $\|f_n - f\|_\infty < \epsilon$ et nous avons $\|(f_n \circ g) - (f \circ g)\|_\infty \leqslant \epsilon$. □

Définition 18.3.
Nous disons qu'une algèbre A de fonctions sur un espace X ***sépare les points*** *de X si pour tout $x_1 \neq x_2$ il existe $g \in A$ telle que $g(x_1) \neq g(x_2)$.*

Nous pouvons maintenant énoncer et démontrer une forme nettement plus générale du théorème de Stone-Weierstrass.

Théorème 18.4 (Stone-Weierstrass[211])**.**
Soit X, un espace compact et Hausdorff et A une sous algèbre de $C(X, \mathbb{R})$ contenant une fonction constante non nulle. Alors A est dense dans $\Big(C(X, \mathbb{R}), \|.\|_\infty\Big)$ si et seulement si A sépare les points de X.

Nous pouvons remplacer $\mathbb{R}$ par $\mathbb{C}$ si de plus l'algèbre A est auto-adjointe : $g \in A$ implique $\bar{g} \in A$.

Démonstration. Nous allons écrire la démonstration en plusieurs étapes (dont la première est le lemme 18.1).

Première étape Pour tout $x \neq y \in X$ et pour tout $\alpha, \beta \in \mathbb{R}$, il existe une fonction $f \in A$ telle que $f(x) = \alpha$ et $f(y) = \beta$.
En effet, vu que A sépare les points nous pouvons considérer une fonction $g \in A$ telle que $g(x) \neq g(y)$ et ensuite poser

$$f(z) = \alpha + \frac{\alpha - \beta}{g(y) - g(x)}\big(g(z) - g(x)\big). \tag{18.5}$$

Les constantes faisant partie de A, cette fonction f est encore dans A.

Seconde étape Pour tout n-uples de fonctions $f_1, \dots, f_n$ dans $\bar{A}$, les fonctions $\min(f_1, \dots, f_n)$ et $\max(f_1, \dots, f_n)$ sont dans $\bar{A}$.
Nous le démontrons pour $n = 2$; le reste allant évidemment par récurrence. Soient $f, g \in \bar{A}$. Étant donné que

$$\max(f, g) = \frac{f + g}{2} + \frac{|f - g|}{2} \tag{18.6a}$$

$$\min(f, g) = \frac{f + g}{2} - \frac{|f - g|}{2}, \tag{18.6b}$$

if suffit de montrer que si $f \in \bar{A}$ alors $|f| \in \bar{A}$. Si f est nulle, c'est évident ; supposons que $f \neq 0$ et posons $M = \|f\|_\infty \neq 0$. Pour tout $x \in X$ nous avons

$$\frac{f(x)^2}{M^2} \in [0, 1]. \tag{18.7}$$

Nous considérons alors la suite

$$h_n = P_n \circ \frac{f^2}{M^2} \tag{18.8}$$

où P_n est une suite de polynômes convergent uniformément vers la racine carré (voir lemme 18.1). Le lemme 18.2 nous assure que h_n converge uniformément vers $\frac{|f|}{M}$ dans $C(X, \mathbb{R})$. Étant donné que $\bar{A}$ est également une algèbre, h_n est dans $\bar{A}$ pour tout n et la limite s'y trouve également (pour rappel, la fermeture $\bar{A}$ est celle de la topologie de la convergence uniforme).

Troisième étape Soit $\epsilon > 0$, $f \in C(X, \mathbb{R})$ et $x \in X$. Il existe une fonction $g_x \in \bar{A}$ telle que

$$\begin{cases} g_x(x) = f(x) & (18.9a) \\ g_x(y) \leqslant f(y) + \epsilon & (18.9b) \end{cases}$$

pour tout $y \in X$.

Soit $z \in X \backslash \{x\}$ et une fonction h_z telle que $h_z(x) = f(x)$ et $h_z(z) = f(z)$. Une telle fonction existe par une des étapes précédentes. Étant donné que f et h_z sont continues, il existe un voisinage ouvert V_z de z sur lequel

$$h_z(y) \leqslant f(y) + \epsilon \tag{18.10}$$

pour tout $y \in V_z$. Nous pouvons sélectionner un nombre fini de points $z_1, \ldots, z_n$ tels que les ouverts $V_{z_1}, \ldots, V_{z_n}$ recouvrent X (parce que X est compact, de tout recouvrement par des ouverts, nous extrayons un sous recouvrement fini.). Nous posons

$$g_x = \min(h_{z_1}, \ldots, h_{z_n}) \in \bar{A}. \tag{18.11}$$

Si $y \in X$, nous sélectionnons le i tel que $h_{z_i}(y) \leqslant f(y) + \epsilon$ et nous avons

$$g_x(y) \leqslant h_{z_i}(y) \leqslant f(y) + \epsilon. \tag{18.12}$$

Étape finale Soit $\epsilon > 0$ et $f \in C(X, \mathbb{R})$. Pour chaque $x \in X$ nous considérons une fonction $g_x \in \bar{A}$ telle que

$$\begin{cases} g_x(x) = f(x) & (18.13a) \\ g_x(y) \leqslant f(y) + \epsilon & (18.13b) \end{cases}$$

pour tout $y \in X$. Les fonctions f et g_x sont continues, donc il existe un voisinage ouvert W_x de x sur lequel

$$g_x(y) \geqslant f(y) - \epsilon. \tag{18.14}$$

De ces W_x nous extrayons un sous recouvrement fini de X : $W_{x_1}, \ldots, W_{x_m}$ et nous posons

$$\varphi = \max(g_{x_1}, \ldots, g_{x_n}) \in \bar{A}. \tag{18.15}$$

Si $y \in X$, il existe un i tel que

$$\varphi(y) \geqslant g_{x_i}(y) \geqslant f(y) - \epsilon. \tag{18.16}$$

La première inégalité est le fait que φ est le maximum des g_{x_k}, et la seconde est le choix de i. Donc pour tout $y \in X$ nous avons

$$f(y) - \epsilon \leqslant \varphi(y) \leqslant f(y) + \epsilon. \tag{18.17}$$

La première inégalité est ce que l'on vient de faire. La seconde est le fait que pour tout i nous ayons $g_{x_i}(y) \leqslant f(y) + \epsilon$; le fait que φ soit le maximum sur les i ne change pas l'inégalité.

Le fait que les inégalités (18.17) soient vraies pour tout $y \in X$ signifie que $\|\varphi - f\|_\infty \leqslant \epsilon$, et donc que $f \in \bar{\bar{A}} = \bar{A}$.

Tout cela prouve que $C(X, \mathbb{R}) \subset \bar{A}$. L'inclusion inverse est le fait que $C(X, \mathbb{R})$ est fermé pour la norme $\|.\|_\infty$, étant donné qu'une limite uniforme de fonctions continues est continue.

□

Corollaire 18.5 ([1])**.**
Soit B, la boule fermée de centre 0 et de rayon 1 dans $\mathbb{R}^n$. La partie $C^\infty(B, \mathbb{R}^n)$ est dense dans $\big(C(B, B), \|.\|_\infty\big)$.

Démonstration. Soit $f \in C(B,B)$ et $\epsilon > 0$. La fonction donnant la composante i est une fonction $f_i \in C(B,\mathbb{R})$ et il existe donc, par le théorème de Stone-Weierstrass 18.4, une fonction $g_i \in C^\infty(B,\mathbb{R})$ telle que $\|g_i - f_i\|_\infty \leqslant \epsilon$.

La fonction g dont les composantes sont les g_i ainsi construits vérifie $\|g - f\|_\infty \leqslant n\epsilon$. □

Attention toutefois que rien n'assure que les fonctions construites par le corollaire 18.5 prennent leurs valeurs dans B.

Le théorème suivant est un des énoncés les plus classiques de Stone-Weierstrass. Il découle évidement du théorème général 18.4 (encore qu'il faut alors bien comprendre qu'il faut traiter la fonction $x \mapsto \sqrt{x}$ séparément). Il en existe cependant une preuve indépendante.

Théorème 18.6.
Soit f, une fonction continue de l'intervalle compact $[a,b]$ à valeurs dans $\mathbb{R}$. Alors pour tout $\epsilon > 0$, il existe un polynôme P tel que $\|P - f\|_\infty < \epsilon$.

Autrement dit, les polynômes sont denses dans $C[a,b]$ pour la norme uniforme.

Corollaire 18.7.
Si $X \subset \mathbb{R}$ est compact et de mesure finie[1]*, alors l'ensemble des polynômes est denses dans $\big(C(X,\mathbb{R}), \|.\|_2\big)$.*

Démonstration. Si f est une fonction dans $C(X,\mathbb{R})$ et si $\epsilon \geqslant 0$ est donné alors nous pouvons considérer un polynôme P tel que $\|f - P\|_\infty \leqslant \epsilon$. Dans ce cas nous avons

$$\|f - P\|_2^2 = \int_X |f(x) - P(x)|^2 dx \leqslant \int_X \epsilon^2 dx = \epsilon^2 \mu(X) \tag{18.18}$$

où $\mu(X)$ est la mesure de X (finie par hypothèse). □

18.2 Primitive de fonction continue

Proposition 18.8 ([212])**.**
Soit un intervalle compact K de $\mathbb{R}$ et une suite (f_n) de fonctions continues sur K telles que $f_n \xrightarrow{unif} f$. Si chacune des fonctions f_n a une primitive sur K alors f également.

Démonstration. Soit $x_0 \in K$ et les primitives F_n choisies[2] pour avoir $F_n' f_n$ et $F_n(x_0) = 0$. Nous allons voir que (F_n) est une suite de Cauchy dans $\big(K, \|.\|_\infty\big)$. Soient $n, m \in \mathbb{N}$ et $x \in K$. Nous avons

$$\begin{aligned}
\|F_n - F_m\|_\infty &\leqslant \|F_n(x) - F_m(x)\| && \text{(18.19a)}\\
&= \|(F_n - F_m)(x)\| && \text{(18.19b)}\\
&\leqslant \|F_n' - F_m'\|_{[x,x_0]} \|x - x_0\| && \text{(18.19c)}
\end{aligned}$$

où nous avons utilisé le théorème des accroissements finis 12.229. Vu que $x \in K$ et que K est borné, $\|x - x_0\|$ est majoré par $\operatorname{diam}(K)$ et

$$\|F_n - F_m\|_K \leqslant \|f_n - f_m\|_K \operatorname{diam}(K). \tag{18.20a}$$

Vu que (f_n) est de Cauchy, si n et m sont assez grands, cela tend vers zéro. La suite (F_n) converge donc vers une certaine fonction F.

Le théorème 12.303 nous permet de permuter la limite et la dérivée pour conclure que $F' = f$ et donc que f a une primitive sur K. □

1. Dans $\mathbb{R}$ cette hypothèse est évidemment superflue par rapport à l'hypothèse de compacité ; mais ça suggère des généralisations ...

2. Les fonctions F_n étant dérivables sont continues.

Proposition 18.9 ([212]).
Soit un intervalle ouvert I de $\mathbb{R}$ et une fonction $f\colon I \to \mathbb{R}$ qui admet une primitive sur tout compact de I. Alors f a une primitive sur I.

Démonstration. Nous considérons une suite exhaustive[3] de compacts K_n pour I et $x_0 \in K_0$. Nous considérons aussi F_n la primitive de f sur K_n telle que $F_n(x_0) = 0$ (possible parce que $x_0 \in K_n$ pour tout n). Les fonctions F_n sont des restrictions les une des autres, et nous pouvons définir

$$\begin{aligned} F\colon I &\to \mathbb{R} \\ x &\mapsto F_n(x) \text{ si } x \in K_n. \end{aligned} \tag{18.21}$$

Nous avons évidemment $F(x_0) = 0$ et nous allons prouver que F est une primitive de f sur I. Soit $x \in I$ vu que I est ouvert, nous pouvons choisir n_0 tel que $x \in \text{Int}(K_{n_0})$. Les fonctions F et F_{n_0} sont égales sur K_n et donc sur un ouvert autour de x. Par conséquent F est dérivable en x et $F'(x) = F'_{n_0}(x) = f(x)$. □

Théorème 18.10.
Soit I un intervalle ouvert de $\mathbb{R}$. Une fonction continue sur I admet une primitive[4] *sur I.*

Démonstration. Sur chaque compact de I, la fonction f est limite uniforme de polynômes[5] (théorème de Stone-Weierstrass 18.6). Donc f est primitivable sur tout compact de I (proposition 18.8) et donc sur I par la proposition 18.9. □

Proposition 18.11.
Soit I un intervalle borné ouvert de $\mathbb{R}$. Une fonction $h \in C_c^\infty(I)$ admet une primitive dans $C_c^\infty(I)$ si et seulement si $\int_I h = 0$.

Démonstration. Si une primitive H de h est à support compact, alors

$$\int_I h = H(b) - H(a) = 0 - 0 = 0. \tag{18.22}$$

Pas de problèmes dans ce sens.

Supposons maintenant que $\int_I h = 0$. Le fait que h admette une primitive dans $C^\infty(I)$ est évident : toute fonction continue admet une primitive[6]. Soit H une telle primitive et $\tilde{H} = H - H(b)$. Alors $\tilde{H}(b) = 0$ et

$$\tilde{H}(a) = H(a) - H(b) = -\int_I h = 0. \tag{18.23}$$

Nous rappelons que le support d'une fonction est *la fermeture* de l'ensemble des points de non-annulation.

Supposons que le support de h soit inclus dans $[m, M] \subset]a, b[$. En prenant des nombres m' et M' tels que $a < m' < m$ et $M < M' < b$ (nous insistons sur le caractère strict de ces inégalités), la fonction h est nulle sur $[a, m']$ et sur $[M', b]$; la fonction $\tilde{H}$ doit donc y être constante. Mais nous avons déjà vu que $\tilde{H}(a) = \tilde{H}(b) = 0$. Donc l'ensemble des points sur lesquels $\tilde{H}$ n'est pas nul est inclus à $]m', M'[$ et donc est strictement (des deux côtés) inclus à I. □

18.2.1 Théorème taubérien de Hardi-Littlewood

Un théorème **taubérien** est un théorème qui compare les modes de convergence d'une série.

3. Voir le lemme 6.156.
4. Définition 12.135.
5. Si tu veux te passer de Stone-Weierstrass, tu peux prouver que toute fonction continue sur un compact est limite uniforme de fonctions affines par morceaux, par exemple. Voir [212].
6. Théorème 18.10.

Lemme 18.12.
Si f et g sont des fonctions continues, alors $s(x) = \max\{f(x), g(x)\}$ est également une fonction continue.

Démonstration. Soit x_0 et prouvons que s est continue en x_0. Si $f(x_0) \neq g(x_0)$ (supposons $f(x_0) > g(x_0)$ pour fixer les idées), alors nous avons un voisinage de x_0 sur lequel $f > g$ et alors $s = f$ sur ce voisinage et la continuité provient de celle de f.

Si au contraire $f(x_0) = g(x_0) = s(x_0)$ alors si (a_n) est une suite tendant vers x_0, nous prenons N tel que $\big|f(a_n) - f(x_0)\big| \leqslant \epsilon$ pour tout $n > N$ et M tel que $\big|g(a_n) - g(x_0)\big| \leqslant \epsilon$ pour tout $n > M$. Alors pour tout $n > \max\{N, M\}$ nous avons

$$\big|s(a_n) - s(x_0)\big| \leqslant \epsilon, \tag{18.24}$$

d'où la continuité de s en x_0. □

La proposition suivante dit que si une fonction connaît un saut, alors on peut le lisser par une fonction continue.

Proposition 18.13.
Soit f continue sur $[a, x_0[$ et sur $[x_0, b]$ avec $f(x_0^-) < f(x_0)$. En particulier nous supposons que $f(x^-)$ existe et est finie. Alors pour tout $\epsilon > 0$, il existe une fonction continue s telle que sur $[a, b]$ on ait $s \leqslant f$ et

$$\int_a^b s(x) - f(x)\,dx \leqslant \epsilon. \tag{18.25}$$

Démonstration. Nous notons A la taille du saut :

$$A = f(x_0) - f(x_0^-). \tag{18.26}$$

Quitte à changer a et b, nous pouvons supposer que

$$f(x) < f(x_0) + \frac{A}{3} \tag{18.27}$$

pour $x \in [a, x_0[$ et

$$f > f(x_0) + \frac{2A}{3} \tag{18.28}$$

pour $x \in [x_0, b]$. C'est le théorème des valeurs intermédiaires qui nous permet de faire ce choix.

Soit $m(x)$ la droite qui joint le point $\big(x_0 - \epsilon, f(x_0 - \epsilon)\big)$ au point $\big(x_0, f(x_0^+)\big)$. Nous posons

$$s(x) = \begin{cases} f(x) & \text{si } x < x_0 - \epsilon \\ \max\{m(x), f(x)\} & \text{si } x_0 - \epsilon \leqslant x \leqslant x_0 \\ f(x) & \text{si } x > x_0. \end{cases} \tag{18.29}$$

En vertu des différents choix effectués, c'est une fonction continue. En effet

$$s(x_0 - \epsilon) = \max\{f(x_0 - \epsilon), f(x_0, \epsilon)\} = f(x_0 - \epsilon) \tag{18.30}$$

et

$$s(x_0) = \max\{m(x_0), f(x_0^+)\} = f(x_0^+) \tag{18.31}$$

parce que $m(x_0) = f(x_0^+)$. En ce qui concerne l'intégrale, si nous posons

$$M = \sup_{x,y \in [a,b]} |f(x) - f(y)|, \tag{18.32}$$

nous avons

$$\int_a^b s - f = \int_{x_0 - \epsilon}^{x_0} s - f \leqslant \epsilon M. \tag{18.33}$$

□

Lemme 18.14.
Pour tout polynôme P, nous avons la formule

$$\lim_{x\to 1^-}(1-x)\sum_{n=0}^{\infty}x^nP(x^n)=\int_0^1P(x)dx. \tag{18.34}$$

Démonstration. D'abord pour $P=1$, la formule se réduit à la série harmonique connue. Ensuite nous prouvons la formule pour le polynôme $P=X^k$ et la linéarité fera le reste pour les autres polynômes. Nous avons

$$(1-x)\sum_n x^nx^{kn}=(1-x)\sum_n(x^{1+k})^n=\frac{1-x}{1-x^{1+k}}=\frac{1}{1+x+\cdots+x^k}. \tag{18.35}$$

Donc

$$\lim_{x\to 1^-}(1-x)\sum_n x^nP(x^n)=\frac{1}{1+k}. \tag{18.36}$$

Par ailleurs, c'est vite vu que

$$\int_0^1x^kdx=\frac{1}{k+1}. \tag{18.37}$$

□

Théorème 18.15 (Hardy-Littlewood[209])**.**
Soit (a_n) une suite réelle telle que

(1) $\frac{a_n}{n}$ tends vers une constante,

(2) $F(x)=\sum_{n=0}^{\infty}a_nx^n$ a un rayon de convergence $\geqslant 1$,

(3) $\lim_{x\to 1^-}F(x)=l$.

Alors $\sum_{n=0}^{\infty}a_n=l$.

Démonstration. Quitte à prendre la suite $b_0=a_0-l$ et $b_n=a_n$, on peut supposer $l=0$.

Soit Γ l'ensemble des fonctions

$$\gamma\colon[0,1]\to\mathbb{R} \tag{18.38}$$

telles que

(1) $\sum_{n=0}^{\infty}a_n\gamma(x^n)$ converge pour $0\leqslant x<1$,

(2) $\lim_{x\to 1^-}\sum_{n\geqslant 0}a_n\gamma(^n)=0$.

Ce Γ est un espace vectoriel.

Les polynômes sont dans Γ Soit $\gamma(t)=t^s$. Pour $0\leqslant x<1$ nous avons

$$\sum_{n=0}^{\infty}a_n\gamma(x^n)=\sum_{n=0}^{\infty}a_nx^{ns}<\sum_{n=0}^{\infty}a_nx^n. \tag{18.39}$$

Donc la condition de convergence est vérifiée. En ce qui concerne la limite,

$$\lim_{x\to 1^-}\sum_{n=0}^{\infty}a_nx^{ns}=\lim_{x\to 1^-}F(x^s)=0 \tag{18.40}$$

parce que par hypothèse, $\lim_{x\to 1^-}F(x)=0$.

Définition de la fonction qui va donner la réponse Nous considérons la fonction $g=\mathbb{1}_{[\frac{1}{2},1]}$, c'est à dire

$$g(t)=\begin{cases}0 & \text{si } 0\leqslant t<1/2\\ 1 & \text{si } 1/2\leqslant t\leqslant 1.\end{cases} \tag{18.41}$$

Nous montrons que si $g \in \gamma$, alors le théorème est terminé. Si $0 \leqslant x \leqslant 1$, on a $0 \leqslant x^n < 1/2$ dès que

$$n > -\frac{\ln(2)}{\ln(x)} \tag{18.42}$$

avec une note comme quoi $\ln(x) < 0$, donc la fraction est positive. Nous désignons par N_x la partie entière de ce n adapté à x. L'idée est que la fonction $g(x^n)$ est la fonction indicatrice de $0 \leqslant n \leqslant N_x$, et donc

$$\sum_{n\geqslant 0} a_n g(x^n) = \sum_{n=0}^{N_x} a_n. \tag{18.43}$$

Mais si $x \to 1^-$, alors $N_x \to \infty$, donc

$$\lim_{N\to\infty} \sum_{n=0}^{N} a_n = \lim_{x\to 1^-} \sum_{n=0}^{N_x} a_n = \lim_{x\to 1^-} \sum_{n\in\mathbb{N}} a_n g(x^n), \tag{18.44}$$

et cela fait zéro si $g \in \Gamma$.

Approximation de g par des polynômes Nous considérons la fonction

$$h(t) = \frac{g(t) - t}{t(1-1)} = \begin{cases} \frac{1}{t-1} & \text{si } t \in [0, 1/2[\\ \frac{1}{t} & \text{si } t \in [1/2, 1]. \end{cases} \tag{18.45}$$

La seconde égalité est au sens du prolongement par continuité. La fonction h est une fonction non continue qui fait un saut de -2 à 2 en $x = 1/2$. En vertu de la proposition 18.13 (un peu adaptée), nous pouvons considérer deux fonctions continues s_1 et s_2 telles que

$$s_1 \leqslant h \leqslant s_2 \tag{18.46}$$

et

$$\int_0^1 s_2 - s_1 \leqslant \epsilon. \tag{18.47}$$

Notons que l'inégalité $s_1 \leqslant s_2$ doit être stricte sur au moins un petit intervalle autour de $x = 1/2$. Soient P_1 et P_2, deux polynômes tels que $\|P_1 - s_1\|_\infty \leqslant \epsilon$ et $\|P_2 - s_2\|_\infty \leqslant \epsilon$ (ici la norme supremum est prise sur $[0, 1]$). C'est le théorème de Stone-Weierstrass (18.6) qui nous permet de le faire.
Nous posons aussi [7]

$$Q_1 = P_1 + \epsilon \tag{18.48a}$$
$$Q_2 = P_2 - \epsilon. \tag{18.48b}$$

Nous avons

$$\int_0^1 Q_1 - Q_2 \leqslant \int_0^1 Q_1 - P_1 + P_1 - P_2 + P_2 - Q_2. \tag{18.49}$$

Pour majorer cela, d'abord $Q_1 - P_1 = P_2 - Q2 = \epsilon$, ensuite,

$$P_1 - P_2 = P_1 - s_1 + s_1 - s_2 + s_2 - P_2 \tag{18.50}$$

dans lequel nous avons $P_1 - s_1 \leqslant \epsilon$, $s_2 - P_2 \leqslant \epsilon$ et $\int_0^1 s_1 - s_2 \leqslant \epsilon$. Au final, nous posons $q = Q_2 - Q_1$ et nous avons

$$\int_0^1 q \leqslant 5\epsilon. \tag{18.51}$$

Enfin nous posons aussi

$$R_i(x) = x + x(1-x)Q_i. \tag{18.52}$$

7. À ce niveau, je crois qu'il y a une faute de frappe dans [209].

Ces polynômes vérifient $R_i(0) = 0$, $R_i(1) = 1$ et

$$R_1 \leqslant g \leqslant R_2 \tag{18.53}$$

parce que

$$Q_1 \leqslant P_1 \leqslant h \leqslant P_2 \leqslant Q_2 \tag{18.54}$$

et

$$t + t(1-t)Q_1 \leqslant \underbrace{t + t(1-t)h(t)}_{g(t)} \leqslant t + t(1-t)Q_2. \tag{18.55}$$

Preuve que g est dans Γ D'abord si $0 \leqslant x < 1$, $x^N < \frac{1}{2}$ pour un certain N, et alors $g(x^N) = 0$. Du coup la série

$$\sum_{n=0}^{\infty} a_n g(x^n) = \sum_{n=0}^{N} a_n \tag{18.56}$$

est une somme finie qui converge donc.

D'autre part nous prenons M tel que $|a_n| < \frac{M}{n}$ pour tout n. Nous majorons $\sum_{n\in\mathbb{N}} a_n g(x^n)$ en utilisant R_1. Mais vu que R_1 est un polynôme, nous pouvons dire que $|\sum_{n=0}^{\infty} a_n R_1(x^n)| \leqslant \epsilon$ en prenant $x \in [\lambda, 1[$ et λ assez grand. Nous avons :

$$\left|\sum_{n=0}^{\infty} a_n g(x^n)\right| \leqslant \left|\sum_{n=0}^{\infty} a_n g(x^n) - \sum_{n=0}^{\infty} a_n R_1(x^n)\right| + \underbrace{\left|\sum_{n=0}^{\infty} a_n R_1(x^n)\right|}_{\leqslant\epsilon} \tag{18.57a}$$

$$\leqslant \epsilon + \sum_{n=0}^{\infty} |a_n|(g - R_1)(x^n) \tag{18.57b}$$

$$\leqslant \epsilon + \sum_{n=0}^{\infty} |a_n|(R_2 - R_1)(x^n) \tag{18.57c}$$

$$\leqslant \epsilon + M \sum_{n=0}^{\infty} \frac{x^n(1-x^n)}{n}(Q_2 - Q_1)(x^n) \qquad R_2 - R_1 = x(1-x)(Q_2 - Q_1) \tag{18.57d}$$

$$= \epsilon + M \sum_{n=0}^{\infty} \frac{x^n(1-x^n)}{n} q(x^n) \tag{18.57e}$$

$$\leqslant \epsilon + M(1-x)\sum_{n} x^n q(x^n) \tag{18.57f}$$

où la ligne (18.57f) provient d'une majoration sauvage de $1/n$ par 1 et de $1 - x^n$ par $1 - x$. Par le lemme 18.14, nous avons alors

$$\lim_{x\to 1^-} |\sum_n a_n g(x^n)| \leqslant \epsilon + M \int_0^1 q \leqslant 6\epsilon. \tag{18.58}$$

□

18.2.2 Théorème de Müntz

Théorème 18.16 (Théorème de Müntz[213, 214, 215])**.**
Soit $C_0\big([0,1]\big)$, l'espace des fonctions continues sur $[0,1]$ muni de la norme $\|.\|_\infty$ ou $\|.\|_2$ et une suite (α_n) strictement croissante de nombres positifs. Nous notons ϕ_λ la fonction $x \mapsto x^\lambda$.

Alors

$$\overline{\text{Span}\{1, \phi_{\alpha_n}\}} \tag{18.59}$$

est dense dans $C_0\big([0,1]\big)$ si et seulement si

$$\sum_{n=2}^{\infty} \frac{1}{\alpha_n} = +\infty. \tag{18.60}$$

Nous prouvons le théorème pour la norme $\|.\|_2$.

Démonstration. Soit $m \in \mathbb{R}^+$; nous notons $\Delta_N(m)$ la distance entre ϕ_m et $\text{Span}\{\phi_{\alpha_1}, \dots, \phi_{\alpha_N}\}$. Cette distance peut être évaluée avec le déterminant de Gram (proposition 9.15)

$$\Delta_N(m)^2 = \frac{G(\phi_m, \phi_{\alpha_1}, \dots, \phi_{\alpha_N})}{G(\phi_{\alpha_1}, \dots, \phi_{\alpha_N})}. \tag{18.61}$$

Pour calculer cela nous avons besoin des produits scalaires [8]

$$\langle \phi_a, \phi_b \rangle = \int_0^1 x^{a+b} dx = \frac{1}{a+b+1}. \tag{18.62}$$

Pour avoir des notation plus compactes, nous notons $\alpha_0 = m$. Donc nous avons à calculer le déterminant

$$G(\phi_m, \phi_{\alpha_1}, \dots, \phi_{\alpha_N}) = \det\left(\frac{1}{\alpha_i + \alpha_j + 1}\right) \tag{18.63}$$

où $i, j = 0, \dots, N$. Nous reconnaissons un déterminant de Cauchy (proposition 9.16) en posant, dans $\frac{1}{\alpha_i+\alpha_j+1}$, $a_i = \alpha_i$ et $b_j = \alpha_j + 1$. Étant donné que $b_j - b_i = a_j - a_i$, nous avons

$$G(\phi_m, \phi_{\alpha_1}, \dots, \phi_{\alpha_N}) = \frac{\prod_{0 \leqslant i < j \leqslant N} (\alpha_j - \alpha_i)^2}{\prod_{i=0}^N \prod_{j=0}^N (\alpha_i + \alpha_j + 1).} \tag{18.64}$$

Nous séparons maintenant les termes où i ou j sont nuls. En ce qui concerne le dénominateur, il faut prendre tous les couples (i, j) avec i et j éventuellement égaux à zéro. Nous décomposant cela en trois paquets. Le premier est $(0, 0)$; le second est $(0, i)$ (chaque couple arrive en fait deux fois parce qu'il y a aussi $(i, 0)$) ; et le troisième sont les i, j tous deux différents de zéro :

$$(2m+1) \prod_{ij} (\alpha_i + \alpha_j + 1) \prod_i (\alpha_i + m + 1)^2. \tag{18.65}$$

Notons que dans le produit central, le carré est contenu dans le fait qu'on écrit $\prod_{ij}$ et non $\prod_{i<j}$. Nous avons donc

$$G(\phi_m, \phi_{\alpha_1}, \dots, \phi_{\alpha_N}) = \frac{\prod_{i<j} (\alpha_i - \alpha_j)^2 \prod_i (\alpha_i - m)^2}{(2m+1) \prod_{ij} (\alpha_i + \alpha_j + 1) \prod_i (\alpha_i + m + 1)^2}. \tag{18.66}$$

Le calcul de $G(\phi_{\alpha_1}, \dots, \phi_{\alpha_N})$ est plus simple [9] :

$$G(\phi_{\alpha_1}, \dots, \phi_{\alpha_N}) = \frac{\prod_{i<j} (\alpha_i - \alpha_j)^2}{\prod_{ij} (\alpha_i + \alpha_j + 1)}. \tag{18.67}$$

En divisant l'un par l'autre il ne reste que les facteurs comprenant m et en prenant la racine carré,

$$\Delta_N(m) = \frac{1}{\sqrt{2m+1}} \prod_{i=1}^N \left| \frac{\alpha_i - m}{\alpha_i + m + 1} \right|. \tag{18.68}$$

Nous passons maintenant à la preuve proprement dite. Supposons que $V = \text{Span}\{\phi_{\alpha_i}, i \in \mathbb{N}\}$ est dense. Si m est un des α_i, il peut évidemment être approché par les ϕ_{α_i}. Mais vue la densité de V, un ϕ_m avec $m \neq \alpha_i$ (pour tout i) alors ϕ_m peut également être arbitrairement approché par les ϕ_{α_i}, c'est à dire que

$$\lim_{N \to \infty} \Delta_N(m) = 0. \tag{18.69}$$

Nous posons

$$u_n = \ln\left(\frac{\alpha_n - m}{\alpha_n + m + 1}\right) \tag{18.70}$$

8. C'est ici qu'on se particularise à la norme $\|.\|_2$.
9. Je crois qu'il y a une faute de frappe dans le dénominateur de [213].

et nous prouvons que la série $\sum_n u_n$ diverge. En effet nous nous souvenons de la formule $\ln(ab) = \ln(a) + \ln(b)$, de telle sorte que la N^{e}somme partielle de $\sum_n u_n$ est

$$\ln\left(\frac{\alpha_1 - m}{\alpha_1 + m + 1} \cdot \ldots \cdot \frac{\alpha_N - m}{\alpha_N + m + 1}\right) = \ln\left(\sqrt{2m+1}\Delta_N(m)\right), \tag{18.71}$$

qui tends vers $-\infty$ lorsque $N \to \infty$.

Si la suite (α_n) est majorée et plus généralement si nous n'avons pas $\alpha_n \to \infty$, alors évidemment la série $\sum_n \frac{1}{\alpha_n}$ diverge. Nous supposons donc que $\lim_{n\to\infty} \alpha_n = \infty$. Nous avons aussi [10]

$$u_n = \ln\left(\frac{\alpha_n - m}{\alpha_n + m + 1}\right) = \ln\left(1 - \frac{2m+1}{\alpha_n + m + 1}\right) \sim -\frac{2m+1}{\alpha_n}. \tag{18.72}$$

Une justification est donné à l'équation (12.164). Ce que nous avons surtout est

$$\sum_n u_n \sim -(2m+1)\sum_n \frac{1}{\alpha_n}. \tag{18.73}$$

Étant donné que la série de gauche diverge, celle de droite diverge [11].

Nous faisons maintenant le sens opposé : nous supposons que la série $\sum_n 1/\alpha_n$ diverge et nous nous posons

$$V = \text{Span}\{\phi_{\alpha_n} \text{ tel que } n \in \mathbb{N}\}. \tag{18.74}$$

Il suffit de prouver que $\phi_m \in \bar{V}$ pour tout m parce qu'un corollaire du théorème de Stone-Weierstrass 18.7 montre que $\text{Span}\{\phi_k \text{ tel que } k \in \mathbb{N}\}$ est dense dans C pour la norme $\|.\|_2$.

Si $\alpha_n \to \infty$, nous avons :

$$u_n \sim \frac{2m+1}{\alpha_n} \to 0 \tag{18.75}$$

et alors $\Delta_N(m) \to 0$. Dans ce cas nous avons immédiatement $\phi_m \in \bar{V}$.

Si par contre α_n ne tend pas vers l'infini, nous repartons de l'expression (18.68), nous posons $0 < \alpha = \sup_i \alpha_i$ et nous calculons :

$$\begin{aligned}
\sqrt{2m+1}\Delta_N(m) &= \prod_{i=1}^{N} \frac{|\alpha_i - m|}{\alpha_i + m + 1} && (18.76a)\\
&\leqslant \prod_{i=1}^{N} \frac{\alpha_i + m}{\alpha_i + m + 1} && (18.76b)\\
&= \prod_{i=1}^{N} \left(1 - \frac{1}{\alpha_i + m + 1}\right) && (18.76c)\\
&\leqslant \prod_{i=1}^{N} \left(1 - \frac{1}{\alpha + m + 1}\right) && (18.76d)\\
&= \left(1 - \frac{1}{\alpha + m + 1}\right)^N. && (18.76e)
\end{aligned}$$

Cette dernière expression tend vers 0 lorsque $N \to \infty$. □

Remarque 18.17.
Certaines sources [12] citent le théorème de Müntz comme ceci (avec un implicite que $\alpha_i \neq 0$) :

$$\overline{\text{Span}\{1, \phi_{\alpha_i}\}} = C\big([0,1]\big) \Leftrightarrow \sum_{i\geqslant 1} \frac{1}{\alpha_i} = +\infty. \tag{18.77}$$

10. Je crois qu'il y a une faute de signe dans la dernière expression de [214].
11. Nous utilisons le fait que si $u_n = \sum v_n$ en tant que suites et si $\sum_n u_n$ diverge, alors $\sum_n v_n$ diverge.
12. Dont le rapport du jury 2014

Que penser de la présence explicite du 1 (c'est à dire de ϕ_0) ou non dans l'ensemble ?

Première chose : la présence éventuelle de ϕ_0 est la raison pour laquelle nous faisons commencer la somme à $i = 2$ et non $i = 1$. Dans le même ordre d'idée, si $\text{Span}\{\phi_{\alpha_i}\}$ est dense, alors en prenant n'importe quelle queue de suite, ça reste dense.

Prouvons donc l'énoncé (18.77). Si $\text{Span}\{1, \phi_{\alpha_i}\}$ est dense, alors en posant $\beta_1 = 0$, $\beta_i = \alpha_{i-1}$ notre théorème prouve que $\sum_{\beta=2}^{\infty} \frac{1}{\beta_i} = +\infty$, cela est exactement que $\sum_{i=1}^{\infty} \frac{1}{\alpha_i} = +\infty$. Dans l'autre sens, si $\sum_{i\geqslant 1} \frac{1}{\alpha_i} = +\infty$, alors nous avons aussi $\sum_{i\geqslant 2} \frac{1}{\alpha_i} = +\infty$ et notre théorème dit que $\text{Span}\{\phi_{\alpha_i}\}$ est dense. A fortiori, $\text{Span}\{1, \phi_{\alpha_i}\}$ est dense.

Exemple 18.18
Nous savons depuis le théorème 15.68 que la somme des inverses des nombres premiers diverge.
△

18.3 Théorèmes de point fixe

18.3.1 Points fixes attractifs et répulsifs

Définition 18.19.
*Soit I un intervalle fermé de $\mathbb{R}$ et $\varphi\colon I \to I$ une application C^1. Soit a un point fixe de φ. Nous disons que a est **attractif** s'il existe un voisinage V de a tel que pour tout $x_0 \in V$ la suite $x_{n+1} = \varphi(x_n)$ converge vers a. Le point a sera dit **répulsif** s'il existe un voisinage V de a tel que pour tout $x_0 \in V$ la suite $x_{n+1} = \varphi(x_n)$ diverge.*

Lemme 18.20 ([216])**.**
Soit a un point fixe de φ.

(1) Si $|\varphi'(a)| < 1$ alors a est attractif et la convergence est au moins exponentielle.

(2) Si $|\varphi'(a)| > 1$ alors a est répulsif et la divergence est au moins exponentielle.

Démonstration. Si $|\varphi'(a) < 1|$ alors il existe k tel que $|\varphi'(a)| < k < 1$ et par continuité il existe un voisinage V de a dans lequel $|\varphi'(x)| < k$ pour tout $x \in V$. En utilisant le théorème des accroissements finis nous avons

$$|x_n - a| = \big|f(x_{n-1} - a)\big| \leqslant k|x_{n-1} - a| \tag{18.78}$$

et par récurrence

$$|x_n - a| \leqslant k^n|x_0 - a|. \tag{18.79}$$

Le cas $|\varphi'(a) > 1|$ se traite de façon similaire. □

Remarque 18.21.
Dans le cas $|\varphi'(a)| = 1$, nous ne pouvons rien conclure. Si $\varphi(x) = \sin(x)$ nous avons $\sin(x) < x$ et le point $a = 0$ est attractif. A contrario, si $\varphi(x) = \sinh(x)$ nous avons $|\sinh(x)| > |x|$ et le point $a = 0$ est répulsif.

18.3.2 Picard

Définition 18.22.
*Une application $f\colon (X, \|.\|_X) \to (Y, \|.\|_Y)$ entre deux espaces métriques est une **contraction** si elle est k-**Lipschitz** pour un certain $0 \leqslant k < 1$, c'est à dire si pour tout $x, y \in X$ nous avons*

$$\|f(x) - f(y)\|_Y \leqslant k\|x - y\|_X. \tag{18.80}$$

Théorème 18.23 (Picard [217, 218] [13].)**.**
Soit X un espace métrique complet et $f\colon X \to X$ une application contractante, de constante de

13. Il me semble qu'à la page 100 de [218], l'hypothèse H1 qui est prouvée ne prouve pas Hn dans le cas $n = 1$. Merci de m'écrire si vous pouvez confirmer ou infirmer. La preuve donnée ici ne contient pas cette «erreur».

Lipschitz k. Alors f admet un unique point fixe, nommé ξ. Ce dernier est donné par la limite de la suite définie par récurrence

$$\begin{cases} x_0 \in X & (18.81a)\\ x_{n+1} = f(x_n). & (18.81b)\end{cases}$$

De plus nous pouvons majorer l'erreur par

$$\|x_n - x\| \leqslant \frac{k^n}{1-k}\|x_n - x_{n-1}\| \leqslant \frac{k^n}{1-k}\|x_1 - x_0\|. \tag{18.82}$$

Soit $r > 0$, $a \in X$ tels que la fonction f laisse la boule $K = \overline{B(a,r)}$ invariante (c'est à dire que f se restreint à $f\colon K \to K$). Nous considérons les suites (u_n) et (v_n) définies par

$$\begin{cases} u_0 = v_0 \in K & (18.83a)\\ u_{n+1} = f(v_n), v_{n+1} \in B(u_n, \epsilon). & (18.83b)\end{cases}$$

Alors le point fixe ξ de f est dans K et la suite (v_n) satisfait l'estimation

$$\|v_n - \xi\| \leqslant \frac{k^n}{1-k}\|u_1 - u_0\| + \frac{\epsilon}{1-k}. \tag{18.84}$$

La première inégalité (18.82) donne une estimation de l'erreur calculable en cours de processus ; la seconde donne une estimation de l'erreur calculable avant de commencer.

Démonstration. Nous commençons par l'unicité du point fixe. Si a et b sont des points fixes, alors $f(a) = a$ et $f(b) = b$. Par conséquent

$$\|f(a) - f(b)\| = \|a - b\|, \tag{18.85}$$

ce qui contredit le fait que f soit une contraction.

En ce qui concerne l'existence, notons que si la suite des x_n converge dans X, alors la limite est un point fixe. En effet en prenant la limite des deux côtés de l'équation $x_{n+1} = f(x_n)$, nous obtenons $\xi = f(\xi)$, c'est à dire que ξ est un point fixe de f. Notons que nous avons utilisé ici la continuité de f, laquelle est une conséquence du fait qu'elle soit Lipschitz. Nous allons donc porter nos efforts à prouver que la suite est de Cauchy (et donc convergente parce que X est complet). Nous commençons par prouver que $\|x_{n+1} - x_n\| \leqslant k^n\|x_0 - x_1\|$. En effet pour tout n nous avons

$$\|x_{n+1} - x_n\| = \|f(x_n) - f(x_{n-1})\| \leqslant k\|x_n - x_{n-1}\|. \tag{18.86}$$

La relation cherchée s'obtient alors par récurrence. Soient $q > p$. En utilisant une somme télescopique,

$$\begin{aligned}\|x_q - x_p\| &\leqslant \sum_{l=p}^{q-1}\|x_{l+1} - x_l\| & (18.87a)\\ &\leqslant \left(\sum_{l=p}^{q-1} k^l\right)\|x_1 - x_0\| & (18.87b)\\ &\leqslant \left(\sum_{l=p}^{\infty} k^l\right)\|x_1 - x_0\|. & (18.87c)\end{aligned}$$

Étant donné que $k < 1$, la parenthèse est la queue d'une série qui converge, et donc tend vers zéro lorsque p tend vers l'infini.

En ce qui concerne les inégalités (18.82), nous refaisons une somme télescopique :

$$\begin{aligned}\|x_{n+p} - x_n\| &\leqslant \|x_{n+p} - x_{n+p-1}\| + \cdots + \|x_{n+1} - x_n\| & (18.88a)\\ &\leqslant k^p\|x_n - x_{n-1}\| + k^{p-1}\|x_n - x_{n-1}\| + \cdots + k\|x_n - x_{n-1}\| & (18.88b)\\ &= k(1 + \cdots + k^{p-1})\|x_n - x_{n-1}\| & (18.88c)\\ &\leqslant \frac{k}{1-k}\|x_n - x_{n-1}\|. & (18.88d)\end{aligned}$$

En prenant la limite $p \to \infty$ nous trouvons

$$\|\xi - x_n\| \leqslant \frac{k}{1-k}\|x_n - x_{n-1}\| \leqslant \frac{k}{1-k}\|x_1 - x_0\|. \tag{18.89}$$

Nous passons maintenant à la seconde partie du théorème en supposant que f se restreigne en une fonction $f\colon K \to K$. D'abord K est encore un espace métrique complet, donc la première partie du théorème s'y applique et f y a un unique point fixe.

Nous allons montrer la relation par récurrence. Tout d'abord pour $n = 1$ nous avons

$$\|v_1 - \xi\| \leqslant \|v_1 - u_1\| + \|u_1 - \xi\| \leqslant \epsilon + \frac{k}{1-k}\|u_1 - u_0\| \tag{18.90}$$

où nous avons utilisé l'estimation (18.89), qui reste valable en remplaçant x_1 par u_1 [14]. Nous pouvons maintenant faire la récurrence :

$$\begin{aligned}
\|v_{n+1} - \xi\| &\leqslant \|v_{n+1} - u_{n+1}\| + \|u_{n+1} - \xi\| && \text{(18.91a)}\\
&\leqslant \epsilon + k\|v_n - \xi\| && \text{(18.91b)}\\
&\leqslant \epsilon + k\left(\frac{k^n}{1-k}\|u_1 - u_0\| + \frac{\epsilon}{1-k}\right) && \text{(18.91c)}\\
&= \frac{\epsilon}{1-k} + \frac{k^{n+1}}{1-k}\|u_1 - u_0\|. && \text{(18.91d)}
\end{aligned}$$

□

Remarque 18.24.
Ce théorème comporte deux parties d'intérêts différents. La première partie est un théorème de point fixe usuel, qui sera utilisé pour prouver l'existence de certaines équations différentielles.

La seconde partie est intéressante d'un point de vie numérique. En effet, ce qu'elle nous enseigne est que si à chaque pas de calcul de la récurrence $x_{n+1} = f(x_n)$ nous commettons une erreur d'ordre de grandeur ϵ, alors le procédé (la suite (v_n)) ne converge plus spécialement vers le point fixe, mais tend vers le point fixe avec une erreur majorée par $\epsilon/(k-1)$.

Remarque 18.25.
Au final l'erreur minimale qu'on peut atteindre est de l'ordre de ϵ. Évidemment si on commet une faute de calcul de l'ordre de ϵ à chaque pas, on ne peut pas espérer mieux.

Remarque 18.26.
Si f elle-même n'est pas contractante, mais si f^p est contractante pour un certain $p \in \mathbb{N}$ alors la conclusion du théorème de Picard reste valide et f a le même unique point fixe que f^p. En effet nommons x le point fixe de f : $f^p(x) = x$. Nous avons alors

$$f^p\big(f(x)\big) = f\big(f^p(x)\big) = f(x), \tag{18.92}$$

ce qui prouve que $f(x)$ est un point fixe de f^p. Par unicité nous avons alors $f(x) = x$, c'est à dire que x est également un point fixe de f.

Si la fonction n'est pas Lipschitz mais presque, nous avons une variante.

Proposition 18.27.
Soit E un ensemble compact [15] *et si $f\colon E \to E$ est une fonction telle que*

$$\|f(x) - f(y)\| < \|x - y\| \tag{18.93}$$

pour tout $x \neq y$ dans E alors f possède un unique point fixe.

14. Elle n'est cependant pas spécialement valable si on remplace x_n par u_n.
15. Notez cette hypothèse plus forte

Démonstration. La suite $x_{n+1} = f(x_n)$ possède une sous suite convergente. La limite de cette sous suite est un point fixe de f parce que f est continue. L'unicité est due à l'aspect strict de l'inégalité (18.93). □

Théorème 18.28 (Équation de Fredholm)**.**
Soit $K\colon [a,b]\times[a,b]\to\mathbb{R}$ et $\varphi\colon [a,b]\to\mathbb{R}$, deux fonctions continues. Alors si λ est suffisamment petit, l'équation

$$f(x) = \lambda\int_a^b K(x,y)f(y)dy + \varphi(x) \tag{18.94}$$

admet une unique solution qui sera de plus continue sur $[a,b]$.

Démonstration. Nous considérons l'ensemble $\mathcal{F}$ des fonctions continues $[a,b]\to[a,b]$ muni de la norme uniforme. Le lemme 12.300 implique que $\mathcal{F}$ est complet. Nous considérons l'application $\Phi\colon \mathcal{F}\to\mathcal{F}$ donnée par

$$\Phi(f)(x) = \lambda\int_a^b K(x,y)f(y)dy + \varphi(x). \tag{18.95}$$

Nous montrons que Φ^p est une application contractante pour un certain p. Pour tout $x\in[a,b]$ nous avons

$$\begin{aligned}\|\Phi(f)-\Phi(g)\|_\infty &\leqslant \|\Phi(f)(x)-\Phi(g)(x)\| &\text{(18.96a)}\\ &= |\lambda|\Big\|\int_a^b K(x,y)\big(f(y)-g(y)\big)dy\Big\| &\text{(18.96b)}\\ &\leqslant |\lambda|\|K\|_\infty|b-a|\|f-g\|_\infty &\text{(18.96c)}\end{aligned}$$

Si λ est assez petit, et si p est assez grand, l'application Φ^p est donc une contraction. Elle possède donc un unique point fixe par le théorème de Picard 18.23. □

18.3.3 Brouwer

Proposition 18.29.
Soit $f\colon [a,b]\to[a,b]$ une fonction continue. Alors f accepte un point fixe.

Démonstration. En effet si nous considérons $g(x) = f(x) - x$ alors nous avons $g(a) = f(a) - a \geqslant 0$ et $g(b) = f(b) - b \leqslant 0$. Si $g(a)$ ou $g(b)$ est nul, la proposition est démontrée ; nous supposons donc que $g(a) > 0$ et $g(b) < 0$. La proposition découle à présent du théorème des valeurs intermédiaires 12.21. □

Exemple 18.30
La fonction $x\mapsto\cos(x)$ est continue entre $[-1,1]$ et $[-1,1]$. Elle admet donc un point fixe. Par conséquent il existe (au moins) une solution à l'équation $\cos(x) = x$. △

Proposition 18.31 (Brouwer dans $\mathbb{R}^n$ version C^∞ via Stokes)**.**
Soit B la boule fermée de centre 0 et de rayon 1 de $\mathbb{R}^n$ et $f\colon B\to B$ une fonction C^∞. Alors f admet un point fixe.

Démonstration. Supposons que f ne possède pas de points fixes. Alors pour tout $x\in B$ nous considérons la ligne droite partant de x dans la direction de $f(x)$ (cette droite existe parce que x et $f(x)$ sont supposés distincts). Cette ligne intersecte ∂B en un point que nous appelons $g(x)$. Prouvons que cette fonction est C^k dès que f est C^k (y compris avec $k=\infty$).

Le point $g(x)$ est la solution du système

$$\begin{cases} g(x) - f(x) = \lambda\big(x - f(x)\big) & \text{(18.97a)}\\ \|g(x)\|^2 = 1 & \text{(18.97b)}\\ \lambda \geqslant 0. & \text{(18.97c)}\end{cases}$$

En substituant nous obtenons l'équation

$$P_x(\lambda) = \|\lambda\big(x - f(x)\big) + f(x)\|^2 - 1 = 0, \tag{18.98}$$

ou encore

$$\lambda^2\|x - f(x)\|^2 + 2\lambda\big(x - f(x)\big) \cdot f(x) + \|f(x)\|^2 - 1 = 0. \tag{18.99}$$

En tenant compte du fait que $\|f(x) < 1\|$ (pare que les images de f sont dans $\mathcal{B}$), nous trouvons que $P_x(0) \leqslant 0$ et $P_x(1) \leqslant 0$. De même $\lim_{\lambda\to\infty} P_x(\lambda) = +\infty$. Par conséquent le polynôme de second degré P_x a exactement deux racines distinctes $\lambda_1 \leqslant 0$ et $\lambda_2 \geqslant 1$. La racine que nous cherchons est la seconde. Le discriminant est strictement positif, donc pas besoin d'avoir peur de la racine dans

$$\lambda(x) = \frac{-\big(x - f(x)\big) \cdot f(x) + \sqrt{\Delta_x}}{\|x - f(x)\|^2} \tag{18.100}$$

où

$$\Delta_x = 4\Big(\big(x - f(x)\big) \cdot f(x)\Big)^2 - 4\|x - f(x)\|^2\big(\|f(x)\|^2 - 1\big). \tag{18.101}$$

Notons que la fonction $\lambda(x)$ est C^k dès que f est C^k ; et en particulier elle est C^∞ si f l'est.

En résumé la fonction g ainsi définie vérifie deux propriétés :

(1) elle est C^∞ ;

(2) elle est l'identité sur ∂B.

La suite de la preuve consiste à montrer qu'une telle rétraction sur B ne peut pas exister [16].

Nous considérons une forme de volume ω sur ∂B : l'intégrale de ω sur ∂B est la surface de ∂B qui est non nulle. Nous avons alors

$$0 < \int_{\partial B} \omega = \int_{\partial B} g^*\omega = \int_B d(g^*\omega) = \int_B g^*(d\omega) = 0 \tag{18.102}$$

Justifications :

— L'intégrale $\int_{\partial B} \omega$ est la surface de ∂B et est donc non nulle.
— La fonction g est l'identité sur ∂B. Nous avons donc $\omega = g^*\omega$.
— Le lemme 14.9.
— La forme ω est de volume, par conséquent de degré maximum et $d\omega = 0$.

□

Un des points délicats est de se ramener au cas de fonctions C^∞. Pour la régularisation par convolution, voir [219] ; pour celle utilisant le théorème de Weierstrass, voir [220].

Théorème 18.32 (Brouwer dans $\mathbb{R}^n$ version continue)**.**
Soit B la boule fermée de centre 0 et de rayon 1 de $\mathbb{R}^n$ et $f\colon B \to B$ une fonction continue [17]*. Alors f admet un point fixe.*

Démonstration. Nous commençons par définir une suite de fonctions

$$f_k(x) = \frac{f(x)}{1 + \frac{1}{k}}. \tag{18.103}$$

Nous avons $\|f_k - f\|_\infty \leqslant \frac{1}{1+k}$ où la norme est la norme uniforme sur B. Par le théorème de Weierstrass 18.5 il existe une suite de fonctions $C^\infty(B, \mathbb{R})$ que nous nommons g_k telles que

$$\|g_k - f_k\|_\infty \leqslant \frac{1}{1+k}. \tag{18.104}$$

16. Notons qu'il n'existe pas non plus de rétractions continues sur B, mais pour le montrer il faut utiliser d'autres méthodes que Stokes, ou alors présenter les choses dans un autre ordre.

17. Une fonction continue sur un fermé de $\mathbb{R}^n$ est à comprendre pour la topologie induite.

Vérifions que cette fonction g_k soit bien une fonction qui prend ses valeurs dans B :

$$\|g_k(x)\| \leqslant \|g_k(x) - f_k(x)\| + \|f_k(x)\| \tag{18.105a}$$
$$\leqslant \frac{1}{1+k} + \frac{\|f(x)\|}{1+\frac{1}{k}} \tag{18.105b}$$
$$\leqslant \frac{1}{1+k} + \frac{1}{1+\frac{1}{k}} \tag{18.105c}$$
$$= 1. \tag{18.105d}$$

Par la version C^∞ du théorème (proposition 18.31), g_k admet un point fixe que l'on nomme x_k.

Étant donné que x_k est dans le compact B, quitte à prendre une sous suite nous supposons que la suite (x_k) converge vers un élément $x \in B$. Nous montrons maintenant que x est un point fixe de f :

$$\|f(x) - x\| = \|f(x) - g_k(x) + g_k(x) - x_k + x_k - x\| \tag{18.106a}$$
$$\leqslant \|f(x) - g_k(x)\| + \underbrace{\|g_k(x) - x_k\|}_{=0} + \|x_k - x\| \tag{18.106b}$$
$$\leqslant \frac{1}{1+k} + \|x_k - x\|. \tag{18.106c}$$

En prenant le limite $k \to \infty$ le membre de droite tend vers zéro et nous obtenons $f(x) = x$. □

18.3.4 Théorème de Schauder

Une conséquence du théorème de Brouwer est le théorème de Schauder qui est valide en dimension infinie.

Théorème 18.33 (Théorème de Schauder[221])**.**
Soit E, un espace vectoriel normé, K un convexe compact de E et $f\colon K \to K$ une fonction continue. Alors f admet un point fixe.

Démonstration. Étant donné que $f\colon K \to K$ est continue, elle y est uniformément continue. Si nous choisissons ϵ alors il existe $\delta > 0$ tel que

$$\|f(x) - f(y)\| \leqslant \epsilon \tag{18.107}$$

dès que $\|x - y\| \leqslant \delta$. La compacité de K permet de choisir un recouvrement fini par des ouverts de la forme

$$K \subset \bigcup_{1 \leqslant i \leqslant p} B(x_j, \delta) \tag{18.108}$$

où $\{x_1, \ldots, x_p\} \subset K$. Nous considérons maintenant $L = \text{Span}\{f(x_j) \text{ tel que } 1 \leqslant j \leqslant p\}$ et

$$K^* = K \cap L. \tag{18.109}$$

Le fait que K et L soient convexes implique que K^* est convexe. L'ensemble K^* est également compact parce qu'il s'agit d'une partie fermée de K qui est compact (lemme 6.75). Notons en particulier que K^* est contenu dans un espace vectoriel de dimension finie, ce qui n'est pas le cas de K.

Nous allons à présent construire une sorte de partition de l'unité subordonnée au recouvrement (18.108) sur K (voir le lemme 14.10). Nous commençons par définir

$$\psi_j(x) = \begin{cases} 0 & \text{si } \|x - x_j\| \geqslant \delta \\ 1 - \frac{\|x - x_j\|}{\delta} & \text{sinon.} \end{cases} \tag{18.110}$$

pour chaque $1 \leqslant j \leqslant p$. Notons que ψ_j est une fonction positive, nulle en-dehors de $B(x_j, \delta)$. En particulier la fonction suivante est bien définie :

$$\varphi_j(x) = \frac{\psi_j(x)}{\sum_{k=1}^{p} \psi_k(x)} \tag{18.111}$$

et nous avons $\sum_{j=1}^{p} \varphi_j(x) = 1$. Les fonctions φ_j sont continues sur K et nous définissons finalement

$$g(x) = \sum_{j=1}^{p} \varphi_j(x) f(x_j). \tag{18.112}$$

Pour chaque $x \in K$, l'élément $g(x)$ est une combinaison des éléments $f(x_j) \in K^*$. Étant donné que K^* est convexe et que la somme des coefficients $\varphi_j(x)$ vaut un, nous avons que g prend ses valeurs dans K^* par la proposition 11.23.

Nous considérons seulement la restriction $g\colon K^* \to K^*$ qui est continue sur un compact contenu dans un espace vectoriel de dimension finie. Le théorème de Brouwer nous enseigne alors que g a un point fixe (proposition 18.32). Nous nommons y ce point fixe. Notons que y est fonction du ϵ choisit au début de la construction, via le δ qui avait conditionné la partition de l'unité.

Nous avons

$$\begin{aligned} f(y) - y &= f(y) - g(y) && \text{(18.113a)}\\ &= \sum_{j=1}^{p} \varphi_j(y) f(y) - \sum_{j=1}^{p} \varphi_j(y) f(x_j) && \text{(18.113b)}\\ &= \sum_{j=1}^{p} \varphi(j)(y)\big(f(y) - f(x_j)\big). && \text{(18.113c)} \end{aligned}$$

Par construction, $\varphi_j(y) \neq 0$ seulement si $\|y - x_j\| \leqslant \delta$ et par conséquent seulement si $\|f(y) - f(x_j)\| \leqslant \epsilon$. D'autre par nous avons $\varphi_j(y) \geqslant 0$; en prenant la norme de (18.113) nous trouvons

$$\|f(y) - y\| \leqslant \sum_{j=1}^{p} \|\varphi_j(y)\big(f(y) - f(x_j)\big)\| \leqslant \sum_{j=1}^{p} \varphi_j(y)\epsilon = \epsilon. \tag{18.114}$$

Nous nous souvenons maintenant que y était fonction de ϵ. Soit y_m le y qui correspond à $\epsilon = 2^{-m}$. Nous avons alors

$$\|f(y_m) - y_m\| \leqslant 2^{-m}. \tag{18.115}$$

L'élément y_m est dans K^* qui est compact, donc quitte à choisir une sous suite nous pouvons supposer que y_m est une suite qui converge vers $y^* \in K$ [18]. Nous avons les majorations

$$\|f(y^*) - y^*\| \leqslant \|f(y^*) - f(y_m)\| + \|f(y_m) - y_m\| + \|y_m - y^*\|. \tag{18.116}$$

Si m est assez grand, les trois termes du membre de droite peuvent être rendus arbitrairement petits, d'où nous concluons que

$$f(y^*) = y^* \tag{18.117}$$

et donc que f possède un point fixe. □

18.3.5 Théorème de Markov-Kakutani et mesure de Haar

Définition 18.34.
*Soit G un groupe topologique. Une **mesure de Haar** sur G est une mesure μ telle que*

(1) $\mu(gA) = \mu(A)$ pour tout mesurable A et tout $g \in G$,

18. Notons que même dans la sous suite nous avons $\|f(y_m) - y_m\| \leqslant 2^{-m}$, avec le même «$m$» des deux côtés de l'inégalité.

(2) $\mu(K) < \infty$ *pour tout compact* $K \subset G$.

Si de plus le groupe G *lui-même est compact nous demandons que la mesure soit normalisée :* $\mu(G) = 1$.

Le théorème suivant nous donne l'existence d'une mesure de Haar sur un groupe compact.

Théorème 18.35 (Markov-Katutani[222])**.**
Soit E *un espace vectoriel normé et* L*, une partie non vide, convexe, fermée et bornée de* E'*. Soit* $T\colon L \to L$ *une application continue. Alors* T *a un point fixe.*

Démonstration. Nous considérons un point $x_0 \in L$ et la suite

$$x_n = \frac{1}{n+1} \sum_{i=0}^{n} T^i x_0. \tag{18.118}$$

La somme des coefficients devant les $T^i(x_0)$ étant 1, la convexité de L montre que $x_n \in L$. Nous considérons l'ensemble

$$C = \bigcap_{n\in\mathbb{N}} \overline{\{x_m \text{ tel que } m \geqslant n\}}. \tag{18.119}$$

Le lemme 6.150 indique que C n'est pas vide, et de plus il existe une sous suite de (x_n) qui converge vers un élément $x \in C$. Nous avons

$$\lim_{n\to\infty} x_{\sigma(n)}(v) = x(v) \tag{18.120}$$

pour tout $v \in E$. Montrons que x est un point fixe de T. Nous avons

$$\|(Tx_{\sigma(k)} - x_{\sigma(k)})v\| = \Big\|T\frac{1}{1+\sigma(k)} \sum_{i=0}^{\sigma(k)} T^i x_0(v) - \frac{1}{1+\sigma(k)} \sum_{i=0}^{\sigma(k)} T^i x_0(v)\Big\| \tag{18.121a}$$

$$= \Big\|\frac{1}{1+\sigma(k)} \sum_{i=0}^{\sigma(k)} T^{i+1} x_0(v) - T^i x_0(v)\Big\| \tag{18.121b}$$

$$= \frac{1}{1+\sigma(k)}\big\|T^{\sigma(k)+1} x_0(v) - x_0(v)\big\| \tag{18.121c}$$

$$\leqslant \frac{2M}{\sigma(k)+1} \tag{18.121d}$$

où $M = \sum_{y\in L} \|y(v)\| < \infty$ parce que L est borné. En prenant $k \to \infty$ nous trouvons

$$\lim_{k\to\infty} \big(Tx_{\sigma(k)} - x_{\sigma(k)}\big)v = 0, \tag{18.122}$$

ce qui signifie que $Tx = x$ parce que T est continue. □

Le théorème suivant est une conséquence du théorème de Markov-Katutani.

Théorème 18.36.
Si G *est un groupe topologique compact possédant une base dénombrable de topologie alors* G *accepte une unique mesure de Haar normalisée. De plus elle est unimodulaire :*

$$\mu(Ag) = \mu(gA) = \mu(A) \tag{18.123}$$

pour tout mesurables $A \subset G$ *et tout élément* $g \in G$.

18.4 Théorèmes de point fixes et équations différentielles

18.4.1 Théorème de Cauchy-Lipschitz

Nous démontrons ici deux théorèmes de Cauchy-Lipschitz. De nombreuses propriétés annexes seront démontrées dans le chapitre sur les équations différentielles, section 30.8.

Théorème 18.37 (Cauchy-Lipschitz[223, 224])**.**
Nous considérons l'équation différentielle

$$\begin{cases} y'(t) = f\big(t, y(t)\big) & \text{(18.124a)} \\ y(t_0) = y_0 & \text{(18.124b)} \end{cases}$$

avec $f\colon U = I \times \Omega \to \mathbb{R}^n$ où I est ouvert dans $\mathbb{R}$ et Ω ouvert dans $\mathbb{R}^n$. Nous supposons que f est continue sur U et localement Lipschitz[19] *par rapport à y.*

Alors il existe un intervalle $J \subset I$ sur lequel la solution au problème est unique. De plus toute solution du problème est une restriction de cette solution à une partie de J. La solution sur J (dite «solution maximale») est de classe C^1.

Démonstration. Nous divisions la preuve en plusieurs étapes (même pas toutes simples).

Cylindre de sécurité et espace fonctionnel Précisons l'espace fonctionnel $\mathcal{F}$ adéquat. Soient V et W les voisinages de t_0 et y_0 sur lesquels f est localement Lipschitz. Nous considérons les quantités suivantes :

(1) $M = \sup_{V\times W} f$;
(2) $r > 0$ tel que $\overline{B(y_0, r)} \subset V$
(3) $T > 0$ tel que $\overline{B(t_0, T)} \subset W$ et $T < r/M$.

Nous considérons alors l'ensemble

$$\mathcal{F} = C^0\big(\overline{B(t_0, T)}, \overline{B(y_0, r)}\big) \tag{18.125}$$

que nous munissons de la norme uniforme. Par le lemme 12.300 l'espace $\big(\mathcal{F}, \|.\|_\infty\big)$ est complet.

Une application $\Phi\colon \mathcal{F} \to \mathcal{F}$ Si y est une solution de l'équation différentielle considérée, elle vérifie

$$y(t) = y_0 + \int_{t_0}^{t} f\big(u, y(u)\big)du. \tag{18.126}$$

Ceci nous incite à considérer l'opérateur $\Phi\colon \mathcal{F} \to \mathcal{F}$ défini par

$$\Phi(y)(t) = y_0 + \int_{t_0}^{t} f\big(u, y(u)\big)du. \tag{18.127}$$

Pour que l'application Φ soit utile nous devons montrer que pour tout $y \in \mathcal{F}$,
— l'application $\Phi(y)$ est bien définie,
— pour tout $t \in \overline{B(y_0, r)}$ nous avons $\Phi(y)(t) \in \overline{B(t_0, T)}$,
— l'application $\Phi(y)\colon \overline{B(t_0, T)} \to \overline{B(y_0, r)}$ est continue.
Attention : nous ne prétendons pas que Φ elle-même soit continue. C'est parti.

$\Phi(y)$ est bien définie Il faut montrer que l'intégrale converge. Le calcul de $\Phi(y)(t)$ ne se fait qu'avec $t \in \overline{B(t_0, T)}$. Vu que u prend ses valeurs dans $[t_0, t]$ et que $y \in \mathcal{F}$, le nombre $y(u)$ est toujours dans $\overline{B(y_0, r)}$. Ceci pour dire que dans l'intégrale, la fonction f n'est considérée que sur $[t_0, t] \times \overline{B(y_0, r)} \subset V \times W$. La fonction f est donc uniformément majorable, et l'intégrale ne pose pas de problèmes.

19. Définition 12.246. Notons que nous ne supposons pas que f soit une contraction.

$\Phi(y)(t) \in \overline{B(t_0,T)}$ Prouvons que $\Phi(y)(t) \in \overline{B(y_0,r)}$. Pour cela, notons que

$$|\Phi(y)(t) - y_0| \leqslant \int_{t_0}^{t} |f\big(u,y(u)\big)|du \leqslant |t-t_0|\|f\|_\infty. \tag{18.128}$$

Étant donné que $t \in \overline{B(t_0,T)}$ nous avons $|t-t_0| \leqslant r/M$ et donc $|\Phi(y)(t) - y_0| \leqslant r$.

$\Phi(y)$ **est continue** Nous pourrions invoquer le théorème 14.62, mais nous allons le faire à la main. Soit $s_0 \in B(t_0,T)$ et prouvons que $\Phi(y)$ est continue en s_0. Pour cela nous prenons $s \in B(s_0,\delta)$ et nous calculons :

$$|\Phi(y)(s) - \Phi(y)(s_0)| \leqslant \int_{s_0}^{s} |f\big(u,y(u)\big)|du \leqslant |s_0-s|\|f\|_\infty. \tag{18.129}$$

C'est le fait que f soit bornée dans le cylindre de sécurité qui fait en sorte que cela tende vers zéro lorsque $s \to s_0$.

L'équation (18.126) signifie que y est un point fixe de Φ. L'espace $\mathcal{F}$ étant complet le théorème de point fixe de Picard (théorème 18.23) s'applique. Nous allons montrer qu'il existe un $p \in \mathbb{N}$ tel que Φ^p soit contractante. Par conséquent Φ^p aura un unique point fixe qui sera également unique point fixe de Φ par la remarque 18.26.

Contractante Prouvons donc que Φ^p est contractante pour un certain p. Pour cela nous commençons par montrer la formule suivante par récurrence :

$$\big\|\Phi^p(x)(t) - \Phi^p(y)(t)\big\| \leqslant \frac{k^p|t-t_0|^p}{p!}\|x-y\|_\infty \tag{18.130}$$

pour tout $x,y \in \mathcal{F}$, et pour tout $t \in \overline{B(t_0,T)}$. Pour $p=0$ la formule (18.130) est vérifiée parce que $\|x-y\|_\infty$ est le supremum de $\|x(t)-y(t)\|$ pour $t \in \overline{B(t_0,T)}$. Supposons que la formule soit vraie pour p et calculons pour $p+1$. Pour tout $t \in \overline{B(t_0,T)}$ nous avons

$$\begin{aligned}
\big\|\Phi^{p+1}(x)(t) - \Phi^{p+1}(y)(t)\big\| &\leqslant \left|\int_{t_0}^{t} \big\|f\big(u,\Phi^p(x)(u)\big) - f\big(u,\Phi^p(y)(u)\big)\big\|du\right| && \text{(18.131a)}\\
&\leqslant \left|\int_{t_0}^{t} k\|\Phi^p(x)(u) - \Phi^p(y)(u)\|du\right| && \text{(18.131b)}\\
&\leqslant \left|\int_{t_0}^{t} k\frac{k^p|t-t_0|}{p!}\|x-y\|_\infty\right| && \text{(18.131c)}\\
&= \frac{k^{p+1}|t-t_0|^{p+1}}{(p+1)!}\|x-y\|_\infty. && \text{(18.131d)}
\end{aligned}$$

Justifications :

— (18.131b) parce que f est Lipschitz.

— (18.131c) par hypothèse de récurrence.

La formule (18.130) est maintenant établie. Nous pouvons maintenant montrer que Φ^p est une contraction pour un certain p. Pour tout $t \in \overline{B(t_0,T)}$ nous avons

$$\|\Phi^p(x)(t) - \Phi^p(y)(t)\| \leqslant \frac{k^p}{t!}|t-t_0|^p\|x-y\|_\infty \leqslant \frac{k^pT^p}{p!}\|x-y\|_\infty \tag{18.132}$$

où nous avons utilisé le fait que $|t-t_0|^p < T^p$. En prenant le supremum sur t des deux côtés il vient

$$\|\Phi^p(x) - \Phi^p(y)\|_\infty \leqslant \frac{k^pT^p}{p!}\|x-y\|_\infty. \tag{18.133}$$

Le membre de droite tend vers zéro lorsque $p \to \infty$ parce que $k<1$ et $T^p/p! \to 0$ [20]. Nous concluons donc que Φ^p est une contraction pour un certain p.

20. C'est le terme général du développement de e^T qui est une série convergente.

Conclusion L'unique point fixe de Φ est alors l'unique solution continue de l'équation différentielle (18.124). Par ailleurs l'équation elle-même $y' = f(t, y)$ demande implicitement que y soit dérivable et donc continue. Nous concluons que l'unique point fixe de Φ est l'unique solution de l'équation différentielle donnée. Cette dernière est automatiquement C^1 parce que si y est continue alors $u \mapsto f(u, y(u))$ est continue, c'est à dire que y' est continue.

Unicité Nous passons maintenant à la partie «prolongement maximum» du théorème. Soient x_1 et x_2 deux solutions maximales du problème (18.124) sur des intervalles I_1 et I_2 respectivement. Les intervalles I_1 et I_2 contiennent $\overline{B(t_0, r)}$ sur lequel $x_1 = x_2$ par unicité.

Nous allons maintenant montrer que pour tout $t \geqslant t_0$ pour lequel x_1 ou x_2 est défini, $x_1(t)$ et $x_2(t)$ sont définis et sont égaux. Le raisonnement sur $t \leqslant t_0$ est similaire.

Supposons que l'ensemble des $t \geqslant t_0$ tels que $x_1 = x_2$ soit ouvert à droite, c'est à dire soit de la forme $[t_0, b[$. Dans ce cas, soit x_1 soit x_2 (soit les deux) cesse d'exister en b. En effet si nous avions les fonctions x_i sur $[t_0, b + \epsilon[$ alors l'équation $x_1 = x_2$ définirait un fermé dans $[t_0, b + \epsilon[$. Supposons pour fixer les idées que x_1 cesse d'exister : le domaine de x_1 (parmi les $t \geqslant 0$) est $[t_0, b[$ et sur ce domaine nous avons $x_1 = x_2$. Dans ce cas x_1 pourrait être prolongé en x_2 au-delà de b. Si x_1 et x_2 s'arrêtent d'exister en même temps en b, alors nous avons bien $x_1 = x_2$.

Nous devons donc traiter le cas où $x_1 = x_2$ sur $[t_0, b]$ alors que x_1 et x_2 existent sur $[t_0, b+\epsilon[$ pour un certain ϵ.

Nous pouvons appliquer le théorème d'existence locale au problème

$$\begin{cases} y' = f(t, y) & \text{(18.134a)} \\ y(b) = x_1(b). & \text{(18.134b)} \end{cases}$$

Il existe un voisinage de b sur lequel la solution est unique. Sur ce voisinage nous devons donc avoir $x_1 = x_2$, ce qui contredit le fait que $x_1 \neq x_2$ en dehors de $[t_0, b]$.

Donc x_1 et x_2 existent et sont égaux sur au moins $I_1 \cup I_2$.

□

Le théorème de Cauchy-Lipschitz donne existence et unicité d'une solution maximale. Cependant cette solution peut ne pas exister partout où les hypothèses sur f sont remplies. En d'autres termes, il peut arriver que f soit Lipschitz jusqu'à t_1, mais que la solution maximale ne soit définie que jusqu'en $t_2 < t_1$. Ce cas fait l'objet du théorème d'explosion en temps fini 30.18.

Sous quelque hypothèses nous pouvons nous assurer de l'existence d'une solution unique sur tout $\mathbb{R}$.

Théorème 18.38 (Cauchy-Lipschitz global[225, 2])**.**
Soit un intervalle I de $\mathbb{R}$, $y_0 \in \mathbb{R}^n$, $t_0 \in I$ et une fonction continue $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ telle que pour tout compact K dans I, il existe $k > 0$ tel que

$$\|f(t, y_1) - f(t, y_2)\| \leqslant k\|y_1 - y_2\| \tag{18.135}$$

pour tout $t \in K$ et $y_1, y_2 \in \mathbb{R}^n$.

Alors le problème

$$\begin{cases} y'(t) = f\big(t, y(t)\big) & \text{(18.136a)} \\ y(t_0) = y_0 & \text{(18.136b)} \end{cases}$$

possède une unique solution $y : I \to \mathbb{R}^n$ sur I.

Démonstration. Soit un intervalle compact K dans I et contenant t_0. Nous notons ℓ le diamètre de K. Sur l'espace $E = C^0(K, \mathbb{R}^n)$ nous considérons la topologie uniforme : $(E, \|.\|_\infty)$. C'est un espace complet par le lemme 12.300 (nous utilisons le fait que $\mathbb{R}^n$ soit complet, théorème 6.57). Nous allons utiliser l'application suivante :

$$\begin{aligned} \Phi &: E \to E \\ \Phi(y)(t) &= y_0 + \int_{t_0}^{t} f\big(s, y(s)\big) ds \end{aligned} \tag{18.137}$$

Démontrons quelque faits à propos de Φ.

La définition fonctionne bien Nous devons commencer par prouver que cette application est bien définie. Si $y \in E$ alors f et y sont continues ; l'application $s \mapsto f\big(s, y(s)\big)$ est donc également continue. L'intégrale de cette fonction sur le compact $[t_0, t]$ ne pose alors pas de problèmes. En ce qui concerne la continuité de $\phi(y)$ sous l'hypothèse que y soit continue,

$$\|\Phi(y)(t) - \Phi(y)(t')\| \leqslant \int_t^{t'} \|f(s, y(s))\| ds \leqslant M|t - t'| \tag{18.138}$$

où M est une majoration de $\|s \mapsto f\big(s, y(s)\big)\|_{\infty,K}$.

Si y est solution alors $\Phi(y) = y$ Supposons que y soit une solution de l'équation différentielle (18.136). Alors, vu que $y'(t) = f\big(t, y(t)\big)$ nous avons :

$$y(t) = y_0 + \int_{t_0}^t y'(s) ds = y_0 + \int_{t_0}^t f\big(s, y(s)\big) ds = \Phi(y)(t). \tag{18.139}$$

Si $\Phi(y) = y$ alors y est solution Nous avons, pour tout t :

$$y(t) = y_0 + \int_{t_0}^t f\big(s, y(s)\big) ds. \tag{18.140}$$

Le membre de droite est dérivable par rapport à t, et la dérivée fait $f\big(t, y(t)\big)$. Donc le membre de gauche est également dérivable et nous avons bien

$$y'(t) = f\big(t, y(t)\big). \tag{18.141}$$

De plus $y(t_0) = y_0 + \int_{t_0}^{t_0} \ldots = y_0$.

Nous sommes encore avec K compact et $E = C^0(K, \mathbb{R}^n)$ muni de la norme uniforme. Nous allons montrer que Φ est une contraction de E pour une norme bien choisie.

Une norme sur E Pour $y \in E$ nous posons

$$\|y\|_k = \max_{t \in K} \big(e^{-k|t-t_0|} \|y(t)\|\big). \tag{18.142}$$

Ce maximum est bien définit et fini parce que la fonction de t dedans est une fonction continue sur le compact K. Cela est également une norme parce que si $\|y\|_k = 0$ alors $e^{-k|t-t_0|}\|y(t)\| = 0$ pour tout t. Étant donné que l'exponentielle ne s'annule pas, $\|y(t)\| = 0$ pour tout t.

Équivalence de norme Nous montrons que les normes $\|.\|_k$ et $\|.\|_\infty$ sont équivalentes[21] :

$$\|y\|_\infty e^{-k\ell} \leqslant \|y\|_k \leqslant \|y\|_\infty \tag{18.143}$$

pour tout $y \in E$. Pour la première inégalité, $\ell \geqslant |t - t_0|$ pour tout $t \in K$, et $k > 0$, donc

$$\|y(t)\| e^{-k\ell} \leqslant e^{-k|t-t_0|} \|y(t)\|. \tag{18.144}$$

En prenant le maximum des deux côtés, $\|y\|_\infty e^{-k\ell} \leqslant \|y\|_k$.

En ce qui concerne la seconde inégalité dans (18.143), $k|t - t_0| \geqslant 0$ et donc $e^{-k|t-t_0|} < 1$.

Vu que les normes $\|.\|_\infty$ et $\|.\|_k$ sont équivalentes, l'espace $(E, \|.\|_k)$ est tout autant complet que $(E, \|.\|_\infty)$. Nous démontrons à présent que Φ est une contraction dans $(E, \|\|_k)$.

Soient $y, z \in E$. Si $t \geqslant t_0$ nous avons

$$\|\Phi(y)(t) - \Phi(z)(t)\| \leqslant \int_{t_0}^t \|f\big(s, y(s)\big) - f\big(s, z(s)\big)\| ds \tag{18.145a}$$

$$\leqslant k \int_{t_0}^t \|y(s) - z(s)\| ds. \tag{18.145b}$$

21. Définition 10.1

Il convient maintenant de remarquer que

$$\|y(t)\| = e^{-k|t-t_0|}e^{k|t-t_0|}\|y(t)\| \leqslant \|y\|_k e^{k|t-t_0|}. \tag{18.146}$$

Nous pouvons avec ça prolonger les inégalités (18.145) par

$$\|\Phi(y)(t) - \Phi(z)(t)\| \leqslant k\|y-z\|_k \int_{t_0}^{t} e^{k|s-t_0|}ds = k\|y-z\|_k \int_{t_0}^{t} e^{k(s-t_0)}ds \tag{18.147}$$

où nous avons utilisé notre supposition $t \geqslant t_0$ pour éliminer les valeurs absolues. L'intégrale peut être faite explicitement, mais nous en sommes arrivés à un niveau de fainéantise tellement inconcevable que

```
1
2  SageMath version 7.3, Release Date: 2016-08-04
3  Type "notebook()" for the browser-based notebook interface.
4  Type "help()" for help.
5
6 sage: var('a,b,k')
7 (a, b, k)
8 sage: f(x)=exp(-k*x)
9 sage: f.integrate(x,a,b)
10 e^(-a*k)/k - e^(-b*k)/k
```

tex/sage/sageSnip014.sage

Au final, si $t \geqslant t_0$,

$$\|\Phi(y)(t) - \Phi(z)(t)\| \leqslant \|y-z\|_k \big(e^{k(t-t_0)} - 1\big). \tag{18.148}$$

Si $t \leqslant t_0$, il faut retourner les bornes de l'intégrale avant d'y faire rentrer la norme parce que $\|\int_0^1 f\| \leqslant \int_0^1 \|f\|$, mais ça ne marche pas avec $\|\int_1^0 f\|$. Pour $t \leqslant t_0$ tout le calcul donne

$$\|\Phi(y)(t) - \Phi(z)(t)\| \leqslant \|y-z\|_k \big(e^{k(t_0-t)} - 1\big). \tag{18.149}$$

Les deux inéquations sont valables a fortiori en mettant des valeurs absolues dans l'exponentielle, de telle sorte que pour tout $t \in K$ nous avons

$$e^{-k|t_0-t|}\|\phi(y)(t) - \Phi(z)(t)\| \leqslant \|y-z\|_k \big(1 - e^{-k|t_0-t|}\big). \tag{18.150}$$

En prenant le supremum sur t,

$$\|\Phi(y) - \Phi(z)\|_k \leqslant \|y-z\|_k (1 - e^{-k\ell}), \tag{18.151}$$

mais $0 < (1 - e^{e-k\ell}) < 1$, donc Φ est contractante pour la norme $\|.\|_k$. Vu que $(E, \|.\|_k)$ est complet, l'application Φ y a un unique point fixe par le théorème de Picard 18.23.

Ce point fixe est donc l'unique solution de l'équation différentielle de départ.

Existence et unicité sur I Il nous reste à prouver que la solution que nous avons trouvée existe sur I : jusqu'à présent nous avons démontré l'existence et l'unicité sur n'importe quel compact dans I.

Soit une suite croissante de compacts K_n contenant t_0 (par exemple une suite exhaustive comme celle du lemme 6.156). Nous avons en particulier

$$I = \bigcup_{n=0}^{\infty} K_n. \tag{18.152}$$

Existence sur I Soit y_n l'unique solution sur K_n. Il suffit de poser

$$y(t) = y_n(t) \tag{18.153}$$

pour n tel que $t \in K_n$. Cette définition fonctionne parce que si $t \in K_n \cap K_m$, il y a forcément un des deux qui est inclus à l'autre et le résultat d'unicité sur le plus grand des deux donne $y_n(t) = y_m(t)$.

Unicité sur I Soient y et z des solutions sur I ; vu que I n'est pas spécialement compact, le travail fait plus haut ne permet pas de conclure que $y = z$.

Soit $t \in I$. Alors $t \in K_n$ pour un certain n et y et z sont des solutions sur K_n qui est compact. L'unicité sur K_n donne $y(t) = z(t)$.

□

18.39.
Il y a d'autres moyens de prouver qu'une solution existe globalement sur $\mathbb{R}$. Si f est globalement bornée, le théorème d'explosion en temps fini donne quelque garanties, voir 30.20.

Le théorème suivant donne une version du théorème de Cauchy-Lipschitz lorsque la fonction f dépend d'un paramètre. Ce théorème n'utilise rien de fondamentalement nouveau. Nous le donnons seulement pour montrer que l'on peut choisir l'espace $\mathcal{F}$ de façon un peu maligne pour élargir le résultat. Si vous voulez un théorème de Cauchy-Lipschitz avec paramètre vraiment intéressant, allez voir le théorème 30.31.

Théorème 18.40 (Cauchy-Lipschitz avec paramètre[1, 226])**.**
Soit un intervalle ouvert I de $\mathbb{R}$, un connexe ouvert Ω de $\mathbb{R}^n$ et un intervalle ouvert Λ de $\mathbb{R}^d$. Soit une fonction $f\colon I \times \Omega \times \Lambda \to \mathbb{R}^n$ continue et localement Lipschitz en Ω. Soient $t_0 \in I$, $y_0 \in \Omega$ et $\lambda_0 \in \Lambda$. Il existe un voisinage compact de (t_0, y_0, λ_0) sur lequel le problème

$$\begin{cases} y'_\lambda(t) = f\big(t, y_\lambda(t), \lambda\big) & \text{(18.154a)} \\ y_\lambda(t_0) = y_0 & \text{(18.154b)} \end{cases}$$

possède une unique solution. De plus $(t, \lambda) \mapsto y_\lambda(t)$ est continue[22]*.*

Démonstration. **Problèmes et choses à faire**
Ceci est une idée de la preuve. Je n'ai pas vérifié toutes les étapes. Soyez prudent.

D'abord nous avons un voisinage compact $V \times \overline{B(y_0, r)} \times \Lambda_0$ de (t_0, y_0, λ_0) sur lequel f est bornée. Ensuite nous récrivons l'équation différentielle sous la forme

$$\begin{cases} \dfrac{\partial y}{\partial t}(t, \lambda) = f\big(t, y(t, \lambda), \lambda\big) & \text{(18.155a)} \\ y(t_0, \lambda) = y_0. & \text{(18.155b)} \end{cases}$$

pour une fonction $y\colon V \times \Lambda_0 \to \mathbb{R}^n$.

Nous posons $\mathcal{F} = C^0\big(V \times \Lambda_0, \mathbb{R}^n\big)$ et nous y définissons l'application

$$\begin{aligned} \Phi\colon \mathcal{F} &\to \mathcal{F} \\ \Phi(y)(t, \lambda) &= y_0 + \int_{t_0}^{t} f\big(s, y(s, \lambda), \lambda\big) ds. \end{aligned} \tag{18.156}$$

Il y a plein de vérifications à faire[226], mais je parie que Φ est bien définie, et que une de ses puissances est une contraction de $(\mathcal{F}, \|.\|_\infty)$. L'unique point fixe est une solution de notre problème et est dans C^0, donc $(t, \lambda) \mapsto y(t, \lambda) = y_\lambda(t)$ est de classe C^0, c'est à dire continue. □

22. Ici, la surprise est que ce soit continu par rapport à λ. Le fait qu'elle le soit par rapport à t est clair depuis le départ parce que c'est finalement rien d'autre que le Cauchy-Lipschitz vieux et connu.

18.41.
Ce théorème marque un peu la limite de ce que l'on peut faire avec la méthode des points fixes dans le cadre de Cauchy-Lipschitz : nous sommes limités à la continuité de la solution parce que les espaces C^p ne sont pas complets[23]. Il n'y a donc pas d'espoir d'adapter la méthode pour prouver que si f est de classe C^p alors $(t,\lambda) \mapsto y_\lambda(t)$ est de classe C^p. On peut, à λ fixé prouver que $t \mapsto y_\lambda(t)$ est de classe C^p (utiliser une récurrence), mais pas plus.

La régularité C^1 de y par rapport à la condition initiale sera l'objet du théorème 30.28. Ce résultat n'est vraiment pas facile et utilise des ingrédients bien autres qu'un point fixe. Ensuite la régularité C^p par rapport à la condition initiale et par rapport à un paramètre seront presque des cadeaux (proposition 30.29 et 30.31).

Exemple 18.42([227])
Nous savons que le théorème de Picard permet de trouver le point fixe par itération de la contraction à partir d'un point quelconque. Tentons donc de résoudre

$$\begin{cases} y'(t) = y(t) & \text{(18.157a)} \\ y(0) = 1 & \text{(18.157b)} \end{cases}$$

dont nous savons depuis l'enfance que la solution est l'exponentielle. Partons donc de la fonction constante $y_0 = 1$, et appliquons la contraction (18.137) :

$$u_1 = 1 + \int_0^1 u_0(s)ds = 1 + t. \tag{18.158}$$

Ensuite

$$u_2 = 1 + \int_0^t (1+s)ds = 1 + t + \frac{t^2}{2}. \tag{18.159}$$

Et on voit que les itérations suivantes vont donner l'exponentielle.

Nous sommes évidemment en droit de se dire que nous avons choisi un bon point de départ. Tentons le coup avec une fonction qui n'a rien à voir avec l'exponentielle : $u_0(x) = \sin(x)$.

Le programme suivant permet de faire de belles investigations numériques en partant d'à peu près n'importe quelle fonction :

```
#! /usr/bin/sage -python
# -*- coding: utf8 -*-

from sage.all import *

x=var('x')

def Phi(f):
    prim=f.integrate()
    return 1+prim(x)-prim(0)

f=sin(x)

for i in range(1,30):
    print(i,f)
    f=Phi(f)

g=f(x)-exp(x)
plot(g,(x,-10,10)).show()
```

23. Par exemple, le théorème de Stone-Weierstrass 18.6 nous dit que la limite uniforme de polynômes (de classe C^∞) peut n'être que continue. Voir aussi le thème 22.

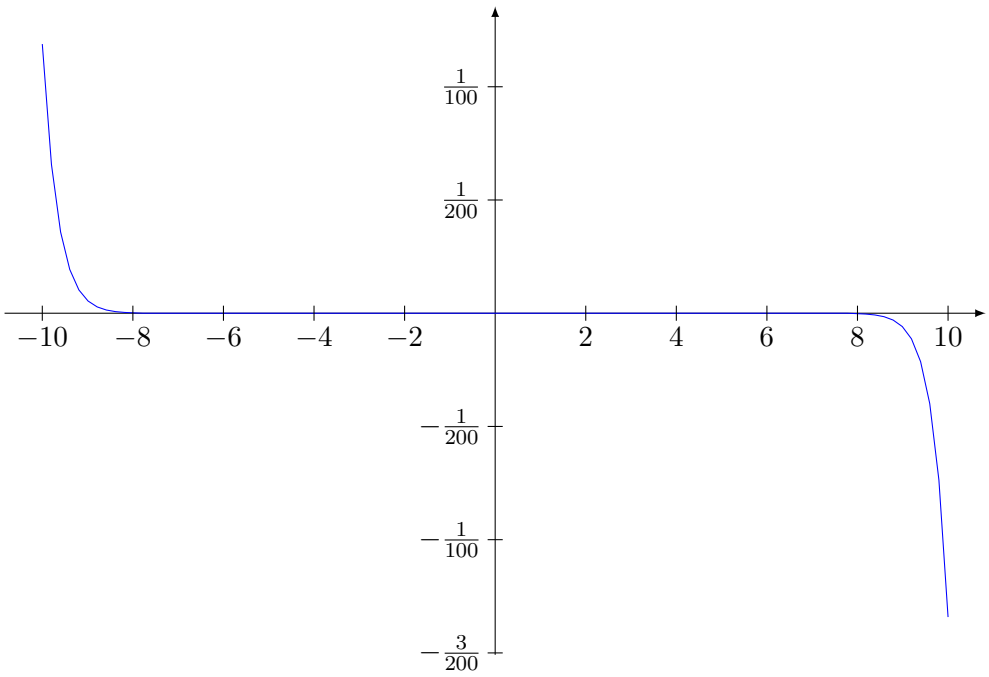
tex/sage/picard_exp.py

Ce programme fait 30 itérations depuis la fonction $\sin(x)$ pour tenter d'approximer $\exp(x)$. Pour donner une idée, après 7 itérations nous avons la fonction suivante :

$$\frac{1}{60}x^5 + \frac{1}{24}x^4 + \frac{1}{2}x^2 + 2x - \sin(x) + 1. \tag{18.160}$$

Nous voyons que les coefficients sont des factorielles, mais pas toujours celles correspondantes à la puissance, et qu'il manque certains termes par rapport au développement de l'exponentielle que nous connaissons. Bref, le polynôme qui se met en face de $\sin(x)$ s'adapte tout seul pour compenser.

Et après 30 itérations, ça donne quoi ? Voici un graphe de l'erreur entre $u_{30}(x)$ et $\exp(30)$:



Pour donner une idée, $\exp(10) \simeq 22000$. Donc il y a une faute de 0.01 sur 22000. Pas mal.

△

18.4.2 Théorème de Cauchy-Arzella

Théorème 18.43 (Cauchy-Arzela[217])**.**
Nous considérons le système d'équation différentielles

$$\begin{cases} y' = f(t,y) & \text{(18.161a)} \\ y(t_0) = y_0. & \text{(18.161b)} \end{cases}$$

avec $f\colon U \to \mathbb{R}^n$, continue où U est ouvert dans $\mathbb{R} \times \mathbb{R}^n$. Alors il existe un voisinage fermé V de t_0 sur lequel une solution C^1 du problème (18.161) *existe.*

Idée de la démonstration. Nous considérons $M = \|f\|_\infty$ et K, l'ensemble des fonctions M-Lipschitz sur U. Nous prouvons que $(K, \|.\|_\infty)$ est compact. Ensuite nous considérons l'application

$$\begin{aligned} \Phi\colon K &\to K \\ \Phi(f)(t) &= x_0 + \int_{t_0}^{t} f\big(u, f(u)\big)du. \end{aligned} \tag{18.162}$$

Après avoir prouvé que Φ était continue, nous concluons qu'elle a un point fixe par le théorème de Schauder 18.33. □

Remarque 18.44.
Quelque remarques.

(1) Les théorème de Cauchy-Lipschitz et Cauchy-Arzella donnent des existences pour des équations différentielles du type $y' = f(t, y)$. Et si nous avons une équation du second ordre ? Alors il y a la méthode de la réduction de l'ordre qui permet de transformer une équation différentielle d'ordre élevé en un système d'ordre 1.

(2) Ces théorèmes posent des *conditions initiales* : la valeur de y est donnée en un point, et la méthode de la réduction de l'ordre permet de donner l'existence de solutions d'un problème d'ordre k en donnant les valeurs de $y(0)$, $y'(0)$, $\dots y^{(k-1)}(0)$. C'est à dire de la fonction et de ses dérivées en un point. Rien n'est dit sur l'existence de *conditions aux bords*.

Ces deux points sont illustrés dans les exemples 30.15 et 30.16.

18.5 Théorèmes d'inversion locale et de la fonction implicite

18.5.1 Mise en situation

Dans un certain nombre de situation, il n'est pas possible de trouver des solutions explicites aux équations qui apparaissent. Néanmoins, l'existence «théorique» d'une telle solution est souvent déjà suffisante. C'est l'objet du théorème de la fonction implicite.

Prenons par exemple la fonction sur $\mathbb{R}^2$ donnée par

$$F(x, y) = x^2 + y^2 - 1. \tag{18.163}$$

Nous pouvons bien entendu regarder l'ensemble des points donnés par $F(x, y) = 0$. C'est le cercle dessiné à la figure 18.1.

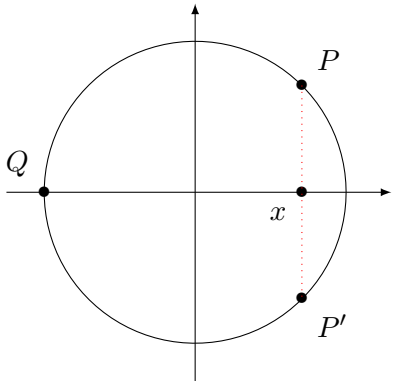


FIGURE 18.1 – Un cercle pour montrer l'intérêt de la fonction implicite. Si on donne x, nous ne pouvons pas savoir si nous parlons de P ou de P'.

Nous ne pouvons pas donner le cercle sous la forme $y = y(x)$ à cause du $\pm$ qui arrive quand on prend la racine carrée. Mais si on se donne le point P, nous pouvons dire que *autour de* P, le cercle est la fonction

$$y(x) = \sqrt{1 - x^2}. \tag{18.164}$$

Tandis que autour du point P', le cercle est la fonction

$$y(x) = -\sqrt{1 - x^2}. \tag{18.165}$$

Autour de ces deux point, donc, le cercle est donné par une fonction. Il n'est par contre pas possible de donner le cercle autour du point Q sous la forme d'une fonction.

Ce que nous voulons faire, en général, est de voir si l'ensemble des points tels que

$$F(x_1, \dots, x_n, y) = 0 \tag{18.166}$$

peut être donné par une fonction $y = y(x_1, \dots, x_n)$. En d'autre termes, est-ce qu'il existe une fonction $y(x_1, \dots, x_n)$ telle que

$$F\big(x_1, \dots, x_n, y(x_1, \dots, x_n)\big) = 0. \tag{18.167}$$

Plus généralement, soit une fonction

$$\begin{aligned} F\colon D \subset \mathbb{R}^n \times \mathbb{R}^m &\to \mathbb{R}^m \\ (x, y) &\mapsto \big(F_1(x, y), \dots, F_m(x, y)\big) \end{aligned} \tag{18.168}$$

avec $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_m)$. Pour chaque x fixé, on s'intéresse aux solutions du système de m équations $F(x, y) = 0$ pour les inconnues y ; en particulier, on voudrait pouvoir écrire $y = \varphi(x)$ vérifiant $F(x, \varphi(x)) = 0$.

18.5.2 Théorème d'inversion locale

Lemme 18.45.
Soit E un espace de Banach (métrique complet) et $\mathcal{O}$ un ouvert de E. Nous considérons une λ-contraction $\varphi\colon \mathcal{O} \to E$. Alors l'application

$$f\colon x \mapsto x + \varphi(x) \tag{18.169}$$

est un homéomorphisme entre $\mathcal{O}$ et un ouvert de E. De plus f^{-1} est Lipschitz de constante plus petite ou égale à $(1 - \lambda)^{-1}$.

Cette proposition utilise le théorème de point fixe de Picard 18.23, et sera utilisée pour démontrer le théorème d'inversion locale 18.47.

Démonstration. Soient $x_1, x_2 \in \mathcal{O}$. Nous posons $y_1 = f(x_1)$ et $y_2 = f(x_2)$. En vertu de l'inégalité de la proposition 6.32 nous avons

$$\|f(x_2) - f(x_1)\| = \|x_2 + \varphi(x_2) - x_1 - \varphi(x_1)\| \tag{18.170a}$$
$$\geqslant \Big|\|x_2 - x_1\| - \|\varphi(x_2) - \varphi(x_1)\|\Big| \tag{18.170b}$$
$$\geqslant (1 - \lambda)\|x_2 - x_1\|. \tag{18.170c}$$

À la dernière ligne les valeurs absolues sont enlevées parce que nous savons que ce qui est à l'intérieur est positif. Cela nous dit d'abord que f est injective parce que $f(x_2) = f(x_1)$ implique $x_2 = x_1$. Donc f est inversible sur son image. Nous posons $A = f(\mathcal{O})$ et nous devons prouver que que $f^{-1}\colon A \to \mathcal{O}$ est continue, Lipschitz de constante majorée par $(1 - \lambda)^{-1}$ et que A est ouvert.

Les inéquations (18.170) nous disent que

$$\|f^{-1}(y_1) - f^{-1}(y_2)\| \leqslant \frac{\|y_1 - y_2\|}{1 - \lambda}, \tag{18.171}$$

c'est à dire que

$$f^{-1}\big(B(y, r)\big) \subset B\big(f^{-1}(y), \frac{r}{1 - \lambda}\big), \tag{18.172}$$

ce qui signifie que f^{-1} est Lipschitz de constante souhaitée et donc continue.

Il reste à prouver que $f(\mathcal{O})$ est ouvert. Pour cela nous prenons $y_0 = f(x_0)$ dans $f(\mathcal{O})$ est nous prouvons qu'il existe ϵ tel que $B(y_0, \epsilon)$ soit dans $f(\mathcal{O})$. Il faut donc que pour tout $y \in B(y_0, \epsilon)$, l'équation $f(x) = y$ ait une solution. Nous considérons l'application

$$L_y\colon x \mapsto y - \varphi(x). \tag{18.173}$$

Ce que nous cherchons est un point fixe de L_y parce que si $L_y(x) = x$ alors $y = x + \varphi(x) = f(x)$. Vu que

$$\|L_y(x) - L_y(x')\| = \|\varphi(x) - \varphi(x')\| \leqslant \lambda\|x - x'\|, \tag{18.174}$$

l'application L_y est une contraction de constante λ. Par ailleurs x_0 est un point fixe de L_{y_0}, donc en vertu de la caractérisation (12.635) des fonctions Lipschitziennes,

$$L_{y_0}\big(\overline{B(x_0,\delta)}\big) \subset \overline{B\big(L_{y_0}(x_0),\lambda\delta\big)} = \overline{B(x_0,\lambda\delta)}. \tag{18.175}$$

Vu que pour tout y et x nous avons $L_y(x) = L_{y_0}(x) + y - y_0$,

$$L_y\big(\overline{B(x_0,\delta)}\big) = L_{y_0}\big(\overline{B(x_0,\delta)}\big) + (y-y_0) \subset \overline{B(x_0,\lambda\delta)} + (y-y_0) \subset \overline{B(x_0),\lambda\delta + \|y-y_0\|}. \tag{18.176}$$

Si $\epsilon < (1-\lambda)\delta$ alors $\lambda\delta + \|y-y_0\| < \delta$. Un tel choix de $\epsilon > 0$ est possible parce que $\lambda < 1$. Pour une telle valeur de ϵ nous avons

$$L_y\big(\overline{B(x_0,\delta)}\big) \subset \overline{B(x_0,\delta)}. \tag{18.177}$$

Par conséquent L_y est une contraction sur l'espace métrique complet $\overline{B(x_0,\delta)}$, ce qui signifie que L_y y possède un point fixe par le théorème de Picard 18.23. □

Le théorème d'inversion locale s'énonce de la façon suivante dans $\mathbb{R}^n$:

Théorème 18.46 (Inversion locale dans $\mathbb{R}^n$)**.**
Soit $f \in C^k(\mathbb{R}^n,\mathbb{R}^n)$ et $x_0 \in \mathbb{R}^n$. Si df_{x_0} est inversible, alors il existe un voisinage ouvert U de x_0 et V de $f(x_0)$ tels que $f\colon U \to V$ soit un C^k-difféomorphisme. (c'est à dire que f^{-1} est également de classe C^k)

Nous allons le démontrer dans le cas un peu plus général (mais pas plus cher[24]) des espaces de Banach en tant que conséquence du théorème de point fixe de Picard 18.23.

Théorème 18.47 (Inversion locale dans un espace de Banach[228])**.**
Soit une fonction $f \in C^p(E,F)$ avec $p \geqslant 1$ entre deux espaces de Banach. Soit $x_0 \in E$ tel que df_{x_0} soit une bijection bicontinue[25]. Alors il existe un voisinage ouvert V de x_0 et W de $f(x_0)$ tels que

(1) $f\colon V \to W$ soit une bijection,

(2) $f^{-1}\colon W \to V$ soit de classe C^p.

Démonstration. Nous commençons par simplifier un peu le problème. Pour cela, nous considérons la translation $T\colon x \mapsto x + x_0$ et l'application linéaire

$$\begin{aligned} L\colon \mathbb{R}^n &\to \mathbb{R}^n \\ x &\mapsto (df_{x_0})^{-1}x \end{aligned} \tag{18.178}$$

qui sont tout deux des difféomorphismes (L en est un par hypothèse d'inversibilité). Quitte à travailler avec la fonction $k = L \circ f \circ T$, nous pouvons supposer que $x_0 = 0$ et que $df_{x_0} = \mathbb{1}$. Pour comprendre cela il faut utiliser deux fois la formule de différentielle de fonction composée de la proposition 12.480 :

$$dk_0(u) = dL_{(f\circ T)(0)}\Big(df_{T(0)}dT_0(u)\Big). \tag{18.179}$$

Vu que L est linéaire, sa différentielle est elle-même, c'est à dire $dL_{(f\circ T)(0)} = (df_{x_0})^{-1}$, et par ailleurs $dT_0 = \mathbb{1}$, donc

$$dk_0(u) = (df_{x_0})^{-1}\Big(df_{x_0}(u)\Big) = u, \tag{18.180}$$

ce qui signifie bien que $dk_0 = \mathbb{1}$. Pour tout cela nous avons utilisé en plein le fait que df_{x_0} était inversible.

Nous posons $g = f - \mathbb{1}$, c'est à dire $g(x) = f(x) - x$, qui a la propriété $dg_0 = 0$. Étant donné que g est de classe C^1, l'application[26]

$$\begin{aligned} dg\colon E &\to \mathrm{GL}(F) \\ x &\mapsto dg_x \end{aligned} \tag{18.181}$$

24. Sauf la justification de la régularité de l'application $A \mapsto A^{-1}$

25. En dimension finie, une application linéaire est toujours continue et d'inverse continu.

26. Ici $\mathrm{GL}(F)$ est l'ensemble des applications linéaires, inversibles et continues de F dans lui-même. Ce ne sont pas spécialement des matrices parce que nous n'avons pas d'hypothèses sur la dimension de F, finie ou non.

est continue. En conséquence de quoi nous avons un voisinage U' de 0 pour lequel

$$\sup_{x\in U'} \|dg_x\| < \frac{1}{2}. \tag{18.182}$$

Maintenant le théorème des accroissements finis 12.229 (12.241 pour la dimension finie) nous indique que pour tout $x, x' \in U'$ nous avons [27]

$$\|g(x') - g(x)\| \leqslant \sup_{a\in[x,x']} \|dg_a\| \cdot \|x - x'\| \leqslant \frac{1}{2}\|x - x'\|, \tag{18.183}$$

ce qui prouve que g est une contraction au moins sur l'ouvert U'. Nous allons aussi donner une idée de la façon dont f fonctionne : si $x_1, x_2 \in U'$ alors

$$\|x_1 - x_2\| = \|g(x_1) - f(x_1) - g(x_2) + f(x_2)\| \tag{18.184a}$$
$$\leqslant \|g(x_1) - g(x_2)\| + \|f(x_1) - f(x_2)\| \tag{18.184b}$$
$$\leqslant \frac{1}{2}\|x_1 - x_2\| + \|f(x_1) - f(x_2)\|, \tag{18.184c}$$

ce qui montre que

$$\|x_1 - x_2\| \leqslant 2\|f(x_1) - f(x_2)\|. \tag{18.185}$$

Maintenant que nous savons que g est contractante de constante $\frac{1}{2}$ et que $f = g + \mathbb{1}$ nous pouvons utiliser la proposition 18.45 pour conclure que f est un homéomorphisme sur un ouvert U (partie de U') de E et f^{-1} a une constante de Lipschitz plus petite ou égale à $(1 - \frac{1}{2})^{-1} = 2$.

Nous allons maintenant prouver que f^{-1} est différentiable et que sa différentielle est donnée par $(df^{-1})_{f(x)} = (df_x)^{-1}$.

Soient $a, b \in U$ et $u = b - a$. Étant donné que f est différentiable en a, il existe une fonction $\alpha \in o(\|u\|)$ telle que

$$f(b) - f(a) - df_a(u) = \alpha(u). \tag{18.186}$$

En notant $y_a = f(a)$ et $y_b = f(b)$ et en appliquant $(df_a)^{-1}$ à cette dernière équation,

$$(df_a)^{-1}(y_b - y_a) - u = (df_a)^{-1}\big(\alpha(u)\big). \tag{18.187}$$

Vu que df_a est bornée (et son inverse aussi), le membre de droite est encore une fonction β ayant la propriété $\lim_{u\to 0} \beta(u)/\|u\| = 0$; en réordonnant les termes,

$$b - a = (df_a)^{-1}(y_b - y_a) + \beta(u) \tag{18.188}$$

et donc

$$f^{-1}(y_b) - f^{-1}(y_a) - (df_a)^{-1}(y_b - y_a) = \beta(u), \tag{18.189}$$

ce qui prouve que f^{-1} est différentiable et que $(df^{-1})_{y_a} = (df_a)^{-1}$.

La différentielle df^{-1} est donc obtenue par la chaine

$$df^{-1}\colon f(U) \xrightarrow{f^{-1}} U' \xrightarrow{df} \mathrm{GL}(F) \xrightarrow{\mathrm{Inv}} \mathrm{GL}(F) \tag{18.190}$$

où l'application $\mathrm{Inv}\colon \mathrm{GL}(F) \to \mathrm{GL}(F)$ est l'application $X \mapsto X^{-1}$ qui est de classe C^∞ par le théorème 12.232. D'autre part, par hypothèse df est une application de classe C^{k-1} et donc au minimum C^0 parce que $k \geqslant 1$. Enfin, l'application $f^{-1}\colon f(U) \to U$ est continue (parce que la proposition 18.45 précise que f est un homéomorphisme). Donc toute la chaine est continue et df^{-1} est continue. Cela entraine immédiatement que f^{-1} est C^1 et donc que toute la chaine est C^1.

Par récurrence nous obtenons la chaine

$$df^{-1}\colon f(U) \xrightarrow[C^{k-1}]{f^{-1}} U' \xrightarrow[C^{k-1}]{df} \mathrm{GL}(F) \xrightarrow[C^\infty]{\mathrm{Inv}} \mathrm{GL}(F) \tag{18.191}$$

qui prouve que df^{-1} est C^{k-1} et donc que f^{-1} est C^k. La récurrence s'arrête ici parce que df n'est pas mieux que C^{k-1}. □

27. Ici nous supposons avoir choisi U' convexe afin que tous les $a \in [x, x']$ soient bien dans U' et donc soumis à l'inéquation (18.182), ce qui est toujours possible, il suffit de prendre une boule.

18.5.3 Théorème de la fonction implicite

Nous énonçons et le démontrons le théorème de la fonction implicite dans le cas d'espaces de Banach.

Théorème 18.48 (Théorème de la fonction implicite dans Banach[141])**.**
Soient E, F et G des espaces de Banach et des ouverts $U \subset E$, $V \subset F$. Nous considérons une fonction $f\colon U \times V \to G$ de classe C^r telle que[28]

$$d_y f_{(x_0,y_0)}\colon F \to G \tag{18.192}$$

soit un isomorphisme pour un certain $(x_0, y_0) \in U \times V$.

Alors nous avons des voisinages U_0 de x_0 dans E et W_0 de $f(x_0, y_0)$ dans G et une fonction de classe C^r

$$g\colon U_0 \times W_0 \to V \tag{18.193}$$

telle que

$$f\big(x, g(x, w)\big) = w \tag{18.194}$$

pour tout $(x, w) \in U_0 \times W_0$.

Cette fonction g est unique au sens suivant : il existe un voisinage V_0 de y_0 tel que si $(x, y) \in U_0 \times V_0$ et $w \in W_0$ satisfont à $f(x, y) = w$ alors $y = g(x, w)$. Autrement dit, la fonction $g\colon U_0 \times W_0 \to V_0$ est unique.

Démonstration. Nous commençons par considérer la fonction

$$\begin{aligned} \Phi\colon U \times V &\to E \times G \\ (x, y) &\mapsto \big(x, f(x, y)\big) \end{aligned} \tag{18.195}$$

et sa différentielle

$$d\Phi_{(x_0,y_0)}(u, v) = \frac{d}{dt}\Big[\big(x_0 + tu, f(x_0 + tu, y_0 + tv)\big)\Big]_{t=0} \tag{18.196a}$$

$$= \left(\frac{d}{dt}\Big[x_0 + tu\Big]_{t=0}, \frac{d}{dt}\Big[f(x_0 + tu, y_0 + tv)\Big]_{t=0}\right) \tag{18.196b}$$

$$= \big(u, df_{(x_0,y_0)}(u, v)\big)\,. \tag{18.196c}$$

Nous utilisons alors la proposition 12.227 pour conclure que

$$d\Phi_{(x_0,y_0)}(u, v) = \big(u, (d_1 f)_{(x_0,y_0)}(u) + (d_2 f)_{(x_0,y_0)}(v)\big), \tag{18.197}$$

mais comme par hypothèse $(d_2 f)_{(x_0,y_0)}\colon F \to G$ est un isomorphisme, l'application $d\Phi_{(x_0,y_0)}\colon E \times F \to E \times G$ est également un isomorphisme. Par conséquent le théorème d'inversion locale 18.47 nous indique qu'il existe un voisinage $\mathcal{O}$ de (x_0, y_0) et $\mathcal{P}$ de $\Phi(x_0, y_0)$ tels que $\Phi\colon \mathcal{O} \to \mathcal{P}$ soit une bijection et $\Phi^{-1}\colon \mathcal{P} \to \mathcal{O}$ soit de classe C^r. Vu que $\mathcal{P}$ est un voisinage de

$$\Phi(x_0, y_0) = \big(x_0, f(x_0, y_0)\big), \tag{18.198}$$

nous pouvons par 10.78 le choisir un peu plus petit de telle sorte à avoir $\mathcal{P} = U_0 \times W_0$ où U_0 est un voisinage de x_0 et W_0 un voisinage de $f(x_0, y_0)$. Dans ce cas nous devons obligatoirement aussi restreindre $\mathcal{O}$ à $U_0 \times V_0$ pour un certain voisinage V_0 de y_0. L'application Φ^{-1} a obligatoirement la forme

$$\begin{aligned} \Phi^{-1}\colon U_0 \times W_0 &\to U_0 \times V_0 \\ (x, w) &\mapsto \big(x, g(x, w)\big) \end{aligned} \tag{18.199}$$

28. La notation d_y est la différentielle partielle de la définition 12.226.

pour une certaine fonction $g\colon U_0 \times W_0 \to V$. Cette fonction g est la fonction cherchée parce qu'en appliquant Φ à (18.199),

$$(x, w) = \Phi\big(x, g(x, w)\big) = \Big(x, f\big(x, g(x, w)\big)\Big), \tag{18.200}$$

qui nous dit que pour tout $x \in U_0$ et tout $w \in W_0$ nous avons

$$f\big(x, g(x, w)\big) = w. \tag{18.201}$$

Si vous avez bien suivi le sens de l'équation (18.199) alors vous avez compris l'unicité. Sinon, considérez $(x, y) \in U_0 \times V_0$ et $w \in W_0$ tels que $f(x, y) = w$. Alors $\big(x, f(x, y)\big) = (x, w)$ et

$$\Phi(x, y) = (x, w). \tag{18.202}$$

Mais vu que $\Phi\colon U_0 \times V_0 \to U_0 \times W_0$ est une bijection, cette relation définit de façon univoque l'élément (x, y) de $U_0 \times V_0$, qui ne sera autre que $g(x, w)$. □

Le théorème de la fonction implicite s'énonce de la façon suivante pour des espaces de dimension finie.

Théorème 18.49 (Théorème de la fonction implicite en dimension finie)**.**
Soit une fonction $F\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ de classe C^k et $(\alpha, \beta) \in \mathbb{R}^n \times \mathbb{R}^m$ tels que

(1) $F(\alpha, \beta) = 0$,

(2) $\frac{\partial(F_1, \ldots, F_m)}{\partial(y_1, \ldots, y_m)} \neq 0$, *c'est à dire que* $(d_y F)_{(\alpha,\beta)}$ *est inversible.*

Alors il existe un voisinage ouvert V de α dans $\mathbb{R}^n$, un voisinage ouvert W de β dans $\mathbb{R}^m$ et une application $\varphi\colon V \to W$ de classe C^k telle que pour tout $x \in V$ on ait

$$F\big(x, \varphi(x)\big) = 0. \tag{18.203}$$

De plus si $(x, y) \in V \times W$ satisfait à $F(x, y) = 0$, alors $y = \varphi(x)$.

Remarque 18.50.
Notons que cet énoncé est tourné un peu différemment en ce qui concerne le nombre de variables dont dépend la fonction implicite : comparez

$$f\big(x, g(x, w)\big) = w \tag{18.204a}$$
$$F\big(x, \varphi(x)\big) = 0. \tag{18.204b}$$

Le deuxième est un cas particulier du premier en posant

$$F(x, y) = f(x, y) - f(x_0, y_0) \tag{18.205}$$

et donc en considérant w comme valant la constante $f(x_0, y_0)$; dans ce cas la fonction g ne dépend plus que de la variable x.

Exemple 18.51
La remarque 18.50 signifie entre autres que le théorème 18.48 est plus fort que 18.49 parce que le premier permet de choisir la valeur d'arrivée. Parlons de l'exemple classique du cercle et de la fonction $f(x, y) = x^2 + y^2$. Nous savons que

$$f(\alpha, \beta) = 1. \tag{18.206}$$

Alors le théorème 18.48 nous donne une fonction g telle que

$$f(x, g(x, r)) = r \tag{18.207}$$

tant que x est proche de α, que r est proche de 1 et que g donne des valeurs proches de β.

L'énoncé 18.49 nous oblige à travailler avec la fonction $F(x,y) = x^2 + y^2 - 1$, de telle sorte que

$$F(\alpha, \beta) = 0, \tag{18.208}$$

et que nous ayons une fonction φ telle que

$$F(x, \varphi(x)) = 0. \tag{18.209}$$

La fonction φ ne permet donc que de trouver des points sur le cercle de rayon 1. △

18.5.4 Exemple

Le théorème de la fonction implicite a pour objet de donner l'existence de la fonction φ. Maintenant nous pouvons dire beaucoup de choses sur les dérivées de φ en considérant la fonction

$$x \mapsto F\big(x, \varphi(x)\big). \tag{18.210}$$

Par définition de φ, cette fonction est toujours nulle. En particulier, nous pouvons dériver l'équation

$$F\big(x, \varphi(x)\big) = 0, \tag{18.211}$$

et nous trouvons plein de choses.

Prenons par exemple la fonction

$$F\big((x,y), z\big) = ze^z - x - y, \tag{18.212}$$

et demandons nous ce que nous pouvons dire sur la fonction $z(x,y)$ telle que

$$F\big(x, y, z(x,y)\big) = 0, \tag{18.213}$$

c'est à dire telle que

$$z(x,y)e^{z(x,y)} - x - y = 0. \tag{18.214}$$

pour tout x et $y \in \mathbb{R}$. Nous pouvons facilement trouver $z(0,0)$ parce que

$$z(0,0)e^{z(0,0)} = 0, \tag{18.215}$$

donc $z(0,0) = 0$.

Nous pouvons dire des choses sur les dérivées de $z(x,y)$. Voyons par exemple $(\partial_x z)(x,y)$. Pour trouver cette dérivée, nous dérivons la relation (18.214) par rapport à x. Ce que nous trouvons est

$$(\partial_x z)e^z + ze^z(\partial_x z) - 1 = 0. \tag{18.216}$$

Cette équation peut être résolue par rapport à $\partial_x z$:

$$\frac{\partial z}{\partial x}(x,y) = \frac{1}{e^z(1+z)}. \tag{18.217}$$

Remarquez que cette équation ne donne pas tout à fait la dérivée de z en fonction de x et y, parce que z apparaît dans l'expression, alors que z est justement la fonction inconnue. En général, c'est la vie, nous ne pouvons pas faire mieux.

Dans certains cas, on peut aller plus loin. Par exemple, nous pouvons calculer cette dérivée au point $(x,y) = (0,0)$ parce que $z(0,0)$ est connu :

$$\frac{\partial z}{\partial x}(0,0) = 1. \tag{18.218}$$

Cela est pratique pour calculer, par exemple, le développement en Taylor de z autour de $(0,0)$.

Exemple 18.52
Est-ce que l'équation $e^y + xy = 0$ définit au moins localement une fonction $y(x)$? Nous considérons la fonction

$$f(x,y) = \begin{pmatrix} x \\ e^y + xy \end{pmatrix} \tag{18.219}$$

La différentielle de cette application est

$$df_{(0,0)}(u) = \frac{d}{dt}\Big[f(tu_1, tu_2)\Big]_{t=0} = \frac{d}{dt}\begin{pmatrix} tu_1 \\ e^{tu_2} + t^2u_1u_2 \end{pmatrix}_{t=0} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \tag{18.220}$$

L'application f définit donc un difféomorphisme local autour des points (x_0, y_0) et $f(x_0, y_0)$. Soit $(u, 0)$ un point dans le voisinage de $f(x_0, y_0)$. Alors il existe un unique (x, y) tel que

$$f(x,y) = \begin{pmatrix} x \\ e^y + xy \end{pmatrix} = \begin{pmatrix} u \\ 0 \end{pmatrix}. \tag{18.221}$$

Nous avons automatiquement $x = u$ et $e^y + xy = 0$. Notons toutefois que pour que ce procédé donne effectivement une fonction implicite $y(x)$ nous devons avoir des points de la forme $(u, 0)$ dans le voisinage de $f(x_0, y_0)$. △

18.6 Décomposition polaire (régularité)

18.53.
Nous allons montrer que l'application

$$\begin{aligned} f\colon S^{++}(n,\mathbb{R}) &\to S^{++}(n,\mathbb{R}) \\ A &\mapsto \sqrt{A} \end{aligned} \tag{18.222}$$

est une difféomorphisme.

Cependant $S^{++}(n,\mathbb{R})$ n'est pas un ouvert de $\mathbb{M}(n,\mathbb{R})$ et nous ne savons pas ce qu'est la différentielle d'une application non définie sur un ouvert. Nous allons donc en réalité montrer que l'application racine carré existe sur un voisinage de chacun des points de $S^{++}(n,\mathbb{R})$. Et comme une union quelconque d'ouverts est un ouvert, la fonction f sera bien définie sur un ouvert de $\mathbb{M}(n,\mathbb{R})$.

Lemme 18.54.
L'application

$$\begin{aligned} f\colon S^{++}(n,\mathbb{R}) &\to S^{++}(n,\mathbb{R}) \\ A &\mapsto A^2 \end{aligned} \tag{18.223}$$

est un C^∞-difféomorphisme.

Démonstration. Prouvons d'abord que f prend ses valeurs dans $S^{++}(n,\mathbb{R})$. Si $A \in S^{++}(n,\mathbb{R})$ alors par la diagonalisation 9.165 elle s'écrit $A = QDQ^{-1}$ où D est diagonale avec des nombres strictement positifs sur la diagonale. Avec cela, $A^2 = QD^2Q^{-1}$ où D^2 contient encore des nombres strictement positifs sur la diagonale.

L'application f étant essentiellement des polynôme en les entrées de A, elle est de classe C^∞.

Passons à l'étude de la différentielle. Comme mentionné en 18.53 nous allons en réalité voir f sur un ouvert de $\mathbb{M}(n,\mathbb{R})$ autour de $A \in S^{++}(n,\mathbb{R})$. Par conséquent si $A \in S^{++}(n,\mathbb{R})$,

$$df\colon S^{++}(n,\mathbb{R}) \to \mathcal{L}\big(\mathbb{M}(n,\mathbb{R}), \mathbb{M}(n,\mathbb{R})\big) \tag{18.224a}$$

$$df_A\colon \mathbb{M}(n,\mathbb{R}) \to \mathbb{M}(n,\mathbb{R}). \tag{18.224b}$$

Le calcul de df_A est facile. Soit $u \in \mathbb{M}(n,\mathbb{R})$ et faisons le calcul en utilisant la formule du lemme (12.178) :

$$df_A(u) = \frac{d}{dt}\Big[f(A+tu)\Big]_{t=0} \tag{18.225a}$$
$$= \frac{d}{dt}\Big[A^2 + tAu + tuA + t^2u^2\Big]_{t=0} \tag{18.225b}$$
$$= Au + uA. \tag{18.225c}$$

Nous allons utiliser le théorème d'inversion locale 18.47 à la fonction f. Dans la suite, A est une matrice de $S^{++}(n,\mathbb{R})$.

$\underline{df_A}$ **est injective** Soit $M \in \mathbb{M}(n,\mathbb{R})$ dans le noyau de df_A. En posant $M' = A^{-1}MQ$ nous avons $M = QM'Q^{-1}$ et on applique df_A à $QM'Q^{-1}$:

$$df_A(QM'Q^{-1}) = Q\big(DM + MD\big)Q^{-1}. \tag{18.226}$$

où $D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ avec $\lambda_i > 0$. La matrice D est inversible. Nous avons $M' = -DM'D^{-1}$, et en coordonnées,

$$M'_{ij} = -\sum_{kl} D_{ik}M'_{kl}D^{-1}_{lj} \tag{18.227a}$$
$$= -\sum_{kl} \lambda_i\delta_{ik}M'_{kl}\frac{1}{\lambda_j}\delta_{lj} \tag{18.227b}$$
$$= -\frac{\lambda_i}{\lambda_i}M'_{ij}. \tag{18.227c}$$

C'est à dire que $M'_{ij} = -\frac{\lambda_i}{\lambda_j}M'_{ij}$ avec $-\frac{\lambda_i}{\lambda_j} < 0$. Cela implique $M' = 0$ et par conséquent $M = 0$.

$\underline{df_A}$ **est surjective** Soit $N \in \mathbb{M}(n,\mathbb{R})$; nous cherchons $M \in \mathbb{M}(n,\mathbb{R})$ tel que $df_A(M) = N$. Nous posons $N' = Q^{-1}NQ$ et $M = QM'Q^{-1}$, ce qui nous donne à résoudre $df_D(M') = N'$. Passons en coordonnées :

$$(DM' + M'D)_{ij} = \sum_k (\delta_{ik}\lambda_i M'_{kj} + M'_{ik}\delta_{kj}\lambda_j) \tag{18.228a}$$
$$= M'_{ij}(\lambda_i + \lambda_j) \tag{18.228b}$$

où $\lambda_i + \lambda_j \neq 0$. Il suffit donc de prendre la matrice M' donnée par

$$M'_{ij} = \frac{1}{\lambda_i + \lambda_j}N'_{ij} \tag{18.229}$$

pour que $df_A(M') = N'$.

Le théorème d'inversion locale donne un voisinage V de A dans $\mathbb{M}(n,\mathbb{R})$ et un voisinage W de A^2 dans $\mathbb{M}(n,\mathbb{R})$ tels que $f\colon V \to W$ soit une bijection et $f^{-1}\colon W \to V$ soit de même régularité, en l'occurrence C^∞. □

Remarque 18.55.
Oui, il y a des matrices non symétriques qui ont une unique racine carré.

La proposition suivante, qui dépend du le théorème d'inversion locale par le lemme 18.54, donne plus de régularité à la décomposition polaire donnée dans le théorème 12.346.

Proposition 18.56 (Décomposition polaire : cas réel (suite))**.**
L'application

$$\begin{aligned} f\colon \mathrm{O}(n,\mathbb{R}) \times S^{++}(n,\mathbb{R}) &\to \mathrm{GL}(n,\mathbb{R}) \\ (Q,S) &\mapsto SQ \end{aligned} \tag{18.230}$$

est un difféomorphisme de classe C^∞.

Démonstration. Si M est donnée dans $\mathrm{GL}(n,\mathbb{R})$ alors la décomposition polaire[29] $M = QS$ est donnée par $S = \sqrt{MM^t}$ et $Q = MS^{-1}$. Autrement dit, si nous considérons la fonction de décomposition polaire

$$f\colon \mathrm{O}(n,\mathbb{R}) \times S^{++}(n,\mathbb{R}) \to \mathrm{GL}(n,\mathbb{R}) \tag{18.231}$$

alors

$$f^{-1}(M) = \big(M(\sqrt{MM^t})^{-1}, \sqrt{MM^t}\big). \tag{18.232}$$

Nous avons vu dans le lemme 18.54 que la racine carré était un C^∞-difféomorphisme. Le reste n'étant que des produits de matrice, la régularité est de mise. □

18.7 Théorème de Von Neumann

Lemme 18.57 ([2])**.**
Soit G, un sous groupe fermé de $\mathrm{GL}(n,\mathbb{R})$ et

$$\mathcal{L}_G = \{m \in \mathbb{M}(n,\mathbb{R}) \text{ tel que } e^{tm} \in G\ \forall t \in \mathbb{R}\}. \tag{18.233}$$

Alors $\mathcal{L}_G$ est un sous-espace vectoriel de $\mathbb{M}(n,\mathbb{R})$.

Démonstration. Si $m \in \mathcal{L}_G$, alors $\lambda m \in \mathcal{L}_G$ par construction. Le point délicat à prouver est le fait que si $a, b \in \mathcal{L}_G$, alors $a + b \in \mathcal{L}_G$. Soit $a \in \mathbb{M}(n,\mathbb{R})$; nous savons qu'il existe une fonction $\alpha_a\colon \mathbb{R} \to \mathbb{M}$ telle que

$$e^{ta} = \mathbb{1} + ta + \alpha_a(t) \tag{18.234}$$

et

$$\lim_{t\to 0} \frac{\alpha(t)}{t} = 0. \tag{18.235}$$

Si a et b sont dans $\mathcal{L}_G$, alors $e^{ta}e^{tb} \in G$, mais il n'est pas vrai en général que cela soit égal à $e^{t(a+b)}$. Pour tout $k \in \mathbb{N}$ nous avons

$$e^{a/k}e^{b/k} = \left(\mathbb{1} + \frac{a}{k} + \alpha_a(\frac{1}{k})\right)\left(\mathbb{1} + \frac{b}{k} + \alpha_b(\frac{1}{k})\right) = \mathbb{1} + \frac{a+b}{2} + \beta\left(\frac{1}{k}\right) \tag{18.236}$$

où $\beta\colon \mathbb{R} \to \mathbb{M}$ est encore une fonction vérifiant $\beta(t)/t \to 0$. Si k est assez grand, nous avons

$$\left\| \frac{a+b}{k} + \beta(\frac{1}{k}) \right\| < 1, \tag{18.237}$$

et nous pouvons profiter du lemme 15.92 pour écrire alors

$$\left(e^{a/k}e^{b/k}\right)^k = e^{k\ln\left(\mathbb{1} + \frac{a+b}{k} + \beta(\frac{1}{k})\right)}. \tag{18.238}$$

Ce qui se trouve dans l'exponentielle est

$$k\left[\frac{a+b}{k} + \alpha(\frac{1}{k}) + \sigma\left(\frac{a+b}{k} + \alpha(\frac{1}{k})\right)\right]. \tag{18.239}$$

Les diverses propriétés vues montrent que le tout tend vers $a + b$ lorsque $k \to \infty$. Par conséquent

$$\lim_{k\to\infty}\left(e^{a/k}e^{b/k}\right)^k = e^{a+b}. \tag{18.240}$$

Ce que nous avons prouvé est que pour tout t, $e^{t(a+b)}$ est une limite d'éléments dans G et est donc dans G parce que ce dernier est fermé. □

29. Proposition 12.346.

Vu que $\mathcal{L}_G$ est un sous-espace vectoriel de $\mathbb{M}(n,\mathbb{R})$, nous pouvons considérer un supplémentaire M.

Lemme 18.58.
Il n'existe pas se suites (m_k) dans $M\backslash\{0\}$ convergeant vers zéro et telle que $e^{m_k} \in G$ pour tout k.

Démonstration. Supposons que nous ayons $m_k \to 0$ dans $M\backslash\{0\}$ avec $e^{m_k} \in G$. Nous considérons les éléments $\epsilon_k = \frac{m_k}{\|m_k\|}$ qui sont sur la sphère unité de $\mathrm{GL}(n,\mathbb{R})$. Quitte à prendre une sous-suite, nous pouvons supposer que cette suite converge, et vu que M est fermé, ce sera vers $\epsilon \in M$ avec $\|\epsilon\| = 1$. Pour tout $t \in \mathbb{R}$ nous avons

$$e^{t\epsilon} = \lim_{k\to\infty} e^{t\epsilon_k}. \tag{18.241}$$

En vertu de la décomposition d'un réel en partie entière et décimale, pour tout k nous avons $\lambda_k \in \mathbb{Z}$ et $|\mu_k| \leqslant \frac{1}{2}$ tel que $t/\|m_k\| = \lambda_k + \mu_k$. Avec ça,

$$e^{t\epsilon} = \lim_{k\to\infty} \exp\Big(\frac{t}{m_k} m_k\Big) = \lim_{k\to\infty} e^{\lambda_k m_k} e^{\mu_k m_k}. \tag{18.242}$$

Pour tout k nous avons $e^{\lambda_k m_k} \in G$. De plus $|\mu_k|$ étant borné et m_k tendant vers zéro nous avons $e^{\mu_k m_k} \to 1$. Au final

$$e^{t\epsilon} = \lim_{k\to\infty} e^{t\epsilon_k} \in G \tag{18.243}$$

Cela signifie que $\epsilon \in \mathcal{L}_G$, ce qui est impossible parce que nous avions déjà dit que $\epsilon \in M\backslash\{0\}$. □

Lemme 18.59.
L'application

$$\begin{aligned} f\colon \mathcal{L}_G \times M &\to \mathrm{GL}(n,\mathbb{R}) \\ l, m &\mapsto e^l e^m \end{aligned} \tag{18.244}$$

est un difféomorphisme local entre un voisinage de $(0,0)$ dans $\mathbb{M}(n,\mathbb{R})$ et un voisinage de $\mathbb{1}$ dans $\exp\big(\mathbb{M}(n,\mathbb{R})\big)$.

Notons que nous ne disons rien de $e^{\mathbb{M}(n,\mathbb{R})}$. Nous n'allons pas nous embarquer à discuter si ce serait tout $\mathrm{GL}(n,\mathbb{R})$ [30] ou bien si ça contiendrait ne fut-ce que G.

Démonstration. Le fait que f prenne ses valeurs dans $\mathrm{GL}(n,\mathbb{R})$ est simplement dû au fait que les exponentielles sont toujours inversibles. Nous considérons ensuite la différentielle : si $u \in \mathcal{L}_G$ et $v \in M$ nous avons

$$df_{(0,0)}(u,v) = \frac{d}{dt}\Big[f\big(t(u,v)\big)\Big]_{t=0} = \frac{d}{dt}\Big[e^{tu}e^{tv}\Big]_{t=0} = u + v. \tag{18.245}$$

L'application df_0 est donc une bijection entre $\mathcal{L}_G \times M$ et $\mathbb{M}(n,\mathbb{R})$. Le théorème d'inversion locale 18.47 nous assure alors que f est une bijection entre un voisinage de $(0,0)$ dans $\mathcal{L}_G \times M$ et son image. Mais vu que df_0 est une bijection avec $\mathbb{M}(n,\mathbb{R})$, l'image en question contient un ouvert autour de $\mathbb{1}$ dans $\exp\big(\mathbb{M}(n,\mathbb{R})\big)$. □

Théorème 18.60 (Von Neumann[2, 229, 230])**.**
Tout sous-groupe fermé de $\mathrm{GL}(n,\mathbb{R})$ est une sous-variété de $\mathrm{GL}(n,\mathbb{R})$.

Démonstration. Soit G un tel groupe ; nous devons prouver que c'est localement difféomorphe à un ouvert de $\mathbb{R}^n$. Et si on est pervers, on ne va pas faire localement difféomorphe à un ouvert de $\mathbb{R}^n$, mais à un ouvert d'un espace vectoriel de dimension finie. Nous allons être pervers.

Étant donné que pour tout $g \in G$, l'application

$$\begin{aligned} L_g\colon G &\to G \\ h &\mapsto gh \end{aligned} \tag{18.246}$$

30. Vu les dimensions y'a tout de même peu de chance.

est de classe C^∞ et d'inverse C^∞, il suffit de prouver le résultat pour un voisinage de $\mathbb{1}$.

Supposons d'abord que $\mathcal{L}_G = \{0\}$. Alors 0 est un point isolé de $\ln(G)$; en effet si ce n'était pas le cas nous aurions un élément m_k de $\ln(G)$ dans chaque boule $B(0, r_k)$. Nous aurions alors $m_k = \ln(a_k)$ avec $a_k \in G$ et donc

$$e^{m_k} = a_k \in G. \tag{18.247}$$

De plus m_k appartient forcément à M parce que $\mathcal{L}_G$ est réduit à zéro. Cela nous donnerait une suite $m_k \to 0$ dans M dont l'exponentielle reste dans G. Or cela est interdit par le lemme 18.58. Donc 0 est un point isolé de $\ln(G)$. L'application ln étant continue [31], nous en déduisons que $\mathbb{1}$ est isolé dans G. Par le difféomorphisme L_g, tous les points de G sont isolés ; ce groupe est donc discret et par voie de conséquence une variété.

Nous supposons maintenant que $\mathcal{L}_G \neq \{0\}$. Nous savons par la proposition 15.91 que

$$\exp\colon \mathbb{M}(n, \mathbb{R}) \to \mathbb{M}(n, \mathbb{R}) \tag{18.248}$$

est une application C^∞ vérifiant $d\exp_0 = \mathrm{Id}$. Nous pouvons donc utiliser le théorème d'inversion locale 18.47 qui nous offre donc l'existence d'un voisinage U de 0 dans $\mathbb{M}(n, \mathbb{R})$ tel que $W = \exp(U)$ soit un ouvert de $\mathrm{GL}(n, \mathbb{R})$ et que $\exp\colon U \to W$ soit un difféomorphisme de classe C^∞.

Montrons que quitte à restreindre U (et donc W qui reste par définition l'image de U par exp), nous pouvons avoir $\exp\big(U \cap \mathcal{L}_G\big) = W \cap G$. D'abord $\exp(\mathcal{L}_G) \subset G$ par construction. Nous avons donc $\exp\big(U \cap \mathcal{L}_G\big) \subset W \cap G$. Pour trouver une restriction de U pour laquelle nous avons l'égalité, nous supposons que pour tout ouvert $\mathcal{O}$ dans U,

$$\exp\colon \mathcal{O} \cap \mathcal{L}_G \to \exp(\mathcal{O}) \cap G \tag{18.249}$$

ne soit pas surjective. Cela donnerait un élément de $\mathcal{O} \cap \complement\mathcal{L}_G$ dont l'image par exp n'est pas dans G. Nous construisons ainsi une suite en considérant une boule $B(0, \frac{1}{k})$ inclue à U et $x_k \in B(0, \frac{1}{k}) \cap \complement\mathcal{L}_G$ vérifiant $e^{x_k} \in G$. Vu le choix des boules nous avons évidemment $x_k \to 0$.

L'élément e^{x_k} est dans $e^{\mathbb{M}(n,\mathbb{R})}$ et le difféomorphisme du lemme 18.59 [32] nous donne $(l_k, m_k) \in \mathcal{L}_G \times M$ tel que $e^{l_k} e^{m_k} = e^{x_k}$. À ce point nous considérons k suffisamment grand pour que e^{x_k} soit dans la partie de l'image de f sur lequel nous avons le difféomorphisme. Plus prosaïquement, nous posons

$$(l_k, m_k) = f^{-1}(e^{x_k}) \tag{18.250}$$

et nous profitons de la continuité pour permuter la limite avec f^{-1} :

$$\lim_{k\to\infty} (l_k, m_k) = f^{-1}\Big(\lim_{k\to\infty} e^{x_k}\Big) = f^{-1}(\mathbb{1}) = (0, 0). \tag{18.251}$$

En particulier $m_k \to 0$ alors que $e^{m_k} = e^{x_k} e^{-l_k} \in G$. La suite m_k viole le lemme 18.58. Nous pouvons donc restreindre U de telle façon à avoir

$$\exp\big(U \cap \mathcal{L}_G\big) = W \cap G. \tag{18.252}$$

Nous avons donc un ouvert de $\mathcal{L}_G$ (l'ouvert $U \cap \mathcal{L}_G$) qui est difféomorphe avec l'ouvert $W \cap G$ de G. Donc G est une variété et accepte $\mathcal{L}_G$ comme carte locale.

□

Remarque 18.61.
En termes savants, nous avons surtout montré que si G est un groupe de Lie d'algèbre de Lie $\mathfrak{g}$, alors l'exponentielle donne un difféomorphisme local entre $\mathfrak{g}$ et G.

31. Par le lemme 15.92.
32. Il me semble que l'utilisation de ce lemme manque à l'avant-dernière ligne de la preuve chez [2].

18.8 Recherche d'extrema

18.8.1 Extrema à une variable

Définition 18.62.
*Soit $f\colon A \subset \mathbb{R} \to \mathbb{R}$ et $a \in A$. Le point a est un **maximum local** de f s'il existe un voisinage $\mathcal{U}$ de a tel que $f(a) \geqslant f(x)$ pour tout $x \in \mathcal{U} \cap A$. Le point a est un **maximum global** si $f(a) \geqslant g(x)$ pour tout $x \in A$.*

La proposition basique à utiliser lors de la recherche d'extrema est la suivante :

Proposition 18.63.
Soit $f\colon A \subset \mathbb{R} \to \mathbb{R}$ et $a \in \text{Int}(A)$. Supposons que f est dérivable en a. Si a est un extremum local, alors $f'(a) = 0$.

La réciproque n'est pas vraie, comme le montre l'exemple de la fonction $x \mapsto x^3$ en $x = 0$: sa dérivée est nulle et pourtant $x = 0$ n'est ni un maximum ni un minimum local.

Cette proposition ne sert donc qu'à sélectionner des *candidats* extremum. Afin de savoir si ces candidats sont des extrema, il y a la proposition suivante.

Proposition 18.64.
Soit $f\colon I \subset \mathbb{R} \to \mathbb{R}$, une fonction de classe C^k au voisinage d'un point $a \in \text{Int}\, I$. Supposons que

$$f'(a) = f''(a) = \ldots = f^{(k-1)}(a) = 0, \tag{18.253}$$

et que

$$f^{(k)}(a) \neq 0. \tag{18.254}$$

Dans ce cas,

(1) Si k est pair, alors a est un point d'extremum local de f, c'est un minimum si $f^{(k)}(a) > 0$, et un maximum si $f^{(k)}(a) < 0$,

(2) Si k est impair, alors a n'est pas un extremum local de f.

Note : jusqu'à présent nous n'avons rien dit des extrema *globaux* de f. Il n'y a pas grand chose à en dire. Si un point d'extremum global est situé dans l'intérieur du domaine de f, alors il sera extremum local (a fortiori). Ou alors, le maximum global peut être sur le bord du domaine. C'est ce qui arrive à des fonctions strictement croissantes sur un domaine compact.

Une seule certitude : si une fonction est continue sur un compact, elle possède une minimum et un maximum global par le théorème 6.126.

Soit une fonction $f\colon I \to \mathbb{R}$, et soit $a \in I$. Si $f'(a) > 0$, alors la tangente au graphe de f au point $\big(a, f(a)\big)$ sera une droite croissante (coefficient directeur positif). Cela ne veut pas spécialement dire que la fonction elle-même sera croissante, mais en tout cas cela est un bon indice.

Exemple 18.65
Si $f(x) = x^2$, il est connu que $f'(x) = 2x$. Nous avons donc que f' est positive si $x \geqslant 0$ et $f' >$ est négative si $x < 0$. Cela correspond bien au fait que x^2 est décroissante sur $]-\infty, 0[$ et croissante sur $]0, \infty[$. △

Sur la figure 18.2, nous avons dessiné la fonction $f(x) = x\cos(x)$ et sa dérivée. Nous voyons que partout où la dérivée est négative, la fonction est décroissante tandis que, inversement, partout où la dérivée est positive, la fonction est croissante.

Les extrema de la fonction f sont donc placés là où f' change de signe. En effet si $f'(x) < 0$ pour $x < a$ et $f'(x) > 0$ pour $x > a$, la fonction est décroissante jusqu'à a et est ensuite croissante. Cela signifie que la fonction connait un creux en a. Le point a est donc un minimum de la fonction.

Attention cependant. Le fait que $f'(a) = 0$ ne signifie pas automatiquement que f a un maximum ou un minimum en a. Nous avons par exemple tracé sur la figure 18.3 les fonctions x^3 et sa dérivée. Il est à noter que, conformément à ce que l'on pense, certes la dérivée s'annule en $x = 0$, mais elle ne change pas de signe.

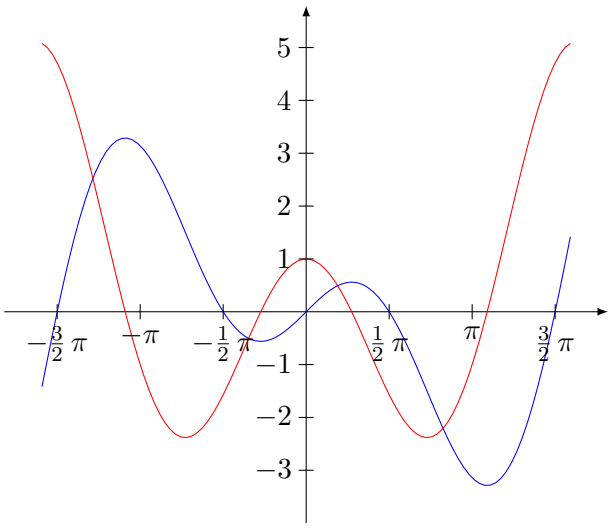


FIGURE 18.2 – La fonction $f(x) = x\cos(x)$ en bleu et sa dérivée en rouge.

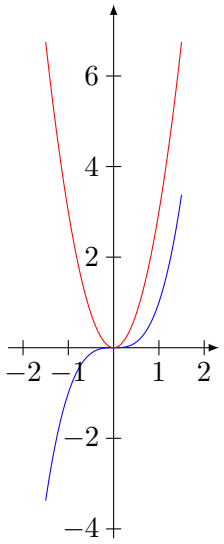


FIGURE 18.3 – La dérivée de x^3 s'annule en $x = 0$, mais ce n'est ni un minimum ni un maximum.

18.8.2 Extrema libre

Définition 18.66.
*Un point a à l'intérieur du domaine d'une fonction $f\colon A \subset \mathbb{R}^n \to \mathbb{R}$ est un **point critique** de f lorsque $df(a) = 0$.*

Ces points sont analogues aux points où la dérivée d'une fonction sur $\mathbb{R}$ s'annule. Les points critiques de f sont dons les candidats à être des points d'extremum.

Dans le cas d'une fonction de deux variables,l la proposition 12.259 nous permet de voir $(d^2f)_a$ comme étant la matrice

$$d^2f(a) = \begin{pmatrix} \frac{d^2f}{dx^2}(a) & \frac{d^2f}{dx\,dy}(a) \\ \frac{d^2f}{dy\,dx}(a) & \frac{d^2f}{dy^2}(a) \end{pmatrix}. \tag{18.255}$$

Dans le cas d'une fonction C^2, cette matrice est symétrique.

Proposition 18.67 ([231])**.**
Soit un ouvert Ω de $\mathbb{R}^n$ et $a \in \Omega$. Soit une fonction $f\colon \Omega \to \mathbb{R}$ différentiable en a. Si a est un extremum local de f, alors a est un point critique de f.

Démonstration. Nous supposons que a est un maximum local (ce sera la même chose si a est un

minimum). Soit $r > 0$ tel que $f(x) \leqslant f(a)$ pour tout $x \in B(a,r)$ (et tel que cette boule reste dans Ω). Soit $u \in \mathbb{R}^n$ assez petit pour que $a \pm u \in B(a,r)$ de sorte que la définition suivante ait un sens :

$$\begin{aligned} g\colon [-1,1] &\to \mathbb{R} \\ t &\mapsto f(a+tu) \end{aligned} \tag{18.256}$$

Cette fonction est différentiable en $t = 0$ (composée de fonctions différentiables, proposition 12.189) et a un maximum local en $t = 0$. Donc $g'(0) = 0$ par la proposition 18.63. Donc

$$0 = \frac{d}{dt}\Big[f(a+tu)\Big]_{t=0} = df_a(u). \tag{18.257}$$

□

18.8.3 Extrema et Hessienne

Proposition 18.68 ([1, 232, 233])**.**
Soit un ouvert Ω de $\mathbb{R}^n$ et une fonction $f\colon \Omega \to \mathbb{R}$ deux fois différentiable ainsi que $a \in \Omega$.

(1) Si a est un point critique de f et si il existe r tel que $(d^2 f_x)$ est semi-définie positive pour tout $x \in B(a,r)$ alors f possède un minimum local en a.

(2) Si a est un point critique[33] de f, et si $d^2 f_a$ est strictement définie positive[34], alors a est un minimum local strict de f,

(3) Si a est un minimum local, alors $(d^2 f)_a$ est semi définie positive.

Démonstration. Nous subdivisons la preuve.

(1) Soit h tel que $a + h \in B(a,r)$. Nous allons montrer que $f(a) \leqslant f(a+h)$; cela montrera que $x = a$ est un minimum local. Pour cela nous utilisons un développement de Taylor : il existe $c \in]a, a+h[$ tel que

$$f(a+h) = f(a) + df_a(h) + \frac{1}{2}(d^2 f)_c(h,h) \geqslant f(a) \tag{18.258}$$

parce que par hypothèse $(d^2 f)_c$ est définie positive et parce que $df_a = 0$.

(2) La forme bilinéaire $d^2 f_a$ est strictement définie positive, donc il existe $\alpha > 0$ tel que

$$d^2 f_a(h,h) > \alpha \|h\|^2 \tag{18.259}$$

pour tout h. Nous écrivons encore Taylor : il existe une fonction ϵ telle que $\lim_{h\to 0} \epsilon(h) = 0$ et

$$f(a+h) = g(a) + df_a(h) + \frac{1}{2}(d^2 f)_a(h,h) + \|h\|^2 \epsilon(h). \tag{18.260}$$

En tenant compte du fait que $df_a = 0$,

$$f(a+h) > f(a) + \|h\|^2 \big(\frac{1}{2}\alpha + \epsilon(h)\big). \tag{18.261}$$

La limite de ϵ nous dit qu'il existe $r > 0$ tel que $\|\epsilon(h)\| < \frac{1}{2}\alpha$ pour tout $h \in B(0,r)$. Pour ces valeurs de h nous avons

$$f(a+h) > f(a). \tag{18.262}$$

Donc a est un minimum local strict de f.

33. Définition 18.66.

34. La fonction f est de classe C^2, donc les dérivées croisées sont égales et $d^2 f$ est symétrique. La définition 9.167 s'applique donc.

(3) Si a est un minimum local, nous savons déjà que $df_a = 0$ par la proposition 18.67. Nous écrivons le développement de Taylor de f à l'ordre 2 de la proposition 12.267 :

$$f(a+h) = f(a) + df_a(h) + \frac{1}{2}(d^2f)_a(h,h) + \|h\|^2\alpha(\|h\|). \tag{18.263}$$

En prenant h assez petit pour que $a+h$ ne sorte pas de la boule dans laquelle a est un minimum, nous avons $f(a+h) - f(a) > 0$. Donc

$$\frac{1}{2}(d^2f)_a(h,h) + \|h\|^2\alpha(\|h\|) > 0 \tag{18.264}$$

Nous divisons cela par $\|h\|^2$ et notons $e_h = h/\|h\|$:

$$\frac{1}{2}(d^2f)_a(e_h,e_h) + \alpha(\|h\|) > 0. \tag{18.265}$$

À la limite $h \to 0$, le premier terme est constant tandis que le deuxième tend vers zéro. À la limite,

$$(d^2f)_a(e_h,e_h) \geqslant 0. \tag{18.266}$$

La caractérisation du lemme 9.170(2) nous dit alors que $(d^2f)_a$ est semi-définie positive.

□

La partie (3) est tout à fait comparable au fait bien connu que, pour une fonction $f\colon \mathbb{R} \to \mathbb{R}$, si le point a est minimum local, alors $f'(a) = 0$ et $f''(a) > 0$.

Notons que le point (3) ne parle pas de minimum strict, et donc pas de matrice *strictement* définie positive.

Exemple 18.69(Proposition 18.68(2) sans point critique)
L'hypothèse de point critique pour l'utilisation de la stricte définition positive de d^2f_a est nécessaire. Soit en effet la fonction

$$f(x) = x^2 + x. \tag{18.267}$$

Elle vérifie $f''(0) = 2$, de telle sorte que sa différentielle seconde en zéro soit strictement définie positive. Le point $x = 0$ n'est cependant même pas un minimum local. Entre autres parce que $f'(0) = 1 \neq 0$. △

La méthode pour chercher les extrema de f est donc de suivre le points suivants :

(1) Trouver les candidats extrema en résolvant $\nabla f = (0,0)$,
(2) écrire $d^2f(a)$ pour chacun des candidats
(3) calculer les valeurs propres de $d^2f(a)$, déterminer si la matrice est définie positive ou négative,
(4) conclure.

Une conséquence de la proposition 9.244(3) [35] est que si $\det M < 0$, alors le point a n'est pas un extrema dans le cas où $M = d^2f(a)$ par le point (3) de la proposition 18.68.

Exemple 18.70
Soit la fonction $f(x,y) = x^4 + y^4 - 4xy$. C'est une fonction différentiable sans problèmes. D'abord sa différentielle est

$$df = \big(4x^3 - 4y; 4y^3 - 4x\big), \tag{18.268}$$

et la matrice des dérivées secondes est

$$M = d^2f(x,y) = \begin{pmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{pmatrix}. \tag{18.269}$$

35. La matrice $d^2f(a)$ est toujours symétrique quand f est de classe C^2.

Nous avons $fd = 0$ pour les trois points $(0,0)$, $(1,1)$ et $-1,-1$.

Pour le point $(0,0)$ nous avons

$$M = \begin{pmatrix} 0 & -4 \\ -4 & 0 \end{pmatrix}, \tag{18.270}$$

dont les valeurs propres sont 4 et -4. Elle n'est donc semi-définie ou définie rien du tout. Donc $(0,0)$ n'est pas un extremum local.

Au contraire pour les points $(1,1)$ et $(-1,-1)$ nous avons

$$M = \begin{pmatrix} 12 & -4 \\ -4 & 12 \end{pmatrix}, \tag{18.271}$$

dont les valeurs propres sont 16 et 8. La matrice d^2f y est donc définie positive. Ces deux points sont donc extrema locaux. △

18.8.4 Un peu de recettes de cuisine

(1) Rechercher les points critiques, càd les (x,y) tels que

$$\begin{cases} \frac{\partial f}{\partial x}(x,y) = 0 \\ \frac{\partial f}{\partial y}(x,y) = 0 \end{cases}$$

En effet, si (x_0,y_0) est un extrémum local de f, alors $\frac{\partial f}{\partial x}(x_0,y_0) = 0 = \frac{\partial f}{\partial y}(x_0,y_0)$.

(2) Déterminer la nature des points critiques : «test» des dérivées secondes :

$$\text{On pose } H(x_0,y_0) = \frac{\partial^2 f}{\partial x^2}(x_0,y_0)\frac{\partial f^2}{\partial y^2}(x_0,y_0) - \left(\frac{\partial^2 f}{\partial x \partial y}(x_0,y_0)\right)^2$$

(a) Si $H(x_0,y_0) > 0$ et $\frac{\partial^2 f}{\partial x^2}(x_0,y_0) > 0 \Longrightarrow (x_0,y_0)$ est un minimum local de f.

(b) Si $H(x_0,y_0) > 0$ et $\frac{\partial^2 f}{\partial x^2}(x_0,y_0) < 0 \Longrightarrow (x_0,y_0)$ est un maximum local de f.

(c) Si $H(x_0,y_0) < 0 \Longrightarrow f$ a un point de selle en (x_0,y_0).

(d) Si $H(x_0,y_0) = 0 \Longrightarrow$ on ne peut rien conclure.

18.8.5 Extrema liés

Soit f, une fonction sur $\mathbb{R}^n$, et $M \subset \mathbb{R}^n$ une variété de dimension m. Nous voulons savoir quelle sont les plus grandes et plus petites valeurs atteintes par f sur M.

Pour ce faire, nous avons un théorème qui permet de trouver des extrema *locaux* de f sur la variété. Pour rappel, $a \in M$ est une **extrema local de f relativement** à l'ensemble M s'il existe une boule $B(a,\epsilon)$ telle que $f(a) \leqslant f(x)$ pour tout $x \in B(a,\epsilon) \cap M$.

Théorème 18.71 (Extrema lié [209])**.**
Soit A, un ouvert de $\mathbb{R}^n$ et

(1) une fonction (celle à minimiser) $f \in C^1(A,\mathbb{R})$,

(2) des fonctions (les contraintes) $G_1,\dots,G_r \in C^1(A,\mathbb{R})$,

(3) $M = \{x \in A \text{ tel que } G_i(x) = 0\, \forall i\}$,

(4) un extrema local $a \in M$ de f relativement à M.

Supposons que les gradients $\nabla G_1(a)$, …,$\nabla G_r(a)$ soient linéairement indépendants. Alors $a = (x_1,\dots,x_n)$ est une solution de $\nabla L(a) = 0$ où

$$L(x_1,\dots,x_n,\lambda_1,\dots,\lambda_r) = f(x_1,\dots,x_n) + \sum_{i=1}^{r} \lambda_i G_i(x_1,\dots,x_n). \tag{18.272}$$

Autrement dit, si a est un extrema lié, alors $\nabla f(a)$ est une combinaisons des $\nabla G_i(a)$, ou encore il existe des λ_i tels que

$$df(a) = \sum_i \lambda_i dG_i(a). \tag{18.273}$$

La fonction L est le **lagrangien** du problème et les variables λ_i sont les **multiplicateurs de Lagrange**.

Démonstration. Si $r = n$ alors les vecteurs linéairement indépendantes $\nabla G_i(a)$ forment une base de $\mathbb{R}^n$ et donc évidemment les λ_i existent. Nous supposons donc maintenant que $r < n$. Nous notons $(z_i)_{i=1\dots n}$ les coordonnées sur $\mathbb{R}^n$.

La matrice

$$\begin{pmatrix} \frac{\partial G_1}{\partial z_1}(a) & \cdots & \frac{\partial G_1}{\partial z_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial G_r}{\partial z_1}(a) & \cdots & \frac{\partial G_r}{\partial z_n}(a) \end{pmatrix} \tag{18.274}$$

est de rang r parce que les lignes sont par hypothèses linéairement indépendantes. Nous nommons $(y_i)_{i=1,\dots,r}$ un choix de r parmi les (z_i) tels que

$$\det \begin{pmatrix} \frac{\partial G_1}{\partial y_1} & \cdots & \frac{\partial G_1}{\partial y_r} \\ \vdots & \ddots & \vdots \\ \frac{\partial G_r}{\partial y_1} & \cdots & \frac{\partial G_r}{\partial y_r} \end{pmatrix} \neq 0. \tag{18.275}$$

Nous identifions $\mathbb{R}^n$ à $\mathbb{R}^s \times \mathbb{R}^r$ dans lequel $\mathbb{R}^r$ est la partie générée par les $(y_i)_{i=1,\dots,r}$. Nous nommons $(x_j)_{j=1,\dots,s}$ les coordonnées sur $\mathbb{R}^s$. Autrement dit, les coordonnées sur $\mathbb{R}^n$ sont $x_1, \dots, x_s, y_1, \dots, y_r$. Dans ces coordonnées, nous nommons $a = (\alpha, \beta)$ avec $\alpha \in \mathbb{R}^s$ et $\beta \in \mathbb{R}^r$.

Si nous notons $G = (G_1, \dots, G_r)$, le théorème de la fonction implicite (théorème 18.48) nous dit qu'il existe un voisinage U' de $\alpha \in \mathbb{R}^n$, un voisinage V' de $\beta \in \mathbb{R}^r$ et une fonction $\varphi\colon U' \to V'$ de classe C^1 telle que si $(x, y) \in U' \times V'$, alors

$$g(x, y) = 0 \tag{18.276}$$

si et seulement si $y = \varphi(x)$. Nous posons maintenant

$$\psi(x) = (x, \varphi(x)) \tag{18.277a}$$
$$h(x) = f\big(\psi(x)\big). \tag{18.277b}$$

Nous avons $\psi(\alpha) = a$ et $\psi(x) \in M$ pour tout $x \in U'$. La fonction h a donc un extrema local en α et donc les dérivées partielles de h y sont nulles. Cela signifie que

$$0 = \frac{\partial h}{\partial x_i}(\alpha) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\frac{\partial x_j}{\partial x_i} + \sum_{k=1}^{r} \frac{\partial f}{\partial y_k}\frac{\partial \varphi_k}{\partial x_i}, \tag{18.278}$$

c'est à dire

$$\frac{\partial f}{\partial x_i}(\alpha) + \sum_{k=1}^{r} \frac{\partial f}{\partial y_k}(a)\frac{\partial \varphi_k}{\partial x_i}(\alpha) = 0 \tag{18.279}$$

pour tout $i = 1, \dots, s$. D'autre part pour tout k, la fonction $l_k(x) = G_k\big(x, \varphi(x)\big)$ est constante et vaut zéro ; ses dérivées partielles sont donc nulles :

$$\frac{\partial l}{\partial x_i}(\alpha) = \frac{\partial G_k}{\partial x_i}(\alpha) + \sum_{k=1}^{r} \frac{\partial G_k}{\partial y_k}(a)\frac{\partial \varphi_k}{\partial x_i}(\alpha) = 0 \tag{18.280}$$

pour tout $i = 1, \dots, s$ et $k = 1, \dots, r$.

Les s premières colonnes de la matrice

$$\begin{pmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_s} & \frac{\partial f}{\partial y_1} & \cdots & \frac{\partial f}{\partial y_r} \\ \frac{\partial G_1}{\partial x_1} & \cdots & \frac{\partial G_1}{\partial x_s} & \frac{\partial G_1}{\partial y_1} & \cdots & \frac{\partial G_1}{\partial y_r} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial G_r}{\partial x_1} & \cdots & \frac{\partial G_r}{\partial x_s} & \frac{\partial G_r}{\partial y_1} & \cdots & \frac{\partial G_r}{\partial y_r} \end{pmatrix} \tag{18.281}$$

s'expriment en terme des r dernières. La matrice est donc au maximum de rang r. Notons que la première ligne est ∇f et les r suivantes sont les ∇G_i. Vu que ces lignes sont des vecteurs liés, il existe $\mu_0, \dots, \mu_r$ tels que

$$\mu_0 \nabla f + \sum_{i=1}^{r} \mu_i \nabla G_i = 0. \tag{18.282}$$

Par hypothèse les ∇G_i sont linéairement indépendants, ce qui nous dit que $\mu_0 \neq 0$. Donc nous avons ce qu'il nous faut :

$$\nabla f(a) = \sum_i \frac{\mu_i}{\mu_0} \nabla G_i(a). \tag{18.283}$$

Notons qu'au vu de l'expression (18.273), le fait que les formes $\{dG_i(a)\}_{1 \leqslant i \leqslant r}$ forment une partie libre dans $(\mathbb{R}^n)^*$ implique que les λ_i sont uniques. □

La proposition suivante est la même que 18.71.

Proposition 18.72.
Soit U, un ouvert de $\mathbb{R}^n$ et des fonctions de classe C^1 $f, g_1, \dots, g_r \colon U \to \mathbb{R}$. Nous considérons

$$\Gamma = \{x \in U \text{ tel que } g_1(x) = \dots = g_r(x) = 0\}. \tag{18.284}$$

Soit a un extrémum de $f|_\Gamma$. Supposons que les formes $dg_1, \dots, dg_r$ soient linéairement indépendantes en a. Alors il existe $\lambda_1, \dots, \lambda_r$ dans $\mathbb{R}$ tel que

$$df_a = \sum_{i=1}^{r} \lambda_i (dg_i)_a. \tag{18.285}$$

En pratique les candidats extrema locaux sont tous les points où les gradients ne sont pas linéairement indépendants, plus tous les points donnés par l'équation $\nabla L = 0$. Parmi ces candidats, il faut trouver lesquels sont maxima ou minima, locaux ou globaux.

L'existence d'extrema locaux se prouve généralement en invoquant de la compacité, et en invoquant le lemme suivant qui permet de réduire le problème à un compact.

Lemme 18.73.
Soit S, une partie de $\mathbb{R}^n$ et C, un ouvert de $\mathbb{R}^n$. Si $a \in \operatorname{Int} S$ est un minimum local relatif à $S \cap C$, alors il est un minimum local par rapport à S.

Démonstration. Nous avons que $\forall x \in B(a, \epsilon_1) \cap S \cap C$, $f(x) \geqslant f(x)$. Mais étant donné que C est ouvert, et que $a \in C$, il existe un ϵ_2 tel que $B(a, \epsilon_2) \subset C$. En prenant $\epsilon = \min\{\epsilon_1, \epsilon_2\}$, nous trouvons que $f(x) \geqslant f(a)$ pour tout $x \in B(a, \epsilon) \cap (S \cap C) = B(a, \epsilon) \cap S$. □

18.9 Fonctions convexes

Définition 18.74 ([234]).
Une fonction f d'un intervalle I de $\mathbb{R}$ vers $\mathbb{R}$ est dite ***convexe*** *lorsque, pour tous x_1 et x_2 de I et tout λ dans $[0, 1]$ nous avons*

$$f\big(\lambda x_1 + (1 - \lambda) x_2\big) \leqslant \lambda f(x_1) + (1 - \lambda) f(x_2) \tag{18.286}$$

Si l'inégalité est stricte, alors nous disons que la fonction f est ***strictement convexe.***

Une fonction est ***concave*** *si son opposée est convexe.*

18.75 ([235]).
Les différents résultats pour les fonctions convexes s'adaptent généralement sans mal aux fonctions strictement convexes. Une nuance cependant : de même que les fonctions dérivables convexes sont celles qui ont une dérivée croissante, les fonctions dérivables strictement convexes sont celles qui ont une dérivée strictement croissante (proposition 18.79). En revanche, il ne faudrait pas croire que la dérivée seconde d'une fonction dérivable strictement convexe est nécessairement une fonction à

valeurs strictement positives (voir théorème 18.80) : la dérivée d'une fonction strictement croissante peut s'annuler occasionnellement, ou plus exactement peut s'annuler sur un ensemble de points d'intérieur vide. Penser à $x \mapsto x^4$ pour un exemple de fonction strictement convexe dont la dérivée seconde s'annule.

18.9.1 Inégalité des pentes

Dans l'étude des fonctions convexes nous allons souvent utiliser la fonction **taux d'accroissement** qui est, pour α dans le domaine de convexité de f définie par

$$\begin{aligned}\tau_\alpha\colon I\backslash\{\alpha\} &\to \mathbb{R}\\ x &\mapsto \frac{f(x)-f(\alpha)}{x-\alpha}.\end{aligned} \tag{18.287}$$

Proposition 18.76 (Inégalité des pentes[236])**.**
Soit f une fonction convexe sur un intervalle $I \subset \mathbb{R}$. Alors pour tout $a < b < c$ dans I nous avons[36]

$$\frac{f(b)-f(a)}{b-a} \leqslant \frac{f(c)-f(a)}{c-a} \leqslant \frac{f(c)-f(b)}{c-b}. \tag{18.288}$$

En d'autres termes,

$$\tau_a(b) \leqslant \tau_a(c) \leqslant \tau_b(c), \tag{18.289}$$

c'est à dire que τ est croissante en ses deux arguments.

Démonstration. D'abord les inégalités $a < b < c$ impliquent $0 < b-a < c-a$ et donc

$$\lambda = \frac{b-a}{c-a} < 1. \tag{18.290}$$

L'astuce est de remarquer que $(1-\lambda)a + \lambda c = b$. Donc λ a toutes les bonnes propriétés pour être utilisé dans la définition de la convexité :

$$f\big((1-\lambda)a + \lambda c\big) \leqslant \lambda f(c) + (1-\lambda)f(a), \tag{18.291}$$

c'est à dire

$$f(b) - f(a) \leqslant \lambda\big(f(c) - f(a)\big) \tag{18.292}$$

ou encore, en remplaçant λ par sa valeur :

$$\frac{f(b)-f(a)}{b-a} \leqslant \frac{f(c)-f(a)}{c-a}. \tag{18.293}$$

Cela fait déjà une des inégalités à savoir.

D'autre part en partant de $-a < -b < -c$ nous posons

$$0 < \lambda = \frac{c-b}{c-a}. \tag{18.294}$$

Nous avons à nouveau $b = (1-\lambda)c + \lambda a$ et nous pouvons obtenir la seconde inégalité

$$\frac{f(c)-f(a)}{c-a} \leqslant \frac{f(c)-f(b)}{c-b}. \tag{18.295}$$

□

36. Les inégalités sont strictes si la fonction f est strictement convexe.

Géométriquement, l'inégalité des pentes se comprend facilement : le coefficient angulaire de la corde du graphe augmente. Donc si $x < y < z$, le coefficient moyen entre x et y est plus petit que celui entre x et z qui est plus petit que celui entre y et z.

Donc si le coefficient angulaire moyen entre a et $b+u$ vaut celui entre a et b, ce coefficient ne peut qu'être constant entra a et b : sinon il serait plus grand entre b et $b+u$ et la moyenne sur $a \to b+u$ serait plus grande que sa moyenne sur $a \to b$. Mais avoir un coefficient angulaire constant signifie être une droite.

En résumé, si une fonction est convexe et non strictement convexe, alors son graphe est une droite. C'est en gros cela que la proposition 18.83 clarifiera.

18.9.2 Convexité et régularité

Lemme 18.77 ([235])**.**
Une fonction convexe sur un ouvert

(1) y admet des dérivées à gauche et à droite en chaque point,

(2) y est continue.

Démonstration. Soit $I =]a,b[$ un intervalle sur lequel f est convexe et $\alpha \in I$. Nous allons prouver que f est continue en α. Nous considérons τ_α le taux d'accroissement définit par (18.287) ; c'est une fonction croissante comme précisé dans l'inégalité des trois pentes 18.76 et de plus $\tau_\alpha(x)$ est bornée supérieurement par $\tau_\alpha(b)$ pour $x < \alpha$ et inférieurement par $\tau_\alpha(a)$ pour $x > \alpha$. Les limites existent donc et sont finies par la proposition 12.88. Autrement dit les limites

$$\lim_{x\to\alpha+} \frac{f(x)-f(\alpha)}{x-\alpha} = \lim_{x\to\alpha^+} \tau_\alpha(x) = \inf_{t>\alpha} \tau_\alpha(t) \tag{18.296a}$$

$$\lim_{x\to\alpha^-} \frac{f(x)-f(\alpha)}{x-\alpha} = \lim_{x\to\alpha^-} \tau_\alpha(x) = \sup_{t<\alpha} \tau_\alpha(t). \tag{18.296b}$$

existent et sont finies, c'est à dire que la fonction f admet une dérivée à gauche et à droite.

Pour tout x nous avons les inégalités

$$\tau_\alpha(a) \leqslant \frac{f(x)-f(\alpha)}{x-\alpha} \leqslant \tau_\alpha(b). \tag{18.297}$$

En posant $k = \max\{\tau_\alpha(a), \tau_\alpha(b)\}$ nous avons

$$\big|f(x)-f(\alpha)\big| \leqslant k|x-\alpha|. \tag{18.298}$$

La fonction est donc Lipschitzienne et par conséquent continue par la proposition 12.245. □

Remarque 18.78.
Les dérivées à gauche et à droite ne sont a priori pas égales. Penser par exemple à une fonction affine par morceaux dont les pentes augmentent à chaque morceau.

18.9.3 Dérivées d'une fonction convexe

Proposition 18.79 ([237, 238, 1])**.**
Une fonction dérivable sur un intervalle I de $\mathbb{R}$

(1) est convexe si et seulement si sa dérivée est croissante sur I.

(2) est strictement convexe si et seulement si sa dérivée est strictement croissante sur I

Démonstration. Pour la preuve de (1) et (2), nous allons démontrer les énoncés «non stricts» et indiquer ce qu'il faut changer pour obtenir les énoncés «stricts».

Sens direct Nous supposons que f est convexe. Soient $a < b$ dans I et $x \in]a,b[$. D'après l'inégalité des pentes 18.76,

$$\frac{f(x)-f(a)}{x-a} \leqslant \frac{f(b)-f(a)}{b-a} \leqslant \frac{f(b)-f(x)}{b-x}. \tag{18.299}$$

En faisant la limite $x \to a$ nous avons

$$f'(a) \leqslant \frac{f(b)-f(a)}{b-a} \tag{18.300}$$

et la limite $x \to b$ donne

$$\frac{f(b)-f(a)}{b-a} \leqslant f'(b). \tag{18.301}$$

Ici les inégalités sont non a priori strictes, même si f est strictement convexe : même avec des inégalités strictes dans (18.299), le passage à la limite rend l'inégalité non stricte. Quoi qu'il en soit nous avons

$$f'(a) \leqslant f'(b). \tag{18.302}$$

Sens direct : strict Nous savons déjà que f' est croissante. Si (18.302) était une égalité, alors f' serait constante sur $]a,b[$ parce qu'en prenant c entre a et b nous aurions $f'(a) \leqslant f'(c) \leqslant f'(b)$ avec $f'(a) = f'(b)$. Donc $f'(a) = f'(c)$. Avoir f' constante sur un intervalle est contraire à la stricte convexité.

Sens réciproque Nous supposons que f' est croissante et nous considérons $a < b$ dans I ainsi que $\lambda \in [0,1]$. Nous posons $x = \lambda a + (1-\lambda)b$, et nous savons que $a \leqslant x \leqslant b$. Le théorème des accroissements finis 12.129 donne $c_1 \in]a,x[$ et $c_2 \in]x,b[$ tels que

$$f'(c_1) = \frac{f(x)-f(a)}{x-a} \tag{18.303}$$

et

$$f'(c_2) = \frac{f(b)-f(x)}{b-x}. \tag{18.304}$$

Et en plus $c_1 < c_2$. Vu que f' est croissante nous avons $f'(c_1) \leqslant f'(c_2)$ et donc

$$\frac{f(x)-f(a)}{x-a} \leqslant \frac{f(b)-f(x)}{b-x}. \tag{18.305}$$

En remplaçant x par sa valeur en termes de λ, a et b nous avons $x-a = (1-\lambda)(b-a)$ et $b-x = \lambda(b-a)$, et l'inégalité (18.305) nous donne

$$f(x) \leqslant \lambda f(a) + (1-\lambda)f(b). \tag{18.306}$$

Sens réciproque : strict Si f' est strictement croissante, nous avons $f'(c_2) < f'(c_2)$ et les inégalité suivantes sont strictes, ce qui donne

$$f(x) < \lambda f(a) + (1-\lambda)f(b). \tag{18.307}$$

□

Théorème 18.80 ([237])**.**
Une fonction f de classe C^2 est convexe si et seulement si f'' est positive.

Démonstration. La fonction est C^2, donc f'' est positive si et seulement si f' est croissante (proposition 12.125) alors que la proposition 18.79 nous jure que f sera convexe si et seulement si f' est croissante. □

Remarque 18.81.
Une fonction peut être strictement convexe sans que sa dérivée seconde ne soit toujours strictement positive. En exemple : $x \mapsto x^4$ est strictement convexe alors que sa dérivée seconde s'annule en zéro.

Exemple 18.82
Quelques exemples utilisant le théorème 18.80

(1) La fonction $x \mapsto x^2$ est convexe parce que sa dérivée seconde est la constante (positive) 2.
(2) La fonction $x \mapsto \frac{1}{x}$ est convexe sur $\mathbb{R}^+\backslash\{0\}$ (sa dérivée seconde est $2x^{-3}$).
(3) La fonction exponentielle est également convexe.
(4) La fonction ln est concave parce que la dérivée seconde de $-\ln$ est $\frac{1}{x^2}$ qui est strictement positif.

△

18.9.4 Graphe d'une fonction convexe

L'idée principale du graphe d'une fonction convexe est qu'il est toujours au dessus du graphe de ses tangentes (lorsqu'elles existent). Lorsqu'elles n'existent pas, le lemme 18.77 donne des coefficients directeurs de droites qui vont rester en dessous du graphe de la fonction.

Proposition 18.83 ([239])**.**
Une fonction convexe est strictement convexe si et seulement s'il n'existe aucun intervalle de longueur non nulle sur lequel elle coïncide avec une fonction affine.

Démonstration. Si sur l'intervalle (non réduit à un point) $[x, y]$, la fonction convexe f coïncide avec une fonction affine, alors $f(t) = at + b$ et pour $\lambda \in]0, 1[$ nous avons

$$f\big(\lambda x + (1-\lambda)y\big) = a\lambda x + a(1-\lambda)y + b = \lambda f(x) + (1-\lambda)f(y) \tag{18.308}$$

où nous avons remplacé b par $\lambda b + (1-\lambda)b$. Par conséquent la fonction n'est pas strictement convexe.

Nous supposons maintenant que la fonction convexe f n'est pas strictement convexe sur l'intervalle I. Il existe $x \neq y \in I$ et $\lambda \in]0, 1[$ tels que

$$f\big(\lambda x + (1-\lambda)y\big) = \lambda f(x) + (1-\lambda)f(y). \tag{18.309}$$

Nous posons $z = \lambda x + (1-\lambda)y$ et $u \in]x, z[$ pour écrire des inégalités des pentes entre $x < u < z < y$. Plus précisément si nous notons $a \to b$ la pente de a à b, c'est à dire $a \to b = \frac{f(b)-f(a)}{b-a}$, alors les inégalités des pentes pour $x < u < z$ puis $u < z < y$ donnent

$$x \to z \leqslant u \to z \leqslant z \to y. \tag{18.310}$$

Voyons maintenant qu'en réalité $z \to y = x \to z$. En effet en replaçant

$$f(y) = \frac{f(z) - \lambda f(x)}{1-\lambda} \tag{18.311}$$

et

$$y = \frac{\lambda x}{1-\lambda} \tag{18.312}$$

dans l'expression $z \to y = \frac{f(y)-f(z)}{y-z}$ nous obtenons

$$z \to y = \frac{f(y) - f(z)}{y - z} = \frac{f(z) - f(x)}{z - x} = x \to z. \tag{18.313}$$

Les inégalités (18.310) sont donc des égalités :

$$\frac{f(z) - f(x)}{z - x} = \frac{f(z) - f(u)}{z - u} = \frac{f(y) - f(z)}{y - z}. \tag{18.314}$$

Nous avons donc montré que le nombre $a = \frac{f(z)-f(u)}{z-u}$ ne dépend pas de u. Nous avons alors

$$f(z) - f(u) = a(z - u) \tag{18.315}$$

ou encore :

$$f(u) = f(z) - a(z - u), \tag{18.316}$$

ce qui signifie que sur $]x, z[$, la fonction f est affine. □

Proposition 18.84.
Une fonction dérivable sur un intervalle I de $\mathbb{R}$ est convexe si et seulement si son graphe est au dessus de chacune de ses tangentes.

Démonstration. **Sens direct** Soient $x, y \in I$. Nous voulons :

$$f(y) \geqslant f(x) + f'(x)(y - x). \tag{18.317}$$

Étant donné que nous aurons besoin, dans le quotient différentiel de quelque chose comme $f(x + t) - f(x)$ nous écrivons la définition (18.286) de la convexité en inversant les rôles de x et y et en manipulant un peu :

$$f\big(ty + (1 - t)x\big) \leqslant tf(y) + (1 - t)f(x) \tag{18.318a}$$
$$f\big(x + t(y - x)\big) \leqslant tf(y) + (1 - t)f(x) \tag{18.318b}$$
$$f\big(x + t(y - x)\big) = f(x) \leqslant tf(y) - tf(x) \tag{18.318c}$$

Nous divisons par t :

$$\frac{f\big(x + t(y - x)\big) - f(x)}{t} \leqslant f(y) - f(x). \tag{18.319}$$

Le passage à la limite $t \to 0$ donne

$$(y - x)f'(x) \leqslant f(y) - f(x), \tag{18.320}$$

ce qu'il fallait.

Sens inverse Pour tout $x, y \in I$ nous supposons avoir

$$f(y) \geqslant f(x) + f'(x)(y - x). \tag{18.321}$$

Si nous supposons $x \neq y$ et si nous posons $z = \lambda x + (1 - \lambda)y$ nous voulons prouver que

$$f(z) \leqslant \lambda f(x) + (1 - \lambda)f(y). \tag{18.322}$$

Pour cela nous écrivons l'inégalité (18.321) avec les couples (x, z) et (y, z) :

$$f(x) \geqslant f(z) + f'(z)'(x - z) \tag{18.323a}$$
$$f(y) \geqslant f(z) + f'(z)'(y - z) \tag{18.323b}$$

En multipliant la première par λ et la seconde par $(1 - \lambda)$ et en sommant,

$$\lambda f(x) + (1 - \lambda)f(y) \geqslant \lambda f(z) + \lambda f'(z)(x - z) + (1 - \lambda)f(z) + (1 - \lambda)f'(z)(y - z) \tag{18.324a}$$
$$= f(z) + f'(z)\big(\lambda(x - z) + (1 - \lambda)(y - z)\big) \tag{18.324b}$$
$$= f(z). \tag{18.324c}$$

□

Proposition 18.85 ([1]).
Soit $f\colon \mathbb{R} \to \mathbb{R}$ une fonction convexe et $a \in \mathbb{R}$. Il existe une constante $c_a \in \mathbb{R}$ telle que pour tout x nous ayons

$$f(x) - f(a) \geqslant c_a(x - a). \tag{18.325}$$

Autrement dit, le graphe de la fonction f est toujours au dessus de la droite d'équation

$$y = f(a) + c_a(x - a). \tag{18.326}$$

Démonstration. Les dérivées à gauche et à droite de f données par le lemme 18.77 sont les candidats tout cuits pour être coefficient directeur de la droite que l'on cherche. Nous allons prouver qu'en posant

$$c_a = \inf_{t>a} \tau_a(t), \tag{18.327}$$

la droite $y = f(a) + c_a(x - a)$ répond à la question [37].

Nous devons prouver que le nombre $\Delta_x = f(x) - \big(f(a) + c_a(x - a)\big)$ est positif pour tout x.

Si $x > a$ Nous divisons par $x - a$ et nous devons prouver que $\frac{\Delta_x}{x-a}$ est positif :

$$\begin{aligned}
\frac{\Delta_x}{x - a} &= \frac{f(x) - f(a)}{x - a} - c_a && (18.328a)\\
&= \tau_a(x) - \inf_{t>a} \tau_a(t) && (18.328b)\\
&\geqslant 0 && (18.328c)
\end{aligned}$$

parce que $t \to \tau_a(t)$ est croissante et que $x > a$.

Si $x < a$ Nous divisons par $x - a$ et nous devons prouver que $\frac{\Delta_x}{x-a}$ est négatif :

$$\begin{aligned}
\frac{\Delta_x}{x - a} &= \frac{f(x) - f(a)}{x - a} - c_a && (18.329a)\\
&= \tau_a(x) - \inf_{t>a} \tau_a(t) && (18.329b)\\
&\leqslant 0 && (18.329c)
\end{aligned}$$

parce que $t \to \tau_a(t)$ est croissante et que $x < a$.

□

Proposition 18.86 ([1]).
Si g est une fonction convexe, il existe deux suites réelles (a_n) et (b_n) telles que

$$g(x) = \sup_{n\in\mathbb{N}}(a_n x + b_n). \tag{18.330}$$

Démonstration. Pour $u \in \mathbb{R}$ nous considérons $a(u)$ et $b(u)$ tels que la droite $y(x) = a(u)x + b(u)$ vérifie $y(u) = g(u)$ et $y(x) \leqslant g(x)$ pour tout x. Cela est possible par la proposition 18.85. Il s'agit d'une droite coupant le graphe de g en $x = u$ et restant en dessous. Nous considérons alors (u_n) une suite quelconque dense dans $\mathbb{R}$ (disons les rationnels pour fixer les idées) et nous posons

$$\begin{cases}
a_n = a(u_n) & (18.331a)\\
b_n = b(u_n). & (18.331b)
\end{cases}$$

Si $q \in \mathbb{Q}$ alors $a_n x + b_n \leqslant g(x)$ pour tout n et $g(q)$ est le supremum qui est atteint pour le n tel que $u_n = q$. Si maintenant x n'est pas dans $\mathbb{Q}$ il faut travailler plus.

37. En prenant l'autre, $c'_a = \sup_{t<a} \tau_a(t)$, ça fonctionne aussi. En pensant à une fonction affine par morceaux, on remarque qu'en choisissant un nombre entre les deux, nous avons plus facilement une inégalité stricte dans (18.325).

Nous prenons $(\tilde{q}_n)$, une sous-suite de (q_n) convergeant vers x et N suffisamment grand pour que pour tout $n \geqslant N$ on ait $|\tilde{q}_n - x| \leqslant \epsilon$ et $|g(\tilde{q}_n) - g(x)| \leqslant \epsilon$; cela est possible grâce à la continuité de g (lemme 18.77). Ensuite les sous-suites $(\tilde{a}_n)$ et $(\tilde{b}_n)$ sont celles qui correspondent :

$$\tilde{a}_n\tilde{q}_n + \tilde{b}_n = g(\tilde{q}_n). \tag{18.332}$$

Nous considérons la majoration

$$|\tilde{a}_n x + \tilde{b}_n - g(x)| \leqslant |\tilde{a}_n x + \tilde{b}_n - (\tilde{a}_n\tilde{q}_n + \tilde{b}_n)| + \underbrace{|\tilde{a}_n\tilde{q}_n + \tilde{b}_n - g(\tilde{q}_n)|}_{=0} + \underbrace{|g(\tilde{q}_n) - g(x)|}_{\leqslant\epsilon} \tag{18.333a}$$

$$\leqslant |\tilde{a}_n||x - \tilde{q}_n| + \epsilon \tag{18.333b}$$

$$= \epsilon\big(|\tilde{a}_n| + 1\big). \tag{18.333c}$$

Il nous reste à montrer que $|\tilde{a}_n|$ est borné par un nombre ne dépendant pas de n (pour les $n > N$).

Vu que la droite de coefficient directeur $\tilde{a}_n$ et passant par le point $\big(\tilde{q}_n, g(\tilde{q}_n)\big)$ reste en dessous du graphe de g, nous avons pour tout n et tout $y \in \mathbb{R}$ l'inégalité

$$g(y) \geqslant \tilde{a}_n(y - \tilde{q}_n) + g(\tilde{q}_n) \in \tilde{a}_n B(y - x, \epsilon) + B\big(g(x), \epsilon\big). \tag{18.334}$$

Si $\tilde{a}_n$ n'est pas borné vers le haut, nous prenons y tel que $B(y - x, \epsilon)$ soit minoré par un nombre k strictement positif et nous obtenons

$$g(y) \geqslant k\tilde{a}_n + l \tag{18.335}$$

avec k et l indépendants de n. Cela donne $g(y) = \infty$. Si au contraire $\tilde{a}_n$ n'est pas borné vers le bas, nous prenons y tel que $B(y - x, \epsilon)$ est majoré par un nombre k strictement négatif. Nous obtenons encore $g(y) = \infty$.

Nous concluons que $|\tilde{a}_n|$ est bornée. □

Lemme 18.87 ([2])**.**
L'application

$$\begin{aligned}\phi\colon S^{++}(n, \mathbb{R}) &\to \mathbb{R}\\ A &\mapsto \det(A)\end{aligned} \tag{18.336}$$

est ***log-convave****, c'est à dire que l'application* $\ln\circ\phi$ *est concave*[38]*. De façon équivalente, si* $A, B \in S^{++}$ *et si* $\alpha + b = 1$*, alors*

$$\det(\alpha A + \beta B) \geqslant \det(A)^\alpha \det(B)^\beta. \tag{18.337}$$

Ici S^{++} est l'ensemble des matrices symétriques strictement définies positives, définition 9.167.

Démonstration. Nous commençons par prouver que l'équation (18.337) est équivalente à la log-concavité du déterminant. Pour cela il suffit de remarquer que les propriétés de croissance et d'additivité du logarithme donnent l'équivalence entre

$$\ln\Big(\det(\alpha A + \beta B)\Big) \geqslant \ln\Big(\det(\alpha A)\Big) + \ln\Big(\det(\beta B)\Big), \tag{18.338}$$

et

$$\det(\alpha A + \beta B) \geqslant \det(A)^\alpha \det(B)^\beta. \tag{18.339}$$

Le théorème de pseudo-réduction simultanée, corollaire 9.173, appliqué aux matrices A et B nous donne une matrice inversible Q telle que

$$\begin{cases} B = Q^t D Q & \text{(18.340a)}\\ A = Q^t Q & \text{(18.340b)}\end{cases}$$

38. La définition 15.57 du logarithme ne fonctionne que pour les réels strictement positifs. C'est le cas du déterminant d'une matrice réelle symétrique strictement définie positive.

avec

$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \tag{18.341}$$

$\lambda_i > 0$. Nous avons alors

$$\det(A)^\alpha \det(B)^\beta = \det(Q)^{2\alpha} \det(Q)^{2\beta} \det(D)^\beta = \det(Q)^2 \det(D)^\beta \tag{18.342}$$

(parce que $\alpha + \beta = 1$) et

$$\det(\alpha A + \beta B) = \det(\alpha Q^t Q + \beta Q^t D Q) = \det\big(Q^t(\alpha \mathbb{1} + \beta D)Q\big) = \det(Q)^2 \det(\alpha \mathbb{1} + \beta D). \tag{18.343}$$

L'inégalité (18.339) qu'il nous faut prouver se réduit donc à

$$\det(\alpha \mathbb{1} + \beta D) \geqslant \det(D)^\beta. \tag{18.344}$$

Vue la forme de D nous avons

$$\det(\alpha \mathbb{1} + \beta D) = \prod_{i=1}^{n} (\alpha + \beta \lambda_i) \tag{18.345}$$

et

$$\det(D)^\beta = \big(\prod_{i=1}^{n} \lambda_i\big)^\beta. \tag{18.346}$$

Il faut donc prouver que

$$\prod_{i=1}^{n} (\alpha + \beta \lambda_i) \geqslant \big(\prod_{i=1}^{n} \lambda_i\big)^\beta. \tag{18.347}$$

Cette dernière égalité de produit sera prouvée en passant au logarithme. Vu que le logarithme est concave par l'exemple 18.82, nous avons pour chaque i que

$$\ln(\alpha + \beta \lambda_i) \geqslant \alpha \ln(1) + \beta \ln(\lambda_i) = \beta \ln(\lambda_i). \tag{18.348}$$

En sommant cela sur i et en utilisant les propriétés de croissance et de multiplicativité du logarithme nous obtenons successivement

$$\sum_{i=1}^{n} \ln(\alpha + \beta \lambda_i) \geqslant \beta \sum_i \ln(\lambda_i) \tag{18.349a}$$

$$\ln\big(\prod_i (\alpha + \beta \lambda_i)\big) \geqslant \ln\Big(\big(\prod_i \lambda_i\big)^\beta\Big) \tag{18.349b}$$

$$\prod_i (\alpha + \beta \lambda_i) \geqslant \big(\prod_i \lambda_i\big)^\beta, \tag{18.349c}$$

ce qui est bien (18.347). □

18.9.5 Convexité et hessienne

Définition 18.88.
Soit une partie convexe U de $\mathbb{R}^n$ et une fonction $f : U \to \mathbb{R}$.

*(1) La fonction f est **convexe** si pour tout $x, y \in U$ avec $x \neq y$ et pour tout $\theta \in]0,1[$ nous avons*

$$f\big(\theta x + (1-\theta)y\big) \leqslant \theta f(x) + (1-\theta) f(y). \tag{18.350}$$

*(2) Elle est **strictement convexe** si nous avons l'inégalité stricte.*

Proposition 18.89 ([233])**.**
Soit Ω ouvert dans $\mathbb{R}^n$ et U convexe dans Ω, et une fonction différentiable $f : U \to \mathbb{R}$.

(1) La fonction f est convexe sur U si et seulement si pour tout $x, y \in U$,

$$f(y) \geqslant f(x) + df_x(y - x). \tag{18.351}$$

(2) La fonction f est strictement convexe sur U si et seulement si pour tout $x, y \in U$ avec $x \neq y$,

$$f(y) > f(x) + df_x(y - x). \tag{18.352}$$

Démonstration. Nous avons quatre petites choses à démontrer.

(1) sens direct Soit une fonction convexe f. Nous avons :

$$f\big((1-\theta)x + \theta y\big) \leqslant (1-\theta)f(x) + \theta f(y), \tag{18.353}$$

donc

$$f\big(x + \theta(y-x)\big) - f(x) \leqslant \theta\big(f(y) - f(x)\big) \tag{18.354}$$

Vu que $\theta > 0$ nous pouvons diviser par θ sans changer le sens de l'inégalité :

$$\frac{f\big(x + \theta(y-x)\big) - f(x)}{\theta} \leqslant f(y) - f(x). \tag{18.355}$$

Nous prenons la limite $\theta \to 0^+$. Cette limite est égale à a limite simple $\theta \to 0$ et vaut (parce que f est différentiable) :

$$\frac{\partial f}{\partial(y-x)}(x) \leqslant f(y) - f(x), \tag{18.356}$$

et aussi

$$df_x(y-x) \leqslant f(y) - f(x) \tag{18.357}$$

par le lemme 12.178.

(1) sens inverse Pour tout $a \neq b$ dans U nous avons

$$f(b) \geqslant f(a) + df_a(b-a). \tag{18.358}$$

Pour $x \neq y$ dans U et pour $\theta \in]0,1[$ nous écrivons (18.358) pour les couples $\big(\theta x + (1-\theta)y, y\big)$ et $\big(\theta x + (1-\theta)y, x\big)$. Ça donne :

$$f(y) \geqslant f\big(\theta x + (1-\theta)y\big) + df_{\theta x+(1-\theta)y}\big(\theta(y-x)\big), \tag{18.359}$$

et

$$f(x) \geqslant f\big(\theta x + (1-\theta)y\big) + df_{\theta x+(1-\theta)y}\big((1-\theta)(x-y)\big). \tag{18.360}$$

La différentielle est linéaire ; en multipliant la première par $(1-\theta)$ et la seconde par θ et en la somme, les termes en df se simplifient et nous trouvons

$$\theta f(x) + (1-\theta)f(y) \geqslant f\big(\theta x + (1-\theta)y\big). \tag{18.361}$$

(2) sens direct Nous avons encore l'équation (18.355), avec une inégalité stricte. Par contre, ça ne va pas être suffisant parce que le passage à la limite ne conserve pas les inégalités strictes. Nous devons donc être plus malins.

Soient $0 < \theta < \omega < 1$. Nous avons $(1-\theta)x + \theta y \in [x, (1-\omega)x + \omega y]$, donc nous pouvons écrire $(1-\theta)x + \theta y$ sous la forme $(1-s)x + s\big((1-\omega)x + \omega y\big)$. Il se fait que c'est bon pour $s = \theta/\omega$ (et aussi que nous avons $\theta/\omega < 1$). Donc nous avons

$$f\big((1-\theta)x + \theta y\big) = f\Big((1 - \frac{\theta}{\omega})x + \frac{\theta}{\omega}\big((1-\omega)x + \omega y\big)\Big) \tag{18.362a}$$

$$< (1 - \frac{\theta}{\omega})f(x) + \frac{\theta}{\omega}f\big((1-\omega)x + \omega y\big). \tag{18.362b}$$

Cela nous permet d'écrire

$$\frac{f\big((1-\theta)x+\theta y\big)-f(x)}{\theta}<\frac{f\big((1-\omega)x+\omega y\big)}{\omega}<f(y)-f(x). \tag{18.363}$$

Le seconde inégalité est le pendant de (18.355). Maintenant en passant à la limite pour θ nous conservons une inégalité stricte par rapport à $f(y)-f(x)$:

$$df_x(y-x)<f(y)-f(x). \tag{18.364}$$

□

Avant de lire la proposition suivante, il faut relire la proposition 12.259 et ce qui s'y rapporte. Lire aussi la remarque 18.81 qui indique qu'il n'y a pas de réciproque dans l'énoncé (2).

Proposition 18.90 ([233])**.**
Soit une fonction $f\colon \Omega\to\mathbb{R}$ deux fois différentiable sur l'ouvert Ω de $\mathbb{R}^n$ et un convexe $U\subset\Omega$.

(1) La fonction f est convexe sur U si et seulement si

$$(d^2f)_x(y-x,y-x)\geqslant 0 \tag{18.365}$$

pour tout $x,y\in U$.

(2) Si pour tout $x\neq y$ dans U nous avons

$$(d^2f)_x(y-x,y-x)>0 \tag{18.366}$$

alors la fonction f est strictement convexe sur U.

Remarque 18.91.
Notons que la condition (18.365) n'est pas équivalente à demander $(d^2f)_x(h,h)\geqslant 0$ pour tout h. En effet nous ne demandons la positivité que dans les directions atteignables comme différence de deux éléments de U. La partie U n'est pas spécialement ouverte ; elle pourrait n'être qu'une droite dans $\mathbb{R}^3$. Dans ce cas, demander que f (qui est C^2 sur l'ouvert Ω) soit convexe sur U ne demande que la positivité de $(d^2f)_x$ appliqué à des vecteurs situés sur la droite U.

Démonstration. Il y a trois parties à démontrer.

(1) sens direct Soit une fonction convexe f sur U. Soient aussi $x,y\in U$ et $h=y-x$. Nous utilisons ma version préférée de Taylor [39] : celui de la proposition 12.267 :

$$f(x+th)=f(x)+tdf_x(h)+\frac{t^2}{2}(d_x^2)(h,h)+t^2\|h\|^2\alpha(th) \tag{18.367}$$

avec $\lim_{s\to 0}\alpha(s)=0$. Le fait que f soit convexe donne

$$0\leqslant f(x+th)-f(x)-tdf_x(h), \tag{18.368}$$

et donc

$$0\leqslant\frac{t^2}{2}(d^2f)_x(h,h)+f^2\|h\|^2\alpha(th). \tag{18.369}$$

En multipliant par 2 et en divisant par t^2,

$$0\leqslant(d^2f)_x(h,h)+2\|h\|^2\alpha(th). \tag{18.370}$$

En prenant $t\to 0$ nous avons bien $(d^2f)_x(y-x,y-x)\geqslant 0$.

39. Si vous présentez ceci au jury d'un concours, vous devriez être capable de raconter ce que signifie d^2f, et pourquoi nous l'utilisons comme une 2-forme.

(1) sens inverse Soient $x, y \in U$. Nous écrivons Taylor en version de la proposition 12.268 :

$$f(y) = f(x) + df_x(y-x) + \frac{1}{2}(d^2f)_z(y-x, y-x) \tag{18.371}$$

pour un certain $z \in]x, y[$. En vertu de ce qui a été dit dans la remarque 18.91 nous ne pouvons pas évoquer l'hypothèse (18.365) pour conclure que $(d^2f)_z(y-x, y-x) \geqslant 0$. Il y a deux manières de nous sortir du problème :

— Trouver $s \in U$ tel que $y - x = s - z$.

— Trouver un multiple de $y - x$ qui soit de la forme $y - x$.

La première approche ne fonctionne pas parce que $s = y - x + z$ n'est pas garanti d'être dans U ; par exemple avec $x = 1$, $z = 2$, $y = 3$ et $U = [0, 3]$. Dans ce cas $s = 4 \notin U$.

Heureusement nous avons $z = \theta x + (1-\theta)y$, donc $z - x = (1-\theta)(y-x)$. Dans ce cas la bilinéarité de $(d^2f)_z$ donne [40]

$$f(y) = f(x) + df_x(y-x) + \underbrace{\frac{1}{2}\frac{1}{(1-\theta)^2}(d^2f)_z(z-x, z-x)}_{\geqslant 0}. \tag{18.372}$$

Nous en déduisons que f est convexe par la proposition 18.89(1).

(2) Le raisonnement que nous venons de faire pour le sens inverse de (1) tient encore, et nous avons

$$f(y) = f(x) + df_x(y-x) + \underbrace{\frac{1}{2}\frac{1}{(1-\theta)^2}(d^2f)_z(z-x, z-x)}_{>0} \tag{18.373}$$

d'où nous déduisons la stricte convexité de f par la proposition 18.89(2).

□

Corollaire 18.92.
Soit un ouvert Ω de $\mathbb{R}^n$ et une fonction deux fois différentiable f sur Ω.

(1) La fonction f est convexe si et seulement si pour tout x, la matrice hessienne d^2f_x est semi-définie positive.

(2) Si pour tout x de Ω, la matrice hessienne d^2f_x est strictement définie positive, alors f est strictement convexe.

Démonstration. Nous pouvons voir ce résultat comme une conséquence directe de la proposition 18.90 en posant $U = \Omega$. Nous allons cependant en donner une démonstration directe.

Soit $a \in \Omega$ et posons la fonction

$$\begin{aligned} g\colon \Omega &\to \mathbb{R} \\ x &\mapsto f(x) - f(a) - (df)_a(x-a). \end{aligned} \tag{18.374}$$

Nous allons calculer des différentielles de f, et une chose importante à comprendre est que la différentielle de la fonction $x \mapsto df_a(x-a)$ ne fait pas intervenir la différentielle seconde de f ; c'est la différentielle de $a \mapsto df_a(x)$ qui demanderait la différentielle seconde de f. Ici la point a étant donné, df_a est une application linéaire sans histoires. En particulier, $df_a(x-a) = df_a(x) - df_a(a)$.

La fonction g vérifie :

(1) $g(a) = 0$,

(2) $dg_x = df_x - df_a$, parce que la différentielle de $x \mapsto df_a(x)$ est $x \mapsto df_a(x)$ en vertu du lemme 12.217.

(3) $dg_a = 0$. Le point a est un point critique de g.

(4) $d^2g_x = d^2f_x$ parce que la différentielle de $x \mapsto df_a$ est nulle.

40. Si vous avez bien suivi, la bilinéarité est contenue dans la proposition 12.259.

Ceci étant dit, nous pouvons commencer avec la preuve.

(1) sens direct Nous supposons que f est convexe. Alors $g(x) \geqslant 0$ pour tout x par la caractérisation 18.89(1). Cela signifie que $x = 0$ est un minimum global de g. Par conséquent la proposition 18.68(3) nous dit que la Hessienne $d^2 f_a$ est semi-définie positive.

(1) sens inverse Nous sommes dans le cas de la proposition 18.68(1). Le point $x = a$ est un minimum local de g, ce qui signifie que $g(x) \geqslant 0$ pour tout x de Ω. Encore une fois la caractérisation 18.89(1) nous permet de conclure.

(2) La fonction g vérifie les conditions de 18.68(2), donc $x = 0$ est un minimum local strict de g. La caractérisation 18.89(2) nous fait conclure que f est strictement convexe.

□

18.9.6 Quelque inégalités

18.9.6.1 Inégalité de Jensen

Proposition 18.93 (Inégalité de Jensen)**.**
Soit $f\colon \mathbb{R} \to \mathbb{R}$ une fonction convexe et des réels $x_1,\dots, x_n$. Soient des nombres positifs $\lambda_1,\dots, \lambda_n$ formant une combinaison convexe[41]*. Alors*

$$f\big(\sum_i \lambda_i x_i\big) \leqslant \sum_i \lambda_i f(x_i). \tag{18.375}$$

Démonstration. Nous procédons par récurrence sur n, en sachant que $n = 2$ est la définition de la convexité de f. Vu que

$$\sum_{k=1}^{n} \lambda_k x_k = \lambda_n x_n + (1 - \lambda_n) \sum_{k=1}^{n-1} \frac{\lambda_k x_k}{1 - \lambda_n}, \tag{18.376}$$

nous avons

$$f\big(\sum_{k=1}^{n} \lambda_k x_k\big) \leqslant \lambda_n f(x_n) + (1 - \lambda_n) f\big(\sum_{k=1}^{n-1} \frac{\lambda_k x_k}{1 - \lambda_n}\big). \tag{18.377}$$

La chose à remarquer est que les nombres $\frac{\lambda_k}{1-\lambda_n}$ avec k allant de 1 à $n - 1$ forment eux-mêmes une combinaison convexe. L'hypothèse de récurrence peut donc s'appliquer au second terme du membre de droite :

$$f\big(\sum_{k=1}^{n} \lambda_k x_k\big) \leqslant \lambda_n f(x_n) + (1 - \lambda_n) \sum_{k=1}^{n-1} \frac{\lambda_k}{1 - \lambda_n} f(x_k) = \lambda_n f(x_n) + \sum_{k=1}^{n-1} \lambda_k f(x_k). \tag{18.378}$$

□

18.9.6.2 Inégalité arithmético-géométrique

La proposition suivante dit que la moyenne arithmétique de nombres strictement positifs est supérieure ou égale à la moyenne géométrique.

Proposition 18.94 (Inégalité arithmético-géométrique[240])**.**
Soient $x_1,\dots, x_n$ des nombres strictement positifs. Nous posons

$$m_a = \frac{1}{n}(x_1 + \cdots + x_n) \tag{18.379}$$

et

$$m_g = \sqrt[n]{x_1 \dots x_n} \tag{18.380}$$

Alors $m_g \leqslant m_a$ et $m_g = m_a$ si et seulement si $x_i = x_j$ pour tout i, j.

41. Définition 11.25.

Démonstration. Par hypothèse les nombres m_a et m_g sont tout deux strictement positifs, de telle sorte qu'il est équivalent de prouver $\ln(m_g) \leqslant \ln(m_a)$ ou encore

$$\frac{1}{n}\big(\ln(x_1) + \cdots + \ln(x_n)\big) \leqslant \ln\left(\frac{x_1 + \cdots + x_n}{n}\right). \tag{18.381}$$

Cela n'est rien d'autre que l'inégalité de Jensen de la proposition 18.93 appliquée à la fonction ln et aux coefficients $\lambda_i = \frac{1}{n}$. □

18.9.6.3 Inégalité de Kantorovitch

Proposition 18.95 (Inégalité de Kantorovitch[241])**.**
Soit A une matrice symétrique strictement définie positive dont les plus grandes et plus petites valeurs propres sont λ_{min} et λ_{max}. Alors pour tout $x \in \mathbb{R}^n$ nous avons

$$\langle Ax, x\rangle\langle A^{-1}x, x\rangle \leqslant \frac{1}{4}\left(\frac{\lambda_{min}}{\lambda_{max}} + \frac{\lambda_{max}}{\lambda_{min}}\right)^2 \|x^4\|. \tag{18.382}$$

Démonstration. Sans perte de généralité nous pouvons supposer que $\|x\| = 1$. Nous diagonalisons [42] la matrice A par la matrice orthogonale $P \in \mathrm{O}(n, \mathbb{R})$: $A = PDP^{-1}$ et $A^{-1} = PD^{-1}P^{-1}$ où D est une matrice diagonale formée des valeurs propres de A.

Nous posons $\alpha = \sqrt{\lambda_{min}\lambda_{max}}$ et nous regardons la matrice

$$\frac{1}{\alpha}A + tA^{-1} \tag{18.383}$$

dont les valeurs propres sont

$$\frac{\lambda_i}{\alpha} + \frac{\alpha}{\lambda_i} \tag{18.384}$$

parce que les vecteurs propres de A et de A^{-1} sont les mêmes (ce sont les valeurs de la diagonale de D). Nous allons quelque peu étudier la fonction

$$\theta(x) = \frac{x}{\alpha} + \frac{\alpha}{x}. \tag{18.385}$$

Elle est convexe en tant que somme de deux fonctions convexes. Elle a son minimum en $x = \alpha$ et ce minimum vaut $\theta(\alpha) = 2$. De plus

$$\theta(\lambda_{max}) = \theta(\lambda_{min}) = \sqrt{\frac{\lambda_{min}}{\lambda_{max}}} + \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}. \tag{18.386}$$

Une fonction convexe passant deux fois par la même valeur doit forcément être plus petite que cette valeur entre les deux [43] : pour tout $x \in [\lambda_{min}, \lambda_{max}]$,

$$\theta(x) \leqslant \sqrt{\frac{\lambda_{min}}{\lambda_{max}}} + \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}. \tag{18.387}$$

Nous sommes maintenant en mesure de nous lancer dans l'inégalité de Kantorovitch.

$$\sqrt{\langle Ax, x\rangle\langle A^{-1}x, x\rangle} \leqslant \frac{1}{2}\left(\frac{\langle Ax, x\rangle}{\alpha} + \alpha\langle A^{-1}x, x\rangle\right) \tag{18.388a}$$

$$= \frac{1}{2}\langle\big(\frac{A}{\alpha} + \alpha A^{-1}\big)x, x\rangle \tag{18.388b}$$

$$\leqslant \frac{1}{2}\Big\|\big(\frac{A}{\alpha} + \alpha A^{-1}\big)x\Big\|\|x\| \tag{18.388c}$$

$$\leqslant \frac{1}{2}\Big\|\frac{A}{\alpha} + \alpha A^{-1}\Big\| \tag{18.388d}$$

Justifications :

42. Théorème spectral 9.165.
43. Je ne suis pas certain que cette phrase soit claire, non ?

— 18.388a par l'inégalité arithmético-géométrique, proposition 18.94. Nous avons aussi inséré $\alpha\frac{1}{\alpha}$ dans le produit sous la racine.
— 18.388c par l'inégalité de Cauchy-Schwarz, théorème 9.25.
— 18.388d par la définition de la norme opérateur de la proposition 10.5

La norme opérateur est la plus grande des valeurs propres. Mais les valeurs propres de $A/\alpha+\alpha A^{-1}$ sont de la forme $\theta(\lambda_i)$, et tous les λ_i sont entre λ_{min} et λ_{max}. Donc la plus grande valeur propre de $A/\alpha + \alpha A^{-1}$ est $\theta(x)$ pour un certain $x \in [\lambda_{min}, \lambda_{max}]$. Par conséquent

$$\sqrt{\langle Ax, x\rangle\langle A^{-1}x, x\rangle} \leqslant \frac{1}{2}\|\frac{A}{\alpha} + \alpha A^{-1}\| \leqslant \sqrt{\frac{\lambda_{min}}{\lambda_{max}}} + \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}. \tag{18.389}$$

□

18.10 Algorithme du gradient à pas optimal

Une idée pour trouver un minimum à une fonction est de prendre un point p au hasard, calculer le gradient $\nabla f(p)$ et suivre la direction $-\nabla f(p)$ tant que ça descend. Une fois qu'on est «dans le creux», recalculer le gradient et continuer ainsi.

Nous allons détailler cet algorithme dans un cas très particulier d'une matrice A symétrique et strictement définie positive.

— Dans la proposition 18.97 nous montrons que résoudre le système linéaire $Ax = -b$ est équivalent à minimiser une certaine fonction.
— La proposition 18.98 donnera une méthode itérative pour trouver ce minimum.

Définition 18.96.
Si X est un espace vectoriel normé et $f\colon X \to \mathbb{R} \cup \{\pm\infty\}$ nous disons que f est ***coercive*** *sur le domaine non borné P de X si pour tout $M \in \mathbb{R}$, l'ensemble*

$$\{x \in P \text{ tel que } f(x) \leqslant M\} \tag{18.390}$$

est borné.

En langage imagé la coercivité de f s'exprime par la limite

$$\lim_{\substack{\|x\|\to\infty \\ x\in P}} f(x) = +\infty. \tag{18.391}$$

Nous rappelons que $S^{++}(n, \mathbb{R})$ est l'ensemble des matrice symétriques strictement définies positives définies en 9.168.

Proposition 18.97.
Soit $A \in S^{++}(n, \mathbb{R})$ et $b \in \mathbb{R}^n$. Nous considérons l'application

$$\begin{aligned} f\colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle. \end{aligned} \tag{18.392}$$

Alors :

(1) Il existe un unique $\bar{x} \in \mathbb{R}^n$ tel que $A\bar{x} = -b$.
(2) Il existe un unique $x^* \in \mathbb{R}^n$ minimisant f.
(3) Ils sont égaux : $\bar{x} = x^*$.

Démonstration. Une matrice symétrique strictement définie positive est inversible, entre autres parce qu'elle se diagonalise par des matrices orthogonales (qui sont inversibles) et que la matrice diagonalisée est de déterminant non nul : tous les éléments diagonaux sont strictement positifs. Voir le théorème spectral symétrique 9.165.

D'où l'unicité du $\bar{x}$ résolvant le système $Ax = -b$ pour n'importe quel b.

f est strictement convexe La fonction f s'écrit

$$f(x) = \frac{1}{2}\sum_{kl} A_{kl}x_l x_k + \sum_k b_k x_k. \tag{18.393}$$

Elle est de classe C^2 sans problèmes, et il est vite vu que $\frac{\partial^2 f}{\partial x_i \partial x_j} = A_{ij}$, c'est à dire que A est la matrice hessienne de f. Cette matrice étant strictement définie positive par hypothèse, la fonction f est strictement convexe par le corollaire 18.92(2).

f est coercive Montrons à présent que f est coercive. Nous avons :

$$|f(x)| = |\frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle| \tag{18.394a}$$
$$\geqslant \frac{1}{2}|\langle Ax, x\rangle| - |\langle b, x\rangle| \tag{18.394b}$$
$$\geqslant \frac{1}{2}\lambda_{max}\|x\|^2 - \|b\|\|x\| \tag{18.394c}$$

Pour la dernière ligne nous avons nommé λ_{max} la plus grande valeur propre de A et utilisé Cauchy-Schwarz pour le second terme. Nous avons donc bien $|f(x)| \to \infty$ lorsque $\|x\| \to \infty$ et la fonction f est coercive.

Soit M une valeur atteinte par f. L'ensemble

$$\{x \in \mathbb{R}^n \text{ tel que } f(x) \leqslant M\} \tag{18.395}$$

est fermé (parce que f est continue) et borné parce que f est coercive. Cela est donc compact [44] et f atteint un minimum qui sera forcément dedans. Cela est pour l'existence d'un minimum.

Pour l'unicité du minimum nous invoquons la convexité : si $\bar{x}_1$ et $\bar{x}_2$ sont deux points réalisant le minimum de f, alors

$$f\left(\frac{\bar{x}_1 + \bar{x}_2}{2}\right) < \frac{1}{2}f(\bar{x}_1) + \frac{1}{2}f(\bar{x}_2) = f(\bar{x}_1), \tag{18.396}$$

ce qui contredit la minimalité de $f(\bar{x}_1)$.

Nous devons maintenant prouver que $\bar{x}$ vérifie l'équation $A\bar{x} = -b$. Vu que $\bar{x}$ est minimum local de f qui est une fonction de classe C^2, le théorème des minima locaux 18.67 nous indique que $\bar{x}$ est solution de $\nabla f(x) = 0$. Calculons un peu cela avec la formule

$$df_x(u) = \frac{d}{dt}\Big[f(x+tu)\Big]_{t=0} = \frac{1}{2}\big(\langle Ax, u\rangle + \langle Au, x\rangle\big) + \langle b, u\rangle = \langle Ax, u\rangle + \langle b, u\rangle = \langle Ax+b, u\rangle. \tag{18.397}$$

Donc demander $df_x(u) = 0$ pour tout u demande $Ax + b = 0$. □

Proposition 18.98 (Gradient à pas optimal)**.**
Soit $A \in S^{++}(n, \mathbb{R})$ (A est une matrice symétrique strictement définie positive) et $b \in \mathbb{R}^n$. Nous considérons l'application

$$\begin{aligned} f\colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle. \end{aligned} \tag{18.398}$$

Soit $x_0 \in \mathbb{R}^n$. Nous définissons la suite (x_k) par

$$x_{k+1} = x_k + t_k d_k \tag{18.399}$$

où

- *$d_k = -(\nabla f)(x_k)$*
- *t_k est la valeur minimisant la fonction $t \mapsto f(x_k + td_k)$ sur $\mathbb{R}$.*

Alors pour tout $k \geqslant 0$ nous avons

$$\|x_k - \bar{x}\| \leqslant K\left(\frac{c_2(A) - 1}{c_2(A) + 1}\right)^k \tag{18.400}$$

où $c_2(A) = \frac{\lambda_{max}}{\lambda_{min}}$ est le rapport ente la plus grande et la plus petite valeur propre [45] de la matrice

44. Théorème 6.123
45. Cela est certainement très lié au conditionnement de la matrice A, voir la proposition 32.104.

A et $\bar{x}$ est l'unique élément de $\mathbb{R}^n$ à minimiser f.

Démonstration. Décomposition en plusieurs points.

Existence de $\bar{x}$ Le fait que $\bar{x}$ existe et soit unique est la proposition 18.97.

Si $(\nabla f)(x_k) = 0$ D'abord si $\nabla f(x_k) = 0$, c'est que $x_{k+1} = x_k$ et l'algorithme est terminé : la suite est stationnaire. Pour dire que c'est gagné, nous devons prouver que $x_k = \bar{x}$. Pour cela nous écrivons (à partir de maintenant «x_k» est la k^{e}composante de x qui est une variable, et non le x_k de la suite)

$$f(x) = \frac{1}{2}\sum_{kl} A_{kl}x_l x_k + \sum_k b_k x_k \tag{18.401}$$

et nous calculons $\frac{\partial f}{\partial x_i}(a)$ en tenant compte du fait que $\frac{\partial x_k}{\partial x_i} = \delta_{ki}$. Le résultat est que $(\partial_i f)(a) = (Ax + b)_i$ et donc que

$$(\nabla f)(a) = Aa + b. \tag{18.402}$$

Vu que A est inversible (symétrique définie positive), il existe un unique $a \in \mathbb{R}^n$ qui vérifie cette relation. Par la proposition 18.97, cet élément est le minimum $\bar{x}$.

Cela pour dire que si $a \in \mathbb{R}^n$ vérifie $(\nabla f)(a) = 0$ alors $a = \bar{x}$. Nous supposons donc à partir de maintenant que $\nabla f(x_k) \neq 0$ pour tout k.

t_k est bien défini Pour $t \in \mathbb{R}$ nous avons

$$f(x_k + td_k) = f(x_k) + \frac{1}{2}t^2\langle Ad_k, d_k\rangle + t\langle \underbrace{Ax_k + b}_{=-d_k}, d_k\rangle = \frac{1}{2}t^2\langle Ad_k, d_k\rangle - t_k\|d_k\|^2 + f(x_k). \tag{18.403}$$

qui est un polynôme du second degré en t. Le coefficient de t^2 est $\frac{1}{2}\langle Ad_k, d_k\rangle > 0$ parce que $d_k \neq 0$ et A est strictement définie positive. Par conséquent la fonction $t \mapsto f(x_k + td_k)$ admet bien un unique minimum. Nous pouvons même calculer t_k parce que l'on connaît pas cœur le sommet d'une parabole :

$$t_k = -\frac{\langle Ax_k + b, d_k\rangle}{\langle Ad_k, d_k\rangle} = \frac{\|d_k\|^2}{\langle Ad_k, d_k\rangle} \tag{18.404}$$

parce que $d_k = -\nabla f(x_k) = -(Ax_k + b)$.

La valeur de d_{k+1} Par définition, $d_{k+1} = -\nabla f(x_{k+1}) = -(Ax_{k+1} + b)$. Mais $x_{k+1} = x_k + t_k d_k$, donc

$$d_{k+1} = -Ax_k - t_k Ad_k - b = d_k - t_k Ad_k \tag{18.405}$$

parce que $-Ax_k - b = d_k$.

Par ailleurs, $\langle d_{k+1}, d_k\rangle = 0$ parce que

$$\langle d_{k+1}, d_k\rangle = \langle d_k, d_k\rangle - t_k\langle d_k, Ad_k\rangle = \|d_k\|^2 - \frac{\|d_k\|^2}{\langle Ad_k, d_k\rangle}\langle d_k, Ad_k\rangle = 0 \tag{18.406}$$

où nous avons utilisé la valeur (18.404) de t_k.

Calcul de $f(x_{k+1})$ Nous repartons de (18.403) où nous substituons la valeur (18.404) de t_k :

$$f(x_{k+1}) = f(x_k) + \frac{1}{2}\frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle} - \frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle} = f(x_k) - \frac{1}{2}\frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle}. \tag{18.407}$$

Encore du calcul ... Vu que le produit $\langle Ad_k, d_k\rangle$ arrive tout le temps, nous allons étudier $\langle A^{-1}d_k, d_k\rangle$. Le truc malin est d'essayer d'exprimer ça en termes de $\bar{x}$ et $\bar{f} = f(\bar{x})$. Pour cela nous calculons $f(\bar{x})$:

$$\bar{f} = f(\bar{x}) = f(-A^{-1}b) = -\frac{1}{2}\langle b, A^{-1}b\rangle. \tag{18.408}$$

Ayant cela en tête nous pouvons calculer :

$$\langle A^{-1}d_k, d_k\rangle = \langle A^{-1}(Ax_k+b), Ax_k+b\rangle \tag{18.409a}$$
$$= \langle x_k, Ax_k\rangle + \langle A^{-1}b, Ax_k\rangle + \langle b, x_k\rangle + \underbrace{\langle A^{-1}b, b\rangle}_{-2\bar{f}} \tag{18.409b}$$
$$= \langle x_k, Ax_k\rangle + 2\langle x_k, b\rangle - 2\bar{f} \tag{18.409c}$$
$$= 2\big(f(x_k)-\bar{f}\big) \tag{18.409d}$$

où nous avons utilisé le fait que $\langle x, Ay\rangle = \langle Ax, y\rangle$ parce que A est symétrique.

Erreur sur la valeur du minimum Nous voulons à présent estimer la différence $f(x_{k+1})-\bar{f}$. Pour cela nous mettons en facteur $f(x_k)-\bar{f}$ dans $f(x_{k+1}-\bar{f})$; et d'ailleurs c'est pour cela que nous avons calculé $\langle A^{-1}d_k, d_k\rangle$: parce que ça fait intervenir $f(x_k)-\bar{f}$.

$$f(x_{k+1})-\bar{f} = f(x_k) - \frac{1}{2}\frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle} - \bar{f} \tag{18.410a}$$
$$= \big(f(x_k)-\bar{f}\big)\left(1 - \frac{1}{2}\frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle\big(f(x_k)-\bar{f}\big)}\right) \tag{18.410b}$$
$$= \big(f(x_k)-\bar{f}\big)\left(1 - \frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle\langle A^{-1}d_k, d_k\rangle}\right). \tag{18.410c}$$

Nous traitons le dénominateur à l'aide de l'inégalité de Kantorovitch 18.95. Nous avons

$$\frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle\langle A^{-1}d_k, d_k\rangle} \geqslant \frac{\|d_k\|^4}{\frac{1}{4}\left(\sqrt{c_2(A)}+\frac{1}{\sqrt{c_2(A)}}\right)^2\|d_k\|^4} = \frac{4c_2(A)}{(c_2(A)+1)^2}. \tag{18.411}$$

Mettre cela dans (18.410c) est un calcul d'addition de fractions :

$$f(x_{k+1})-\bar{f} \leqslant \big(f(x_k)-\bar{f}\big)\left(\frac{c_2(A)-1}{c_2(A)+1}\right)^2. \tag{18.412}$$

Par récurrence nous avons alors

$$f(x_k)-\bar{f} \leqslant \big(f(x_0)-\bar{f}\big)\left(\frac{c_2(A)-1}{c_2(A)+1}\right)^{2k}. \tag{18.413}$$

Notons qu'il n'y a pas de valeurs absolues parce que $\bar{f}$ étant le minimum de f, les deux côtés de l'inégalité sont automatiquement positifs.

Erreur sur la position du minimum Nous voulons à présent étudier la norme de $x_k-\bar{x}$. Pour cela nous l'écrivons directement avec la définition de f en nous souvenant que $b=-A\bar{x}$:

$$f(x_k)-\bar{f} = \frac{1}{2}\langle Ax_k, x_k\rangle + \langle A\bar{x}, x_k\rangle + \frac{1}{2}\langle A\bar{x}, \bar{x}\rangle + \langle A\bar{x}, \bar{x}\rangle \tag{18.414a}$$
$$= \frac{1}{2}\langle Ax_k, x_k\rangle - \langle A\bar{x}, x_k\rangle + \frac{1}{2}\langle A\bar{x}, \bar{x}\rangle \tag{18.414b}$$
$$= \frac{1}{2}\langle Ax_k, x_k\rangle - \frac{1}{2}\langle A\bar{x}, x_k\rangle - \frac{1}{2}\langle A\bar{x}, x_k\rangle + \frac{1}{2}\langle A\bar{x}, \bar{x}\rangle \tag{18.414c}$$
$$= \frac{1}{2}\Big(\langle A(x_k-\bar{x}), x_k\rangle + \langle A\bar{x}, \bar{x}-x_k\rangle\Big) \tag{18.414d}$$
$$= \frac{1}{2}\Big(\langle A(x_k-\bar{x}), (x_k-\bar{x})\rangle\Big) \tag{18.414e}$$

où à la dernière ligne nous avons fait $\langle A\bar{x}, \bar{x}-x_k\rangle = \langle \bar{x}, A(\bar{x}-x_k)\rangle$ en vertu de la symétrie de A.

Les produits de la forme $\langle Ay, y\rangle$ sont majorés par $\lambda_{min}\|y\|^2$ parce que λ_{min} est la plus grande valeur propre de A. Dans notre cas,

$$f(x_k)-\bar{f} \geqslant \frac{1}{2}\lambda_{min}\|x_k-\bar{x}\|^2 \tag{18.415}$$

Conclusion En combinant les inéquations (18.415) et (18.413) nous trouvons

$$\frac{1}{2}\lambda_{min}\|x_k-\bar{x}\|^2 \leqslant f(x_k)-\bar{f} \leqslant \big(f(x_0)-\bar{f}\big)\left(\frac{c_2(A)-1}{c_2(A)+1}\right)^{2k}, \tag{18.416}$$

c'est à dire

$$\|x_k-\bar{x}\| \leqslant \sqrt{\frac{2\big(f(x_0)-\bar{f}\big)}{\lambda_{min}+1}}^{2k}. \tag{18.417}$$

□

Notons que lorsque $c_2(A)$ est proche de 1 la méthode converge rapidement. Par contre si $c_2(A)$ est proche de zéro, la méthode converge lentement.

18.11 Ellipsoïde de John-Loewer

Soit q une forme quadratique sur $\mathbb{R}^n$ ainsi que $\mathcal{B}$ une base orthonormée de $\mathbb{R}^n$ dans laquelle la matrice de q est diagonale. Dans cette base, la forme q est donnée par la proposition 9.246 :

$$q(x)=\sum_i \lambda_i x_i \tag{18.418}$$

où les λ_i sont les valeurs propres de q.

Plus généralement nous notons $mat_{\mathcal{B}}(q)$ la matrice de q dans la base $\mathcal{B}$ de $\mathbb{R}^n$.

Proposition 18.99.
Soit $\mathcal{B}$ une base orthonormée de $\mathbb{R}^n$ et l'application[46]

$$\begin{aligned} D\colon Q(\mathbb{R}^n) &\to \mathbb{R} \\ q &\mapsto \det\big(mat_{\mathcal{B}}(q)\big). \end{aligned} \tag{18.419}$$

Alors :

(1) La valeur et D ne dépend pas du choix de la base orthonormée $\mathcal{B}$.

(2) La fonction D est donnée par la formule $D(q)=\prod_i \lambda_i$ où les λ_i sont les valeurs propres de q.

(3) La fonction D est continue.

Démonstration. Soit q une forme quadratique sur $\mathbb{R}^n$. Nous considérons $\mathcal{B}$ une base de diagonalisation de q :

$$q(x)=\sum_i \lambda_i x_i \tag{18.420}$$

où les x_i sont les composantes de x dans la base $\mathcal{B}$. Par définition, la matrice $mat_{\mathcal{B}}(q)$ est la matrice diagonale contenant les valeurs propres de q.

Nous considérons aussi $\mathcal{B}_1$, une autre base orthonormées de $\mathbb{R}^n$. Nous notons $S=mat_{\mathcal{B}_1}(q)$; étant symétrique, cette matrice se diagonalise par une matrice orthogonale : il existe $P\in \mathrm{O}(n,\mathbb{R})$ telle que

$$S=Pmat_{\mathcal{B}}(q)P^t; \tag{18.421}$$

donc $\det(S)=\det(PP^t)\det\big(\mathrm{diag}(\lambda_1,\dots,\lambda_n)\big)=\lambda_1\dots\lambda_n$. Ceci prouve en même temps que D ne dépend pas du choix de la base et que sa valeur est le produit des valeurs propres.

Passons à la continuité. L'application déterminant $\det\colon S_n(\mathbb{R}^n)\to\mathbb{R}$ est continue car polynôme en les composantes. D'autre par l'application $mat_{\mathcal{B}}\colon Q(\mathbb{R}^n)\to S_n(\mathbb{R})$ est continue par la proposition 9.245. L'application D étant la composée de deux applications continues, elle est continue. □

46. L'ensemble $Q(E)$ est l'ensemble des formes quadratiques sur E.

Proposition 18.100 (Ellipsoïde de John-Loewner[2])**.**
Soit K compact dans $\mathbb{R}^n$ et d'intérieur non vide. Il existe une unique ellipsoïde[47] *(pleine) de volume minimal contenant K.*

Démonstration. Nous subdivisons la preuve en plusieurs parties.

À propos de volume d'un ellipsoïde Soit $\mathcal{E}$ un ellipsoïde. La proposition 9.258 et son corollaire 9.259 nous indiquent que

$$\mathcal{E} = \{x \in \mathbb{R}^n \text{ tel que } q(x) \leqslant 1\} \tag{18.422}$$

pour une certaine forme quadratique strictement définie positive q. De plus il existe une base orthonormée $\mathcal{B} = \{e_1, \dots, e_n\}$ de $\mathbb{R}^n$ telle que

$$q(x) = \sum_{i=1}^{n} a_i x_i^2 \tag{18.423}$$

où $x_i = \langle e_i, x \rangle$ et les a_i sont tous strictement positifs. Nous nommons $\mathcal{E}_q$ l'éllipsoïde associée à la forme quadratique q et V_q son volume que nous allons maintenant calculer[48] :

$$V_q = \int_{\sum_i a_i x_i^2 < 1} dx \tag{18.424}$$

Cette intégrale est écrite de façon plus simple en utilisant le C^1-difféomorphisme

$$\begin{aligned} \varphi\colon \mathcal{E}_q &\to B(0,1) \\ x &\mapsto \Big(x_1\sqrt{a_1}, \dots, x_n\sqrt{a_n}\Big). \end{aligned} \tag{18.425}$$

Le fait que φ prenne bien ses valeurs dans $B(0,1)$ est un simple calcul : si $x \in \mathcal{E}_q$, alors

$$\sum_i \varphi(x)_i^2 = \sum_i a_i x_i^2 < 1. \tag{18.426}$$

Cela nous permet d'utiliser le théorème de changement de variables 13.224 :

$$V_q = \int_{\sum_i a_i x_i^2 < 1} dx = \frac{1}{\sqrt{a_1 \dots a_n}} \int_{B(0,1)} dx. \tag{18.427}$$

La dernière intégrale est le volume de la sphère unité dans $\mathbb{R}^n$; elle n'a pas d'importance ici et nous la notons V_0. La proposition 18.99 nous permet d'écrire V_q sous la forme

$$V_q = \frac{V_0}{\sqrt{D(q)}}. \tag{18.428}$$

Existence de l'ellipsoïde Nous voulons trouver un ellipsoïde contenant K de volume minimal, c'est à dire une forme quadratique $q \in Q^{++}(\mathbb{R}^n)$ telle que
— $D(q)$ soit maximal
— $q(x) \leqslant 1$ pour tout $x \in K$.
Nous considérons l'ensemble des candidats semi-définis positifs.

$$A = \{q \in Q^+ \text{ tel que } q(x) \leqslant 1 \forall x \in K\}. \tag{18.429}$$

Nous allons montrer que A est convexe, compact et non vide dans $Q(\mathbb{R}^n)$; il aura ainsi un maximum de la fonction continue D définie sur $Q(\mathbb{R}^n)$. Nous montrerons ensuite que le maximum est dans Q^{++}. L'unicité sera prouvée à part.

47. Définition 9.256.
48. Le volume ne change pas si nous écrivons l'inégalité stricte au lieu de large dans le domaine d'intégration ; nous le faisons pour avoir un domaine ouvert.

Non vide L'ensemble K est compact et donc borné par $M > 0$. La forme quadratique $q\colon x \mapsto \|x\|^2/M^2$ est dans A parce que si $x \in K$ alors

$$q(x) = \frac{\|x\|^2}{M^2} \leqslant 1. \tag{18.430}$$

Convexe Soient $q, q' \in A$ et $\lambda \in [0,1]$. Nous avons encore $\lambda q + (1-\lambda)q' \in Q^+$ parce que

$$\lambda q(x) + (1-\lambda)q'(x) \geqslant 0 \tag{18.431}$$

dès que $q(x) \geqslant 0$ et $q'(x) \geqslant 0$. D'autre part si $x \in K$ nous avons

$$\lambda q(x) + (1-\lambda)q'(x) \leqslant \lambda + (1-\lambda) = 1. \tag{18.432}$$

Donc $\lambda q + (1-\lambda)q' \in A$.

Fermé Pour rappel, la topologie de $Q(\mathbb{R}^n)$ est celle de la norme (9.564). Nous considérons une suite (q_n) dans A convergeant vers $q \in Q(\mathbb{R}^n)$ et nous allons prouver que $q \in A$, de sorte que la caractérisation séquentielle de la fermeture (proposition 6.139) conclue que A est fermé. En nommant e_x le vecteur unitaire dans la direction x nous avons

$$\big|q(x)\big| = \big|\|x\|^2 q(e_x)\big| \leqslant \|x\|^2 N(q), \tag{18.433}$$

de sorte que notre histoire de suite convergente donne pour tout x :

$$\big|q_n(x) - q(x)\big| \leqslant \|x\|^2 N(q_n - q) \to 0. \tag{18.434}$$

Vu que $q_n(x) \geqslant 0$ pour tout n, nous devons aussi avoir $q(x) \geqslant 0$ et donc $q \in Q^+$ (semi-définie positive). De la même manière si $x \in K$ alors $q_n(x) \leqslant 1$ pour tout n et donc $q(x) \leqslant 1$. Par conséquent $q \in A$ et A est fermé.

Borné La partie K de $\mathbb{R}^n$ est borné et d'intérieur non vide, donc il existe $a \in K$ et $r > 0$ tel que $\overline{B(a,r)} \subset K$. Si par ailleurs $q \in A$ et $x \in \overline{B(0,r)}$ nous avons $a + x \in K$ et donc $q(a+x) \leqslant 1$. De plus $q(-a) = q(a) \leqslant 1$, donc

$$\sqrt{q(x)} = \sqrt{q\big(x + a - a\big)} \leqslant \sqrt{q(x+a)} + \sqrt{q(-a)} \leqslant 2 \tag{18.435}$$

par l'inégalité de Minkowski 9.254. Cela prouve que si $x \in \overline{B(0,r)}$ alors $q(x) \leqslant 4$. Si par contre $x \in \overline{B(0,1)}$ alors $rx \in \overline{B(0,r)}$ et

$$0 \leqslant q(x) = \frac{1}{r^2} q(rx) \leqslant \frac{4}{r^2}, \tag{18.436}$$

ce qui prouve que $N(q) \leqslant \frac{4}{r^2}$ et que A est borné.

L'ensemble A est compact parce que fermé et borné, théorème de Borel-Lebesgue 6.123. L'application continue $D\colon Q(\mathbb{R}^n) \to \mathbb{R}$ de la proposition 18.99 admet donc un maximum sur le compact A. Soit q_0 ce maximum.

Nous montrons que $q_0 \in Q^{++}(\mathbb{R}^d)$. Nous savons que l'application $f\colon x \mapsto \frac{\|x\|^2}{M^2}$ est dans A et que $D(f) > 0$. Vu que q_0 est maximale pour D, nous avons

$$D(q_0) \geqslant D(f) > 0. \tag{18.437}$$

Donc $q_0 \in Q^{++}$.

Unicité S'il existe une autre ellipsoïde de même volume que celle associée à la forme quadratique q_0, nous avons une forme quadratique $q \in Q^{++}$ telle que $q(x) \leqslant 1$ pour tout $x \in K$. C'est à dire que nous avons $q_0, q \in A$ tels que $D(q_0) = D(q)$.

Nous considérons la base canonique $\mathcal{B}_c$ de $\mathbb{R}^n$ et nous posons $S = mat_{\mathcal{B}_c}(q)$, $S_0 = mat_{\mathcal{B}_c}(q_0)$. Étant donné que A est convexe, $(q_0 + q)/2 \in A$ et nous allons prouver que cet élément de A contredit la maximalité de q_0. En effet

$$D\left(\frac{q+q_0}{2}\right) = \det\left(\frac{S+S_0}{2}\right) \tag{18.438}$$

Nous allons utiliser le lemme 18.87 qui dit que le logarithme est log-concave sous la forme de l'équation (18.337) avec $\alpha = \beta = \frac{1}{2}$:

$$D\left(\frac{q+q_0}{2}\right) = \det\left(\frac{S+S_0}{2}\right) > \sqrt{\det(S)}\sqrt{\det(S_0)} = \det(S_0) = D(q_0). \tag{18.439}$$

Nous avons utilisé le fait que $D(q_0) = D(q)$ qui signifie que $\det(S_0) = \det(S)$. L'inéquation (18.439) contredit la maximalité de $D(q_0)$ et donne donc l'unicité.

□

18.12 Formes quadratiques, signature, et lemme de Morse

Soit $(E, \|.\|_E)$ un espace vectoriel réel normé de dimension finie n. L'ensemble des formes quadratiques réelles [49] sur E est vu comme l'ensemble des matrices symétriques $S_n(\mathbb{R})$; il sera noté $Q(E)$ et le sous-ensemble des formes quadratiques non dégénérées est $S_n(\mathbb{R}) \cap \mathrm{GL}(n, \mathbb{R})$ qui sera noté $\Omega(E)$. Nous rappelons que la correspondance est donnée de la façon suivante. Si $A \in S_n(\mathbb{R})$, la forme quadratique associée est q_A donnée par $q_A(x) = x^t Ax$.

Nous noterons encore $Q^+(E)$ les formes quadratiques positives sur E et $Q^{++}(E)$ les formes quadratiques strictement définies positives sur E.

Sur $Q(E)$ nous mettons la norme

$$N(q) = \sup_{\|x\|_E=1} |q(x)|, \tag{18.440}$$

qui du point de vue de $S_n(\mathbb{R})$ est

$$N(A) = \sup_{\|x\|_E=1} |x^t Ax|. \tag{18.441}$$

Notons que à droite, c'est la valeur absolue usuelle sur $\mathbb{R}$.

Nous savons par le théorème de Sylvester (théorème 16.96) que dans $\mathbb{M}(n, \mathbb{R})$, toute matrice symétrique de signature (p, q) est semblable à la matrice

$$\mathbb{1}_{p,q} = \begin{pmatrix} \mathbb{1}_p & & \\ & \mathbb{1}_p & \\ & & 0_{n-p-q} \end{pmatrix}. \tag{18.442}$$

Donc deux matrices de S_n sont semblables si et seulement si elles ont la même signature (même si elles ne sont pas de rang maximum, cela soit dit au passage). Si nous notons $S_n^{p,q}(\mathbb{R})$ l'ensemble des matrices réelles symétriques de signature (p, q), alors

$$S_n^{p,q}(\mathbb{R}) = \{P^t AP \text{ tel que } P \in \mathrm{GL}(n, \mathbb{R})\} \tag{18.443}$$

où A est une quelconque ce ces matrices.

Nous voudrions en savoir plus sur ces ensembles. En particulier nous aimerions savoir si la signature est une notion «stable» au sens où ces ensembles seraient ouverts dans S_n. Pour cela nous considérons l'action de $\mathrm{GL}(n, \mathbb{R})$ sur S_n définie par

$$\begin{aligned} \alpha\colon \mathrm{GL}(n, \mathbb{R}) \times S_n(\mathbb{R}) &\to S_n(\mathbb{R}) \\ (P, A) &\mapsto P^t AP \end{aligned} \tag{18.444}$$

49. Définition 9.242.

faite exprès pour que les orbites de cette action soient les ensembles $S_n^{p,q}(\mathbb{R})$.

La proposition suivante montre que lorsque $p+q=n$, c'est à dire lorsqu'on parle de matrices de rang maximum, les ensembles $S_n^{p,q}(\mathbb{R})$ sont ouverts, c'est à dire que la signature d'une forme quadratique est une propriété «stable» par petite variations des éléments de matrice. Notons tout de suite que si le rang n'est pas maximum, le théorème de Sylvester dit qu'elle est semblable à une matrice diagonale avec des zéros sur la diagonale ; en modifiant un peu ces zéros, on peut modifier évidemment la signature.

Proposition 18.101 ([2])**.**
Soit $(E, \|.\|_E)$ un espace vectoriel normé de dimension finie. Alors

(1) les formes quadratiques non dégénérées forment un ouvert dans l'ensemble des formes quadratiques,

(2) les ensembles $S_n^{p,q}(\mathbb{R})$ avec $p+q=n$ sont ouverts dans $S_n(\mathbb{R})$,

(3) les composantes connexes de $\Omega(E)$ sont les $S_n^{p,q}(\mathbb{R})$ avec $p+q=n$,

(4) les $S_n^{p,q}(\mathbb{R})$ non dégénérés sont connexes par arc.

Démonstration. Cette preuve est donnée du point de vue des matrices. La différence entre le point (3) et (4) est que dans le premier nous prouvons la connexité de $S_n^{p,q}(\mathbb{R})$ à partir de la connexité de $\mathrm{GL}^+(n,\mathbb{R})$, tandis que dans le second nous prouvons la connexité par arc de $S_n^{p,q}(\mathbb{R})$ à partir de la connexité par arc de $\mathrm{GL}^+(n,\mathbb{R})$. Bien entendu le second implique le premier.

(1) Il s'agit simplement de remarquer que $Q(E) = S_n(\mathbb{R})$, que $\Omega(E) = S_n(\mathbb{R}) \cap \mathrm{GL}(n,\mathbb{R})$ et que le déterminant est une fonction continue sur $\mathbb{M}(n,\mathbb{R})$.

(2) Soit $A_0 \in S_n^{p,q}(\mathbb{R})$. Le théorème de Sylvester 16.96 nous donne une matrice inversible P telle que $P^t A_0 P = \mathbb{1}_{p,q}$. Nous allons montrer qu'il existe un voisinage $\mathcal{U}$ de $\mathbb{1}_{p,q}$ contenu dans $S_n^{p,q}(\mathbb{R})$. À partir de là, l'ensemble $(P^{-1})^t \mathcal{U} P^{-1}$ sera un voisinage de A_0 contenu dans $S_n^{p,q}(\mathbb{R})$.

Nous considérons les espaces vectoriels

$$F = \mathrm{Span}\{e_1, \dots, e_p\} \tag{18.445a}$$
$$G = \mathrm{Span}\{e_{p+1}, \dots, e_n\} \tag{18.445b}$$

La norme euclidienne $\|.\|_p$ sur F est équivalente à la norme $|.|_E$ par le théorème 10.3. Donc il existe une constante $k_1 > 0$ telle que pour tout $x \in F$,

$$\|x\|_p \geqslant k_1 \|x\|_E. \tag{18.446}$$

De la même façon sur G, il existe une constante $k_2 > 0$ telle que

$$\|x\|_q \geqslant k_2 \|x\|_E. \tag{18.447}$$

Si nous posons $k = \min\{k_1^2, k_2^2\}$, alors nous avons

$$\forall x \in F, \quad \|x\|_p^2 \geqslant k_1^2 \|x\|_E^2 \geqslant k\|x\|_E^2 \tag{18.448a}$$
$$\forall x \in G, \quad \|x\|_q^2 \geqslant k_2^2 \|x\|_E^2 \geqslant k\|x\|_E^2. \tag{18.448b}$$

Soit une matrice $A \in S_n(\mathbb{R})$ telle que $N(A - \mathbb{1}_{p,q}) < k$, c'est à dire que A est dans un voisinage de $\mathbb{1}_{p,q}$ pour la norme sur $S_n(\mathbb{R})$ donné par (18.441). Si x est non nul dans E, nous avons

$$\left|x^t(A - \mathbb{1}_{p,q})x\right| \leqslant N(\mathbb{1}_{p,q} - A)\|x\|^2 \leqslant k\|x\|^2. \tag{18.449}$$

En déballant la valeur absolue, cela signifie que

$$-k\|x\|_E^2 \leqslant x^t(A - \mathbb{1}_{p,q})x \leqslant k\|x\|^2. \tag{18.450}$$

Si $x \in F$, alors la première inéquation et (18.446) donnent

$$x^tAx \geqslant \|x\|_p^2 - k\|x\|_E^2 > 0 \tag{18.451}$$

Si $x \in G$, alors la seconde inéquation et (18.447) donnent

$$x^tAx \leqslant k\|x\|_E^2 - \|x\|_q^2 < 0. \tag{18.452}$$

Nous avons donc montré que $x \mapsto x^tAx$ est positive sur F et négative sur G, ce qui prouve que A est bien de signature (p,q) et appartient donc à $S_n^{p,q}(\mathbb{R})$. Autrement dit nous avons

$$B(\mathbb{1}_{p,q}, k) \subset S_n^{p,q}(\mathbb{R}). \tag{18.453}$$

(3) Cette partie de la preuve provient essentiellement de [242], et fonctionne pour tous les $S_n^{p,q}(\mathbb{R})$, même pour ceux qui ne sont pas de rang maximum.
Soit $A \in S_n^{p,q}(\mathbb{R})$. Nous savons que $\mathrm{GL}(n,\mathbb{R})$ a deux composantes connexes (proposition 12.335). Vu que l'application

$$\begin{aligned}\alpha\colon \mathrm{GL}(n,\mathbb{R}) &\to S_n\\ P &\mapsto P^tAP\end{aligned} \tag{18.454}$$

est continue, l'image d'un connexe de $\mathrm{GL}(n,\mathbb{R})$ par α est connexe (proposition 6.86). En particulier, $\alpha\big(\mathrm{GL}^{\pm}(n,\mathbb{R})\big)$ sont deux connexes et nous savons que $S_n^{p,q}(\mathbb{R})$ a au plus ces deux composantes connexes.
Notre but est maintenant de trouver une intersection entre $\alpha\big(\mathrm{GL}^{+}(n,\mathbb{R})\big)$ et $\alpha\big(\mathrm{GL}^{-}(n,\mathbb{R})\big)$ [50]. Soit par le théorème de Sylvester, soit par le théorème de diagonalisation des matrices symétriques réelles 9.165, il existe une matrice $P \in \mathrm{GL}(n,\mathbb{R})$ diagonalisant A. En suivant la remarque 9.166, et en notant Q la matrice obtenue à partir de P en changeant le signe de sa première ligne, nous avons

$$\alpha(Q) = Q^tAQ = P^tAP = \alpha(P). \tag{18.455}$$

Or si $P \in \mathrm{GL}^{+}(n,\mathbb{R})$, alors $Q \in \mathrm{GL}^{-}(n,\mathbb{R})$ et inversement. Donc nous avons trouvé une intersection entre $\alpha\big(\mathrm{GL}^{+}(n,\mathbb{R})\big)$ et $\alpha\big(\mathrm{GL}^{-}(n,\mathbb{R})\big)$.

(4) Soient A et B dans $S_n^{p,q}(\mathbb{R}) \cap \mathrm{GL}(n,\mathbb{R})$. Par le théorème de Sylvester, il existe P et Q dans $\mathrm{GL}(n,\mathbb{R})$ telles que $A = P^t\mathbb{1}_{p,q}P$ et $B = Q^t\mathbb{1}_{p,q}Q$. Par la remarque 9.166 nous pouvons choisir P et Q dans $\mathrm{GL}^{+}(n,\mathbb{R})$. Ce dernier groupe étant connexe par arc, il existe un chemin

$$\gamma\colon [0,1] \to \mathrm{GL}^{+}(n,\mathbb{R}) \tag{18.456}$$

tel que $\gamma(0) = P$ et $\gamma(1) = Q$. Alors le chemin

$$s \mapsto \gamma(s)^t\mathbb{1}_{p,q}\gamma(s) \tag{18.457}$$

est un chemin continu dans $S_n^{p,q}(\mathbb{R})$ joignant A à B.

□

Nous savons déjà de la proposition 18.101 que les ensembles $S_n^{p,q}(\mathbb{R})$ (pas spécialement de rang maximum) sont ouverts dans $S_n(\mathbb{R})$. Le lemme suivant nous donne une précision à ce sujet, dans le cas des matrices de rang maximum, en disant que la matrice qui donne la similitude entre A_0 et A est localement un C^1-difféomorphisme de A.

Lemme 18.102.
Soit $A_0 \in \Omega(\mathbb{R}^n) = S_n \cap \mathrm{GL}(n,\mathbb{R})$, une matrice symétrique inversible. Alors il existe un voisinage V de A_0 dans S_n et une application $\phi\colon V \to \mathrm{GL}(n,\mathbb{R})$ qui

50. À ce point, il me semble que [242] fait erreur parce que la matrice $-\mathbb{1}_n$ est de déterminant 1 lorsque n est pair. L'argument donné ici provient de [2]

(1) est de classe C^1,

(2) est telle que pour tout $A \in V$, $\varphi(A)^t A_0 \phi(A) = A$.

Démonstration. Nous considérons l'application

$$\begin{aligned} \varphi\colon \mathbb{M}(n,\mathbb{R}) &\to S_n \\ M &\mapsto M^t A_0 M. \end{aligned} \tag{18.458}$$

Étant donné que les composantes de $\varphi(M)$ sont des polynômes en les entrées de M, cette application est de classe C^1 – et même plus. Soit maintenant $H \in \mathbb{M}(n,\mathbb{R})$ et calculons $d\varphi_{\mathbb{1}}(H)$ par la formule (12.418) :

$$d\varphi_{\mathbb{1}}(H) = \frac{d}{dt}\Big[\varphi(\mathbb{1}+tH)\Big]_{t=0} \tag{18.459a}$$
$$= \frac{d}{dt}\Big[(\mathbb{1}+tH^t)A_0(\mathbb{1}+tH)\Big]_{t=0} \tag{18.459b}$$
$$= \frac{d}{dt}\Big[A_0 + tA_0H + tH^tA_0 + t^2H^tA_0H\Big]_{t=0} \tag{18.459c}$$
$$= A_0H + H^tA_0. \tag{18.459d}$$

Donc

$$d\varphi_{\mathbb{1}}(H) = (A_0H) + (A_0H)^t. \tag{18.460}$$

Par conséquent

$$\ker(d\varphi_{\mathbb{1}}) = \{H \in \mathbb{M}(n,\mathbb{R}) \text{ tel que } A_0H \text{ est antisymétrique}\}, \tag{18.461}$$

et si nous posons

$$F = \{H \in \mathbb{M}(n,\mathbb{R}) \text{ tel que } A_0H \text{ est symétrique}\} \tag{18.462}$$

nous avons

$$\mathbb{M}(n,\mathbb{R}) = F \oplus \ker(d\varphi_{\mathbb{1}}) \tag{18.463}$$

parce que toute matrice peur être décomposée de façon unique en partir symétrique et antisymétrique. De plus l'application

$$\begin{aligned} f\colon F &\to S_n \\ H &\mapsto A_0H \end{aligned} \tag{18.464}$$

est une bijection linéaire. D'abord $A_0H = 0$ implique $H = 0$ parce que A_0 est inversible, et ensuite si $X \in S_n$, alors $X = A_0A_0^{-1}X$, ce qui prouve que X est l'image par f de $A_0^{-1}X$ et donc que f est surjective.

Maintenant nous considérons la restriction $\psi = \varphi_{|_F}$, $\psi\colon F \to S_n$. Remarquons que $\mathbb{1} \in F$ parce que $A_0 \in S_n$. L'application $d\psi_{\mathbb{1}}$ est une bijection. En effet d'abord

$$d(\varphi_{|_F})_{\mathbb{1}} = (d\varphi_{\mathbb{1}})_{|_F}, \tag{18.465}$$

ce qui prouve que

$$\ker(d\psi_{\mathbb{1}}) = \ker(d\varphi_{\mathbb{1}}) \cap F = \{0\}, \tag{18.466}$$

ce qui prouve que $d\psi_{\mathbb{1}}$ est injective. Pour montrer que $d\psi_{\mathbb{1}}$ est surjective, il suffit de mentionner le fait que $\dim F = \dim S_n$ du fait que l'application (18.464) est une bijection linéaire.

Nous pouvons utiliser le théorème d'inversion locale (théorème 18.47) et conclure qu'il existe un voisinage ouvert U de $\mathbb{1}$ dans F tel que ψ soit un difféomorphisme C^1 entre U et $V = \psi(U)$. Vu que $\mathrm{GL}(n,\mathbb{R})$ est ouvert dans $\mathbb{M}(n,\mathbb{R})$, nous pouvons prendre $U \cap \mathrm{GL}(n,\mathbb{R})$ et donc supposer que $U \subset \mathrm{GL}(n,\mathbb{R})$.

Pour tout $A \in V$, il existe une unique $M \in U$ telle que $\psi(M) = A$, c'est à dire telle que $A = M^t A_0 M$. Cette matrice M est $\psi^{-1}(A)$ et est une matrice inversible. Bref, nous posons

$$\begin{aligned} \phi\colon V &\to \mathrm{GL}(n,\mathbb{R}) \\ A &\mapsto \psi^{-1}(A), \end{aligned} \tag{18.467}$$

et ce ϕ est de classe C^1 sur V parce que c'est ce que dit le théorème d'inversion locale. Cette application répond à la question parce que V est un voisinage de $\varphi(\mathbb{1}) = A_0$ et pour tout $A \in V$ nous avons

$$\phi(A)^t A_0 \phi(A) = \varphi^{-1}(A)^t A_0 \varphi^{-1}(A) = A. \tag{18.468}$$

□

18.12.1 Lemme de Morse

Lemme 18.103 (Lemme de Morse)**.**
Soit $f \in C^3(\mathcal{U},\mathbb{R})$ où $\mathcal{U}$ est un ouvert de $\mathbb{R}^n$ contenant 0. Nous supposons que $df_0 = 0$ et que $d^2 f_0$ est non dégénérée[51] *et de signature $(p, n-p)$. Alors il existe un C^1-difféomorphisme φ entre deux voisinages de 0 dans $\mathbb{R}^n$ tel que*

(1) $\varphi(0) = 0$,

(2) si $\varphi(x) = u$ alors

$$f(x) - f(0) = u_1^2 + \cdots + u_p^2 - u_{p+1}^2 - \ldots - u_n^2. \tag{18.469}$$

Une autre façon de dire est qu'il existe un C^1-difféomorphisme local ψ tel que

$$(f \circ \psi)(x) - f(0) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \ldots - x_n^2. \tag{18.470}$$

Démonstration. Nous allons noter Hf la matrice Hessienne de f, c'est à dire $Hf_a = d^2 f_a \in \mathcal{L}^{(2)}(\mathbb{R}^n,\mathbb{R})$. Écrivons la formule de Taylor avec reste intégral (proposition 12.272 avec $p = 0$ et $m = 2$) :

$$f(x) - f(0) = \underbrace{df_0(x)}_{=0} + \int_0^1 (1-t) \underbrace{d^2 f_{tx}(x,x)}_{x^t (Hf)_{tx} x = \langle Hf_{tx}x, x\rangle} dt = x^t Q(x) x \tag{18.471}$$

avec

$$Q(x) = \int_0^1 (1-t)(Hf)_{tx} dt \tag{18.472}$$

qui est une intégrale dans $\mathcal{L}^{(2)}(\mathbb{R}^n,\mathbb{R})$. Nous prouvons à présent que Q est de classe C^1 en utilisant le résultat de différentiabilité sous l'intégrale 14.73. Pour cela nous passons aux composantes (de la matrice) et nous considérons

$$\begin{aligned} h_{kl}\colon U \times [0,1] &\to \mathbb{R} \\ h_{kl}(x,t) &= (1-t)\frac{\partial^2 f}{\partial x_k \partial x_l}(tx). \end{aligned} \tag{18.473}$$

Étant donné que f est de classe C^3, la dérivée de h_{kl} par rapport à x_i ne pose pas de problèmes :

$$\frac{\partial h_{kl}}{\partial x_i} = t(t-1)\frac{\partial^3 f}{\partial x_i \partial x_k \partial x_l}(tx), \tag{18.474}$$

qui est encore continue à la fois en t et en x. La proposition 14.73 nous montre à présent que

$$Q_{kl}(x) = \int_0^1 (1-t) h_{kl}(tx) dt \tag{18.475}$$

51. En tant qu'application bilinéaire.

est une fonction C^1. Étant donné que les composantes de Q sont C^1, la fonction Q est également C^1.

Nous avons $Q(0) = \frac{1}{2}(Hf)_0 \in S_n \cap \mathrm{GL}(n,\mathbb{R})$, d'abord parce que f est C^2 (et donc la matrice hessienne est symétrique), ensuite par hypothèse d^2f_0 est non dégénérée.

À partir de là, le lemme 18.102 donne un voisinage V de $Q(0)$ dans S_n et une application ϕ de classe C^1

$$\phi\colon V \to \mathrm{GL}(n,\mathbb{R}) \tag{18.476}$$

telle que pour tout $A \in V$,

$$\phi(A)^tQ(0)\phi(A) = A. \tag{18.477}$$

Si on pose $M = \phi \circ Q$, et si x est dans un voisinage de zéro, Q étant continue nous avons $Q(x) \in V$ et donc

$$Q(x) = M(x)^tQ(0)M(x). \tag{18.478}$$

Notons que l'application $\mathbb{M}\colon \mathbb{R} \to \mathrm{GL}(n,\mathbb{R})$ est de classe C^1 parce que Q et ϕ le sont.

Nous avons

$$f(x) - f(0) = x^tQ(x)x = x^tM(x)^tQ(0)M(x)x = y(x)^tQ(0)y(x) \tag{18.479}$$

où $y(x) = M(x)x = (\phi \circ Q)(x)x$ est encore une fonction de classe C^1 parce que la multiplication est une application C^∞.

D'un autre côté le théorème de Sylvester 16.96 nous donne une matrice inversible P telle que

$$Q(0) = P^t \begin{pmatrix} \mathbb{1}_p & \\ & -\mathbb{1}_{n-p} \end{pmatrix} P. \tag{18.480}$$

Et nous posons enfin $u = \varphi(x) = Py(x)$ qui est toujours de classe C^1 et qui donne

$$\begin{aligned} f(x) - f(0) &= y^tQ(0)y && (18.481a)\\ &= y^tP^t \begin{pmatrix} \mathbb{1} & \\ & -\mathbb{1} \end{pmatrix} Py && (18.481b)\\ &= u^t \begin{pmatrix} \mathbb{1} & \\ & -\mathbb{1} \end{pmatrix} u && (18.481c)\\ &= u_1^2 + \cdots + u_p^2 - u_{p+1}^2 - \ldots - u_n^2. && (18.481d) \end{aligned}$$

Nous devons maintenant montrer que, quitte à réduire son domaine à un ouvert plus petit, φ est un C^1-difféomorphisme. Dans la chaine qui donne φ, seule l'application

$$\begin{aligned} g\colon U \subset \mathbb{R}^n &\to \mathbb{R}^n\\ x &\mapsto M(x)x \end{aligned} \tag{18.482}$$

est sujette à caution. Nous allons appliquer le théorème d'inversion locale. Nous savons que g est de classe C^1 et donc différentiable ; calculons la différentielle en utilisant la formule (12.418) :

$$dg_0(x) = \frac{d}{dt}\Big[g(tx)\Big]_{t=0} = \frac{d}{dt}\Big[tM(tx)x\Big]_{t=0} = M(0)x. \tag{18.483}$$

Note que nous avons utilisé la règle de Leibnitz pour la dérivée d'un produit, mais le second terme s'est annulé. Donc $dg_0 = M(0) \in \mathrm{GL}(n,\mathbb{R})$ et g est localement un C^1-difféomorphisme.

Il suffit de restreindre φ au domaine sur lequel g est un C^1-difféomorphisme pour que φ devienne lui-même un C^1-difféomorphisme.

□

Définition 18.104.
*Un point a est un **point critique** de la fonction différentiable f si $df_a = 0$.*

Corollaire 18.105 ([243])**.**
Les points critiques non dégénérés d'une fonction C^3 sont isolés.

Démonstration. Soit a un point critique non dégénéré. Par le lemme de Morse 18.103, il existe un C^1-difféomorphisme ψ et un entier p tel que

$$(f \circ \psi)(x) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \ldots - x_n^2 + f(a) \tag{18.484}$$

sur un voisinage $\mathcal{U}$ de a. Vue la formule générale $df_x(u) = \nabla f(x) \cdot u$, si x est un point critique de f, alors $\nabla f(x) = 0$. Dans notre cas, les points critiques de $f \circ \psi$ dans $\mathcal{U}$ doivent vérifier $x_i = 0$ pour tout i, et donc $x = a$.

Nous devons nous assurer que la fonction f elle-même n'a pas de points critiques dans $\mathcal{U}$. Pour cela nous utilisons la formule générale de dérivation de fonction composée :

$$\nabla(f \circ \psi)(x) = \sum_k \frac{\partial f}{\partial y_k}\big(g(x)\big)\nabla g_k(x). \tag{18.485}$$

Si $\psi(x)$ est une point critique de f, alors le membre de droite est le vecteur nul parce que tous les $\partial_k f\big(\psi(x)\big)$ sont nuls. Par conséquent le membre de gauche est également nul, et x est un point critique de $f \circ \psi$. Or nous venons de voir que $f \circ \psi$ n'a pas de points critiques dans $\mathcal{U}$.

Donc f n'a pas de points critiques dans un voisinage d'un point critique non dégénéré. □

18.13 Variétés

18.13.1 Introduction

Soit $f : S^2 \to \mathbb{R}$ une fonction définie sur la sphère usuelle $S^2 \subset \mathbb{R}^3$. Une question naturelle est d'estimer la régularité de f ; est-elle continue, dérivable, différentiable ? Il n'existe pas de dérivée directionnelle étant donné que le quotient différentiel

$$\frac{f(x + \epsilon u_1, y + \epsilon u_2) - f(x, y)}{\epsilon}$$

n'a pas de sens pour un point $(x + \epsilon u_1, y + \epsilon u_2)$ qui n'est pas –sauf valeurs particulières– dans la surface. Pour la même raison il n'est pas possible de parler de différentiabilité de cette manière. Comment faire, sans devoir étendre le domaine de définition de f à un voisinage de la sphère ? Une solution possible est de parler de la notion de variété.

Une variété est un objet qui ressemble, vu de près, à $\mathbb{R}^m$ pour un certain m. En d'autres termes, on imagine une variété comme un recollement de morceaux de $\mathbb{R}^m$ vivant dans un espace plus grand $\mathbb{R}^n$. Ces morceaux sont appelés des ouverts de carte, et l'application qui exprime la ressemblance à $\mathbb{R}^m$ est l'application de carte.

18.13.2 Définition et propriétés

Définition 18.106.
Soit $\varnothing \neq M \subset \mathbb{R}^n$, $1 \leqslant m < n$ et $k \geqslant 1$. M est une variété de classe C^k de dimension m *si pour tout $a \in M$, il existe un voisinage ouvert U de a dans $\mathbb{R}^n$, et un ouvert V de $\mathbb{R}^m$ tel que $U \cap M$ soit le graphe d'une fonction $f : V \subset \mathbb{R}^m \to \mathbb{R}^{n-m}$ de classe C^1, c'est-à-dire qu'il existe un réagencement des coordonnées $(x_{i_1}, \ldots, x_{i_m}, x_{i_{m+1}}, \ldots, x_{i_n})$ avec*

$$M \cap U = \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n \text{ tel que } (x_{i_1}, \ldots, x_{i_m}) \in V \quad \left\{ \begin{array}{rcl} x_{i_{m+1}} & = & f_1(x_{i_1}, \ldots, x_{i_m}) \\ \vdots & = & \vdots \\ x_{i_n} & = & f_{n-m}(x_{i_1}, \ldots, x_{i_m}) \end{array} \right. \right\}$$

où V est un voisinage ouvert de $(a_{i_1}, \ldots, a_{i_m}) \in \mathbb{R}^m$.

La littérature regorge de théorèmes qui proposent des conditions équivalentes à la définition d'une variété. Celle que nous allons le plus utiliser est la suivante

Proposition 18.107.
Soit $M \subset \mathbb{R}^n$ et $1 \leqslant m \leqslant n-1$. L'ensemble M est une variété si et seulement si $\forall a \in M$, il existe un voisinage ouvert $\mathcal{U}$ de a dans $\mathbb{R}^n$ et une application $F: W \subset \mathbb{R}^m \to \mathbb{R}^n$ où W est un ouvert tels que

(1) F est un homéomorphisme de W vers $M \cap \mathcal{U}$,

(2) $F \in C^1(W, \mathbb{R}^n)$,

(3) Le rang de $dF(w) \in L(\mathbb{R}^m, \mathbb{R}^n)$ est de rang maximum (c'est à dire m) en tout point $w \in W$.

Pour rappel, si $T: \mathbb{R}^m \to \mathbb{R}^n$ est une application linéaire, son rang[52] est la dimension de son image. Si A est la matrice d'une application linéaire, alors le rang de cette application linéaire est égal à la taille de la plus grande matrice carré de déterminant non nul contenue dans A[53].

La condition de rang maximum sert à éviter le genre de cas de la figure 18.4 qui représente l'image de l'ouvert $]-1, 1[$ par l'application $F(t) = (t^2, t^3)$.

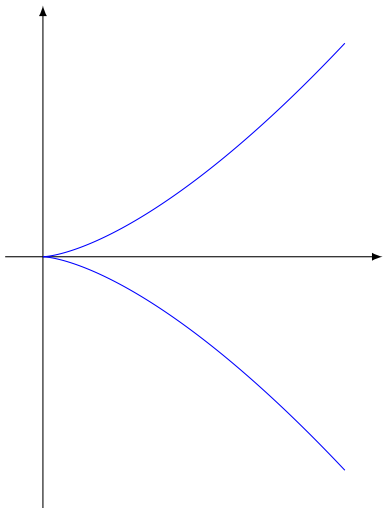


FIGURE 18.4 – Quelque chose qui n'est pas de rang maximum et qui n'est pas une variété.

La différentielle a pour matrice

$$dF(t) = (2t, 3t^2). \tag{18.486}$$

Le rang maximum est 1, mais en $t = 0$, la matrice vaut $(0, 0)$ et son rang est zéro. Pour toute autre valeur de t, c'est bon.

Une autre caractérisation des variétés est donnée par la proposition suivante

Proposition 18.108.
Soit $M \in \mathbb{R}^n$ et $1 \leqslant m \leqslant n-1$. L'ensemble M est une variété si et seulement si $\forall a \in M$, il existe un voisinage ouvert $\mathcal{U}$ de a dans $\mathbb{R}^n$ tel et une application $G \in C^1(\mathcal{U}, \mathbb{R}^{n-m})$ tel que

(1) le rang de $dG(a) \in L(\mathbb{R}^n, \mathbb{R}^{n-m})$ soit maximum (c'est à dire $n-m$) en tout $a \in M$,

(2) $M \cap \mathcal{U} = \{x \in \mathcal{U}$ tel que $G(x) = 0\}$.

18.13.3 Espace tangent

Soit M, une variété dans $\mathbb{R}^n$, et considérons un chemin $\gamma: I \to \mathbb{R}^n$ tel que $\gamma(t) \in M$ pour tout $t \in I$ et tel que $\gamma(0) = a$ et que γ est dérivable en 0. La **tangente** au chemin γ au point $a \in M$ est la droite

$$s \mapsto a + s\gamma'(0). \tag{18.487}$$

L'**espace tangent** de M au point a est l'ensemble décrit par toutes les tangentes en a pour tous les chemins γ possibles.

52. Définition 7.22.
53. Proposition 7.39

Proposition 18.109.
Une variété de dimension m dans $\mathbb{R}^n$ a un espace tangent de dimension m en chacun de ses points.

18.14 Prolongement de fonctions et complétion d'espaces métriques

Sources : [244]

Lemme 18.110.
Soit E, un espace vectoriel normé complet et (A_n) une suite emboité de fermés non vides dont le diamètre tend vers zéro. Alors l'intersection $\bigcap_{n\in\mathbb{N}} A_n$ contient exactement un point.

Démonstration. Si l'intersection contenait deux points distincts a et b, alors nous aurions $\operatorname{diam}(A_n) \geqslant \|a-b\|$, ce qui contredirait la limite.

Soit une suite (x_n) avec $x_k \in A_k$ pour tout $k \in \mathbb{N}$. C'est une suite de Cauchy. En effet si $\epsilon > 0$, considérons N tel que $\operatorname{diam}(A_N) < \epsilon$. Dans ce cas dès que $n, m > N$ nous avons $x_n, x_m \in A_N$ et donc $\|x_n - x_m\| \leqslant \epsilon$. La suite x_n converge donc vers un élément dans E.

Nous devons montrer que $x \in A_k$ pour tout k. La queue de suite $(x_n)_{n\geqslant k}$ est une suite de Cauchy dans A_k qui converge donc vers un élément de A_k (ici nous utilisons le fait que A_k est fermé). Par unicité de la limite, cette dernière doit être x. Par conséquent $x \in \bigcap_{n\in\mathbb{N}} A_n$. □

Théorème 18.111 ([245])**.**
Soient X et Y des espaces vectoriels normés. Pour une application linéaire $f\colon X \to Y$, les assertions suivantes sont équivalentes :

(1) f est continue sur X,

(2) f est continue en un point de X,

(3) f est bornée.

Proposition 18.112.
Soit un espace normé X, un espace de Banach F et une partie dense A de X. Si l'application linéaire

$$f\colon \big(A, \|.\|_X\big) \to F \tag{18.488}$$

est continue[54]*, alors il existe une unique application linéaire continue $\tilde{f}\colon X \to F$ prolongeant f. De plus $\|\tilde{f}\| = \|f\|$.*

Démonstration. Soit $x \in X$ et la suite d'ensemble

$$A_n = \{y \in A \text{ tel que } \|x-y\| \leqslant 2^{-n}\}. \tag{18.489}$$

Étant donné que A est dense, ces ensembles sont tous non vides. De plus $\operatorname{diam} A_n \to 0$ parce que si $y, y' \in A_n$ alors

$$\|y-y'\| \leqslant \|y-x\| + \|x-y'\| \leqslant 2^{-n+1}. \tag{18.490}$$

Vu que f est bornée, la suite d'ensembles $f(A_n)$ est une suite emboitée d'ensembles non vides de F. De plus leur diamètre tend vers zéro. En effet si $z, z' \in f(A_n)$, nous posons $z = f(y)$, $z' = f(y')$ et nous avons

$$\|z-z'\| \leqslant \|f(y)-f(x)\| + \|f(x)-f(y')\| \leqslant \|f\|\big(\|y-x\| + \|x-y'\|\big), \tag{18.491}$$

ce qui montre que $\operatorname{diam} f(A_n) \leqslant \|f\|2^{-n+1}$. Notons que nous avons utilisé la linéarité de f. Par le lemme 18.110, l'intersection $\bigcap_{n\in\mathbb{N}} \overline{f(A_n)}$ contient exactement un point. Nous posons

$$S(x) = \bigcap_{n\in\mathbb{N}} \overline{f(A_n)}. \tag{18.492}$$

54. Nous avons bien mis sur A la topologie induite de X. Notons que ce n'est pas toujours celle qui est la plus naturelle sur A.

Nous allons montrer que l'application $x \mapsto S(x)$ ainsi définie est l'application que nous cherchons.

Nous commençons par montrer que pour toute suite $y_k \to x$ avec $y_k \in A$ nous avons

$$f(y_k) \to S(x). \tag{18.493}$$

Pour cela nous considérons $n_0 \in \mathbb{N}$ et k_0 tel que $y_{k_0} \in A_{n_0}$. Avec cela nous avons

$$\|f(y_k) - S(x)\| \leqslant \operatorname{diam}(A_{n_0}) \leqslant \|f\|2^{-n_0+1}. \tag{18.494}$$

Pour montrer que S est linéaire, nous considérons deux suites dans A : $y_k \to x$ et $y'_k \to x'$ ainsi que la somme $y_k + y'k \to x + x'$. Nous écrivons la relation (18.493) pour ces trois suites :

$$f(y_k) \to S(x) \tag{18.495a}$$
$$f(y'_k) \to S(x') \tag{18.495b}$$
$$f(y_k + y'_x) \to S(x + x'). \tag{18.495c}$$

Cependant, étant donné que f est linéaire, pour tout k nous avons $f(y_k + y'_k) = f(y_k) + f(y'_k)$ et par conséquent

$$f(y_k + y'_k) \to S(x) + S(x'). \tag{18.496}$$

Par unicité de la limite, $S(x + x') = S(x) + S(x')$. Le même genre de raisonnement montre que $S(\lambda x) = \lambda S(x)$. L'application S est donc linéaire.

En ce qui concerna la continuité, nous avons

$$\|S(x)\| = \lim \|f(y_k)\| \leqslant \|f\| \lim y_k\| = \|f\|\|x\|, \tag{18.497}$$

donc $\|S\| \leqslant \|f\|$, c'est à dire que S est borné et donc continue parce que linéaire (théorème 18.111).

Nous montrons maintenant que S prolonge f. Si $x \in A$, alors nous avons $\bigcap_{n\in\mathbb{N}} f(A_n) = f(x)$, et donc $S(x) = f(x)$. Cela montre du même coup que $\|f\| \leqslant \|S\|$ et que par conséquent $\|f\| = \|S\|$.

Passons à la partie sur l'unicité. Soient donc S et T, deux prolongements continus de f sur X. Soit $x \in X$ et une suite $x_n \to x$ dans A. Par continuité nous avons $T(x_n) \to T(x)$ et $S(x_n) \to S(x)$. Étant donné que par ailleurs pour tout n nous avons $S(x_n) = T(x_n)$, l'unicité de la limite montre que $T(x) = S(x)$. □

Définition 18.113.
*Soit une application $f\colon X \to Y$. Le **module de continuité** de f est la fonction $\omega_f\colon \mathbb{R} \to \mathbb{R}$ définie comme suit. On pose $\omega_f(x) = 0$ pour $x \leqslant 0$ et si $h > 0$,*

$$\omega_f(h) = \sup_{\substack{x,y\in X\\ d_X(x,y)<h}} d_Y\big(f(x), f(y)\big). \tag{18.498}$$

Lemme 18.114.
Soit $f \in C^0\big([0,1], \mathbb{C}\big)$ et ω son module de continuité. Si λ et h sont strictement positifs avec $\lambda h \in [0,1]$ alors

$$\phi(\lambda h) \leqslant (\lambda + 1)\omega(h). \tag{18.499}$$

Démonstration. La fonction ω est décroissante, et pour $h, k > 0$ nous avons $\omega(h+k) \leqslant \omega(h)+\omega(k)$. Par récurrence pour tout $k \in \mathbb{N}$ nous avons

$$\omega(kh) \leqslant k\omega(h). \tag{18.500}$$

En écrivant cela pour $k = \lceil\lambda\rceil$, nous avons

$$\omega(\lambda h) \leqslant \omega(kh) \leqslant k\omega(h) \leqslant (\lambda + 1)\omega(h). \tag{18.501}$$

□

Lemme 18.115.
Une fonction f est uniformément continue si et seulement si son module de continuité est continu en zéro.

Dans la même veine que la proposition 18.112 nous avons ce résultat.

Théorème 18.116 ([246])**.**
Soient E et F, deux espaces métriques complets ainsi que A dense dans E. Si $u\colon A \to F$ est uniformément continue, alors elle se prolonge de façon unique en une fonction continue $\tilde{u}\colon E \to F$. De plus ce prolongement est uniformément continu.

Démonstration. Soit $x \in E\backslash A$ et une suite (x_n) contenue dans A et convergente vers x. Nous voulons définir

$$\tilde{u}(x) = \lim_{n\to\infty} u(x_n) \tag{18.502}$$

mais pour ce faire nous devons prouver que la suite $\big(u(x_n)\big)$ converge dans F et que la limite ne dépend pas de la suite choisie parmi les suites de A qui convergent (dans E) vers x.

Commençons par montrer que $\big(u(x_n)\big)$ est de Cauchy dans F. Pour cela nous prenons $\epsilon > 0$ et $\eta > 0$ telle que $d_E(a,b) < \eta$ implique $d_F\big(u(a),u(b)\big) < \epsilon$ (uniforme continuité de u). Après, il suffit de choisir N tel que pour tout $n, m > N$ nous ayons $d(x_m, x_n) < \eta$ (parce que u_n est de Cauchy). Avec tout ça nous avons

$$d_F\big(u(x_m),u(x_n)\big) < \epsilon, \tag{18.503}$$

ce qui signifie que $\big(u(x_n)\big)$ est de Cauchy et donc convergente dans F.

Nous voulons montrer maintenant que si (x_n) et (y_n) sont deux suites dans A convergentes vers x alors $\lim_{n\to\infty} u(x_n) = \lim_{n\to\infty} u(y_n)$. Pour cela nous considérons la suite $z = (x_1, y_1, x_2, y_2, \ldots)$. Nous avons évidemment $z_n \to x$, et donc $u(z_n)$ converge dans F par ce qui a été dit plus haut. Mais $u(x_n)$ et $u(y_n)$ en sont deux sous-suites convergentes. Donc leurs limites sont égales.

Il reste à montrer que ce $\tilde{u}$ est continue et uniformément continue. Pour cela nous utilisons le module de continuité et le lemme 18.115. Étant donné que $\tilde{u}$ prolonge u nous avons

$$\omega_{\tilde{u}}(h) \geqslant \omega_u(h). \tag{18.504}$$

Soit $h > 0$ et $\epsilon > 0$; soit aussi $x, y \in E$ tels que $d(x,y) < h$. Nous prenons des suites $(a_n) \to x$ et $(y_n) \to y$ tout en choisissant n assez grand pour avoir $d_E(a_n, b_n) < h$. Nous avons

$$d_F\big(\tilde{u}(x),\tilde{u}(y)\big) \leqslant d_F\big(\tilde{u}(x),u(a_n)\big) + d\big(u(a_n),u(b_n)\big) + d_F\big(u(b_n),\tilde{u}(y)\big). \tag{18.505}$$

Si n est assez grand, par construction de $\tilde{u}$, le premier et le dernier terme sont plus petits que ϵ. Par définition du module de continuité nous avons d'autre part $d_F\big(u(a_n),u(b_n)\big) \leqslant \omega_u(h)$. Du coup

$$d_F\big(\tilde{u}(x),\tilde{u}(y)\big) \leqslant \omega_u(h) + 2\epsilon. \tag{18.506}$$

Si nous prenons le supremum sur les x et y vérifiant $d_E(x,y) < h$, à gauche nous obtenons $\omega_{\tilde{u}}(h)$ tandis que le membre de droite ne dépend pas de x et$y$. Donc pour tout ϵ, nous avons

$$\omega_{\tilde{u}}(h) \leqslant \omega_u(h) + 2\epsilon. \tag{18.507}$$

En comparaison avec (18.504), nous trouvons

$$\omega_{\tilde{u}(h)} \leqslant \omega_u(h). \tag{18.508}$$

Les fonctions u et $\tilde{u}$ ayant le même module de continuité, le lemme 18.115 nous enseigne que l'une est uniformément continue si et seulement si l'autre l'est. Vu que u est uniformément continue par hypothèse, le prolongement $\tilde{u}$ est uniformément continu. □

Définition 18.117.
*Un **plongement** de l'espace topologique X dans Y est une application $f\colon X \to Y$ telle que $f\colon X \to f(X)$ soit un homéomorphisme.*

Théorème 18.118 (Extensiton des isométries)**.**
Soit $\tilde{M}$ un espace métrique complet et une application isométrique

$$f\colon A \to \tilde{M} \tag{18.509}$$

où A est une partie dense d'un espace métrique M (pas spécialement complet). Alors f accepte une une unique extension isométrique

$$\tilde{f}\colon M \to \tilde{M} \tag{18.510}$$

Supposons de plus que M soit complet[55]*. Alors $\tilde{f}\colon M \to \tilde{M}$ est une bijection si et seulement si $f(A)$ est dense dans $\tilde{M}$.*

Démonstration. Nous commençons par prouver l'unicité. Soit $\tilde{f}_1$ et $\tilde{f}_2$, deux extensions de f et $x \in M$. Si (a_n) est une suite dans A convergeant vers x (possible parce que A est dense dans M), alors nous avons

$$\tilde{f}_1(a_n) = \tilde{f}_2(a_n) \tag{18.511}$$

et donc $\tilde{f}_1(x) = \tilde{f}_2(x)$ par continuité (une application isométrique est continue (proposition 6.135)).

Nous démontrons à présent l'existence.

Construction de $\tilde{f}$ Soit $x \in M$ et (a_n) une suite dans A qui converge vers x. Nous définissons

$$\tilde{f}(x) = \lim_{k\to\infty} f(a_k). \tag{18.512}$$

Note : nous pouvons prouver que cette définition ne dépend pas du choix de la suite (a_n) convergeant vers x, mais ce serait superflu parce que nous avons déjà prouvé l'unicité de $\tilde{f}$. Par contre nous devons expliquer pourquoi la limite du membre de droite de (18.512) existe dans $\tilde{M}$. D'abord la suite (a_n) est de Cauchy parce qu'elle est convergente (attention : M n'étant pas complet le fait d'être de Cauchy n'implique pas la convergence). Donc, étant donné que f est une isométrie, la suite $\big(f(a_n)\big)$ est de Cauchy dans $\tilde{M}$. Or ce dernier étant complet, la suite des images converge.

Montrons que cette application $\tilde{f}\colon M \to \tilde{M}$ répond à la question.

$\tilde{f}$ est isométrique Soient $a, b \in M$ et des suites dans A convergeant vers eux : $a_n \to a$, $b_n \to b$.

Nous avons, par continuité de l'application distance,

$$\begin{aligned}
d\big(\tilde{f}(a), \tilde{f}(b)\big) &= \lim_{k\to\infty} d\big(\tilde{f}(a_k), \tilde{f}(b)\big) && \text{(18.513a)}\\
&= \lim_{k\to\infty}\lim_{l\to\infty} d\big(\tilde{f}(a_k), \tilde{f}(b_l)\big) && \text{(18.513b)}\\
&= \lim_{k\to\infty}\lim_{l\to\infty} d\big(f(a_k), f(b_l)\big) && \text{(18.513c)}\\
&= \lim_{k\to\infty}\lim_{l\to\infty} d\big(a_k), b_l\big) && \text{(18.513d)}\\
&= d(a, b). && \text{(18.513e)}
\end{aligned}$$

Cela prouve que $\tilde{f}$ est une isométrie.

Pour la suite nous supposons que M est complet. Notons tout de suite que $\tilde{f}$ est injective parce qu'elle est isométrique.

Bijection (premier sens) Nous supposons que $\tilde{f}\colon M \to \tilde{M}$ est une bijection. Par l'absurde nous supposons que $f(A)$ n'est pas dense dans $\tilde{M}$, c'est à dire que nous avons un point $x \in \tilde{M}$ et une boule n'intersectant par $f(A)$:

$$B(x, r) \cap f(A) = \varnothing. \tag{18.514}$$

Étant donné que $\tilde{f}$ a pour image des limites de suites dans $f(A)$, l'image de $\tilde{f}$ est contenue dans $\overline{f(A)}$. Donc si $\tilde{f}$ est surjective, c'est que $\tilde{M} \subset \overline{f(A)}$ et donc que $\overline{f(A)} = \tilde{M}$. Cela prouve que si $\tilde{f}$ est bijective, alors $f(A)$ est dense dans $\tilde{M}$.

55. Il me semble que cette hypothèse manque dans [74].

Bijection (l'autre sens) Nous supposons que $\overline{f(A)} = \tilde{M}$ et nous devons prouver que $\tilde{f}$ est surjective. Soit $x \in \tilde{M}$ et $f(a_n)$ une suite dans $f(A)$ qui converge vers x ; une telle suite existe parce que $f(A)$ est dense dans $\tilde{M}$. Cette suite est de Cauchy dans $\tilde{M}$ parce que dans un espace métrique, une suite convergente est de Cauchy. La suite (a_n) est elle-même également de Cauchy parce que

$$d(a_n, a_m) = d\big(f(a_n), f(a_m)\big). \tag{18.515}$$

Étant donné que (a_n) est de Cauchy dans M, elle converge vers un élément que nous nommons $a \in M$. Par continuité de f nous avons alors

$$\tilde{f}(a) = \lim_{k\to\infty} f(a_k) = x. \tag{18.516}$$

Cela prouve que x est bien dans l'image de $\tilde{f}$ et donc que $\tilde{f}$ est surjective.

□

Une conséquence du théorème de prolongement est le théorème suivant qui permet de compléter un espace métrique.

Théorème 18.119 (Complétion d'un espace métrique[247, 74])**.**
Tout espace métrique se plonge par une isométrie à image dense dans un espace métrique complet. De plus ce dernier est unique à isométrie près.

Plus précisément, soit (M, d) un espace métrique. Il existe un espace métrique complet $\tilde{M}$ muni d'un plongement isométrique $\varphi\colon M \to \tilde{M}$ tel que $\varphi(M)$ soit dense dans $\tilde{M}$.

Ce complété de M est unique au sens suivant. Si $\tilde{M}_1$ et $\tilde{M}_2$ sont deux espaces métriques complets munis de plongements isométriques $f_i\colon M \to \tilde{M}_1$ dont les images sont denses, alors il existe une bijection isométrique $\phi\colon \tilde{M}_1 \to \tilde{M}_2$ telle que $\phi \circ f_1 = f_2$.

Démonstration. Nous ne prouvons que l'existence.

Soit C_M l'ensemble des suites de Cauchy de M. Nous définissons

$$\begin{aligned} f\colon C_M \times C_M &\to \mathbb{R} \\ u, v &\mapsto \lim_{n\to\infty} d(u_n, v_n). \end{aligned} \tag{18.517}$$

Notre première tâche est de nous assurer que cela est bien défini, c'est à dire que la limite existe toujours. En effet, si u et v sont des suites de Cauchy dans M, nous avons

$$|d(u_n, v_n) - d(u_m, v_m)| \leqslant d(u_n, v_n) + d(u_m, v_m) \leqslant 2\epsilon \tag{18.518}$$

dès que m et n sont assez grand. Cela prouve que la suite $n \mapsto d(u_n, v_n)$ est de Cauchy dans $\mathbb{R}$. Par complétude de $\mathbb{R}$, elle converge [56].

Nous considérons la relation d'équivalence $u \sim v$ si et seulement si $f(u, v) = 0$. Nous posons $\tilde{M} = C_M/\sim$ et nous y mettons la distance

$$d([u], [v]) = f(u, v) \tag{18.519}$$

et nous devons encore vérifier que cela est bien défini. Prenons $u' \sim u$ et $v' \sim v$. Alors nous avons

$$d(u'_n, v'_n) \leqslant d(u'_n, u_n) + d(u_n, v_n) + d(v_n, v'_n), \tag{18.520}$$

et donc

$$d(u', v') = \lim_{n\to\infty} d(u'_n, v'n) \leqslant \lim_{n\to\infty} d(u_n, v_n) = d(u, v). \tag{18.521}$$

Le même argument en inversant les primes et les non primes montre l'inégalité inverse. Donc $d(u, v) = d(u', v')$ dans C_M, et donc la distance (18.519) est bien définie sur $\tilde{M}$.

Afin de s'assurer que $\tilde{M}$ répond bien à la question du théorème, il faut encore démontrer les points suivants :

56. Ici nous utilisons la complétude de $\mathbb{R}$. Cette dernière doit donc être démontrée indépendamment. De plus nous ne pouvons pas définir $\mathbb{R}$ comme étant le complété de $\mathbb{Q}$ en utilisant ce théorème.

— M se plonge isométriquement dans $\tilde{M}$.
— l'image de M par le plongement est dense dans $\tilde{M}$.
— $\tilde{M}$ est complet.

Nous allons maintenant considérer l'application

$$\begin{aligned}\varphi\colon M &\to \tilde{M}\\ x &\mapsto \text{la classe de la suite constante } x.\end{aligned} \tag{18.522}$$

Plongement isométrique Nous allons montrer que cela est une isométrie bijective et que $\varphi(M)$ est dense dans $\tilde{M}$. Le fait que φ soit bijective entre M et $\varphi(M)$ est évident. C'est une isométrie parce que

$$d\big(\varphi(x),\varphi(y)\big) = \lim_{n\to\infty} d\big(\varphi(x)_n,\varphi(y)_n\big) = d(x,y). \tag{18.523}$$

Densité Soit $[u]\in\tilde{M}$. Tous les termes u_n sont des éléments de M. Nous considérons la suite dans $\varphi(M)$ donnée par

$$a_n = \varphi(u_n) \tag{18.524}$$

Chaque a_n est un élément [57] de $\tilde{M}$. Montrons que (a_n) converge dans $\tilde{M}$ vers u. Nous avons

$$\begin{aligned}d(a_n,u) &= \lim_{k\to\infty} d\big((a_n)_k,u_k\big) && (18.525a)\\ &= \lim_{k\to\infty} d(u_n,u_k) && (18.525b)\\ &= d(u_n,\ell) && (18.525c)\end{aligned}$$

en notant ℓ la limite de la suite (u_n). Ici nous avons utilisé le fait que la fonction distance était continue pour l'inverser avec la limite, par le théorème 12.87. Nous avons alors

$$\lim_{n\to\infty} d(a_n,[u]) = \lim_{n\to\infty} d(u_n,\ell) = 0. \tag{18.526}$$

Complétude Nous passons maintenant à la preuve du fait que $\tilde{M}$ est complet. Soit (y_n) une suite de Cauchy dans $\tilde{M}$. Soit $\epsilon>0$; nous définissons $K(n)$ par

$$d\big((y_n)_k,(y_n)_l\big)<\epsilon \tag{18.527}$$

dès que $k,l\geqslant K(n)$. Cette définition fonctionne parce que pour chaque n, y_n est une suite de Cauchy dans M. Nous posons

$$x_n = (y_n)_{K(n)}\in M \tag{18.528}$$

et nous allons montrer que (x_n) est de Cauchy dans M –donc est un élément de $\tilde{M}$– et que $y_k\to(x_n)$ dans $\tilde{M}$.

Nous commençons par montrer que (x_n) est de Cauchy dans M. Nous avons

$$\begin{aligned}d(x_n,x_m) &= d\big((y_n)_{K(n)},(y_m)_{K(m)}\big) && (18.529a)\\ &\leqslant d\big((y_n)_{K(n)},(y_n)_l\big)+d\big((y_n)_l,(y_m)_l\big)+d\big((y_m)_l,(y_m)_{K(m)}\big) && (18.529b)\end{aligned}$$

Nous choisissons n,m tels que $d(y_n,y_m)<\epsilon$, ce qui nous permet de choisir l de telle façon à avoir $d\big((y_n)_k,(y_m)_k\big)<\epsilon$ pour tout $k\geqslant l$. De plus, quitte à encore augmenter l, nous supposons que $l>K(m)$ et $l>K(m)$. Avec ces choix nous voyons que $d(x_n,x_m)<3\epsilon$, ce qui signifie que la suite (x_n) est de Cauchy dans M.

En ce qui concerne la convergence $y_n\to(x)$, on a

$$\begin{aligned}d\big(y_n,(x)\big) &= \lim_{k\to\infty} d\big((y_n)_k,(y_k)_{K(k)}\big) && (18.530a)\\ &\leqslant \lim_{k\to\infty} d\big((y_n)_k,(y_n)_l\big)+\lim_{k\to\infty} d\big((y_n)_l,(y_k)_l\big)+\lim_{k\to\infty} d\big((y_k)_l,(y_k)_{K(k)}\big) && (18.530b)\end{aligned}$$

Nous devons trouver un n tel que si k est suffisamment grand, le tout est majoré par ϵ. Voici nos choix :

57. À partir de maintenant nous n'écrivons plus explicitement la classe d'équivalence.

— n tel que $d(y_n, y_m) < \epsilon$ dès que $m \geqslant n$,
— $k > n$,
— $k > K(n)$,
— $l > k$,
— $l > K(k)$,
— l suffisamment grand pour que $d\big((y_n)_l, (y_k)_l\big) < \epsilon$.
Avec tous ces choix, les trois termes de (18.530b) sont plus petits que ϵ.
Ceci prouve que $\tilde{M}$ est complet.

□

Théorème 18.120 (Principe du prolongement analytique)**.**
Soit U un ouvert connexe. Si deux fonctions analytiques coïncident sur un sous-ensemble D de U contenant un point d'accumulation dans U, alors elles sont égales sur U.

18.15 Un petit extra

Soit f une fonction de $\mathbb{R}$ dans $\mathbb{R}$. Supposons que

(1) $f(1) = 1$,

(2) $f(x + y) = f(x) + f(y)$ pour tout réels x et y.

Nous pouvons montrer [58] que la seule fonction *continue* qui possède ces propriétés est la fonction identité $f(x) = x$ pour tout $x \in \mathbb{R}$.

De la même manière, il est aisé de voir que les seules applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^n$ sont de la forme

$$f(x) = Ax \tag{18.531}$$

pour une constante réelle A. Une question naturelle qu'on peut alors se poser est la suivante :

Est-il possible de définir une fonction non continue ayant les propriétés (1) et (2) ?

En fait, il est possible de démontrer que si E est un espace vectoriel de dimension finie, alors toute application linéaire $f : E \to F$ (où F est un espace vectoriel) sera continue. Ceci ne reste plus vrai si l'espace vectoriel E est de dimension infinie. Donc une manière de trouver une réponse positive à la question posée plus haut, serait de voir $\mathbb{R}^n$ comme espace vectoriel de dimension infinie. Après un peu de réflexion, la réponse est venue à nous (merci à Nicolas et à Samuel).

Si nous admettons l'axiome du choix, alors nous pouvons appliquer le théorème de Zorn et nous savons que tout espace vectoriel admet une base. En particulier, l'ensemble des réels vu comme espace vectoriel sur $\mathbb{Q}$ admet une base, i.e. $\exists (e_i)_{i\in I}$ des éléments de $\mathbb{R}^n$ tels que tout réel s'écrit comme combinaison linéaire à coefficients rationnels de ces e_i, i.e.

$$\forall r \in \mathbb{R}^n, \exists (\lambda_i)_{i\in I} \text{ des éléments de } \mathbb{Q} \text{ tels que } r = \sum_{i\in I} \lambda_i e_i. \tag{18.532}$$

Utilisons cette base pour définir une fonction h de la manière suivante.

$$\forall i \in I, \text{ on définit } h(e_i) = \alpha_i \tag{18.533}$$

où les α_i doivent être bien choisis dans $\mathbb{R}^n$. Pour satisfaire la propriété (1), choisissons sans perte de généralité $e_1 = 1$ et $h(e_1) = 1$. Ajoutons à cette propriété la linéarité en imposant que

$$h(\sum \lambda_i e_i) = \sum \lambda_i \alpha_i. \tag{18.534}$$

Les équations (1) et (2) nous permettent de voir que, moyennant le choix des α_i, la fonction h est bien définie sur $\mathbb{R}^n$ et linéaire. Il est clair que si nous prenons par exemple

$$\alpha_i = e_i \ \forall i \in I$$

58. et toi, tu le peux ?

nous obtenons que la fonction h est en fait la fonction identité sur $\mathbb{R}^n$. Par contre, si nous définissons la fonction h comme satisfaisant la propriété (2) et si nous choisissons les α_i dans (1) de la manière suivante

$$\begin{aligned} h(e_1) &= e_2 \\ h(e_2) &= e_1 \\ h(e_i) &= e_i \quad \forall i \in I\backslash\{1,2\} \end{aligned} \tag{18.535}$$

alors la fonction ainsi obtenue est linéaire et bien définie mais n'est plus l'identité. Donc nous avons trouvé une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^n$ qui n'est pas continue.

Exercice 1 Trouver d'autres exemples d'applications linéaires non continues (pas nécessairement des transformations de $\mathbb{R}^n$).

Chapitre 19

Arc paramétré

La structure de ce chapitre, comme beaucoup de choses dans le Frido, est fortement liée au choix de présenter toutes les matières dans l'ordre mathématiquement logique. Nous devons donc le placer après la trigonométrie ; les propriétés principales des fonctions trigonométriques étant dans la proposition 16.28, et c'est la proposition 16.29 qui nous permet de dire que $\big(\cos(t), \sin(t)\big)$ décrit le cercle.

Et enfin nous n'avons pas encore calculé la circonférence du cercle, et pour cause : nous n'avons pas encore donné de définition à la longueur d'un chemin dans $\mathbb{R}^2$. C'est pourquoi ce chapitre va aller droit à la longueur avant de donner des exemples.

19.1 Définitions

Définition 19.1.
*Un **arc paramétré** dans $\mathbb{R}^p$ est un couple (I, γ) où I est un intervalle de $\mathbb{R}$ et γ est une application continue de I dans $\mathbb{R}^p$. Nous disons que (I, γ) est un arc paramétré **compact** (ou un **chemin** dans $\mathbb{R}^p$) lorsque I est compact dans $\mathbb{R}$.*

L'intervalle I d'un arc paramétré compact est toujours de la forme $[a, b]$, étant donné que tous les intervalles compacts de $\mathbb{R}$ sont de cette forme. Un **sous arc** de (I, γ) est un arc de la forme (I_0, γ) avec $I_0 \subset I$.

Définition 19.2.
*Un **chemin** dans $\mathbb{R}$ est une application continue*

$$\begin{aligned} \sigma\colon [a,b] &\to \mathbb{R}^3 \\ t &\mapsto \sigma(t). \end{aligned} \tag{19.1}$$

La fonction $\sigma'(t)$ est la **vitesse** du chemin σ. Si la fonction $t \mapsto \sigma(t)$ est dérivable, on dit que $\sigma''(t)$ est l'**accélération**. Les points $\sigma(a)$ et $\sigma(b)$ sont les extrémités du chemin. L'ensemble

$$\{\sigma(t) \text{ tel que } t \in [a,b]\} \tag{19.2}$$

est la **courbe** σ.

19.2 Longueur d'arc

Nous voulons définir et étudier la notion de longueur d'un arc paramétré. Pour cela, le plus raisonnable est d'approcher l'arc par des petits segments de droites (dont les longueurs sont évidentes), et d'extraire la «meilleure» approximation.

Une des notions clefs pour la suite est celle de subdivision d'intervalles. Cette notion sera encore utilisée par la suite à propos des intégrales.

Définition 19.3.
Si I est un intervalle d'extrêmes a et b avec $a < b$, nous appelons ***subdivision finie*** *de I un choix de nombres t_i tels que*

$$a = t_0 < t_1 < \ldots < t_n = b. \tag{19.3}$$

Nous disons qu'une subdivision σ' est ***plus fine*** *que la subdivision σ si l'ensemble des points de σ est inclus à celui des points de σ'. Dans ce cas, la subdivision σ' est un* ***raffinement*** *de σ. Nous désignons par $\mathcal{S}(I)$ l'ensemble des subdivisions finies de l'intervalle I.*

Dans la suite, toutes les subdivisions que nous considérons seront des subdivisions finies. Aussi nous parlerons simplement de *subdivisions* sans préciser. Nous allons souvent noter $\sigma = (t_i)_{i=1}^n$ pour désigner la subdivision formée par les nombres t_i. Il faut garder en tête que dans une subdivision, les nombres *sont ordonnés*.

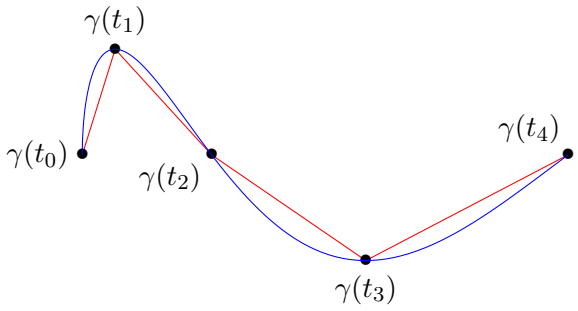


FIGURE 19.1 – La longueur d'un découpage. La somme des longueurs des segments droits est facile à calculer.

Définition 19.4.
Soit un arc paramétré compact (I, γ) et une subdivision $\sigma = (t_i)_{i=1}^n$ de $I = [a, b]$. À partir de γ et du découpage σ nous définissons le nombre (voir figure 19.1)

$$l_\sigma(\gamma) = \sum_{i=1}^{n} \big\|\gamma(t_i) - \gamma(t_{i-1})\big\|. \tag{19.4}$$

On appelle ***longueur*** *de l'arc γ le nombre*

$$l(\gamma) = \sup_\sigma l_\sigma \in [0, \infty]. \tag{19.5}$$

Nous disons que γ est ***rectifiable*** *lorsque $l(\gamma) < \infty$.*

Lorsque nous voulons spécifier sur quel intervalle nous considérons l'arc, nous noterons $l(I, \gamma)$ au lieu de $l(\gamma)$ pour être plus précis.

Par l'inégalité triangulaire, si σ_1 est plus fine que σ, nous avons

$$l_\sigma(\gamma) \leqslant l_{\sigma_1}(\gamma), \tag{19.6}$$

Comme cela peut être vu sur la figure 19.2.

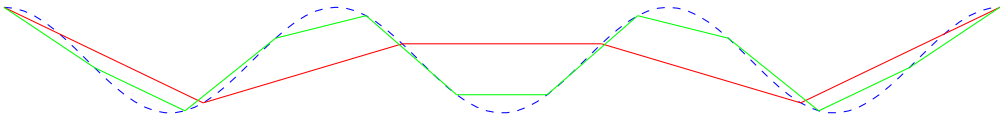


FIGURE 19.2 – Il est visible que la longueur donnée par l'approximation par des petits segments (verts) est plus longue et plus précise que celle donnée par les longs segments (rouge).

Dans la vie réelle, il est souvent difficile et peu pratique de calculer le supremum «à la main». C'est pourquoi nous allons travailler à exprimer la longueur d'un arc à l'aide d'une intégrale (théorème 19.9).

Lemme 19.5.
Nous avons $l(\gamma) = 0$ si et seulement si $\gamma(t)$ est un vecteur constant.

Démonstration. Si l'application $\gamma(t)$ est constante, le résultat est évident. Supposons maintenant que γ ne soit pas constante. Cela signifie qu'il existe t_1 et t_2 dans I tels que $\gamma(t_1) \neq \gamma(t_2)$. Dans ce cas, si nous prenons le découpage $\sigma = \{a, t_1, t_2, b\}$, la somme (19.4) contient au moins le terme non nul $\|\gamma(t_2) - \gamma(t_1)\|$, et donc $l_\sigma(\gamma) > 0$. Par définition du supremum, nous avons alors $l(\gamma) \geqslant l_\sigma(\gamma) > 0$. □

Proposition 19.6.
Soit (I, γ) un arc paramétré compact.

(1) Si $\gamma' = (I', \gamma)$ avec $I' \subset I$, alors $l(\gamma') \leqslant l(\gamma)$.

(2) Soit $c \in [a, b]$, et considérons les arcs $\gamma_1 = \big([a, c], \gamma\big)$ et $\gamma_2 = \big([c, b], \gamma\big)$. Alors

$$l(\gamma) = l(\gamma_1) + l(\gamma_2). \tag{19.7}$$

En particulier, γ est rectifiable si et seulement si γ_1 et γ_2 le sont.

Démonstration. (1) Nous notons $I = [a, b]$ et $I' = [a', b']$. Étant donné que $I' \subset I$, nous avons

$$a \leqslant a' < b' \leqslant b. \tag{19.8}$$

Pour chaque subdivision $\sigma_0 : a' = t_0 < t_1 < \ldots < t_n = b'$ de I', nous pouvons construire une subdivision de I en «ajoutant» les points a et b, c'est à dire

$$\sigma : a \leqslant t_0 < \ldots < t_n \leqslant b. \tag{19.9}$$

Si nous calculons $l_\sigma(\gamma)$, nous avons tous les termes qui arrivent dans $l_{\sigma_0}(\gamma')$ plus le premier et dernier terme : $\|\gamma(t_0) - \gamma(a)\|$ et $\|\gamma(b) - \gamma(t_n)\|$. Nous avons donc

$$l_{\sigma_0}(\gamma') \leqslant l_\sigma(\gamma) \leqslant \sup_\sigma l_\sigma(\gamma) = l(\gamma). \tag{19.10}$$

Étant donné que pour toute subdivision σ_0 nous avons $l_{\sigma_0}(\gamma') \leqslant l(\gamma)$, en prenant le supremum sur les subdivision σ_0 de I', nous avons comme annoncé

$$l(\gamma') \leqslant l(\gamma). \tag{19.11}$$

(2) Soit $\sigma = \{t_i\}$ une subdivision de $[a, b]$. Nous considérons les subdivision σ_1 et σ_2 définies comme suit :

$$\begin{aligned} \sigma_1 &: \{t_i \text{ tel que } t_i < c\} \cup \{c\}, \\ \sigma_2 &: \{t_i \text{ tel que } t_i > c\} \cup \{c\}. \end{aligned} \tag{19.12}$$

L'inégalité triangulaire implique que

$$l_\sigma(\gamma) \leqslant l_{\sigma\cup\{c\}}(\gamma) = l_{\sigma_1}(\gamma_1) + l_{\sigma_2}(\gamma_2) \leqslant l(\gamma_1) + l(\gamma_2). \tag{19.13}$$

Nous avons donc

$$l(\gamma) \leqslant l(\gamma_1) + l(\gamma_2). \tag{19.14}$$

Nous prouvons maintenant l'inégalité inverse. Soit $\varepsilon > 0$. Étant donné que $l(\gamma_1)$ est le supremum des quantités $l_{\sigma_1}(\gamma_1)$ lorsque σ_1 parcours toutes les subdivision possibles, il existe une partition σ_1^ε telle que (idem pour γ_2)

$$\begin{aligned} l_{\sigma_1^\varepsilon}(\gamma_1) + \frac{\varepsilon}{2} &> l(\gamma_1), \\ l_{\sigma_2^\varepsilon}(\gamma_2) + \frac{\varepsilon}{2} &> l(\gamma_2), \end{aligned} \tag{19.15}$$

où σ_1^ε est une subdivision de $[a,c]$ et σ_2^ε en est une de $[c,b]$. En faisant la somme des deux équations (19.15), nous trouvons

$$l(\gamma_1)+l(\gamma_2)<l_{\sigma_1^\varepsilon}(\gamma_1)+l_{\sigma_2^\varepsilon}(\gamma_2)+\varepsilon=l_{\sigma_1^\varepsilon\cup\sigma_2^\varepsilon}(\gamma)\leqslant l(\gamma)+\varepsilon. \tag{19.16}$$

L'inégalité $l(\gamma_1)+l(\gamma_2)<l(\gamma)+\varepsilon$ étant valable pour tout ε, nous avons

$$l(\gamma_1)+l(\gamma_2)\leqslant l(\gamma). \tag{19.17}$$

Cette inégalité, combinée avec l'inégalité (19.14), donne bien $l(\gamma)=l(\gamma_1)+l(\gamma_2)$.

□

19.3 Abscisse curviligne

Définition 19.7.
Soit (I,γ) un arc rectifiable compact avec $I=[a,b]$. L'application

$$\begin{aligned}\varphi\colon [a,b]&\to\mathbb{R}^+\\ t&\mapsto l\big([a,t],\gamma\big)\end{aligned} \tag{19.18}$$

*est la **longueur d'arc** de γ.*

Cette fonction nous permet de calculer la distance (suivant la courbe) entre deux point arbitraires parce que si $a\leqslant t<u\leqslant b$, nous avons

$$l\big([t,u],\gamma\big)=\varphi(u)-\varphi(t). \tag{19.19}$$

En effet,

$$\varphi(u)-\varphi(t)=l\big([a,u],\gamma\big)-l\big([a,t],\gamma\big), \tag{19.20}$$

mais en utilisant la proposition 19.6, nous avons

$$l\big([a,u],\gamma\big)=l\big([a,t],\gamma\big)+l\big([t,u],\gamma\big). \tag{19.21}$$

Proposition 19.8.
La longueur d'arc d'un arc rectifiable compact est une fonction continue et croissante.

Démonstration. Soit (I,γ) un arc paramétré rectifiable compact avec $I=[a,b]$. Afin de montrer que φ est croissante, prenons $t\in I$ ainsi que $h>0$ et montrons que $\varphi(t+h)\geqslant\varphi(t)$. La proposition 19.6 implique que

$$l\big([a,t+h],\gamma\big)=l\big([a,t],\gamma\big)+l\big([t,t+h],\gamma\big), \tag{19.22}$$

c'est à dire

$$\varphi(t+h)=\varphi(t)+l\big([t,t+h],\gamma\big)\geqslant\varphi(t). \tag{19.23}$$

Pour la continuité, soit t fixé dans $[a,b]$ et $\varepsilon>0$. Il nous faut démontrer qu'il existe $\eta>0$ tel que si s est dans $[0,\eta]$ alors

$$|\varphi(t+s)-\varphi(t)|\leqslant\varepsilon,\qquad\forall t\in[a,b].$$

Étant donné que $l\big([t,b],\gamma\big)$ est le supremum des $l_\sigma\big([t,b],\gamma\big)$, il existe une subdivision $\sigma:t,t_1,\cdots,t_{n-1},b$ telle que

$$l_\sigma\big([t,b],\gamma\big)>l\big([t,b],\gamma\big)-\frac{\varepsilon}{2}=\varphi(b)-\varphi(t)-\frac{\varepsilon}{2}. \tag{19.24}$$

La continuité de γ implique qu'il existe un η tel que

$$s\in[0,\eta]\Rightarrow\|\gamma(t+s)-\gamma(t)\|<\frac{\varepsilon}{2} \tag{19.25}$$

Quitte à prendre η encore plus petit, nous supposons que $t+\eta<t_1$. Soit $s\in[0,\eta]$ et considérons la subdivision de $[t,b]$ donnée par $\sigma'=\sigma\cup\{t+s\}$. Étant donné que σ' est plus fine que σ, le nombre $l_\sigma\big([t,b],\gamma\big)$ est inférieur ou égal à $l_{\sigma'}\big([t,b],\gamma\big)$. Nous avons donc les inégalités

$$\begin{aligned}\varphi(b)-\varphi(t)-\frac{\varepsilon}{2}&\leqslant l_\sigma\big([t,b],\gamma\big)\\&\leqslant l_{\sigma'}\big([t,b],\gamma\big)\\&=\big\|\gamma(t+s)-\gamma(t)\big\|+l_{\sigma'\setminus\{t\}}\big([t+s,b]\gamma\big)\\&\leqslant\|\gamma(t+s)-\gamma(t)\|+\varphi(b)-\varphi(t+s)\\&\leqslant\frac{\varepsilon}{2}+\varphi(b)-\varphi(t+s).\end{aligned}\tag{19.26}$$

Au final, nous avons trouvé que

$$\varphi(t+s)-\varphi(t)\leqslant\varepsilon,\tag{19.27}$$

ce qui prouve que φ est continue au point t. □

En guise de paramètre sur un arc, nous pouvons utiliser la longueur d'arc elle-même. En effet si (I,γ) est un arc de longueur l, nous pouvons donner le même arc avec le couple $\big([0,l],g\big)$ où g est la fonction qui au réel s fait correspondre l'élément $\gamma\big(\varphi^{-1}(s)\big)$ de $\mathbb{R}^n$. Dire

$$P=(\gamma\circ\varphi^{-1})(s)\tag{19.28}$$

revient à dire que le point P est le point sur la courbe sur lequel on tombe après avoir marché une distance s sur la courbe.

Nous allons revenir sur ce «changement de paramètre» plus tard, en particulier dans la section 19.7.

19.3.1 Formule intégrale de la longueur

Nous pouvons voir un chemin γ comme étant la trajectoire d'une particule en fonction du temps. Sa vitesse à l'instant t est le vecteur $\gamma'(t)$, tandis que sa vitesse *scalaire* est le nombre $\|\gamma'(t)\|$. Une question naturelle est de savoir quelle est la longueur de la trajectoire parcourue entre $t=a$ et $t=b$.

Si nous prenons un petit intervalle de temps dt, nous pouvons supposer que le mobile avance à la vitesse constante $\|\gamma'(t)\|$. Cela ferait un trajet parcouru de longueur $\|\gamma'(t)\|dt$. Nous nous attendons donc à une formule de la forme suivante pour la longueur de γ :

$$l(\gamma)=\int_a^b\|\gamma'(t)\|dt.\tag{19.29}$$

Plus explicitement, si $\gamma(t)=\big(x(t),y(t),z(t)\big)$, alors nous aurions la formule

$$l(\gamma)=\int_a^b\sqrt{x'(t)^2+y'(t)^2+z'(t)^2}dt.\tag{19.30}$$

Théorème 19.9.
Soit (I,γ) un arc paramétré compact de classe $\mathcal{C}^1$. Alors γ est rectifiable et

$$l(\gamma)=\int_a^b\|\gamma'(t)\|dt=\int_\gamma 1,\tag{19.31}$$

où $I=[a,b]$.

Démonstration. L'égalité avec l'intégrale le long de γ de la fonction 1 est simplement la définition 14.17 de l'intégrale curviligne.

Si $\sigma = \{t_i\}$ est une subdivision de l'intervalle $[a, b]$, alors

$$\begin{aligned} l_\sigma(\gamma) &= \sum_{i=1}^{n} \|\gamma(t_i) - \gamma(t_{i-1})\| \\ &= \sum_{i=1}^{n} \| \int_{t_{i-1}}^{t_i} \gamma'(t)dt\| \\ &\leqslant \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \|\gamma'(t)\|dt \\ &= \int_a^b \|\gamma'(t)\|dt. \end{aligned} \tag{19.32}$$

Cela prouve déjà que

$$l(\gamma) = \sup_\sigma l_\sigma(\gamma) \leqslant \int_a^b \|\gamma'(t)\|dt. \tag{19.33}$$

Nous devons maintenant prouver l'inégalité inverse.

Notons φ l'abscisse curviligne $\varphi(t) = l\big([a, t], \gamma\big)$. Cette dernière vérifie

$$\varphi(t + h) - \varphi(t) = l\big([t, t + h], \gamma\big) \geqslant \|\gamma(t + h) - \gamma(t)\|, \tag{19.34}$$

et en particulier

$$\left\| \frac{\gamma(t + h) - \gamma(t)}{h} \right\| \leqslant \frac{\varphi(t + h) - \varphi(t)}{h}. \tag{19.35}$$

D'autre part, en utilisant (19.33) sur le segment $[t, t + h]$, nous avons

$$\varphi(t + h) - \varphi(t) = l\big([t, t + h], \gamma\big) \leqslant \int_t^{t+h} \|\gamma'(u)\|du. \tag{19.36}$$

Cela nous permet de continuer l'inéquation (19.35) en

$$\left\| \frac{\gamma(t + h) - \gamma(t)}{h} \right\| \leqslant \frac{\varphi(t + h) - \varphi(t)}{h} \leqslant \frac{1}{h} \int_t^{t+h} \|\gamma'(u)\|du. \tag{19.37}$$

Prenons la limite $h \to 0$. À gauche nous reconnaissons la formule de la dérivée, et nous obtenons $\|\gamma'(t)\|$; au centre nous avons $\varphi'(t)$ et à droite, si $n(u)$ représente une primitive de la fonction $u \mapsto \|\gamma'(u)\|$,

$$\lim_{h\to 0} \frac{n(t + h) - n(t)}{h} = n'(t) = \|\gamma'(t)\|. \tag{19.38}$$

Au final,

$$\|\gamma'(t)\| \leqslant \varphi'(t) \leqslant \|\gamma'(t)\|, \tag{19.39}$$

c'est à dire $\varphi'(t) = \|\gamma'(t)\|$ et donc par le théorème fondamental du calcul intégral 13.205,

$$\varphi(t) - \varphi(a) = \int_a^t \|\gamma'(u)\|du. \tag{19.40}$$

Par construction de la longueur d'arc, $\varphi(a) = 0$ et en posant $t = b$ nous obtenons la relation recherchée :

$$l(\gamma) = \varphi(b) = \int_a^b \|\gamma'(u)\|du. \tag{19.41}$$

□

Remarque 19.10.
Cela est cohérent avec 14.178, mais il faut garder en tête que $l(\gamma)$ n'est pas le mesure de Lebesgue de l'image de γ dans $\mathbb{R}^2$. Cette dernière est nulle.

Exemple 19.11
Soient donc a et b deux points de $\mathbb{R}^m$, et γ la droite joignant a à b, c'est à dire

$$\gamma(t) = (1-t)a + tb \tag{19.42}$$

avec $t \in [0,1]$. Le théorème 19.9 nous enseigne que la longueur de ce chemin est

$$l\big([0,1],\gamma\big) = \int_0^1 \|\gamma'(t)\|dt = \int_0^1 \| -a+b\| = \|b-a\|, \tag{19.43}$$

qui est bien la distance entre a et b. △

Exemple 19.12(Circonférence du cercle)
Nous savons que l'image de

$$\begin{aligned}\gamma\colon [0,2\pi[&\to \mathbb{R}^2\\ t &\mapsto \big(R\cos(t), R\sin(t)\big)\end{aligned} \tag{19.44}$$

est le cercle de centre $(0,0)$ et de rayon $R>0$. Et de plus cet arc est de classe C^1 (et même C^∞) par la proposition 16.19. La longueur sera, d'après la formule (19.9)

$$l_\gamma = \int_0^{2\pi} \|\gamma'(t)\|dt = 2\pi R \tag{19.45}$$

grâce à la formule $\sin^2 + \cos^2 = 1$ du lemme 16.21.

Mais tout cela n'est pas satisfaisant parce que nous n'avons pas encore de valeur numérique de π.

Il y a une autre façon de faire en considérant le quart de cercle dont la longueur en fonction de π est vite calculée par

$$l_\gamma = \int_0^{pi/2} \|\gamma'(t)\|dt = \frac{\pi R}{2}. \tag{19.46}$$

Cette même longueur est calculée en termes de fonctions plus courantes avec le chemin

$$\begin{aligned}\sigma\colon]0,1[&\to \mathbb{R}^2\\ t &\mapsto \left(\frac{t}{\sqrt{R^2-t^2}}\right)\end{aligned} \tag{19.47}$$

La longueur s'exprime avec

$$l_\sigma = \int_0^1 \sqrt{\frac{R^2}{R^2+t^2}}dt. \tag{19.48}$$

Notons que le changement de variables $t = R\sin(u)$ permet de retrouver l'expression $l_\sigma = \pi R/2$.

Pour avoir une approximation de π, il est loisible de calculer une approximation numérique de l'intégrale (19.48) (avec $R=1$) et de l'égaler à $\pi/2$. △

Exemple 19.13
Considérons l'arc de cercle de rayon R interceptée par l'angle θ présenté sur la figure 19.3.

Par définition, cette longueur sera

$$\int_{\theta_0}^{\theta_1} \sqrt{R^2\sin^2(t) + R^2\cos^2(t)}dt = R(\theta_1-\theta_0). \tag{19.49}$$

Le radian comme unité de mesure d'angle est donc l'unité parfaite : elle est la longueur d'arc interceptée (si le rayon est $R=1$). △

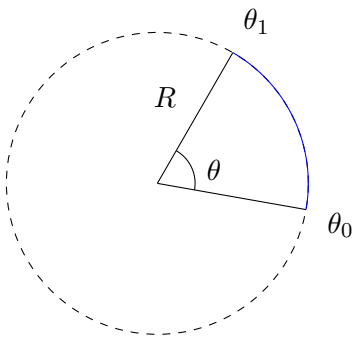


FIGURE 19.3 – Quelle est la longueur de la partie bleue de ce cercle de rayon R ?

19.14.
Si on veut savoir la longueur d'une courbe donnée sous la forme d'une fonction $y = y(x)$, un chemin qui trace la courbe est évidement donné par

$$\gamma(t) = (t, y(t)), \tag{19.50}$$

et le vecteur tangent au chemin est $\gamma'(t) = (1, y'(t))$. Donc

$$\|\gamma'(t)\| = \sqrt{1 + y'(t)^2}, \tag{19.51}$$

et

$$L = \int_a^b \sqrt{1 + y'(t)^2}. \tag{19.52}$$

Exemple 19.15
La longueur de l'hélice

$$\sigma(t) = \begin{pmatrix} \cos(2t) \\ \sin(2t) \\ \sqrt{5}t \end{pmatrix} \tag{19.53}$$

pour $t \in [0, 2\pi]$ est donnée par

$$l(\sigma) = \int_0^{4\pi} \sqrt{4\sin^2(2t) + 4\cos^2(2t) + 5}dt = \int_0^{4\pi} \sqrt{9} = 12\pi. \tag{19.54}$$

△

Définition 19.16.
*Soit $\sigma_1 \colon [a, b] \to \mathbb{R}^3$, un chemin et $\sigma_2 \colon [c, d] \to \mathbb{R}^3$, un autre chemin. On dit que ces chemins sont **équivalents** s'il existe une fonction $\varphi \colon [a, b] \to [c, d]$ strictement croissante telle que $\sigma_1(t) = \sigma_2\big(\varphi(t)\big)$.*

Deux chemins équivalents parcourent la même courbe dans le même sens. Ils ne le parcourent toutefois pas à la même vitesse. On dit que les chemins sont **opposée** si la fonction φ de la définition est strictement décroissante. Dans ce cas, ils ont la même image, mais parcourue dans le sens opposés. Nous disons que deux chemins équivalents sont un **changement de paramétrisation** pour la même courbe.

Dans le cas d'une paramétrisation équivalente, nous avons $\varphi(a) = c$ et $\varphi(b) = d$. Les points de départ et d'arrivée des deux paramètres coïncident. Dans le cas d'un paramètre qui va dans le sens opposé par contre nous avons automatiquement $\varphi(a) = d$ et $\varphi(b) = c$.

Proposition 19.17.
La longueur d'une courbe ne dépend pas du paramètre (équivalent ou opposé) choisi.

Démonstration. Soient $\sigma_1 : [a,b] \to \mathbb{R}^3$ et $\sigma_2 : [c,d] \to \mathbb{R}^3$ tels que

$$\sigma_1(t) = \sigma_2\big(\varphi(t)\big) \tag{19.55}$$

où $\varphi : [a,b] \to [a,d]$ est une bijection strictement monotone. Par définition on a

$$l(\sigma_1) = \int_a^b \|\sigma_1'(t)\| dt. \tag{19.56}$$

Nous pouvons exprimer la dérivée de σ_1 en termes de celle de σ_2 en dérivant la relation (19.55) :

$$\sigma_1'(t) = \varphi'(t)\sigma_2'\big(\varphi(t)\big). \tag{19.57}$$

En ce qui concerne la norme,

$$\|\sigma_1'(t)\| = |\varphi'(t)|\|\sigma_2'(t)\|. \tag{19.58}$$

Notez dans cette relation que $\varphi'(t)$ est un nombre (et non un vecteur). Étant donné que nous avons supposé que φ était monotone, soit elle est monotone croissante et $\|\varphi'(t)\| = \varphi'(t)$ pour tout t, soit elle est monotone décroissante et $\|\varphi'(t)\| =' \varphi(t)$ pour tout t.

Considérons d'abord le premier cas, c'est à dire $\|\varphi'(t)\| = \varphi'(t)$. Nous posons $s = \varphi(t)$, $ds = \varphi'(t)dt$. En remplaçant cela dans la formule de la longueur est

$$\begin{aligned} l(\sigma_1) &= \int_a^b \varphi'(t)\|\sigma_2\big(\varphi(t)\big)\|dt \\ &= \int_{\varphi(a)}^{\varphi(b)} \|\sigma_2'(s)\|ds \\ &= \int_c^d \|\sigma_2'(s)\|ds \\ &= l(\sigma_2). \end{aligned} \tag{19.59}$$

Si nous considérons maintenant une paramétrisation strictement décroissante. Dans ce cas, $\varphi'(t) \leqslant 0$ et $\|\varphi'(t)\| = -\varphi'(t)$. Nous posons encore une fois $s = \varphi(t)$, $ds = \varphi'(t)ds$. Ici il ne faut pas oublier que $\varphi(a) = d$ et $\varphi(b) = c$. Le calcul est à part cela le même en faisant attention au singe :

$$\begin{aligned} l(\sigma_1) &= \int_a^b \varphi'(t)\|\sigma_2\big(\varphi(t)\big)\|dt \\ &= -\int_{\varphi(a)}^{\varphi(b)} \|\sigma_2'(s)\|ds \\ &= -\int_d^c \|\sigma_2'(s)\|ds \\ &= \int_c^d \|\sigma_2'(s)\|ds \\ &= l(\sigma_2). \end{aligned} \tag{19.60}$$

Nous avons changé le signe en changeant l'ordre des bornes. □

19.4 Suite du chapitre

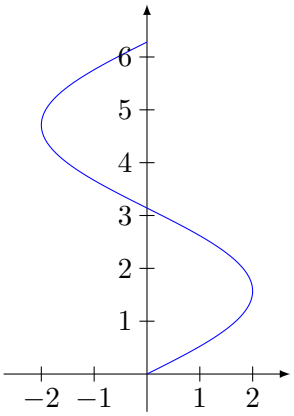
Le grand avantage des arcs paramétrés par rapports aux graphes de fonctions est le le graphe peut «faire des retours en arrière», ou bien des auto intersections. Outre les deux exemples typiques de la la figure 19.4, un exemple classique est la droite verticale. Les fonctions $y = ax + b$ permettent de décrire toutes les droites, sauf les droites verticales. Dans le cadre des courbes paramétrées, les droites verticales et horizontales sont sur pied d'égalité. Quelques exemples classiques :

Droite horizontale Une droite horizontale à la hauteur a est donnée par la courbe paramétrée $\gamma(t) = (t,a)$, avec $t \in I = \mathbb{R}$.

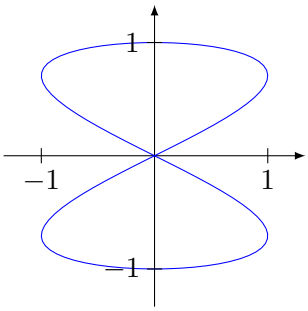
Droite verticale Une droite verticale à la distance b de l'origine est donnée par la courbe paramétrée $\gamma(t) = (b, t)$, avec $t \in I = \mathbb{R}$.

Graphe d'une fonction Le graphe d'une fonction $f\colon \mathbb{R} \to \mathbb{R}$ est donné par l'arc $\gamma(t) = \big(t, f(t)\big)$.

Un cercle Le cercle de rayon R est donné par l'arc $\gamma(t) = \big(R\cos(t), R\sin(t)\big)$.



(a) $\gamma(t) = (2\sin(t), t)$, avec $0 \leqslant t \leqslant 2\pi$



(b) $\gamma(t) = (sin(2t), sin(t))$ avec $-\pi \leqslant t \leqslant \pi$

FIGURE 19.4 – Des exemples d'arcs paramétrées. Ceux ne sont pas des graphes.

Remarque 19.18.
Afin d'alléger la notation, nous allons le plus souvent désigner l'arc (I, γ) simplement par la fonction γ. Il est cependant toujours *très* important de savoir sur quel intervalle nous considérons le chemin. Cela dépendra le plus souvent du contexte, et nous indiquerons l'intervalle I explicitement lorsqu'une ambigüité est à craindre.

Par exemple, lorsque nous considérons le cercle $\gamma(t) = \big(R\cos(t), R\sin(t)\big)$, le plus souvent l'intervalle de variation de t sera $I = [0, 2\pi]$. Par contre, si nous considérons la droite $\gamma(t) = (t, 2t)$, l'intervalle de variation de t sera naturellement $I = \mathbb{R}$.

19.5 Courbes paramétrés

19.5.1 Définitions et exemples

Exemple 19.19
Soit $v \in \mathbb{R}^3$ et $x_0 \in \mathbb{R}^3$. Le chemin

$$\sigma(t) = x_0 + tv \tag{19.61}$$

est une droite. Sa vitesse est $\sigma'(t) = v$. △

Exemple 19.20
La courbe

$$\sigma(t) = \begin{pmatrix}\cos(t)\\ \sin(t)\end{pmatrix} \in \mathbb{R}^2 \tag{19.62}$$

avec $t \in [0, 2\pi[$ est le cercle unité parcouru une fois dans le sens trigonométrique.

Notez que si on prend $t \in [0, 4\pi[$, nous avons un *autre* chemin ; c'est le même cercle unité, mais parcouru *deux* fois. Même si le «dessin» (le graphe) des deux est le même, le chemin n'est pas le même.

Le chemin

$$\gamma(t) = \begin{pmatrix}\cos(2\pi - t)\\ \sin(2\pi - t)\end{pmatrix} \tag{19.63}$$

est le cercle unité parcouru une fois dans le sens inverse. Encore une fois le «dessin» est le même, mais le chemin n'est pas le même. △

Exemple 19.21
Le chemin

$$\sigma(t) = \begin{pmatrix} t \\ t^2 \end{pmatrix} \tag{19.64}$$

est un chemin dont l'image est la parabole d'équation $y = x^2$. △

L'importance de la dérivée du chemin réside en le fait qu'elle donne la tangente. En effet le vecteur $\sigma'(t)$ est tangent au graphe de σ au point $\sigma(t)$.

Corollaire 19.22.
La tangente à un cercle est perpendiculaire au rayon.

Démonstration. Nous savons que pour un cercle,

$$y'(x) = \frac{-x}{\sqrt{R^2 - x^2}}. \tag{19.65}$$

Un point général du cercle a pour abscisse $x = R\cos(\theta)$. En remplaçant nous trouvons le coefficient directeur suivant pour la tangente :

$$y'\big(R\cos(\theta)\big) = -\frac{1}{\tan(\theta)}. \tag{19.66}$$

Par conséquent une droite perpendiculaire à la tangente aurait comme coefficient directeur le nombre $\tan(\theta)$. Or cela est bien le coefficient directeur du rayon qui joint le point $(0, 0)$ au point $\big(R\cos(\theta), R\sin(\theta)\big)$.

□

Exemple 19.23
Pour le cercle,

$$\sigma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}, \tag{19.67}$$

la dérivée est donnée par

$$\sigma'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t). \end{pmatrix} \tag{19.68}$$

Le produit scalaire $\sigma(t) \cdot \sigma'(t)$ est nul. Le vecteur $\sigma'(t)$ est donc bien tangent (corollaire 19.22).
△

Exemple 19.24
Le courbe donnée par le chemin

$$\sigma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \\ t \end{pmatrix} \tag{19.69}$$

est une hélice. Sa vitesse est

$$\sigma'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \\ 1 \end{pmatrix}. \tag{19.70}$$

Notez que pour tout $t \in \mathbb{R}$, nous avons $\|\sigma'(t)\| = \sqrt{2}$. △

Remarque 19.25.
Lorsqu'on parle d'une courbe dans l'espace, l'intervalle sur lequel on considère la variation du paramètre est une donné fondamentale. Elle fait partie intégrante de la définition de la courbe.

19.6 Élément de longueur

19.6.1 Élément de longueur : cartésiennes

Étant donné que la longueur d'arc d'une courbe paramétrée (I, γ) est donnée par l'intégrale de $\|\gamma'(t)\|$, il est naturel d'appeler le nombre $\|\gamma'(t)\|\,dt$, **l'élément de longueur** de la courbe γ au point $\gamma(t)$.

En coordonnées cartésiennes dans le plan, une courbe paramétré est donnée par

$$\gamma(t) = \big(x_1(t), x_2(t)\big), \tag{19.71}$$

et l'élément de longueur est

$$\|x'(t)\|\,dt = \sqrt{(x_1')^2 + (x_2')^2}\,dt. \tag{19.72}$$

19.6.2 Élément de longueur : polaires (1)

En coordonnées polaires, une courbe est donnée par

$$\gamma(t) = \big(\rho(t), \theta(t)\big), \tag{19.73}$$

et le passage aux cartésiennes se fait via les formules

$$\begin{cases} x(t) = \rho(t)\cos\big(\theta(t)\big) & \text{(19.74a)} \\ y(t) = \rho(t)\sin\big(\theta(t)\big). & \text{(19.74b)} \end{cases}$$

L'élément de longueur se trouve directement en remplaçant $x(t)$ et $y(t)$ dans la formule (19.72). Les dérivées sont données par

$$\begin{aligned} x'(t) &= \rho'(t)\cos\theta(t) - \rho(t)\theta'(t)\sin\theta(t) \\ y'(t) &= \rho'(t)\sin\theta(t) + \rho(t)\theta'(t)\cos\theta(t), \end{aligned} \tag{19.75}$$

et un calcul montre que

$$\big(x'(t)\big)^2 + \big(y'(t)\big)^2 = \big(\rho'(t)\big)^2 + \big(\rho(t)\big)^2\big(\theta'(t)\big)^2. \tag{19.76}$$

Nous reviendrons plus en détail sur le concept de changement de paramétrisation (ici, les polaires) à la section 19.7.

19.6.3 Élément de longueur : polaires (2)

Parfois on utilise θ comme paramètre. L'équation de la courbe est alors donnée en coordonnées polaires sous la forme

$$\rho(\theta) = f(\theta), \tag{19.77}$$

où f est une fonction réelle et il faut comprendre que nous parlons de la courbe $\big(\rho(\theta), \theta\big)$ en coordonnées polaires. En coordonnées cartésiennes, cette courbe est donnée par

$$\begin{cases} x(t) = \rho(t)\cos(t) & \text{(19.78a)} \\ y(t) = \rho(t)\sin(t) & \text{(19.78b)} \end{cases}$$

avec t qui parcours le plus souvent l'intervalle $[0, 2\pi]$. Notez qu'il se peut que le domaine ne soit pas toujours $[0, 2\pi]$; cela peut dépendre des circonstances. Quoi qu'il en soit, la donnée du domaine fait partie de la donnée d'une courbe, et il ne peut donc pas y avoir d'équivoques à ce niveau.

Nous utilisons à nouveau la formule (19.72) en y mettant les valeurs (19.78) :

$$\begin{cases} x'(t) = \rho'(t)\cos(t) - \rho(t)\sin(t) & \text{(19.79a)} \\ y'(t) = \rho'(t)\sin(t) + \rho(t)\cos(t), & \text{(19.79b)} \end{cases}$$

et

$$\big(x'(t)\big)^2 + \big(y'(t)\big)^2 = \rho'(t)^2 + \rho(t)^2. \tag{19.80}$$

Remarque 19.26.
N'oubliez pas, en utilisant ces formules, que ce qui rentre dans l'intégrale est la *racine carré* de $(x')^2 + (y')^2$.

Exemple 19.27
Calculons la circonférence du cercle. En coordonnées polaires, le graphe du cercle correspond à l'équation

$$\big(\rho(t), \theta(t)\big) = (R, t) \tag{19.81}$$

où R est constante (le rayon du cercle) et t va de 0 à 2π. En substituant dans l'équation (19.76), l'élément de longueur à intégrer est seulement

$$\sqrt{R^2} = R \tag{19.82}$$

parce que $\rho'(t) = 0$ et $\theta'(t) = 1$. La longueur du cercle est alors directement donnée par

$$l = \int_0^{2\pi} R dt = 2\pi R. \tag{19.83}$$

Nous pouvions aussi faire le calcul en coordonnées cartésiennes. Alors la courbe est donnée par les équations

$$\begin{aligned} x(t) &= R\cos(t) \\ y(t) &= R\sin(t) \end{aligned} \tag{19.84}$$

et $t \in [0, 2\pi]$. La circonférence du cercle est alors

$$l = \int_0^{2\pi} \sqrt{R^2\sin^2(t) + R^2\cos^2(t)}\, dt = \int_0^{2\pi} R\, dt = 2\pi R. \tag{19.85}$$

△

Remarque 19.28.
Il faut bien comprendre que quand on parle de courbes paramétrées en coordonnées cartésiennes on pense à une courbe dont le paramètre est, par exemple, t et les équations de la courbe sont $(x(t), y(t))$. Cela ne veut pas dire que x ou y soit le paramètre. La cas où x ou y est le paramètre est un cas particulier qui est possible seulement pour certaines courbes et notamment pour les graphes. Le cercle de rayon 1 n'est pas un graphe, donc si on veut utiliser x ou y comme paramètre il faut d'abord découper la courbe en deux morceaux, par exemple, la moitié inférieure ($y < 0$) et la moitié supérieure ($y > 0$).

Exemple 19.29
Une **cycloïde** est une courbe paramétrée par

$$\begin{cases} x(t) = a(t - \sin(t)) & \text{(19.86a)} \\ y(t) = a(1 - \cos(t)) & \text{(19.86b)} \end{cases}$$

avec $a > 0$ et $t \in \mathbb{R}$. Comme montré sur la figure 19.5, la cycloïde donne lieu à un graphe périodique. Il est possible de montrer (le faire) que le premier arc correspond à $t \in [0, 2\pi]$. Nous voulons donc calculer la longueur de l'arc sur cet intervalle.

Nous avons $x'(t) = a(1 - \cos(t))$ et $y'(t) = a\sin(t)$, de telle façon que

$$\sqrt{(x')^2 + (y')^2} = a\sqrt{2 - 2\cos(t)} = a\sqrt{4\sin^2\left(\frac{t}{2}\right)} = 2a\Big|\sin\frac{t}{2}\Big|. \tag{19.87}$$

La longueur est donc donnée par

$$\int_0^{2\pi} 2a|\sin\frac{t}{2}|dt = 4a\int_0^{\pi} \sin(t)dt = 8a. \tag{19.88}$$

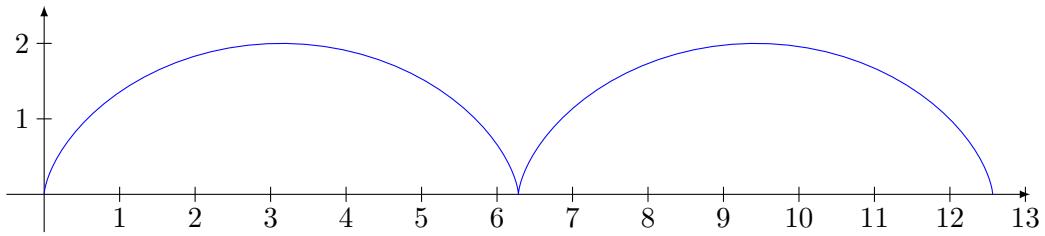


FIGURE 19.5 – La cycloïde de paramètre $a = 1$ entre 0 et 4π.

△

Exemple 19.30
La **cardioïde** est la courbe donnée par

$$\rho(\theta) = a(1 + \cos(\theta)). \tag{19.89}$$

avec $\theta \in [-\pi, \pi]$. Le nom de cette courbe provient de son graphe illustré à la figure 19.6.

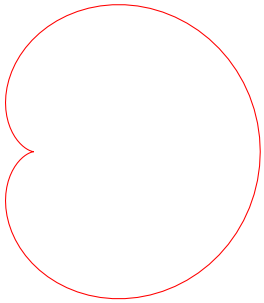


FIGURE 19.6 – Une cardioïde, $\rho = 1 + \cos(\theta)$.

L'équation (19.89) est donnée sous la forme (19.77), c'est à dire que $\theta(t) = t$ et $\theta'(t) = 1$, et par conséquent l'élément de longueur est donné par

$$\begin{aligned}(\rho')^2 + (\rho)^2 &= \big(-a\sin(\theta)\big)^2 + a^2\big(1+\cos(\theta)\big)^2 \\ &= a^2\sin^2(\theta) + a^2\big(1 + 2\cos(\theta) + \cos^2(\theta)\big) \\ &= a^2\big(1 + 1 + 2\cos(\theta)\big) \\ &= 2a^2\big(1+\cos(\theta)\big) \\ &= 4a^2\cos^2\frac{\theta}{2}.\end{aligned} \tag{19.90}$$

La longueur d'arc est donc donnée par

$$l = \int_{-\pi}^{\pi} 2a\cos\frac{\theta}{2}d\theta = 2a\int_{-\pi/2}^{\pi/2}\cos(t)2dt = 8a. \tag{19.91}$$

△

19.6.4 Approximation de la longueur par des cordes

Définition 19.31.
Soit un arc paramétré (I, γ). Un point $t \in I$ est dit ***régulier*** *si $\gamma'(t) \neq 0$, et il est dit* ***critique*** *si $\gamma'(t) = 0$. Le point $t \in I$ est dit* ***birégulier*** *si les vecteurs $\gamma'(t)$ et $\gamma''(t)$ sont linéairement indépendants et non nuls.*

Par extension, nous dirons également que le point $\gamma(t)$ lui-même est régulier, critique ou birégulier. Un arc est dit régulier *lorsque tous ses points sont réguliers.*

Note : dans le lemme 19.65 et ses dépendances, nous utilisons effectivement que l'arc γ est de classe C^2.

Nous savons que la longueur d'une courbe est donné par le supremum sur toutes les subdivision de la longueur des cordes correspondantes. De plus l'inégalité triangulaire nous enseigne que plus la subdivision est fine, plus la longueur sera grande. Il est donc naturel de penser que sur un petit intervalle, la longueur de la courbe ne doit pas être très différente de la longueur de la corde correspondante.

La proposition suivante est un énoncé précis et quantitatif de ce fait.

Proposition 19.32.
Soit (I,γ) un arc de classe $\mathcal{C}^1$ et $t_0 \in I$ un point régulier (c'est à dire $\gamma'(t_0) \neq 0$). Alors pour tout $\varepsilon > 0$, il y a un $\delta > 0$ tel que on trouve $t, t' \in I \cap (t_0, \delta)$ tels que

$$\left| \int_t^{t'} \|\gamma'(u)\| du - \|\gamma(t) - \gamma(t')\| \right| \leqslant 2\varepsilon |t - t'|. \tag{19.92}$$

Intuitivement, cette proposition signifie qu'au voisinage de t_0, la longueur d'arc est équivalente à celle de la corde.

Démonstration. Par la continuité de γ' (parce que γ est $\mathcal{C}^1$), pour tout ε, il existe un δ tel que

$$|t - t_0| < \delta \Rightarrow \|\gamma'(t) - \gamma'(t_0)\| \leqslant \varepsilon. \tag{19.93}$$

Nous considérons la fonction

$$u \mapsto \gamma(u) - \gamma(t_0) - (u - t_0)\gamma'(t_0), \tag{19.94}$$

dont la dérivée (par rapport à u) est

$$\gamma'(u) - \gamma'(t_0). \tag{19.95}$$

Nous y appliquons la formule des accroissements finis entre t et t' choisis dans $I \cap]t_0 - \delta, t_0 + \delta[$. Il existe un u entre t et t' tel que

$$\begin{aligned} \|\gamma(t) - \gamma(t_0) - (t - t_0)\gamma'(t_0) - \gamma(t') + \gamma(t_0) + (t' - t_0)\gamma'(t_0)\| & \\ = |t - t'| \|\gamma'(u) - \gamma'(t_0)\| & \\ \leqslant \varepsilon |t - t'|. & \end{aligned} \tag{19.96}$$

En simplifiant ce qui peut être simplifié dans le membre de gauche, nous trouvons

$$\|\gamma(t) - \gamma(t') - (t - t')\gamma'(t_0)\| \leqslant \varepsilon |t - t'|. \tag{19.97}$$

Le membre de gauche peut être minoré en utilisant la proposition 6.32 :

$$\Big| \|\gamma(t) - \gamma(t')\| - \|(t - t')\gamma'(t_0)\| \Big| \leqslant \varepsilon |t - t'|. \tag{19.98}$$

D'autre part, les inégalités (6.17) montrent que

$$-\|\gamma'(u) - \gamma'(t_0)\| \leqslant \|\gamma'(u)\| - \|\gamma'(t_0)\| \leqslant \|\gamma'(u) - \gamma'(t_0)\|. \tag{19.99}$$

Si de plus u est compris entre t et t', ces inégalités sont encore coincées entre $-\varepsilon$ et ε. En intégrant (19.99) par rapport à u entre t et t', nous obtenons

$$\left| \int_t^{t'} \|\gamma'(u)\| - (t - t')\|\gamma'(t_0)\| \right| \leqslant \varepsilon |t - t'|. \tag{19.100}$$

Afin d'alléger les notations pour la ligne suivante, nous notons A le nombre positif $\int_t^{t'} \|\gamma'(u)\| du$. Nous avons

$$\begin{aligned} \Big| A - \|\gamma(t) - \gamma(t')\| \Big| &= \Big| A - |t - t'| \, \|\gamma'(t_0)\| + |t - t'| \, \|\gamma'(t_0)\| - \|\gamma(t) - \gamma(t')\| \Big| \\ &\leqslant \Big| A - |t - t'| \, \|\gamma'(t_0)\| \Big| + \Big| |t - t'| \, \|\gamma'(t_0)\| - \|\gamma(t) - \gamma(t')\| \Big|. \end{aligned} \tag{19.101}$$

L'équation (19.98) montre que le second terme est plus petit ou égal à $\varepsilon|t-t'|$. En ce qui concerne le premier terme, étant donné que A est positif,

$$\Big|A - |t-t'|\,\|\gamma'(t_0)\|\Big| \leqslant \Big|A - (t-t')\,\|\gamma'(t_0)\|\Big| \leqslant \varepsilon|t-t'|. \tag{19.102}$$

Au final, l'inéquation (19.101) donne

$$\Big|A - \|\gamma(t)-\gamma(t')\|\Big| \leqslant 2\varepsilon\,|t-t'|, \tag{19.103}$$

ce qu'il fallait démontrer. □

19.7 Arc géométrique

Définition 19.33.
Soient (I,γ) et (J,g) deux arcs paramétrés de classe $\mathcal{C}^k$. On dit qu'il sont ***équivalents*** *s'il existe une bijection $\theta\colon I \to J$ de classe $\mathcal{C}^k$, d'inverse de classe $\mathcal{C}^k$ telle que $g = \gamma\circ\theta$. Nous notons $\gamma \sim g$ lorsque γ et g sont équivalents (les ensembles I et J sont sous-entendus).*

Le passage d'une paramétrisation (I,γ) à une autre (J,g) se fait selon le diagramme suivant :

$$\begin{array}{ccc} I & \xrightarrow{\ \gamma\ } & \mathbb{R}^n \\ {\scriptstyle\theta}\uparrow & \nearrow_{g} & \\ J & & \end{array} \tag{19.104}$$

Proposition 19.34.
La relation donnée dans la définition 19.33 est une relation d'équivalence.

Démonstration. Les trois points d'une relation d'équivalence se vérifient en utilisant le fait que θ est inversible, et que l'inverse θ^{-1} jouit des mêmes propriétés de continuité ($\mathcal{C}^k$) que θ.

Réflexivité Nous avons $\gamma \sim \gamma$ avec $\theta = \mathrm{Id}$.

Symétrie Si $\gamma \sim g$, alors nous avons une application θ telle que $g = \gamma\circ\theta$, et donc $\gamma = g\circ\theta^{-1}$, ce qui montre que $g \sim \gamma$.

Transitivité Si $\gamma \sim g$ et $g \sim h$ avec $g = \gamma\circ\theta$ et $h = g\circ\omega$, alors $h = \gamma\circ(\theta\circ\omega)$, ce qui montre que $\gamma \sim h$.

□

Si les arcs (I,γ) et (J,g) sont équivalents, les images dans $\mathbb{R}^n$ sont identiques, et décrivent donc «le même dessin». Nous allons préciser cette notion plus loin.

Définition 19.35.
Pour cette raison les classes d'équivalences sont appelées des ***arcs géométriques*** *(de classe $\mathcal{C}^k$).*

Si Γ est une arc géométrique, ses représentants sont dits des **paramétrages admissibles** ou, plus simplement *paramétrisation*. On dit que l'application $\theta\colon J \to I$ est un **changement de variable**. Nous disons que un arc géométrique est *compact* quand ses représentants sont compacts.

Lemme 19.36.
Dans le cas d'un arc $\mathcal{C}^1$, les changements de variables sont strictement monotones (croissants ou décroissants).

Démonstration. Nous considérons (I,γ) et (J,g), deux paramétrisations différentes du même arc géométrique, et $\theta \in \mathcal{C}^1(J,I)$ le changement de variable. Nous allons noter t la variable sur I et s la variable sur J. Par définition, $\theta\big(\theta^{-1}(t)\big) = t$, et par conséquent,

$$\theta'\big(\theta^{-1}(t)\big)(\theta^{-1})'(t) = 1. \tag{19.105}$$

En particulier $\theta'\big(\theta^{-1}(t)\big)$ ne s'annule pas pour aucune valeur de t. Mais $\theta^{-1}(t)$ peut prendre n'importe quelle valeur dans J, donc nous avons $\theta'(s) \neq 0$ pour tout $s \in J$. Cela signifie bien que θ est strictement monotone. En effet, θ' étant continue, elle ne peut pas changer de signe sans passer par zéro (théorème 12.21 des valeurs intermédiaires). □

Théorème 19.37.
La longueur d'un arc est indépendante de sa paramétrisation, c'est à dire que les représentants d'un arc géométrique compact de classe $\mathcal{C}^1$ ont même longueur.

Démonstration. Nous utilisons les mêmes notations que celles du lemme 19.36. Nous savons déjà que le changement de variable $\theta\colon J \to I$ est strictement monotone. Supposons que θ soit croissante. En effectuant un changement de variable dans l'intégrale qui donne la longueur [1] nous avons

$$\begin{aligned} l(\gamma) &= \int_I \|\gamma'(t)\|dt \\ &= \int_J \|\gamma'\big(\theta(s)\big)\|\theta'(s)ds \\ &= \int_J \|\gamma'\big(\theta(s)\big)\theta'(s)\|ds \\ &= \int_J \|\frac{d}{ds}(\gamma\circ\theta)(s)\|ds \\ &= \int_J \|g'(s)\|ds \\ &= l(J,g). \end{aligned} \tag{19.106}$$

□

Définition 19.38.
*Nous nommons **longueur** d'un arc géométrique la longueur commune de tous ses représentants. On dit que l'arc géométrique est **rectifiable** si sa longueur est $< \infty$.*

19.7.1 Abscisse curviligne et paramétrisation normale

Définition 19.39.
*Soit (I,γ) un arc paramétré continu rectifiable. Nous appelons **abscisse curviligne** de γ toute application $\phi\colon I \to \mathbb{R}$ telle que pour tout $t,t' \in I$ avec $t < t'$, nous ayons*

$$l\big([t,t'],\gamma\big) = \big|\phi(t') - \phi(t)\big|. \tag{19.107}$$

*Si il existe un $t_0 \in I$ tel que $\phi(t_0) = 0$, alors nous disons que t_0 est l'**origine** de l'abscisse ϕ.*

*Un arc paramétré (I,γ_N) continu rectifiable est dit **normal** si l'identité est une abscisse curviligne. Dans ce cas, pour tout choix de t et t' dans I avec $t < t'$, nous avons*

$$l\big([t,t'],\gamma_N\big) = t' - t. \tag{19.108}$$

19.40.
Si γ est de classe C^1 et est une paramétrisation normale, alors pour tout x_1, x_2 dans le domaine nous avons

$$l\big([x_1,x_2],\gamma\big) = \int_{x_1}^{x_2} \|\gamma'(t)\|dt = x_2 - x_1. \tag{19.109}$$

Cela implique $\|\gamma'(t)\| = 1$ pour tout t.

1. Théorème 19.9.

Exemple 19.41
Le cercle unitaire est donné par l'arc

$$\gamma(t) = \big(\cos(t), \sin(t)\big) \tag{19.110}$$

et $t \in [0, 2\pi]$. Pour tout choix de t et t' dans $[0, 2\pi]$, nous avons

$$l\big([t, t'], \gamma\big) = \int_t^{t'} \sqrt{\sin^2(u) + \cos^2(u)}du = t' - t. \tag{19.111}$$

Les angles exprimés en radians forment donc une paramétrisation normale du cercle de rayon 1.
△

Lemme 19.42.
Pour un arc paramétré compact, la longueur d'arc est une abscisse curviligne.

Démonstration. Par définition de la longueur d'arc φ, nous avons

$$\varphi(t') - \varphi(t) = l\big([a, t'], \gamma\big) - l\big([a, t], \gamma\big) = \diamondsuit. \tag{19.112}$$

Supposons pour fixer les idées que $t' > t$. En utilisant la proposition 19.6, nous avons

$$l\big([a, t'], \gamma\big) = l\big([a, t], \gamma\big) + l\big([t, t'], \gamma\big), \tag{19.113}$$

et donc après simplification de deux termes,

$$\diamondsuit = l\big([t, t'], \gamma\big), \tag{19.114}$$

ce qui est précisément la propriété demandée pour être une abscisse curviligne. □

Proposition 19.43.
Pour tout arc paramétré C^1 sans points critiques, il existe un changement de coordonnées qui rend l'arc normal.

Démonstration. Soit (I, γ) un arc de classe $\mathcal{C}^1$. Nous devons montrer qu'il existe un intervalle J et une application $\theta\colon J \to I$ de classe $\mathcal{C}^1$ et d'inverse $\mathcal{C}^1$ tel que l'arc (J, γ_N) soit $\mathcal{C}^1$ où $\gamma_N = \gamma \circ \theta$.

Si $I = [a, b]$, nous considérons la fonction

$$\begin{aligned} \phi\colon I &\to \mathbb{R}^+ \\ t &\mapsto \int_a^t \|\gamma'(u)\|du. \end{aligned} \tag{19.115}$$

Étant définie par l'intégrale d'une fonction $\mathcal{C}^0$, la fonction ϕ est $\mathcal{C}^1$, et nous avons $\phi'(t) = \|\gamma'(t)\| > 0$ pour tout $t \in I$. Vue comme application $\phi\colon [a, b] \to [0, l(\gamma)]$, l'application ϕ est bijective et d'inverse $\mathcal{C}^1$. Voyons cela point par point.

(1) La fonction ϕ est injective parce que strictement croissante.

(2) Elle est surjective parce que $\phi(a) = 0$ et $\phi(b) = l(\gamma)$.

(3) La continuité de l'inverse est plus délicate. Soit $l \in [0, l(\gamma)]$ et $\varepsilon > 0$. Pour prouver la continuité de ϕ^{-1} en s, nous devons trouver un δ tel que

$$|s - s'| < \delta \Rightarrow \big|\phi^{-1}(s) - \phi^{-1}(s')\big| < \varepsilon. \tag{19.116}$$

Étant donné que s et s' sont dans l'image de ϕ, nous considérons les uniques t et t' tels que $s = \phi(t)$ et $s' = \phi(t')$. La quantité $\phi(t) - \phi(t')$ devient

$$\int_a^t \|\gamma'(u)\|du - \int_a^{t'} \|\gamma'(u)\|du = \int_t^{t'} \|\gamma'(u)\|du. \tag{19.117}$$

D'autre part, $\phi^{-1}(s) = t$ et $\phi^{-1}(s') = t'$, donc la condition (19.117) devient

$$|\int_{t'}^{t} \|\gamma'(u)\| du| \leqslant \delta \Rightarrow |t - t'| < \varepsilon. \tag{19.118}$$

Cela revient à la continuité des fonctions définies par des intégrales.

(4) La dérivée de son inverse est donnée par [2]

$$(\phi^{-1})'(s) = \frac{1}{\phi'\big(\phi^{-1}(s)\big)}. \tag{19.119}$$

Nous avons vu que ϕ^{-1} et ϕ' étaient continues. La fonction $(\phi^{-1})'$ étant exprimée en termes de ces deux fonctions elle est également continue.

Nous considérons l'arc paramétré (J, γ_N) avec $J = [0, l(\gamma)]$ et

$$\gamma_N(s) = (\gamma \circ \phi^{-1})(s). \tag{19.120}$$

Nous montrons maintenant que cette nouvelle paramétrisation est normale. Soient $0 \leqslant s \leqslant s' \leqslant l(\gamma)$,

$$\begin{aligned}
l\big([s, s'], g\big) &= \int_s^{s'} \|\gamma'_N(u)\| du \\
&= \int_{\phi^{-1}(s)}^{\phi^{-1}(s')} \|(\gamma'_N \circ \phi)(t)\| \phi'(t) dt \\
&= \int_{\phi^{-1}(s)}^{\phi^{-1}(s')} \|(\gamma_N \circ \phi)'(t)\| dt \\
&= \int_{\phi^{-1}(s)}^{\phi^{-1}(s')} \|\gamma'(t)\| dt \\
&= \int_0^{\phi^{-1}(s')} \|\gamma'(t)\| \, dt - \int_0^{\phi^{-1}(s)} \|\gamma'(t)\| \, dt \\
&= \phi\big(\phi^{-1}(s')\big) - \phi\big(\phi^{-1}(s)\big) \\
&= s' - s,
\end{aligned} \tag{19.121}$$

ce qui prouve que la paramétrisation (J, γ_N) est normale. □

Nous retenons que la paramétrisation normale de γ est donnée par (J, γ_N) avec $J = [0, l(\gamma)]$ et

$$\gamma_N(s) = (\gamma \circ \phi^{-1})(s) \tag{19.122}$$

où

$$\begin{aligned}
\phi\colon I &\to \mathbb{R}^+ \\
t &\mapsto \int_a^t \|\gamma'(u)\| du.
\end{aligned} \tag{19.123}$$

Notons aussi que ϕ est une fonction croissante, étant l'intégrale d'une fonction positive.

Exemple 19.44
Trouvons les coordonnées normales pour la cycloïde donnée par

$$\begin{cases} x(t) = a(t - \sin(t)), & \text{(19.124a)} \\ y(t) = a(1 - \cos(t)) & \text{(19.124b)} \end{cases}$$

et $t \in]0, 2\pi[$. Relire l'exemple 19.29.

2. Pour obtenir cette formule, dérivez les deux membres de l'équation $\phi\big(\phi^{-1}(s)\big) = s$.

D'abord nous trouvons ϕ avec la formule (19.123) avec $a = 0$. En utilisant le bout de calcul (19.87), nous avons

$$\phi(t) = 2a \int_0^t \sin \frac{u}{2} du = 4a \left(1 - \cos \frac{t}{2}\right). \tag{19.125}$$

Pour trouver $\phi^{-1}(s)$, nous résolvons l'équation

$$s = \phi\big(\phi^{-1}(s)\big) \tag{19.126}$$

par rapport à $\phi^{-1}(s)$. Dans un premier temps, nous trouvons

$$1 - \frac{s}{4a} = \cos \frac{\phi^{-1}(s)}{2}, \tag{19.127}$$

donc $\frac{\phi^{-1}(s)}{2} = \arccos(\frac{4a-s}{4a})$, et finalement

$$\phi^{-1}(s) = 2 \arccos \left(\frac{4a - s}{4a}\right). \tag{19.128}$$

Il nous reste à injecter cela dans les expressions de $x(t)$ et $y(t)$ pour trouver $(\gamma_N)_x(s)$ et $(\gamma_N)_y(s)$. D'abord,

$$(\gamma_N)_x(s) = a\big[\phi^{-1}(s) - \sin\big(\phi^{-1}(s)\big)\big]. \tag{19.129}$$

Nous utilisons maintenant la formule trigonométrique $\sin(x) = 2 \sin \frac{x}{2} \cos \frac{x}{2}$ afin de simplifier les expression :

$$\begin{aligned}(\gamma_N)_x &= a\Big[2 \arccos \left(\frac{4a - s}{4a}\right) - 2 \sin\big(\arccos \left(\frac{4a - s}{4a}\right)\big) \cos\big(\arccos \left(\frac{4a - s}{4a}\right)\big)\Big] \\ &= a\Big[2 \arccos \left(\frac{4a - s}{4a}\right) - \frac{4a - s}{2a} \sqrt{1 - \left(\frac{4a - s}{4a}\right)^2}\Big] \\ &= 2a \arccos \left(\frac{4a - s}{4a}\right) - \sqrt{8as - s^2}\, \frac{4a - s}{8a}\end{aligned} \tag{19.130}$$

où nous avons utilisé la formule $\sin\big(\arccos(x)\big) = \sqrt{1 - x^2}$. Ensuite, pour obtenir $(\gamma_N)_y$ nous devons calculer

$$(\gamma_N)_y(s) = a\big[1 - \cos\big(\phi^{-1}(s)\big)\big]. \tag{19.131}$$

Encore une fois, il est intéressant d'exprimer le cosinus en termes des angles divisés par deux : $\cos(x) = \cos^2 \frac{x}{2} - \sin^2 \frac{x}{2}$.

$$\begin{aligned}(\gamma_N)_y &= a\Big[1 - \cos^2 \frac{\phi^{-1}(s)}{2} + \sin^2 \frac{\phi^{-1}(s)}{2}\Big] \\ &= a\Big[2 - 2 \cos^2 \frac{\phi^{-1}(s)}{2}\Big] \\ &= 2a\Big[1 - \left(\frac{4a - s}{4a}\right)^2\Big].\end{aligned} \tag{19.132}$$

Dans cette paramétrisation, $s \in]0, 8a[$. △

Exemple 19.45
La cardioïde $\rho(\theta) = a\big(1 + \cos(\theta)\big)$ avec θ entre $-\pi$ et π. Avant d'utiliser la formule (19.123), nous devons trouver l'élément de longueur de la cardioïde. Étant donné la façon dont l'équation de la cardioïde nous est donnée, l'élément de longueur est donné par[3] (19.80) :

$$\begin{aligned}\|\gamma'(u)\|^2 &= a^2 \sin^2(u) + a^2(1 + \cos(u))^2 \\ &= 2a^2\big(1 + \cos(u)\big),\end{aligned} \tag{19.133}$$

3. Nous vous déconseillons d'étudier cette formule par cœur. Sachez cependant la retrouver assez vite.

et par conséquent [4]

$$\begin{aligned}\phi(t) &= \int_0^t \sqrt{2a^2\big(1+\cos(u)\big)}du \\ &= \int_0^t \sqrt{2a^2\left(1+\cos^2\frac{u}{2}-\sin^2\frac{u}{2}\right)}du \\ &= 2a\int_0^t \cos\frac{u}{2}du \\ &= 4a\sin\frac{t}{2}.\end{aligned} \tag{19.134}$$

Pour trouver l'inverse, nous résolvons $\phi\big(\phi^{-1}(s)\big)=s$ par rapport à $\phi^{-1}(s)$:

$$\begin{aligned}4a\sin\left(\frac{\phi^{-1}(s)}{2}\right)&=s,\\ \phi^{-1}(s)&=2\arcsin\left(\frac{s}{4a}\right).\end{aligned} \tag{19.135}$$

Avant d'écrire trop brutalement $\gamma_N(s)=(\gamma\circ\phi^{-1})(s)$, il faut comprendre comment est γ. Nous avons reçu la courbe sous forme polaire, c'est à dire

$$\gamma(t)=\big(\gamma_r(t),\gamma_\theta(t)\big)=\Big(a\big(1+\cos(t)\big),t\Big). \tag{19.136}$$

C'est comme cela qu'il faut comprendre la donnée $\rho(\theta)=a\big(1+\cos(\theta)\big)$. Maintenant la formule $\gamma_N(s)=(\gamma\circ\phi^{-1})(s)$ devient

$$\begin{cases}(\gamma_N)_r(s)=\gamma_r\big(\phi^{-1}(s)\big) & \text{(19.137a)}\\ (\gamma_N)_\theta(s)=\gamma_\theta\big(\phi^{-1}(s)\big). & \text{(19.137b)}\end{cases}$$

Étant donné que $\gamma_\theta(t)=t$, la seconde est facile :

$$(\gamma_N)_\theta(s)=2\arcsin\left(\frac{s}{4a}\right). \tag{19.138}$$

Pour la première,

$$(\gamma_N)_r(s)=a\big[1+\cos\big(2\arcsin\frac{s}{4a}\big)\big]=\frac{16a^2-s^2}{8a}. \tag{19.139}$$

Nous écrivons donc le nouveau paramétrage en coordonnées polaires sous la forme

$$\left(\frac{16a^2-s^2}{8a},2\arcsin\frac{s}{4a}\right). \tag{19.140}$$

La question qui arrive maintenant est de savoir quel intervalle parcours la nouvelle variable s. D'après le résultat de l'exemple 19.89, la longueur de la cardioïde est de $8a$ et nous avons donc $s\in[0,8a]$. Cependant, la condition d'existence de arcsin nous interdit d'avoir s plus grand que $4a$ en valeur absolue. Où est le problème ?

Le problème est que nous avons changé l'origine de notre paramètre en donnant $\phi(t)$ comme une intégrale à partir de 0 au lieu de $-\pi$. Cela se voit en regardant de quel point nous partons : en $s=0$ nous sommes sur le point $(2a,0)$ tandis qu'avec le paramètre original, c'est à dire $\theta\in[-\pi,\pi]$, nous avons pour $\theta=-\pi$ le point $(0,-\pi)$.

Il se passe donc que si nous commençons à parcourir la cardioïde avec $s=0$, nous partons du milieu, et nous ne parcourons donc pas tout. Étant donné que le «premier» point de la cardioïde est le point $(0,-\pi)$, le paramètre s commence en $s=-4a$, et nous avons comme intervalle :

$$s\in[-4a,4a], \tag{19.141}$$

ce qui est en accord avec la conditions d'existence. △

Quel enseignement tirer de cet exemple ? Lorsqu'on calcule $\phi(t)$ pour trouver les coordonnées normales, il y a deux solutions.

4. L'utilisation stricte de la formule (19.123) demanderait d'intégrer à partir de $-\pi$. Pour plus de simplicité, nous intégrons à partir de zéro, et nous verrons plus tard comment adapter l'intervalle du nouveau paramètre.

(1) Utiliser strictement la formule $\phi(t) = \int_a^t \|\gamma'(u)\|du$, en prenant bien comme borne de départ le point de départ de la paramétrisation de γ. À ce moment la coordonnée normale construite aura $[0, l(\gamma)]$ comme intervalle de variation.

(2) Faire commencer l'intervalle d'intégration en zéro (ou ailleurs). Un bon choix peut simplifier quelque calculs, mais alors il faudra bien choisir la valeur de départ de la nouvelle coordonnées pour que le «premier» point de la courbe soit correct. Dans ce cas, la longueur de l'intervalle sera quand même $l(\gamma)$. Il n'y a donc pas de problèmes pour trouver la valeur du bout de l'intervalle de variation du paramètre normal.

Dans tous les cas, il faut bien préciser l'intervalle de variation du paramètre lorsqu'on donne une courbe paramétrée.

19.7.2 Tangente à une courbe paramétrée

Définition 19.46.
*Soit (I, γ) un arc paramétré de classe $\mathcal{C}^k$ avec $k \geqslant 1$. Nous disons que la courbe admet une **tangente** en $\gamma(t_0) \in \mathbb{R}^n$ lorsque les deux conditions suivantes sont remplies*

(1) $\gamma(t) \neq \gamma(t_0)$ pour tout t dans un voisinage de t_0 ;

(2) la direction de la droite qui passe par $\gamma(t)$ et $\gamma(t_0)$ admet une limite lorsque $t \to t_0$.

Dans ce cas, la tangente sera la droite passant par le point $\gamma(t_0)$ et dont la direction est donnée par la limite.

Dans cette définition, par **direction** d'une droite, nous entendons le vecteur de norme 1 parallèle à celle-ci sans tenir compte du signe. La tangente sera donc la droite passant par $\gamma(t_0)$ et parallèle au vecteur

$$\lim_{t\to t_0} \frac{\gamma(t) - \gamma(t_0)}{\|\gamma(t) - \gamma(t_0)\|}. \tag{19.142}$$

Évidement si nous avions écrit $\gamma(t_0) - \gamma(t)$, ça n'aurait pas changé la droite. Par abus de langage, nous parlerons souvent de «la direction u» même lorsque u n'est pas de norme 1.

Formellement, une direction est une classe d'équivalence de vecteurs pour la relation $u \sim v$ s'il existe $\lambda \neq 0$ tel que $u = \lambda v$, mais nous n'aurons pas besoin de cette précision ici.

Sans surprises, la tangente est à peu près toujours donnée par la dérivée lorsqu'elle existe. Plus précisément nous avons le

Théorème 19.47.
Soit (I, γ), un arc paramétré de classe $\mathcal{C}^k$ ($k \geqslant 1$) et $t_0 \in I$ tel que

$$\gamma'(t_0) = \gamma''(t_0) = \ldots = \gamma^{(q-1)}(t_0) = 0 \tag{19.143}$$

et

$$\gamma^{(q)}(t_0) \neq 0 \tag{19.144}$$

pour un entier $1 \leqslant q \leqslant k$. Alors γ admet une tangente en $\gamma(t_0)$ de direction $\gamma^{(q)}(t_0)$.

Démonstration. Le développement de $\gamma(t_0)$ en série de Taylor autour de t jusqu'à l'ordre q est

$$\begin{aligned}\gamma(t) = \gamma(t_0) + \gamma'(t_0)|t - t_0| + \frac{\gamma' t(t_0)}{2}|t - t_0|^2 + \cdots + \frac{\gamma^{(q)}(t_0)}{q!}|t - t_0|^q \\ + \varepsilon(t)|t - t_0|^q\end{aligned} \tag{19.145}$$

où ε est une application $\varepsilon\colon \mathbb{R} \to \mathbb{R}^n$ telle que $\lim_{t\to t_0} \varepsilon(t) = 0$. En utilisant les hypothèses, nous éliminons la majorité des termes dans le développement (19.145) :

$$\gamma(t) - \gamma(t_0) = \frac{1}{q!}\gamma^{(q)}(t_0)|t - t_0|^q + \varepsilon(t)|t - t_0|^q. \tag{19.146}$$

La direction de la droite qui joint $\gamma(t)$ à $\gamma(t_0)$ est donc donnée par

$$\frac{\gamma(t)-\gamma(t_0)}{\|\gamma(t)-\gamma(t_0)\|}=\frac{\frac{1}{q!}\gamma^{(q)}(t_0)|t-t_0|^q+\varepsilon(t)|t-t_0|^q}{\|\frac{1}{q!}\gamma^{(q)}(t_0)|t-t_0|^q+\varepsilon(t)|t-t_0|^q\|} \tag{19.147}$$

et la limite lorsque $t\to t_0$ donne $\gamma^{(q)}(t_0)$ comme direction de la tangente.

□

Lorsque le théorème s'applique, le vecteur

$$\tau=\frac{\gamma^{(q)}(t_0)}{\|\gamma^{(q)}(t_0)\|} \tag{19.148}$$

est appelé le **vecteur unitaire tangent** en $\gamma(t_0)$ à l'arc paramétré γ.

Corollaire 19.48.
Si (I,γ) est un arc paramétré de classe $\mathcal{C}^1$ régulier (c'est à dire $\gamma'(t)\neq 0$ pour tout t) alors l'arc admet une tangente en tout point et le vecteur unitaire de la tangente est donné par

$$\tau(t)=\frac{\gamma'(t)}{\|\gamma'(t)\|}, \tag{19.149}$$

pour tout t dans I.

Corollaire 19.49.
Si $\gamma=(J,\gamma_N)$ est un arc paramétré de classe $\mathcal{C}^1$, normal, alors le vecteur unitaire de la tangente au point $\gamma_N(s)$ est donné par $\tau(s)=\gamma'_N(s)$.

Démonstration. Nous devons démontrer que dans le cas d'une paramétrisation normale nous avons $\|\gamma'_N(s)\|=1$ pour tout s. Par définition,

$$l\big([s,s'],g\big)=\int_s^{s'}\|\gamma'_N(u)\|du=s'-s. \tag{19.150}$$

Par conséquent,

$$\lim_{h\to 0}\frac{1}{h}\int_s^{s+h}\|\gamma'_N(u)\|du=\lim_{y\to 0}\frac{s+h-s}{h}=1. \tag{19.151}$$

Cela implique que $\|\gamma'_N(s)\|=1$, et donc en particulier que (J,γ_N) est un arc régulier. Le corollaire précédent montre alors que $\tau(s)=\gamma'_N(s)/\|\gamma'_N(s)\|=\gamma'_N(s)$. □

Exemple 19.50
Considérons la courbe $\gamma(t)=(t^2,t^3)$, et cherchons la tangente en $t_0=0$. En dérivant nous avons successivement

$$\begin{aligned}\gamma(t)&=(t^2,t^3)\\ \gamma'(t)&=(2t,3t^2)\\ \gamma''(t)&=(2,6t).\end{aligned} \tag{19.152}$$

En posant $t=0$, nous trouvons que $\gamma'(0)=0$ mais $\gamma''(0)=(2,0)\neq 0$. Le théorème nous dit donc que la direction de la tangente est horizontale. Nous pouvons faire le calcul directement :

$$\frac{\gamma(t)-\gamma(t_0)}{\|\gamma(t)-\gamma(t_0)\|}=\frac{(t^2,t^3)}{\sqrt{t^4+t^6}}=\frac{(t^2,t^3)}{t^2\sqrt{1+t^2}}=\frac{(1,t)}{\sqrt{1+t^2}}, \tag{19.153}$$

dont la limite $t\to 0$ est bien le vecteur horizontal $(1,0)$.

La figure 19.7 montre quelque tangente, c'est à dire quelques vecteurs dans la direction $\gamma'(t)$ (pour les $t\neq 0$, il ne faut pas aller à la dérivée seconde). Nous remarquons que de part et d'autres du sommet, les vecteurs ne sont pas dirigés dans le même sens. *En tant que vecteurs* de norme 1,

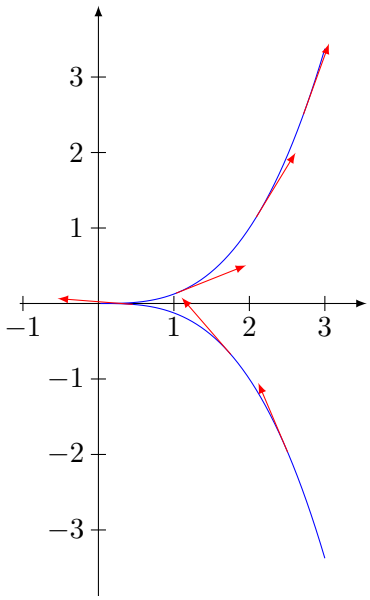


FIGURE 19.7 – Quelques tangentes de la courbe $\gamma(t) = (t^2, t^3)$.

ces vecteurs n'ont pas de limites quand $t \to 0$. Ce sont bien les *directions* qui ont une limite, parce que la direction ne tient pas compte du sens.

△

19.8 Repère de Frenet

Dans cette section, nous ne considérons que des courbes dans $\mathbb{R}^3$.

Proposition 19.51.
Soit $\gamma = (J, \gamma_N)$ un arc paramétré normal de classe $\mathcal{C}^2$. Alors pour toute valeur de s dans J, nous avons

$$\tau(s) \cdot \tau'(s) = 0 \tag{19.154}$$

où $\tau(s) = \gamma'_N(s)$. C'est à dire que la dérivée seconde est perpendiculaire à la dérivée première.

Démonstration. La paramétrisation étant normale, nous avons

$$\|\gamma'_N(s)\|^2 = \sum_{i=1}^{n} x'_i(s)^2 = 1; \tag{19.155}$$

ce qui implique, en dérivant les deux membres, que

$$0 = 2\sum_{i=1}^{n} x'_i(s)x''_i(s), \tag{19.156}$$

c'est à dire exactement $\gamma'_N(s) \cdot \gamma''_N(s) = 0$; d'où la thèse. □

Remarque 19.52.
Si nous n'utilisons pas des coordonnées normales, la proposition 19.51 n'est pas spécialement vraie. Prenons par exemple la courbe qui donne la parabole :

$$\gamma(t) = (t, t^2) \tag{19.157a}$$
$$\gamma'(t) = (1, 2t) \tag{19.157b}$$
$$\gamma''(t) = (0, 2) \tag{19.157c}$$

Nous avons $\gamma'(t) \cdot \gamma''(t) = 4t$. Par conséquent, la dérivée seconde n'est la normale à la courbe que en $t = 0$. Cela est une propriété très intéressante des coordonnées normales : la dérivée seconde d'une coordonnées normale donne un vecteur normal à la courbe, c'est à dire perpendiculaire à la tangente.

Définition 19.53.
Soit $\gamma = (J, \gamma_N)$ un arc paramétré normal de classe $\mathcal{C}^2$.

*(1) Le **vecteur unitaire tangent** est donné par le corollaire 19.48 : $\tau(t) = \gamma'(t)/\|\gamma'(t)\|$.*

*(2) La **normale principale** est le vecteur $\tau'(s)$. Le **vecteur unitaire normal** est le vecteur*

$$\nu(s) = \frac{\tau'(s)}{\|\tau'(s)\|} = \frac{\gamma_N''(s)}{\|\gamma_N''(s)\|}. \tag{19.158}$$

Nous déduirons une formule plus pratique en dehors des coordonnées normales en (19.186).

*(3) La **courbure** au point $\gamma_N(s)$ est le réel*

$$c(s) = \|\tau'(s)\| = \|\gamma_N''(s)\|. \tag{19.159}$$

Note : il y a une notion de courbure signée qui sera donnée dans la définition 19.67.

*(4) Le **rayon de courbure** est le réel*

$$R(s) = \frac{1}{c(s)} = \frac{1}{\|\gamma_N''(s)\|}. \tag{19.160}$$

Par la proposition 19.51, nous avons $\nu(s) \cdot \tau(s) = 0$. En combinant toutes les formules, nous avons les différentes expressions suivantes pour le vecteur normal unitaire :

$$\nu(s) = \frac{\gamma_N''(s)}{c(s)} = \frac{\tau'(s)}{\|\tau'(s)\|} = \frac{\tau'(s)}{c(s)} = R(s)\tau'(s) = R(s)\gamma_N''(s). \tag{19.161}$$

Proposition 19.54.
La fonction courbure s'écrit $c = \|\gamma_N' \times \gamma_N''\|$.

Démonstration. Par le point (4) de la proposition 9.54, nous avons

$$\langle \gamma_N', \gamma_N'' \rangle^2 + \|\gamma_N' \times \gamma_N''\|^2 = \|\gamma_N'\|^2 \|\gamma_N''\|^2 = \|\gamma_N''\|^2 \tag{19.162}$$

parce que, la paramétrisation étant normale, $\|\gamma_N'\| = 1$. Mais $\langle \gamma_N', \gamma_N'' \rangle = 0$, donc il reste $\|\gamma_N' \times \gamma_N''\|^2 = \|\gamma_N''\|^2$, d'où

$$c(s) = \|\gamma_N''(s)\| = \|\gamma_N'(s) \times \gamma_N''(s)\| \tag{19.163}$$

pour chaque s dans J. □

Définition 19.55.
*Soit s un point birégulier (c'est à dire $\gamma_N'(s) \neq 0$ et $\gamma_N''(s) \neq 0$) de l'arc normal $\gamma = (J, \gamma_N)$. Le **vecteur unitaire de la binormale** est le vecteur*

$$\beta(s) = \tau(s) \times \nu(s) \tag{19.164}$$

Par leurs définitions, τ et ν sont unitaires, tandis que la proposition 19.51 montre qu'ils sont également orthogonaux. Les propriétés du produit vectoriel font que β est également unitaire, et simultanément orthogonal à τ et à ν.

Définition 19.56.
*Le repère orthonormal $\{\gamma_N(s), \tau(s), \beta(s)\}$ est le **repère de Frenet** au point $\gamma_N(s)$.*

Lemme 19.57.
Le vecteur unitaire normal est donné par $\nu(s) = \beta(s) \times \tau(s)$.

Démonstration. Ceci est une application de la formule d'expulsion (9.163) et de l'orthonormalité de la base de Frenet :

$$\beta \times \tau = (\tau \times \nu) \times \tau = -(\nu \cdot \tau)\tau + (\tau \cdot \tau)\nu = \nu. \tag{19.165}$$

□

19.8.1 Torsion

Décomposons le vecteur $\beta'(s)$ dans la base de Frenet. Pour cela nous allons utiliser la proposition 9.264 et montrer que $\beta'(s) \cdot \tau(s) = \beta'(s) \cdot \beta(s) = 0$, ce qui voudra dire que, dans la base de Frenet, les composantes de β' le long de τ et β sont nulles. Le vecteur β' sera donc colinéaire à ν.

D'abord, étant donné que la norme de $\beta(s)$ est constante par rapport à s, nous avons

$$0 = \frac{d}{ds}\|\beta(s)\|^2 = 2\beta'(s) \cdot \beta(s). \tag{19.166}$$

Ensuite, nous dérivons la définition $\beta(s) = \tau(s) \times \nu(s)$ en utilisant la formule de Leibnitz (9.162) :

$$\beta'(s) = \tau'(s) \times \nu(s) + \tau(s) \times \nu'(s). \tag{19.167}$$

Mais $\tau'(s) = \gamma_N''(s)$ tandis que $\nu(s) = \frac{\gamma_N''(s)}{\|\gamma_N''(s)\|}$, de telle sorte que $\tau'(s) \times \nu(s) = 0$. Nous restons donc avec $\beta'(s) = \tau(s) \times \nu'(s)$, ce qui prouve que $\beta'(s)$ est perpendiculaire à $\tau(s)$ et donc que $\beta'(s) \cdot \tau(s) = 0$.

Le vecteur $\beta'(s)$ est donc un multiple de $\nu(s)$. Nous notons $t(s)$ le facteur de proportionnalité :

$$\beta'(s) = t(s)\nu(s). \tag{19.168}$$

Définition 19.58.
*Soit $\gamma = (J, \gamma_N)$ un arc paramétré normal de classe $\mathcal{C}^3$. La **torsion** de γ au point $\gamma_N(s)$ est le réel*

$$t(s) = \|\beta'(s)\| = \|\tau(s) \times \nu'(s)\|. \tag{19.169}$$

*Lorsque $t(s) \neq 0$, le réel $T(s) = \frac{1}{t(s)}$ est le **rayon de torsion** de γ en $\gamma_N(s)$.*

Étant donné que pour chaque s, l'ensemble $\{\tau(s), \nu(s), \beta(s)\}$ est une base, il est naturel de vouloir décomposer leurs dérivées dans cette base. D'abord, par définition de c et de t, nous avons

$$\begin{aligned} \tau'(s) &= c(s)\nu(s) \\ \beta'(s) &= t(s)\nu(s). \end{aligned} \tag{19.170}$$

Il reste à décomposer $\nu'(s)$. Définissons α_τ, α_ν et α_β (qui peuvent dépendre de s) par

$$\nu'(s) = \alpha_\tau\tau(s) + \alpha_\nu\nu(s) + \alpha_\beta\beta(s). \tag{19.171}$$

En vertu de la proposition 9.264, nous avons

$$\begin{aligned} \alpha_\tau &= \langle \nu'(s), \tau(s)\rangle = -\langle \nu(s), \tau'(s)\rangle = -\langle \nu(s), c(s)\nu(s)\rangle = -c(s), \\ \alpha_\nu &= \langle \nu'(s), \nu(s)\rangle = 0, \\ \alpha_\beta &= \langle \nu'(s), \beta(s)\rangle = -\langle \nu(s), \beta'(s)\rangle = -t(s), \end{aligned} \tag{19.172}$$

où nous avons utilisé le fait que $\langle \nu(s), \nu(s)\rangle = \|\nu(s)\|^2 = 1$. Si nous mettons ces résultats sous forme matricielle, nous avons les **formules de Frenet** :

$$\begin{pmatrix} \tau'(s) \\ \nu'(s) \\ \beta'(s) \end{pmatrix} = \begin{pmatrix} 0 & c(s) & 0 \\ -c(s) & 0 & -t(s) \\ 0 & t(s) & 0 \end{pmatrix} \begin{pmatrix} \tau(s) \\ \nu(s) \\ \beta(s) \end{pmatrix}. \tag{19.173}$$

Proposition 19.59.
Si s est un point birégulier, alors la torsion est donnée par

$$t(s) = -\frac{(\gamma_N' \times \gamma_N'') \times \gamma_N'''}{\|\gamma_N'(s) \times \gamma_N''(s)\|^2}. \tag{19.174}$$

Démonstration. Par l'équation (19.161), nous avons $\gamma_N''(s) = c'(s)\nu(s)$, et par conséquent

$$\gamma_N'''(s) = c'(s)\nu(s) + c(s)\nu'(s) = c'(s)\nu(s) + c(s)\big[-c(s)\tau(s) - t(s)\beta(s)\big], \tag{19.175}$$

où nous avons utilisé la formule de Frenet pour $\nu'(s)$. Par ailleurs, sachant le corollaire 19.49 et la formule de Frenet pour τ', nous avons

$$\gamma_N' \times \gamma_N'' = \tau(s) \times \tau'(s) = \tau(s) \times c(s)\nu(s) = c(s)\beta(s). \tag{19.176}$$

En combinant les deux dernières équations, et en se souvenant que la base de Frenet et orthonormale,

$$(\gamma_N' \times \gamma_N'') \cdot \gamma_N'''(s) = -c(s)^2 t(s), \tag{19.177}$$

et donc, en remplaçant $c(s)$ par la formule (19.163),

$$t(s) = -\frac{(\gamma_N' \times \gamma_N'') \cdot \gamma_N'''}{\|\gamma_N' \times \gamma_N''\|^2}. \tag{19.178}$$

□

19.9 Hors des coordonnées normales

Remarque 19.60.
Notons que la définition de τ est donnée pour tout arc $\mathcal{C}^1$ régulier (I, γ) par $\tau(t) = \gamma'(t)/\|\gamma'(t)\|$. La propriété $\tau = \gamma_N'$ n'est valable que lorsque la paramétrisation est normale. Les autres définitions ont toutes été données dans le cas d'une paramétrisation normale.

La remarque 19.60 nous incite à exprimer toute la base de Frenet en terme de γ lorsque la paramétrisation n'est pas normale. Étant donné que nous pouvons toujours faire le changement de variable $\gamma(t) = \gamma_N\big(\phi(t)\big)$ (proposition 19.43), il est possible d'exprimer les vecteurs τ, ν et β ainsi que les réels c et t en fonction de γ et de ses dérivées.

Nous allons maintenant travailler à écrire les formules.

Pour plus de facilité, nous collectons les définitions. Afin d'alléger la notation, nous n'exprimons pas explicitement les dépendances en s :

Vecteur unitaire tangent Par le corollaire 19.49, τ est donné par $\tau = \gamma_N'$.

Vecteur unitaire normal Par la définition 19.53, ν est donné par $\nu = \frac{\tau'}{\|\tau'\|}$.

Vecteur unitaire de la binormale Par la définition 19.53, β est donné par $\beta = \tau \times \nu$.

Courbure Par la définition 19.53, c est donné par $c = \|\tau'\|$.

Torsion Par la définition 19.58, t est donné par $t = \|\beta'\|$.

Le schéma du changement de variable est

$$\begin{array}{ccc} t \in I & \xrightarrow{f} & \mathbb{R}^3 \\ {\scriptstyle\phi}\big\downarrow & \nearrow_{g} & \\ s \in J & & \end{array} \tag{19.179}$$

La difficulté ne sera pas d'éliminer γ_N de toutes les formule, mais bien de se débarrasser des fonctions ϕ qui arrivent quand nous exprimons γ_N en termes de γ, et en particulier lorsque nous voulons exprimer les dérivées de γ_N en termes de γ et de ses dérivées.

Regardons d'abord comment les dérivées de γ_N s'expriment en termes de γ. En utilisant le fait que $\gamma_N(s) = (\gamma \circ \phi^{-1})(s)$ et que $\|\gamma_N'(s)\| = 1$, nous avons

$$\gamma_N'(s) = \frac{\gamma_N'(s)}{\|\gamma_N'(s)\|} = \frac{(\gamma \circ \phi^{-1})'(s)}{\|(\gamma \circ \phi^{-1})'(s)\|} = \frac{\gamma'\big(\phi^{-1}(s)\big)(\phi^{-1})'(s)}{\|\gamma'\big(\phi^{-1}(s)\big)\||(\phi^{-1})'(s)|} = \frac{\gamma'(t)}{\|\gamma'(t)\|} \tag{19.180}$$

où nous avons utilisé le fait que ϕ^{-1} étant croissante (parce que l'inverse d'une fonction croissante est croissante), $(\phi^{-1})'(s) = |(\phi^{-1})'(s)|$. Pourquoi écrivons nous $|\phi^{-1}(s)|$ et non $\|\phi^{-1}(s)\|$?

Pour la dérivée seconde, nous dérivons la relation (19.180) :

$$\gamma_N''(s) = \frac{\gamma''\big(\phi^{-1}(s)\big)(\phi^{-1})'(s)}{\|\gamma'\big(\phi^{-1}(s)\big)\|} + \gamma'\big(\phi^{-1}(s)\big)\frac{d}{ds}\Big[\|\gamma'\big(\phi^{-1}(s)\big)\|\Big]. \tag{19.181}$$

Le petit calcul suivant va nous permettre de simplifier cette expression :

$$(\phi^{-1})'(s) = (\phi^{-1})'\big(\phi(t)\big) = \frac{1}{\phi'(t)} = \frac{1}{\|\gamma'(t)\|}. \tag{19.182}$$

Donc

$$\gamma_N''(s) = \frac{\gamma''(t)}{\|\gamma'(t)\|^2} + \gamma'(t)\frac{d}{ds}\Big[\|\gamma'(t)\|\Big] \tag{19.183}$$

où il est entendu que $t = \phi^{-1}(s)$. Avec cette expression, nous ne nous sommes pas encore débarrassés de la fonction ϕ, mais nous allons voir que cela nous sera suffisant.

Pour le vecteur unitaire tangent $\tau(s)$, nous avons donc immédiatement

$$\tau(s) = \gamma_N'(s) = \frac{\gamma'(t)}{\|\gamma'(t)\|}. \tag{19.184}$$

Ici encore il est sous-entendu que le t dans le membre de droite est lié au s du membre de gauche par $t = \phi^{-1}(s)$. Il est donc naturel de nous demander si nous avons gagné quelque chose, étant donné que la formule (19.184) contient encore la fonction ϕ.

Géométriquement, le vecteur $\tau(s)$ est le vecteur normal unitaire de la courbe au point $\gamma_N(s)$. En utilisant les relations du diagramme (19.179), nous avons en réalité $\gamma_N(s) = \gamma_N\big(\phi(t)\big) = \gamma(t)$. Le vecteur $\frac{\gamma'(t)}{\|\gamma'(t)\|}$ représente donc le vecteur normal tangent au point $\gamma(t)$.

Pour calculer la courbure, nous devons d'abord calculer le produit vectoriel

$$\begin{aligned}\gamma_N'(s) \times \gamma_N''(s) &= \frac{\gamma'(t)}{\|\gamma'(t)\|} \times \left(\frac{\gamma''(t)}{\|\gamma'(t)\|^2} + \gamma'(t)\frac{d}{ds}\Big[\|\gamma'(t)\|\Big]\right) \\ &= \frac{\gamma'(t) \times \gamma''(t)}{\|\gamma'(t)\|^3}\end{aligned} \tag{19.185}$$

parce que le deuxième terme dans la parenthèse est un multiple de $\gamma'(t)$, de telle sorte à ce que son produit vectoriel avec $\gamma'(t)/\|\gamma'(t)\|$ soit nul. En prenant la norme,

$$c(s) = \frac{\|\gamma'(t) \times \gamma''(t)\|}{\|\gamma'(t)\|^3}. \tag{19.186}$$

Encore une fois, cette équation nous enseigne que la courbure au point $\gamma(t) \in \mathbb{R}^3$ est donnée par le membre de droite, qui ne dépend que de t.

Le vecteur unitaire binormal est donné par $\beta(s) = \tau(s) \times \nu(s)$. En utilisant (19.184) et (19.161),

$$\beta(s) = \tau(s) \times \nu(s) = \gamma_N'(s) \times \frac{\gamma_N''(s)}{c(s)}. \tag{19.187}$$

Les formules (19.185) pour le produit vectoriel et (19.186) pour la courbure donnent ensuite

$$\beta(s) = \frac{\gamma'(t) \times \gamma''(t)}{\|\gamma'(t)\|^3} \cdot \frac{1}{c(s)} = \frac{\gamma'(t) \times \gamma''(t)}{\|\gamma'(t) \times \gamma''(t)\|}. \tag{19.188}$$

Cela donne le vecteur unitaire binormal au point $\gamma(t)$ en terme de $\gamma'(t)$ et $\gamma''(t)$.

La torsion demande d'utiliser la dérivée troisième de γ_N. Nous avons

$$\begin{aligned}\gamma_N'''(s) &= (\gamma\circ\phi^{-1})'''(s)\\ &= \Big(\gamma'\big(\phi^{-1}(s)\big)(\phi^{-1})'(s)\Big)''\\ &= \Big(\gamma''\big(\phi^{-1}(s)\big)(\phi^{-1})'(s)^2+\gamma'\big(\phi^{-1}(s)\big)(\phi^{-1})''(s)\Big)'\\ &= \gamma'''\big(\phi^{-1}(s)\big)(\phi^{-1})'(s)^3+v\\ &= \frac{\gamma'''\big(\phi^{-1}(s)\big)}{\|\gamma'(t)\|^3}+v \qquad \text{par (19.182)}\end{aligned} \tag{19.189}$$

où v est un élément de $\langle\gamma''\big(\phi^{-1}(s)\big),\gamma'\big(\phi^{-1}(s)\big)\rangle$. Le vecteur v est donc perpendiculaire à $\gamma'\times\gamma''$ et donc à $\gamma_N'\times\gamma_N''$ à cause de la relation (19.185) qui montre que $\gamma'\times\gamma''$ est parallèle à $\gamma_N'\times\gamma_N''$. De ce fait, lorsque nous calculons $(\gamma_N'\times\gamma_N'')\cdot\gamma_N'''$, la partie v de γ_N''' n'entre pas en ligne de compte.

Nous avons donc le calcul suivant, en remplaçant les diverses occurrences de $\gamma_N'\times\gamma_N''$ par sa valeur (19.185) en termes de γ,

$$\begin{aligned}t(s) &= -\frac{(\gamma_N'\times\gamma_N'')\cdot\gamma_N'''}{\|\gamma_N'\times\gamma_N''\|^2}\\ &= -\frac{(\gamma_N'\times\gamma_N'')\cdot\gamma'''(t)}{\|\gamma_N'\times\gamma_N''\|^2\,\|\gamma'(t)\|^2}\\ &= -\frac{(\gamma'\times\gamma'')\cdot\gamma'''}{\|\gamma'\times\gamma''\|^2}.\end{aligned} \tag{19.190}$$

Dans cette expression, il est sous-entendu que tous les γ_N sont fonctions de s et tous les γ sont fonction de t où s et t sont liés par $s=\phi(t)$.

Ce que nous avons prouvé est le

Théorème 19.61.
Pour tout représentant (I,γ), les éléments métriques (τ,ν,β,c,t) au point $\gamma(t)$ s'expriment en fonction de $\gamma(t)$, $\gamma'(t)$, $\gamma''(t)$ et $\gamma'''(t)$.

Lemme 19.62.
Si γ est le graphe de la fonction y alors la courbure de γ est donnée par la formule

$$c\big(\gamma(t)\big)=\frac{|y''(t)|}{\big(1+y'(t)^2\big)^{3/2}} \tag{19.191}$$

Démonstration. Nous avons :

$$\gamma(t)=\big(t,y(t)\big) \tag{19.192a}$$
$$\gamma'(t)=\big(1,y'(t)\big) \tag{19.192b}$$
$$\gamma''(t)=\big(0,y''(t)\big). \tag{19.192c}$$

Il s'agit maintenant seulement d'utiliser la formule (19.186) en se souvenant comment on calcule un produit vectoriel [5] :

$$\gamma'\times\gamma''=\begin{vmatrix}e_1 & e_2 & e_3\\ 1 & y' & 0\\ 0 & y'' & 0\end{vmatrix}=y''. \tag{19.193}$$

□

5. Définition 9.52.

19.10 Tracer des courbes paramétriques dans $\mathbb{R}^2$

Nous allons maintenant voir comment les concepts introduits nous aident à effectivement tracer des courbes dans le plan. Les courbes que nous regardons sont de la forme $\gamma(t) = \big(x(t), y(t)\big)$, et nous supposons que ces fonctions soient suffisamment régulières (disons trois fois continument dérivables). Nous ne supposons pas que la courbe soit donnée en coordonnées normales, en particulier, $\gamma''(t)$ n'est pas le vecteur normal en $\gamma(t)$.

La notion clef qui va jouer est le **cercle osculateur** de la courbe γ au point $\gamma(t)$. Sans rentrer dans les détails, disons que c'est le cercle qui «colle» le mieux possible la courbe. Le rayon de ce cercle est le rayon de courbure :

$$R(t) = \frac{\|\gamma(t)\|^3}{\|\gamma'(t) \times \gamma''(t)\|}. \tag{19.194}$$

En pratique, le produit vectoriel se calcule comme ceci :

$$\gamma'(t) \times \gamma''(t) = \begin{vmatrix} e_x & e_y & e_z \\ x'(t) & y'(t) & 0 \\ x''(t) & y''(t) & 0 \end{vmatrix} = (x'y'' - x''y')e_z. \tag{19.195}$$

Le centre du cercle osculateur va se trouver quelque part sur la normale. Le vecteur normal est donné par

$$n(t) = J\frac{\gamma'(t)}{\|\gamma'(t)\|} \tag{19.196}$$

où J est la rotation d'angle $\frac{\pi}{2}$:

$$J\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} y'(t) \\ -x'(t) \end{pmatrix}. \tag{19.197}$$

Cela nous laisse deux possibilités pour le centre du cercle osculateur : $\gamma(t) + R(t)n(t)$ ou bien $\gamma(t) - R(t)n(t)$. Il faut savoir de quel côté de la courbe est situé le centre du cercle osculateur. Il faut choisir le côté de la concavité, c'est à dire le côté de la dérivée seconde.

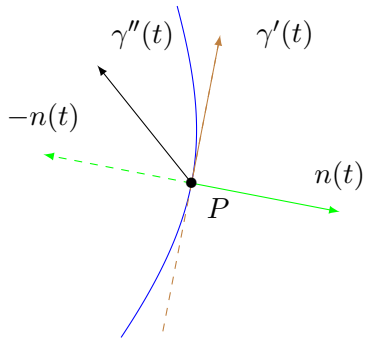


FIGURE 19.8 – De quel coté de $\gamma'(t)$ se trouvent $n(t)$ et $-n(t)$?

La difficulté maintenant est de savoir qui de $n(t)$ ou $-n(t)$ est du côté de $\gamma''(t)$. Il faut savoir si $n(t)$ est du même côté de la droite tangente que $\gamma''(t)$ ou non. Par construction, si nous regardons la figure 19.8, le vecteur $n(t)$ sera toujours à gauche de $\gamma'(t)$. Le fait que $\gamma''(t)$ soit à gauche ou à droite de $\gamma'(t)$ est donné par le signe du produit vectoriel $\gamma'(t) \times \gamma''(t)$. Si ce produit vectoriel est positif, il faut choisir $-n(t)$ et s'il est négatif, il faut choisir $n'(t)$.

Le truc pour obtenir le signe de $x'y'' - x''y'$ est de faire

$$\frac{(\gamma' \times \gamma'') \cdot e_z}{\|\gamma' \times \gamma''\|}. \tag{19.198}$$

Le centre de courbure sera donc situé à la position

$$\Omega(t) = \gamma(t) - n(t)\frac{\|\gamma(t)\|^3}{\|\gamma'(t) \times \gamma''(t)\|^2}(\gamma' \times \gamma'') \cdot e_z \tag{19.199}$$

Nous pouvons écrire cela plus explicitement en nous souvenant que $\gamma' \times \gamma'' = (x'y'' - x''y')e_z$, par conséquent $\frac{(\gamma' \times \gamma'') \cdot e_z}{\|\gamma' \times \gamma''\|^2} = \frac{1}{x'y'' - x''y'}$. Nous avons

$$\Omega_x(t) = x(t) - y'(t)\frac{x'^2 + y'^2}{x'y'' - x''y'} \tag{19.200a}$$

$$\Omega_y(t) = y(t) + x'(t)\frac{x'^2 + y'^2}{x'y'' - x''y'}. \tag{19.200b}$$

Quelques exemples de cercles osculateurs sont sur la figure 19.9.

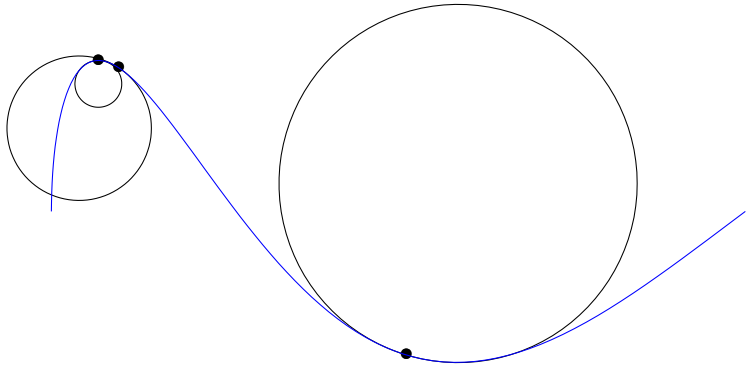


FIGURE 19.9 – Exemple de cercles osculateurs.

19.11 Courbes planes

Définition 19.63.
*Une courbe $\gamma\colon [a,b] \to \mathbb{R}^n$ est **fermée** si $\gamma(a) = \gamma(b)$. Elle est **simple** si $\gamma(t) \neq \gamma(t')$ dès que $t, t' \in]a,b[$ et $t \neq t'$.*

Définition 19.64.
Nous disons qu'une courbe fermée est continue, de classe C^1, de classe C^2 ou autre condition de régularité si son extension périodique comme application $\gamma\colon \mathbb{R} \to \mathbb{R}^2$ a cette régularité.

19.11.1 Angle

Lemme 19.65 ([198]).
Soient des courbes régulières γ et σ de classe C^2 de l'intervalle ouvert I vers $\mathbb{R}^2$. Soit $\theta_0 \in \mathbb{R}$ tel que

$$\frac{\gamma'(t_0) \cdot \sigma'(t_0)}{\|\gamma'(t_0)\|\|\sigma'(t_0)\|} = \cos(\theta_0) \tag{19.201a}$$

$$\frac{\gamma'(t_0) \cdot J\sigma'(t_0)}{\|\gamma'(t_0)\|\|\sigma'(t_0)\|} = \cos(\theta_0). \tag{19.201b}$$

Alors il existe une unique fonction différentiable $\theta\colon I \to \mathbb{R}$

$$\frac{\gamma'(t) \cdot \sigma'(t)}{\|\gamma'(t)\|\|\sigma'(t)\|} = \cos\big(\theta(t)\big) \tag{19.202a}$$

$$\frac{\gamma'(t) \cdot J\sigma'(t)}{\|\gamma'(t)\|\|\sigma'(t)\|} = \sin\big(\theta(t)\big). \tag{19.202b}$$

Démonstration. Il suffit de prendre

$$f(t) = \frac{\gamma'(t) \cdot \sigma'(t)}{\|\gamma'(t)\|\|\sigma'(t)\|} \tag{19.203}$$

et

$$g(t) = \frac{\gamma'(t) \cdot J\sigma'(t)}{\|\gamma'(t)\|\|\sigma'(t)\|} \tag{19.204}$$

dans la proposition 16.34. Ces courbes sont de classe C^1 parce que γ et σ sont de classe C^2. □

19.11.2 Courbure signée

Nous avons déjà défini la courbure d'une courbe en la définition 19.53. Nous introduisons maintenant la courbure signée qui est propre à la dimension deux.

Définition 19.66.
*La **structure complexe** sur $\mathbb{R}^2$ est l'application*

$$\begin{aligned} J\colon \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x,y) &\mapsto (-y,x). \end{aligned} \tag{19.205}$$

Définition 19.67.
*La **courbure signée** de la courbe $\gamma\colon I \to \mathbb{R}^2$ (I est un intervalle dans $\mathbb{R}$) est la fonction*

$$\kappa(t) = \frac{\gamma''(t) \cdot J\gamma'(t)}{\|\gamma'(t)\|^3} \tag{19.206}$$

où J est la structure complexe de la définition 19.66.

Cette définition est motivée par le fait qu'en identifiant $\mathbb{R}^2$ à $\mathbb{C}$, l'application J revient à l'application $z \mapsto iz$.

Si $v, w \in \mathbb{R}^2$ nous avons formellement

$$v \times w = -(v \cdot Jw)e_3. \tag{19.207}$$

En particulier pour tout $v \in \mathbb{R}^2$ nous avons

$$v \cdot Jv = 0. \tag{19.208}$$

Lemme 19.68.
Soir une courbe régulière $\gamma\colon [a,b] \to \mathbb{R}^2$ et un difféomorphisme $h\colon [c,d] \to [a,b]$. Si nous posons $\sigma = \gamma \circ h$ alors

$$\kappa_\sigma(u) = \operatorname{sgn}\big(h'(u)\big)\kappa_\gamma\big(h(u)\big). \tag{19.209}$$

Démonstration. Nous utilisons la définition (19.206) de la courbure signée. La règle de dérivation en chaîne donne :

$$\sigma'(u) = \gamma'\big(h(u)\big)h'(u) \tag{19.210a}$$

$$\sigma''(u) = \gamma''\big(h(u)\big)h'(u)^2 + \gamma'\big(h(u)\big)h''(u). \tag{19.210b}$$

La numérateur de $\kappa_\sigma(u)$ est :

$$(\gamma'' \circ h)h'^2 \cdot J(\gamma' \circ h)h' + h''(\gamma' \circ h) \cdot J(\gamma' \cdot h)h' \tag{19.211}$$

dont le second terme est nul parce que $v \cdot Jv = 0$. Il nous reste donc

$$\kappa_\sigma(u) = \frac{(h')^3}{|h'|^3}\frac{(\gamma'' \circ h) \cdot J(\gamma' \circ h)}{\|\gamma' \circ h\|^3} = \operatorname{sgn}(h')\kappa_\gamma\big(h(u)\big). \tag{19.212}$$

□

Lemme 19.69 ([198])**.**
Si γ_N est un arc paramétré normal, alors

$$\gamma_N''(s) = \kappa(s)J\gamma_N'(s). \tag{19.213}$$

Démonstration. Vu que la paramétrisation est normale, $\gamma'_N \cdot \gamma'_N = 1$, et en dérivant, $\gamma''_N \cdot \gamma'_N = 0$. Donc γ''_N est un multiple de $J\gamma'$. En tenant compte du fait que la paramétrisation est normale, la courbure est

$$\kappa(s) = \gamma''_N(s) \cdot J\gamma'(s). \tag{19.214}$$

En y injectant $\gamma''(s) = \lambda(s)J\gamma'(s)$ nous trouvons

$$\kappa(s) = \lambda(s)J\gamma'(s) \cdot J\gamma'(s) = \lambda(s). \tag{19.215}$$

Donc le facteur de proportionnalité est $\kappa(s)$. □

Théorème 19.70.
Soit une courbe $\gamma\colon I \to \mathbb{R}^2$ de classe C^2.

(1) γ est une partie de droite si et seulement si $\kappa(t) = 0$ pour tout t.

(2) γ est une partie d'un cercle de rayon $r > 0$ si et seulement si $|\kappa(s)| = \frac{1}{r}$.

Démonstration. Si γ est une droite, la dérivée seconde est nulle et la courbure est nulle. Supposons pour la réciproque que $\kappa(t) = 0$ pour tout t. Nous utilisant une paramétrisation normale de γ, ce qui ne change pas que la courbure reste nulle. Nous avons par le lemme 19.69 que $\gamma''(s) = 0$ et donc l'existence de $a, b \in \mathbb{R}^2$ tels que $\gamma(t) = at + b$.

Passons au cas du cercle. Si γ est un cercle, la paramétrisation normale est

$$\gamma(t) = R\begin{pmatrix}\cos(\frac{t}{R})\\ \sin(\frac{t}{R})\end{pmatrix} \tag{19.216a}$$

$$\gamma'(t) = \begin{pmatrix}-\sin(\frac{t}{R})\\ \cos(\frac{t}{R})\end{pmatrix} \tag{19.216b}$$

$$\gamma''(t) = -\frac{1}{R}\begin{pmatrix}\cos(\frac{t}{R})\\ \sin(\frac{t}{R})\end{pmatrix}. \tag{19.216c}$$

Avec tout cela nous avons $\kappa(s) = \gamma''(s) \cdot J\gamma'(s) = \frac{1}{R}$.

Nous supposons enfin que $\kappa(t) = 1/R$ et que le paramétrage soit normal (encore une fois, un reparamétrage ne change pas la courbure lorsqu'elle est constante). Nous définissons la courbe

$$\begin{aligned}\beta\colon [b, c] &\to \mathbb{R}^2\\ t &\mapsto \gamma(t) + rJ\gamma'(t).\end{aligned} \tag{19.217}$$

Nous avons $\beta'(t) = \gamma'(t) + rJ\gamma''(t)$. Mais par le lemme 19.69 nous avons $\gamma'' = kJ\gamma' = \frac{1}{r}J\gamma'$. Donc

$$\beta'(t) = \gamma'(t) - \gamma'(t) = 0. \tag{19.218}$$

Du coup β est constante : $\beta(t) = a$. Alors $a = \gamma(t) + rJ\gamma'(t)$ et en particulier

$$\|\gamma(t) - a\| = \|rJ\gamma'(t)\| = r. \tag{19.219}$$

Donc effectivement γ reste sur un cercle de rayon r et de centre a. □

Définition 19.71 ([198])**.**
*La **courbure totale** de la courbe $\gamma\colon [a, b] \to \mathbb{R}^2$ est le nombre*

$$K = \int_a^b \kappa(t)\|\gamma'(t)\|dt. \tag{19.220}$$

Lemme 19.72.
La courbure signée ne change pas sous reparamétrisation positive, et change de signe sous reparamétrisation négative.

Démonstration. Soit la courbe $\gamma\colon [a,b] \to \mathbb{R}^2$ et un difféomorphisme $h\colon [a,b] \to [c,d]$. Il s'agit d'intégrer la relation (19.209) en effectuant le changement de variables $t = h(u)$:

$$K_\gamma = \int_a^b \kappa_\gamma(t)\|\gamma'(t)\|dt \tag{19.221a}$$

$$= \int_c^d \underbrace{\kappa_\gamma\big(h(u)\big)}_{\kappa_\sigma(u)} \|\gamma'\big(h(u)\big)\|h'(u)du. \tag{19.221b}$$

En utilisant le fait que $\sigma'(u) = \gamma'\big(h(u)\big)h'(u)$ nous avons alors

$$K_\gamma = \int_c^d \kappa_\sigma(u) \left\| \frac{\sigma'(u)}{h'(u)} \right\| h'(u)du \tag{19.222a}$$

$$= \int_c^d \kappa_\sigma(u)\|\sigma'(u)\| \frac{h'(u)}{|h'(u)|}du \tag{19.222b}$$

$$= \operatorname{sgn}(h')K_\sigma. \tag{19.222c}$$

□

Lemme-définition 19.73 ([198])**.**
Soit une courbe régulière $\gamma\colon I \to \mathbb{R}^2$ de classe C^2 et t_0 dans l'intérieur de I. Soit $\theta_0 \in \mathbb{R}$ tel que

$$\frac{\gamma'(t_0)}{\|\gamma'(t_0)\|} = \big(\cos(\theta_0), \sin(\theta_0)\big). \tag{19.223}$$

Alors il existe une unique fonction différentiable $\theta\colon I \to \mathbb{R}$ telle que $\theta(t_0) = \theta_0$ et

$$\frac{\gamma'(t)}{\|\gamma'(t)\|} = \begin{pmatrix} \cos\big(\theta(t)\big) \\ \sin\big(\theta(t)\big) \end{pmatrix} \tag{19.224}$$

pour tout $t \in I$.

*Cette fonction est l'**angle** de γ déterminé par θ_0.*

Démonstration. Soit $\beta(t) = (t, 0)$; alors $\beta'(t) = (1, 0)$ et nous avons

$$\gamma' \cdot \beta' = \gamma'_x \tag{19.225a}$$

$$\gamma' \cdot J\beta' = \gamma'_y. \tag{19.225b}$$

Par la proposition 19.65 il existe une unique fonction $\theta\colon I \to \mathbb{R}$ telle que $\theta(t_0) = \theta_0$ et

$$\begin{cases} \cos\big(\theta(t)\big) = \dfrac{\gamma' \cdot \beta'}{\|\gamma'\|\|\beta'\|} = \dfrac{\gamma'_x}{\|\gamma'\|} & (19.226a) \\ \sin\big(\theta(t)\big) = \dfrac{\gamma' \cdot J\beta'}{\|\gamma'\|\|\beta'\|} = \dfrac{\gamma'_y}{\|\gamma'\|}. & (19.226b) \end{cases}$$

Une telle fonction est bien celle que l'on demande ici. □

Lemme 19.74.
Si γ est une courbe régulière de classe C^2, alors sa courbure et son angle vérifient la relation

$$\theta'(t) = \|\gamma'(t)\|\kappa(t). \tag{19.227}$$

Démonstration. Par définition de l'angle (lemme 19.73) nous avons

$$\frac{\gamma'(t)}{\|\gamma'(t)\|} = \big(\cos\theta(t), \sin\theta(t)\big). \tag{19.228}$$

Dérivant cela,

$$\frac{\gamma''(t)}{\|\gamma'(t)\|} + \gamma'(t)\frac{d}{dt}\left(\frac{1}{\|\gamma'(t)\|}\right) = \theta'(t)\begin{pmatrix}-\sin\big(\theta(t)\big)\\ \cos\big(\theta(t)\big)\end{pmatrix} \tag{19.229a}$$

$$= \theta'(t)J\begin{pmatrix}\cos\big(\theta(t)\big)\\ \sin\big(\theta(t)\big)\end{pmatrix} \tag{19.229b}$$

$$= \theta'(t)\frac{J\gamma'(t)}{\|\gamma'(t)\|}. \tag{19.229c}$$

Nous prenons le produit scalaire de cette égalité avec $J\gamma'$ en tenant compte du fait que $\gamma' \cdot J\gamma' = 0$:

$$\frac{\gamma'' \cdot J\gamma'}{\|\gamma'\|} = \frac{\theta'}{\|\gamma'\|}J\gamma' \cdot J\gamma'. \tag{19.230}$$

En remarquant que $Jv \cdot Jv = \|v\|^2$ nous trouvons $\theta'\|\gamma'\| = \|\gamma'\|^2\kappa_\gamma$ et donc

$$\theta'(t) = \|\gamma'(t)\|\kappa_\gamma(t), \tag{19.231}$$

ce qu'il fallait prouver. □

Ce lemme nous fournit la formule attendue pour la courbure totale.

Lemme 19.75 ([198])**.**
Soit une courbe régulière de classe C^2 $\gamma\colon [a,b] \to \mathbb{R}^2$. Sa courbure totale est donnée par

$$K = \theta(b) - \theta(a). \tag{19.232}$$

Démonstration. Il suffit de remplacer dans la définition (19.220) de la courbure totale l'intégrante par son expression du lemme 19.74 :

$$K = \int_a^b \kappa(t)\|\gamma'(t)\|dt = \int_a^b \theta'(t)dt = \theta(b) - \theta(a) \tag{19.233}$$

par le théorème 13.205. □

19.11.3 Degré, indice et homotopie

Définition 19.76 ([198])**.**
*Le **nombre de tours** d'une courbe fermée de classe C^1 $\gamma\colon [a,b] \to \mathbb{R}^2$ est le nombre*

$$\mathrm{Turn}(\gamma) = \frac{1}{2\pi}\int_a^b \kappa(t)\|\gamma'(t)\|dt \tag{19.234}$$

où κ est la courbure signée définie en 19.67.

Lemme-définition 19.77.
*Soit une application continue $\phi\colon S^1 \to S^1$. Un **relèvement** de σ est une application $\varphi\colon \mathbb{R} \to \mathbb{R}$ une application continue telle que*

$$\phi\big(\cos(t), \sin(t)\big) = \begin{pmatrix}\cos\big(\varphi(t)\big)\\ \sin\big(\varphi(t)\big)\end{pmatrix}. \tag{19.235}$$

*Le **degré** est l'entier $\deg(\phi)$ tel que*

$$\varphi(2\pi) - \varphi(0) = 2\deg(\phi)\pi. \tag{19.236}$$

Ce nombre ne dépend pas du choix du relèvement φ.

Démonstration. Soient φ_1 et φ_2, deux application qui satisfont les contraintes. Nous avons une fonction continue $n\colon S^1 \to \mathbb{R}$ telle que

$$\varphi_1(t) - \varphi_2(t) = 2\pi n(t). \tag{19.237}$$

La fonction n ne pouvant prendre que des valeurs entières et étant continue, elle est constante. Par conséquent

$$\varphi_1(2\pi) - \varphi_1(0) = \varphi_2(2\pi) - \varphi_2(0), \tag{19.238}$$

ce que nous voulions. □

Définition 19.78.
Soit une courbe fermée $\gamma\colon [0,L] \to \mathbb{R}^2$ de classe C^1. Nous posons

$$\Phi_\gamma(t) = \frac{\tilde{\gamma}'(t)}{\|\tilde{\gamma}'(t)\|} \tag{19.239}$$

*où $\tilde{\gamma}(u) = \gamma\big(\frac{L}{2\pi}u\big)$. L'**indice de rotation** de γ est le degré de Φ_γ, c'est à dire*

$$\mathrm{Ind}(\gamma) = \deg(\Phi_\gamma). \tag{19.240}$$

Notons que dans cette définition, Φ_γ n'est rien d'autre que le vecteur unitaire tangent à γ, ramené à $[0, 2\pi]$.

Proposition 19.79.
Pour une courbe fermée de classe[6] C^2, l'indice de rotation est égal au nombre de tours.

Démonstration. Soit $\gamma\colon [0,L] \to \mathbb{R}^2$ une courbe vérifiant les hypothèses. Par définition, $\mathrm{Ind}(\gamma) = \deg(\Phi_\gamma)$ où

$$\Phi_\gamma(t) = \frac{\tilde{\gamma}'(t)}{\|\tilde{\gamma}'(t)\|} = \begin{pmatrix} \cos\tilde{\theta}(t) \\ \sin\tilde{\theta}(t) \end{pmatrix} \tag{19.241}$$

où nous avons noté $\tilde{\theta}$ l'angle tournant[7] de $\tilde{\gamma}$ et θ celui de γ. Vu que $\tilde{\theta}$ vérifie les hypothèses de la définition 19.77 nous pouvons calculer le degré de Φ_γ par

$$\deg(\Phi_\gamma) = \frac{1}{2\pi}\big(\tilde{\theta}(2\pi) - \tilde{\theta}(0)\big) = \frac{1}{2\pi}\int_0^{2\pi} \tilde{\theta}'(s)ds. \tag{19.242}$$

Il faut trouver le lien entre θ et $\tilde{\theta}$. Pour cela nous notons que

$$\frac{\tilde{\gamma}'(t)}{\|\tilde{\gamma}'(t)\|} = \frac{\gamma'\left(\frac{L}{2\pi}t\right)}{\|\gamma'\left(\frac{L}{2\pi}\right)\|}. \tag{19.243}$$

En comparant avec (19.241) il vient

$$\tilde{\theta}(t) = \theta\big(\frac{L}{2\pi}t\big) \tag{19.244}$$

et

$$\tilde{\theta}'(s) = \frac{L}{2\pi}\theta'\left(\frac{L}{2\pi}s\right). \tag{19.245}$$

Le changement de variables $t = \frac{L}{2\pi}s$ est donc tout vu dans l'intégrale (19.242) :

$$\deg(\Phi_\gamma) = \frac{1}{2\pi}\int_0^{2\pi} \frac{L}{2\pi}\theta'\big(\frac{L}{2\pi}s\big)ds = \frac{1}{2\pi}\int_0^{L} \theta'(t)dt = \frac{1}{2\pi}\int_0^{L} \kappa_\gamma(t)\|\gamma'(t)\|dt = \mathrm{Turn}(\gamma) \tag{19.246}$$

où nous avons aussi utilisé le lemme 19.74 qui donne le lien entre θ' et κ. □

6. La dérivée seconde arrive dans la définition de la courbure ; il faudrait donc supposer au moins C^2 pour avoir la continuité de la courbure.
7. Définition 19.73.

Définition 19.80 (homotopie de chemins fermés)**.**
Les courbes fermées $\gamma_0, \gamma_1 \colon [0, L] \to Y$ (Y est un espace topologique) sont ***homotopes*** *s'il existe une application continue*

$$F \colon [0,1] \times [0, L] \to Y \tag{19.247}$$

telle que

(1) $F(0,t) = \gamma_0(t)$ *pour tout* t*,*

(2) $F(1,t) = \gamma_1(t)$ *pour tout* t*,*

(3) $F(u, L) = F(u, 0)$ *pour tout* u*.*

L'application F est l'homotopie entre γ_0 et γ_1.

Proposition 19.81 (Homotopie, degré et indice[198])**.**
Il y a deux résultats à ne pas confondre.

(1) Si $\phi_i \colon S^1 \to S^1$ *sont homotopes, alors* $\deg(\phi_1) = \deg(\phi_2)$*.*

(2) Si $\gamma_i \colon [0, L] \to \mathbb{R}^2$ *sont homotopes, alors* $\mathrm{Ind}(\phi_1) = \mathrm{Ind}(\phi_2)$*.*

Démonstration. Nous allons décomposer la preuve en deux parties.

Le degré pour les applications $S^1 \to S^1$ Nous supposons avoir deux applications homotopes $\gamma_i \colon S^1 \to S^1$. Si F est l'homotopie [8] entre γ_0 et γ_1, nous posons $\gamma_u(t) = F(u,t)$ qui est encore une courbe fermée $\gamma_u \colon S^1 \to S^2$. Nous pouvons donc considérer le degré de γ_u. C'est un entier n_u qui vérifie

$$\tilde{\gamma}_u(2\pi) - \tilde{\gamma}_u(0) = 2\pi n_u. \tag{19.248}$$

Le membre de gauche est une fonction continue de u ; dans le membre de droite n_u ne pouvant prendre que des valeurs entières, elle est alors constante.

Indice de rotation pour des courbes dans $\mathbb{R}^2$ Soient maintenant γ_0 et γ_1 deux courbes homotopes de $[0, L]$ dans $\mathbb{R}^2$. Par définition, $\mathrm{Ind}(\gamma_i) = \deg(\Phi_{\gamma_i})$. Prouvons alors que Φ_{γ_0} et homotope à Φ_{γ_1}. De cette façon, la première partie de la preuve conclura à

$$\mathrm{Ind}(\gamma_0) = \deg(\Phi_{\gamma_0}) = \deg(\Phi_{\gamma_1}) = \mathrm{Ind}(\gamma_1). \tag{19.249}$$

Nous savons que

$$\Phi_\gamma = \frac{\gamma'\left(\frac{L}{2\pi}t\right)}{\|\gamma'\left(\frac{L}{2\pi}t\right)\|}, \tag{19.250}$$

donc en posant

$$\Phi_F(u,t) = \frac{\gamma_u'\left(\frac{L}{2\pi}t\right)}{\|\gamma_u'\left(\frac{L}{2\pi}t\right)\|} \tag{19.251}$$

nous avons une homotopie entre Φ_{γ_0} et Φ_{γ_1}.

□

Théorème 19.82 ([198])**.**
Le nombre de tours d'une courbe simple fermée de classe C^2 dans $\mathbb{R}^2$ est ± 1.

Démonstration. Soit une telle courbe γ et son image Γ.

Choix d'un point et d'une tangente Si ℓ est une droite dans le plan, soit p le point de Γ le plus proche de ℓ. Il n'est peut-être pas unique, mais la parallèle à ℓ passant par p est une tangente à Γ telle que tout Γ se trouve d'un seul côté de ℓ_p. Pour voir cela, il suffit de choisir un système d'axes pour lequel ℓ est l'axe $y = 0$. La distance entre ℓ et les points de Γ est donnée par γ_y, et donc les extrema sont atteints là où $\gamma_y' = 0$, c'est à dire pour les points sur lesquels la tangente est parallèle à ℓ.

Étant donné qu'il y a (au moins) un maximum et un minimum distincts, pour pouvons choisir le point p de telle sorte que pour tout point $q \in \Gamma$, l'angle du vecteur $q - p$ soit entre

8. Définition 19.80.

α_0 et $\alpha_0 + \pi$ et non entre $\alpha_0 - pi$ et α_0. Ce choix revient à choisir p de telle sorte que Γ soit d'un côté ou de l'autre de ℓ_p.

Pour simplifier les notations plus tard nous choisissons ℓ_p horizontale, de telle sorte que $\alpha_0 = 0$, et que Γ est au dessus de ℓ_p. Les angles des vecteurs $q - p$ pour $q \in \Gamma$ sont donc tous entre 0 et π.

Définition de Σ Soit L la longueur[9] de γ et β, une paramétrisation de γ telle que $\beta(0) = p$. Nous considérons le triangle

$$\mathcal{T} = \{(t_1, t_2) \in \mathbb{R}^2 \text{ tel que } 0 \leqslant t_1 \leqslant t_2 \leqslant L\} \tag{19.252}$$

et l'application sécante $\Sigma\colon \mathcal{T} \to S^1$ définie par

$$\Sigma(t_1, t_2) = \begin{cases} \dfrac{\beta'(t)}{\|\beta'(t)\|} & \text{si } t_1 = t_2 = t \\ -\dfrac{\beta'(0)}{\|\beta'(0)\|} & \text{si } t_1 = 0 \text{ et } t_2 = L \\ \dfrac{\beta(t_2) - \beta(t_1)}{\|\beta(t_2) - \beta(t_1)\|} & \text{sinon.} \end{cases} \tag{19.253}$$

Continuité de Σ Nous devons prouver les limites[10] suivantes :

$$\lim_{\substack{(t_1,t_2)\to(t,t)\\ 0\leqslant t_1\leqslant t\leqslant t_2\leqslant L}} \frac{\beta(t_1) - \beta(t_2)}{\|\beta(t_1) - \beta(t_2)\|} = \frac{\beta'(t)}{\|\beta'(t)\|} \tag{19.254a}$$

$$\lim_{\substack{(t_1,t_2)\to(0,L)\\ 0\leqslant t_1\leqslant t\leqslant t_2\leqslant L}} \frac{\beta(t_1) - \beta(t_2)}{\|\beta(t_1) - \beta(t_2)\|} = -\frac{\beta'(0)}{\|\beta'(0)\|} \tag{19.254b}$$

Nous commençons par (19.254a) en multipliant et divisant par $t_2 - t_1$:

$$\frac{\beta(t_2) - \beta(t_1)}{\|\beta(t_2) - \beta(t_1)\|} = \frac{\beta(t_2) - \beta(t_1)}{t_2 - t_1} \frac{t_2 - t_1}{\|\beta(t_2) - \beta(t_1)\|}. \tag{19.255}$$

Si chacune des limites des deux facteurs existent dans $\mathbb{R}$, la limite du produit sera le produit des limites. Pour le premier facteur nous développons $\beta(t_2)$ autour de $t = t_1$ via la formulation (12.678) :

$$\beta(t_2) = \beta(t_1) + (t_2 - t_1)\beta'(t_1) + \alpha(t_2 - t_1) \tag{19.256}$$

où α est une fonction ayant la propriété $\lim_{t\to 0} \frac{\alpha(t)}{t} = 0$. Nous avons à calculer la limite de

$$\frac{(t_1 - t_1)\beta'(t_1) + \alpha(t_2 - t_1)}{t_2 - t_1} = \beta'(t_1) + \frac{\alpha(t_2 - t_1)}{t_2 - t_1}. \tag{19.257}$$

Prendre la limite $(t_1, t_2) \to (t, t)$ donne bien $\beta'(t)$ parce que β est de classe C^1. En ce qui concerne la limite de la deuxième partie, nous allons la faire plus en détail pour la limite (19.254b).

La limite (19.254b). Nous prenons la prolongation périodique de β. Alors si $t_2 = L - \epsilon$ nous pouvons écrire $\beta(-\epsilon)$ au lieu de $\beta(t_2)$. Nous développons $\beta(t_1)$ autour de $t = -\epsilon$ (parce que t_1 est petit) :

$$\beta(t_1) = \beta(-\epsilon) + (t_1 + \epsilon)\beta'(-\epsilon) + \alpha(t_1 + \epsilon). \tag{19.258}$$

Après multiplication et division par $t_1 + \epsilon$, la première limite à calculer est celle de

$$\frac{\beta(-\epsilon) - \beta(t_1)}{t_1 + \epsilon} = -\beta'(-\epsilon) + \frac{\alpha(t_1 + \epsilon)}{t_1 + \epsilon}, \tag{19.259}$$

9. Définition 19.4.
10. Limite au sens de la définition 6.65, en sachant qu'elle est unique par la proposition 6.68.

pour $(\epsilon, t_1) \to (0,0)$. Cela donne bien $-\beta'(0)$. La seconde limite à calculer est celle de

$$\frac{t_1+\epsilon}{\|\beta(-\epsilon)-\beta(t_1)\|} = \frac{t_1+\epsilon}{\|-(t_1+\epsilon)\beta'(-\epsilon)-\alpha(t_1+\epsilon)\|} = \frac{t_1+\epsilon}{\|(t_1+\epsilon)\beta'(-\epsilon)+\alpha(t_1+\epsilon)\|}. \tag{19.260}$$

Nous calculons la limite de l'inverse (qui, si elle est non nulle donnera la réponse en inversant à nouveau) en nous souvenant de la formule

$$\big|a-|b|\big| \leqslant |a+b| \leqslant \big|a+|b|\big|. \tag{19.261}$$

Nous avons l'encadrement

$$\frac{\big|(t_1+\epsilon)\beta'(-\epsilon)-|\alpha(t_1+\epsilon)|\big|}{t_1+\epsilon} \leqslant \frac{\|(t_1+\epsilon)\beta'(-\epsilon)+\alpha(t_1+\epsilon)\|}{t_1+\epsilon} \leqslant \frac{\big|(t_1+\epsilon)\beta'(-\epsilon)+|\alpha(t_1+\epsilon)|\big|}{t_1+\epsilon} \tag{19.262}$$

Les limites des deux extrêmes existent et valent $\beta'(0)$; la règle de l'étau 12.93 conclu.

Deux chemins homotopes Nous considérons dans $\mathcal{T}$ les points $A=(0,0)$, $B=(0,L)$ et $C=(L,L)$. Nous allons considérer les chemins direct de A à C et celui passant via B. Et comme ces chemins doivent être paramétrés de 0 à 2π, il faut faire un peu attention. Nous définissons les chemins

$$\sigma_i \colon S^1 \to S^1 \tag{19.263}$$

par

$$\sigma_1(t) = \Sigma\left(\frac{L}{2\pi}t, \frac{L}{2\pi}t\right) = \frac{\beta'(tL/2\pi)}{\|\beta'(tL/2\pi)\|} = \frac{\tilde{\beta}'(t)}{\|\tilde{\beta}'(t)\|} = \Phi_\beta \tag{19.264a}$$

$$\sigma_2(t) = \begin{cases} \Sigma(0, \frac{L}{\pi}t) & \text{si } 0 \leqslant t \leqslant \pi \\ \Sigma\big(\frac{L}{\pi}(t-\pi), L\big) & \text{si } \pi \leqslant t \leqslant 2\pi \end{cases} \tag{19.264b}$$

où nous avons repris les notations de la définition 19.78. Notons que pour σ_2, en $t=\pi$ les deux expressions donnent

$$\sigma_2(\pi) = \Sigma(0,L) = -\frac{\beta'(0)}{\|\beta'(0)\|}. \tag{19.265}$$

Ce sont des chemins fermés parce que

$$\sigma_1(2\pi) = \Sigma(L,L) = \frac{\beta'(L)}{\|\beta'(L)\|} = \frac{\beta'(0)}{\|\beta'(0)\|} = \Sigma(0,0) = \sigma_1(0). \tag{19.266}$$

Notons que dans toutes ces définitions et calculs, nous avons utilisé de façon assez cruciale la définition 19.64 pour définir la dérivée de β en $t=0$.

Les chemins σ_1 et σ_2 sont homotopes par construction.

Indices et degrés La proposition 19.81(1) nous donne $\deg(\sigma_1) = \deg(\sigma_2)$. Nous avons alors la chaîne d'égalités

$$\deg(\sigma_2) = \deg(\sigma_1) = \deg(\Phi_\beta) = \mathrm{Ind}(\beta) = \mathrm{Ind}(\gamma) = \mathrm{Turn}(\gamma) \tag{19.267}$$

où les justifications sont :

(1) $\deg(\sigma_2) = \deg(\sigma_1)$ par homotopie : proposition 19.81.

(2) $\deg(\sigma_1) = \deg(\Phi_\beta)$ parce que $\sigma_1 = \Phi_\beta$.

(3) $\deg(\Phi_\beta = \mathrm{Ind}(\beta)$ par définition 19.78 de l'indice.

(4) $\mathrm{Ind}(\beta) = \mathrm{Ind}(\gamma)$ par invariance de l'indice sous reparamétrisation.

(5) $\mathrm{Ind}(\gamma) = \mathrm{Turn}(\gamma)$ par la proposition 19.79.

Il nous reste à montrer que $\deg(\sigma_2) = \pm 1$.

La géométrie de σ_2 Notons que par définition les valeurs de $\sigma_2(t)$ sont les vecteurs (unitaires) joignant p aux points de Γ lorsque $0 \leqslant t \leqslant \pi$, et les vecteurs inverses pour $\pi \leqslant t \leqslant 2\pi$. Plus précisément nous avons, si $0 < a < \pi$:

$$\sigma_2(a) = \Sigma(0, \frac{L}{\pi}a) = \frac{\beta(\frac{L}{\pi}a) - \beta(L)}{\|\beta(\frac{L}{\pi}a) - \beta(L)\|} \tag{19.268}$$

et

$$\sigma_2(\pi + a) = \Sigma(\frac{L}{\pi}a, L) = \frac{\beta(L) - \beta(\frac{L}{\pi}a)}{\|\beta(L) - \beta(\frac{L}{\pi}a)\|}. \tag{19.269}$$

En sachant que $\beta(0) = \beta(L)$ nous avons alors

$$\sigma_2(\pi + a) = -\sigma_2(a). \tag{19.270}$$

Notons aussi que le fait que γ soit une courbe simple assure que le numérateur et le dénominateur de σ_2 ne s'annulent pas autrement que pour $t = L$ ou $t = 0$.

Un relèvement pour σ_2 Ces propriétés motivent cette idée pour le relèvement de σ_2 :

$$\begin{aligned} \theta\colon [0, 2\pi[&\to [0, 2\pi[\\ t &\mapsto \text{l'angle du vecteur } \sigma_2(t), \end{aligned} \tag{19.271}$$

avec $\theta(2\pi)$ définit par continuité. Nous allons cependant voir, en étant plus prudent, que cette définition n'assure pas la continuité (surtout en $t = \pi$).

Soyons donc plus prudent et construisons θ petit à petit.

Étant donné que ℓ_p est tangente à Γ au point p, la droite ℓ_p est parallèle à $\beta'(0)$, et l'angle entre ℓ_p et $\beta'(0)$ est soit 0 soit π. Donc $\theta(0)$ devrait valoir soit 0 soit π.

Nous commençons par définir ceci :

$$\begin{aligned} \theta\colon [0, \pi[&\to [0, \pi] \\ t &\mapsto \text{angle de } \sigma_2(t) = \arccos\big(\sigma_2(t)_x\big). \end{aligned} \tag{19.272}$$

C'est le choix d'avoir ℓ_p horizontale et Γ au dessus de ℓ_p qui nous assure que pour tout $t \in [0, \pi]$, l'angle de $\sigma_2(t)$ peut être choisi entre 0 et π. De plus cette fonction est continue en tant que partie de la fonction $\arccos\colon [-1, 1] \to [0, \pi]$ qui est elle-même continue par la proposition 6.188.

Le fait que θ soit continue est assuré par le fait que σ_2 est C^∞.

Si $\theta(0) = 0$ Cela correspond à la situation des vecteurs rouges sur la figure 19.10.

Vu que $\sigma_2(\pi) = -\sigma_2(0)$, l'angle de σ_2 en $t = \pi$ est le supplémentaire de celui en $t = 0$. Mais pour $0 \leqslant t < \pi$, $\theta(t)$ prend ses valeurs entre 0 et π, le seul supplémentaire de 0 à être disponible est $\theta(\pi) = \pi$ (et non $\theta(\pi) = -\pi$ par exemple). Nous définissons donc $\theta(\pi) = \pi$ pour la continuité.

En ce qui concerne les $\pi < t < 2\pi$ nous savons que $\sigma_2(\pi + a) = -\sigma_2(a)$. Les angles sont donc les supplémentaires de ceux pour $0 \leqslant t \leqslant \pi$. Pour assurer la continuité en $t = pi$ nous sélectionnons la place $]\pi, 2\pi]$; et nous définissons

$$\begin{aligned} \theta\colon]\pi, 2\pi[&\to]\pi, 2\pi] \\ t &\mapsto \text{angle de } \sigma_2(t). \end{aligned} \tag{19.273}$$

Le nombre $\theta(2\pi)$ est défini par continuité. Il doit valoir $\theta(2\pi) = 2\pi$.

La fonction θ ainsi définie est un relèvement pour σ_2, et le degré peut être calculé :

$$\deg(\sigma_2) = \frac{1}{2\pi}\big(\theta(2\pi) - \theta(0)\big) = 1. \tag{19.274}$$

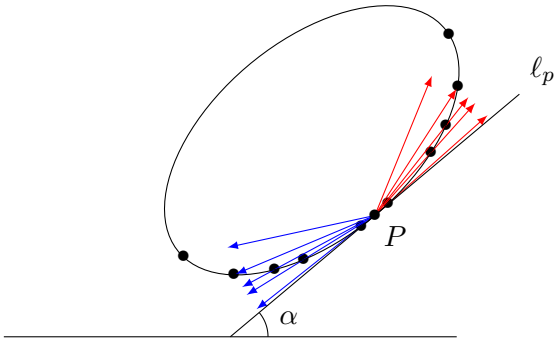


FIGURE 19.10 – Les vecteurs représenant σ_2 dans le cas où $\beta'(0)$ est dans le sens de ℓ_p ou dans le sens inverse. Pour le sport nous avons dessiné la situation avec une droite ℓ quelconque plutôt que horizontale.

Si $\theta(0) = \pi$ Cela correspond à la situation des vecteurs bleus sur la figure 19.10.

Alors $\theta(\pi) = 0$ est obligatoire parce qu'il doit être supplémentaire à $\theta(0)$. Les angles atteints par $\sigma_2(t)$ pour $t \in]\pi, 2\pi[$ sont encore les complémentaires, mais cette fois la continuité en $t = \pi$ nous impose de les chercher dans $]-\pi, 0]$ et nous définissons

$$\begin{aligned} \theta\colon]\pi, 2\pi[&\to]-\pi, 0] \\ t &\mapsto \text{angle de } \sigma_2(t). \end{aligned} \tag{19.275}$$

Par continuité nous devons avoir $\theta(2\pi) = \theta(\pi) - \pi$ et donc $\theta(2\pi) = -\pi$.

Le degré de σ_2 est alors

$$\deg(\sigma_2) = \frac{1}{2\pi}\big(\theta(2\pi) - \theta(0)\big) = -1. \tag{19.276}$$

Conclusion Nous avons montré que le degré de σ_2 est 1 ou -1, et en remontant les égalités (19.267) nous déduisons que $\text{Turn}(\gamma) = \pm 1$.

□

19.12 Courbes fermées planes

19.12.1 Cercle circonscrit

La proposition suivante est dans le même esprit que l'ellipse de John-Loewer [11].

Proposition-définition 19.83 (Cercle circonscrit[248, 1])**.**
Soit une courbe fermée simple et continue $\gamma\colon [0,1] \to \mathbb{R}^2$. Soit Γ son image. Il existe un unique cercle de rayon minimum contenant Γ. Ce cercle est le ***cercle circonscrit*** *à γ.*

Il a les propriétés suivantes :

(1) Le cercle circonscrit à γ coupe Γ en au moins deux points distincts.

(2) Tout arc du cercle circonscrit plus grand que le demi-cercle intersection Γ.

Démonstration. Division de la preuve.

Existence L'application γ étant continue, l'ensemble Γ est compact (théorème 6.77). Nous considérons l'ensemble Q des formes quadratiques de la forme

$$q_{a,r}(x) = \|a - x\|^2 - r^2 \tag{19.277}$$

avec $a \in \mathbb{R}^2$ et $r \in \mathbb{R}$. Nous mettons sur cet ensemble la topologie de $\mathbb{R}^2 \times \mathbb{R} = \mathbb{R}^3$. Le nombre $q_{a,r}(x)$ est continu en a, r et x. Soit A l'ensemble des formes quadratiques de cette forme et vérifiant

$$q\big(\gamma(t)\big) \leqslant 0 \tag{19.278}$$

11. Proposition 18.100.

pour tout $t \in [0,1]$.

Cet ensemble A est non vide parce que Γ est compact et donc borné ; il existe donc une boule qui contient Γ en son intérieur.

Montrons que A est fermé dans Q. Si $q_{a,r} \notin A$ alors il existe t_0 tel que $q_{a,r}\big(\gamma(t_0)\big) > 0$. Par continuité, il existe un voisinage de (a,r) dans $\mathbb{R}^3$ et donc de $q_{a,r}$ dans Q tel que $q_{a',r'}\big(\gamma(t_0)\big)$ reste strictement positif pour tout (a',r') dans ce voisinage.

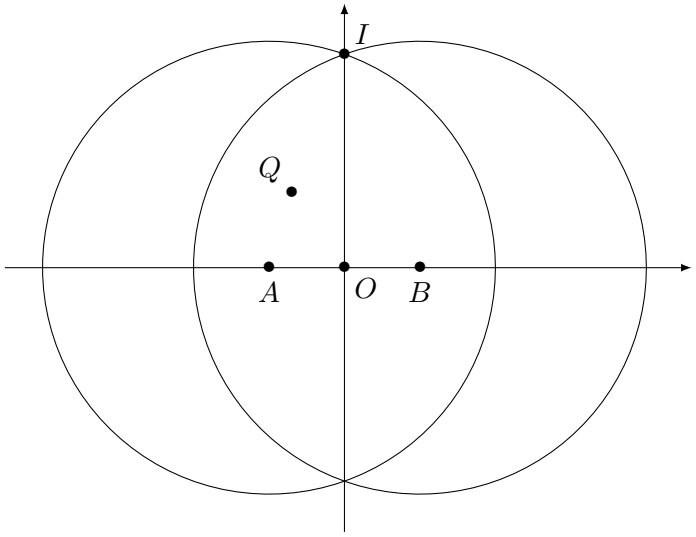
En particulier l'ensemble

$$\{r \in \mathbb{R}^+ \text{ tel que } \exists\,(a,r) \text{ tel que } q_{a,r} \in A\} \tag{19.279}$$

est fermé et borné vers le base. De plus $r = 0$ n'est pas dans cet ensemble. Il possède donc un minimum strictement positif. Le cercle correspondant donne l'existence.

Unicité En ce qui concerne l'unicité, si Γ est contenu dans les boules $B(a,R)$ et $B(b,R)$ alors

$$\Gamma \subset B(a,R) \cap B(b,R). \tag{19.280}$$



Nous choisissons les axes comme indiqué sur le dessin et nous montrons que l'intersection est dans le cercle de centre O et de rayon $\|OI\| < R$. Soit Q un point de l'intersection ; par symétrie il est suffisant de supposer $Q_x < 0$ et $Q_y > 0$. Vu que Q est dans le cercle de centre B et de rayon R, il doit satisfaire

$$(B_x - Q_x)^2 + Q_y^2 < R. \tag{19.281}$$

D'autre part le point I est d'abscisse $I_x = 0$ et d'ordonnée donnée par $I_y^2 = R^2 - B_x^2$.

Nous devons prouver que $Q_x^2 + Q_y^2 \leqslant I_y^2$. Il s'agit simplement de calculer

$$Q_x^2 + Q_y^2 \leqslant Q_x^2 + R^2 - B_x^2 + 2B_xQ_x - Q_x^2 = R^2 - B_x^2 + 2B_xQ_x \leqslant R^2 - B_x^2 = I_y^2 \tag{19.282}$$

parce que $B_x > 0$ et $Q_x < 0$.

Nous concluons que Γ est inclus à un cercle de rayon plus petit que R et donc que R n'est pas minimum. D'où l'unicité.

Au moins deux intersections Nous nommons C le cercle circonscrit à γ, et nous écrivons, pour $p \in \Gamma$

$$r(p) = d(p,C) \tag{19.283}$$

la distance entre p et C. Cela est une fonction continue sur le compact Γ. Elle atteint donc ses bornes sur Γ.

Si $C \cap \Gamma = \varnothing$ alors $r(p) > 0$ pour tout p et le minimum est également strictement positif. Soit r_0 ce minimum. Alors le cercle C_2 même centre que C mais de rayon $r_0/2$ n'intersecte pas non plus Γ parce que

$$d(p,C) \leqslant d(p,C_2) + d(C_2,C) \tag{19.284}$$

où $d(p,C) = r(p)$ et $d(C_2, C) = r_0/2$.

$$d(p, C_2) \geqslant r(p) - \frac{r_0}{2} \geqslant \frac{r_0}{2} > 0. \tag{19.285}$$

Si $C \cap \Gamma = \{q\}$ alors un choix d'axe place le centre de C en $(0,0)$, le point q en $(1,0)$ et fixe le rayon de C à 1. Pour $p \in \Gamma$ nous notons $r(p)$ la distance entre p et le point de C situé sur sa gauche, c'est à dire, si $p = (p_x, p_y)$,

$$r(p) = \sqrt{1 - p_y^2} + p_x. \tag{19.286}$$

Cela est encore une fonction continue sur Γ qui atteint son minimum valant r_0. Alors le cercle de centre $(\frac{r_0}{2}, 0)$ et de même rayon contient encore Γ mais n'a plus de points d'intersection avec Γ.

Enfin, tout nombre de points d'intersection entre C et Γ est possible à partir de 2. Pour en avoir deux, prendre une ellipse, et pour en avoir plus, prendre des polynôme dont les angles sont un peu modifiés de façon à rester C^1.

Intersection avec les demi-arcs Supposons, en fixant encore les axes, que le cercle circonscrit soir encore centré en $(0,0)$ et que Γ n'intersecte pas le demi-cercle $x < 0$. Alors pour tout $p \in \Gamma$ la distance entre p et ce demi-cercle est strictement positive. Il y a un minimum r_0. En décalant le centre du cercle de $r_0/2$ vers la droite, nous obtenons un nouveau cercle contenant Γ mais ne l'intersectant pas.

□

19.12.2 Description locale

Définition 19.84.
Une courbe plane différentiable est ***convexe*** *si son graphe est en tout point d'un seul côté de sa tangente.*

Lemme 19.85 ([1])**.**
Soit une courbe simple, convexe $\gamma\colon [0,1] \to \mathbb{R}^2$ que nous supposons être de classe C^1. Nous notons Γ l'image de γ. Alors Γ est localement le graphe d'une fonction convexe (définition 18.74).

Démonstration. Soit $p \in \Gamma$ et l_p la tangente à Γ en p. Nous considérons un système d'axe centré en p de telle sorte que l_p soit l'axe des abscisses et l'axe des ordonnées soit dirigé de tel manière que Γ se trouve dans la partie $y > 0$. De plus nous paramétrons γ de telle sorte à avoir $\gamma(0) = p = (0,0)$.

Vu que $l_p \equiv y = 0$ est la tangente à Γ nous avons $\gamma_y' = 0$ et $\gamma_x'(0) > 0$. Nous en déduisons que $\gamma_x'(t) > 0$ pour tout $t \in B(0,\delta)$ pour δ suffisamment petit. Nous posons alors

$$g(x) = \gamma_y\big(\gamma_x^{-1}(x)\big) \tag{19.287}$$

qui est bien définie parce que γ_x est une bijection entre $B(0,\delta)$ et son image. La fonction g est continue par la proposition 6.189 et même dérivable par la proposition 12.117. De plus, vu la formule (12.240), la fonction g^{-1} est de classe C^1 parce que $(g^{-1})'$ est une composée d'applications continues.

Si g est C^2 Dans ce cas, g'' ne peut pas changer de signe, sinon la tangente coupe le graphe. Par positivité de g (et le fait que $g(0) = g'(0) = 0$), il n'est pas possible d'avoir $g'' < 0$ partout. Donc $g'' \geqslant 0$ partout. Cela prouve que g est convexe par la caractérisation 18.80.

Si g est seulement de classe C^1 Le graphe de g correspond au graphe de Γ. Nous montrons que g est convexe en utilisant la caractérisation de la proposition 18.84.

La tangente au graphe de g en $x = x_0$, que nous notons l_0, est la tangente à Γ en $t = \gamma_x^{-1}(x_0)$. Le graphe de g, qui est une partie de Γ se trouve donc d'un seule côté de l_0.

Nous nous restreignons g à un compact I et nous considérons la fonction

$$d_a(x) = g(x) - l_a(x) \tag{19.288}$$

qui donne la distance entre le graphe de g et la tangente à g en $x = a$. Cela est une fonction continue en x et en a. Le graphe de g est au dessus de la tangente en $x = 0$ (par construction des axes). Supposons que le graphe de g soit en dessous de la tangente en $x = x_2$. Alors nous avons, pour tout x :

$$\begin{cases} d_0(x) \geqslant 0 & (19.289a) \\ d_{x_1}(x) \leqslant 0. & (19.289b) \end{cases}$$

Nous posons

$$s(a) = \sup_{x \in I} d_a(x), \tag{19.290}$$

qui est une fonction continue par la proposition 12.37. Vu le définitions, $s(0) \geqslant 0$ et $s(x_2) \leqslant 0$. Il existe donc $m \in [0, m_2]$ tel que $s(m) = 0$. À ce moment nous avons $g(x) = l_m(x)$ pour tout $x \in I$ et donc g est une droite, et en réalité toutes les inégalités sont des égalités. La fonction g est alors bien convexe (mais pas strictement).

□

19.12.3 Enveloppe convexe

Proposition 19.86 ([249, 1])**.**
Soit une courbe simple, fermée et convexe $\gamma\colon [0,1] \to \mathbb{R}^2$ que nous supposons être de classe C^2. Nous notons Γ l'image de γ. Alors il existe un convexe D tel que $\partial D = \Gamma$.

Démonstration. Pour $p \in \Gamma$ nous notons l_p la tangente à Γ en p et H_p le demi-plan (fermé) contenant Γ. Nous posons

$$D = \bigcap_{p \in \Gamma} H_p. \tag{19.291}$$

Cet ensemble est convexe comme intersection de convexes et fermé comme intersection de fermés. Nous prouvons que $\Gamma = \partial D$.

L'inclusion $\Gamma \subset \partial D$ est la plus facile. Si $p \in \Gamma$ alors p est dans chacun des H_q et donc dans D. De plus tout voisinage de p contient des points en dehors de H_p, donc p n'est pas dans l'intérieur de D. Ce dernier étant fermé, un point hors de l'intérieur est sur le bord. Ergo $p \in \partial D$.

Pour l'inclusion inverse, soit $p \in \partial D$.

Il existe q tel que $p \in l_q$ Vu que D est fermé, le point p est dans D, et donc dans tous les H_q. Supposons qu'il soit dans l'intérieur de tous les H_q. Alors nous considérons la fonction

$$r(q) = \frac{d(p, H_q)}{2} \tag{19.292}$$

définie sur Γ. C'est une fonction continue[12] strictement positive définie sur le compact Γ qui possède donc un minimum strictement positif. Si r_0 est ce minimum, alors $B(p, r_0)$ est inclue à tous les H_q, ce qui ferait que p est à l'intérieur de D. Nous concluons à l'existence de $q \in \Gamma$ tel que $p \in l_q$.

Le point où ça décolle Nous supposons que $p \notin \Gamma$, sinon ce serait trop facile. Nous paramétrons γ de telle sorte à avoir $q = \gamma(0)$ et nous posons

$$T = \{t \in \mathbb{R}^+ \text{ tel que } l_{\gamma(t)} = l_q\}. \tag{19.293}$$

Cela est un fermé dans $\mathbb{R}$ parce que γ est de classe C^1. Nous posons $t_0 = \inf(T^c)$ et pour tout ϵ suffisamment petit,

$$\begin{cases} l_{\gamma(t_0+\epsilon)} \neq l_q, & (19.294a) \\ l_{\gamma(t_0-\epsilon)} = l_q. & (19.294b) \end{cases}$$

La seconde est parce que si $l_{\gamma(t_0-\epsilon)} \neq l_q$ nous aurions $t_0 - \epsilon \in T^c$. Soit $r = \gamma(t_0)$; nous avons $l_r = l_q$ parce que si $l_r \neq l_q$ alors par continuité de γ' nous aurions $l_{r'} \neq l_q$ pour tout $r' \in \gamma\big(B(t_0, \delta)\big)$.

12. L'équation de la droite l_q a des coefficients continus parce que γ est de classe C^1.

Graphe d'une fonction strictement convexe En suivant le lemme 19.85, l'ensemble Γ est localement (autour de r) le graphe d'une fonction convexe au-dessus de l_r. Soit $g\colon B(0,\delta) \to \mathbb{R}^2$ cette fonction convexe.

Soit $\epsilon > 0$. Si $g''(x) = 0$ sur $]0,\epsilon]$ alors g' y est constante. Mais $g'(0) =$, ce qui signifierait que sur $[0,\epsilon]$ nous ayons $g'(x) = 0$ et donc $g(x) = 0$. Cela ferait que $l_{r'} \equiv y = 0$ pour tout r' de la forme $(x,0)$ avec $x \in [0,\epsilon]$ (qui sont des points de Γ). Cela est en contradiction avec la définition de r. Donc il existe un point $x \in [0,\epsilon[$ tel que $g''(x) > 0$.

Rappelons que $p \in l_q$, ce qui fait que p a pour coordonnées $(p_x, 0)$. Nous restreignons δ et ϵ de telle sorte que p_x soit plus grand à la fois que ϵ et δ.

Il existe donc un intervalle $[a,b]$ avec $a, b \geqslant 0$ et $a, b < p_x$ sur lequel g est strictement convexe.

La tangente qui tue En particulier $g(b) > 0$ et le théorème de des accroissements finis 12.129(1) nous donne l'existence de $m \in [0,b]$ tel que la tangente à g en $x = m$ est parallèle au segment joignant $(0,0)$ à $(b, f(b))$. Cette tangente, que nous nommons l_m, est en dessous de la corde, par strict convexité. En particulier, son point d'intersection avec $y = 0$ est strictement entre 0 et m.

L'ensemble D est d'un seul côté de l_m. Ce côté est forcément celui de $q = (0,0)$ (parce que $q \in \Gamma \subset D$), et donc le points de coordonnées $(x,0)$ avec $x > m$ ne sont pas dans D. Pas de chance, p est un point de ce type.

La contradiction Nous avons prouvé que si $p \in \partial D \backslash \Gamma$ alors p n'est pas dans D, ce qui est impossible parce que, l'ensemble D étant fermé, nous avons $\partial D \subset D$.

□

Remarque 19.87.
Bien que cela puisse paraître évident dès le début, nous ne démontrerons que dans la proposition 19.89 que D est l'enveloppe convexe de Γ.

Corollaire 19.88.
Si $p \in \mathrm{Int}(D)$ alors toute droite passant par p intersecte Γ en exactement 2 points.

Démonstration. Vu que la partie D est bornée, toute droite passant par son intérieur coupe ∂D en au moins deux points (un dans chaque sens, et en utilisant le lemme de passage de douane 6.177).

Soit ℓ une droite passant par p et supposons qu'elle coupe Γ en trois points distincts. Alors au moins deux d'entre eux sont du même côté de p. Soient q_1 et q_2 ces points. Nous avons donc dans l'ordre $p \in \mathrm{Int}(D)$, $q_1 \in \partial D$ et $q_2 \in \partial D$.

Vu que tout Γ est d'un seul côté de ses tangentes, lesdites tangentes ne passent pas par l'intérieur de D. Ni p ni q_2 ne sont sur ℓ_{q_1}, parce que si $q_2 \in \ell_{q_1}$ alors $\ell_{q_1} = \ell$, ce qui est impossible parce que ℓ passe par $p \in \mathrm{Int}(D)$.

Or p et q_2 sont de part et d'autres de ℓ_{q_1}, ce qui est impossible parce que Γ est d'un seul côté de cette droite. □

Proposition 19.89.
Soit D l'ensemble définit en (19.291).

(1) $D = \mathcal{C}\mathrm{onv}(\Gamma)$ ($\mathcal{C}\mathrm{onv}(\Gamma)$ désigne l'enveloppe convexe de Γ)

(2) $\partial\, \mathcal{C}\mathrm{onv}(\Gamma) = \Gamma$.

Pour la définition d'enveloppe convexe, voir la définition 11.32.

Démonstration. L'ensemble D est un convexe contenant Γ. Donc $\mathcal{C}\mathrm{onv}(\Gamma) \subset D$. L'inclusion inverse est à prouver.

Soit $x \in D$. Si $x \in \partial D$ alors $x \in \Gamma$ (proposition 19.86) et donc $x \in \mathcal{C}\mathrm{onv}(\Gamma)$. Nous ne devons donc traiter que le cas $x \in \mathrm{Int}(D)$.

Le corollaire 19.88 nous dit que toute droite passant par x coupe Γ en exactement deux points. Soient p et q ces points. Alors $p, q \in \Gamma \subset \mathcal{C}\mathrm{onv}(\Gamma)$. Vu que $\mathcal{C}\mathrm{onv}(\Gamma)$ est convexe, tout le segment $[p,q]$ est dans $\mathcal{C}\mathrm{onv}(\Gamma)$, et en particulier $p \in \mathcal{C}\mathrm{onv}(\Gamma)$.

Nous passons à la seconde affirmation. Nous savons que $D = \mathcal{C}\text{onv}(\Gamma)$. En prenant le bord des deux côtés, $\partial D = \partial\,\mathcal{C}\text{onv}(\Gamma)$, donc $\Gamma = \partial\,\mathcal{C}\text{onv}(\Gamma)$. □

Lemme 19.90 (Des tangentes parallèles[250])**.**
Une courbe fermée γ de classe C^1 est convexe si et seulement si elle ne possède pas 3 tangentes parallèles distinctes.

Démonstration. Si le graphe Γ possédait trois tangentes parallèles distinctes, une serait entre les deux autres et le graphe Γ serait de part et d'autres des cette tangente. Dans ce cas, γ n'est pas convexe.

Nous montrons maintenant que si γ n'est pas convexe, alors elle possède trois tangentes parallèles distinctes. Pour $p \in \Gamma$ tel que Γ soit des deux côtés de l_p.

Soient Γ_1 et Γ_2 les parties de Γ délimitées par l_p. Nous notons q_i le point de Γ_i le plus éloigné de la droite l_p, il existe parce que G est compact et que la fonction distance à l_p est continue sur Γ. Les points p, q_1 et q_2 sont distincts (sinon l_p ne couperait pas Γ en deux parties).

Montrons que $l_{q_i} \parallel l_p$. Pour cela nous choisissons un système d'axe dans lequel $l_p \equiv y = 0$. Dans ce système, la distance entre $\gamma(t)$ et l_p est $\gamma_y(t)$ et les extrema de cette fonction ont lieu aux points t tels que $\gamma_y'(t) = 0$, c'est à dire aux points sur lesquels la tangente est parallèle à l_p.

À quel moment avons nous utilisé le fait que la courbe soit fermée ? Au moment de dire que le point le plus éloigné devait vérifier $\gamma_y'(t) = 0$. En effet un extrema peut ne pas vérifier cette condition s'il n'est pas à l'intérieur du domaine. Dans notre cas, nous avons $\gamma\colon [0,1] \to \mathbb{R}^2$ qui est fermée : $\gamma(0) = \gamma(1)$. Donc en réalité nous pouvons considérer $\gamma\colon \mathbb{R} \to \mathbb{R}^2$ et tous les points du domaine sont intérieurs au domaine. L'extremum doit donc vérifier la condition d'annulation de la dérivée. □

Exemple 19.91
Si la courbe n'est pas fermée, alors le lemme 19.90 ne tient pas comme le montre le contre-exemple du graphe de $f(x) = x^3 - x$. Il est non convexe et pourtant ne présente que 2 tangentes parallèles (dans chaque directions). △

Lemme 19.92.
Soit γ une courbe convexe et ℓ une droite qui intersecte Γ mais qui n'est pas tangente. Alors l'intersection entre ℓ et Γ comprend au maximum 2 points.

Démonstration. Supposons que ℓ coupe Γ en trois points p, q, r (dans cet ordre). Vu que ℓ n'est pas tangente à Γ, la tangente ℓ_q est distincte de ℓ (et intersecte ℓ en l'unique point q). Les points p et r sont dans Γ et sont pourtant de deux côtés différents de ℓ_q. Contradiction avec la convexité de γ. □

La proposition suivante nous dit que si deux points de Γ ont la même tangente, alors entre ces deux points, Γ est le segment de droite les joignant.

Proposition 19.93 ([1])**.**
Soit $\gamma\colon [0, L] \to \mathbb{R}^2$ une courbe fermée simple et convexe de classe C^1. Si la tangente en $p = \gamma(s_p)$ et la tangente en $q = \gamma(s_q)$ sont identiques (pas seulement parallèles), alors soit

$$\gamma\big([s_p, s_q]\big) = [p, q] \tag{19.295}$$

soit

$$\gamma\big([s_q, L]\big) = [p, q] \tag{19.296}$$

Démonstration. Nous considérons un système d'axe dans lequel $p = (0,0)$, $\ell_p = \ell_q \equiv y = 0$ et tel que $\gamma_y(t) \geqslant 0$ pour tout t. Nous choisissons enfin une paramétrisation de γ telle que $p = \gamma(0)$.

Si $\gamma_y\big([0, s_q]\big) = \{0\}$ alors par le théorème des valeurs intermédiaires 12.21, tous les points du type $(t, 0)$ avec $0 \leqslant t \leqslant q_x$ sont atteints par $\gamma(t)$ avec $0 \leqslant t \leqslant s_q$. De plus aucun autre point ne peut être atteint parce que γ étant simple, elle ne peut pas faire de retour en arrière.

Nous supposons donc que γ_y n'est pas identiquement nulle sur $[0, s_q]$; il existe donc un s_M avec $0 < s_M < q_q$ qui maximise γ_y sur $[0, s_q]$. La tangente à Γ en $\gamma(s_M)$ est horizontale. Les droites $y = 0$ et $y = \gamma_y(s_M)$ sont donc deux tangentes parallèles à Γ. Par le lemme des tangentes parallèles 19.90, il n'y a pas d'autres tangentes horizontales. Donc pour tout $n \in \mathbb{N}$ la droite $y = \frac{1}{n}$ n'est tangente nulle part [13] à Γ.

Par le lemme 19.92, la droite $y = \frac{1}{n}$ ne peut intersecter Γ qu'en seulement deux points. Or le théorème des valeurs intermédiaires appliqué à γ_y sachant que $\gamma_y(0) = \gamma_y(s_q) = 0$ et $\gamma_y(s_M) > 0$ nous donne a_n, b_n tels que

$$0 < a_n < s_M \tag{19.297a}$$
$$s_M < b_n < s_q \tag{19.297b}$$

et $\gamma_y(a_n) = \gamma_y(b_n) = \frac{1}{n}$. Donc la droite $y = \frac{1}{n}$ intersecte γ deux fois dans $[0, s_q]$. En conséquence de quoi $\gamma_y(t) < \frac{1}{n}$ pour tout $t \in [q_q, L]$. Cela étant valable pour tout n nous avons $\gamma_y\big([s_q, L]\big) = \{0\}$ et nous sommes ramenés essentiellement au premier cas. □

Proposition 19.94.
Soit une courbe fermée simple γ de classe C^1 et une droite ℓ. Alors il existe au moins deux points distincts q_1, q_2 tels que $\ell_{q_1} \parallel \ell_{q_2} \parallel \ell$ avec $\ell_p \neq \ell_q$.

Démonstration. Pour $p \in \Gamma$ nous considérons le nombre $r(p) = d(p, \ell)$. En tant que fonction continue sur un compact, elle possède un minimum et un maximum. Dans un système d'axe pour lequel $\ell \equiv y = 0$, la fonction r s'écrit $r(p) = p_y$ et les extrema arrivent en $\gamma(s)$ avec $\gamma_y'(s) = 0$, ce qui signifie que les tangentes aux extrema sont parallèles à ℓ. Vu que le maximum et le minimum ne peuvent pas être égaux (sinon la courbe serait horizontale et pas simple), les tangentes en ces points sont distinctes. □

Corollaire 19.95.
Soit une courbe convexe fermée simple γ de classe C^2. L'ensemble $\mathcal{C}\text{onv}(\Gamma)$ est compact.

Démonstration. Nous savons que $\mathcal{C}\text{onv}(\Gamma)$ n'est autre que D par la proposition 19.89. Nous savons déjà que D est fermé. Il nous suffit donc de prouver qu'il est borné (théorème de Borel-Lebesgue 6.123). Nous considérons deux droites perpendiculaires et les 4 tangentes correspondantes par la proposition 19.94. Vu que Γ est d'un seul côté de chacune de ces tangentes, elle est contenue dans le rectangle délimité par ces 4 droites. □

19.12.4 Courbure et convexité

Lemme 19.96.
Soit $\gamma\colon [0, L] \to \mathbb{R}^2$ une courbe simple, fermée en paramétrisation normale. Alors l'application

$$\begin{aligned} \sigma\colon [0, L] &\to S^1 \\ s &\mapsto \gamma(s) \end{aligned} \tag{19.298}$$

est surjective.

Démonstration. Le théorème 19.82 nous dit que si $\theta(0) = a$ alors $\theta(2\pi) = a + 2\pi$ ou $a - 2\pi$. Le théorème des valeurs intermédiaires nous dit alors que θ prend toutes les valeurs entre a et $a + 2\pi$ ou $a - 2\pi$. □

Proposition 19.97 ([250])**.**
Une courbe fermée simple de classe C^2 est convexe si et seulement si sa courbure est soit toujours positive, soit toujours négative.

13. À part $n = 0$ et si par manque de chance, $\gamma_y(s_M)$ est un nombre de la forme $1/n$.

Démonstration. Nous considérons la courbe γ munie d'un paramétrage de vitesse 1, c'est à dire avec $\|\gamma'(t)\| = 1$ pour tout t. Si θ est sa fonction d'angle, alors nous avons $\theta' = \kappa$ par le lemme 19.74. Donc la fonction θ est monotone si et seulement si la courbure ne change pas de signe. Nous allons montrer que θ est monotone si et seulement si γ est convexe.

$\Rightarrow$ Nous supposons que θ est monotone et γ non convexe. Soient $p \in \Gamma$ tel que Γ soit des deux côtés de ℓ_p, et soient les points q_1, q_2 donnés par le lemme des tangentes parallèles 19.90 tels que $\ell_p \parallel \ell_{q_1} \parallel \ell_{q_2}$. Parmi les vecteurs tangents en p, q_1 et q_2, deux au moins ont la même direction ; supposons que ce sont q_1 et q_2. C'est à dire que si $p = \gamma(s_0)$, $q_1 = \gamma(s_1)$ et $q_2 = \gamma(s_2)$ alors nous avons $\gamma'(s_1) = \gamma'(s_2)$ et donc aussi

$$\theta(s_1) = \theta(s_2) + 2n\pi \tag{19.299}$$

pour un certain n. Mais θ est monotone et la différence entre sa première et sa dernière valeur doit valoir 2π ou -2π par le théorème 19.82. Donc n ne peut valoir que -1, 0 et 1.

Si $n = 0$ alors θ est constante sur $[s_1, s_2]$. Si $n = 1$ alors $\theta(s_1) = \sigma(s_2) + 2\pi$ alors que sur toute la courbe, θ ne peut faire que 2π. Donc θ est constant sur $[0, s_1]$ et sur $[s_2, L]$ (où L est le bord de la paramétrisation). Si $n = -1$, même conclusion.

Dans tous les cas, Γ contient une ligne droite, soit de q_1 à q_2, soit de q_2 à q_1. Et dans ces cas nous avons $\ell_{q_1} = \ell_{q_2}$, ce qui est contraire à la construction de q_i.

Nous concluons que γ est convexe.

$\Leftarrow$ Nous supposons que γ est convexe, mais que θ n'est pas monotone. Il existe donc $s_1 < s_0 < s_2$ tels que

$$\theta(s_1) = \theta(s_2) \neq \theta(s_0). \tag{19.300}$$

Et vu le lemme 19.96, il existe s_3 tel que $\gamma'(s_3) = -\gamma'(s_1)$.

Donc en s_1, s_2 et s_3 nous avons trois tangentes parallèles. La proposition 19.90 est alors formelle, γ étant convexe, deux de ces tangentes doivent être identiques.

La proposition 19.93 dit qu'entre deux points dont les tangentes sont identiques, la courbe doit être un segment de droite. Or sur un segment de droite, $\kappa = 0$ et θ est constante.

— La partie $\gamma\big([s_1, s_2]\big)$ ne peut pas être droite parce que nous avions supposé l'existence d'un $s_0 \in]s_1, s_2[$ tel que $\theta(s_1) = \theta(s_2) \neq \theta(s_0)$.
— La partie $\gamma\big([s_1, s_3]\big)$ ne peut pas être droite parce que $\theta(s_3) \neq \theta(s_1)$.
— La partie $\gamma\big([s_2, s_3]\big)$ ne peut pas être droite parce que $\theta(s_2) \neq \theta(s_1)$.

Nous sommes donc devant une contradiction.

Nous en concluons que θ doit être monotone.

□

19.12.5 Théorème des quatre sommets

Lemme 19.98.
Soit une droite ℓ du plan. Il existe $a, c \in \mathbb{R}^2$ avec $c \neq 0$ tels que $z \in \ell$ si et seulement si $(z-a) \cdot c = 0$.

Démonstration. Une droite est paramétrée par $\gamma(t) = p + tq$. En posant $a = p$ et $c = Jq$ nous avons la réponse. En effet nous allons montrer qu'avec ces valeurs de a et c, nous avons $z \in \Gamma$ si et seulement si $(z - a) \cdot c = 0$.

D'abord un point de γ est de la forme $z = \gamma(t) = p + tq$. Nous avons :

$$\big(\gamma(t) - a\big) \cdot c = \big(\gamma(t) - p\big) \cdot Jq = tq \cdot Jq = 0. \tag{19.301}$$

Et dans l'autre sens, si $(z-a) \cdot c = 0$ nous devons prouver que $z \in \Gamma$. Nous avons : $(z-p) \cdot Jq = 0$, ce qui fait que $z - p$ est un multiple de q. Autrement dit : $z - p = \lambda q$ ou encore $z = \alpha q + p$, qui est sur la droite Γ. □

Définition 19.99.
*Un **sommet** d'une courbe est un point d'extremum local de la courbure.*

Théorème 19.100 (Théorème des quatre sommets[2, 198])**.**
Soit un arc paramétrique $\gamma\colon \mathbb{R} \to \mathbb{R}^2$ fermé, simple et convexe[14] *de classe C^3 et T-périodique.*
Alors γ possède au moins 4 points critiques sur chaque période.

Démonstration. Nous supposons que la paramétrisation de γ soit normale.

Si la courbure κ est constante sur une partie ouverte de la (paramétrisation de la) courbe, alors tous les points de cette partie sont des sommets et le théorème est fait. Nous supposons que κ n'est pas constante et en particulier que Γ ne contient ni bouts de droites ni bouts de cercles (théorème 19.70).

La fonction κ étant de classe C^1 sur le compact Γ, elle admet au moins un maximum et un minimum distincts. Vu que ces points sont intérieurs, ils correspondent au changement de signe de κ'. Soient p et q ces points. Pour la simplicité nous supposons que γ est paramétré de telle sorte que $\gamma(0) = p$, et $q = \gamma(s_q)$ avec $0 < s_q < T$.

Nous supposons que p et q sont les seuls points de changement de signe de κ'.

Soit ℓ la droite passant par p et q. Tous les points du segment $[p, q]$ (qui sont dans $\mathcal{C}\text{onv}(\Gamma)$) ne peuvent pas être sur Γ (sinon nous aurions un morceau de droite). Donc certains points sont dans l'intérieur de $\mathcal{C}\text{onv}(\Gamma)$. Donc la droite ℓ passe par l'intérieur de $\mathcal{C}\text{onv}(\Gamma)$ et le corollaire 19.88 nous dit que la droite ℓ ne coupe Γ en seulement deux points.

Par conséquent, les ensembles $\gamma\big([0, s_q]\big)$ et $\gamma\big([s_q, T]\big)$ sont de part et d'autre de Γ. Vu qu'en ces points, κ' change de signe et qu'il ne change de signe en aucun autre points, la fonction κ' est positive d'un côté de ℓ et négative de l'autre.

D'autre part par le lemme 19.98, il existe $a \in \mathbb{R}^2$ et $c \neq 0$ tels que $z \in \ell$ si et seulement si $(z - a) \cdot c = 0$. La fonction $z \mapsto (z - a) \cdot c$ est donc positive d'un côté de ℓ et négative de l'autre.

En résumé les fonctions

$$s \mapsto \kappa'(s) \tag{19.302a}$$
$$s \mapsto \big(\gamma(s) - a\big) \cdot c \tag{19.302b}$$

changent de signe en même temps et le produit a donc un signe constant. Ce produit n'est de plus pas nul parce que κ' n'est nul sur aucun intervalle (sinon κ y serait constant et Γ un segment de droite) et $\big(\gamma(s) - a\big) \cdot c$ ne s'annule pour aucun s sauf ceux qui correspondent à p et q.

Nous avons donc

$$0 \neq \int_0^T \kappa'(s)\big(\gamma(s) - a\big) \cdot c\, ds \tag{19.303a}$$
$$= \underbrace{\Big[\big(\gamma(s) - a\big) \cdot c\kappa(s)\Big]_0^T}_{A=0} - \int_0^T \kappa(s)\gamma'(s) \cdot c\, ds \tag{19.303b}$$
$$= -\int_0^T \kappa(s)\big(\gamma'(s) \cdot c\big) ds \tag{19.303c}$$
$$= \int_0^T J\gamma''(s) \cdot c\, ds \tag{19.303d}$$
$$= J\int_0^T \gamma''(s) \cdot c\, ds \tag{19.303e}$$
$$= J\big[\gamma'(s) \cdot c\big]_0^T \tag{19.303f}$$
$$= 0. \tag{19.303g}$$

Justifications :

14. Par la proposition 19.97 nous pouvons aussi bien demander à la courbure d'être toujours strictement positive, comme le fait [2].

— L'expression A est nulle parce que les valeurs en 0 et en T sont identiques.
— Nous utilisons le lemme 19.69 pour faire $-\kappa(s)\gamma'(s) = J\gamma''(s)$.

Le tout est une contradiction de la forme $0 \neq a = 0$.

Nous avons donc au moins un troisième point de changement de signe de κ'. Vu que la courbe est périodique, il en faut un nombre pair et donc un quatrième. □

L'exemple de l'ellipse montre qu'il n'y a pas lieu de chercher d'autres extrema de κ à part les 4 déjà trouvés.

Exemple 19.101
Nous trouvons les sommets de l'ellipse.

$$\gamma(t) = \big(a\cos(t), b\sin(t)\big) \tag{19.304a}$$
$$\gamma'(t) = \big(-a\sin(t), b\cos(t)\big) \tag{19.304b}$$
$$\gamma''(t) = -\big(a\cos(t), b\sin(t)\big) \tag{19.304c}$$
$$\tag{19.304d}$$

La courbure est

$$\kappa(t) = \frac{\gamma'' \cdot J\gamma'}{\|\gamma'\|^3} \tag{19.305a}$$
$$= \frac{-1}{[a^2\sin^2(t) + b^2\cos^2(t)]^{3/2}} \begin{pmatrix} a\cos(t) \\ b\sin(t) \end{pmatrix} \cdot \begin{pmatrix} -b\cos(t) \\ -a\sin(t) \end{pmatrix} \tag{19.305b}$$
$$= \frac{ab}{[a^2\sin^2(t) + b^2\cos^2(t)]^{3/2}}. \tag{19.305c}$$

Vu que $ab > 0$, les extrema de cela sont ceux du dénominateur et il suffit donc d'étudier les extrema de

$$f(t) = a^2\sin^2(t) + b^2\cos^2(t). \tag{19.306}$$

Nous avons

$$f'(t) = 2(a^2 - b^2)\cos(t)\sin(t), \tag{19.307}$$

fonction qui s'annule effectivement 4 fois sur une période. Deux maxima et deux minima. △

19.12.6 Le théorème de Jordan

Définition 19.102 ([251]).
*Une **courbe de Jordan** est une courbe simple fermée dans le plan.*

Le théorème suivant a un énoncé relativement simple, mais la démonstration est en réalité très longue.

Théorème 19.103 (Théorème de Jordan[251]).
Le complémentaire d'une courbe de Jordan Γ dans un plan affine réel est formé de exactement deux composantes connexes distinctes, dont l'une est bornée et l'autre non. Toutes deux ont pour frontière la courbe Γ.

Chapitre 20

Géométrie hyperbolique

20.1 Inversion

Lemme-définition 20.1.
Soit un cercle C et un point A du plan $\mathbb{R}^2$. Soit une droite passant par A et coupant C en deux points P et P' (pas spécialement distincts). Alors le nombre

$$\|AP\|\|AP'\| \tag{20.1}$$

ne dépend pas du choix de la droite et est nommé la ***puissance*** *du point A par rapport au cercle C.*

20.1.1 Cercles perpendiculaires

Définition 20.2.
Deux cercles sont ***perpendiculaires*** *lorsque leurs tangentes aux points d'intersections sont perpendiculaires.*

Lemme 20.3 ([1])**.**
Soient deux cercles perpendiculaires C_1 et C_2. Alors

(1) le centre de C_1 est hors de C_2.

(2) Si ℓ est une droite passant par ce le centre de C_1 (nommé O) et si ℓ coupe C_2 en les deux points P et P', alors P et P' sont situés du même côté de O.

Démonstration. Nous nommons O_1 le centre de C_1 ainsi que Q et Q' les points d'intersection de C_1 avec C_2. Si ℓ_Q et $\ell_{Q'}$ sont les tangentes à C_1 en Q et Q', alors ce sont des rayons de C_2 (parce que les cercles sont perpendiculaires). Par conséquent le centre O_2 de C_2 est le point d'intersection $Q' = \ell_Q \cap \ell_{Q'}$.

Le triangle O_1O_2Q est rectangle en Q et donc $\|O_1O_2\| > \|QO_2\|$. Or le nombre $\|QO_2\|$ est le rayon de C_2, donc O_1 est en dehors de C_2.

Ceci achève de prouver le point (1) ; nous démontrons le point (2). Les points P et P' sont sur le cercle C_2, donc tous les points du segment $[PP']$ sont dans le cercle. Or le centre de C_1 doit être en dehors de C_2 ; il ne peut donc pas être dans le segment $[PP']$, ce qui prouve que P et P' ne sont pas de part et d'autre de O sur la droite (OPP'). □

Proposition 20.4 ([1])**.**
Soit un cercle C_1 de centre O et de rayon R.

(1) Un cercle C_2 est perpendiculaire à C_1 si et seulement s'il existe une droite ℓ passant par O telle que les points d'intersection $\{P, P'\} = \ell \cap C_2$ soient situés du même côté de O et vérifient

$$\|OP\|\|OP'\| = R^2. \tag{20.2}$$

(2) Dans ce cas, toutes les droites passant par O et coupant C_2 en deux points P, P' vérifient le fait que P et P' soient du même côté de O et $\|OP\|\|OP'\| = R^2$.

Démonstration. Nous commençons par prouver le point (2). Le fait que les points P et P' soient du même côté de O est le lemme 20.3(2). Pour la relation sur les distances, soit $Q \in C_1 \cap C_2$. Vu que C_1 et C_2 sont perpendiculaires, la droite (OQ) ne coupe C_2 qu'en ce point, et la puissance de O par rapport au cercle C_2 est $\|OQ\|^2 = R^2$.

La même puissance peut être calculée via la droite ℓ :

$$\|OP\|\|OP'\|. \tag{20.3}$$

Donc $\|OP\|\|OP'\| = R^2$.

Soit un cercle C_2 passant par P et P'. Notons que P et P' ne sont pas sur C_1 parce qu'ils ne pourraient pas être alignés avec O. De plus l'un est à l'intérieur de C_1 et l'autre à l'extérieur de C_1. Les cercles C_1 et C_2 possèdent donc deux points distincts d'intersections. La puissance de O par rapport à C_2 est :

$$\|OP\|\|OP'\| = R^2 \tag{20.4}$$

parce que (OP) est une droite coupant C_2 en les points P et P'.

Soit Q un point d'intersection de C_1 et C_2, et Q' l'autre point d'intersection de C_2 avec la droite (OQ). La puissance de O par rapport à C_2 peut également être calculée à partir de cette droite (lemme 20.1) et nous avons

$$\|OQ\|\|OQ'\| = R^2, \tag{20.5}$$

mais $Q \in C_1$, donc $\|OQ\| = R$ et partant $\|OQ'\| = R$. Nous en déduisons que $Q' \in C_1$ également. Or Q' ne peut pas être l'autre point d'intersection de C_1 avec C_2 (sinon O, Q, Q' ne seraient pas alignés). Donc $Q = Q'$ et nous déduisons que la droite (OQ) est tangente à C_2.

Le rayon de C_1 est tangent à C_2. Cela signifie que C_1 est perpendiculaire à C_2. □

20.1.2 Inversion

Proposition-définition 20.5 ([195])**.**
Soit un cercle C de centre O dans $\mathbb{R}^2$. Il existe une unique application

$$i_C \colon \mathbb{R}^2 \backslash O \to \mathbb{R}^2 \backslash O \tag{20.6}$$

telle que

(1) $i_C(x) = x$ pour tout $x \in C$

(2) i_C échange l'intérieur et l'extérieur de C.

(3) i_C laisse invariants les droites et les cercles orthogonales à C.

*Cette application est l'**inversion** de cercle C.*

Démonstration. Soit P à l'intérieur de C, mais différent de O. Nous notons ℓ la droite (OP) et nous considérons un cercle C_2 passant par P et perpendiculaire à C (existence par la proposition 20.4). Nous avons $\ell \perp C$ (parce que ℓ est un rayon) et $C_2 \perp C$. Donc $i_C(\ell) = \ell$ et $i_C(C_2) = C_2$ par l'exigence (3). Mais comme $P \in \ell \cap C_2$ nous avons aussi

$$i_C(P) \in \ell \cap C_2. \tag{20.7}$$

Mais ℓ et C_2 se coupent en exactement deux points. Vu que $i_C(P)$ doit être hors de C (exigence (2)), avoir $i_C(P) = P$ est impossible. Nous en concluons que $i_C(P)$ doit être l'autre intersection.

Nous avons prouvé que les conditions (2) et (3) fixent l'image d'un point situé dans l'intérieur de C.

Si P est extérieur au cercle C, la même procédure fonctionne : nous considérons la droite $\ell = (OP)$ et un cercle C_2 perpendiculaire à C et passant par P. Encore une fois, ces deux objets sont fixés par i_C, et vu que $i_C(P)$ doit être à l'intérieur de C, il est fixé.

L'unicité est montrée.

En ce qui concerne l'existence, si $P \neq O$, la procédure suivante donne P' :

— Soit la droite $\ell = (OP)$.
— Soit un cercle C_2 perpendiculaire à C et passant par P.
— Le point P' est le point de l'intersection $C_2 \cap \ell$ qui n'est pas P.

Il est aisé de vérifier que poser $i_C(P) = P'$ donne une application qui vérifie toute les propriétés. □

20.6.
Nous recopions la construction de l'inversion d'un point par rapport à un cercle. Si C_1 est un cercle de centre O et P est un point différent de O, alors la procédure suivante construit $P' = i_{C_1}(P)$:

— Soit la droite $\ell = (OP)$.
— Soit un cercle C_2 perpendiculaire à C_1 et passant par P.
— Le point P' est le point de l'intersection $C_2 \cap \ell$ qui n'est pas P.

Que se passe-t-il si ℓ et C_2 n'ont qu'une seule intersection ? Alors la droite $\ell = (OP)$ est tangente à C_2. Or de O il n'existent que deux droites tangentes à C_2, et ce sont les rayons passant par les intersections parce que les cercles sont perpendiculaires. En d'autres mots, cette situation se présente lorsque P est sur le cercle C_1. Dans ce cas, $i_{C_1}(P) = P$.

Remarque 20.7.
Lorsque nous disons qu'une inversion «conserve les droites passant par O», il y a pour sous-entendu que nous considérons la droites privée du point O, parce que de toutes façons, l'inversion n'est pas définie sur O.

Nous allons résoudre cet intéressant problème en 21.6.1, en ajoutant le point ∞ à toutes les droites.

Corollaire 20.8.
L'inversion i_C est une involution ($i_C^2 = \mathrm{Id}$).

Démonstration. Soit un cercle C de centre O et un point P. Si C_2 est un cercle perpendiculaire à C passant par P, alors nous avons vu en 20.6 que $P' = i_C(P)$ est l'autre intersection entre C_2 et la droite (OP).

Pour construire l'image de P', il faut un cercle perpendiculaire à C passant par P'. Le cercle C_2 déjà utilisé fait l'affaire. Ensuite, la droite (OP') est la même que la droite (OP). Donc l'image de P' est P. □

Soit C_1 le cercle dans $\mathbb{C}$ de centre 0 et de rayon 1. Nous notons $\alpha\colon \mathbb{C} \to \mathbb{C}$ la dilatation de rapport R.

Proposition 20.9 ([1])**.**
Soit C le cercle de rayon R centré en 0 ($C = \alpha(C_1)$). Alors

$$i_{\alpha(C_1)} \circ \alpha = \alpha \circ i_{C_1}. \tag{20.8}$$

Démonstration. Soit $z = re^{i\theta}$; nous devons prouver que

$$i_C\big(\alpha(z)\big) = \alpha\big(i_{C_1}(z)\big). \tag{20.9}$$

Nous avons :

$$i_C\big(\alpha(z)\big) = i_C(Rre^{i\theta}) = R^2\frac{1}{Rr}e^{i\theta} = R\frac{1}{r}e^{i\theta} = \alpha\big(\frac{1}{r}e^{i\theta}\big) = \alpha\big(i_{C_1}(z)\big). \tag{20.10}$$

□

La définition 21.67 pour l'inversion sur $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ sera basée sur la proposition suivante.

Proposition 20.10 ([1])**.**
Soit le cercle de rayon R et de centre $a \in \mathbb{C}$. Alors

$$i_C(z) = \frac{R^2}{\bar{z} - \bar{a}} + a. \tag{20.11}$$

De plus si t_a est la translation de vecteur a, nous avons la décomposition

$$i_{t_a(C_R)} = t_a \circ i_{C_R} \circ t_{-a} \tag{20.12}$$

où C_R est le cercle de rayon R centré en 0.

Démonstration. Si $z \in \mathbb{C}$ et z' est son image par i_C, alors non seulement

$$\|a - z'\|\|a - z\| = R^2, \tag{20.13}$$

mais en plus $a - z' = \lambda(a - z)$ pour un certain $\lambda > 0$. Cela est l'expression du fait que z' est sur la demi-droite qui joint a à z. Nous avons donc

$$\lambda\|a - z\|^2 = R^2 \tag{20.14}$$

et alors

$$\lambda = \frac{R^2}{(a - z)(\bar{a} - \bar{z})}. \tag{20.15}$$

En récrivant $a - z' = \alpha(a - z)$ avec cette valeur de λ nous trouvons

$$a - z' = \frac{R^2}{\overline{a - z}}, \tag{20.16}$$

ce qu'il fallait démontrer.

La décomposition demandée est une simple vérification en utilisant $i_{C_R}(z) = \frac{R^2}{\bar{z}}$ qui découle de la proposition 20.9. □

Avant d'aller plus loin, donnons l'équation d'un cercle dans $\mathbb{C}$. Si C est un cercle de entre ω et de rayon r, alors $z \in C$ si et seulement si $d(z, \omega) = r$. En développant, et en passant au carré sans perte d'information (les deux membres sont positifs), $z \in C$ si et seulement si

$$(z - \omega)(\bar{z} - \bar{\omega}) = r^2. \tag{20.17}$$

Proposition 20.11 ([195, 252])**.**
Soit un cercle C de centre O. L'inversion

$$i_C \colon \mathbb{R}^2\backslash\{0\} \to \mathbb{R}^2\backslash\{0\} \tag{20.18}$$

transforme

(1) les droites passant par O sur elles-mêmes ;

(2) les cercles passant par O en des droites ne passant pas par O ;

(3) les droites ne passant pas par O en des cercles passant par O ;

(4) les cercles ne passant pas par O en des cercles ne passant pas par O.

Démonstration. Point par point.

(1) Une droite passant par O est une droite perpendiculaire à C. Par le point (3) de la définition 20.5, elle est invariante.

(2) Nous construisons successivement :
- Un cercle C_1 de centre O_1 et passant par O. Le but est de déterminer l'image de ce cercle.
- Le point P de C_1 tel que $[OP]$ en soit un diamètre.
- Le point $P' = i_C(P)$.
- La droite ℓ perpendiculaire à (OP) et passant par P'.

Nous montrons maintenant que $i_C(C_1\backslash\{O\}) = \ell$. Soit $Q \in C_1\backslash\{P, O\}$. Nous posons

$$Q' = (QO) \cap \ell. \tag{20.19}$$

Vu que $[OP]$ est un diamètre de C_1 et que $Q \in C_1$, le triangle OPQ est rectangle en Q. Et étant donné que Q' est sur ℓ nous savons que $OP'Q'$ est rectangle en P'.

De plus les angles en O de ces deux triangles sont identiques (parce que c'est l'angle formé par les droites (OQ) et (OP)) ; les triangles OPQ et $QP'Q'$ sont donc semblables et nous pouvons utiliser le théorème de Thalès [1] :

$$\frac{OP}{OQ} = \frac{OQ'}{OP'}. \tag{20.20}$$

Donc

$$\|OP\|\|OP'\| = \|OQ\|\|OQ'\|, \tag{20.21}$$

mais P' est l'image de P par l'inversion du cercle C, c'est à dire $\|OP\|\|OP'\| = R^2$. Nous en déduisons que

$$\|OQ\|\|OQ'\| = R^2, \tag{20.22}$$

c'est à dire que Q' est l'image de Q par i_C, et donc que

$$i_C\big(C_1\backslash\{O\}\big) \subset \ell. \tag{20.23}$$

Pour avoir l'inclusion inverse, il faut remarquer que ℓ est parallèle à la tangente à C_1 en O. Donc si $Q \in \ell$, la droite (OQ) intersecte le cercle C_1 en un point Q'. En refaisant le cheminement du résultat (20.23) à l'envers, il est loisible de prouver que $i_C(Q') = Q$ et donc que ℓ est bien inclue à l'image de C_1 par i_C.

(3) Nous commençons par prouver que toutes les droites ne passant pas par O sont des images de cercles passant par O.

Nous considérons :
— une droite ℓ ne passant pas par O.
— la droite d, perpendiculaire à ℓ passant par O
— le point $P' = \ell \cap d$,
— le point $P = i_C(P')$,
— le cercle C_1 dont $[OP]$ est un diamètre.

Par tout ce que nous avons fait jusqu'à présent, la droite ℓ est l'image du cercle C_1. Or si $\|OP\| = r$ alors

$$\|OP'\| = \frac{R^2}{r}. \tag{20.24}$$

Donc quelle que soit la valeur de $\|OP'\|$ dans $]0, \infty[$, il existera un point P tel que le cercle passant par O et P ait pour image la droite perpendiculaire à (OP) passant par $i_C(P)$.

Étant donné que i_C est une involution surjective des cercles passant par O vers les droites ne passant pas par O, elle transforme également toutes les droites ne passant pas par O en un cercle passant par O.

(4) Pour cette partie, nous allons utiliser un peu de géométrie analytique dans $\mathbb{C}$ [2] .

Le cas centré Nous supposons que C est centré en 0 et de rayon, 1, ce telle sorte que $i_C(z) = \frac{1}{\bar{z}}$. Soit C_1 un cercle de centre ω et de rayon r, ne passant pas par 0, en particulier tel que $|\omega| \neq r$.

Si $z \in i(C_1)$ alors $i(z) \in C_1$ et nous avons

$$\big(i(z) - \omega\big)\overline{\big(i(z) - \omega\big)} = r^2. \tag{20.25}$$

1. Faites bien le dessin : ce n'est pas une situation de Thalès ultra-standard de collège.
2. Principalement parce que je ne comprends pas le raisonnement fait dans [195].

En développant et en multipliant par $z\bar{z}$ nous trouvons

$$\bar{\omega}z + \omega\bar{z} + z\bar{z}(r^2 - \omega\bar{\omega}) = 1. \tag{20.26}$$

Nous pouvons diviser par $(r^2 - |\omega|^2)$ parce que C_1 ne passe pas par 0. En remettant en ordre, et en notant $s = r^2 - \omega\bar{\omega}$ pour plus de clarté,

$$\big(z - \frac{\omega}{s}\big)(\bar{z} - \frac{\bar{\omega}}{s}) - \frac{\omega\bar{\omega}}{s^2} = \frac{1}{s^2}, \tag{20.27}$$

ou encore

$$\big(z - \frac{\omega}{s}\big)(\bar{z} - \frac{\bar{\omega}}{s}) = \frac{1}{s} + \frac{\omega\bar{\omega}}{s}. \tag{20.28}$$

Pour que cela soit l'équation du cercle de centre $\frac{\omega}{s}$ et de rayon $\sqrt{\frac{1}{s} + \frac{\omega\bar{\omega}}{s^2}}$, il faut vérifier que

$$\frac{1}{s} + \frac{\omega\bar{\omega}}{s^2} \geqslant 0. \tag{20.29}$$

En multipliant par s^2, il s'agit de vérifier que $s + \omega\bar{\omega} \geqslant 0$, ce qui est correct parce que $s + \omega\bar{\omega} = r$.

En résumé, si $z \in i(C_1)$ alors z est dans le cercle C_2 de centre $\frac{\omega}{s}$ et de rayon $\frac{r}{s}$. Étant donné que $r \neq \omega$ nous savons que ce dernier cercle ne passe pas par 0.

Nous avons prouvé que $i(C_1) \subset C_2$. Pour prouver l'inclusion inverse, vu que i est une involution, il faut prouver $i(C_2) \subset C_1$. Pour cela nous écrivons l'équation qui donne $i(z) \in C_2$ et en développant nous devons conclure que $z \in C_1$. Nous ne le faisons pas ici.

Le cas de rayon non unité Pour un cercle quelconque, il faut passer par la formule (20.11).

Si $z \in i_C(C_1)$ alors $i(z) \in C_1$ et nous pouvons écrire

$$\big(i(z) - a\big)\big(\overline{i(z)} - \bar{a}\big) = R^2. \tag{20.30}$$

De là il faut déduire que z est sur un cercle ne passant pas par 0. Bons calculs...

□

Chapitre 21

Espaces projectifs

Sur les espaces projectifs : [253].

Définition 21.1.
*Soit E un espace vectoriel de dimension finie sur le corps commutatif $\mathbb{K}$. Nous définissons sur $E\backslash\{0\}$ la relation d'équivalence $u \sim v$ si et seulement si $u = \lambda v$ pour un certain $\lambda \in \mathbb{K}$. Cette relation est la relation de **colinéarité**. L'ensemble des classes d'équivalence de $\sim$ est l'**espace projectif** de E et sera noté $P(E)$.*

Définition 21.2.
*Si $\dim E = 2$, l'ensemble $P(E)$ est la **droite projective**, et si $\dim E = 3$ nous parlons du **plan projectif**.*

Étant donné que tous les $\mathbb{K}$-espaces vectoriels de dimensions $n+1$ sont isomorphes à $\mathbb{K}^{n+1}$, nous noterons $P_n(\mathbb{K})$ ou P_n l'espace projectif $P(\mathbb{K}^{n+1})$.

Exemple 21.3
Si $n = 1$ et $\mathbb{K} = \mathbb{R}$, l'espace projectif est l'ensemble des droites vectorielles dans le plan usuel. Il y en a une pour chaque point du type $(x, 1)$ avec $x \in \mathbb{R}$ et ensuite une horizontale, passant par le point $(1, 0)$. Nous avons donc

$$P_1(\mathbb{R}) = \{(1,0)\} \cup \{(x,1) \text{ tel que } x \in \mathbb{R}\}. \tag{21.1}$$

Le point $(1, 0)$ est dit «point à l'infini». △

21.1 Sous espaces projectifs

Un **sous-espace projectif** de $P(E)$ est une partie de la forme $P(F)$ où F est un sous-espace vectoriel de E.

Proposition 21.4.
Si F et G sont des sous-espaces vectoriels de E, alors

$$P(F) \cap P(G) = P(F \cap G) \tag{21.2}$$

et nous avons

$$\dim P(F) + \dim P(G) = \dim P(F+G) + \dim P(F \cap G). \tag{21.3}$$

Démonstration. Nous avons

$$P(F) = \{[v] \text{ tel que } v \in F\} \tag{21.4}$$

où les crochets signifient la classe par rapport à la relation de colinéarité. Nous avons alors

$$P(F) \cap P(G) = \{[v] \text{ tel que } v \in F \cap G\} = P(F \cap G). \tag{21.5}$$

Cela prouve le premier point.

En ce qui concerne l'équation (21.3), en considérant $\dim P(E) = \dim E - 1$ nous devons prouver l'égalité

$$\dim F + \dim G = \dim(F + G) + \dim(F \cap G) \tag{21.6}$$

concernant les dimensions des espaces vectoriels usuelles. Si nous considérons une base de E telle que $B_1 = \{e_1, \dots, e_{k_1}\}$ est une base de $F \cap G$, $B_2 = \{e_{k_1+1}, \dots, e_{k_2}\}$ complète B_1 en une base de F et $B_3 = \{e_{k_2+1}, \dots, e_n\}$ complète $B_1 \cup B_2$ en une base de G.

Nous avons alors

$$\dim F + \dim G = 2\operatorname{Card}(B_1) + \operatorname{Card}(B_2) + \operatorname{Card}(b_3) \tag{21.7a}$$
$$\dim(F + G) = \operatorname{Card}(B_1) + \operatorname{Card}(b_2) + \operatorname{Card}(B_3) \tag{21.7b}$$
$$\dim(F \cap G) = \operatorname{Card}(B_1). \tag{21.7c}$$

De là la relation (21.3) se déduit immédiatement. □

Théorème 21.5 (incidence)**.**
Soient F et F deux sous-espaces vectoriels de E tels que

$$\dim P(F) + \dim P(G) \geqslant \dim P(E). \tag{21.8}$$

Alors $P(F) \cap P(G) \neq \varnothing$.

Démonstration. En utilisant les hypothèses et la proposition 21.4 nous avons

$$\dim P(E) + \dim P(G) = \dim P(F + G) + \dim P(F \cap G) \geqslant \dim P(E). \tag{21.9}$$

En passant aux espaces vectoriels correspondants,

$$\dim(F + G) + \dim(F \cap G) \geqslant \dim(E) + 1. \tag{21.10}$$

Mais nous avons aussi $\dim(F + G) \leqslant \dim(E)$ et par conséquent $\dim(F \cap G) \geqslant 1$. Au final, $\dim P(F \cap G) \geqslant 0$. Cela prouve que $P(F \cap G)$ contient au moins un élément (nous rappelons que lorsqu'un espace projectif contient un seul élément, sa dimension est zéro). □

Exemple 21.6
Soient les plans $\Pi_1 \equiv x = 0$ et $\Pi_2 \equiv y = 0$. Nous avons

$$P(\Pi_1) = \{[0, y, 1]\} \cup \{[0, 1, 0]\} \tag{21.11a}$$
$$P(\Pi_2) = \{[x, 0, 1]\} \cup \{[1, 0, 0]\} \tag{21.11b}$$

où le crochet signifie la classe pour la colinéarité. Ces deux droites projectives ont comme point d'intersection le point $[0, 0, 1]$. △

Définition 21.7.
*Un **hyperplan projectif** est un sous-espace projectif de $P(E)$ de la forme $P(V)$ où V est un hyperplan de E.*

Définition 21.8.
*Soit E un espace vectoriel de dimension au moins 3. Nous disons que $d \subset P(E)$ est une **droite projective** de $P(E)$ si $d = P(D)$ pour une plan vectoriel $D \subset E$.*

*Nous disons que trois points de $P(E)$ sont **alignés** lorsqu'il existe une droite projective les contenant.*

21.9.
Dans la définition 21.8 nous voyons $P(D)$ comme inclus à $P(E)$ dès que D est un sous-espace vectoriel de E. Cela est possible parce que si la direction de $v \in D$, c'est à dire la classe $[v]$ est également une direction dans E.

Le lemme suivant peut paraître idiot, mais ce qui serait sûrement idiot est de l'utiliser sans s'en rendre compte.

Lemme 21.10.
Deux points dans $P(E)$ sont toujours alignés.

Démonstration. Soient deux points $A, B \in P(E)$. Si $A = \pi(a)$ et $B = \pi(b)$ alors le plan D passant par a, b et 0 est vectoriel et $P(D)$ contient A et B.

Note : si a, b et 0 sont trois points alignés, alors $A = B$. Il suffit de prendre les points a, c et 0 où $c \in E$ est un point quelconque non aligné avec 0 et a. Nous avons de toutes façons $A = B = \pi(a)$. □

Lemme 21.11.
Trois points distincts A, B, C dans $P(E)$ sont alignés si et seulement si il existe trois points non alignés $a, b, c \in E$ tels que

(1) le plan passant par a, b et c est vectoriel (c'est à dire passe par 0),

(2) $A = \pi(a)$, $B = \pi(b)$, et $C = \pi(c)$.

Démonstration. Deux implications à montrer.

Sens direct Soient A, B, C distincts et alignés dans $P(E)$. Alors il existe un plan vectoriel D tel que $A, B, D \in P(D)$.

La condition $A \in P(D)$ implique qu'il existe $a \in D$ tel que $A = \pi(A)$. Idem pour B et C. Les points a, b et c ainsi construits sont distincts parce que A, B et C sont distincts. Si par malheur ces trois points étaient alignés, ce n'est pas grave : il suffit de remplacer a par λa avec $\lambda \neq 0$ pour qu'ils ne le soient plus (cette manipulation ne change pas le fait que le nouveau choix de point a reste dans D parce que D est vectoriel). Nous avons donc trois points non alignés a, b et c tous contenus dans D. Le plan D répond à la question.

Sens réciproque Soient a, b et c non alignés dans E tels que $A = \pi(a)$, $B = \pi(b)$ et $C = \pi(c)$. Le plan D les contenant tous trois est vectoriel par hypothèse. Nous avons $A, B, C \in P(D)$ et donc A, B et C sont alignés dans $P(E)$.

□

Proposition 21.12.
Soit $H = P(V)$ un hyperplan projectif de $P(E)$ et soit m hors de H. Alors toute droite projective passant par m coupe H en un et un seul point.

Démonstration. Si $\dim E = n$ nous avons $\dim V = n - 1$. Soit $d = P(D)$ une droite projective passant par m, c'est à dire que D est de dimension 2 dans E. Si $D \subset V$ alors $m \in P(D) \subset P(V)$; or nous avons demandé que m soit hors de $P(V)$. Par conséquent D n'est pas inclus à V et en particulier $\dim(D + V) = \dim(E)$.

Nous recopions la formule (21.3) pour notre cas :

$$\underbrace{\dim d}_{=1} + \underbrace{\dim H}_{=n-2} = \underbrace{\dim P(D + V)}_{=n-1} + \dim P(D \cap V). \tag{21.12}$$

Nous avons donc $\dim P(D \cap V) = 0$, ce qui signifie que l'ensemble $P(D \cap V) = P(D) \cap P(V) = d \cap H$ contient un et un seul point. □

21.2 Espace projectifs comme «complétés» d'espaces affines

Soit E un espace vectoriel de dimension 2 et $P(E)$ la droite projective correspondante, et soit $\{e_1, e_2\}$ une base de E. Nous considérons la droite affine $d \equiv y = 1$. Nous avons la bijection

$$\begin{aligned} \phi\colon d \cup \{\infty\} &\to P(E) \\ (x, 1) &\mapsto \text{la droite vectorielle passant par } (x, 1) \\ \infty &\mapsto \text{la droite vectorielle passant par } (1, 0). \end{aligned} \tag{21.13}$$

Lemme 21.13.
Si nous munissons l'ensemble $d\cup\{\infty\}$ de la topologie compactifiée d'Alexandroff, la bijection (21.13) *est un homéomorphisme.*

Soient maintenant les plans affines dans l'espace vectoriel E de dimension 3

$$\Pi_1 \equiv z = 0 \tag{21.14a}$$
$$\Pi_2 \equiv z = 1. \tag{21.14b}$$

Une droite (vectorielle) de E coupe Π_2 en un et un seul point, sauf si elle est contenue dans Π_1. Nous avons donc une bijection

$$\begin{aligned}\phi\colon P(E) &\to \Pi_2 \cup P(\Pi_1)\\ d &\mapsto \begin{cases}\Pi_2 \cap d & \text{si cette intersection est non vide}\\ d & \text{sinon.}\end{cases}\end{aligned} \tag{21.15}$$

La droite projective $P(\Pi_1)$ est la droite à l'infini du plan projectif $P(E)$. Nous voyons que le plan projectif $P(E)$ peut être vu comme un plan affine (Π_2) «complété» par une droite affine $P(\Pi_1)$. Cette dernière droite est elle-même une droite affine complétée par un point à l'infini.

Nous pouvons généraliser cette démarche en considérant un espace affine $\mathcal{E}$ de direction E sur le corps $\mathbb{K}$. Nous construisons $F = E \times \mathbb{K}$ et nous considérons un repère affine sur F tel que $E \equiv x_{n+1} = 0$. Nous pouvons donc identifier $\mathcal{E}$ à l'hyperplan affine d'équation $x_{n+1} = 1$ dans F.

Une droite vectorielle de F non contenue dans E coupe $\mathcal{E}$ en un unique point ; nous avons donc une bijection

$$\mathcal{E} \cup P(E) \to P(F). \tag{21.16}$$

Dans ce cadre, $P(E)$ est l'hyperplan à l'infini et nous disons que $P(E)$ est la **complétion projective** de $\mathcal{E}$.

Exemple 21.14
Nous considérons les plans affines

$$\Pi_1 \equiv z = 0 \tag{21.17a}$$
$$\Pi_2 \equiv z = 1 \tag{21.17b}$$

et nous avons la bijection

$$P(E) = \Pi_2 \cup P(\Pi_1). \tag{21.18}$$

Un plan affine D a deux possibilités : soit il coupe Π_2 en une droite, soit il est égal à Π_1. Si $D \cap \Pi_2 = d$ (d est une droite affine), alors nous avons

$$P(D) = d \cup \{\infty_D\}, \tag{21.19}$$

ce qui justifie la terminologie comme quoi $P(D)$ est une droite dans $P(E)$. △

Soit E un espace vectoriel de dimension 3 et le plan projectif $P(E)$. Nous avons deux types de droites projectives :

(1) D'abord nous avons la droite à l'infini, donnée[1] par $P(z = 0)$.
(2) Ensuite nous avons toutes les droites affines du plan $z = 1$. Chacune de ces droites est complétée par un point à l'infini.

Exemple 21.15
Étudions un peu le second type de droites. D'abord si deux droites sont parallèles, leurs points à

1. Dans notre représentation usuelle du plan projectif $z = 1$.

l'infini sont identiques. Prenons par exemple les droites $d = \{z = 1, x = 1\}$ et $d' = \{z = 1, x = 2\}$. Elles décrivent les directions des vecteurs

$$\begin{pmatrix}1\\y\\1\end{pmatrix} \quad \text{et} \quad \begin{pmatrix}2\\y\\1\end{pmatrix}. \tag{21.20}$$

En normalisant, ce sont les vecteurs

$$\frac{1}{\sqrt{2+y^2}}\begin{pmatrix}1\\y\\1\end{pmatrix} \quad \text{et} \quad \frac{1}{\sqrt{5+y^2}}\begin{pmatrix}2\\y\\1\end{pmatrix}, \tag{21.21}$$

et toutes deux tendent vers le vecteur $(0, 1, 0)$ pour $y \to \infty$. △

Lemme 21.16.
Deux droites d'un plan projectif ont toujours une intersection.

Démonstration. Si les deux droites sont des droites affines non parallèles, le résultat est évident. Si elles sont parallèles, alors l'intersection est donnée par le point à l'infini comme indiqué dans l'exemple 21.15.

Supposons que d est la droite à l'infini tandis que d' est une droite affine. Dans notre représentation usuelle du plan affine, la droite à l'infini d a contient les vecteurs $(1, y, 0)$ et le point à l'infini $(0, 1, 0)$. La droite affine d' a pour équation paramétriques

$$\begin{cases} x = at + c & \text{(21.22a)}\\ y = bt + d & \text{(21.22b)}\\ z = 1. & \text{(21.22c)}\end{cases}$$

Les directions données par la droite d' sont donc

$$\frac{1}{a^2t^2 + b^2t^2 + c^2 + d^2}\begin{pmatrix}at + c\\bt + d\\1\end{pmatrix} \tag{21.23}$$

Son point à l'infini est la direction du vecteur $(a, b, 0)$, qui est bien un point de la droite à l'infini (éventuellement son point à l'infini [2]). □

La plupart du temps nous considérons le plan projectif comme étant le plan affine $z = 1$ de l'espace affine de dimension 3 complété par la droite affine $x = 1, z = 0$, elle-même complétée par le point $(0, 1, 0)$. Ce n'est évidemment pas la seule manière. Tout plan peut être considéré comme le plan à l'infini et pour une droite projective, tout point peut être considéré comme point à l'infini.

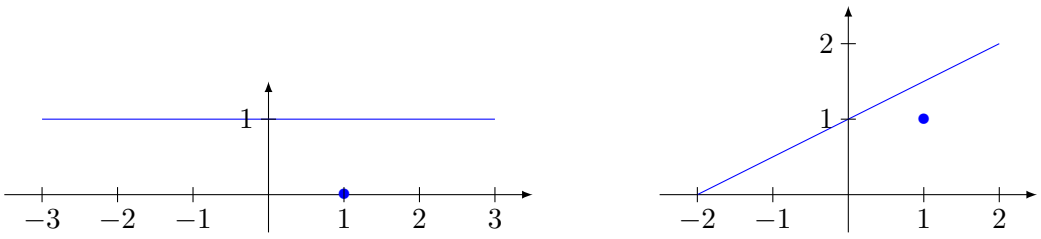
Sur la figure 21.1(a), le point à l'infini est la direction $(1, 0)$ tandis que la direction $(1, 1)$ n'a rien de spécial. À l'inverse sur la figure 21.1(b), la direction à l'infini est $(1, 1)$ tandis que la direction $(1, 0)$ est une direction usuelle.

Remarque 21.17.
Du point de vue de la topologie, si nous mettons celle de la compactification d'Alexandroff, tous les points de la droite projective sont équivalents.

Du point de vue de la géométrie différentielle, c'est la même chose. En effet nous pouvons mettre sur la droite projective un système de deux cartes en pensant aux angles. La première sur $]-a, a[$ avec par exemple $a < \pi/4$. La seconde carte serait $]a/2, \pi[$. Dans ce cas la direction $\theta = 0$ semble jouer un rôle spécial, mais il n'en est rien.

Nous pouvons également considérer les cartes $]\pi/4 - a, \pi/4 + a[$ et $]\pi/4 + a/2, 5\pi/4[$. Dans ces cartes, c'est plutôt le point $\theta = \pi/4$ qui semble différent (encore qu'il soit bien centré dans une carte).

2. D'accord, aller chercher le point à l'infini de la droite à l'infini, c'est chercher loin, mais n'empêche que ça existe.



(a) Ici le point à l'infini est la direction $(1,0)$.

(b) Ici le point à l'infini est la direction $(1,1)$.

FIGURE 21.1 – Deux façons de voir la droite projective. Étant donné que les points de la droite projective doivent être interprétés comme des directions (des classes d'équivallence), en réalité les deux dessins représentent les mêmes ensembles.

21.3 Théorème de Pappus

Théorème 21.18.
Soient deux droites d et d' dans un plan affine. Soient $A,B,C \in d$ et $A',B',C' \in d'$ tels que $AB' \parallel BA'$ et $BC' \parallel B'C$. Alors $AC' \parallel A'C$.

Démonstration. Si d et d' ne sont pas parallèles nous considérons o, le point d'intersection. Les relations de parallélisme des hypothèses impliquent qu'il existe λ_1 et λ_2 tels que

$$\begin{cases} A = \lambda_1 B & (21.24a)\\ B' = \lambda_1 A' & (21.24b) \end{cases}$$

et

$$\begin{cases} B' = \lambda_2 C' & (21.25a)\\ C = \lambda_2 B. & (21.25b) \end{cases}$$

En substituant nous trouvons

$$\begin{cases} C = \dfrac{\lambda_2}{\lambda_1} A & (21.26a)\\ A' = \dfrac{\lambda_2}{\lambda_1} C', & (21.26b) \end{cases}$$

ce qui implique que $A'C \parallel AC'$.

Si les droites d et d' sont parallèles, alors nous avons les translations

$$\begin{cases} B = A + x & (21.27a)\\ A' = B' + x & (21.27b) \end{cases}$$

et

$$\begin{cases} B = C + y & (21.28a)\\ C' = B' + y, & (21.28b) \end{cases}$$

ce qui montre que

$$\begin{cases} C = A + x - y & (21.29a)\\ A' = C' + x - y, & (21.29b) \end{cases}$$

et donc que $A'C \parallel AC'$. □

Le théorème suivant est une version projective.

Théorème 21.19.
Soient d et d' deux droites projectives d'un plan projectif. Soient $A,B,C \in d$ et $A',B',C' \in d'$. Alors les points $B'C \cap C'B$, $C'A \cap A'C$ et $A'B \cap B'A$ sont alignés.

Démonstration. Soient $E = BC' \cap C'B$ et $E' = C'A \cap A'C$. Ces deux points existent parce que deux droites projectives distinctes ont toujours un unique point d'intersection. Nous allons prendre EE' comme droite à l'infini et prouver que le point $A'B \cap B'A$ est dessus. Étant donné que le point d'intersection de $B'C$ et $C'B$ est à l'infini nous avons $B'C \parallel C'B$ (cela est un exemple de la flexibilité de la notion de parallélisme en géométrie projective). De la même façon nous avons $C'A \parallel A'C$.

Par le théorème de Pappus affine nous avons alors $A'B \parallel B'A$ et par conséquent le point d'intersection est sur la droite à l'infini, c'est à dire sur la droite EE'. □

21.4 Homographies

21.4.1 Homographies

Définition 21.20.
Soient E et F deux espaces vectoriels avec leurs projections naturelles

$$\pi_E \colon E\backslash\{0\} \to P(E) \tag{21.30a}$$
$$\pi_F \colon F\backslash\{0\} \to P(F). \tag{21.30b}$$

*Une application $g \colon P(E) \to P(F)$ est une **homographie** s'il existe un isomorphisme d'espaces vectoriels $\bar{g} \colon E \to F$ tel que le diagramme*

$$\begin{array}{ccc} E\backslash\{0\} & \xrightarrow{\bar{g}} & F\backslash\{0\} \\ {\scriptstyle \pi_E}\downarrow & & \downarrow{\scriptstyle \pi_F} \\ P(E) & \xrightarrow[g]{} & P(F) \end{array} \tag{21.31}$$

commute, c'est à dire s'il existe $\bar{g} \colon E \to F$ telle que

$$\pi_F\big(\bar{g}(v)\big) = g\big(\pi_E(v)\big) \tag{21.32}$$

pour tout $v \in E$.

Lemme 21.21.
Si $\bar{g} \colon E \to F$ est linéaire et si $\ker \bar{g} = \{0\}$ alors l'application g définie par

$$g\big(\pi_E(v)\big) = \pi_F\big(\bar{g}(v)\big) \tag{21.33}$$

est une homographie.

Démonstration. Nous devons simplement vérifier que l'équation (21.33) définit bien une application. Soient $v, w \in E$ tels que $\pi_E v = \pi_E w$; nous devons montrer que

$$\pi_F \bar{g} v = \pi_F \bar{g} w. \tag{21.34}$$

L'équation (21.34) sera vérifiée si et seulement s'il existe $\lambda \in \mathbb{R}$ tel que $\bar{g}v = \lambda \bar{g}w$, c'est à dire si et seulement si $\bar{g}(v - \lambda w) = 0$. Étant donné que nous supposons que le noyau de $\bar{g}$ est réduit à $\{0\}$, l'équation (21.34) sera vérifiée si et seulement si $v = \lambda w$, ce qui signifie exactement $\pi_E(v) = \pi_E(w)$. □

La proposition suivante donne les premières propriétés des homographies.

Proposition 21.22.
Quelques propriétés des homographies.

(1) Une homographie est bijective.

(2) Si deux espaces projectifs sont homographes, alors ils ont même dimension.

(3) L'ensemble des homographies $P(E) \to P(F)$ est un groupe (pour la composition).

(4) Une homographie conserve l'alignement des points.

Démonstration. Nous considérons une homographie $g\colon P(E) \to P(F)$, et $\bar{g}$ l'isomorphisme d'espaces vectoriels correspondant.

(1) Pour l'injectivité, si $g\big([v]\big) = g\big([w]\big)$ alors en utilisant la définition d'une homographie, $\pi_F\bar{g}v = \pi_F\bar{g}w$, ce qui implique que $\bar{g}v = \lambda\bar{g}w$, et donc $v = \lambda w$, ce qui signifie $[v] = [w]$.
Pour la surjectivité, un élément général de $P(F)$ prend la forme $\pi_F\bar{g}v$ pour un certain $v \in E$. Nous avons $g\big(\pi_E v\big) = \pi_F\bar{g}v$. Par conséquent l'élément $\pi_F\bar{g}v$ est bien dans l'image de g.

(2) Une homographie $P(E) \to P(F)$ n'existe que s'il existe un isomorphisme $E \to F$. Les dimensions sont donc automatiquement égales.

(3) Il suffit de vérifier que l'application

$$\begin{aligned}\varphi\colon P(E) &\to P(E)\\ \pi_F\bar{g}v &\mapsto \pi_E v\end{aligned} \tag{21.35}$$

est bien définie et donne l'inverse de g.

(4) Soient les points A, B, C alignés dans $P(E)$; ils correspondent à des directions de E qui sont données par des vecteurs situés sur la même droite affine. Autrement dit, il existe trois points $a, b, c \in E$ situés sur la même droite affine tels que $A, B, C = \pi_E(a, b, c)$. Les images par g sont données par $\pi_F\bar{g}a$, $\pi_F\bar{g}b$, et $\pi_F\bar{g}c$.
Étant donné qu'un isomorphisme d'espaces vectoriels conserve l'alignement affin, les points $\bar{g}a$, $\bar{g}b$ et $\bar{g}c$ sont alignés dans F. Cela implique que les projections par π_F sont alignés dans $P(F)$.

□

21.4.2 Le groupe projectif

Définition 21.23.
Le groupe des homographies de l'espace $P(E)$ est le ***groupe projectif****, noté* $\mathrm{PGL}(E)$.

Nous avons une surjection naturelle

$$\begin{aligned}\mathrm{GL}(E) &\to \mathrm{PGL}(E)\\ \bar{g} &\mapsto g\end{aligned} \tag{21.36}$$

qui s'avère être un morphisme de groupes.

Proposition 21.24.
Nous avons l'isomorphisme de groupes

$$\frac{\mathrm{GL}(E)}{\{\textit{homothéties}\}} \simeq \mathrm{PGL}(E). \tag{21.37}$$

Démonstration. Nous devons prouver que le noyau de l'application (21.36) est constitué des homothéties. Considérons un automorphisme d'espace vectoriel $f\colon E \to E$ dont l'homographie associée est l'identité, et prouvons que f est une homothétie. Nous avons le diagramme commutatif suivant :

$$\begin{array}{ccc} E\backslash\{0\} & \xrightarrow{f} & E\backslash\{0\} \\ {\scriptstyle\pi_E}\downarrow & & \downarrow{\scriptstyle\pi_E} \\ P(E) & \xrightarrow[\mathrm{Id}]{} & P(E). \end{array} \tag{21.38}$$

Pour tout vecteur $v \in E$ nous avons $\pi_E(v) = \pi_E\big(f(v)\big)$. Cela implique qu'il existe $\lambda \in \mathbb{R}$ tel que $f(v) = \lambda v$. Tous les vecteurs de E sont donc des vecteurs propres de f. Cela n'est possible que si toutes les valeurs propres sont identiques, c'est à dire que f est une homothétie. □

21.4.3 Repères projectifs

Si nous avons une base $\{e_i\}$ de $\mathbb{R}^n$ nous associons à $M \in P(E)$ les coordonnées $(X : Y : T)$. Mais si on prend la base $\{2e_1, e_2, \ldots, e_n\}$, les coordonnées du même point deviennent $(X/2 : Y : T)$ alors que du point de vue de l'espace projectif, rien n'a été changé : la classe de e_1 est la même que celle de $2e_1$. Les coordonnées homogènes ne sont donc pas intrinsèques.

Définition 21.25 ([254])**.**
*Des éléments $\{P_i\}_{i\in I}$ sont **projectivement independents** si en choisissant $v_i \in \pi^{-1}(P_i)$ nous obtenons des vecteurs $\{v_i\}_{i\in I}$ linéairement indépendants.*

Définition 21.26.
*Soit E un espace vectoriel de dimension $n+1$. Un **repère projectif** de $P(E)$ est la donnée de $n+2$ points $m_0, \ldots, m_{n+1}$ tels que*

(1) les vecteurs m_i, $i \neq 0$, sont les images d'une base $\{e_i\}$ de E

(2) $m_0 = \pi_E(e_1 + e_2 + \ldots + e_{n+1})$.

Note que si $m_k = \pi_E(v_k)$ $(k = 0, \ldots, n+1)$, alors tout choix de $n+1$ vecteurs parmi les v_k est une base de E.

Exemple 21.27
Un repère projectif de l'espace $P(\mathbb{R}^3)$ est par exemple les éléments $\{m_i\}_{i=1,\ldots,3}$ donnés par

$$m_1 = \pi(e_1) \tag{21.39a}$$
$$m_2 = \pi(e_2) \tag{21.39b}$$
$$m_3 = \pi(e_3) \tag{21.39c}$$
$$m_0 = \pi(e_1 + e_2 + e_3). \tag{21.39d}$$

△

Exemple 21.28
Pour $P(\mathbb{C}^2)$, un repère projectif possible est $m_1 = [1, 0]$, $m_2 = [0, 1]$, $m_0 = [1, 1]$. △

21.29.
Pourquoi voulons nous des repères projectifs ? Pourquoi demander un quatrième élément alors que trois devraient suffire ? Le fait est que si E est de dimension 3, nous voudrions pouvoir identifier E et $P(\mathbb{R}^3)$.

Plus précisément, si E est de dimension $n+1$ et possède une base $\{f_i\}_{i=1,\ldots,n+1}$, il existe un unique isomorphisme d'espaces vectoriels $E \to \mathbb{R}^{n+1}$ qui envoie cette base sur la base canonique de $\mathbb{R}^{n+1}$. La base de E étant fixée, nous pouvons donner à un point de E les coordonnées de son image dans $\mathbb{R}^{n+1}$ par cet isomorphisme *qui est unique*.

Dans le cas des espaces projectifs, nous voudrions avoir une unique homographie $\phi\colon P(E) \to P(\mathbb{R}^{n+1})$ qui permet de donner à un point $A \in E$ les coordonnées de $\pi^{-1}\big(\phi(A)\big)$. Bien entendu ce dernier n'est pas un élément bien défini de $\mathbb{R}^{n+1}$ parce qu'il y a toute une droite d'éléments de $\mathbb{R}^n$ qui se projettent sur $\phi(A)$.

L'idée d'imposer un point de plus est la bonne. Si nous imposons un point de plus, nous pouvons dire que les coordonnées de $A \in P(E)$ sont celles dans $\mathbb{R}^{n+1}$ de l'élément de $\pi^{-1}\big(\phi(A)\big)$ dont la dernière coordonnée est par exemple 1.

Nous allons maintenant mettre ça en musique.

D'abord nous donnons un exemple de non unicité.

Exemple 21.30
Soit un espace vectoriel E de dimension 2 et une base $\{b_1, b_2\}$ de E. Nous considérons également l'espace $\mathbb{R}^2$ muni de sa base canonique $\{e_1, e_2\}$.

Soit une homographie $\phi\colon P(E) \to P(\mathbb{R}^2)$ telle que

$$\phi\big(\pi(b_i)\big) = \pi(e_i) \tag{21.40}$$

pour $i = 1, 2$. Nous allons facilement construire une autre homographie qui vérifie les mêmes conditions.

L'idée est la suivante. L'espace $P(E)$ peut être vu comme la droite complétée $\{(x,1)\}_{x\in\mathbb{R}} \cup \{(1,0)\}$ et l'espace $P(\mathbb{R}^2)$ également. Une homographie respectant (21.40) doit envoyer le $(1,0)$ de E vers le $(1,0)$ de $\mathbb{R}^2$ et le $(1,1)$ de E vers le $(1,1)$ de $\mathbb{R}^2$. Mais en ce qui concerne le reste de la droite, l'homographie peut la parcourir à la vitesse qu'elle veut.

Il faut envoyer

Soit donc une homographie $\phi\colon P(E) \to P(\mathbb{R}^2)$, et nous définissons

$$\begin{aligned} \phi'\colon P(E) &\to P(\mathbb{R}^2) \\ \pi(xb_1 + yb2)&\phi\big(\pi(xb_1 + \lambda y b_2)\big) \end{aligned} \tag{21.41}$$

pour un certain $\lambda \neq 1$. En ce qui concerne le relèvement, l'application $\bar{\phi}'\colon E \to \mathbb{R}^2$ donnée par

$$\bar{\phi}'(xb_1 + yb_2) = \bar{\phi}(xb_1 + \lambda y b_2) \tag{21.42}$$

est bien définie et vérifie

$$\pi_{\mathbb{R}^2} \circ \bar{\phi} = \phi \circ \pi_E. \tag{21.43}$$

Donc ϕ' est une homographie. De plus

$$\phi'\big(\pi(b_1)\big) = \phi\big(\pi(b_1)\big) \tag{21.44a}$$

$$\phi'\big(\pi(b_2)\big) = \phi\big(\pi(\lambda b_2)\big) = \phi\big(\pi(b_2)\big) \tag{21.44b}$$

parce que $\pi(\lambda b_2) = \pi(b_2)$.

Nous n'avons donc pas l'unicité. △

C'est pour rétablir cette unicité que nous demandons d'avoir un point de plus pour avoir un repère projectif. De cette façon nous aurons une unique homographie $\phi\colon P(E) \to P(\mathbb{R}^{n+1})$ vérifiant $\phi\big(\pi_E(b_i)\big) = \pi_{\mathbb{R}^{n+1}}(e_i)$ pour tout $i = 0, \dots, n+1$.

Lemme 21.31.
Soit un espace vectoriel E de dimension $n+1$ muni de deux bases $\{e_i\}_{i=1,\dots,n+1}$ et $\{f_i\}_{i=1,\dots,n+1}$. Soit un repère projectif $\{m_0, m_i\}_{i=1,\dots,n+1}$ de $P(E)$.

Si $\pi(e_i) = \pi(f_i) = m_i$ pour tout $i = 1, \dots, n+1$ et si

$$\pi(e_1 + \dots + e_{n+1}) = \pi(f_1 + \dots + f_{n+1}) \tag{21.45}$$

alors les deux bases sont proportionnelles : il existe λ tel que $f_i = \lambda e_i$ pour $i = 1, \dots, n+1$.

Démonstration. Nous avons $\pi(e_i) = \pi(f_i)$ pour tout $i = 1, \dots, n+1$. Donc pour chaque $i = 1, \dots, n+1$ il existe $\lambda_i \in \mathbb{K}$ tel que $e_i = \lambda f_i$. Nous devons voir que les λ_i sont en réalité tous égaux.

Pour cela nous avons aussi l'égalité pour $i = 0$:

$$\pi(e_1 + \dots + e_{n+1}) = \pi(f_1 + \dots + f_{n+1}), \tag{21.46}$$

ce qui donne un $\mu \in \mathbb{K}$ tel que $e_1 + \dots + e_{n+1} = \mu(f_1 + \dots + f_{n+1})$, c'est à dire

$$\lambda_1 f_1 + \dots + \lambda_{n+1} f_{n+1} = \mu f_1 + \dots + \mu f_{n+1}. \tag{21.47}$$

Du fait que les f_i forment une base, cette égalité impose à tous les λ_i d'être égal à μ. □

Théorème 21.32 ([255]).
Soient $P(E)$ et $P(F)$ deux espaces projectifs de dimensions n.

(1) Une homographie $P(E) \to P(F)$ envoie un repère projectif sur un repère projectif.

(2) Si $(m_0, \dots, m_{n+1})$ est un repère projectif de $P(E)$, si $(m'_0, \dots, m'_{n+1})$ est un repère projectif de $P(F)$ alors il existe une unique homographie $g\colon P(E) \to P(F)$ telle que $g(m_i) = m'_i$ pour tout $i = 0, 1, \dots, n+1$

Démonstration. Un point à la fois.

(1) Soit une homographie $\phi\colon P(E) \to P(F)$ et un repère projectif $\{m_0, m_1, \dots, m_{n+1}\}$ de $P(E)$. Nous posons $m'_i = \phi(m_i)$ pour tout $i = 0, \dots, n+1$. Nous devons prouver que ces m'_i forment un repère projectif de $P(F)$.
D'abord pour $i = 1, \dots, n+1$ nous avons $m'_i = \phi\big(\pi_E(e_i)\big) = \pi_F\big(\bar{\phi}(e_i)\big)$, mais $\{\bar{\phi}(e_i)\}_{i=1,\dots,n+1}$ est une base de F parce que $\bar{\phi}$ est un isomorphisme d'espaces vectoriels. Donc oui : les m'_i $(i = 1, \dots, n+1)$ sont les projetés d'une base de F.
Nous posons au passage $f_i = \bar{\phi}(e_i)$. En ce qui concerne m_0 nous savons que $m_0 = \pi_E(e_1 + \dots + e_{n+1})$ et

$$m'_0 = \phi\big(\pi_E(e_1 + \dots + e_{n+1})\big) = \pi_F\big(\bar{\phi}(e_1 + \dots + e_{n+1})\big) = \pi_F(f_1 + \dots + f_{n+1}), \quad (21.48)$$

ce qui termine de montrer que $\{m'_i\}_{i=0,\dots,n+1}$ est un repère projectif de $P(F)$.

(2) Soient un repère projectif $(m_0, \dots, m_{n+1})$ de $P(E)$ et un repère projectif $(m'_0, \dots, m'_{n+1})$ de $P(F)$. Nous choisissons des bases $\{e_i\}$ de E et $\{f_i\}$ de F telles que

$$m_i = \pi_E(e_i) \quad (21.49a)$$
$$m'_i = \pi_F(f_i) \quad (21.49b)$$

pour $i = 1, \dots, n+1$ et

$$m_0 = \pi_E(e_1 + \dots + e_{n+1}) \quad (21.50a)$$
$$m'_0 = \pi_F(f_1 + \dots + f_{n+1}). \quad (21.50b)$$

Nous considérons un isomorphisme d'espace vectoriel $\bar{\phi}\colon E \to F$ tel que $\bar{\phi}(e_i) = f_i$ pour tout i, et nous voulons définir $\phi\colon P(E) \to P(F)$ par

$$\phi\big(\pi_E(v)\big) = \pi_F\big(\bar{\phi}(v)\big). \quad (21.51)$$

Cela est bien définit parce que si $\pi_E(v) = \pi_E(w)$ alors $w = \lambda v$ et

$$\pi_F\big(\bar{\phi}(\lambda v)\big) = \pi_F\big(\lambda\bar{\phi}(v)\big) = \pi_F\big(\bar{\phi}(v)\big). \quad (21.52)$$

L'application définie par (21.51) est une homographie qui envoie m_i sur m'_i pour tout $i = 0, \dots, n+1$. Ceci prouve la partie «existence» du point (2).
Pour l'unicité, soient des homographies

$$\phi_1\colon P(E) \to P(F) \quad (21.53a)$$
$$\phi_2\colon P(E) \to P(F) \quad (21.53b)$$

telles que $\phi_1(m_i) = \phi_2(m_i)$ pour tout $i = 0, \dots, n+1$. Soit aussi une base $\{e_i\}_{i=1,\dots,n+1}$ de E adaptée au repère projectif, c'est à dire $m_i = \pi_E(e_i)$ pour $i = 1, \dots, n+1$ et $\pi_E(e_1 + \dots + e_{n+1}) = m_0$. Nous considérons aussi les isomorphismes d'espaces vectoriels $\bar{\phi}_1$ et $\bar{\phi}_2$. Avec tout ce beau monde nous avons

$$\phi_1(m_i) = \pi_E\big(\bar{\phi}_1(e_i)\big) \quad (21.54a)$$
$$\phi_2(m_i) = \pi_E\big(\bar{\phi}_2(e_i)\big). \quad (21.54b)$$

Mais nous savons que $\phi_1(m_i) = \phi_2(m_i)$, donc nous savons que $\pi_E\big(\bar{\phi}_1(e_i)\big) = \pi_E\big(\bar{\phi}_2(e_i)\big)$, ce qui nous fait conclure que

$$\bar{\phi}_1(e_i) = \lambda_i \bar{\phi}_2(e_i) \tag{21.55}$$

pour certaines constantes $\lambda_i \in \mathbb{K}$. Le même raisonnement appliqué à m_0 nous donne un $\mu \in \mathbb{K}$ tel que

$$\bar{\phi}_1(e_1) + \ldots + \bar{\phi}_1(e_{n+1}) = \mu\big(\bar{\phi}_2(e_1) + \ldots + \bar{\phi}_2(e_{n+1})\big). \tag{21.56}$$

En mettant l'un dans l'autre :

$$\lambda_1\bar{\phi}_2(e_1) + \ldots + \lambda_{n+1}\bar{\phi}_2(e_{n+1}) = \mu\big(\bar{\phi}_2(e_1) + \ldots + \bar{\phi}_2(e_{n+1})\big). \tag{21.57}$$

Sachant que $\{\bar{\phi}_2(e_i)\}_{i=1,\ldots,n+1}$ est une base de F et nous souvenant de l'unicité de la décomposition d'un élément dans une base[3], nous en déduisons que tous les λ_i doivent être égaux à μ. Donc pour tout $v \in E$ nous avons $\bar{\phi}_1(v) = \lambda\bar{\phi}_2(v)$.
Cela a pour conséquence que $\phi_1 = \phi_2$.

□

21.33.
Si nous avons une droite projective, trois points sont nécessaires pour créer un repère et donc pour construire une homographie de la droite sur elle-même. Soit E un espace vectoriel de dimension 2 et $P(E)$ la droite projective qui lui est associée. Soit une homographie $f\colon P(E) \to P(E)$ et $\bar{f}\colon E \to E$, l'isomorphisme d'espaces vectoriels associé (par $f \circ \pi_E = \pi_E \circ \bar{f}$). Si $\{e_1, e_2\}$ est une base de E alors l'application $\bar{f}$ a une matrice

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{M}(2, \mathbb{K}) \tag{21.58}$$

avec $\det A \neq 0$ parce que $\bar{f}$ est un isomorphisme.

La plupart des points de $P(E)$ sont représentés par des points de la forme $(z, 1)$. Nous voudrions savoir quelle est la direction représentée par le point $\bar{f}(z, 1)$; c'est à dire que nous voudrions savoir $f([z, 1])$ sous la forme $[z', 1]$ (si possible). Nous avons

$$\bar{f}(z, 1) = (a_{11}z + a_{12}, a_{21}z + a_{22}). \tag{21.59}$$

Nous posons $\lambda = a_{21}z + a_{22}$ et nous avons

$$\bar{f}(z, 1) = \lambda\left(\frac{a_{11}z + a_{12}}{\lambda}, 1\right). \tag{21.60}$$

Il y a plusieurs possibilités suivant les valeurs de λ et de z.

(1) Si $\lambda = 0$ c'est que nous avons $\bar{f}(z, 1) = (a_{11}z + a_{12}, 0)$. L'application f envoie donc le point $(z : 1)$ sur le point à l'infini.
(2) Si $\lambda \neq$, alors f envoie le point $(z : 1)$ vers un autre point «normal».
(3) Si le point de départ est le point à l'infini alors $\bar{f}(1, 0) = (a_{11}, a_{21})$. Cela peut être le point à l'infini ou non selon les valeurs des a_{ij}.

Dans tous les cas si nous posons

$$\begin{cases} \varphi_f(z) = \dfrac{a_{11}z + a_{12}}{a_{21}z + a_{22}} & \text{(21.61a)} \\ \varphi_f(\infty) = \dfrac{a_{11}}{a_{21}} & \text{(21.61b)} \end{cases}$$

alors nous avons

$$\bar{f}(z, 1) = \big(\varphi_f(z), 1\big). \tag{21.62}$$

Si nous prenons la convention que $\frac{1}{0} = \infty$ et que $(\infty, 0)$ est le point à l'infini, alors cette application φ_f donne bien toutes les valeurs de f, y compris les cas à l'infini.

3. Proposition 7.6.

Lemme 21.34 ([256]).
Trois points distincts d'une droite projective forment un repère projectif.

Démonstration. Soit une droite projective $d = P(E)$ où E est un espace vectoriel de dimension 2 sur le corps $\mathbb{K}$. Soient trois points distincts A, B et C de d. Nous avons $a, b, c \in E$ tels que $A = \pi(a)$, $B = \pi(b)$ et $C = \pi(c)$. Vu que $A \neq B$, les vecteurs a et b ne sont pas proportionnels et la partie $\{a, b\}$ est libre dans E. Autrement dit, c'est une base [4].

Il existe donc $\alpha, \beta \in \mathbb{K}$ tels que $c = \alpha a + \beta b$. De plus α et β ne sont pas nuls parce que $C \neq A$ et $C \neq B$. En prenant $a' = \alpha a$, $b' = \beta b$ et $c' = c$ nous avons : $A = \pi(a')$, $B = \pi(b')$, $C = \pi(c')$ en même temps que $\{a', b'\}$ est une base de E et $c' = a' + b'$. Donc A, B, C est un repère projectif de $d = P(E)$. □

Corollaire 21.35 ([257]).
Soient E et F des espaces vectoriels de dimension 2. Soient A_i ($i = 1, 2, 3$) des points distincts sur $P(E)$ et B_i distincts sur $P(F)$. Alors il existe une unique homographie $P(E) \to P(F)$ portant A_i sur B_i pour tout $i = 1, 2, 3$.

Démonstration. Il s'agit de mettre en conjonction le lemme 21.34 qui dit que les A_i forment un repère projectif de $P(E)$ (idem : les B_i forment un repère projectif de $P(F)$) et le théorème 21.32(2) qui dit que l'une va sur l'autre par une unique homographie. □

21.4.4 Identifications $P(\mathbb{K}^2)$ vers $\mathbb{K} \cup \{\infty\}$

Pour rappel, une droite projective est l'espace projectif modelé sur un espace vectoriel de dimension deux (définition 21.8).

21.36.
Nous allons faire un usage assez intense de bijections entre $P(\mathbb{K}^2)$ et $\hat{\mathbb{K}} = \mathbb{K} \cup \{\infty\}$. Une possible est

$$\begin{aligned} \varphi_0 \colon P(\mathbb{K}^2) &\to \mathbb{K} \cup \{\infty\} \\ [k_1, k_2] &\mapsto \begin{cases} \frac{k_1}{k_2} & \text{si } k_2 \neq 0 \\ \infty & \text{si } k_2 = 0. \end{cases} \end{aligned} \tag{21.63}$$

Notons que nous utilisons ici le fait que $\mathbb{K}$ soit commutatif, sinon il aurait fallu choisir $k_1 k_2^{-1}$ ou $k_2^{-1} k_1$ au lieu d'écrire gentiment k_1/k_2.

Corollaire 21.37.
Soit une bijection $\varphi \colon P(\mathbb{K}^2) \to \hat{\mathbb{K}}$. Les points $\{\varphi^{-1}(\infty), \varphi^{-1}(0), \varphi^{-1}(1)\}$ forment un repère projectif de $P(\mathbb{K}^2)$.

Démonstration. Il s'agit seulement d'une application du lemme 21.34.

Juste pour l'amusement, nous allons le prouver explicitement pour la bijection $\varphi = \varphi_0$ donnée en (21.63). Un repère projectif est la définition 21.26. Nous avons

$$\varphi_0^{-1}(\infty) = [1, 0] = \pi_{K^2}\big((1, 0)\big) \tag{21.64a}$$
$$\varphi_0^{-1}(0) = [0, 1] = \pi_{K^2}\big((0, 1)\big) \tag{21.64b}$$
$$\varphi_0^{-1}(1) = [1, 1] = \pi_{K^2}\big((1, 1)\big) \tag{21.64c}$$

Les points $(1, 0)$ et $(0, 1)$ forment un base de $\mathbb{K}^2$ et nous avons bien $(1, 1) = (1, 0) + (0, 1)$. Donc le tout vérifie bien la définition d'un repère projectif. □

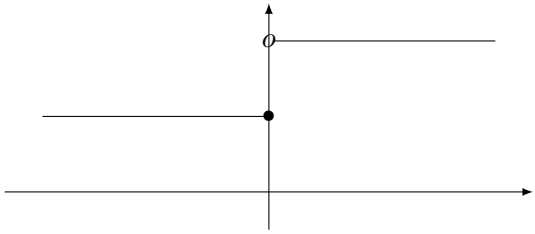
D'autre part, comme il est plus agréable de travailler avec $\hat{\mathbb{K}}$ qu'avec $P(\mathbb{K}^2)$ nous avons envie de voir $\hat{\mathbb{K}}$ comme un espace projectif (qu'il n'est pas). Il y a cependant nombre d'autres identifications

4. Il convient de citer ici le théorème 7.14(3).

possibles. En voici un autre :

$$\begin{aligned}\varphi_1\colon P(\mathbb{K}^2) &\to \hat{\mathbb{K}}\\ [k_1,k_2] &\mapsto \begin{cases}\frac{k_2}{k_1} & \text{si } k_1\neq 0\\ \infty & \text{si } k_1=0.\end{cases}\end{aligned} \tag{21.65}$$

Vous en voulez une plus compliquée ? En voici une pour $\mathbb{K}=\mathbb{R}$, basée sur le dessin suivant :



La bijection associée est :

$$\begin{aligned}\varphi_d\colon P(\mathbb{R}^2) &\to \hat{\mathbb{R}}\\ [k_1,k_2] &\mapsto \begin{cases}\frac{k_1}{k_2} & \text{si } k_1k_2\leqslant 0\\ 2\frac{k_1}{k_2} & \text{si } k_1k_2<0\\ \infty & \text{si } k_2=0.\end{cases}\end{aligned} \tag{21.66}$$

Pour mettre un peu d'ordre dans toutes ces identifications possibles, nous introduisons une classe.

Définition 21.38.
Pour une bijection $\varphi\colon P(\mathbb{K}^2)\to\hat{\mathbb{K}}$ nous définissons

$$A(\varphi)=\left\{\varphi_a\colon P(\mathbb{K}^2)\to\hat{\mathbb{K}} \text{ tel que } \varphi_a^{-1}\circ\varphi \text{ soit une homographie }\right\} \tag{21.67}$$

Les classes sont assez larges parce que pour toute homographie $\phi\colon P(\mathbb{K}^2)\to P(\mathbb{K}^2)$, nous avons $\varphi\circ\phi^{-1}\in A(\varphi)$. Mieux, nous avons le lemme suivant.

Lemme 21.39.
L'application

$$\begin{aligned}\psi\colon A(\varphi) &\to \mathrm{PGL}(\mathbb{K}^2)\\ \varphi_a &\mapsto \varphi_a^{-1}\circ\varphi\end{aligned} \tag{21.68}$$

est une bijection.

Démonstration. Pour rappel, $\mathrm{PGL}(E)$ est le groupe des homographies de E, voir la définition 21.23.

Surjectif Si $\phi\in\mathrm{PGL}(\mathbb{K}^2)$ nous avons $\phi=\psi(\varphi\circ\phi^{-1})$.

Injectif Si $\psi(\varphi_a)=\psi(\varphi_b)$ alors

$$\varphi_a^{-1}\circ\varphi=\varphi_b^{-1}\circ\varphi, \tag{21.69}$$

d'où nous déduisons $\varphi_a^{-1}=\varphi_b^{-1}$ parce que φ est une bijection.

□

21.4.5 Birapport

Proposition-définition 21.40.
Soit une droite projective $d=P(E)$ et trois points distincts A, B et C sur cette droite. Soit une bijection $\varphi\colon P(\mathbb{K}^2)\to\hat{\mathbb{K}}$. Si X est un point de d alors nous nommons le ***birapport*** *de X par rapport à A, B et C l'élément de $\hat{\mathbb{K}}$ donné par*

$$[A,B,C,X]_\varphi=(\varphi\circ\phi)(X) \tag{21.70}$$

où $\phi\colon d \to P(\mathbb{K}^2)$ *est l'unique homographie telle que*

$$\phi(A) = \varphi^{-1}(\infty) \tag{21.71a}$$
$$\phi(B) = \varphi^{-1}(0) \tag{21.71b}$$
$$\phi(C) = \varphi^{-1}(1). \tag{21.71c}$$

Démonstration. Nous devons prouver qu'il existe effectivement une unique homographie vérifiant les conditions (21.71).

A, B, C **est un repère projectif de** $P(E)$ Voir le lemme 21.34.

$\varphi^{-1}(\infty), \varphi^{-1}(0), \varphi^{-1}(1)$ **est un repère projectif de** $P(\mathbb{K}^2)$ Lemme 21.34 ou corollaire 21.37 au choix.

Conclusion Le théorème 21.32(2) nous donne existence et unicité d'une homographie $P(E) \to P(\mathbb{K}^2)$ envoyant le premier repère sur le second.

□

Remarque 21.41.
La majorité des sources ne parlent pas de la dependence du birapport en le choix de φ parce que tout le monde ne semble ne considérer que $\varphi = \varphi_0$ définie en (21.63). Il est cependant naturel de se demander si la définition dépend effectivement du choix de φ. La réponse est oui : ça dépend du choix.

Exemple 21.42(Une autre identification qui ne va pas bien)
Nous montrons que l'identification $\varphi_d\colon P(\mathbb{K}^2) \to \hat{\mathbb{K}}$ donnée en (21.66) ne donne pas lieu au même birapport que celui de φ_0.

Nous travaillons le birapport sur la droite projective la plus simple : $P(\mathbb{K}^2)$ avec $\mathbb{K} = \mathbb{R}$. Prenons pour la simplicité $A = [1, 0]$, $B = [0, 1]$ et $C = [1, 1]$. Alors l'homographie demandée dans la définition de $[.,.,.,.]_{\varphi_0}$ est $\varphi_0 = \mathrm{Id}$. Par conséquent,

$$\big[A, B, C, [k_1, k_2]\big]_{\varphi_0} = (\varphi_0 \circ \phi_0)[k_1, k_2] = \varphi_0[k_1, k_2] = \frac{k_1}{k_2}. \tag{21.72}$$

En ce qui concerne le birapport définit par φ_d nous avons

$$\varphi_d^{-1}(\infty) = [1, 0] \tag{21.73a}$$
$$\varphi_d^{-1}(0) = [0, 1] \tag{21.73b}$$
$$\varphi_d^{-1}(1) = [1, 2], \tag{21.73c}$$

de telle sorte que nous cherchons une homographie $\phi_d\colon P(\mathbb{K}^2) \to P(\mathbb{K}^2)$ telle que

$$\phi_d[1, 0] = [1, 0] \tag{21.74a}$$
$$\phi_d[0, 1] = [0, 1] \tag{21.74b}$$
$$\phi_d[1, 1] = [1, 2] \tag{21.74c}$$

L'homographie $\phi_d[k_1, k_2] = [k_1, 2k_2]$ convient et nous avons

$$\big[A, B, C, D, [k_1, k_2]\big]_{\varphi_d} = (\varphi_d \circ \phi_d)[k_1, k_2] = \varphi_d[k_1, 2k_2] = \frac{k_1}{2k_2} \tag{21.75}$$

dès que $k_1 k_2 < 0$. Nous avons donc bien trouvé

$$[A, B, C, X]_{\varphi_0} \neq [A, B, C, X]_{\varphi_d}. \tag{21.76}$$

△

Le birapport n'est pas un objet tout à fait canonique parce qu'il dépend effectivement du choix de l'identification entre $P(\mathbb{K}^2)$ et $\hat{\mathbb{K}}$.

Proposition 21.43.
Soit une bijection $\varphi\colon P(\mathbb{K}^2) \to \hat{\mathbb{K}}$. *Si* $\varphi_a \in A(\varphi)$ [5] *alors les birapports construits sur* φ *et* φ_a

5. $A(\varphi)$ définie en 21.38.

coïncident.

Démonstration. Soient trois points distincts $A, B, C \in P(E)$, et $X \in P(E)$. Nous avons

$$[A, B, C, X]_\varphi = (\varphi \circ \phi)(X) \tag{21.77}$$

où $\phi\colon P(E) \to P(\mathbb{K}^2)$ est l'unique homographie telle que

$$\phi(A) = \varphi^{-1}(\infty) \tag{21.78a}$$
$$\phi(B) = \varphi^{-1}(0) \tag{21.78b}$$
$$\phi(C) = \varphi^{-1}(1). \tag{21.78c}$$

Et

$$[A, B, C, X]_{\varphi_a} = (\varphi_a \circ \phi_a)(X) \tag{21.79}$$

où $\phi_a\colon P(E) \to P(\mathbb{K}^2)$ est l'unique homographie telle que

$$\phi_a(A) = \varphi_a^{-1}(\infty) \tag{21.80a}$$
$$\phi_a(B) = \varphi_a^{-1}(0) \tag{21.80b}$$
$$\phi_a(C) = \varphi_a^{-1}(1). \tag{21.80c}$$

Il est facile de voir que $\phi_a = \varphi_a^{-1} \circ \varphi\phi$. En effet, cela est une homographie parce que $\varphi_a \in A(\varphi)$, et parce que la composée d'homographies est une homographie. De plus,

$$(\varphi_a^{-1} \circ \varphi \circ \phi)(A) = (\varphi_a^{-1} \circ \varphi)\phi^{-1}(\infty) = \varphi_a(\infty) \tag{21.81a}$$
$$(\varphi_a^{-1} \circ \varphi \circ \phi)(B) = (\varphi_a^{-1} \circ \varphi)\phi^{-1}(0) = \varphi_a(0) \tag{21.81b}$$
$$(\varphi_a^{-1} \circ \varphi \circ \phi)(C) = (\varphi_a^{-1} \circ \varphi)\phi^{-1}(1) = \varphi_a(1). \tag{21.81c}$$

Au final nous avons :

$$[A, B, C, X]_{\varphi_a} = (\varphi_a \circ \varphi_a^{-1} \circ \varphi \circ \phi)(X) = (\varphi \circ \phi)(X) = [A, B, C, X]_\varphi. \tag{21.82}$$

□

Remarque 21.44.
Tout le monde semble ne considérer que l'identification usuelle $\varphi_0\colon P(\mathbb{K}^2) \to \hat{\mathbb{K}}$ donnée par $\varphi_0[k_1, k_2] = k_1/k_2$. Toute la discussion concernant la dépendance du birapport en le choix de l'identification (y compris la définition des classes $A(\varphi)$) peut être sautée en disant qu'on ne considéra que φ_0.

Et c'est ce que nous allons faire : sauf avis contraire, nous utiliserons le birapport associé à l'identification φ_0.

Lemme 21.45 ([255])**.**
Nous avons

$$[A, B, C, X]_\varphi = \begin{cases} \infty & \text{si et seulement si } X = A \\ 0 & \text{si et seulement si } X = B \\ 1 & \text{si et seulement si } X = C. \end{cases} \tag{21.83}$$

Démonstration. Par définition $[A, B, C, X]_\varphi = \varphi \circ \phi(X)$. Nous avons donc équivalence entre les affirmations suivantes :

— $[A, B, C, X]_\varphi = \infty$
— $(\varphi \circ \phi)(X) = \infty$
— $\phi(X) = \varphi^{-1}(\infty)$
— $\phi(X) = \phi(A)$
— $A = X$

parce que ϕ et φ sont des bijections.

Le même raisonnement tient pour les deux autres. □

Proposition 21.46.
Autres petites propriétés faciles ... Soit une droite projective $d = P(E)$ et trois points distincts $A, B, C \in d$.

(1) Les points A, B, C et X sont distincts si et seulement si $[A, B, C, X] \in \mathbb{K}\backslash\{0, 1\}$.

(2) Pour tout $k \in \hat{\mathbb{K}}$, il existe un unique $X \in d$ tel que $[A, B, C, X] = k$.

Démonstration. Notons pour le point (1) que l'énoncé demande déjà que A, B et C soient distincts. Sinon le birapport n'est pas définit.

(2) Les points A, B et C sont distincts par hypothèse. Vu le lemme 21.45, pour que X soit distincts de A, B et C il faut et il suffit que le birapport ne soit ni ∞ ni 1 ni 0. Donc $\mathbb{K}\backslash\{0, 1\}$.

(2) Nous avons $[A, B, C, X] = \phi(X)$ où $\phi\colon P(E) \to P(\mathbb{K}^2)$ est une homographie et donc une bijection par la proposition 21.22(1). Donc oui, pour tout éléments de $P(\mathbb{K}^2)$ il existe un unique élément de $P(E)$ dont le birapport par rapport à A, B et C soit cet élément.

Notons encore une fois que nous avons identifié $P(\mathbb{K}^2)$ à $\hat{\mathbb{K}}$ par la bijection (21.63). □

Proposition 21.47 ([255])**.**
Soient deux droites projectives d et d' ainsi que 4 points sur chacune : $A_1, A_2, A_3, A_4 \in d$, $A'_1, A'_2, A'_3, A'_4 \in d'$. Nous supposons que A_1, A_2, A_3 sont distincts et que $A'_1, A'_2, A'3$ également. Alors il y a équivalence entre

(1) Il existe une homographie $\phi\colon d \to d'$ telle que $\phi(A_i) = A'_i$ pour $i = 1, 2, 3, 4$,

(2) égalité des birapports :

$$[A_1, A_2, A_3, A_4] = [A'_1, A'_2, A'_3, A'_4]. \tag{21.84}$$

Dans ce cas, l'homographie est unique.

Démonstration. Nous divisons la preuve en trois parties évidentes.

(1) implique (2) Nous avons une homographie $\mu'\colon d' \to \hat{\mathbb{K}}$ telle que $\mu'(A'_1) = \infty$, $\mu'(A'_2) = 0$ et $\mu'(A'_3) = 1$. En composant [6] avec l'homographie $\phi\colon d \to d'$ de l'hypothèse nous avons une homographie $\mu' \circ \phi\colon d \to \hat{\mathbb{K}}$ qui vérifie

$$(\mu' \circ \phi)(A_1) = \mu'(A'_1) = \infty \tag{21.85a}$$
$$(\mu' \circ \phi)(A_2) = \mu'(A'_2) = 0 \tag{21.85b}$$
$$(\mu' \circ \phi)(A_3) = \mu'(A'_3) = 1, \tag{21.85c}$$

ce qui signifie que $\mu' \circ \phi$ est l'homographie qui définie le birapport sur d. Par conséquent

$$[A_1, A_2, A_3, A_4] = (\mu' \circ \phi)(A_4) = \mu'(A'_4) = [A'_1, A'_2, A'_3, A'_4]. \tag{21.86}$$

(2) implique (1) La partie $\{A_1, A_2, A_3\}$ est un repère projectif de d par le lemme 21.34. Idem pour $\{A'_1, A'_2, A'_3\}$. Nous considérons les homographies $\mu\colon d \to \hat{\mathbb{K}}$ et $\mu'\colon d' \to \hat{\mathbb{K}}$ définissant les birapports par rapport à ces repères. Ces homographies vérifient, en utilisant l'hypothèse :

$$\mu(A_4) = [A_1, A_2, A_3, A_4] = [A'_1, A'_2, A'_3, A'_4] = \mu'(A'_4). \tag{21.87}$$

Et de plus

$$\mu(A_1) = [A_1, A_2, A_3, A_1] = \infty = \mu'(A'_1) \tag{21.88a}$$
$$\mu(A_2) = [A_1, A_2, A_3, A_2] = 0 = \mu'(A'_2) \tag{21.88b}$$
$$\mu(A_3) = [A_1, A_2, A_3, A_3] = 1 = \mu'(A'_3). \tag{21.88c}$$

6. Proposition 21.22(3), la composition est encore une homographie.

Autrement dit : $\mu(A_i) = \mu'(A'_i)$ pour tout $i = 1, 2, 3, 4$. Nous nous inspirons de ce diagramme :

$$\begin{array}{ccc} d & \overset{\phi}{\dashrightarrow} & d' \\ & {\scriptstyle\mu}\searrow \quad \swarrow{\scriptstyle\mu'} & \\ & \mathbb{K} \cup \{\infty\} & \end{array} \tag{21.89}$$

La composée $\phi = \mu'^{-1} \circ \mu$ vérifie

$$(\mu'^{-1} \circ \mu)(A_i) = A'_i \tag{21.90}$$

pour tout i, et est une homographie.

Unicité Le fait qu'une homographie vérifiant $\phi(A_i) = A'_i$ pour $i = 1, 2, 3$ soit unique découle du fait qu'il existe une unique homographie portant un repère projectif sur un autre. A fortiori la condition $\phi(A_4) = A'_4$ ne retire rien à l'unicité.

□

Théorème 21.48 ([255])**.**
Une bijection entre deux droites projectives est une homographie si et seulement si elle conserve le birapport.

Démonstration. Chacun des deux sens séparément.

$\Rightarrow$ Soit une homographie $\phi\colon d \to d'$ entre deux droites projectives. Nous devons prouver que pour tout choix 4 points A, B, C, X dans d (dont A, B et C sont distincts) nous avons

$$[A, B, C, X] = [\phi(A), \phi(B), \phi(C), \phi(X)]. \tag{21.91}$$

Nous nommons $\mu\colon d \to \hat{\mathbb{K}}$ l'homographie qui donne le birapport sur d par rapport à A, B et C, et $\mu'\colon d' \to \hat{\mathbb{K}}$ celle qui donne le birapport sur d' par rapport à $\phi(A)$, $\phi(B)$, $\phi(C)$. Voici un diagramme de la situation :

$$\begin{array}{ccc} d & \overset{\phi}{\dashrightarrow} & d' \\ & {\scriptstyle\mu}\searrow \quad \swarrow{\scriptstyle\mu'} & \\ & \mathbb{K} \cup \{\infty\} & \end{array} \tag{21.92}$$

Nous prouvons maintenant que $\mu' = \mu \circ \phi^{-1}$. En effet :

$$(\mu \circ \phi^{-1})\big(\phi(A)\big) = \mu(A) = \infty \tag{21.93a}$$
$$(\mu \circ \phi^{-1})\big(\phi(B)\big) = \mu(B) = 0 \tag{21.93b}$$
$$(\mu \circ \phi^{-1})\big(\phi(C)\big) = \mu(C) = 1. \tag{21.93c}$$

Par conséquent le birapport à droite dans (21.91) peut se calculer à l'aide de $\mu \circ \phi^{-1}$:

$$[\phi(A), \phi(B), \phi(C), \phi(X)] = \mu'\big(\phi(X)\big) = (\mu \circ \phi^{-1})\big(\phi(X)\big) = \mu(X) = [A, B, C, X]. \tag{21.94}$$

La première implication est prouvée.

$\Leftarrow$ Soit une bijection $f\colon d \to d'$ conservant le birapport, ainsi que trois points distincts A, B et C dans d. Vu que f est une bijection les points $f(A)$, $f(B)$ et $f(C)$ sont distincts dans d'. Par le lemme 21.34 et le théorème 21.32(2), il existe une unique homographie $\phi\colon d \to d'$ telle que $\phi(A) = f(a)$, $\phi(B) = f(B)$ et $\phi(C) = f(C)$. Pour tout $X \in d$ nous avons

$$[f(A), f(B), f(C), f(X)] = [A, B, C, X] \tag{21.95a}$$
$$= [\phi(A), \phi(B), \phi(C), \phi(X)] \tag{21.95b}$$
$$= [f(A), f(B), f(C), \phi(X)]. \tag{21.95c}$$

Justifications :

— (21.95a) parce que f conserve le birapport par hypothèse.
— (21.95b) parce que ϕ conserve le birapport étant une homographie (c'est le premier sens du présent théorème)

Nous nommons $\mu'\colon d' \to \hat{\mathbb{K}}$ l'homographie donnant le birapport par rapport aux points $f(A)$, $f(B)$, $f(C)$. Alors le résultat (21.95) se lit

$$\mu'\big(f(X)\big) = \mu'\big(\phi(X)\big). \tag{21.96}$$

Mais comme μ' est une bijection (proposition 21.22(1)) cela implique $f(X) = \phi(X)$. Vu que nous avons fait ce raisonnement pour un X quelconque dans d nous avons $f = \phi$, ce qui prouve que f est une homographie.

□

Lemme 21.49.
Soient a, b, c distincts sur la droite projective $D = P(E)$. Soient $x, y \in E$ tels que $\pi_E(x) = a$, $\pi_E(y) = b$, $\pi_E(x + y) = c$. Alors

$$d = \pi_E(\lambda x + \mu y) \tag{21.97}$$

si et seulement si

$$[a, b, c, d] = \pi_{\mathbb{K}^2}(\lambda, \mu). \tag{21.98}$$

Démonstration. Étant donné que a et b sont distincts, les vecteurs x et y forment une base de E. Soit $f\colon E \to \mathbb{K}^2$ un isomorphisme qui envoie (x, y) sur e_1, e_2 où e_i sont les vecteurs de base de $\mathbb{K}^2$. Ensuite nous considérons $g\colon P(E) \to P(\mathbb{K}^2)$, l'homographie associée à f. Par définition $f\big(\pi_E z\big) = \pi_{\mathbb{K}^2}\big(f(z)\big)$. Par f nous avons

$$a \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad b \mapsto \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{21.99}$$

Donc par g nous avons

$$a \mapsto \infty \quad b \mapsto 0. \tag{21.100}$$

Nous avons aussi $f(\lambda x + \mu y) = (\lambda, \mu)$ et

$$\begin{aligned} g(c) &= g\big(\pi_E(x + y)\big) && \text{(21.101a)}\\ &= \pi_F f(x + y) && \text{(21.101b)}\\ &= \pi_F(f(x) + f(y)) && \text{(21.101c)}\\ &= \pi_F \begin{pmatrix} 1 \\ 1 \end{pmatrix} && \text{(21.101d)}\\ &= 1. && \text{(21.101e)} \end{aligned}$$

La dernière inégalité est le fait que la direction $(1, 1)$ dans $\mathbb{R}^2$ est représentée par le point $x = 1$ sur la droite $y = 1$ qui est notre «représentation» de la droite affine. L'application g a donc toutes les propriétés qu'il faut pour être l'application qui définit le birapport. Nous avons donc bien $g(d) = [a, b, c, d]$.

D'une part si $d = \pi_E(\lambda x + \mu y)$ alors

$$g(d) = \pi_{\mathbb{K}^2} f(\lambda x + \mu y) = \pi_{\mathbb{K}^2}(\lambda, \mu). \tag{21.102}$$

Dans l'autre sens si $[a, b, c, d] = \pi_{\mathbb{K}^2}(\lambda, \mu)$ alors supposons que $g(d) = \pi_{\mathbb{K}^2}(\lambda, \mu)$ avec $d = \pi_E(v)$ alors

$$g\pi_E v = \pi_{\mathbb{K}^2} f(v), \tag{21.103}$$

ce qui implique $f(v) = \alpha(\lambda, \mu)$ pour un certain $\alpha \in \mathbb{K}$. Par conséquent $v = \alpha(\lambda x + \mu y)$ et $d = \pi_E(\lambda x + \mu y)$. □

21.5 Coordonnées homogènes

Soit E un espace vectoriel de dimension $n+1$ et une base $\{e_0, \dots, e_n\}$ de E. Soit $M \in P(E)$ et $u \in E$ un élément engendrant M. Au point M nous voudrions associer les coordonnées $(x_0, \dots, x_n)$ de u dans E. Notons que toutes les coordonnées de u ne sont jamais nulles en même temps parce que u doit indiquer une direction. Nous savons par ailleurs que les coordonnées $(x_0, \dots, x_n)$ indiquent le même point de $P(E)$ que les coordonnées $(x'_0, \dots, x'_n)$ si et seulement si $x_i = \lambda x_i$.

Définition 21.50.
*La classe d'équivalence de $(x_0, \dots, x_n)$ est la **coordonnées homogène** de M. Nous la notons $(x_0 : \dots : x_n)$.*

21.5.1 Dualité

Soit E un espace vectoriel de dimension $n+1$. Une forme linéaire non nulle est un élément de E^*, mais aussi un représentant d'un élément de $P(E^*)$.

Le noyau d'une forme linéaire ω est un hyperplan. Le noyau de la forme linéaire $\lambda\omega$ étant le même hyperplan, l'hyperplan est donné par toute la classe de ω dans $P(E^*)$. Nous avons donc une bijection

$$P(E^*) \leftrightarrow \{\text{hyperplans vectoriels de } E\}. \tag{21.104}$$

Soit E de dimension 3 et une base $\{e_1, e_2, e_3\}$. L'espace dual E^* possède la base duale $\{e_1^*, e_2^*, e_3^*\}$. À un élément $m \in P(E^*)$ nous associons la droite

$$H_m\{(X : Y : T) \text{ tel que } m(X, Y, T) = 0\} \tag{21.105}$$

dans $P(E)$. Si les coordonnées homogènes de m étaient $(u : v : w)$ alors l'équation de la droite H_m est

$$uX + vY + wT = 0. \tag{21.106}$$

En effet si $\omega \in E^*$ est un représentant de m alors $\omega = \lambda(ue_1^* + ve_2^* + we_3^*)$ et l'équation (21.106) est indépendante de λ ainsi que du choix du représentant dans E du point $(X : Y : T)$ dans $P(E)$.

Si les points m_1 et m_2 sont distincts dans $P(E^*)$, ils donnent deux droites $m_1(X, Y, T) = 0$ et $m_2(X, Y, T) = 0$. Les points de la droite qui joint m_1 à m_2 dans $P(E^*)$ sont de la forme $\lambda m_1 + \mu m_2$ et ils sont associés à l'équation

$$\lambda m_1(X, Y, T) + \mu m_2(X, Y, T) = 0 \tag{21.107}$$

qui sont encore des droites dans $P(E)$. Toutes ces droites passent par le point d'intersection des droits associées à m_1 et m_2. Nous avons donc

$$\bigcap_{\lambda,\mu} H_{\lambda m_1 + \mu m_2} = H_{m_1} \cap H_{m_2}. \tag{21.108}$$

Lemme 21.51.
L'application

$$\begin{aligned} P(E^*) &\to \{\textit{droites dans } P(E)\} \\ m &\mapsto H_m \end{aligned} \tag{21.109}$$

est une bijection.

Démonstration. Une droite dans $P(E)$ est donnée en coordonnées homogènes par une équation $aX + bY + cT = 0$. Cette droite est décrite par le point $(a : b : c)$ dans $P(E^*)$. Ce dernier correspond à la direction de la forme $ae_1^* + be_2^* + ce_3^*$. Cela prouve que l'application est surjective.

Pour l'injectivité, si $m_1 \neq m_2$ dans $P(E^*)$, les formes ω_1 et ω_2 associées dans E^* ne sont pas multiples l'une de l'autre. Donc les équations

$$a_1X + b_1Y + z_1T = 0 \tag{21.110}$$

et

$$a_2X + b_2Y + z_2T = 0 \tag{21.111}$$

n'ont pas de solutions communes et décrivent donc des droites distinctes. □

Lemme 21.52.
Trois points distincts m_1, m_2 et m_3 dans $P(E^*)$ sont alignés si et seulement si les droites H_{m_1}, H_{m_2} et H_{m_3} sont distinctes et concourantes.

Démonstration. Supposons avoir trois points alignés, c'est à dire

$$m_3 = m_1 + \mu(m_2 - m_1). \tag{21.112}$$

Soit $X : Y : T$ le point d'intersection de H_{m_1} avec H_{m_2}. Alors $m_1(X, Y, T) = m_2(X, Y, T) = 0$. En tenant compte de (21.112) nous avons alors évidemment $m_3(X, Y, T) = 0$.

Supposons maintenant que les trois droites H_{m_i} soient concourantes. Nous avons donc un point $(X : Y : T)$ dans $P(E)$ tel que $m_i(X, Y, T) = 0$. Si m_i est la classe de $a_ie_1^* + b_ie_2^* + c_ie_3^*$ alors nous avons le système

$$\begin{cases} a_1X + b_1Y + c_1T = 0 & \text{(21.113a)} \\ a_2X + b_2Y + c_2T = 0 & \text{(21.113b)} \\ a_3X + b_3Y + c_3T = 0. & \text{(21.113c)} \end{cases}$$

Afin que cela ait une solution non triviale nous devons avoir

$$\det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \neq 0, \tag{21.114}$$

c'est à dire que les points (a_i, b_i, c_i) soient alignés. □

En tenant compte de ce qui a été dit, une droite dans $P(E^*)$ est constituée de points qui fournissent des droites concourantes dans $P(E)$. Donc une droite de $P(E^*)$ se caractérise par un point de $P(E)$ (l'intersection) de la façon suivante. Un point $M_d \in P(E)$ donne lieu à un **faisceau de droites** passant par M_d. Chacune de ces droites donne lieu à un point de $P(E^*)$ et tous ces points sont alignés. Nous avons ainsi construit la droite d dans $P(E^*)$ correspondante au point M_d de $P(E)$.

21.5.2 Polynômes

Soit l'espace projectif de dimension n avec ses coordonnées homogènes $(X_0 : \ldots : X_n)$. Nous considérons l'espace affine $H \equiv X_n = 1$ dans l'espace vectoriel E de dimension $n + 1$. Nous considérons pour H un repère affine ayant pour origine le point $(0, \ldots, 0, 1)$. Considérons un polynôme homogène P sur le corps $\mathbb{K}$. L'équation

$$P(X_0, \ldots, X_n) = 0 \tag{21.115}$$

sur l'espace vectoriel E descend immédiatement à l'espace projectif : étant donné que P est homogène nous avons $P(u) = 0$ si et seulement si $P(\lambda u) = 0$.

Nous essayons de décrire l'ensemble A des points de $P(E)$ satisfaisant $P(X_0, \ldots, X_n) = 0$. Nous savons que les éléments de $P(E)$ ont chacun un représentant soit dans H soit sur la droite à l'infini. Ceux de A ayant un représentant dans H sont d'équation

$$Q(x_0, \ldots, x_{n-1}) = 0 \tag{21.116}$$

où Q est le polynôme donné par $Q(X_0, \ldots, x_{n-1}) = P(x_0, \ldots, x_{n-1}, 1)$. Les points de A ayant un représentant sur la droite à l'infini s'obtiennent par l'équation

$$R(x_0, \ldots, x_{n-1}) = 0 \tag{21.117}$$

où R est le polynôme donné par $R(x_0, \ldots, x_{n-1}) = P(x_0, \ldots, x_{n-1}, 0)$.

Exemple 21.53
Nous considérons la conique projective

$$X^2 - XT - Y^2 - T^2 = 0. \tag{21.118}$$

Elle est décomposée en deux partie : une dans l'espace affine «normale» et une à l'infini. La première s'obtient en posant $T = 1$ dans (21.118) :

$$x^2 - x - y^2 - 1 = 0. \tag{21.119}$$

L'autre est obtenue en posant $T = 0$:

$$x^2 - y^2 = 0. \tag{21.120}$$

La partie à l'infini est donc composée de deux points : $(1 : 1 : 0)$ et $(1 : -1 : 0)$.

Le graphique de l'équation (21.119) est donné à la figure 21.2. Nous y voyons que les asymptotes sont effectivement données par les directions $(1, 1)$ et $(1, -1)$ dans le plan.

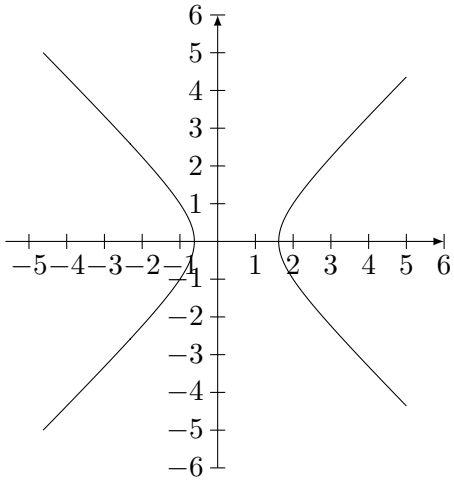


FIGURE 21.2 – Le graphique de $x^2 - x - y^2 - 1 = 0$.

△

Nous pouvons tenter de faire l'exercice inverse : considérer une conique dans $\mathbb{R}^2$, la voir comme une partie d'une conique dans l'espace projectif et trouver les points à l'infini qui la complètent.

Exemple 21.54
La droite projective usuelle est donnée par la droite affine $y - 1 = 0$. L'homogénéisation donne $y - z = 0$ et par conséquent la partie à l'infini est donnée par $y = 0$, c'est à dire la direction $(1, 0)$ comme il se doit. △

Exemple 21.55
Prenons la conique

$$x^2 + xy + y^3 - 2 = 0. \tag{21.121}$$

D'abord nous homogénéisons cette équation pour la voir dans $\mathbb{R}^3$:

$$x^2z + xyz + y^3 - 2z^3 = 0. \tag{21.122}$$

Les points à l'infini sont ceux qui correspondent à $z = 0$, c'est à dire la droite donnée en coordonnées homogènes par $(1 : 0 : 0)$. △

21.6 La sphère de Riemann $P_1(\mathbb{C})$

Définition 21.56.
*La **sphère de Riemann** est l'espace projectif modelé sur $\mathbb{C}^2$: en vertu des notations données à la page 1183, c'est*

$$P_1(\mathbb{C}) = P(\mathbb{C}^2). \tag{21.123}$$

L'ensemble $P_1(\mathbb{C})$ est le quotient $\mathbb{C}^2\backslash\{(0,0)\}/\sim$ où $\sim$ est la relation d'équivalence de $\mathbb{C}$-colinéarité dans $\mathbb{C}^2$.

Lemme 21.57.
L'application

$$\begin{aligned}\varphi_0\colon P_1(\mathbb{C}) &\to \mathbb{C}\cup\{\infty\}\\ [z_1,z_2] &\mapsto \begin{cases}\frac{z_1}{z_2} & si\ z_2\neq 0\\ \infty & si\ z_2=0\end{cases}\end{aligned} \tag{21.124}$$

est une bijection qui respecte la conjugaisons complexe : $\varphi_0\big([z_1,z_2]^*\big)=\varphi_0\big([z_1,z_2]\big)^*$.

Démonstration. Notons d'abord que la définition a un sens parce que si un représentant que $[z_1,z_2]$ est de la forme $(z,0)$ alors ils sont tous de cette forme. L'affirmation «$z_1\neq 0$ dans $[z_1,z_2]$» a donc un sens.

Injectif Supposons $\varphi_0\big([z_1,z_2]\big)=\varphi_0\big([t_1,t_2]\big)$.

Si les deux membres sont égaux à ∞ alors nous avons $z_2=t_2=0$, et alors avec $\lambda=z_1/t_1$ nous avons $(z_1,z_2)=\lambda(t_1,t_2)$, ce qui prouve que $[z_1,z_2]=[t_1,t_2]$.

Si les deux membres sont égaux à zéro alors $z_1=t_1=0$ et le même raisonnement tient.

Sinon nous avons $z_1/z_2=t_1/t_2$ où tous les nombres sont non nuls. Cela donne

$$z_2=\frac{t_2}{t_1}z_1, \tag{21.125}$$

et donc

$$\frac{t_1}{z_1}(z_1,z_2)=(t_1,t_2), \tag{21.126}$$

qui montre qu'au niveau des classes, $[z_1,z_2]=[t_1,t_2]$.

Surjectif Nous avons

$$\infty=\varphi_0\big([1,0]\big) \tag{21.127}$$

et si $z\neq\infty$ nous avons $z=\varphi_0\big([z,1]\big)$.

□

21.6.1 Éléments de géométrie dans $P_1(\mathbb{C})$

Étant donné que nous sommes partis pour faire de la géométrie dans $\mathbb{C}$ et même dans $\hat{\mathbb{C}}=\mathbb{C}\cup\{\infty\}$, autant nous armer des équations de cercles et de droites dans $\mathbb{C}$, ainsi que de quelque notions adjacentes.

Remarque 21.58.
La définition 21.2 parle de plan et de droites projectives. Ici nous ne sommes pas dans ce cadre parce que nous travaillons sur $P_1(\mathbb{C})$ où $\mathbb{C}$ n'est certainement pas un espace de dimension 3. Les droites dont nous allons parler ne sont pas des droites projectives avec leur point à l'infini.

21.6.1.1 Équation complexe d'une droite

L'équation d'une droite dans $\mathbb{R}^2$ est $d \equiv ax + by = c$ avec $a, b, c \in \mathbb{R}$ et a, b non nuls en même temps. En posant $z = x + iy$ nous voulons exprimer l'équation en termes de z au lieu de x et y. Nous avons[258]

$$x = \frac{z + \bar{z}}{2}, \quad y = \frac{z - \bar{z}}{2i}, \tag{21.128}$$

et nous pouvons écrire $d \equiv a(z + \bar{z}) - ib(z - \bar{z}) = 2c$, ou encore $d \equiv (a - bi)z + (a + ib)\bar{z} = 2c$. En posant $\omega = a + bi \in \mathbb{C}^*$ et $k = 2c \in \mathbb{R}$ nous avons l'équation

$$\bar{\omega}z + \omega z = k. \tag{21.129}$$

Définition 21.59.
*Une **droite** est une partie de $\hat{\mathbb{C}}$ de la forme*

$$d(\omega, k) = \{z \in \mathbb{C} \text{ tel que } \bar{\omega}z + \omega\bar{z} = k\} \cup \{\infty\} \tag{21.130}$$

avec $\omega \in \mathbb{C}^*$ et $k \in \mathbb{R}$.

Dans $\hat{\mathbb{C}}$, toutes les droites contiennent le point ∞.

Problèmes et choses à faire

La proposition 20.11 montre que toute inversion transforme un cercle-droite en un cercle-droite, nonobstant d'accepter de prolonger toute droite par ∞.

Est-ce que l'on peut dire que toutes les droites contiennent le point ∞ ?
En donnant ∞ à toutes les droites et à aucun cercle, la proposition 20.11 fonctionne partout en posant $i_C(O) = \infty$ et $i_C(\infty) = O$.

De plus en pensant à la projection stéréographique, ce serait logique : quelle que soit la direction dans laquelle un point s'éloigne de $z = 0$, son image par l'inverse de la projection stéréographique s'approche du pôle nord.

21.6.1.2 Équation complexe d'un cercle

Un cercle de centre $\omega \in \mathbb{C}$ et de rayon r a pour équation $|z - \omega| = r$, et nous avons les équivalences suivantes :

$$|z - \omega| = r \Leftrightarrow |z - \omega|^2 = r^2 \Leftrightarrow (z - \omega)(\bar{z} - \bar{\omega}) = r^2 \Leftrightarrow z\bar{r} - \bar{\omega}z - \omega\bar{z} = r^2 - |\omega|^2. \tag{21.131}$$

Donc un cercle de centre $\omega \in \mathbb{C}$ et de rayon $r \in \mathbb{R}$ a pour équation

$$z\bar{z} - \bar{\omega}z - \omega\bar{z} = r^2 - |\omega|^2. \tag{21.132}$$

Définition 21.60.
*Un **cercle** dans $\hat{\mathbb{C}}$ est une partie de la forme*

$$C(\omega, r) = \{z \in \mathbb{C} \text{ tel que } z\bar{z} - \bar{\omega}z - \omega\bar{z} = r^2 - |\omega|^2\} \tag{21.133}$$

avec $\omega \in \mathbb{C}$ et $r \in \mathbb{R}$.

Dans la sphère de Riemann, aucun cercle ne contient le point ∞.

Exemple 21.61
Trouvons le centre et le rayon du cercle d'équation

$$\bar{\omega}z + \omega\bar{z} = kz\bar{z} \tag{21.134}$$

avec $k \neq 0$. En divisant par k et en posant $\sigma = \omega/k$ nous avons :

$$z\bar{z} - \bar{\sigma}z - \sigma\bar{z} = 0. \tag{21.135}$$

Cela est un cercle de centre σ et de rayon $|\sigma|$. En effet si $z \in \mathbb{C}$ vérifie cette équation,

$$|z - \sigma|^2 = (z - \sigma)(\bar{z} - \bar{\sigma}) = \underbrace{z\bar{z} - \bar{\sigma}z - \sigma\bar{z}}_{=0} + |\sigma|^2 = |\sigma|^2, \tag{21.136}$$

c'est à dire que tous les points de $\mathbb{C}$ qui vérifient l'équation donnée sont à la distance $|\sigma|$ de σ. En particulier $z = 0$ est sur le cercle. △

21.6.1.3 Cercle-droite

Une chose de bien avec les équations complexes, c'est que nous pouvons écrire les droites et les cercles avec le même type d'équations.

Lemme-définition 21.62 ([258])**.**
*Un **cercle-droite** est l'ensemble des points $z \in \mathbb{C}$ tels que*

$$az\bar{z} - \bar{\omega}z - \omega\bar{z} = k \tag{21.137}$$

avec $a, k \in \mathbb{R}$ et $\omega \in \mathbb{C}$.

(1) Si $a = 0$, cela est une droite ;

(2) si $a \neq 0$, cela est un cercle.

(3) Un cercle-droite peut être l'ensemble vide.

Démonstration. Si $a = 0$ alors nous tombons tout de suite sur l'équation (21.129). Si $a \neq 0$ alors nous pouvons diviser par a, poser $\sigma = \omega/a$ et $l = k/a$ pour obtenir

$$z\bar{z} - \bar{\sigma} - \sigma\bar{z} = l, \tag{21.138}$$

qui est l'équation (21.132) d'un cercle ... ou pas tout à fait. En effet, (21.138) serait l'équation du cercle de centre σ et de rayon r donné par $l = r^2 - |\sigma|^2$, c'est) dire

$$r^2 = l + |\sigma|^2, \tag{21.139}$$

alors que rien n'assure que le nombre $l + |\sigma|^2$ soit positif. Dans le cas où c'est positif, nous avons bien un cercle. Sinon c'est l'ensemble vide. □

Remarque 21.63.
Lorsque nous parlons de cercle-droite, nous parlons de partie de $\mathbb{C}$ et non de $\hat{\mathbb{C}}$ parce que l'équation (21.137) a du mal à traiter le cas $z = \infty$. À cause du fait que nous avons décider de donner le point ∞ à toutes les droites, la fusion des notions de droites et de cercles n'est pas totale ; en tout cas pas en une seule équation.

Exemple 21.64([258])
Soit le cercle de centre $\omega = ir$ et de rayon r. Quelle que soit la valeur de $r > 0$, ce cercle passe par le point 0 et l'axe réel lui est tangent. L'équation de ce cercle est :

$$z\bar{r} + irz - ir\bar{z} = 0. \tag{21.140}$$

Vu que $ir \neq 0$ nous pouvons diviser et obtenir

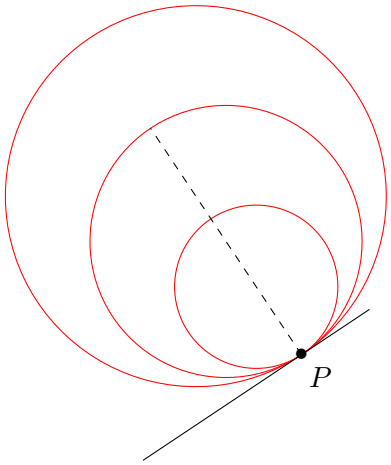
$$\frac{z\bar{z}}{ir} + z - \bar{z} = 0. \tag{21.141}$$

En faisant tendre r vers ∞ nous obtenons $z - \bar{z} = 0$, c'est à dire l'équation de la droite réelle.

Cela explique pourquoi il est souvent dit qu'une droite est un cercle dont le rayon est à l'infini. △

21.65.
Notons que l'exemple 21.64 est générique : prenez une droite ℓ, un point P sur ℓ, et considérez un cercle dont le centre est situé sur la perpendiculaire à ℓ passant par P, et dont le rayon est tel que le cercle passe par P. En prenant $|\omega - P| \to \infty$, l'équation du cercle devient celle de la droite ℓ.

Cela est particulièrement pratique lorsque nous travaillons dans $\hat{\mathbb{C}}$ parce nous y avons une notion précise du point à l'infini. Notons que (peut-être contre intuitivement), il existe un seul point à l'infini dans $\hat{\mathbb{C}}$. Et ce point est le centre de tous les cercles que l'on veut transformer en droites. Cela pose évidemment la question de savoir comment on définit précisément un cercle dont le centre est réellement ∞.



21.6.1.4 Rotation-homothétie

Définition 21.66.
*Une **rotation-homothétie** est une application $\hat{\mathbb{C}} \to \hat{\mathbb{C}}$ de la forme $z \mapsto \lambda z$ avec $\lambda \in \mathbb{C}$.*

Le nom provient du fait que si λ est réel, alors $z \mapsto \lambda z$ est une vraie homothétie, et si $\lambda = e^{i\theta}$ alors $z \mapsto e^{i\theta}z$ est une vraie rotation. Pour une valeur $\lambda \in \mathbb{C}$ générique, l'application $z \mapsto \lambda z$ est une composée des deux.

21.6.1.5 Application linéaire

Nous nous en voudrions de ne pas parler d'applications linéaires lorsque nous parlons de géométrie sur $\hat{\mathbb{C}}$. Soit $\alpha \in \mathbb{C}^*$ et $\beta \in \mathbb{C}$. Lorsque nous parlons de l'application linéaire

$$\begin{aligned} f\colon \hat{\mathbb{C}} &\to \hat{\mathbb{C}} \\ z &\mapsto \alpha z + \beta, \end{aligned} \tag{21.142}$$

nous entendons implicitement que $f(\infty) = \infty$.

21.6.1.6 Inversion

L'inversion d'un cercle de $\mathbb{R}^2$ est définie par la proposition 20.5. De nombreuses propriétés y sont décrites, y compris son écriture complexe dans la proposition 20.10. Tout cela était du temps de $\mathbb{R}^2$ ou de $\mathbb{C}$, mais maintenant nous sommes dans $\hat{\mathbb{C}}$ et nous voulons plus.

Définition 21.67 ([258])**.**
*Soit $\omega \in \mathbb{C}$ et $R \in \mathbb{R}^*$. L'**inversion** de centre ω et de **puissance** R^2 est l'application*

$$\begin{aligned} i\colon \hat{\mathbb{C}} &\to \hat{\mathbb{C}} \\ z &\mapsto \begin{cases} \dfrac{R^2}{\bar{z} - \bar{\omega}} + \omega & si\ z \in \mathbb{C}\backslash\{\omega\} \\ \infty & si\ z = \omega \\ \omega & si\ z = \infty. \end{cases} \end{aligned} \tag{21.143}$$

Notons que grâce aux conventions type $1/0 = \infty$ et $1/\infty = 0$, nous pouvons nous contenter de la première formule pour tout $z \in \hat{\mathbb{C}}$, et nous n'avons en réalité pas besoin de décrire $i(\infty)$ et $i(\omega)$ séparément.

Exemple 21.68
L'inversion de cercle de centre 0 et de rayon 1 est l'application $z \mapsto \frac{1}{\bar{z}}$, que l'on prolonge avec $0 \mapsto \infty$ et $\infty \mapsto 0$. △

21.6.2 Homographies

La notion d'homographie est la définition 21.20. Pour une homographie $\phi\colon P_1(\mathbb{C}) \to P_1(\mathbb{C})$ nous avons un isomorphisme d'espace vectoriel $\bar{\phi}\colon \mathbb{C}^2 \to \mathbb{C}^2$. Vue la bijection (21.124), nous voulons plutôt travailler avec $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ qui est un ensemble avec lequel nous sommes plus familier. Nous allons donc travailler avec

$$\begin{aligned}&\tilde{\phi}\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}\\ &\tilde{\phi} = \varphi_0 \circ \phi \circ \varphi_0^{-1}.\end{aligned} \tag{21.144}$$

Proposition 21.69.
L'application $\phi = \varphi_0^{-1} \circ \tilde{\phi} \circ \varphi_0$ est une homographie de $P(\mathbb{C}^2)$ si et seulement si l'application $\tilde{\phi}\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ est de la forme

$$\tilde{\phi}(z) = \frac{az + b}{cz + d} \tag{21.145}$$

avec $a, b, c, d \in \mathbb{C}$ tels que $ad - bc \neq 0$.

Par convention nous posons $z/0 = \infty$ dès que $z \neq 0$; en particulier $\tilde{\phi}(\infty) = a/c$ et $\tilde{\phi}(-d/c) = \infty$.

Démonstration. Nous séparons la condition suffisante de la condition nécessaire.

$\Rightarrow$ La condition $\pi \circ \bar{\phi} = \phi \circ \pi$ (de (21.32)) nous dit que

$$\big[\bar{\phi}(z_1, z_2)\big] = \phi\big([z_1, z_2]\big), \tag{21.146}$$

et comme $\bar{\phi}$ est un isomorphisme d'espace vectoriel nous avons $a, b, c, d \in \mathbb{C}$ vérifiant $ad - cb \neq 0$ pour lesquels

$$\bar{\phi}(z_1, z_2) = \begin{pmatrix} az_1 + bz_2 \\ cz_1 + dz_2 \end{pmatrix}. \tag{21.147}$$

Soit $z \in \mathbb{C}$. Alors nous avons

$$\tilde{\phi}(z) = (\varphi_0 \circ \phi)[z, 1] = \varphi_0\Big(\big[\bar{\phi}(z, 1)\big]\Big) = \varphi_0\big([az + b, cz + d]\big) = \frac{az + b}{cz + d}. \tag{21.148}$$

Il est important de comprendre que cette formule fonctionne pour tout $z \in \mathbb{C}$. En effet nous pourrions avoir un doute sur $z = -d/c$. D'abord si $c = 0$ alors $d \neq 0$ et ce problème n'existe pas : le dénominateur est toujours non nul. Nous avons donc seulement un doute lorsque $c \neq 0$. Dans ce cas,

$$\tilde{\phi}(-d/c) = \frac{-\frac{ad}{c} + b}{0}. \tag{21.149}$$

Mais $c \neq 0$, donc le numérateur est non nul. Or lorsque $z \neq 0$ nous avons posé $z/0 = \infty$, donc dans notre cas,

$$\tilde{\phi}(-d/c) = \infty \tag{21.150}$$

automatiquement, et cela est encodé dans la formule (21.148).

Il nous reste à déterminer $\tilde{\phi}(\infty)$. Nous avons :

$$\tilde{\phi}(\infty) = (\varphi_0 \circ \phi)\big([1, 0]\big) = \varphi_0\big[\bar{\phi}(1, 1)\big] = \varphi_0\big([a, c]\big) = \begin{cases} a/c & \text{si } c \neq 0 \\ \infty & \text{si } c = 0 \end{cases} \tag{21.151}$$

où la distinction entre les deux cas n'est pas fondamentale parce que si $c = 0$ alors $a \neq 0$ et $a/c = \infty$.

$\Leftarrow$ En notant $Z = [z_1, z_2]$ (pour $z_1, z_2 \in \mathbb{C}$) nous avons

$$\phi\big([z_1, z_2]\big) = \varphi^{-1}\left(\frac{a\varphi_0(Z) + b}{c\varphi_0(Z) + d}\right) = \big[a\varphi_0(Z) + b, c\varphi_0(Z) + d\big]. \tag{21.152}$$

Nous définissons $\bar{\phi}$ par son action sur les vecteurs de base : $\bar{\phi}(1, 0) = (a, c)$ et $\bar{\phi}(0, 1) = (b, d)$. Nous avons bien l'isomorphisme d'espace vectoriel $\bar{\phi}\colon \mathbb{C}^2 \to \mathbb{C}^2$ dont nous avons besoin pour

la définition 21.20. D'abord le fait que $ad-cb$ soit non nul assure que le $\bar{\phi}$ ainsi défini est bien bijectif[7]. Et de plus ce $\bar{\phi}$ vérifie la condition

$$\begin{aligned}\pi\big(\bar{\phi}(z_1,z_2)\big) &= \pi\big((az_1,cz_1)+(bz_2,dz_2)\big) && (21.153a)\\ &= \big[az_1+bz_2,cz_1+dz_2\big] && (21.153b)\\ &= \big[a\varphi_0(Z)+b,c\varphi_0(Z)+d\big] && (21.153c)\\ &= \phi\big(\pi(z_1,z_2)\big). && (21.153d)\end{aligned}$$

Nous avons utilisé la notion de classe pour diviser par z_2 et faire apparaître $\varphi_0(Z)$.

□

Remarque 21.70.
En prenant les conventions relativement claires $\infty\times a=\infty$ (pour $a\neq 0$) et $\infty\pm a=\infty$ (avec $a\neq\infty$), alors tout est dans la formule

$$\bar{\phi}(z)=\frac{az+b}{cz+d} \tag{21.154}$$

avec $ad-cb\neq 0$. Il n'y a pas besoin de traiter séparément le cas $z=\infty$ ou $z=-d/c$.

Définition 21.71.
*Nous aimons tellement l'identification $\varphi_0\colon P(\mathbb{C}^2)\to\hat{\mathbb{C}}$ que nous allons parler d'**homographie** sur $\hat{\mathbb{C}}$ pour les applications $\tilde{\phi}$ de la forme*

$$\tilde{\phi}(z)=\frac{az+b}{cz+d} \tag{21.155}$$

avec $ad-cb\neq 0$.

21.72.
La proposition 21.69 nous indique que les homographies de $\hat{\mathbb{C}}$ sont de la forme $\phi=\varphi_0^{-1}\circ\tilde{\phi}\circ\varphi_0$ pour une homographie $\phi\colon P(\mathbb{C}^2)\to P(\mathbb{C}^2)$.

Proposition 21.73 ([258])**.**
L'application $h\colon\hat{\mathbb{C}}\to\hat{\mathbb{C}}$ associée à une homographie est soit linéaire, soit de la forme $h=l_1\circ\iota\circ l_2$ où l_i sont linéaires et ι est l'application $z\mapsto 1/z$.

Démonstration. Commençons par une remarque : lorsque nous parlons d'une application linéaire, c'est au sens de la note 21.6.1.5 qui explique qu'une application linéaire sur $\mathbb{C}$ est automatiquement prolongée à $\hat{\mathbb{C}}$ par $f(\infty)=\infty$.

Soit donc l'application

$$h(z)=\frac{az+b}{cz+d}. \tag{21.156}$$

Si $c=0$, alors c'est une application linéaire et la preuve est terminée. Nous supposons que $c\neq 0$. Nous posons

$$l_2(z)=cz+d, \tag{21.157}$$

et ensuite $l_1(z)=\alpha z+\beta$ avec α et β à déterminer. Un peu de calcul :

$$(l_1\circ\iota\circ l_2)(z)=l_1\left(\frac{1}{cz+d}\right)=\frac{\alpha+\beta cz+\beta d}{cz+d}, \tag{21.158}$$

et en imposant que cela soit égal à $\frac{az+b}{cz+d}$ nous trouvons $\beta=a/c$ et $\alpha=b-ad/c$. Il est vite vérifier que ces choix donnent le bon résultat. □

7. Si vous n'en êtes pas convaincu, écrivez la matrice de l'application qui envoie $(1,0)$ sur (a,c) et $(0,1)$ sur (b,d), et demandez-vous sous quelle condition elle est inversible.

21.74.
Vu que les applications linéaires sont des composées d'une translation et d'une rotation-homothétie, et que l'application i est une composée d'une inversion $z \mapsto 1/\bar{z}$ et d'une réflexion $z \mapsto \bar{z}$, toutes les homographies sont des composées des éléments suivants :

— inversion [8] $z \mapsto 1/\bar{z}$, prolongée par $i(\infty) = 0$ et $i(0) = \infty$;
— réflexion $z \mapsto \bar{z}$;
— translation $z \mapsto z + \alpha$ avec $\alpha \in \mathbb{C}$;
— rotation-homothétie $z \mapsto \lambda z$ avec $\lambda \in \mathbb{C}$.

Toutes ces opérations sont prolongées à $\hat{\mathbb{C}}$ par $1/\infty = 0$, $\lambda \cdot \infty = \omega$ (si $\lambda \neq 0$) et $\infty + \lambda = \infty$. Nous ne définissons pas $0 \cdot \infty$ et $\infty - \infty$.

Certes nous pouvons construire des homographies à partir d'ingrédients dont la conjugaison complexe. Il ne faudrait cependant pas déduire que cette conjugaison est une homographie.

Lemme 21.75.
La conjugaison complexe n'est pas une homographie.

Démonstration. Si elle l'était nous aurions des nombres $a, b, c, d \in \mathbb{C}$ tels que $ad - bc \neq 0$ et

$$\frac{az + b}{cz + d} = \bar{z} \tag{21.159}$$

pour tout $z \in \mathbb{C}$.

En posant $z = 0$ nous avons déjà $b/d = 0$, c'est à dire $b = 0$. Avec $z = 1$ nous trouvons alors $a/(c + d) = 1$, c'est à dire

$$a = c + d. \tag{21.160}$$

Si $ci + d \neq 0$ Dans ce cas nous pouvons évaluer (21.159) en $z = i$ et avoir $a = -ci + d$. Mais comme nous avions déjà $a = c + d$ nous déduisons $c = 0$. Nous restons donc avec

$$\frac{a}{d} z = \bar{z} \tag{21.161}$$

pour tout z. En prenant $z = 1$ puis $z = i$, il est vite remarqué que cela n'est pas possible.

Si $ci + d = 0$ Nous rappelons que $ad \neq 0$. Nous écrivons l'équation avec $z = -i$ pour trouver

$$\frac{-ai}{-ci + d} = i, \tag{21.162}$$

qui donne immédiatement $a = ci - d$. Nous avons donc les trois équations

$$\begin{cases} c = id & \text{(21.163a)} \\ a = ci - d & \text{(21.163b)} \\ a = c + d. & \text{(21.163c)} \end{cases}$$

Une tentative de résolution tombe rapidement sur une impossibilité (en substituant la première dans les deux autres et en comparant les deux valeurs de a par exemple).

□

La proposition suivante ressemble à s'y méprendre à la proposition 20.11, mais elle diffère en deux points. D'abord elle ne traite que de l'inversion par rapport à l'origine, mais surtout, elle traite le point $z = \infty$. C'est un avantage de travailler sur $\hat{\mathbb{C}}$ plutôt que sur $\mathbb{R}^2$.

Proposition 21.76 (Inversion de cercles et de droites)**.**
L'inversion dans $\hat{\mathbb{C}}$ envoie

(1) une droite passant par 0 sur elle-même

8. Oui, c'est l'inversion de la géométrie hyperbolique, voir 21.6.1.6.

(2) une droite ne passant pas par 0 *sur un cercle passant par* 0.

(3) un cercle ne passant pas par 0 *en un cercle ne passant pas par* 0.

(4) un cercle passant par 0 *en une droite ne passant pas par* 0.

Démonstration. Décomposition en tous les cas possibles.

Droite passant par 0 La façon la plus simple de traiter la droite passant par 0 est de l'écrire sous forme paramétrique :

$$z(t) = te^{i\theta} \tag{21.164}$$

pour θ fixé et $t \in \mathbb{R} \cup \{\infty\}$. En appliquant l'inversion :

$$\iota\big(z(t)\big) = \frac{1}{(te^{i\theta})^*} = \frac{1}{t}e^{i\theta}. \tag{21.165}$$

Notons que les cas particuliers fonctionnent : pour $t = 0$ nous avons le point ∞ et pour $t = \infty$ nous avons 0.

Droite ne passant pas par 0 Une droite ne passant par par $z = 0$ est un ensemble de la forme

$$d(\omega, k) = \{z \in \mathbb{C}^* \text{ tel que } \bar{\omega}z + \omega\bar{z} = k\} \cup \{\infty\} \tag{21.166}$$

avec $k \neq 0$. Étant donné que ι est une bijection et même une involution nous avons $z \in \iota\big(d(\omega, k)\big)$ si et seulement si $\tau(z) \in d(\omega, k)$. L'équation est donc, pour $z \neq 0$:

$$\frac{\bar{\omega}}{z} + \frac{\omega}{\bar{z}} = k. \tag{21.167}$$

Et comme $z \neq 0$ nous pouvons multiplier par $z\bar{z}$ pour trouver $\bar{z}\bar{\omega} + z\omega = kz\bar{z}$. Donc

$$\iota\big(d(\omega, k)\big) = \{z \in \mathbb{C}^* \text{ tel que } \bar{z}\bar{\omega} + z\omega = kz\bar{z}\} \cup \{0\}. \tag{21.168}$$

Dans l'ensemble, nous pouvons renommer z et $\bar{z}$ pour avoir une forme plus symétrique. De plus il se fait que $z = 0$ vérifie l'équation donnée ; nous pouvons donc lever la condition $z \in \mathbb{C}^*$ et ne plus ajouter $\{0\}$ à coté :

$$\iota\big(d(\omega, k)\big) = \{z \in \mathbb{C} \text{ tel que } \bar{\omega}z + \omega\bar{z} = kz\bar{z}\}. \tag{21.169}$$

Cela est l'équation d'un cercle passant par l'origine (définition 21.60).

Cercle ne passant pas par 0 Nous considérons le cercle $C(\omega, r)$ avec $r^2 \neq |\omega|^2$. Il ne contient ni ∞ ni 0 et nous avons alors

$$\iota\big(C(\omega, r)\big) = \{z \in \mathbb{C}^* \text{ tel que } \frac{1}{z\bar{z}} - \bar{\omega}\frac{1}{\bar{z}} - \omega\frac{1}{z} = r^2 - |\omega|^2\}. \tag{21.170}$$

Vu que z n'est jamais nul nous pouvons multiplier l'équation par $z\bar{z}$:

$$\iota\big(C(\omega, r)\big) = \{z \in \mathbb{C} \text{ tel que } \big(|\omega|^2 - r^2\big)z\bar{z} - \bar{\omega}z - \omega\bar{z} = -1\}. \tag{21.171}$$

Le coefficient $|\omega|^2 - r^2$ est non nul par hypothèse et cet ensemble est un cercle par le lemme 21.62. Il ne passe manifestement pas par $z = 0$.

Cercle passant par 0 Le cercle passe par 0, et donc son image par ∞. Nous écrivons alors

$$\iota\big(C(\omega, |\omega|)\big) = \iota\big(C(\omega, |\omega|)\backslash\{0\}\big) \cup \{\infty\}. \tag{21.172}$$

Nous avons

$$C(\omega, |\omega|)\backslash\{0\} = \{z \in \mathbb{C}^* \text{ tel que } z\bar{z} - \bar{\omega}z - \omega\bar{z} = 0\}, \tag{21.173}$$

et un calcul usuel donne

$$\iota\big(C(\omega, |\omega|)\backslash\{0\}\big) = \{z \in \mathbb{C}^* \text{ tel que } 1 - \bar{\omega}z - \omega\bar{z} = 0\}, \tag{21.174}$$

et donc

$$\iota\big(C(\omega, |\omega|)\big) = d(\omega, 1), \tag{21.175}$$

en nous souvenant que le point ∞ est contenu dans $d(\omega, 1)$.

□

Proposition 21.77.
Une homographie conserve l'ensemble des droites et cercles de $\hat{\mathbb{C}}$.

Attention : cela ne veut pas dire qu'une homographie transforme une droite en une droite et un cercle en un cercle. Ça veut dire qu'une homographie transforme une droite en une droite ou un cercle et un cercle en une droite ou un cercle.

Démonstration. Nous savons par la proposition 20.11 et 21.74 que les homographies se décomposent en inversion, réflexion, translation et rotation-homothétie.

À part pour l'inversion, tout est clair comment ça fonctionne hein. En ce qui concerne l'inversion, nous avons la proposition 21.76 qui donne déjà toutes les réponses.

□

21.78.
Les homographies préservent les angles, c'est l'objet du théorème suivant. Il ne faudrait cependant pas croire que si A, B et C sont trois points distincts, l'angle entre $\overline{AC}$ et $\overline{BC}$ est le même que celui entre $\overline{f(A)f(C)}$ et $\overline{f(B)f(C)}$ dès que f est une homographie. Cela serait préserver les angles globalement, c'est à dire préserver les angles lorsque les points sont déplacés par f.

Nous allons regarder les angles locaux, c'est à dire lorsque les points sont déplacés par df.

Définition 21.79.
Nous disons qu'une application $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ préserve localement les angles non orientés lorsque

$$\cos\big(df_a(u), df_a(v)\big) = \cos(u, v) \tag{21.176}$$

pour tout $a \in \mathbb{R}^2$ et $u, v \in \mathbb{R}^2$. Ici il est mieux de penser à $u, v \in T_a\mathbb{R}^2$ pour qui sait les espaces tangents en géométrie différentielle.

Voir la définition de l'angle 16.47.

Théorème 21.80.
Les homographies de $P(\mathbb{C}^2)$ préservent localement les angles non orientés.

Démonstration. En ce qui concerne les translations, dilatations et rotations, les choses sont claires. Vérifions pour l'inversion, qu'il faut interpréter dans $\mathbb{R}^2$. Pour $z = x + iy$ nous avons

$$\iota(z) = \frac{1}{z} = \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2}. \tag{21.177}$$

Nous devons donc étudier la fonction

$$\begin{aligned} F\colon \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x, y) &\mapsto \Big(\frac{x}{r^2}, -\frac{y}{r^2}\Big) \end{aligned} \tag{21.178}$$

où nous avons posé $r^2 = x^2 + y^2$ pour simplifier les notations.

Soient deux vecteurs $u, v \in \mathbb{R}^2$ et un point $a \in \mathbb{R}^2$. Nous devons prouver que

$$\frac{u \cdot v}{\|u\|\|v\|} = \frac{df_a(u) \cdot df_a(v)}{\|df_a(u)\|\|df_a(v)\|}. \tag{21.179}$$

Pour cela, nous pourrions calculer dF_a et passer en coordonnées polaires[255] mais nous préférons faire les calculs à la dure parce que nous avons Sage avec nous.

Nous notons A la matrice de dF en $a = (x, y)$ et nous avons

$$Au \cdot Av = A^tAu \cdot v \tag{21.180}$$

ainsi que $\|u\| = \sqrt{u \cdot u} = \sqrt{A^tAu \cdot u}$, de telle sorte qu'il devienne urgent de calculer A^tA. Voici le calcul :

```
var('y')

# les fonctions coordonnées
F1(x,y)=x/(x**2+y**2)
F2(x,y)=-y/(x**2+y**2)

# La matrice différentielle :
A=matrix( [  [F1.diff(x).simplify_full(),F1.diff(y).simplify_full()],[F2.diff(x).simplify_full(),F2.diff(y).simplify_full()]  ] )

# Quelqu'un peut expliquer pourquoi ceci ne fonctionne pas ?
# A=matrix( [  [F1.diff(x),F1.diff(y)],[F2.diff(x),F2.diff(y)]  ] ).simplify_full()

# A^tA
S=A.transpose()*A
S=S.simplify_full()      # Mais ça, ça marche !!
print(S)
```

tex/sage/sageSnip009.sage

Le résultat est que

$$A^tA = \begin{pmatrix} \frac{1}{r^4} & 0 \\ 0 & \frac{1}{r^4} \end{pmatrix} = \frac{1}{r^4}\mathrm{Id}\,. \tag{21.181}$$

La vérification de (21.179) est alors immédiate. □

21.6.3 Birapport

Nous introduisons maintenant quelque chose qui s'appelle le «birapport» et qui n'est a priori pas du tout lié au birapport définit en 21.40.

Définition 21.81 (Birapport dans $\hat{\mathbb{C}}$[195])**.**
*Soient $a, b, c, x \in \hat{\mathbb{C}}$ où a, b et c sont distincts. Le **birapport** de ces quatre nombres est l'élément de $\hat{\mathbb{C}}$ donné par, si $a, b, c \neq \infty$:*

$$[a, b, c, x] = \frac{(a-c)(b-x)}{(b-c)(a-x)}, \tag{21.182}$$

et

$$[\infty, b, c, x] = \frac{b-x}{a-x} \tag{21.183a}$$

$$[a, \infty, c, x] = \frac{a-c}{a-x} \tag{21.183b}$$

$$[a, b, \infty, x] = \frac{b-x}{a-x} \tag{21.183c}$$

21.82.
Notons la «logique» des cas particuliers. Pour le premier, si $a \to \infty$ tandis que les autres restent dans $\mathbb{C}$ alors $a - c$ et $a - x$ deviennent du même ordre de grandeur et se simplifient. Il reste les deux autres parties de la fraction.

C'est cette même logique qui, partant de $[a, b, \infty, x] = \frac{b-x}{a-x}$ donne

$$[a, b, \infty, \infty] = 1 \tag{21.184}$$

comme il se doit si nous avons l'intention de ressembler au lemme 21.45.

L'objet «birapport» introduit ici est évidemment lié au birapport sur $P(\mathbb{C}^2)$ définit plus haut. Le lien est la proposition suivante.

Proposition 21.83.
Soit l'application $\varphi_0 \colon P(\mathbb{C}^2) \to \hat{\mathbb{C}}$ définie en 21.36. Si $A, B, C, X \in P(\mathbb{C}^2)$ alors

$$[A, B, C, X]_{\varphi_0} = [\varphi_0(A), \varphi_0(B), \varphi_0(C), \varphi_0(X)]. \tag{21.185}$$

Cela est une égalité dans $\hat{\mathbb{C}}$.

Démonstration. Nous écrivons $A = [a_1, a_2]$, $B = [b_1, b_2]$, $C = [c_1, c_2]$ avec $a_1, a_2, b_2, b_2, c_1, c_2 \in \mathbb{C}$. Par définition $[A, B, C, X]_{\varphi_0} = (\varphi_0 \circ \phi)(X)$ où $\phi \colon P(\mathbb{C}^2) \to P(\mathbb{C}^2)$ est l'unique homographie telle que

$$\phi[a_1, a_2] = [1, 0] \tag{21.186a}$$
$$\phi[b_1, b_2] = [0, 1] \tag{21.186b}$$
$$\phi[c_1, c_2] = [1, 1] \tag{21.186c}$$

Une des difficultés de cette preuve va être de calculer ce ϕ. D'abord nous pouvons introduire $\tilde{\phi} = \varphi_0 \circ \phi \circ \varphi^{-1}$ qui est obligatoirement (proposition 21.69) de la forme

$$\tilde{\phi}(z) = \frac{\alpha z + \beta}{\gamma a + \delta}. \tag{21.187}$$

Nous allons imposer les relations (21.186) pour déterminer les coefficients α, β, γ et δ.

D'abord

$$(\varphi_0^{-1} \circ \tilde{\phi} \circ \varphi_0)\big([a_1, a_2]\big) = (\varphi_0^{-1} \circ \tilde{\phi})(a_1/a_2) \tag{21.188a}$$
$$= \varphi_0^{-1}\left(\frac{\alpha\frac{a_1}{a_2} + \beta}{\gamma\frac{a_1}{a_2} + \delta}\right) \tag{21.188b}$$
$$= \big[\alpha\frac{a_2}{a_2} + \beta, \gamma\frac{a_1}{a_2} + \delta\big]. \tag{21.188c}$$

Égaler cela à $[1, 0]$ donne

$$\begin{cases} \alpha\dfrac{a_1}{a_2} + \beta \neq 0 & \text{(21.189a)} \\ \gamma\dfrac{a_1}{a_2} + \delta = 0. & \text{(21.189b)} \end{cases}$$

Donc nous avons déjà

$$\tilde{\phi}(z) = \frac{\alpha z + \beta}{\gamma(z - \frac{a_1}{a_2})} = \frac{\alpha z + \beta}{\gamma\big(z - \varphi_0(A)\big)}. \tag{21.190}$$

En y imposant la contrainte $(\varphi^{-1} \circ \tilde{\phi} \circ \varphi_0)([b_1, b_2]) = [0, 1]$ nous trouvons les contraintes

$$\begin{cases} \gamma\big(\varphi_0(B) - \varphi_0(A)\big) \neq 0 & \text{(21.191a)} \\ \beta = -\alpha\varphi_0(B). & \text{(21.191b)} \end{cases}$$

Nous avons décidé d'écrire $\varphi_0(A)$ au lieu de a_1/a_2 à la fois pour un soucis de simplification d'écriture et dans le but de ressembler à (21.185). En substituant :

$$\tilde{\phi}(z) = \frac{\alpha\big(z - \varphi_0(B)\big)}{\gamma\big(z - \varphi_0(A)\big)}. \tag{21.192}$$

La condition pour $[c_1, c_2]$ donne

$$\big[\alpha\big(\varphi_0(C) - \varphi_0(B)\big), \gamma\big(\varphi_0(C) - \varphi_0(A)\big)\big] = [1, 1], \tag{21.193}$$

ce qui donne

$$\alpha\big(\varphi_0(C)-\varphi_0(B)\big)=\gamma\big(\varphi_0(C)-\varphi_0(A)\big). \tag{21.194}$$

Nous avons alors

$$\gamma=\alpha\frac{\varphi_0(C)-\varphi_0(B)}{\varphi_0(C)-\varphi_0(A)}, \tag{21.195}$$

et les α se simplifient dans la formule pour $\tilde{\phi}$:

$$\tilde{\phi}(z)=\frac{\big(z-\varphi_0(B)\big)\big(\varphi_0(C)-\varphi_0(A)\big)}{\big(z-\varphi_0(A)\big)\big(\varphi_0(C)-\varphi_0(B)\big)}. \tag{21.196}$$

Par la proposition 21.69, l'application $\varphi_0^{-1}\circ\tilde{\phi}\circ\varphi_0$ est une homographie. Nous pouvons donc calmement calculer le birapport $[A,B,C,X]_{\varphi_0}$ de la façon suivante :

$$\begin{aligned}
[A,B,C,X]_{\varphi_0}&=(\varphi_0\circ\phi)(X) && (21.197a)\\
&=(\varphi_0\circ\varphi_0^{-1}\circ\tilde{\phi}\circ\varphi_0)(X) && (21.197b)\\
&=\tilde{\phi}\big(\varphi_0(X)\big) && (21.197c)\\
&=\frac{\big(\varphi_0(X)-\varphi_0(B)\big)\big(\varphi_0(C)-\varphi_0(A)\big)}{\big(\varphi_0(X)-\varphi_0(A)\big)\big(\varphi_0(C)-\varphi_0(B)\big)} && (21.197d)\\
&=[\varphi_0(A),\varphi_0(B),\varphi_0(C),\varphi_X(A)]. && (21.197e)
\end{aligned}$$

□

Proposition 21.84.
Les homographies de $\hat{\mathbb{C}}$ conservent le birapport.

Démonstration. Ici le mot «homographie» réfère à la définition 21.71 et le birapport à 21.81. Soient $a,b,c,x\in\hat{\mathbb{C}}$ et une homographie $\tilde{\phi}\colon\hat{\mathbb{C}}\to\hat{\mathbb{C}}$. Il existe une homographie $\phi\colon P(\mathbb{C}^2)\to P(\mathbb{C}^2)$ telle que $\tilde{\phi}=\varphi_0\circ\phi\varphi_0^{-1}$. Alors

$$\begin{aligned}
\big[\tilde{\phi}(a),\tilde{\phi}(b),\tilde{\phi}(c),\tilde{\phi}(x)\big]&=\big[(\phi\circ\varphi_0^{-1})(a),(\phi\circ\varphi_0^{-1})(b),(\phi\circ\varphi_0^{-1})(c),(\phi\circ\varphi_0^{-1})(x),\big]_{\varphi_0} && (21.198a)\\
&=\big[\varphi_0^{-1}(a),\varphi_0^{-1}(b),\varphi_0^{-1}(c),\varphi_0^{-1}(x),\big]_{\varphi_0} && (21.198b)\\
&=[a,b,c,x]. && (21.198c)
\end{aligned}$$

Justifications :

— Identification des birapports sur $P(\mathbb{C}^2)$ et sur $\hat{\mathbb{C}}$, proposition 21.83.
— Invariance du birapport sour les homographies (dans $P(\mathbb{C}^2)$), proposition 21.47.

□

Proposition 21.85.
Soient des points a,b,c,x dans $\hat{\mathbb{C}}$ avec a,b,c distincts. Ils sont alignés ou cocycliques si et seulement si $[a,b,c,x]\in\hat{\mathbb{R}}$.

Démonstration. Nous allons faire plusieurs cas. Mais dans tous les cas vous pouvez relire la définition des angles orientés 16.116 et la partie sur les angles dans les nombres complexes 16.5.2.

Tous les points sont distincts et dans $\mathbb{C}$ D'une part, nous savons que le nombre complexe $re^{i\theta}$ est réel si et seulement si $\theta\in[0]_\pi$, et d'autre part l'argument du birapport (21.182) est

$$[\overrightarrow{ca},\overrightarrow{cb}]+[\overrightarrow{xb},\overrightarrow{xa}]. \tag{21.199}$$

Le birapport est réel si et seulement si

$$[\overrightarrow{ca},\overrightarrow{cb}]-[\overrightarrow{xa},\overrightarrow{xb}]\in[0]_\pi. \tag{21.200}$$

À gauche nous avons une classe modulo 2π et à droite une classe modulo π. L'égalité signifie qu'il y a un représentant du membre de gauche qui appartient au membre de droite. Si vous

aimez faire très attention à ce que signifient les notations, voici trois manières d'écrire la condition, par ordre croissant de précision :

$$[\overrightarrow{ca}, \overrightarrow{cb}] - [\overrightarrow{xa}, \overrightarrow{xb}] = [0]_\pi, \tag{21.201a}$$
$$[\overrightarrow{ca}, \overrightarrow{cb}] - [\overrightarrow{xa}, \overrightarrow{xb}] \in [0]_\pi, \tag{21.201b}$$
$$[\overrightarrow{ca}, \overrightarrow{cb}] - [\overrightarrow{xa}, \overrightarrow{xb}] \subset [0]_\pi. \tag{21.201c}$$

Nous avons donc que le birapport est réelle si et seulement si la condition (21.201c) est vérifiée. D'après le théorème 16.124, cette dernière condition est équivalente à dire que les points a, b, c et x sont alignés.

Pas quatre points distincts, dans $\mathbb{C}$ Nous supposons encore que a, b, c et x sont dans $\mathbb{C}$. Mais nous supposons que x est un de a, b ou c. Vu que par hypothèse a, b et c sont distincts, c'est le seul cas à considérer dans la catégorie des 4 points non distincts.

Trois points sont toujours alignés ou cocycliques[9]. Donc nous devons seulement montrer que dans ce cas le birapport est toujours dans $\hat{\mathbb{R}}$. Par définition,

— Si $x = a$ alors $[a, b, c, x] = \infty$,
— Si $x = b$ alors $[a, b, c, x] = 0$,
— Si $a = c$ alors $[a, b, c, x] = 1$.

Dans tous les cas de figure le birapport est dans $\hat{\mathbb{R}}$.

À ce niveau de la preuve nous devons encore vérifier les cas où a, b, c ou x valent ∞. Si l'un de a, b ou c est ∞ et si x l'est aussi, alors, comme ∞ est aligné avec tout, nous avons seulement une droite passant par deux points. Il nous faut donc seulement regarder les cas où un seul des 4 points est ∞.

Si $a = \infty$ Le birapport est alors

$$[\infty, b, c, x] = \frac{b - x}{a - x}, \tag{21.202}$$

qui est un nombre a priori complexe dont le dénominateur est supposé non nul parce que le cas $a = x$ est déjà traité. L'argument de ce nombre est dans la classe de l'angle orienté

$$\arg\left(\frac{b - x}{a - x}\right) \in [\overrightarrow{xa}, \overrightarrow{xb}]. \tag{21.203}$$

Le birapport est réel si et seulement si le membre de gauche est dans $[0]_\pi$. Et cela est justement le cas où le membre de droite donne des points alignés.

Les cas $b = \infty$, $c = \infty$ et $x = \infty$ se traitent de la même manière. □

21.6.4 Division harmonique

Définition 21.86.
*Nous disons que les éléments a, b, c et x de $\hat{\mathbb{C}}$ sont en **division harmonique** lorsque* $[a, b, c, x] = -1$.

21.87.
Une chose qui sera utile par la suite est de remarquer que $[a, b, c, \infty] = -1$ lorsque $c = \frac{a+b}{2}$.

Nous allons maintenant voir comment, pour $a, b, c \in \mathbb{C}$ donnés nous pouvons construire x tels que a, b, c, x soient en division harmonique. Vu que trois points sont soit cocycliques soit alignés nous divisons la construction en deux parties.

Notons que si nous trouvons une construction qui donne une point x vérifiant $[a, b, c, x] = -1$ alors nous prouvons au passage que la construction ne dépend pas des choix intermédiaires parce que il n'existe qu'un unique x tel que $[a, b, c, x] = -1$ lorsque a, b, c sont donnés.

9. Si ils ne sont pas alignés, prendre la médiatrice du segment $[a, b]$ et celle de $[b, c]$, et l'intersection vous donnera le centre d'un cercle passant par a, b et c.

Lemme 21.88 ([259, 260])**.**
Soient a, b, c cocycliques dans $\mathbb{C}$. Nous nommons $\mathcal{C}$ le cercle contenant a, b et c ainsi que T_a et T_b les tangentes à $\mathcal{C}$ en a et b. Soit $m = T_a \cap T_b$ et la droite $L = (mc)$. Alors le point

$$x = (mc) \cap \mathcal{C} \tag{21.204}$$

vérifie $[a, b, c, x] = -1$.

Si $m = \infty$ (arrive lorsque $T_a \parallel T_b$) alors en guise de L nous prenons la parallèle à T_a passant par c.

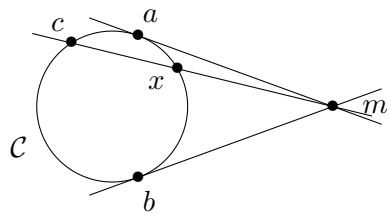
Démonstration. Nous séparons les cas suivant que $m = \infty$ ou non.

$\underline{m = \infty}$ Les tangentes à $\mathcal{C}$ en a et en b sont parallèles, c'est à dire que ces points sont diamétralement opposés sur $\mathcal{C}$. Les homographies préservent le birapport (proposition 21.84), et les rotations, dilatations et translations sont des homographies (voir 21.74).

Nous pouvons donc nous ramener au cas où $\mathcal{C}$ est centré en 0 et de rayon 1 avec $a = i$ et $b = -i$. Dans ce cas, $c = e^{i\theta}$. Vu que x est donné par l'intersection entre le cercle et la droite horizontale passant par c nous avons $x = e^{i(\pi-\theta)}$. Le birapport se calcule explicitement :

$$[a, b, c, x] = \frac{(i - e^{i\theta})(-i - e^{i(\pi-\theta)})}{(-i - e^{i\theta})(i - e^{i(\pi-\theta)})} = -1. \tag{21.205}$$

$\underline{m \neq \infty}$ Nous sommes dans la situation suivante où à une translation près nous supposons $x = 0$:



Nous allons prouver que dans ce cas, $[a, b, c, x] = -1$. Pour cela nous considérons l'inversion de centre x (qui est $x = 0$ par translation). Soit $\tilde{\phi}$ cette homographie. Elle conserve le birapport, il nous allons voir que calculer $[\tilde{\phi}(a), \tilde{\phi}(b), \tilde{\phi}(c), \tilde{\phi}(x)]$ se révèle être plus facile [10].

Nous nommons $A = (am)$, $B = (bm)$, $C = (cm)$ et $\mathcal{C}$, le cercle. Nous allons maintenant faire intensément usage de la proposition 21.76. Nous avons :

— $\tilde{\phi}(A)$ est un cercle passant par 0.
— $\tilde{\phi}(B)$ est un cercle passant par 0.
— $\tilde{\phi}(C)$ est la droite C.
— $\tilde{\phi}(\mathcal{C})$ est une droite ne passant pas par 0.

Les droites A et B se coupent en m et en ∞ (qui sont des points distincts). Donc les cercles $\tilde{\phi}(A)$ et $\tilde{\phi}(B)$ se coupent en 0 et $\tilde{\phi}(m)$, aucun de ces deux points n'est sur la droite $\tilde{\phi}(\mathcal{C})$.

Par tangence, la droite A et le cercle $\mathcal{C}$ se coupent en un seul point (a). Donc $\tilde{\phi}(A)$ coupe $\tilde{\phi}(\mathcal{C})$ en un seul point, $\tilde{\phi}(a)$. Idem pour le cercle $\tilde{\phi}(B)$.

Nous avons donc que les cercles $\tilde{\phi}(A)$ et $\tilde{\phi}(B)$ sont tangents à la droite $\tilde{\phi}(\mathcal{C})$ et se coupent en exactement deux points distincts (qui sont donc du même côté de la droite).

Nous nous intéressons à la droite $\tilde{\phi}(C)$. C'est une droite parce que c'est l'image d'une droite passant par 0. Elle passe par 0, par $\tilde{\phi}(c)$ et $\tilde{\phi}(m)$. Le fait qu'elle passe par 0 et $\tilde{\phi}(m)$ fait que c'est la droite passant par les deux intersections des cercles. Vu que $c \in \mathcal{C} \cap C$, le point d'intersection $\tilde{\phi}(C) \cap \tilde{\phi}(\mathcal{C})$ est $\tilde{\phi}(c)$.

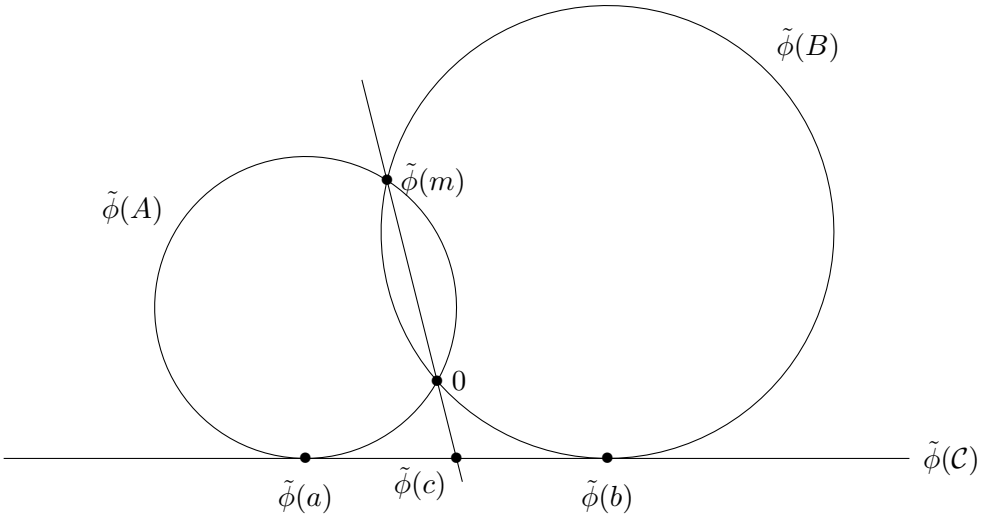
Quelle est la puissance du point $\tilde{\phi}(c)$ par rapport au cercle $\tilde{\phi}(A)$? En la calculant avec la droite $\tilde{\phi}(C)$, qui intersecte les deux cercles aux points déjà étudiés, la puissance est :

$$k = d\big(\tilde{\phi}(c), 0\big) \times d\big(\tilde{\phi}(c), \tilde{\phi}(m)\big). \tag{21.206}$$

10. Si vous n'avez peur d'aucun calculs, il suffit de poser $a = e^{i\theta}$, $b = e^{-i\theta}$ et $c = e^{i\sigma}$ et vous êtes théoriquement capable de calculer les coordonnées de tous les points, y compris de x en termes de θ et σ. Ensuite le calcul du birapport est explicite.

Vu que ces points sont également sur le cercle $\tilde{\phi}(B)$, la puissance de $\tilde{\phi}(c)$ par rapport à ce second point est la même.

Tout cela justifie le dessin suivant [11] :



Mais la droite passant par $\tilde{\phi}(a)$ et $\tilde{\phi}(c)$ (qui est tangente au cercle) permet également de calculer cette puissance :

$$k = d\big(\tilde{\phi}(a), \tilde{\phi}(c)\big). \tag{21.207}$$

Idem pour la puissance par rapport à l'autre cercle :

$$k = d\big(\tilde{\phi}(a), \tilde{\phi}(b)\big). \tag{21.208}$$

Nous en déduisons que $\tilde{\phi}(c)$ est le milieu entre $\tilde{\phi}(a)$ et $\tilde{\phi}(b)$.

Du coup

$$\big[\tilde{\phi}(a), \tilde{\phi}(b), \tilde{\phi}(c), \tilde{\phi}(x)\big] = \big[\tilde{\phi}(a), \tilde{\phi}(b), \tilde{\phi}(c), \infty\big] = -1 \tag{21.209}$$

en vertu de ce que nous avons raconté en 21.87.

□

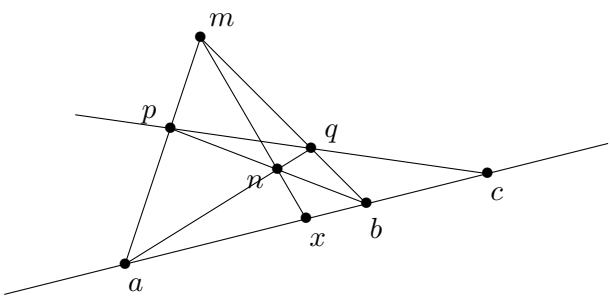
Lemme 21.89 ([261])**.**
Soit $a, b, c \in \mathbb{C}$ colinéaires. Soit m un point hors de cette droite. Nous considérons une droite issue de c coupant $[ma]$ en p et $[mb]$ en q.

Nous construisons $n = (aq) \cap (pb)$ et finalement $x = (mn) \cap (ab)$.

À la fin nous avons

$$[a, b, c, x] = -1. \tag{21.210}$$

Démonstration. Commençons par un dessin de la situation :



11. Bien que ce ne soit pas strictement nécessaire à la preuve, est-ce que vous savez si les deux cercles ont le même rayon ? Et si par hasard la droite $\big(\tilde{\phi}(m)\tilde{\phi}(c)\big)$ n'arrive pas perpendiculairement à $\tilde{\phi}(\mathcal{C})$?

Les points a, b et m ne sont pas alignés, et nous pouvons les utiliser comme repère barycentrique (voir 11.42 pour savoir en deux mots ce que c'est). Nous nommons (α, β, γ) les coordonnées de n dans ce système, c'est à dire que

$$\alpha\overrightarrow{na} + \beta\overrightarrow{nb} + \gamma\overrightarrow{nm} = 0. \tag{21.211}$$

Dans notre contexte, nous pouvons voir le vecteur $\overrightarrow{st}$ comme une façon d'écrire le nombre $t - s$. Par la proposition 11.47 nous savons les coordonnées barycentriques de p, x et q en regardant le triangle acb. Voici les coordonnées et les relations qu'elles signifient :

$$n = (\alpha, \beta, \gamma), \qquad \alpha\overrightarrow{na} + \beta\overrightarrow{nb} + \gamma\overrightarrow{nm} = 0 \tag{21.212a}$$
$$p = (\alpha, 0, \gamma), \qquad \alpha\overrightarrow{pa} + \gamma\overrightarrow{pm} = 0 \tag{21.212b}$$
$$q = (0, \beta, \gamma), \qquad \beta\overrightarrow{qb} + \gamma\overrightarrow{qm} = 0 \tag{21.212c}$$
$$x = (\alpha, \beta, 0). \qquad \alpha\overrightarrow{xa} + \beta\overrightarrow{xb} = 0. \tag{21.212d}$$

Nous voudrions maintenant voir les coordonnées de c. Nous posons $c = (\lambda, \mu, \sigma)$:

$$\lambda\overrightarrow{ca} + \mu\overrightarrow{cb} + \sigma\overrightarrow{cm} = 0. \tag{21.213}$$

Mais a, b et c sont alignés, donc $\overrightarrow{ca}$ et $\overrightarrow{cb}$ sont colinéaires, alors que $\overrightarrow{cm}$ n'est pas aligné avec les deux autres. L'annulation (21.213) demande donc l'annulation séparément

$$\begin{cases} \sigma\overrightarrow{cm} = 0 & \text{(21.214a)} \\ \lambda\overrightarrow{ca} + \mu\overrightarrow{cb} = 0. & \text{(21.214b)} \end{cases}$$

Nous en déduisons que $\sigma = 0$ et aussi que λ et μ ne sont pas nuls. Nous posons arbitrairement $\lambda = 1$ parce que les coordonnées barycentriques sont définies à coefficient multiplicatif près.

Nous imposons à présent le fait que p, q et c sont alignés. Pour cela nous devons faire apparaître les vecteurs $\overrightarrow{pq}$, $\overrightarrow{pc}$, $\overrightarrow{pc}$. Vu le dessin et les relations disponibles (21.212) le mieux est d'utiliser les relations de Chasles (proposition 11.3) pour faire $\overrightarrow{ca} = \overrightarrow{cp} + \overrightarrow{pq}$ et $\overrightarrow{cb} = \overrightarrow{cq} + \overrightarrow{qb}$. La relation (21.214b) devient :

$$\overrightarrow{cp} + \overrightarrow{pa} + \mu(\overrightarrow{cq} + \overrightarrow{qb}) = 0. \tag{21.215}$$

Les vecteurs $\overrightarrow{cp}$ et $\overrightarrow{cq}$ sont alignés, donc nous les écrivons ensemble. Les vecteurs $\overrightarrow{pa}$ et $\overrightarrow{qb}$ se transforment en utilisant les relations (21.212) :

$$\overrightarrow{cp} + \mu\overrightarrow{cq} - \frac{\gamma}{\alpha}\overrightarrow{pm} - \mu\frac{\gamma}{\beta}\overrightarrow{qm}. \tag{21.216}$$

Enfin nous voulons faire la somme du terme $\overrightarrow{pm}$ avec le terme $\overrightarrow{qm}$. D'abord on change le signe :

$$\overrightarrow{cp} + \mu\overrightarrow{cq} - \frac{\gamma}{\alpha}\overrightarrow{pm} + \frac{\mu\gamma}{\beta}\overrightarrow{mq} \tag{21.217}$$

ensuite nous écrivons

$$\frac{\mu\gamma}{\beta} = \frac{\mu\gamma}{\beta} + \frac{\gamma}{\alpha} - \frac{\gamma}{\alpha}, \tag{21.218}$$

et

$$\overrightarrow{cp} + \mu\overrightarrow{cq} - \frac{\gamma}{\alpha}(\overrightarrow{pm} + \overrightarrow{mq}) + \left(\frac{\mu\gamma}{\beta} + \frac{\gamma}{\alpha}\right)\overrightarrow{mq} = 0. \tag{21.219}$$

Tout cela pour

$$\overrightarrow{cp} + \mu\overrightarrow{cq} - \frac{\gamma}{\alpha}\overrightarrow{pq} + \left(\frac{\mu\gamma}{\beta} + \frac{\gamma}{\alpha}\right)\overrightarrow{mq} = 0. \tag{21.220}$$

Le dernier terme n'est pas colinéaire aux deux premiers et s'annule donc séparément :

$$\frac{\mu\gamma}{\beta} + \frac{\gamma}{\alpha} = 0. \tag{21.221}$$

Cela donne $\mu = -\beta/\alpha$.

Au finale nous avons

$$\overrightarrow{ca} - \frac{\beta}{\alpha}\overrightarrow{cb} = 0 \tag{21.222}$$

et donc

$$\alpha\overrightarrow{ca} - \beta\overrightarrow{cb} = 0, \tag{21.223}$$

ce qui donne les coordonnées $(\alpha, -\beta, 0)$ pour le point c.

Vu que nous somme dans l'espace vectoriel $\mathbb{C}$, ce que nous notons $\overrightarrow{AB}$ n'est rien d'autre que la différence $B - A$ dans $\mathbb{C}$ [12]. La relation (21.223) signifie donc

$$\alpha(a - c) = \beta(b - c). \tag{21.224}$$

Nous avons alors :

$$\frac{a - c}{b - c} = \frac{\beta}{\alpha}. \tag{21.225}$$

Par ailleurs, la relation (21.212d) à propos des coordonnées de x donne

$$\frac{a - x}{b - x} = -\frac{\beta}{\alpha}. \tag{21.226}$$

En égalisant les deux valeurs de β/α nous trouvons :

$$\frac{a - c}{b - c} = \frac{x - a}{b - x}, \tag{21.227}$$

ce qui donne (via un petit jeu de signes)

$$\frac{(c - a)(x - b)}{(c - b)(x - a)} = -1. \tag{21.228}$$

C'est cela que nous voulions. □

21.6.5 Groupe circulaire

Nous avons vu que les homographies présent l'ensemble des cercles et droites. Nous pouvons nous demander quel est le groupe maximum préservant l'ensemble des cercles et droites.

Définition 21.90.
*Le **groupe circulaire** de $\mathbb{C}$ est le groupe de transformations de $\hat{\mathbb{C}}$ engendré par les homographies* [13] *et la conjugaison complexe. Le groupe circulaire de l'espace projectif est l'ensemble des applications de la forme $\varphi_0^{-1} \circ f \circ \varphi_0$ où f est un élément du groupe circulaire de eC.*

Plusieurs remarques à propos de cette définition.

(1) Vu le lemme 21.75, la conjugaison complexe n'est pas une homographie. Donc cette définition n'est pas stupide : le groupe circulaire est strictement plus grand que le groupe des homographies.

(2) Vous vous souvenez de la définition d'un sous-groupe engendré ? C'est la définition 3.8.

Lemme 21.91.
Soit une application $\alpha\colon \mathbb{C} \to \mathbb{C}$ fixant 1 et 0 et préservant les divisions harmoniques (c'est à dire tel que son prolongement à $\hat{\mathbb{C}}$ donné par $\alpha(\infty) = \infty$ préserve les divisions harmoniques). Alors α est un automorphisme de corps [14].

12. Les mauvaise langues diront que tout le chapitre sur les espaces affines, et surtout la partie sur les barycentres ne sont rien d'autres que le snobisme d'écrire $\overrightarrow{xy}$ au lieu de $y - x$. C'est aussi une facilité d'écriture.

13. Homographie de $\hat{\mathbb{C}}$: définition 21.71.

14. Définition 2.23.

Démonstration. Nous savons que si $a, b, c \in \mathbb{C}$ nous avons $c = (a + b)/2$ si et seulement si $[a, b, c, \infty] = -1$. Vu que α préserve les divisions harmoniques nous avons équivalence entre les affirmations suivantes :

$$c = \frac{a+b}{2} \tag{21.229a}$$
$$[a, b, c, \omega] = -1 \tag{21.229b}$$
$$[\alpha(a), \alpha(b), \alpha(c), \alpha(\infty)] = -1 \tag{21.229c}$$
$$\frac{\alpha(a) + \alpha(b)}{2} = \alpha(c). \tag{21.229d}$$

Donc α préserve les milieux : pour tout $a, b \in \mathbb{C}$ nous avons

$$\alpha\left(\frac{a+b}{2}\right) = \frac{\alpha(a) + \alpha(b)}{2}. \tag{21.230}$$

En particulier, cette relation avec $b = 0$ donne (parce que $\alpha(0) = 0$) : $\alpha(a/2) = \alpha(a)/2$. Nous avons au final, en utilisant cela en conjonction avec (21.230) :

$$\frac{\alpha(a) + \alpha(b)}{2} = \alpha\left(\frac{a+b}{2}\right) = \frac{\alpha(a+b)}{2}. \tag{21.231}$$

Cela démontre déjà que

$$\alpha(a + b) = \alpha(a) + \alpha(b). \tag{21.232}$$

En particulier $\alpha(-a) = \alpha(0 - a) = \alpha(0) - \alpha(a) = -\alpha(a)$.

Nous passons maintenant à la démonstration du fait que $\alpha(ab) = \alpha(a)\alpha(b)$. Pour tout a différent de 0 et ± 1 nous avons

$$[a, -a, a^2, 1] = \frac{(a - a^2)(-a - 1)}{(-a - a^2)(a - 1)} = -1. \tag{21.233}$$

Et en prenant $\alpha(a)$ en guise de a nous avons aussi

$$[\alpha(a), -\alpha(a), \alpha(a)^2, \alpha(1)] = -1. \tag{21.234}$$

Vu que α préserve les divisions harmoniques, l'équation (21.233) donne aussi

$$\big[\alpha(a), \alpha(-a), \alpha(a^2), \alpha(1)\big] = -1, \tag{21.235}$$

c'est à dire

$$\big[\alpha(a), -\alpha(a), \alpha(a^2), 1\big] = -1. \tag{21.236}$$

Comparant (21.234) avec (21.236) et en tenant compte de l'unicité du birapport [15] nous avons

$$\alpha(a^2) = \alpha(a)^2. \tag{21.237}$$

Avec cela nous pouvons y aller en remarquant que

$$ab = \left(\frac{a+b}{2}\right)^2 - \left(\frac{a-b}{2}\right)^2. \tag{21.238}$$

Nous appliquons α à cette dernière équations en tenant compte de ce que nous savons déjà

$$\alpha(ab) = \left(\frac{\alpha(a) + \alpha(b)}{2}\right)^2 - \left(\frac{\alpha(a) - \alpha(b)}{2}\right)^2 = \alpha(a)\alpha(b). \tag{21.239}$$

□

15. C'est à dire que si trois éléments du birapport sont donnés, le quatrième est fixé. C'est une variation sur la thème de la proposition 21.46(2).

Théorème 21.92 ([261, 262]).
Le groupe circulaire de $\mathbb{C}$ est le groupe des bijections $\hat{\mathbb{C}} \to \hat{\mathbb{C}}$ préservant l'ensemble des cercles-droites.

Démonstration. L'inclusion dans un sens est facile : les homographies conservent l'ensemble des cercles et droites par la proposition 21.77. Et la conjugaison complexe aussi.

Soit une bijection $f\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ préservant les cercles-droites. Nous supposons dans un premier temps que $f(0) = 0$, $f(1) = 1$ et $f(\infty) = \infty$.

Pour f vérifiant $f(0, 1, \infty) = 0, 1, \infty$ Si $\mathcal{C}$ est un cercle alors $f(\mathcal{C})$ est un cercle ou une droite, mais vu que $\mathcal{C}$ ne contient pas ∞, l'ensemble $f(\mathcal{C})$ ne le contient pas non plus. Donc f transforme un cercle en un cercle et une droite en une droite.

f préserve les divisions harmoniques Soient a, b, c, x dans $\hat{\mathbb{C}}$ tels que $[a, b, c, x] = -1$. Nous allons prouver que $[f(a), f(b), f(c), f(x)] = -1$.

Si a, b et c sont colinéaires, nous suivons la construction du lemme 21.89. Soit m hors de la droite (ab) et une droite D passant par c et coupant $[ma]$ en p et $[mb]$ en q. Nous posons $n = (pb) \cap (qa)$. Alors $x = (mn) \cap (ac)$.

Vues les propriétés de f (en particulier c'est une bijection qui respecte les intersectons, tangences, cercles et droites). Le point $f(m)$ est hors de la droite $\big(f(a)f(b)\big)$. La droite $f(D)$ passe par $f(c)$ et coupe les segments $[f(m)f(a)]$ en $f(p)$ et $[f(m)f(b)]$ en $f(q)$. Alors nous avons

$$f(n) = \big(p(p)f(b)\big) \cap \big(f(q)f(a)\big) \tag{21.240}$$

et aussi

$$f(x) = \big(p(m)f(n)\big) \cap \big(f(a)f(c)\big) \tag{21.241}$$

Donc $f(x)$ se construit à partir de $f(a)$, $f(b)$ et $f(c)$ en suivant la même construction que x à partir de a, b et c. Nous en concluons que $[f(a), f(b), f(c), f(x)] = -1$.

Si a, b et c sont cocycliques, le même raisonnement, en suivant le lemme 21.88 nous donne le même résultat.

f est un automorphisme du corps $\mathbb{C}$ C'est le lemme 21.91.

Et enfin ... Notre application f est un automorphisme du corps $\mathbb{C}$ qui fixe $\mathbb{R}$ parce qu'elle laisse invariante les droites dans $\mathbb{C}$. Donc la proposition 5.6 nous dit que f est soit l'identité soit la conjugaison complexe. Dans les deux cas, f est dans le groupe circulaire.

Pour f plus générale Nous ne supposons plus que f fixe 0, 1 et ∞. En tout cas les nombres $f^{-1}(1)$, $f^{-1}(0)$ et $f^{-1}(\infty)$ sont distincts parce que f est une bijection. Nous pouvons considérer une homographie[16] $\phi\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ telle que $\phi(1) = f^{-1}(1)$, $\phi(0) = f^{-1}(0)$ et $\phi(\infty) = f^{-1}(\infty)$. Dans ce cas l'application

$$g = f \circ \phi \tag{21.242}$$

vérifie $g(1) = 1$, $g(0) = 1$ et $g(\infty) = \infty$ tout en continuant à transformer un cercle-droite en un cercle-droite. Donc $f \circ \phi$ est soit l'identité soit la conjugaison complexe. Avec ça, l'application

$$f = g \circ \phi^{-1} \tag{21.243}$$

est la composée d'une homographie avec soit l'identité soit la conjugaison complexe. Elle est donc dans le groupe circulaire.

□

16. Attention : ici nous parlons d'homographies de $\hat{\mathbb{C}}$, pas de $P(\mathbb{C}^2)$. L'existence d'une telle application demande de composer le corollaire 21.35 avec l'application φ_0 et la définition 21.71 et 21.72.

21.6.6 Action du groupe modulaire

Le **demi-plan de Poincaré** est l'ensemble

$$P = \{z \in \mathbb{C} \text{ tel que } \Im(z) > 0\}. \tag{21.244}$$

Le **groupe modulaire** est le quotient de groupes

$$\mathrm{PSL}(2,\mathbb{Z}) = \frac{\mathrm{SL}(2,\mathbb{Z})}{\mathbb{Z}_2}. \tag{21.245}$$

Ce sont donc les matrices au signe près de la forme

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{21.246}$$

où a, b, c et d sont entiers tels que $ad - cb = 1$.

Théorème 21.93 ([263])**.**
Le groupe modulaire agit fidèlement (définition 3.75) sur le demi-plan de Poincaré par

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} * z = \frac{az + b}{cz + d}. \tag{21.247}$$

L'ensemble $D = D_1 \cup D_2$ avec

$$D_1 = \{z \in P \text{ tel que } |z| > 1, -\frac{1}{2} \leqslant \Re(z) < \frac{1}{2}\} \tag{21.248a}$$

$$D_2 = \{z \in P \text{ tel que } |z| = 1, -\frac{1}{2} \leqslant \Re(z) \leqslant 0\} \tag{21.248b}$$

est un domaine fondamental (définition 3.82) de cette action.

De plus si nous notons

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \tag{21.249}$$

alors pour tout $z \in P$, il existe $A \in \mathrm{gr}(S,T)$ telle que $A * z \in D$.

Démonstration. Nous divisions la preuve en plusieurs étapes.

Bien définie D'abord il faut remarquer que l'action (21.247) est bien définie par rapport au quotient : $A * z = (-A) * z$. La vérification est immédiate.

Interne Montrons que si $A \in \mathrm{PSL}(2,\mathbb{Z})$ et $z \in P$ alors $A * z \in P$. Nous avons

$$A * z = \frac{az + b}{cz + d} = \frac{(az + b)(c\bar{z} + d)}{|cz + d|^2} = \frac{a|z|c + azd + bc\bar{z} + bd}{|cz + d|^2}, \tag{21.250}$$

et donc en décomposant $z = \Re(z) + i\Im(z)$,

$$\Im(A * z) = \Im\left(\frac{azd + bc\bar{z}}{|cz + d|^2}\right) = \frac{ad - bc}{|cz + d|^2}\Im(z) = \frac{\Im(z)}{|cz + d|^2} \tag{21.251}$$

où nous avons tenu compte de $ad - bc = 1$. Donc l'action respecte la (stricte) positivité de la partie imaginaire.

Action Nous vérifions maintenant que la formule donne bien une action : $A*(B*z)=(AB)*z$. Cela est un bon calcul :

$$A*(B*z)=A*\left(\frac{a'z+b}{c'z+d'}\right) \tag{21.252a}$$
$$=\frac{a\left(\frac{a'z+b}{c'z+d'}\right)+b}{c\left(\frac{a'z+b}{c'z+d'}\right)+d} \tag{21.252b}$$
$$=\frac{a(a'z+b')+b(c'z+d')}{c(a'z+b')+d(c'z+d')} \tag{21.252c}$$
$$=\frac{(aa'+bc')z+(ab'+bd')}{(ca'+dc')z+(cb'+dd')} \tag{21.252d}$$
$$=\begin{pmatrix}aa'+bc' & ab'+bd'\\ a'c+dc' & cb'+dd'\end{pmatrix}*z \tag{21.252e}$$
$$=(AB)*z. \tag{21.252f}$$

Fidèle Soit $A\in \mathrm{PSL}(2,\mathbb{Z})$ tel que pour tout $z\in P$ nous ayons

$$\frac{az+b}{cz+d}=z. \tag{21.253}$$

Alors nous avons

$$cz^2+(d-a)z+b=0. \tag{21.254}$$

Cela est donc un polynôme en z qui s'annule sur un ouvert [17] (le demi-plan de Poincaré). Il doit donc être identiquement nul, donc $c=b=a-d=0$. Si vous n'y croyez pas, écrivez pour $z=\epsilon i$ (avec $\epsilon>0$) :

$$-c\epsilon^2+\epsilon(d-a)i+b=0 \tag{21.255}$$

pour tout ϵ. Le fait d'avoir $c\epsilon^2=b$ pour tout ϵ implique que $c=b=0$. Donc A est de la forme

$$A=\begin{pmatrix}a & 0\\ 0 & d\end{pmatrix}, \tag{21.256}$$

avec la contrainte supplémentaire que $ad=1$, les nombres a et b étant entiers. Nous avons donc soit $a=d=1$ soit $a=d=-1$. Étant donné le quotient par $\mathbb{Z}_2$, ces deux possibilités donnent le même élément de $\mathrm{PSL}(2,\mathbb{Z})$.

Les orbites intersectent D Soit $z\in P$. Nous devons trouver $A\in\mathrm{PSL}(2,\mathbb{Z})$ tel que $A*z\in D$. Nous savons déjà que

$$\Im(A*z)=\frac{\Im(z)}{|cz+d|^2}. \tag{21.257}$$

Nous notons $\mathcal{O}_z$ l'orbite de z sous le groupe modulaire et nous posons

$$I_z=\{\Im(u)\text{ tel que }u\in\mathcal{O}_z\}=\{\Im(A*z)\text{ tel que }A\in\mathrm{PSL}(2,\mathbb{Z})\}, \tag{21.258}$$

l'ensemble des parties imaginaires des éléments de l'orbite de z. Nous allons montrer que cet ensemble est borné vers le haut en montrant que la quantité $|cz+d|$ ne peut,) z donné, prendre qu'un nombre fini de valeurs plus grandes que $\Im(z)$ [18]. Nous cherchons donc les couples $(c,d)\in\mathbb{Z}^2$ tels que $|cz+d|<1$.

17. On ne peut pas dire que $b=0$ simplement en justifiant qu'on l'obtient en posant $z=0$ parce que $z=0$ n'est pas dans le demi-plan de Poincaré.

18. Bien que cela ne soit pas indispensable pour la preuve, remarquons que I_z ne comprend qu'une quantité au plus dénombrable de valeurs. Le fait que, à z donné, la quantité $|cz+d|^2$ puisse être rendue aussi grande que l'on veut est évident. Donc I_z est borné vers le bas par zéro (qui n'est pas atteint, mais qui est une valeur d'adhérence).

Nous avons $\Im(cz+d)=c\Im(z)$, donc $|cz+d|\geqslant|c\Im(z)|$, mais il n'y a qu'un nombre fini de $c\in\mathbb{Z}$ tels que $|c\Im(z)|<1$. De la même façon, pour la partie réelle nous avons

$$\Re(cz+d)=c\Re(z)+d, \tag{21.259}$$

et pour chaque c, il n'y a qu'un nombre fini de $d\in\mathbb{Z}$ qui laissent cette quantité plus petite que 1 (en valeur absolue).

Donc I_z possède un maximum. Soit $A_1\in\mathrm{PSL}(2,\mathbb{Z})$ tel que $\Im(A_1*z)=\max I_z$. Nous notons $z_1=A_1*z$, et que nous n'avons a priori pas l'unicité. Nous allons maintenant agir sur z_1 avec l'élément

$$T=\begin{pmatrix}1&1\\0&1\end{pmatrix} \tag{21.260}$$

pour ramener z_1 dans le domaine D. Si $u\in P$ nous avons $T*u=u+1$ et donc

$$T^n*u=u+n. \tag{21.261}$$

Vu que D est de largeur 1, il existe un n (éventuellement négatif) tel que

$$\Re(T^n*z_1)\in[-\frac{1}{2},\frac{1}{2}[. \tag{21.262}$$

Notons qu'ici le fait d'être ouvert d'un côté et fermé de l'autre joue de façon essentielle (pour l'unicité aussi). Nous notons $z_2=T^n*z_1$.

Supposons un instant que $|z_2|<1$. Nous considérons l'élément

$$S=\begin{pmatrix}0&-1\\1&0\end{pmatrix} \tag{21.263}$$

qui fait

$$\Im(S*z)=\frac{\Im z}{|z|^2}. \tag{21.264}$$

Donc si $|z_2|<1$ alors $\Im(S*z_2)>\Im(z_2)$, ce qui contredit la maximalité de $\Im(z_2)$ dans I_z. Nous en déduisons que $|z_2|\geqslant 1$. Nous en déduisons que $|z_2|\geqslant 1$.

Si $|z_2|>1$, alors $z_2\in D_1$ et c'est bon. Si $|z_2|=1$, alors il faut encore un peu travailler. Si $z_2\pm 1$ est à l'intérieur du disque, alors en agissant avec T ou T^{-1} nous retrouvons la même contradiction que précédemment. En écrivant $z_2=e^{i\theta}$, nous devons donc avoir $2\cos(\theta)\leqslant 1$ ou encore $|\Re(z_2)|\leqslant\frac{1}{2}$. Donc si $\Re(z_2)\leqslant 0$ alors $z_2\in D_2$.

Le dernier cas à traiter est $\Re(z_2)\in\,]0,\frac{1}{2}]$, c'est à dire $\theta\in[\frac{\pi}{3},\frac{\pi}{2}[$. Dans ce cas l'action avec S ramène l'angle dans la bonne zone parce que $S*z=-\frac{1}{z}$ et donc $S*(\rho e^{-i\theta})=-\frac{1}{\rho}e^{-i\theta}$.

Unicité Nous voulons à présent montrer que si $z\in D$, alors $A*z$ n'est plus dans D (sauf si $A=\pm\mathbb{1}$). Nous supposons que $z\in D$ et $A\in\mathrm{PSL}(2,\mathbb{Z})$ soient tels que $A*z\in D$, et nous prouvons qu'alors soit nous arrivons à une contradiction soit nous arrivons à $A=\mathbb{1}$. Pour cela nous allons décomposer en de nombreux cas.

(1) Nous commençons par $\Im(A*z)\geqslant\Im(z)$. Dans ce cas nous avons $|cz+d|\leqslant 1$ et en particulier $|c||\Im(z)|\leqslant 1$. Étant donné que le point de D qui a la partie imaginaire la plus petite est $-\frac{1}{2}+\frac{2}{\sqrt{3}}i$, nous trouvons $|c|\leqslant 2/\sqrt{3}$. Vu que c doit être entier, nous avons trois cas : $c=-1,0,1$.

(a) Soit $c=0$. Alors $A=\begin{pmatrix}a&b\\0&d\end{pmatrix}$ et la condition de déterminant est $ad=1$, ce qui signifie $a=d=1$ (la possibilité $a=b=-1$ est «éliminée» le quotient par $\mathbb{Z}_2$ définissant $\mathrm{PSL}(2,\mathbb{Z})$). La matrice A doit alors être de la forme

$$A=\begin{pmatrix}1&b\\0&1\end{pmatrix} \tag{21.265}$$

et $A*z=z+b$. Si $z\in D$, alors le seul $z+b$ à être (peut-être) encore dans D est $b=0$, mais alors A est l'identité.

(b) Soit $c = 1$. Alors la condition $|cz+d| \leqslant 1$ nous donne trois possibilités [19] : $d = -1, 0, 1$.

i. Si $d = -1$, alors nous devons avoir $|z - 1| \leqslant 1$. Il est instructif de faire un dessin, mais le point d'intersection entre les cercles $|z| = 1$ et $|z - 1| = 1$ est le point $\frac{1}{2} + \frac{\sqrt{3}}{2}i$, qui n'est pas dans D. Bref, il n'y a pas de points dans D vérifiant $|z - 1| \leqslant 1$.

ii. Si $d = 1$, alors (et c'est maintenant que la dissymétrie de D intervient) nous avons le point

$$z = -\frac{1}{2} + \frac{\sqrt{3}}{2}i \tag{21.266}$$

qui est dans D et qui vérifie $|z + 1| \leqslant 1$. Voyons à quoi ressemble la matrice A dans ce cas. Son déterminant est $a - b = 1$. Nous écrivons donc

$$A = \begin{pmatrix} b+1 & b \\ 1 & 1 \end{pmatrix}, \tag{21.267}$$

et en tenant compte du fait que $z\bar{z} = |z + 1| = 1$, nous calculons

$$A * z = \frac{(b+1)z + b}{z+1} \tag{21.268a}$$

$$= \frac{(bz + z + b)(\bar{z} + 1)}{|z+1|^2} \tag{21.268b}$$

$$= z + b + 1. \tag{21.268c}$$

La seule façon de ne pas quitter D est d'avoir $b = -1$, mais alors nous avons

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \tag{21.269}$$

et $A * z = z$. Donc au final z est quand même le seul de son orbite à être dans D. Notons au passage cette très intéressante propriété du point

$$z_0 = -\frac{1}{2} + \frac{\sqrt{3}}{2}i. \tag{21.270}$$

C'est un point de qui vérifie $z_0 = A * z_0$ pour un élément non trivial A de $\mathrm{PSL}(2, \mathbb{Z})$. L'existence d'un tel élément est ce qui va nous coûter un peu de sueur pour prouver que $PSL(2, \mathbb{Z})$ est engendré par S et T.

iii. Le cas $d = 0$ nous fait écrire $1 = \det A = -b$, donc $b = -1$ et

$$A = \begin{pmatrix} a & -1 \\ 1 & 0 \end{pmatrix}. \tag{21.271}$$

Nous avons alors $A * z = a - \frac{1}{z}$. De plus la condition $|z| \leqslant 1$ revient à $|z = 1|$. Pour les nombres complexes de module 1, l'opération $z \to -1/z$ est la symétrie autour de l'axe des imaginaires purs. Le seul à ne pas sortir de D est le fameux $z = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$, qui revient sur lui-même avec $a = -1$.

Nous passons à la possibilité $c = -1$. Dans ce cas la matrice est de la forme

$$A = \begin{pmatrix} a & b \\ -1 & d \end{pmatrix}, \tag{21.272}$$

et nous revenons au cas $c = 1$ en prenant $-A$ au lieu de A.

19. Je ne rigolais pas quand je disais qu'on allait avoir de nombreux cas.

(2) Nous passons au cas $\Im(A*z)<\Im(z)$. Nous récrivons cette condition avec

$$\Im(A*z)<\Im\big(A^{-1}*(A*z)\big). \tag{21.273}$$

Si nous supposons que z et A sont tels que z et $A*z$ soient tous deux dans D, alors $z'=A*z$ est un élément de D tel que

$$\Im(z')<\Im(A^{-1}*z'). \tag{21.274}$$

Or nous avons vu qu'aucun élément de D vérifiant cette condition n'existait sans être trivial (celui qui ne bouge pas). Pour cela il suffit d'appliquer tout ce que nous venons de dire avec A^{-1} au lieu de A.

Quelque conclusions Après avoir passé tous les cas en revue, le fameux point $z_0=-\frac{1}{2}+\frac{\sqrt{3}}{2}i$ est l'unique point de D à accepter une matrice non triviale $A\in\mathrm{PSL}(2,\mathbb{Z})$ telle que $z_0=A*z_0$.

Nous remarquons aussi que tous les points de P sont ramenés dans D par une matrice obtenue comme produit de T, S, T^{-1} et S^{-1}.

□

Corollaire 21.94 ([264])**.**
Les matrices S et T génèrent le groupe modulaire au sens où toute matrice de $\mathrm{PSL}(2,\mathbb{Z})$ *s'écrit comme*

$$T^{m_1}S^{p_1}\dots T^{m_k}S^{p_k} \tag{21.275}$$

pour un certain k et des nombres $m_i,p_i\in\mathbb{Z}$. Autrement dit, $\mathrm{PSL}(2,\mathbb{Z})=\mathrm{gr}(S,T)$.

Démonstration. Soit z, un point de D autre que z_0. Alors si $A\in\mathrm{PSL}(2,\mathbb{Z})$ est non trivial nous avons $A*z$ hors de D. Du coup, comme vu dans la démonstration du théorème 21.93, il existe $B\in\mathrm{gr}(S,T)$ tel que $B*(A*z)\in D$. Vu que D ne contient qu'un seul point de chaque orbite, nous avons

$$B*A*z=z, \tag{21.276}$$

et donc $BA=\pm\mathbb{1}$, ce qui prouve que[20] $A=B^{-1}$, c'est à dire que $A\in\mathrm{gr}(S,T)$. □

20. Dans $\mathrm{PSL}(2,\mathbb{Z})$, nous n'avons pas besoin de mettre $\pm$ parce qu'il est compris dans la définition.

Chapitre 22

Analyse vectorielle

22.1 Le théorème de Green

Soit un champ de vecteurs

$$F(x,y,z) = \begin{pmatrix} F_1(x,y,z) \\ F_2(x,y,z) \\ F_2(x,y,z) \end{pmatrix} \tag{22.1}$$

et un chemin $\sigma\colon [a,b] \to \mathbb{R}^3$ donné par

$$\sigma(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}. \tag{22.2}$$

Nous avons défini la circulation de F le long de σ par

$$\begin{aligned} \int_\sigma F \cdot d\sigma &= \int_a^b F\big(\sigma(t)\big) \cdot \sigma'(t)dt \\ &= \int_a^b \Big[F_1\big(\sigma(t)\big)x'(t) + F_2\big(\sigma(t)\big)y'(y) + F_3\big(\sigma(t)\big)z'(t)\Big]dt \\ &= \int_\sigma F_1dx + F_2dy + F_3dz. \end{aligned} \tag{22.3}$$

La dernière ligne est juste une notation compacte[1]. Elle sert à se souvenir qu'on va mettre x' à côté de F_1, y' à côté de F_2 et z' à côté de F_3. L'avantage de cette notation est qu'on peut écrire d'autres combinaisons.

Si f et g sont deux fonctions sur $\mathbb{R}^3$, nous pouvons écrire

$$\int_\sigma f dy + g dz. \tag{22.4}$$

Cela signifie

$$\int_a^b \Big[f\big(\sigma(t)\big)y'(t) + g\big(\sigma(t)\big)z'(t)\Big]dt. \tag{22.5}$$

Soit D une région du plan et σ, son contour que nous prenons, par convention[2], dans l'orientation trigonométrique, comme indiqué sur la figure 22.1. Nous supposons également que le domaine D n'a pas de trous intérieurs.

Nous notons par $\sigma = \partial D$ le bord de D, c'est à dire le contour dont nous venons de parler.

1. Il y aurai beaucoup de choses à dire là-dessus, mais la vie est trop courte pour parler de formes différentielles, et c'est dommage.

2. Il y aurait beaucoup de choses à dire sur ça aussi, mais...

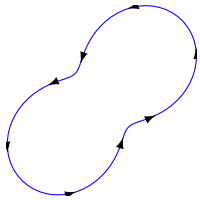


FIGURE 22.1 – Un contour avec son ortientation.

Théorème 22.1 (Théorème de Green)**.**
Soient $P, Q\colon D \to \mathbb{R}$ deux fonctions de classe C^1. Alors

$$\int_{\partial D} P dx + Q dy = \int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy. \tag{22.6}$$

Pour rappel, l'intégrale du membre de gauche signifie

$$\int_a^b \Big[P\big(\sigma(t)\big)\sigma'_x(t) + Q\big(\sigma(t)\big)\sigma'_y(t)\Big]dt. \tag{22.7}$$

Ce n'est d'ailleurs rien d'autre que l'intégrale du champ de vecteurs $\begin{pmatrix} P \\ Q \end{pmatrix}$.

Corollaire 22.2.
L'aire du domaine D est donnée par

$$A = \frac{1}{2}\int_{\partial D} (xdy - ydx). \tag{22.8}$$

Démonstration. L'intégrale $\int_{\partial D}(xdy - ydx)$ se traite avec le théorème de Green où l'on pose $P = -y$ et $Q = x$. Nous avons donc

$$\begin{aligned} \int_{\partial D} -ydx + xdy &= \int_D \left(\frac{\partial x}{\partial x} - \frac{\partial(-y)}{\partial y}\right) dxdy \\ &= \int_D 2\, dxdy. \end{aligned} \tag{22.9}$$

La dernière ligne est bien le double de la surface. □

Exemple 22.3
Calculons (encore une fois) l'aire du disque de rayon R. Il s'agit de calculer l'intégrale

$$I = \frac{1}{2}\int_\sigma (xdt - ydx) \tag{22.10}$$

où σ est le cercle donné par

$$\sigma(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} R\cos(t) \\ R\sin(t) \end{pmatrix} \tag{22.11}$$

Le calcul est

$$\begin{aligned} I &= \frac{1}{2}\int_0^{2\pi} \underbrace{R\cos(\theta)}_{x}\underbrace{R\cos(\theta)}_{y'} - \underbrace{R\sin(\theta)}_{y}\underbrace{(-R\sin(\theta))}_{x'}\, d\theta \\ &= \frac{R^2}{2}\int_0^\pi d\theta \\ &= \pi R^2. \end{aligned} \tag{22.12}$$

△

Exemple 22.4
Calculons l'aire de l'ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} \leqslant 1 \tag{22.13}$$

dont le bord est donné par

$$\begin{cases} x(t) = a\cos(t) & \text{(22.14a)} \\ y(t) = b\sin(t). & \text{(22.14b)} \end{cases}$$

Le terme xdy devient $a\cos(t)b\cos(t) = ab\cos^2(t)$ et le terme ydx devient $b\sin(t)(-a\sin(t)) = -ab\sin^2(t)$. L'intégrale qui donne la surface est donc

$$\frac{1}{2}\int_{\partial D}(xdy - ydx) = \frac{1}{2}\int_0^{2\pi} ab = \pi ab. \tag{22.15}$$

△

Le théorème de Green peut être mis sous une autre forme.

Théorème 22.5 (Théorème de Green, forme vectorielle)**.**
Si G est un champ de vecteurs sur D, nous avons

$$\int_{\partial D} G \cdot d\sigma = \int_D (\nabla \times G) \cdot dS \tag{22.16}$$

où le second membre est le flux de $\nabla \times G$ sur la surface D.

Démonstration. Analysons le membre de droite. Nous savons que D est une surface dans le plan $\mathbb{R}^2$. Le vecteur normal à la surface est donc simplement le vecteur (constant) e_z. Le produit scalaire $(\nabla \times F) \cdot dS$ est donc $(\nabla \times F) \cdot e_z$ et se réduit à la troisième composante du rotationnel, c'est à dire

$$\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}. \tag{22.17}$$

Cela est bien le membre de droite de l'équation (22.6). Le membre de gauche de cette dernière est bien le membre de gauche de (22.16). □

Exemple 22.6
Soit le champ de vecteurs $F(x,y) = \begin{pmatrix} xy^2 \\ y + x \end{pmatrix}$, et soit à calculer

$$\int_D \nabla \times F \cdot dS \tag{22.18}$$

où D est la région comprise entre les courbes $y = x^2$ et $y = x$ pour $x \geqslant 0$ (voir la figure 22.2).

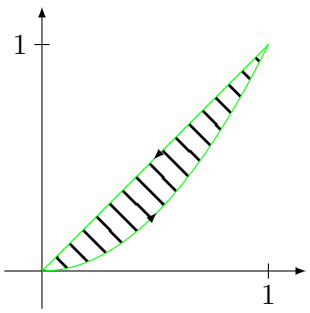


FIGURE 22.2 – Le contour d'intégration pour l'exemple 22.6.

Nous pouvons calculer cette intégrale directement en calculant le rotationnel de F :

$$\nabla \times F = \begin{pmatrix} 0 \\ 0 \\ 1-2xy \end{pmatrix}. \tag{22.19}$$

Par conséquent l'intégrale à effectuer est

$$I = \int_0^1 dx \int_{x^2}^x (1-2xy)dy = \frac{1}{12}. \tag{22.20}$$

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y)=1-2*x*y
sage: f.integrate(y,x**2,x).integrate(x,0,1)
(x, y) |--> 1/12
```

L'autre façon de calculer l'intégrale est d'utiliser le théorème de Green et de calculer la circulation de F le long de ∂D :

$$I = \int_{\partial D} F \cdot \sigma. \tag{22.21}$$

Le chemin $\sigma = \partial D$ est composé de la parabole $y = x^2$ et du segment de droite $x = y$. Attention : il faut respecter l'orientation. Nous avons

$$\sigma_1(t) = (t, t^2) \tag{22.22}$$

et

$$\sigma_2(t) = (1-t, 1-t). \tag{22.23}$$

Notez bien que le second chemin est $(1-t, 1-t)$ et non (t, t) parce qu'il faut le parcourir dans le bon sens (voir le dessin).

Commençons par le premier chemin :

$$\begin{aligned} \sigma_1(t) &= (t, t^2) \\ \sigma_1'(t) &= (1, 2t) \\ F\big(\sigma_1(t)\big) &= \begin{pmatrix} t^5 \\ t+t^2 \end{pmatrix}, \end{aligned} \tag{22.24}$$

et par conséquent

$$F\big(\sigma_1(t)\big) \cdot \sigma_1'(t) = t^5 + 2t^2 + 2t^3, \tag{22.25}$$

et le premier morceau de la circulation vaut

$$\int_{\sigma_1} F \cdot d\sigma_1 = \int_0^1 t^5 + 2t^2 + 2t^3 = \frac{4}{3}. \tag{22.26}$$

Pour le second chemin :

$$\begin{aligned} \sigma_2(t) &= (1-t, 1-t) \\ \sigma_2'(t) &= (-1, -1) \\ F\big(\sigma_2(t)\big) &= \begin{pmatrix} (1-t)^3 \\ 2(1-t) \end{pmatrix}. \end{aligned} \tag{22.27}$$

Par conséquent

$$F\big(\sigma_2(t)\big) \cdot \sigma_2(t) = -(1-t)^2 - 2(1-t). \tag{22.28}$$

Le second morceau de la circulation est par conséquent

$$\int_0^1 -(1-t)^2 - 2(1-t)dt = -\frac{5}{4}. \tag{22.29}$$

La circulation de F le long de σ est donc égale à

$$\frac{4}{3} - \frac{5}{4} = \frac{1}{12}. \tag{22.30}$$

Comme prévu, nous obtenons le même résultat. △

22.2 Théorème de la divergence dans le plan

22.2.1 La convention de sens de parcours

Soient D, un domaine dans le plan et une paramétrisation

$$\begin{aligned} \sigma\colon [a,b] &\to \mathbb{R}^2 \\ t &\mapsto \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \end{aligned} \tag{22.31}$$

une paramétrisation du bord ∂D de D. La normale à σ est perpendiculaire à la tangente, donc la normale extérieure de norme 1 vaut

$$n = \frac{\big(y'(t), -x'(t)\big)}{\sqrt{\big(x'(t)\big)^2 + \big(y'(t)\big)^2}} \quad \text{ou} \quad n- = \frac{\big(y'(t), -x'(t)\big)}{\sqrt{\big(x'(t)\big)^2 + \big(y'(t)\big)^2}}. \tag{22.32}$$

Comment faire le choix ?

Nous prenons comme convention que le sens *du chemin* doit être tel que le vecteur normal extérieur soit

$$n = \frac{\big(y'(t), -x'(t)\big)}{\sqrt{\big(x'(t)\big)^2 + \big(y'(t)\big)^2}}. \tag{22.33}$$

Donc si le chemin σ donne lieu à un vecteur n pointant vers l'intérieur, il faut utiliser le chemin qui va dans le sens contraire : $\tilde{\sigma}(t) = \sigma(1-t)$.

Les vecteurs tangents et normaux d'un contour sont dessinés sur la figure 22.3.

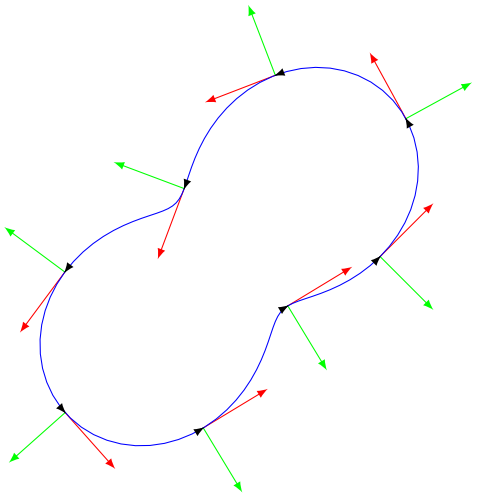


FIGURE 22.3 – Le champ de vecteurs tangents est dessiné en rouge tandis qu'en vert nous avons le champ de vecteurs normaux extérieurs.

22.2.2 Théorème de la divergence

Théorème 22.7 (Théorème de la divergence)**.**
Soit F un champ de vecteurs sur $\mathbb{R}^2$. Le flux de F à travers le bord de D est égal à l'intégrale de la divergence de F sur D. En formule :

$$\int_{\partial D} F \cdot n \, d\sigma = \int_D \nabla \cdot F \, dxdy. \tag{22.34}$$

Démonstration. Tant $F \cdot n$ que $\nabla \times F$ sont des fonctions. Le membre de gauche est donc l'intégrale d'une fonction sur un chemin et le membre de droite est l'intégrale d'une fonction sur une surface. Notre convention de sens de parcours du chemin permet d'écrire le produit scalaire $F \cdot n$ sous la forme suivante :

$$\begin{aligned} F \cdot n &= \frac{1}{\|\sigma'\|} \begin{pmatrix} F_x \\ F_y \end{pmatrix} \cdot \begin{pmatrix} y' \\ -x' \end{pmatrix} \\ &= \frac{1}{\|\sigma'\|} (F_x y' - F_y x') \\ &= \frac{1}{\|\sigma'\|} \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} \\ &= \frac{1}{\|\sigma'\|} \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \sigma'. \end{aligned} \tag{22.35}$$

Par conséquent, la *fonction*

$$F \cdot n \tag{22.36}$$

est la même que la *fonction*

$$\frac{1}{\|\sigma'\|} \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \sigma'. \tag{22.37}$$

L'intégrale de cette dernière fonction sur le chemin σ est

$$\begin{aligned} I &= \int_\sigma F \cdot n \\ &= \int_\sigma \frac{1}{\|\sigma'\|} \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \sigma' \\ &= \int_a^b \frac{1}{\|\sigma'(t)\|} \begin{pmatrix} -F_y\big(\sigma(t)\big) \\ F_x\big(\sigma(t)\big) \end{pmatrix} \cdot \sigma'(t) \|\sigma'(t)\| dt \\ &= \int_a^b \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \sigma'(t) dt. \end{aligned} \tag{22.38}$$

Cette dernière intégrale est la circulation du champ de vecteurs $\begin{pmatrix} -F_y \\ F_x \end{pmatrix}$ sur le chemin σ. Le théorème de Green 22.5 nous enseigne que la circulation le long d'un chemin est égale au flux du rotationnel à travers la surface. Par conséquent,

$$I = \int_D \left(\nabla \times \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \right) \cdot dS = \int_D \nabla \cdot F \, dxdy \tag{22.39}$$

□

22.3 Théorème de Stokes

Nous nous mettons maintenant dans $\mathbb{R}^3$, et nous y considérons une surface paramétrée S donc le bord est ∂S.

Théorème 22.8 (Théorème de Stokes)**.**
Alors le flux du rotationnel de F à travers S est égal à la circulation de F le long du bord. En formule :

$$\int_S \nabla \times F \cdot dS = \int_{\partial S} F \cdot d\sigma. \tag{22.40}$$

Nous pouvons nous donner une idée du pourquoi ce théorème est vrai. D'abord, si la surface est plate, cela est exactement le théorème de Green 22.5. Supposons maintenant que le bord reste plat, mais que la surface se déforme un petit peu. Le chemin

$$\sigma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \\ 0 \end{pmatrix} \tag{22.41}$$

est tout autant le bord du disque plat de rayon 1 que celui de la demi-sphère

$$\phi(x, y) = \begin{pmatrix} x \\ y \\ \sqrt{1 - x^2 - y^2} \end{pmatrix}. \tag{22.42}$$

Le champ de vecteur que nous considérons est $G = \nabla \times F$. Il a un certain flux à travers le disque plat, et ce plus est égal à la circulation de F sur σ. Quel est le flux de G à travers la demi-sphère ? Étant donné que $\nabla \cdot G = \nabla \cdot (\nabla \times F) = 0$, le champ de vecteurs G est incompressible, de telle façon que tout ce qui rentre dans la demi-sphère doit en sortir. Le flux de G à travers la demi-sphère doit par conséquent être égal à celui à travers le disque plat.

Exemple 22.9

Soit C l'intersection entre le cylindre $x^2 + y^2 = 1$ et le plan $x + y + z = 1$. Calculer la circulation de

$$F(x, y, z) = \begin{pmatrix} -y^3 \\ x^3 \\ -z^3 \end{pmatrix} \tag{22.43}$$

le long de C.

Au lieu de calculer directement

$$\int_C F \cdot d\sigma, \tag{22.44}$$

nous allons calculer

$$\int_S \nabla \times F \cdot dS \tag{22.45}$$

où S est une surface dont C est le bord. Cette intégrale est à calculer avec la formule (14.209).

La première chose à faire est de trouver une surface dont le bord est C et en trouver une paramétrisation ϕ. Le plus simple est de prendre le graphe du plan sur le cercle $x^2 + y^2 + 1$. Une paramétrisation de cette surface est simplement

$$\begin{aligned} \phi\colon D &\to \mathbb{R}^3 \\ (x, y) &\mapsto \begin{pmatrix} x \\ y \\ 1 - x - y \end{pmatrix} \end{aligned} \tag{22.46}$$

où D est le disque de rayon 1. Étant donné que cela paramètre le plan $x + y + z - 1 = 0$, le vecteur normal est $n = e_x + e_y + z_z$. Nous pouvons cependant calculer ce vecteur normal en suivant la recette usuelle. D'abord les vecteurs tangents sont

$$\frac{\partial \phi}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \frac{\partial \phi}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}. \tag{22.47}$$

Et le vecteur normal est donné par le produit vectoriel :

$$\begin{aligned} n &= \frac{\partial\phi}{\partial x} \times \frac{\partial\phi}{\partial y} \\ &= \begin{vmatrix} e_x & e_y & e_z \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{vmatrix} \\ &= e_x + e_y + z_z. \end{aligned} \tag{22.48}$$

Ensuite, le rotationnel de F est donné par

$$\nabla \times F = 3(x^2 + y^2)e_z. \tag{22.49}$$

Par conséquent,

$$\nabla \times F \cdot \left(\frac{\partial\phi}{\partial x} \times \frac{\partial\phi}{\partial y}\right) = 3(x^2 + y^2). \tag{22.50}$$

L'intégrale à calculer est donc

$$\begin{aligned} \int_S \nabla \times F \cdot dS &= \int_D (\nabla \times F)\big(\phi(x,y)\big) \cdot \left(\frac{\partial\phi}{\partial x} \times \frac{\partial\phi}{\partial y}\right) dxdy \\ &= 3\int_D (x^2 + y^2)dxdy. \end{aligned} \tag{22.51}$$

Cette dernière intégrale est l'intégrale d'une fonction sur le disque de rayon 1. Elle s'effectue en passant aux coordonnées polaires :

$$3\int_D (x^2 + y^2)dxdy = \int_0^{2\pi} d\theta \int_0^1 (r^2)r\,dr = \frac{3\pi}{2}. \tag{22.52}$$

△

22.4 Théorème de Gauss

Soit V une partie de $\mathbb{R}^3$ délimitée par une surface S sur laquelle nous considérons la normale extérieure. Soit F un champ de vecteurs sur $\mathbb{R}^3$.

Théorème 22.10 (Théorème de la divergence ou de Gauss)**.**
Le flux d'un champ de vecteur F à travers une surface fermée est égale à l'intégrale de la divergence sur le volume correspondant :

$$\int_{\partial V} F \cdot dS = \int_V \nabla \cdot F\,dxdydz. \tag{22.53}$$

Ce théorème signifie que la quantité de fluide qui s'accumule dans le volume (le flux est ce qui rentre moins ce qui sort) est égal à l'intégrale de $\nabla \cdot F$ sur le volume, alors que nous savons que, localement, la quantité $\nabla \cdot F(x,y,z)$ est la quantité de fluide qui s'accumule au point (x,y,z).

Remarque 22.11.
Ce théorème ne fonctionne qu'avec des surfaces fermées. Essayer de l'appliquer au calcul de flux à travers des surfaces ouvertes n'a pas de sens parce qu'une surface ouverte ne délimite pas un volume.

22.12.
La formule de la divergence peut être utilisée comme intégration par partie. Si u est une fonction et F un champ de vecteurs, $\nabla(uF) = \nabla(u) \cdot F + u\nabla \cdot F$ et alors

$$\int_{\partial V} uF \cdot n = \int V\nabla(uF) = \int_V u\nabla \cdot F + \int_V F \cdot \nabla u \tag{22.54}$$

où n est le champ de vecteurs normal extérieur à V. En remettant les termes dans un ordre qui ressemble plus à l'intégration par partie :

$$\int_V F \cdot \nabla u = \int_{\partial V} uF \cdot n - \int_V u \nabla F. \tag{22.55}$$

Exemple 22.13
Calculer le flux du champ de vecteurs

$$F(x,y,z) = \begin{pmatrix} 2x \\ y^2 \\ z^2 \end{pmatrix} \tag{22.56}$$

à travers la sphère de rayon 1 centrée à l'origine. Nous utilisons le théorème de la divergence

$$\int_S F \cdot n \, dS = \int_B \nabla \cdot F \, dxdydz \tag{22.57}$$

où S est la sphère et B est la boule (la sphère pleine). La divergence de F se calcule :

$$\nabla \cdot F = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} = 2 + 2x + 2y. \tag{22.58}$$

L'intégrale est donc en trois termes :

$$\begin{aligned} \int_B 2 &= 2\text{Volume(B)} = \frac{8\pi}{3} \\ \int_B y \, dxdydz &= 0 \\ \int_B z \, dxdydz &= 0. \end{aligned} \tag{22.59}$$

△

Dans certains cas le théorème de Gauss permet de simplifier le calcul de l'intégrale d'une fonction sur une surface.

Exemple 22.14
Soit à calculer l'intégrale

$$I = \int_{\partial B} (x^2 + y + z) dS, \tag{22.60}$$

c'est à dire l'intégrale de la fonction $x^2 + y + z$ sur la sphère. Le vecteur normal à la sphère est

$$n = xe_x + ye_y + ze_z. \tag{22.61}$$

Étant donné que nous sommes sur la sphère de rayon 1, ce vecteur est même normé. La fonction que nous regardons n'est rien d'autre que $F \cdot n$ avec

$$F = \begin{pmatrix} x \\ 1 \\ 1 \end{pmatrix}. \tag{22.62}$$

Nous pouvons donc simplement intégrer $\nabla \cdot F$ sur toute la boule :

$$I = \int_B \nabla \cdot F \, dxdydz = \int_B 1 \, dxdudz = \frac{4\pi}{3}. \tag{22.63}$$

△

22.5 Coordonnées curvilignes

22.5.1 Base locale

Les coordonnées sphériques et cylindriques sont deux systèmes de coordonnées «un peu courbe» qui existent sur $\mathbb{R}^3$. Il en existe de nombreux autres, que nous appelons **coordonnées curvilignes**. Des coordonnées curvilignes sur $\mathbb{R}^3$ est n'importe quel [3] système qui permet de repérer un point de $\mathbb{R}^3$ à partir de trois nombres.

Il s'agit donc d'un ensemble de trois applications

$$x_i\colon \mathbb{R}^3 \to \mathbb{R}. \tag{22.64}$$

Les coordonnées cylindriques sont

$$\begin{cases} x_1(r,\theta,z) = r\cos\theta & \text{(22.65a)}\\ x_2(r,\theta,z) = r\sin\theta & \text{(22.65b)}\\ x_3(r,\theta,z) = z & \text{(22.65c)} \end{cases}$$

Soit donc un système général $q = (q_1, q_2, q_3)$ et

$$M(q) = \begin{pmatrix} x_1(q)\\ x_2(q)\\ x_3(q) \end{pmatrix}. \tag{22.66}$$

Si nous fixons q_2 et q_3 et que nous laissons varier q_1, nous obtenons une courbe [4] dont nous pouvons considérer le vecteur vitesse, c'est à dire le vecteur tangent. En chaque point nous avons ainsi trois vecteurs

$$\frac{\partial M}{\partial q_i}(q). \tag{22.67}$$

Nous disons que le système de coordonnées curviligne est **orthogonal** si ces trois vecteurs sont orthogonaux. Dans la suite nous supposerons que c'est toujours le cas.

Nous posons

$$h_i = \left\| \frac{\partial M}{\partial q_i} \right\| \tag{22.68}$$

et nous considérons les trois vecteurs normés

$$e_i = h_i^{-1} \frac{\partial M}{\partial q_i}. \tag{22.69}$$

Les trois vecteurs $\{e_1, e_2, e_3\}$ forment une base orthonormée dite **base locale**. Ce sont des vecteurs liés [5] au point M.

22.5.2 Importance de l'orthogonalité

Nous avons dit que nous nous restreignons au cas où les vecteurs e_i sont orthogonaux. En termes de produits scalaires, cela signifie

$$e_i \cdot e_j = \delta_{ij}. \tag{22.70}$$

Nous en étudions maintenant quelque conséquence. L'équation (22.69) peut s'écrire plus explicitement sous la forme

$$e_i = \sum_k h_i^{-1} \frac{\partial x_k}{\partial q_i} 1_k. \tag{22.71}$$

3. Nous n'entrons pas dans les détails de régularité.
4. Dans le cas des sphériques, c'est une demi-droite horizontale d'angle q_2 et de hauteur q_3.
5. En géométrie différentielle on dira que ce sont des élément de l'espace tangent, mais c'est une toute autre histoire.

Notez que pour chaque k et i, la quantité $h_i^{-1}\frac{\partial x_k}{\partial q_i}$ est un simple nombre. Nous allons les mettre dans une matrice :

$$A_{ki} = h_i^{-1}\frac{\partial x_k}{\partial q_i}. \tag{22.72}$$

Cela nous donne le changement de base

$$e_i = \sum_k A_{ki} 1_k. \tag{22.73}$$

Le produit $e_i \cdot e_j$ s'écrit alors

$$\begin{aligned} e_i \cdot e_j &= \sum_{kl} A_{ki}A_{lj} \underbrace{1_k \cdot 1_l}_{=\delta_{kl}} \\ &= \sum_{kl} A_{ki}A_{lj}\delta_{kl} \\ &= \sum_k A_{ki}A_{kj} \\ &= \sum_k (A^T)_{ik}A_{kj}. \end{aligned} \tag{22.74}$$

Or cela doit valoir δ_{ij}. Par conséquent

$$A^T = A^{-1}. \tag{22.75}$$

Le fait que les coordonnées curvilignes considérées soient orthogonales s'exprime donc par la fait que la matrice de changement de base est une matrice orthogonale.

Cette circonstance nous permet d'inverser le changement de base (22.73) en multipliant cette équation par $(A^{-1})_{il}$ des deux côtés et en faisant la somme sur i :

$$\sum_i (A^{-1})_{il} e_i = \sum_{kl} \underbrace{A_{ki}(A^{-1})_{il}}_{=\delta_{kl}} 1_k, \tag{22.76}$$

par conséquent

$$\sum_i (A^T)_{il} e_i = 1_l, \tag{22.77}$$

et

$$1_l = \sum_i A_{li} e_i = \sum_i h_i^{-1}\frac{\partial x_l}{\partial q_i} e_i. \tag{22.78}$$

Armés de cette importante formule, nous pouvons exprimer les quantités que nous connaissons dans la base canonique en termes de la base locale.

Une autre conséquence du fait que e_1, e_2 et e_3 est une base orthonormée est que, éventuellement en réordonnant les vecteurs, on a

$$\begin{aligned} e_1 \times e_2 &= e_3 \\ e_2 \times e_3 &= e_1 \\ e_3 \times e_1 &= e_2 \end{aligned} \tag{22.79}$$

Ces trois relations s'écrivent en une seule avec

$$e_i \times e_j = \sum_k \epsilon_{ijk} e_k \tag{22.80}$$

où

$$\epsilon_{ijk} = \begin{cases} 0 & \text{si } i,j,k \text{ ne sont pas tous différents} \\ 1 & \text{si } ijk \text{ se ramène à 123 par un nombre pair de permutations} \\ -1 & \text{si } ijk \text{ se ramène à 123 par un nombre impair de permutations} \end{cases} \tag{22.81}$$

est le **symbole de Levi-Civita**. La formule du produit vectoriel peut également être utilisée à l'envers sous la forme

$$e_k = \frac{1}{2}\sum_{ij} \epsilon_{ijk}\, e_i \times e_j. \tag{22.82}$$

Le symbole de Levi-Civita possède de nombreuses formules. En voici certaines, facilement démontrables en considérant tous les cas :

$$\epsilon_{ijk}\epsilon_{ijl} = \delta_{kl}|\epsilon_{ijk}|. \tag{22.83}$$

Grâce au symboles de Levi-Civita, le produit mixte des vecteurs de base a une belle forme :

$$e_l \cdot (e_i \times e_j) = \sum_k \epsilon_{ijk} e_l \times e_k = \sum_k \epsilon_{ijk}\delta_{lk} = \epsilon_{ijl}. \tag{22.84}$$

22.5.3 Coordonnées polaires

Les coordonnées curvilignes polaires sont données par

$$M(r,\theta) = \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}, \tag{22.85}$$

et par conséquent

$$\frac{\partial M}{\partial r} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \quad \frac{\partial M}{\partial \theta} = \begin{pmatrix} -r\sin(\theta) \\ r\cos(\theta) \end{pmatrix}. \tag{22.86}$$

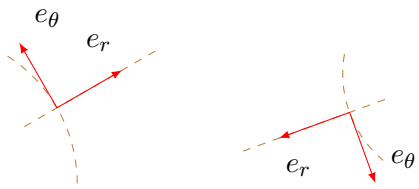
Nous avons les normes $h_r = 1$ et $h_\theta = r$, et donc les vecteurs de la base locale en (r,θ) sont

$$e_r = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = \cos(\theta)e_x + \sin(\theta)e_y \tag{22.87}$$

ainsi que

$$e_\theta = \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix} = -\sin(\theta)e_x + \cos(\theta)e_y. \tag{22.88}$$

Ces vecteurs sont représentés à la figure 22.4. Notez qu'il y en a une paire différente en chaque point.



(a) Base locale. (b) Base locale.

FIGURE 22.4 – En brun, les lignes que le point suivrait si on ne variait qu'une coordonnées polaire à la fois. Les vecteurs rouges sont les vecteurs e_r et e_θ.

22.5.4 Coordonnées cylindriques

Les coordonnées cylindriques sont les mêmes que les coordonnées polaires à part qu'il faut écrire

$$M(r,\theta,z) = \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \\ z \end{pmatrix}, \tag{22.89}$$

et nous avons le vecteur de base supplémentaire

$$e_z = \frac{\partial M}{\partial z} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{22.90}$$

parce que $h_z = 1$.

22.5.5 Coordonnées sphériques

Les coordonnées curvilignes sphériques sont données par

$$M(\rho,\theta,\varphi)=\begin{pmatrix}\rho\sin(\theta)\cos(\varphi)\\ \rho\sin(\theta)\sin(\varphi)\\ \rho\cos(\theta)\end{pmatrix}, \tag{22.91}$$

dont les dérivées sont données par

$$\begin{aligned}\frac{\partial M}{\partial r}&=\begin{pmatrix}\sin(\theta)\cos(\varphi)\\ \sin(\theta)\sin(\varphi)\\ \cos(\theta)\end{pmatrix},\qquad \frac{\partial M}{\partial\theta}=\begin{pmatrix}\rho\cos(\theta)\cos(\varphi)\\ \rho\cos(\theta)\sin(\varphi)\\ -\rho\sin(\theta)\end{pmatrix},\\ \frac{\partial M}{\partial\varphi}&=\begin{pmatrix}-\rho\sin(\theta)\sin(\varphi)\\ \rho\sin(\theta)\cos(\varphi)\\ 0\end{pmatrix}\end{aligned} \tag{22.92}$$

Les normes de ces vecteurs sont $h_\rho=1$, $h_\theta=\rho$ et $h_\varphi=\rho\sin(\theta)$. Les vecteurs de la base locale en (ρ,θ,φ) sont donc

$$\begin{aligned}e_r&=\begin{pmatrix}\sin(\theta)\cos(\varphi)\\ \sin(\theta)\sin(\varphi)\\ \cos(\theta)\end{pmatrix},\quad e_\theta=\begin{pmatrix}\cos(\theta)\cos(\varphi)\\ \cos(\theta)\sin(\varphi)\\ -\sin(\theta)\end{pmatrix},\\ e_\varphi&=\begin{pmatrix}-\sin(\varphi)\\ \cos(\varphi)\\ 0\end{pmatrix}\end{aligned} \tag{22.93}$$

22.5.6 Gradient en coordonnées curvilignes

Soit $(x,y,z)\mapsto f(x,y,z)$ une fonction sur $\mathbb{R}^3$. Nous pouvons la composer avec les coordonnées curvilignes q pour obtenir la fonction

$$\tilde{f}(q_1,q_2,q_3)=f\big(x_1(q),x_2(x),x_3(q)\big). \tag{22.94}$$

Nous disons que $\tilde{f}$ est l'expression de f dans les coordonnées q. Nous savons déjà comment calculer le gradient de f en coordonnées cartésiennes :

$$F(x,y,z)=\nabla f(x,y,z)=\begin{pmatrix}\partial_x f(x,y,z)\\ \partial_y f(x,y,z)\\ \partial_z f(x,y,z)\end{pmatrix}. \tag{22.95}$$

Cela est un vecteur lié au point (x,y,z). Nous voudrions exprimer ce vecteur dans la base $\{e_1,e_2,e_3\}$. En d'autres termes, nous voudrions trouver les nombres $\tilde{F}_1$, $\tilde{F}_2$ et $\tilde{F}_3$ tels que

$$F(x,y,z)=F\big(x(q),y(q),z(q)\big)=\tilde{F}_1e_1+\tilde{F}_2e_2+\tilde{F}_3e_3. \tag{22.96}$$

Ces nombres seront des fonctions de (q_1,q_2,q_3).

Par définition,

$$\nabla f=\sum_l\frac{\partial f}{\partial x_l}1_l. \tag{22.97}$$

En remplaçant 1_l par sa valeur en termes des e_i par la formule (22.78),

$$\begin{aligned}\nabla f&=\sum_l\frac{\partial f}{\partial x_l}1_l\\ &=\sum_l\frac{\partial f}{\partial x_l}\sum_i h_i^{-1}\frac{\partial x_l}{\partial q_i}e_i\\ &=\sum_{il}\frac{1}{h_i}\frac{\partial f}{\partial x_l}\frac{\partial x_l}{\partial q_i}e_i\\ &=\sum_i\frac{1}{h_i}\frac{\partial\tilde{f}}{\partial q_i}e_i.\end{aligned} \tag{22.98}$$

Plus explicitement,

$$\nabla f\big(x(q),y(q),z(q)\big)=\sum_i\frac{1}{h_i(q)}\frac{\partial\tilde{f}}{\partial q_i}(q)e_i \tag{22.99}$$

où

$$h_i(q)=\left\|\frac{\partial M}{\partial q_i}(q)\right\|. \tag{22.100}$$

Le plus souvent nous n'allons pas noter explicitement la dépendance de h_i en q.

22.5.6.1 Coordonnées sphériques

Nous pouvons exprimer le gradient d'une fonction en coordonnées sphériques en utilisant la formule (22.99) :

$$\nabla\tilde{f}(\rho,\theta,\varphi)=\frac{\partial\tilde{f}}{\partial\rho}e_\rho+\frac{1}{\rho}\frac{\partial\tilde{f}}{\partial\theta}e_\theta+\frac{1}{\rho\sin(\theta)}\frac{\partial\tilde{f}}{\partial\varphi}r_\varphi. \tag{22.101}$$

Cette expression peut paraître peu pratique parce que les vecteurs e_ρ, e_θ et e_φ eux-mêmes changent en chaque point. Elle est effectivement peu adaptée au dessin, mais elle est très pratique pour des fonctions ayant des symétries.

Exemple 22.15
Le potentiel de la gravitation est la fonction

$$V(x,y,z)=\frac{1}{\sqrt{x^2+y^2+z^2}}. \tag{22.102}$$

En coordonnées sphériques elle s'écrit

$$\tilde{V}(\rho,\theta,\varphi)=\frac{1}{\rho}. \tag{22.103}$$

En voila une fonction qu'elle est facile à dériver, contrairement à V ! En suivant la formule (22.101), nous avons immédiatement

$$\nabla\tilde{V}=-\frac{1}{\rho^2}e_\rho. \tag{22.104}$$

Nous voyons immédiatement que cela est un champ de vecteurs dont la norme diminue comme le carré de la distance à l'origine et qui est en permanence dirigé vers l'origine. △

22.5.7 Divergence en coordonnées curvilignes

Nous savons que

$$\nabla\tilde{f}=\sum_j\frac{1}{h_j}\frac{\partial\tilde{f}}{\partial q_j}e_j. \tag{22.105}$$

Nous pouvons en particulier considérer la fonction $f(q)=q_i$. De la même manière que nous avions noté x_i la fonction $x\mapsto x_i$, nous notons q_i la fonction $q\mapsto q_i$. Le gradient de cette fonction est donné par

$$\nabla q_i=\sum_j\frac{1}{h_j}\frac{\partial q_i}{\partial q_j}e_j, \tag{22.106}$$

mais $\frac{\partial q_i}{\partial q_j}=\delta_{ij}$, donc

$$\nabla q_i=\frac{e_i}{h_i}, \tag{22.107}$$

ou encore

$$e_i=h_i\nabla q_i. \tag{22.108}$$

Cela n'est pas étonnant : la direction dans laquelle la coordonnées q_i varie le plus est le vecteur e_i qui donne la tangente à la courbe obtenue lorsque *seul* q_i varie.

Commençons par calculer la divergence de e_i. En utilisant la formule (22.82),

$$\nabla \cdot e_k = \frac{1}{2}\sum_{ij}\epsilon_{ijk}\nabla \cdot (e_i \times e_j). \tag{22.109}$$

Nous avons, en utilisant les règles de Leibnitz (12.827),

$$\begin{aligned}\nabla \cdot (e_i \times e_j) &= \nabla \cdot (h_i\nabla q_i \times h_j\nabla q_j)\\ &= \nabla(h_ih_j) \cdot \big(\nabla q_i \times \nabla q_j\big) + h_ih_j\nabla \cdot \big(\nabla q_i \times \nabla q_j\big)\\ &= \nabla(h_ih_j) \cdot \big(\nabla q_i \times \nabla q_j\big)\\ &\quad + h_ih_j\nabla q_j \cdot \big(\underbrace{\nabla \times \nabla q_i}_{=0}\big)\\ &\quad + h_ih_j\nabla q_i \cdot \big(\underbrace{\nabla \times \nabla q_j}_{=0}\big)\end{aligned} \tag{22.110}$$

Cela nous fait

$$\nabla \cdot e_k = \sum_{ij}\epsilon_{ijk}\frac{\nabla(h_ih_j)}{h_ih_j} \cdot (e_i \times e_j). \tag{22.111}$$

parce que $\nabla q_i = h_i^{-1}e_i$. Nous pouvons développer le gradient qui intervient :

$$\nabla(h_ih_j) = \sum_l \frac{1}{h_l}\frac{\partial}{\partial q_l}(h_ih_j)e_l. \tag{22.112}$$

Nous voyons donc arriver le produit mixte $e_l \cdot (e_i \times e_j)$. En utilisant la formule (22.84), cela s'exprime directement sous la forme ϵ_{ijl}.

Nous avons alors

$$\begin{aligned}\nabla \cdot e_k &= \frac{1}{2}\sum_{ijl}\frac{1}{h_ih_jh_l}\frac{\partial}{\partial q_l}(h_ih_j)\epsilon_{ijk}\epsilon_{ijl}\\ &= \frac{1}{2}\sum_{ijl}\delta_{kl}|\epsilon_{ijk}|\frac{\partial}{\partial q_l}(h_ih_j)\\ &= \frac{1}{2}\sum_{ij}\frac{|\epsilon_{ijk}|}{h_ih_jh_k}\frac{\partial}{\partial q_k}(h_ih_j).\end{aligned} \tag{22.113}$$

Par exemple,

$$\nabla \cdot e_1 = \frac{1}{h_1h_2h_3}\frac{\partial}{\partial q_1}(h_2h_3). \tag{22.114}$$

Nous devons maintenant chercher le gradient d'un champ général

$$F(q) = \sum_k F_k(q)e_k. \tag{22.115}$$

La première chose à faire est d'utiliser la formule de Leibnitz :

$$\nabla \cdot F = \sum_k \nabla F_k(q) \cdot e_k + \sum_k F_k(q)\nabla \cdot e_k. \tag{22.116}$$

Afin d'alléger les notations, nous allons nous concentrer sur le terme numéro k et ne pas écrire la somme. Si i et j sont les nombres tels que $\epsilon_{ijk} = 1$, alors ce que la formule (22.113) signifie, c'est que

$$\nabla \cdot e_k = \frac{1}{h_1h_2h_3}\frac{\partial}{\partial q_k}(h_ih_j). \tag{22.117}$$

Nous savons déjà par la formule (22.99) que

$$\nabla F_k = \sum_l \frac{1}{h_l}\frac{\partial F_k}{\partial q_l}e_l, \tag{22.118}$$

par conséquent

$$\nabla F_k \cdot e_k = \sum_l \frac{1}{h_l} \frac{\partial F_k}{\partial q_l} \delta_{kl} = \frac{1}{h_k} \frac{\partial F_k}{\partial q_k}. \tag{22.119}$$

Pour obtenir cela nous avons utilisé le fait que $e_l \cdot e_k = \delta_{lk}$. Le terme numéro k de la somme (22.116) est donc

$$\frac{1}{h_k} \frac{\partial F_k}{\partial q_k} + \frac{F_k}{h_k h_i h_j} \frac{\partial (h_i h_j)}{\partial q_k} = \frac{1}{h_i h_j h_k} \frac{\partial (F_k h_i h_j)}{\partial q_k} \tag{22.120}$$

où il est entendu que i et j représentent les nombres tels que $\epsilon_{ijk} = 1$.

Au final, nous avons

$$\nabla \cdot F = \frac{1}{h_1 h_2 h_3} \sum_{ijk} |\epsilon_{ijk}| \frac{\partial (F_k h_i h_j)}{\partial q_k}. \tag{22.121}$$

Ici, la somme sur i et j consiste seulement à sélectionner les termes tels que i et j ne sont pas k. En écrivant la somme explicitement,

$$\nabla \cdot F = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial q_1}(F_1 h_2 h_3) + \frac{\partial}{\partial q_2}(F_2 h_1 h_3) + \frac{\partial}{\partial q_3}(F_3 h_1 h_2) \right]. \tag{22.122}$$

22.5.7.1 Coordonnées cylindriques

En coordonnées cylindriques, nous avons déjà vu que $h_r = 1$, $h_\theta = r$ et $h_z = 1$. La divergence est donc donnée par

$$\nabla \cdot F = \frac{1}{r} \left[\frac{\partial}{\partial r}(r F_r) + \frac{\partial}{\partial \theta}(F_\theta) + \frac{\partial}{\partial z}(r F_z) \right]. \tag{22.123}$$

Par exemple si

$$F(r, \theta, z) = r e_\theta + e_z, \tag{22.124}$$

nous avons

$$(\nabla \cdot F)(r, \theta, z) = \frac{1}{r} \left[\frac{\partial}{\partial \theta}(r) + \frac{\partial}{\partial z}(r) \right] = 0. \tag{22.125}$$

Cela est logique parce que $r e_\theta$ est à peu près le champ dont nous avons parlé dans l'exemple (12.314), qui était à divergence nulle. En réalité, le champ dont on parlait dans cet exemple était exactement $-e_\theta$. Le champ e_z est également à divergence nulle parce qu'il est constant.

22.5.7.2 Coordonnées sphériques

En coordonnées sphériques, nous avons $h_\rho = 1$, $h_\theta = r$ et $h_\varphi = r \sin\theta$, donc

$$\nabla \cdot F = \frac{1}{r^2 \sin\theta} \left[\frac{\partial}{\partial \rho}(\rho^2 \sin\theta F_\rho) + \frac{\partial}{\partial \theta}(\rho \sin\theta F_\theta) + \frac{\partial}{\partial \varphi}(\rho F_\varphi) \right]. \tag{22.126}$$

si $F(\rho, \theta, \varphi) = F_\rho e_\rho + F_\theta e_\theta + F_\varphi e_\varphi$.

22.5.8 Laplacien en coordonnées curvilignes orthogonales

Soit une fonction $f : \mathbb{R}^3 \to \mathbb{R}$. Le **laplacien** de f est donné par

$$\Delta f = \nabla \cdot (\nabla f). \tag{22.127}$$

En utilisant les formules données, nous avons

$$\Delta f = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial q_1} \left(\frac{h_2 h_3}{h_1} \frac{\partial f}{\partial q_1} \right) + \frac{\partial}{\partial q_2} \left(\frac{h_1 h_3}{h_2} \frac{\partial f}{\partial q_2} \right) + \frac{\partial}{\partial q_3} \left(\frac{h_1 h_2}{h_3} \frac{\partial f}{\partial q_3} \right) \right]. \tag{22.128}$$

Dans cette expression, la fonction f est donnée comme fonction de q_1, q_2 et q_3.

En coordonnées cylindriques, cela s'écrit

$$\begin{aligned}\Delta f &= \frac{1}{r}\left[\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right)+\frac{\partial}{\partial\theta}\left(\frac{1}{r}\frac{\partial f}{\partial\theta}\right)+\frac{\partial}{\partial z}\left(r\frac{\partial f}{\partial z}\right)\right]\\ &= \frac{\partial^2 f}{\partial r^2}+\frac{1}{r}\frac{\partial f}{\partial r}+\frac{1}{r^2}\frac{\partial^2 f}{\partial\theta^2}+\frac{\partial^2 f}{\partial z^2}.\end{aligned} \tag{22.129}$$

Dans cette expression, f est fonction de r, θ et z.

En coordonnées sphériques, cela devient

$$\Delta f = \frac{1}{\rho^2\sin\theta}\left[\frac{\partial}{\partial\rho}\left(\rho^2\sin\theta\frac{\partial f}{\partial\rho}\right)+\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right)+\frac{\partial}{\partial\varphi}\left(\frac{1}{\sin\theta}\frac{\partial f}{\partial\varphi}\right)\right]. \tag{22.130}$$

Dans cette expression, f est fonction de ρ, θ et φ.

22.5.9 Rotationnel en coordonnées curvilignes orthogonales

Nous voulons calculer le rotationnel de $F(q)=\sum_k F_k(q)e_k$. Pour cela nous commençons par écrire $e_k = h_k\nabla q_k$ et nous utilisons la formule (12.828) avec $F_k h_k$ en guise de f :

$$\begin{aligned}\nabla\times F_k e_k &= \nabla\times(F_k h_k\nabla q_k)\\ &= F_k h_k\underbrace{\nabla\times(\nabla q_k)}_{=0}+\nabla(F_k h_k)\times\nabla q_k\\ &= \frac{1}{h_k}\nabla(F_k h_k)\times e_k.\end{aligned} \tag{22.131}$$

Nous utilisons à présent la formule (22.99) du gradient et le formule $e_j\times e_k=\sum_l\epsilon_{jkl}e_l$:

$$\begin{aligned}\nabla\times(F_k e_k) &= \sum_j\frac{1}{h_j h_k}\frac{\partial}{\partial q_j}(F_k h_k)e_j\times e_k\\ &= \sum_{jl}\frac{1}{h_j h_k}\epsilon_{jkl}\frac{\partial}{\partial q_j}(F_k h_k)e_l.\end{aligned} \tag{22.132}$$

Le rotationnel s'écrit donc

$$\nabla\times F=\sum_{jkl}\frac{1}{h_j h_k}\epsilon_{jkl}\frac{\partial}{\partial q_j}(F_k h_k)e_l. \tag{22.133}$$

Devant e_1 par exemple nous avons seulement les termes $j=2$, $k=3$ et $j=3$, $k=2$. Étant donné que $\epsilon_{231}=1$ et $\epsilon_{321}=-1$, le coefficient de e_1 sera simplement

$$\frac{1}{h_2 h_3}\left(\frac{\partial}{\partial q_2}(F_3 h_3)-\frac{\partial}{\partial q_3}(F_2 h_2)\right). \tag{22.134}$$

La formule complète devient

$$\begin{aligned}\nabla\times\sum_k F_k e_k = &\frac{1}{h_2 h_3}\left(\frac{\partial}{\partial q_2}(F_3 h_3)-\frac{\partial}{\partial q_3}(F_2 h_2)\right)\\ &+\frac{1}{h_1 h_3}\left(\frac{\partial}{\partial q_3}(F_1 h_1)-\frac{\partial}{\partial q_1}(F_3 h_3)\right)\\ &+\frac{1}{h_2 h_1}\left(\frac{\partial}{\partial q_1}(F_2 h_2)-\frac{\partial}{\partial q_2}(F_1 h_1)\right).\end{aligned} \tag{22.135}$$

22.5.9.1 Coordonnées cylindriques

En utilisant $h_r = 1$, $h_\theta = r$ et $h_z = 1$, nous trouvons

$$\begin{aligned}\nabla \times (F_r e_r + F_\theta e_\theta + F_z e_z) = &\frac{1}{r}\left(\frac{\partial F_z}{\partial \theta} - \frac{\partial (F_\theta r)}{\partial z}\right) e_r \\ &+ \left(\frac{\partial F_r}{\partial z} - \frac{\partial F_z}{\partial r}\right) e_\theta \\ &+ \left(\frac{\partial (F_\theta r)}{\partial r} - \frac{\partial F_r}{\partial \theta}\right) e_z.\end{aligned} \tag{22.136}$$

22.5.9.2 Coordonnées sphériques

En utilisant $h_\rho = 1$, $h_\theta = \rho$ et $h_\varphi = \rho \sin\theta$, nous trouvons

$$\begin{aligned}\nabla \times (F_\rho e_\rho + F_\theta e_\theta + F_\varphi e_\varphi) = &\frac{1}{\rho \sin\theta}\left(\frac{\partial (F_\varphi)\sin\theta}{\partial \theta} - \frac{\partial F_\theta}{\partial \varphi}\right) e_\rho \\ &+ \frac{1}{\rho \sin\theta}\left(\frac{\partial F_\rho}{\partial \varphi} - \frac{\partial (F_\varphi \rho \sin\theta)}{\partial \rho}\right) e_\theta \\ &+ \frac{1}{\rho}\left(\frac{\partial F_\theta \rho}{\partial \rho} - \frac{\partial F_r}{\partial \theta}\right) e_\varphi.\end{aligned} \tag{22.137}$$

Note : dans le premier terme, il y a une simplification par ρ.

22.6 Les formules

22.6.1 Coordonnées polaires

Les vecteurs de base :

$$e_r = \begin{pmatrix}\cos(\theta) \\ \sin(\theta)\end{pmatrix} = \cos(\theta) e_x + \sin(\theta) e_y \tag{22.138a}$$

$$e_\theta = \begin{pmatrix}-\sin(\theta) \\ \cos(\theta)\end{pmatrix} = -\sin(\theta) e_x + \cos(\theta) e_y. \tag{22.138b}$$

Le gradient :

$$\nabla \tilde{f}(r,\theta) = \frac{\partial \tilde{f}}{\partial r}(r,\theta) e_r + \frac{1}{r}\frac{\partial \tilde{f}}{\partial \theta}(r,\theta) e_\theta. \tag{22.139}$$

La divergence :

$$\nabla \cdot F = \frac{1}{r}\left[\frac{\partial}{\partial r}(r F_r) + \frac{\partial}{\partial \theta}(F_\theta)\right]. \tag{22.140}$$

Le rotationnel :

$$\nabla \times (F_r e_r + F_\theta e_\theta) = \left(\frac{\partial (F_\theta r)}{\partial r} - \frac{\partial F_r}{\partial \theta}\right) e_z. \tag{22.141}$$

Notons que le rotationnel n'existe pas vraiment en deux dimensions. Ici nous avons vu le champ $F(r,\theta)$ comme un champs dans $\mathbb{R}^3$ ne dépendant pas de z et n'ayant pas de composante z. Le résultat est un rotationnel qui est dirigé selon l'axe z.

22.6.2 Coordonnées cylindriques

Les vecteurs de base : idem qu'en coordonnées polaires, et on ajoute e_z sans modifications.
Le gradient :

$$\nabla \tilde{f}(r,\theta,z) = \frac{\partial \tilde{f}}{\partial r}(r,\theta,z) e_r + \frac{1}{r}\frac{\partial \tilde{f}}{\partial \theta}(r,\theta,z) e_\theta + \frac{\partial \tilde{f}}{\partial z}(r,\theta,z) e_z. \tag{22.142}$$

La divergence :

$$\nabla \cdot F = \frac{1}{r}\left[\frac{\partial}{\partial r}(rF_r) + \frac{\partial}{\partial \theta}(F_\theta) + \frac{\partial}{\partial z}(rF_z)\right]. \tag{22.143}$$

Le rotationnel :

$$\begin{aligned}\nabla \times (F_r e_r + F_\theta e_\theta + F_z e_z) = &\frac{1}{r}\left(\frac{\partial F_z}{\partial \theta} - \frac{\partial (F_\theta r)}{\partial z}\right) e_r \\ &+ \left(\frac{\partial F_r}{\partial z} - \frac{\partial F_z}{\partial r}\right) e_\theta \\ &+ \left(\frac{\partial (F_\theta r)}{\partial r} - \frac{\partial F_r}{\partial \theta}\right) e_z.\end{aligned} \tag{22.144}$$

Note : les formules concernant les coordonnées polaires se réduisent de celles-ci en enlevant toutes les références à z.

22.6.3 Coordonnées sphériques

Les vecteurs de base :

$$\begin{aligned} e_r &= \begin{pmatrix}\sin(\theta)\cos(\varphi)\\ \sin(\theta)\sin(\varphi)\\ \cos(\theta)\end{pmatrix}, \quad e_\theta = \begin{pmatrix}\cos(\theta)\cos(\varphi)\\ \cos(\theta)\sin(\varphi)\\ -\sin(\theta)\end{pmatrix}, \\ e_\varphi &= \begin{pmatrix}-\sin(\varphi)\\ \cos(\varphi)\\ 0\end{pmatrix}\end{aligned} \tag{22.145}$$

Le gradient :

$$\nabla \tilde{f}(\rho, \theta, \varphi) = \frac{\partial \tilde{f}}{\partial \rho} e_\rho + \frac{1}{\rho}\frac{\partial \tilde{f}}{\partial \theta} e_\theta + \frac{1}{\rho \sin(\theta)}\frac{\partial \tilde{f}}{\partial \varphi} r_\varphi. \tag{22.146}$$

La divergence :

$$\nabla \cdot F = \frac{1}{\rho^2 \sin\theta}\left[\frac{\partial}{\partial \rho}(\rho^2 \sin\theta F_\rho) + \frac{\partial}{\partial \theta}(\rho \sin\theta F_\theta) + \frac{\partial}{\partial \varphi}(\rho F_\varphi)\right]. \tag{22.147}$$

Le rotationnel :

$$\begin{aligned}\nabla \times (F_\rho e_\rho + F_\theta e_\theta + F_\varphi e_\varphi) = &\frac{1}{\rho \sin\theta}\left(\frac{\partial (F_\varphi)\sin\theta}{\partial \theta} - \frac{\partial F_\theta}{\partial \varphi}\right) e_\rho \\ &+ \frac{1}{\rho \sin\theta}\left(\frac{\partial F_\rho}{\partial \varphi} - \frac{\partial (F_\varphi \rho \sin\theta)}{\partial \rho}\right) e_\theta \\ &+ \frac{1}{\rho}\left(\frac{\partial F_\theta \rho}{\partial \rho} - \frac{\partial F_r}{\partial \theta}\right) e_\varphi.\end{aligned} \tag{22.148}$$

Chapitre 23

Espaces de Hilbert

23.1 Espaces de Hilbert

Définition 23.1.
*Un **espace de Banach** est un espace vectoriel normé complet pour la topologie de la norme.*

Définition 23.2.
*Un espace vectoriel muni d'un produit scalaire est une espace **préhilbertien**. Si il est complet*[1] *pour la norme induite par le produit scalaire alors il est de **Hilbert**.*

Dans les deux cas nous considérons la topologie métrique dérivant du produit scalaire.

La différence entre un espace de Hilbert et un espace de Banach est que dans le cas d'un espace de Hilbert, nous demandons que la norme dérive d'un produit scalaire.

Dans le cas des espaces de dimension finie, le fait d'être complet est automatique, comme le montre la proposition suivante.

Proposition 23.3.
Soit $\big(E, \|.\|\big)$ un espace vectoriel normé de dimension finie sur un corps $\mathbb{K}$ qui est complet (typiquement $\mathbb{K}$ est $\mathbb{R}$ ou $\mathbb{C}$). Alors E est complet.

Démonstration. Nous considérons une suite de Cauchy (f_n) dans E et si $\{e_\alpha\}$ est une base orthonormée de E nous définissons les coefficients $f_n = \sum_\alpha a_{n\alpha}e_\alpha$. La somme sur α est finie par hypothèse sur la dimension de E.

Nous avons

$$\|f_n - f_m\| = \|\sum_\alpha (a_{n\alpha} - a_{m\alpha})e_\alpha\| = \sum_\alpha |a_{n\alpha} - a_{m\alpha}|^2. \tag{23.1}$$

Pour tout ϵ, il existe N tel que si $m, n > N$ alors $|a_{n\alpha} - a_{m\alpha}| < \sqrt{\epsilon}$. Autrement dit, pour chaque α, la suite $(a_{n\alpha})_{\alpha\in\mathbb{N}}$ est de Cauchy dans $\mathbb{K}$ et converge donc dans $\mathbb{K}$. Soit a_α la limite et définissons $f = \sum_\alpha a_\alpha e_\alpha$. Nous avons alors

$$\|f_n - f\| = \|\sum_\alpha (a_{n\alpha} - a_\alpha)e_\alpha\|, \tag{23.2}$$

dont la limite $n \to \infty$ est bien zéro. Donc la suite (f_n) converge vers $f \in E$. L'espace E est alors complet. □

Proposition 23.4.
Si $\mathscr{H}$ est un espace de Hilbert réel, alors

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle. \tag{23.3}$$

Si $\mathscr{H}$ est un espace de Hilbert complexe alors

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\,\mathrm{Re}\langle x, y\rangle. \tag{23.4}$$

1. Définition 6.38.

Dans les deux cas nous avons l'inégalité de Cauchy-Schwarz :

$$|\langle x, y\rangle| \leqslant \|x\|\|y\|. \tag{23.5}$$

Dans un espace vectoriel de dimension infinie, tous les opérateurs linéaires ne sont pas continus.

Exemple 23.5
Soit un espace vectoriel V engendré par la base $\{e_k\}_{k\in\mathbb{N}}$ et l'application linéaire $T\colon V \to V$ donnée par

$$Te_k = ke_k. \tag{23.6}$$

Nous allons montrer que l'image inverse de la boule unité ouverte $\mathcal{O}$ n'est pas ouverte. En effet si (ϵ_k) est une suite de réels strictement positifs tendant vers zéro, les vecteurs

$$a_k = \left(\frac{1}{k} + \epsilon_k\right) e_k \tag{23.7}$$

sont hors de $T^{-1}\mathcal{O}$ parce que

$$Ta_k = (1 + k\epsilon_k)e_k. \tag{23.8}$$

Mais la suite (a_k) converge vers 0 qui fait partie de $T^{-1}\mathcal{O}$. Donc le complémentaire de $T^{-1}\mathcal{O}$ n'est pas fermé, ce qui prouve que $T^{-1}\mathcal{O}$ n'est pas ouvert. △

23.1.1 Sous-espace vectoriel fermé ? ? ?

Nous verrons que beaucoup de résultats demandent un sous-espace vectoriel *fermé*. Une question légitime est : est-ce qu'il existe des sous-espaces vectoriels qui ne soient pas fermés ? En dimension finie, tous les sous-espaces vectoriels sont fermés, mais cela n'est pas vrai en dimension infinie.

Soit en effet une partie libre infinie $A = \{v_i\}_{i\in\mathbb{N}}$ dans un espace de Hilbert $\mathscr{H}$. L'ensemble $\mathrm{Span}(A)$ des combinaisons linéaires d'éléments de A est un sous-espace vectoriel de $\mathscr{H}$, mais il n'est pas fermé.

En effet, supposons pour simplifier les notations que $\|v_i\| = 1$ pour tout i. Alors nous considérons la combinaison

$$a = \sum_{k=1}^{\infty} \alpha_k v_k \tag{23.9}$$

où les α_k sont suffisamment décroissants pour assurer les convergences [2]. Le vecteur a n'est pas dans $\mathrm{Span}(A)$, mais la suite $a_n = \sum_{k=1}^{n} \alpha_k v_k$ est dans $\mathrm{Span}(A)$ et converge vers a : $a_n \xrightarrow{\mathscr{H}} a$. En effet,

$$\|a_n - a\| = \|\sum_{k=n+1}^{\infty} \alpha_k v_k\| \leqslant \sum_{k=n+1}^{\infty} |\alpha_k| \xrightarrow{n\to\infty} 0. \tag{23.10}$$

Vous voulez des détails sur la dernière limite ? Vu que la somme $\sum_k |\alpha_k|$ converge, la suite des sommes de queues de suites [3] converge vers zéro :

$$\lim_{n\to\infty} \sum_{k=n+1}^{\infty} |\alpha_k| = 0. \tag{23.11}$$

23.2 Théorème de la projection

Théorème 23.6 (Projection sur partie fermée convexe[265, 266])**.**
Soit $\mathscr{H}$ un espace de Hilbert, $x \in \mathscr{H}$, et C un sous ensemble fermé convexe de $\mathscr{H}$.

(1) Les deux conditions suivantes sur $y \in \mathscr{H}$ sont équivalentes :

2. Par exemple $\alpha_k = 1/k^2$ si les v_k sont orthonormaux.
3. On se comprend hein.

(a) $\|x-y\| = \inf\{\|x-z\| \text{ tel que } z \in C\}$,

(b) pour tout $z \in C$, $\text{Re}\langle x-y, z-y\rangle \leqslant 0$.

(2) Il existe un unique $y \in \mathscr{H}$*, 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.

Nous nommons d l'infimum en question de la première condition.

Existence Soit (y_n) une suite dans C telle que

$$\lim_{n\to\infty} \|x-y_n\| = \inf\{\|x-y\| \text{ tel que } z \in C\} = d. \tag{23.12}$$

Nous allons montrer que cette suite peut être choisie de Cauchy. Elle convergera donc dans $\mathscr{H}$ parce que ce dernier est complet. Mais C étant supposé fermé dans $\mathscr{H}$, la limite appartiendra à C. Soient $r, s \in \mathbb{N}$. D'abord nous avons

$$\|y_r - y_s\|^2 = \langle y_r - y_s + x - x, y_r - y_s + x - x\rangle \tag{23.13a}$$
$$= \|y_r - x\|^2 + \|y_s - x\|^2 - 2\langle y_r - x, y_s - x\rangle. \tag{23.13b}$$

Ensuite,

$$4\left\|\frac{y_r+y_s}{2} - x\right\|^2 = \langle y_r + y_s - 2x, y_r - y_s - 2x\rangle \tag{23.14a}$$
$$= \|y_r - x\|^2 + \|y_s - x\|^2 + 2\langle y_r - x, y_s - x\rangle. \tag{23.14b}$$

Si nous égalisons les valeurs de $2\langle y_r - x, y_s - x\rangle$ nous trouvons

$$\|y_r - y_s\|^2 = -4\left\|\frac{y_r+y_s}{2} - x\right\|^2 + 2\|y_r - x\|^2 + 2\|y_s - x\|^2. \tag{23.15}$$

La distance infimum étant d, nous pouvons choisir y_n de telle façon à avoir

$$\|y_n - x\| \leqslant d + \frac{1}{n}. \tag{23.16}$$

D'autre part étant donné que C est convexe, $(y_r + y_s)/2$ est dans C et nous avons

$$\left\|\frac{y_r+y_s}{2} - x\right\| \leqslant d. \tag{23.17}$$

En mettant ces majorations dans (23.15) nous trouvons

$$\|y_r - y_s\|^2 \leqslant -4d + 2\left(d + \frac{1}{r}\right) + 2\left(d + \frac{1}{s}\right) = \frac{1}{r} + \frac{1}{s}. \tag{23.18}$$

La suite (y_n) est donc de Cauchy et la limite est un élément de C. Prouvons que cet élément y réalise l'infimum. Pour cela nous avons les inégalités

$$d \leqslant \|x-y\| \leqslant \|x-y_n\| + \|y_n - y\|. \tag{23.19}$$

En prenant le limite $n \to \infty$ nous trouvons

$$d \leqslant \|x-y\| \leqslant d. \tag{23.20}$$

Unicité Même preuve que pour le théorème en dimension finie 12.233.

(1)a⇒ (1)b Même preuve que pour le théorème en dimension finie 12.233.

(1)b⇒ (1)a Même preuve que pour le théorème en dimension finie 12.233.

□

Proposition 23.7.
Soit C une partie convexe et fermée de l'espace de Hilbert $\mathscr{H}$. Alors pour tout $x_1, x_2 \in \mathscr{H}$ nous avons

$$\| \texttt{proj}_C(x_1) - \texttt{proj}_C(x_2)\| \leqslant \|x_1 - x_2\|. \tag{23.21}$$

En particulier la projection est une application continue.

Démonstration. Nous posons $y_1 = \texttt{proj}_C(x_1)$ et $y_2 = \texttt{proj}_C(x_2)$. Par la partie (1)b du théorème 23.6, nous avons, pour tout $z, z' \in C$ les inégalités

$$\mathrm{Re}\langle x_1 - y_1, z - y_1\rangle \leqslant 0 \tag{23.22a}$$
$$\mathrm{Re}\langle x_2 - y_2, z' - y_2\rangle \leqslant 0. \tag{23.22b}$$

En prenant $z = y_2$ et $z' = y_1$ et en sommant nous trouvons

$$\mathrm{Re}\langle (x_1 - y_1) + (y_2 - x_2), y_2 - y_1\rangle \leqslant 0. \tag{23.23}$$

Nous pouvons maintenant calculer

$$\begin{aligned}
\|y_1 - y_2\|^2 &= \mathrm{Re}\, \|y_1 - y_2\| \\
&= \mathrm{Re}\langle y_1 - y_2, (y_1 - x_1) + x_1 - x_2 + (x_2 - y_2)\rangle \\
&= \mathrm{Re}\langle y_1 - y_2, x_1 - x_2\rangle + \mathrm{Re}\langle y_2 - y_1, (x_1 - y_1) + (y_2 - x_2)\rangle \\
&\leqslant \langle y_1 - y_2, x_1 - x_2\rangle \\
&\leqslant \big\|\langle y_1 - y_2, x_1 - x_2\rangle\big\| \\
&=\leqslant \|y_1 - y_2\|\|x_1 - x_2\|.
\end{aligned} \tag{23.24}$$

En simplifiant par $\|y_1 - y_2\|$ [4] nous trouvons le résultat

$$\|y_1 - y_2\| \leqslant \|x_1 - x_2\|. \tag{23.25}$$

□

Théorème 23.8 (Projection orthogonale)**.**
Soit $\mathscr{H}$ un espace de Hilbert et K, un sous-espace vectoriel fermé non réduit à $\{0\}$ et $x \in \mathscr{H}$. L'élément $y = \texttt{proj}_K(x)$ est l'unique élément de K tel que

$$x - y \in K^\perp. \tag{23.26}$$

De plus l'application $x \mapsto \texttt{proj}_K(x)$ est linéaire, continue et de norme 1.

L'élément y ainsi définit est la **projection orthogonale** de x sur K et sera noté $\texttt{proj}_K(f)$.

Démonstration. La continuité de la projection est donné par la proposition 23.7.

Soit z un élément de K tel que $\langle z - x, a\rangle = 0$ pour tout $a \in K$. Nous avons

$$\|x - a\|^2 = \|z - x\|^2 + \|a - z\|^2 + 2\underbrace{\langle z - x, a - z\rangle}_{=0} \tag{23.27a}$$
$$\geqslant \|z - x\|^2. \tag{23.27b}$$

Le produit scalaire est nul parce que $a - z \in K$. La distance $\|z - x\|$ est donc bien la plus petite distance entre x et les éléments de K.

Dans l'autre sens, nous supposons que $y \in K$ minimise la distance à x dans K. Par hypothèse pour tout a et pour tout $\lambda \in \mathbb{R}$, la différence

$$\|(y + \lambda a) - x\|^2 - \|y - x\|^2 \tag{23.28}$$

4. Si c'est nul, alors la preuve est évidente.

est positive. En développant les produits scalaires nous trouvons la conditions suivante

$$\lambda^2\|a\|^2 + 2\lambda\langle a, y-x\rangle \geqslant 0 \tag{23.29}$$

qui doit être vraie pour tout $\lambda \in \mathbb{R}$. En tant que polynôme du second degré en λ, cela n'aura pas deux racines réelles distinctes uniquement si $\langle a, y-x\rangle = 0$.

Nous montrons maintenant la linéarité de la projection orthogonale. Soient $x_1, x_2 \in \mathscr{H}$. L'élément $y = \mathtt{proj}_K x_1 + \mathtt{proj}_K x_2$ satisfait à la condition d'orthogonalité : pour tout $z \in K$,

$$\langle x_1 + x_2 - \mathtt{proj}_K x_1 - \mathtt{proj}_K x_2, z\rangle = \langle x_1 - \mathtt{proj}_K x_1, z\rangle + \langle x_2 - \mathtt{proj}_K x_2, z\rangle = 0. \tag{23.30}$$

Étant donné que K est un sous-espace vectoriel, la condition de minimalité est automatiquement vérifiée (seconde partie du théorème 23.6).

En ce qui concerne la norme opérateur de $\mathtt{proj}_K$, la décomposition de $x \in \mathscr{H}$ en composantes dans K et $K^\perp$ est

$$x = x + (x - \mathtt{proj}_K x). \tag{23.31}$$

Étant deux parties orthogonales nous avons

$$\|\mathtt{proj}_K x\|^2 = \|x\|^2 - \|x - \mathtt{proj}_K x\|^2. \tag{23.32}$$

En prenant $\|x\| = 1$ nous trouvons $\|\mathtt{proj}_K x\|^2 \leqslant 1$ et par conséquent $\|\mathtt{proj}_K\| \leqslant 1$. Mais d'autre part en prenant $x \in K$ nous avons automatiquement $\|\mathtt{proj}_K\| \geqslant 1$. □

Proposition 23.9.
Soit $\mathscr{H} = L^2(\Omega, \mathcal{A}, \mu)$, $\mathcal{F}$ une sous tribu de $\mathcal{A}$ et K l'ensemble de fonctions $\mathcal{F}$-mesurables dans $L^2(\Omega, \mathcal{A}, \mu)$. Si $f \in L^2(\Omega, \mathcal{A}, \mu)$ est positive, alors $\mathtt{proj}_K f$ est positive (presque partout).

Démonstration. L'ensemble $A = \{\mathtt{proj}_K f < 0\}$ est dans $\mathcal{F}$. En effet

$$A = (\mathtt{proj}_K f)^{-1}\big(]-\infty, 0[\big) \tag{23.33}$$

alors que, par construction, $\mathtt{proj}_K f$ est $\mathcal{F}$-mesurable. La fonction indicatrice $\mathbb{1}_A$ est alors $\mathcal{F}$-mesurable (c'est à dire $1_A \in K$) et nous avons

$$0 \leqslant \int_\Omega f\mathbb{1}_A = \int_\Omega \mathtt{proj}_K f\,\mathbb{1}_A \leqslant 0. \tag{23.34}$$

Étant donné que nous avons supposé $f \geqslant 0$ nous avons alors $\mu(A) = 0$. D'où le fait que $\mathtt{proj}_K f$ est presque partout positive. □

23.3 Systèmes orthogonaux et bases

Dans cette partie nous noterons $\mathbb{K}$ le corps de base de l'espace $\mathscr{H}$. Seuls deux cas sont envisagés : $\mathbb{K} = \mathbb{C}$ ou $\mathbb{K} = \mathbb{R}$.

Pour chaque $x \in \mathscr{H}$ nous considérons l'application

$$\begin{aligned}\Phi_x\colon \mathscr{H} &\to \mathbb{K}\\ y &\mapsto \langle x, y\rangle.\end{aligned} \tag{23.35}$$

Ce sont des applications continues.

23.3.1 Orthogonal d'une partie

Définition 23.10.
*Soit une partie A de l'espace de Hilbert $\mathscr{H}$. L'**orthogonal** de A est l'ensemble*

$$A^\perp = \{v \in \mathscr{H} \text{ tel que } \langle v, x\rangle = 0 \forall x \in A\}. \tag{23.36}$$

Proposition 23.11.
Si $\mathscr{H}$ est une préhilbert et si $A \subset \mathscr{H}$, alors l'ensemble $A^\perp$ est un sous-espace fermé de $\mathscr{H}$.

Démonstration. L'application Φ_x définie en (23.35) est continue pour chaque $x \in \mathscr{H}$ et par conséquent l'ensemble $\ker \Phi_x = \Phi_x^{-1}(\{0\})$ est fermé. L'ensemble

$$A^\perp = \bigcap_{x\in A} \ker \Phi_x \tag{23.37}$$

est donc fermé. □

Proposition 23.12 (wikipédia)**.**
Soit V un espace vectoriel normé et une décomposition en somme directe $V = F \oplus G$. Alors les trois conditions suivantes sont équivalentes.

(1) L'isomorphisme naturel $F \times G \to v$ est un homéomorphisme.

(2) F et G sont fermés et la restriction à G de la projection $V \to V/F$ est un homéomorphisme.

(3) F et G sont fermés et la projection de E sur F est continue.

Lorsqu'une décomposition en somme directe vérifie la proposition 23.12, nous disons que la décomposition est **topologique**.

Théorème 23.13.
Soit $\mathscr{H}$ un espace de Hilbert et F un sous-espace fermé de $\mathscr{H}$. Alors

(1) Nous avons la décomposition

$$\mathscr{H} = F \oplus F^\perp. \tag{23.38}$$

(2) La projection sur F par rapport à la somme directe (23.38) *est la projection $\texttt{proj}_F$ du théorème de projection.*

(3) La décomposition (23.38) *est topologique.*

Proposition 23.14.
Si F est un sous-espace de l'espace de Hilbert $\mathscr{H}$ alors on a $F^{\perp\perp} = \bar{F}$.

Démonstration. Nous savons par la proposition 23.11 que $F^\perp$ est fermé, par conséquent le théorème 23.13 donne la somme directe

$$\mathscr{H} = F^\perp \oplus F^{\perp\perp}. \tag{23.39}$$

Mais $\bar{F}$ étant également fermé nous avons la somme directe

$$\mathscr{H} = \bar{F} \oplus (\bar{F})^\perp. \tag{23.40}$$

Montrons que $(\bar{F})^\perp = F^\perp$. En effet si $x \in F^\perp$ et si $y \in \bar{F}$, alors il existe une suite y_n dans F qui converge vers y. Pour chaque n nous avons $\langle x, y_n\rangle = 0$ et dons $\langle x, y\rangle = 0$ par continuité du produit scalaire.

Nous avons donc

$$\mathscr{H} = F^\perp \oplus F^{\perp\perp} = \bar{F} \oplus F^\perp. \tag{23.41}$$

Mais $\bar{F} \subset F^{\perp\perp}$ (prendre une suite). Les espaces $\bar{F}$ et $F^{\perp\perp}$ étant tous deux des supplémentaires de $F^\perp$, nous déduisons qu'ils doivent être égaux. □

Proposition 23.15.
Soit $\mathscr{H}$ un espace de Hilbert et F un sous-espace vectoriel de $\mathscr{H}$. Alors F est dense si et seulement si $F^\perp = \{0\}$.

Démonstration. Nous savons que
$$\mathscr{H} = F^\perp \oplus \bar{F}. \tag{23.42}$$
Donc nous avons $H = \bar{F}$ si et seulement si $F^\perp = \{0\}$. □

Pour vérifier si un sous-espace vectoriel d'un espace de Hilbert est dense, il suffit donc de montrer que son orthogonal est réduit à zéro.

23.3.2 Dual, théorème de représentation de Riesz

Lemme 23.16.
L'application $\Phi\colon \mathscr{H} \to \mathcal{L}(\mathscr{H}, \mathbb{K})$ donnée par $y \mapsto \Phi_y$ est une isométrie : nous avons $\|\Phi_y\| = \|y\|$. De plus pour chaque y, l'application Φ_y est continue.

Démonstration. En utilisant la définition de la norme opérateur et l'inégalité de Cauchy-Schwarz,
$$\|\Phi_y\| = \sup_{\|x\|=1} \|\Phi_y(x)\| = \sup |\langle x, y\rangle| \leqslant \sup \|x\|\|y\| = \|y\|. \tag{23.43}$$
Par conséquent $\|\Phi_y\| \leqslant \|y\|$. Mais d'autre part le fait que $\Phi_y(y) = \|y\|^2$ montre que $\|\Phi_y\| \geqslant \|y\|$.

En ce qui concerne la continuité de Φ_y, elle est garantie par le fait que c'est une application linéaire bornée via la proposition 10.18. □

Définition 23.17.
*Le **dual** de l'espace de Hilbert $\mathscr{H}$ est l'ensemble*
$$\mathscr{H}' = \{f\colon \mathscr{H} \to \mathbb{K} \text{ linéaire et continue}\}. \tag{23.44}$$

Notons que dans le contexte des espaces de Hilbert nous demandons la continuité des éléments du dual parce qu'elle n'est pas automatique par la linéarité dans les cas de dimension infinie. En principe nous devrions préciser dual *topologique*, mais nous ne le ferons pas systématiquement lorsque le contexte parle clairement de topologie (ce qui est le cas lorsqu'on parle d'espaces de Hilbert). De temps en temps le dual *algébrique* d'un espace est noté E^* ; dans ce cas la continuité n'est pas demandée.

Théorème 23.18 (Théorème de représentation de Riesz)**.**
L'application
$$\begin{aligned} \Phi\colon \mathscr{H} &\to \mathscr{H}' \\ y &\mapsto \Phi_y \end{aligned} \tag{23.45}$$
est une bijection isométrique.

Démonstration. Nous savons du lemme 23.16 que Φ est une isométrique. Nous devons seulement montrer que Φ est surjective. Soit $f \in \mathscr{H}'$. Par continuité nous savons que $F = \ker(f)$ est fermé, donc
$$\mathscr{H} = \ker(f) \oplus (\ker f)^\perp \tag{23.46}$$
par le théorème 23.13. Nous considérons une base de H adaptée à cette décomposition, c'est à dire
— $(u_s)_{s\in S}$ une base de $\ker f$,
— $(v_t)_{t\in T}$ une base de $(\ker f)^\perp$.
Il se fait que T se réduit à un seul élément parce que si $v \in (\ker f)^\perp$, alors $f(v) \neq 0$, et par conséquent $\{f(v)\}$ est déjà une base de $\text{Image } f = \mathbb{K}$. Le théorème du rang (théorème 7.23) assure alors que
$$\{v\} \cup \{u_s\}_{s\in S} \tag{23.47}$$

est une base de $\mathscr{H}$ avec $v \in (\ker f)^\perp$ et $u_s \in \ker f$. Nous choisissons v pour avoir $\|v\| = 1$.

Nous pouvons maintenant prouver l'existence de y tel que $f_y = f$. En effet si nous posons $y = \overline{f(v)}v$ nous avons $f_y = f$, en effet

$$f_y(v) = \langle v, y\rangle = f(v)\langle v, v\rangle = f(v) \tag{23.48a}$$
$$f_y(u_s) = \langle u_s, y\rangle = 0. \tag{23.48b}$$

Par conséquent f_y et f coïncident sur une base de $\mathscr{H}$.

En ce qui concerne l'unicité, d'abord si $f_y = f$ alors nous devons avoir $y \in (\ker f)^\perp$ et par conséquent $y = \lambda v$ pour un certain $\lambda \in \mathbb{K}$. Nous avons alors

$$f_y(v) = \bar{\lambda}\langle v, v\rangle = \bar{\lambda}. \tag{23.49}$$

Pour que cela soit égal à $f(v)$, nous fixons $\lambda = \overline{f(v)}$.

□

Notons que nous avons réellement utilisé le théorème du rang pour l'unicité. Si nous ne demandions pas l'unicité, alors nous n'avions pas besoin du fait que $\dim(\ker f)^\perp = 1$, et nous aurons donc pu parler de formes plus générales à valeurs dans $\mathbb{K}^n$.

23.3.3 Séparabilité

Définition 23.19.
Un espace topologique est ***séparable*** *s'il possède une partie dénombrable dense.*

Définition 23.20 ([267])**.**
Si E est un espace vectoriel normé nous disons que Δ est une partie ***totale*** *de E si* $\mathrm{Span}(\Delta)$ *est dense dans E. Attention : nous rappelons que* $\mathrm{Span}(\Delta)$ *est l'ensemble des combinaisons linéaires* finies *d'éléments de Δ.*

Proposition 23.21.
Un espace vectoriel normé est séparable si et seulement s'il possède une partie totale dénombrable.

Démonstration. Si Δ est une partie dénombrable de E, alors le $\mathbb{Q}$-espace vectoriel $\mathrm{Span}_{\mathbb{Q}}\Delta$ est dénombrable, et sa fermeture est la même que celle de $\mathrm{Span}_{\mathbb{K}}\Delta$. □

Définition 23.22.
Une famille $(u_i)_{i\in I}$ d'éléments de $\mathscr{H}$ indicée par un ensemble quelconque I est un ***système orthonormé*** *si*

(1) $\|u_i\| = 1$ pour tout $i \in I$,

(2) $u_i \perp u_j$ pour tout $i \neq j$.

Notons que si $(u_n)_{n\in\mathbb{N}}$ est un système orthonormé dénombrable alors en utilisant les formules de la proposition 23.4, nous avons

$$\Big\| \sum_{k=1}^{n} \xi_k u_k \Big\|^2 = \sum_{k=1}^{n} |\xi_k|^2. \tag{23.50}$$

Exemple 23.23
Dans l'ensemble $L^2(a, b)$ avec $b - a = L$ l'ensemble des fonctions

$$e_n(t) = \frac{1}{\sqrt{b-a}} \exp(\frac{2\pi}{L} int) \tag{23.51}$$

avec $n \in \mathbb{Z}$ forme un système orthonormé. En effet

$$\langle e_n, e_n\rangle = \frac{1}{b-a}\int_a^b \exp(\frac{2\pi}{L} int)\exp(-\frac{2\pi}{L} int) = 1 \tag{23.52}$$

et

$$\langle e_n, e_m \rangle = \frac{1}{b-a} \int_a^b e^{\frac{2\pi}{L}(n-m)t} dt = 0. \tag{23.53}$$

Dans cette intégrale nous utilisons le fait que $b = a + (b-a)$ pour simplifier les expressions en cours de calcul.

La famille (23.51) est le **système trigonométrique** de $L^2(a,b)$. On en parle aussi dans [220]. △

Proposition 23.24.
Une partie orthonormée est libre.

Démonstration. Soit $(u_i)_{i\in I}$ une famille orthonormée de $\mathscr{H}$ et une combinaison linéaire *finie* nulle :

$$\sum_{k=1}^{n} a_k u_k = 0. \tag{23.54}$$

Nous développons la somme en utilisant les formules de la proposition 23.4 :

$$0 = \| \sum_k a_k u_k \|^2 = \sum_k |a_k|^2 \|u_k\|^2 + 2 \sum_{k<l} \langle a_k u_k, a_l u_l \rangle. \tag{23.55}$$

La famille étant orthonormée, les choses se simplifient en

$$\sum_k |a_k|^2 = 0, \tag{23.56}$$

ce qui signifie que $a_k = 0$ pour tout k. □

Avant de continuer nous devons définir comment nous calculons des sommes sur des ensembles quelconques. Si I est un ensemble et si pour chaque $i \in I$ nous avons un nombre réel positif a_i, alors nous définissons

$$\sum_{i\in I} a_i = \sup_{\substack{J\subset I \\ J \text{ fini}}} \sum_{j\in J} a_j. \tag{23.57}$$

Cela est discuté dans la section 9.10.

Proposition 23.25 (Inégalités de Bessel)**.**
Soit $\mathscr{H}$ un préhilbert. Si $(u_i)_{i\in I}$ est un système orthonormé et si $x \in \mathscr{H}$, alors

$$\sum_{i\in I} \big|\langle x, u_i \rangle\big|^2 \leqslant \|x\|^2. \tag{23.58}$$

Démonstration. Les éléments de la somme étant des réels positifs, la notion de somme à utiliser est celle de la définition 9.178.

Posons $c_i(x) = \langle x, u_i \rangle$. Pour toute partie finie $J \subset I$ nous avons

$$0 \leqslant \big\| x - \sum_{j\in J} c_j(x) u_j \big\|^2 = \|x\|^2 - 2\,\mathrm{Re} \sum_j \langle x, c_j(x) u_j \rangle + \sum_j |c_j(x)|^2. \tag{23.59}$$

Mais en tenant compte du fait que

$$\langle x, c_j(x) u_j \rangle = \overline{c_j(x)} \langle x, u_j \rangle = |c_j(x)|^2, \tag{23.60}$$

nous restons avec

$$\| x - \sum_{j\in J} c_j(x) u_j \| = \|x\|^2 - \sum_{j\in J} |c_j(x)|^2. \tag{23.61}$$

Finalement,

$$\sum_{j\in J} |c_j(x)|^2 \leqslant \|x\|^2. \tag{23.62}$$

Ayant cette inégalité pour toute partie finie de I, nous l'avons encore pour le supremum. □

Proposition 23.26.
Soit $\mathscr{H}$ un préhilbert et une famille orthonormé $(u_i)_{i\in I}$. Si

$$x = \sum_{i\in I} \xi_i u_i \tag{23.63}$$

alors $\xi_i = \langle x, u_i\rangle$.

Démonstration. Nous appliquons l'application Φ_{u_k} du théorème de représentation de Riesz[5] à l'équation (23.63).

Si I est dénombrable, alors permuter Φ_{u_k} avec la somme consiste à invoquer la continuité, et permuter la limite des sommes partielles avec Φ_{u_k} (l'application Φ_{u_k} est continue parce qu'isométrique).

Sinon, il faut utiliser la proposition 9.187. Il faut donc montrer que la famille $\Phi_{u_k}\big(\xi_i u_i\big)$ est sommable. Cela est fort vrai parce que cette famille ne contient en réalité qu'en seul élément non nul, celui avec $i = k$, qui vaut ξ_k. Au final nous avons :

$$\langle x, u_k\rangle = \sum_{i\in I} \Phi_{u_k}(\xi_i u_i) = \xi_k. \tag{23.64}$$

□

23.3.4 Base hilbertienne

Définition 23.27.
*Une **base orthonormée** est une famille dénombrable orthonormé et totale*[6]*. Cela sera souvent aussi appelé une **base hilbertienne**.*

23.28.
Cette définition demande quelque remarques.

(1) La notion de base hilbertienne pas la même notion de base qu'en algèbre. En effet pour avoir une base algébrique d'un espace vectoriel, nous demandons que les éléments soient des combinaisons linéaires *finies* des éléments de la base, tandis qu'ici en demandant que la partie soit totale nous demandons simplement que les combinaisons linéaires finies soient denses.

(2) Nous allons voir qu'un espace de Hilbert est généré par les sommes *infinies* de vecteurs d'une base hilbertienne avec des coefficients qui forment une suite dans ℓ^2.

(3) Nous ne demandons pas que la famille soit libre ? La belle affaire. Une famille orthogonale est toujours libre, proposition 23.24.

Lemme 23.29 ([268])**.**
Si $\mathcal{E}$ est une base hilbertienne[7] *de l'espace de Hilbert $\mathscr{H}$ et si $x \in \mathscr{H}$ alors l'ensemble*

$$\{e \in \mathcal{E} \text{ tel que } \langle x, e\rangle \neq 0\} \tag{23.65}$$

est au plus dénombrable.

Démonstration. Notons qu'ici, $\mathscr{H}$ n'est pas supposé séparable. Nous savons par l'inégalité de Bessel (23.58) que

$$\sum_{e\in\mathcal{E}} |\langle x, e\rangle|^2 \leqslant \|x\|^2. \tag{23.66}$$

Donc si $\epsilon > 0$ est donné, l'ensemble $\{e \in \mathcal{E} \text{ tel que } |\langle x, e\rangle| > \epsilon\}$ est fini. Or

$$\{e \in \mathcal{E} \text{ tel que } \langle x, e\rangle \neq 0\} = \{e \in \mathcal{E} \text{ tel que } |\langle x, e\rangle| > 0\} = \bigcup_{n=1}^{\infty} \{e \in \mathcal{E} \text{ tel que } |\langle x, e\rangle| > \frac{1}{n}\}. \tag{23.67}$$

Bref, cet ensemble est une union dénombrable d'ensembles finis. Il est donc dénombrable. □

5. Théorème 23.18.
6. Définition 23.20
7. Définition 23.27.

Remarque 23.30.
Le lemme 23.29 ne signifie pas que la base $\mathcal{E}$ doive être dénombrable. Il signifie seulement que pour chaque x séparément, seule une partie dénombrable de $\mathcal{E}$ est nécessaire.

Corollaire 23.31 ([268])**.**
Si un espace de Hilbert possède une base hilbertienne dénombrable, alors toutes ses bases hilbertiennes sont dénombrables.

Démonstration. Soient $\mathcal{E}$ et $\mathcal{F}$ des bases hilbertiennes de l'espace de Hilbert $\mathscr{H}$, en supposant que $\mathcal{E}$ soit dénombrable. Pour chaque $f \in \mathcal{F}$ nous avons $\langle f, e\rangle \neq 0$ pour au moins un $e \in \mathcal{E}$, sinon en vertu de la décomposition (23.63) de f dans la base $\mathcal{E}$, nous aurions $f = 0$. Nous avons donc

$$\mathcal{F} \subset \bigcup_{e\in\mathcal{E}} \{f \in \mathcal{F} \text{ tel que } \langle f, e\rangle \neq 0\}, \tag{23.68}$$

alors que le lemme 23.29 indique que chacun des ensembles de l'union est au plus dénombrable. La partie $\mathcal{F}$ est donc une union dénombrable d'ensemble dénombrables. Elle est dénombrable. □

Remarque 23.32.
En travaillant un peu plus sur la notion de cardinalité, le corollaire 23.31 indique que toutes les bases hilbertiennes ont même cardinalité. En effet en laissant tomber l'hypothèse de dénombrabilité sur $\mathcal{E}$, l'inclusion (23.68) donne que $\mathcal{F}$ est une union de $\text{Card}(\mathcal{E})$ ensembles dénombrables et est alors de cardinalité $\text{Card}(\mathcal{E})$.

Définition 23.33.
*Une partie orthonormale B est **maximale** si le seul $x \in \mathscr{H}$ vérifiant $\langle x, b\rangle = 0$ pour tout $b \in B$ est $x = 0$.*

Lemme 23.34 ([1, 269])**.**
Tout espace de Hilbert possède une partie orthonormale maximale.

Démonstration. Soit $\mathscr{H}$ un espace de Hilbert et $\mathcal{O}$, l'ensemble de parties orthonormales de $\mathscr{H}$, ordonné [8] par l'inclusion. Cet ensemble est inductif. En effet soit une partie totalement ordonnée $\{A_i\}_{i\in I}$ de $\mathcal{O}$: chaque A_i est une partie orthonormale de $\mathscr{H}$, et de plus pour $i, j \in I$ nous avons soit $A_i \subset A_j$ soit $A_j \subset A_i$. Nous pouvons considérer

$$A = \bigcup_{i\in I} A_i. \tag{23.69}$$

Cela est encore une partie orthonormé de $\mathscr{H}$ parce que si $x, y \in A$, alors il existe $i, j \in I$ tels que $x \in A_i$ et $y \in A_j$. Vu que $\{A_i\}_{i\in I}$ est totalement ordonné nous supposons pour fixer les idées que $A_i \subset A_j$. Alors x et y sont dans A_j qui est une partie orthonormée ; nous en déduisons que $\langle x, y\rangle = 0$ et donc que A est une partie orthonormée. C'est à dire : $A \in \mathcal{O}$. Par ailleurs, A est un majorant de $\{A_i\}_{i\in I}$ pour l'inclusion parce que $A_i \subset A$ pour tout i.

Nous avons prouvé que $\mathcal{O}$ est un ensemble inductif. Le lemme de Zorn 2.10 nous dit alors que $\mathcal{O}$ possède un maximum. Ce maximum est une partie orthonormale inclue dans aucune autre partie orthonormale. C'est à dire que ce maximum est une partie orthonormale maximale au sens de la définition 23.33. □

Proposition 23.35 ([269])**.**
Si un espace de Hilbert possède une partie orthonormale maximale dénombrable, alors toutes les parties orthonormales sont dénombrables (ou finies).

Démonstration. Soit B une partie orthonormale maximale, et A une partie orthonormale. Pour chaque $b \in B$ nous notons

$$A(b) = \{a \in A \text{ tel que } \langle a, b\rangle \neq 0\} \cup \{0\}. \tag{23.70}$$

8. Définition 2.3.

Un élément de A qui ne serait dans aucun des $A(b)$ serait perpendiculaire à tous les éléments de B, et serait donc l'élément nul. Mais l'élément nul est dans tous les $A(b)$; donc nous avons

$$A \subset \bigcup_{b\in B} A(b). \tag{23.71}$$

Or par l'inégalité de Bessel, l'ensemble $A(b)$ est dénombrable. Par conséquent A est inclus à une union dénombrable d'ensembles dénombrables ; A est dénombrable. □

Le fait que tout espace de Hilbert possède une base hilbertienne est vrai. Nous allons démontrer ce résultat d'abord pour les espaces séparables et ensuite, indépendamment, en général. Si vous vous la sentez de maîtriser la proposition 23.37, vous pouvez sauter la 23.36.

Proposition 23.36.
Tout espace de Hilbert séparable possède une bases hilbertienne.

Démonstration. Vu que nous supposons avoir un espace de Hilbert séparable, il possède une partie totale dénombrable par la proposition 23.21. Soit $(v_n)_{n\in\mathbb{N}}$ une telle partie. Quitte à supprimer les v_i qui sont combinaisons linéaires des précédents, nous pouvons supposer que cette partie est libre. Nous considérons l'espace vectoriel

$$F_n = \text{Span}\{v_1, \dots, v_n\}. \tag{23.72}$$

Sur F_n nous pouvons appliquer un procédé de Gram-Schmidt pour construire une base orthonormée $\{u_1, \dots, u_n\}$ de F_n au sens usuel. En considérant F_{n+1} et en recommençant, les vecteurs $u_1, \dots, u_n$ ne changent pas, mais nous obtenons un vecteur u_{n+1}.

Nous construisons ainsi une suite (u_n) qui est alors orthonormée au sens des espaces de Hilbert. Nous devons encore prouver qu'il s'agit d'un ensemble total. Cela est simplement dû au fait que tout élément de $\text{Span}\{v_n\}$ est contenu dans $\text{Span}\{u_n\}$ parce que Span ne considère que des combinaisons linéaires finies. □

Proposition 23.37 ([269])**.**
À propos de parties orthonormales maximales.

1. *Une partie d'un espace de Hilbert est orthonormale maximale si et seulement si elle est une base hilbertienne.*
2. *Tout espace de Hilbert possède une base hilbertienne.*

Démonstration. Soit une partie orthonormale maximale B et la fermeture de son espace engendré : $F = \overline{\text{Span}(B)}$. Pour que B soit une base, nous devons démontrer que $F = \mathscr{H}$. Pour cela nous considérons $x \in \mathscr{H}$ et nous utilisons le théorème de projection orthogonale 23.8 pour mentionner le fait que

$$x - \texttt{proj}_F(x) \perp F \tag{23.73}$$

Cela dit que $x - \texttt{proj}_F(x)$ est un vecteur orthogonal en particulier à tous les éléments de B ; par maximalité nous avons $x - \texttt{proj}_F(x) = 0$, c'est à dire $x = \texttt{proj}_F(x)$ ou encore $x \in F$. Cela prouve que $F = \mathscr{H}$.

Note : pour être pointilleux, nous aurions dû travailler non avec $x - \texttt{proj}_F(x)$, mais avec le vecteur normalisé à 1.

En ce qui concerne le second point, nous invoquons le lemme 23.34 pour dire que tout espace de Hilbert possède une partie orthonormale maximale. Ensuite la première partie de cette proposition nous dit que cette dernière est une base hilbertienne. □

Voici un petit résumé de ce que nous avons vu en termes de dénombrabilité et séparabilité.

Théorème 23.38.
Pour un espace de Hilbert, les choses suivantes sont équivalentes.

(1) L'espace est séparable[9]*.*

(2) L'espace possède au moins une base hilbertienne[10] *dénombrable.*

(3) Toutes les bases hilbertiennes sont dénombrables.

(4) Toute partie libre est dénombrable.

Démonstration. Plein de résultats à citer ...

(1) implique (2) est la proposition 23.36.

(2) implique (3) est la proposition 23.31.

(3) implique (4) Le procédé de Gram-Schmidt met en bijection une partie libre avec une partie orthonormale (qui engendre le même espace, mais c'est une autre affaire). Or lorsque l'espace de Hilbert possède une base dénombrable, toutes les parties orthonormales sont dénombrables par la proposition 23.35.

(4) implique (1) Tout espace de Hilbert possède des bases hilbertiennes par la proposition 23.37(2). Une telle base est forcément une partie libre parce que toute famille orthonormale est libre (proposition 23.24), et donc dénombrable par hypothèse. À ce point nous avons montré que notre espace de Hilbert possédait une base hilbertienne dénombrable. Cela implique qu'il est séparable par la proposition 23.21.

□

Remarque 23.39.
À mon avis il doit exister un théorème de complétion de base hilbertienne disant que si on a une famille orthonormée, alors elle se prolonge en base. Utilisant cela, nous trouvons une nouvelle démonstration de la proposition 23.25 en disant que la somme sur la «partie de base» est plus petite que la somme sur la «base complète».

23.40.
Vu que nous n'avons l'intention de ne travailler qu'avec des espaces de Hilbert séparables et que toutes leurs bases sont dénombrables, nous n'allons travailler qu'avec des bases dénombrables, et donc des systèmes orthonormés dénombrables. Nous allons conventionnellement les indicer par $\mathbb{N}$.

La proposition suivante explique que la notion de projection est compatible avec la décomposition d'un vecteur dans un système orthonormé.

Proposition 23.41.
Soit $(u_k)_{k\geqslant 1}$ un système orthonormé d'un préhilbert $\mathscr{H}$. Soient

$$x = \sum_{k=1}^{\infty} \xi_k u_k \tag{23.74}$$

et

$$F = \text{Span}\{u_1, \dots, u_n\}. \tag{23.75}$$

Alors

$$\texttt{proj}_F(x) = \sum_{k=1}^{n} \xi_k u_k. \tag{23.76}$$

Démonstration. Nous allons dans un premier temps montrer que

$$y = x - \sum_{k=1}^{n} \xi_k u_k \tag{23.77}$$

9. Définition 23.19.
10. Définition 23.27.

est dans $F^\perp$. Pour cela nous calculons

$$\langle x - \sum_{k=1}^{n} \xi_k u_k, u_j, \rangle = \langle \sum_{k=n+1}^{\infty} \xi_k u_k, u_j \rangle = 0 \tag{23.78}$$

où nous avons utilisé la continuité du produit scalaire pour permuter la somme (infinie) et le produit. Étant donné que $y \in F^\perp$ nous avons $\mathtt{proj}_F y = 0$ par le point (2) du théorème 23.13.

D'autre part $\mathtt{proj}_F y$ peut être calculé selon

$$\mathtt{proj}_F y = \mathtt{proj}_F x - \sum_{k=1}^{n} \xi_k \mathtt{proj}_F u_k \tag{23.79}$$

tandis que $\mathtt{proj}_F u_k = u_k$ lorsque $1 \leqslant k \leqslant n$. Par conséquent l'annulation de $\mathtt{proj}_F y$ donne

$$\mathtt{proj}_F x = \sum_{k=1}^{n} \xi_k u_k, \tag{23.80}$$

donc le résultat. □

Théorème 23.42 (Meilleur approximation)**.**
Soit $\{u_i\}_{i\in I}$ une famille orthonormé de l'espace de Hilbert $\mathscr{H}$. Alors pour tout $x \in \mathscr{H}$ et pour tout J fini dans I et pour toute famille de nombres complexes $(a_j)_{j\in J}$ nous avons

$$\| \sum_{j\in J} \langle x, u_j \rangle u_j - x \| \leqslant \| \sum_{j\in J} a_j u_j - x \|. \tag{23.81}$$

Ce théorème exprime le fait que les nombres $\langle x, u_i \rangle$ sont les meilleurs coefficients à mettre devant les u_i pour approximer x.

Corollaire 23.43.
Soit $\{u_i\}_{i\in I}$ une famille orthonormé de $\mathscr{H}$. Pour tout $x \in \mathscr{H}$ et pour toutes parties finies J, K de I avec $J \subset K$ nous avons

$$\| \sum_{j\in K} \langle x, u_j \rangle u_j - x \| \leqslant \| \sum_{j\in J} \langle x, u_j \rangle u_j - x \|. \tag{23.82}$$

Ce corollaire exprime le fait que plus on prend de termes de la forme $\langle , u_i \rangle u_i$, mieux c'est.

Proposition 23.44.
Soit $\mathscr{H}$ un espace de Hilbert et (u_n) un système orthonormé dans $\mathscr{H}$. Si $(\xi_n)_{n\in\mathbb{N}}$ est une suite dans ℓ^2 alors la série

$$\sum_{n=1}^{\infty} \xi_n u_n \tag{23.83}$$

converge dans $\mathscr{H}$.

Autrement dit l'application

$$\begin{aligned} S\colon \mathscr{H} &\to \ell^2 \\ x &\mapsto \big(\langle x, u_n \rangle\big)_{n\geqslant 1} \end{aligned} \tag{23.84}$$

est surjective.

Démonstration. Nous allons montrer que la série $\sum_{n=1}^{\infty} \xi_n u_n$ est de Cauchy, c'est à dire que la limite

$$\lim_{n\to\infty} \| \sum_{k=n}^{n+p} \xi_k u_k \| = 0 \tag{23.85}$$

est uniforme en p. Cela est un corollaire de la formule (23.50) parce que

$$\| \sum_{k=n}^{n+p} \xi_k u_k \|^2 = \sum_{k=n}^{n+p} |x_k|^2. \tag{23.86}$$

Mais si (ξ_n) est dans ℓ^2, pour tout ϵ, il existe N tel que si $n > N$ alors le membre de droite est inférieur à ϵ indépendamment de p. □

23.3.5 Décomposition dans une base hilbertienne

Étant donné une base hilbertienne de $\mathscr{H}$, nous notons

$$c_k(x) = \langle x, u_k \rangle. \tag{23.87}$$

Dans le théorème suivant (et d'ailleurs partout), les sommes sur I sont prises au sens de la définition 9.179.

Théorème 23.45 (Décomposition dans une base orthogonale)**.**
Soit $\mathscr{H}$ un espace de Hilbert séparable $\{u_i\}_{i\in I}$ une base orthonormée (I est un ensemble dénombrable quelconque).

(1) Pour tout $x \in \mathscr{H}$ nous avons

$$x = \sum_{i\in I} \langle x, u_i \rangle u_i \tag{23.88}$$

où la somme est prise au sens de la définition 9.179. En particulier, la somme converge de façon commutative.

(2) Si $\{e_i\}_{i\in I}$ est une famille orthonormée qui satisfait la décomposition (23.88) *pour tout $x \in \mathscr{H}$ alors $\{e_i\}$ est une base hilbertienne.*

(3) Nous avons l'identité de ***Plancherel***

$$\|x\|^2 = \sum_{i\in I} |\langle x, u_i \rangle|^2. \tag{23.89}$$

Le point (5) nous indiquera que cette égalité est en fait suffisante pour dire que nous avons une base.

(4) Nous avons l'identité de ***Parseval***

$$\langle x, y \rangle = \sum_{i\in I} \langle x, u_i \rangle \overline{\langle y, u_i \rangle}. \tag{23.90}$$

(5) Si $\{e_i\}$ est une famille de vecteurs unitaires vérifiant l'identité de Plancherel, alors c'est une base hilbertienne.

(6) Si $\{u_i\}_{i\in I}$ est une base hilbertienne, la suite $n \mapsto |\langle x, e_n \rangle|$ appartient à $\ell^2(I)$.

Démonstration. (1) Étant donné que le système $\{u_i\}_{i\in I}$ est total, nous pouvons considérer une suite de combinaisons linéaires finies des u_i qui converge vers x. Nous écrivons

$$x_n = \sum_{x\in J_n} a_{n,k} u_j \tag{23.91}$$

et $x_n \to x$ dans $\mathscr{H}$. Les ensembles J_n sont des sous-ensembles finis de I. Nous pouvons les choisir de telle sorte que $J_n \subset J_{n+1}$ et $\bigcup_{n\in\mathbb{N}} J_n = I$. Ce choix correspond à éventuellement prendre $a_{n,j} = 0$ pour toutes les valeurs de j «en trop».
Soit $\epsilon > 0$ et N tel que $\|x_n - x\| < \epsilon$ pour tout $n \geqslant N$. Nous allons montrer que pour tout J fini tel que $J_N \subset J$ nous avons $\| \sum_{j\in J} \langle x, u_j \rangle u_j - x\| < \epsilon$. Étant donné que

$$\sum_{j\in J_N} \langle x, u_j \rangle u_j = \sum_{j\in J} a_j u_j \tag{23.92}$$

avec

$$a_j = \begin{cases} \langle x, u_j \rangle & \text{si } j \in J_N \\ 0 & \text{sinon,} \end{cases} \tag{23.93}$$

le théorème de meilleure approximation 23.42 nous enseigne que

$$\| \sum_{j\in J} \langle x, u_j \rangle u_j - x\| \leqslant \| \sum_{j\in J_N} \langle x, u_j \rangle u_j - x\| < \epsilon \tag{23.94}$$

par conséquent la somme $\sum_{i\in I}\langle x,u_i\rangle u_i$ converge vers x au sens général, et en particulier commutativement.

Les sommes étant commutatives (en particulier $x=\sum_{i\in I}\langle x,u_i\rangle u_i$), et les bases hilbertiennes étant dénombrables, nous ne perdons aucune généralité en ne considérant que des bases indexées par $\mathbb{N}$.

(2) Nous devons montrer que l'ensemble des combinaisons linéaires finies est dense dans $\mathscr{H}$. Par hypothèse, pour tout ϵ, il existe un ensemble fini J tel que

$$\|\sum_{j\in J}\langle x,u_j\rangle u_j-x\|<\epsilon. \tag{23.95}$$

Cela prouve la densité dont nous avions besoin.

(3) La norme étant une fonction continue, elle commute avec les sommes infinies, de telle sorte que l'égalité de Plancherel donne

$$\|x\|^2=\|\sum_{i\in I}\langle x,u_i\rangle u_i\|^2. \tag{23.96}$$

Le système des u_i étant orthonormé,

$$\|x\|^2=\|\sum_{i\in I}\langle x,u_i\rangle u_i\|^2=\sum_{i\in I}|\langle x,u_i\rangle|^2. \tag{23.97}$$

(4) Au tour de Parseval. Nous commençons par prouver que la somme du membre de droite converge. En utilisant l'inégalité $|zz'|\leqslant|z|^2+|z'|^2$ (valable pour $z,z'\in\mathbb{C}$) et Plancherel, nous avons

$$\begin{aligned}\sum_{i\in I}|\langle x,u_i\rangle\overline{\langle y,u_i\rangle}|&\leqslant\sum_{i\in I}|\langle x,u_i\rangle|^2+|\langle y,u_i\rangle|^2\\&\leqslant\|x\|^2+\|y\|^2.\end{aligned} \tag{23.98}$$

Nous en déduisons que la famille $\langle x,u_i\rangle\overline{\langle y,u_i\rangle}$ est (commutativement) sommable en utilisant la proposition 9.186. Par ailleurs nous savons que

$$x=\sum_{i\in I}c_i(x)u_i \tag{23.99a}$$

$$y=\sum_{i\in I}c_i(y)u_i, \tag{23.99b}$$

et le produit scalaire étant une forme bilinéaire continue,

$$\langle x,y\rangle=\sum_{i\in I}\sum_{j\in I}\langle c_i(x)u_i,\overline{c_j(y)u_j}\rangle \tag{23.100a}$$

$$=\sum_{i\in I}\sum_{j\in J}c_i(x)\overline{c_j(y)}\delta_{ij} \tag{23.100b}$$

$$=\sum_{i\in I}c_i(x)\overline{c_j(y)}. \tag{23.100c}$$

(5) Nous utilisons l'égalité de Plancherel avec $x=u_j$:

$$\|u_j\|^2=\|u_j\|+\sum_{i\in I\setminus\{j\}}|\langle u_j,u_i\rangle|^2. \tag{23.101}$$

Par conséquent $\langle u_j,u_i\rangle=0$ dès que $i\neq j$. Cela prouve que le système $\{u_i\}_{i\in I}$ est orthonormé. Nous devons encore prouver que le système est total. Pour cela nous repartons de l'équation (23.61) que nous avions déduites dans la démonstration de l'inégalité de Bessel :

$$\|x-\sum_{j\in J}c_j(x)u_j\|=\|x\|^2-\sum_{j\in J}|c_j(x)|^2. \tag{23.102}$$

Par hypothèse le membre de droite peut être rendu aussi petit que l'on veut en prenant J grand (mais fini) dans I. Le membre de gauche indique alors que le système $\{u_i\}_{i\in I}$ est total.

(6) L'identité de Plancherel signifie entre autres que si $x \in \mathscr{H}$ alors $\sum_{i\in I} |\langle x, u_i\rangle|^2$ converge. Du coup la suite $(\langle x, u_i\rangle)_{i\in I}$ est dans $\ell^2(I)$.

□

Remarque 23.46.
Nous avons décidé d'indexer les bases hilbertiennes par $\mathbb{N}$; cela est légitime parce que les sommes sont commutatives. Il ne faut cependant pas perdre de vue qu'en pratique l'ensemble naturel avec lequel on indexe une base est parfois $\mathbb{Z}$. Un tel cas est donné par la base trigonométrique de L^2. Indexer cette dernière par $\mathbb{N}$ plutôt que par $\mathbb{Z}$ serait une contorsion inutile.

Remarque 23.47.
L'égalité de Parseval est la raison pour laquelle les physiciens écrivent souvent

$$\mathrm{Id} = \sum_{n=1}^{\infty} |u_n\rangle\langle u_n| \tag{23.103}$$

dans les livres de mécanique quantique par exemple. Dans certains, nous lisons même

$$\int_{-\infty}^{+\infty} dq|q\rangle\langle q| = \hat{I}. \tag{23.104}$$

Notons que ces personnes travaillent avec un espace de Hilbert dont la base n'est pas dénombrable. Pour dire que la physique, ça n'utilise pas des mathématiques pour rire !

Remarque 23.48.
Par définition une base orthonormée est donc une partie dénombrable dont l'espace vectoriel engendré est dense. Un espace de Hilbert possédant une base orthonormée est donc séparable. C'est ce fait qui nous pousse à ne considérer que des espaces de Hilbert séparables ; nous n'allons donc pas étudier ce qu'il se passerait par exemple en considérant l'espace vectoriel librement engendré par les éléments de $\mathbb{R}$.

Exemple 23.49
L'identité de Parseval (23.90), dans le cas de l'espace des fonctions continues périodiques de période 2π signifie qu'en posant

$$c_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(s)e^{-ins}, \tag{23.105}$$

nous avons

$$\frac{1}{2\pi}\int_0^{2\pi} |f(s)|^2 = \sum_{n=-\infty}^{\infty} |c_n(f)|^2. \tag{23.106}$$

△

Corollaire 23.50.
Soit $\mathscr{H}$ un espace de Hilbert et (u_n) une base orthonormée. L'application

$$\begin{aligned} S\colon \mathscr{H} &\to \ell^2 \\ x &\mapsto \big(\langle x, u_n\rangle\big)_{n\in\mathbb{N}} \end{aligned} \tag{23.107}$$

est un isomorphisme d'espaces de Hilbert.

De plus l'isomorphisme réciproque est

$$\begin{aligned} S^{-1}\colon \ell^2 &\to \mathscr{H} \\ (\xi_n) &\mapsto \sum_{n=1}^{\infty} \xi_n u_n. \end{aligned} \tag{23.108}$$

Démonstration. Nous devons prouver que l'application est bijective et qu'elle vérifie

$$S(x) \cdot S(y) = \langle x, y \rangle \tag{23.109}$$

où le point dénote le produit dans ℓ^2.

Pour la surjectivité, si $(\xi_n) \in \ell^2$ alors nous savons que la somme $\sum_n \xi_n u_n$ converge par la proposition 23.44 et par conséquent (ξ_n) est l'image par S de ce vecteur de $\mathscr{H}$.

Pour l'injectivité, si $S(x) = S(y)$ alors

$$x = \sum_n \langle x, u_n \rangle u_n = \sum_n \langle y, u_n \rangle u_n = y \tag{23.110}$$

en utilisant la décomposition (23.88).

Le fait que S soit une isométrie est contenu dans Parseval. □

Proposition 23.51 ([1])**.**
Soit un espace de Hilbert séparable $\mathscr{H}$, un sous-espace vectoriel fermé V et une base orthonormée $\{b_i\}_{i\in I}$ de $\mathscr{H}$. En posant $v_i = \texttt{proj}_V(b_i)$ alors

$$C = \overline{\text{Span}\{b_i - v_i\}} \tag{23.111}$$

est l'orthogonal de V.

Démonstration. Juste pour rappel, lorsque nous écrivons $v_i = \texttt{proj}_V(b_i)$, nous parlons de la projection orthogonale du théorème 23.8. Donc tous les vecteurs $b_i - v_i$ sont dans $V^\perp$. En passant aux limites, $C \subset V^\perp$.

Soit $x \in V^\perp$ que nous décomposons dans la base $\{b_i\}_{i\in I}$ comme $x = \sum_{i\in I} x_i b_i$. Posons $c_i = b_i - v_i = b_i - \texttt{proj}_V(b_i)$. Alors d'une part

$$0 = \texttt{proj}_V(x) = \sum_i x_i \texttt{proj}_V(b_i) = \sum_i x_i(b_i - c_i). \tag{23.112}$$

Nous avons utilisé la continuité de $\texttt{proj}_V$ pour permuter avec la somme. D'autre part,

$$x = \sum_i x_i b_i = \sum_i x_i(b_i - c_i + c_i) = \underbrace{\sum_i x_i(b_i - c_i)}_{=0} + \sum_i x_i c_i \in C. \tag{23.113}$$

Notons que pour la dernière appartenance, il est important de prendre la fermeture pour définir C. □

Proposition 23.52 ([1, 270])**.**
Toute partie orthonormée d'un espace de Hilbert séparable se prolonge en une base hilbertienne.

Démonstration. Soit une partie orthonormée $\{u_i\}_{i\in I}$ de l'espace de Hilbert $\mathscr{H}$. Nous mentionnons que cette partie est libre et donc, par le théorème 23.38, dénombrable [11]. Montrons pour commencer que la partie $V = \overline{\text{Span}\{u_i\}_{i\in I}}$ est un sous-espace vectoriel fermé de $\mathscr{H}$.

Vectoriel Soit $v, w \in V$ et $\epsilon > 0$. Il existe $a \in \text{Span}\{u_i\}$ tel que $\|a - v\| < \epsilon$ et $b \in \text{Span}\{u_i\}$ tel que $\|b - w\| < \epsilon$. Dans ce cas,

$$\|v + w - (a + b)\| \leqslant 2\epsilon. \tag{23.114}$$

Cela prouve que $v + w \in V$. Nous procédons de même pour λv.

Fermé Par construction.

11. Avec un peu de mauvaise foi, vous pouvez quand même dire que cela n'implique pas que I lui-même soit dénombrable, mais vous pouvez le supposer pour fixer les idées.

L'orthogonal[12] de V est un sous-espace vectoriel de $\mathscr{H}$, et nous pouvons donc en considérer une base hilbertienne $C = \{c_\alpha\}_{\alpha\in A}$. Nous prétendons que $C \cup U$ est une base hilbertienne[13] de $\mathscr{H}$.

$C \cup U$ **est libre** Bing ! Il ne faut pas le démontrer : ça ne fait pas partie de la définition d'une base hilbertienne.

$C \cup U$ **est générateur** Bang ! Il ne faut pas le démontrer : ça ne fait pas partie de la définition d'une base hilbertienne.

$C \cup U$ **est orthogonal** Ah, voila quelque chose à démontrer. Nous devons vérifier que les produits sont nuls. Soient $u, v \in U$ et $a, b \in C$. Nous avons :
— $\langle u, v\rangle = 0$ par hypothèse.
— $\langle u, a\rangle = 0$ parce que les éléments de C sont orthogonaux à V et que $u_i \in V$.
— $\langle a, b\rangle = 0$ parce que C est une base hilbertienne de $V^\perp$.

$C \cup U$ **est dénombrable** L'ensemble C est dénombrable parce que c'est une base hilbertienne. Quant à $\{u_i\}_{i\in I}$, nous avons déjà mentionné le fait qu'il doive être dénombrable. L'union deux parties dénombrables est dénombrable.

$C \cup U$ **est total** Nous devons prouver que $\overline{\text{Span}(C \cup U)} = \mathscr{H}$ parce qu'il y a bien la fermeture qui intervient dans la définition 23.20. Pour cela nous utilisons la proposition 23.15. Si $x \in \mathscr{H}$ alors nous avons

$$x = \texttt{proj}_V(x) + \big(x - \texttt{proj}_V(x)\big) = \sum_{i\in I} x_i u_i + \sum_{\alpha\in A} x_\alpha c_\alpha \tag{23.115}$$

pour des coefficients x_i et x_α. Notons que I et A sont deux ensembles différents. Aucun des x_α n'est un des x_i, ni inversement. Supposons que $x \in \text{Span}(C \cup U)^\perp$; alors

$$0 = \langle x, u_j\rangle = \sum_{i\in I} x_i \underbrace{\langle u_i, u_j\rangle}_{=\delta_{ij}} + \sum_{\alpha\in A} \underbrace{\langle c_\alpha, u_j\rangle}_{=0} = x_j. \tag{23.116}$$

Donc $x_j = 0$. En faisant de même avec $0 = \langle x, c_\beta\rangle = x_\beta$ nous déduisons $x = 0$.

□

23.3.6 Digression sur les normes opérateurs

Le théorème 23.45 nous indique que si $\{u_i\}_{i\in I}$ est une base hilbertienne, alors pour tout $x \in \mathscr{H}$ nous avons

$$\texttt{proj}_{u_i}(x) = \langle x, u_i\rangle u_i, \tag{23.117}$$

et donc

$$\sum_{i\in I} \texttt{proj}_{u_i} x = x. \tag{23.118}$$

Nous ne pouvons cependant pas conclure que

$$\sum_{i\in I} \texttt{proj}_{u_i} = \text{Id} \tag{23.119}$$

au sens de la norme opérateur de la définition 10.5. En effet en prenant $I = \mathbb{N}$, l'égalité (23.119) demanderait d'avoir

$$\lim_{N\to\infty} \left\| \sum_{i=1}^{N} \texttt{proj}_{u_i} - \text{Id} \right\|_\infty = 0, \tag{23.120}$$

or pour tout N, le vecteur u_{N+1} réalise

$$\sum_{i=1}^{N} \texttt{proj}_{u_i}(u_{N+1}) - u_{N+1} = -u_{N+1}. \tag{23.121}$$

12. Voir la définition 23.10 et la proposition 23.51.
13. Définition 23.27.

Par conséquent pour tout N nous avons

$$\sup_{\|x\|=1} \left\| \sum_{i=1}^{N} \texttt{proj}_{u_i} x - x \right\| \geqslant 1. \tag{23.122}$$

Nous ne pouvons donc pas dire

$$\sum_{n=1}^{\infty} \texttt{proj}_{u_i} = \mathrm{Id} \tag{23.123}$$

au sens de la norme opérateur.

Nous avons cependant la convergence au sens faible.

Proposition 23.53.
Soit $\mathscr{H}$ un espace de Hilbert et $\{u_i\}_{i\in\mathbb{N}}$ une base hilbertienne de $\mathscr{H}$. Au sens de la topologie faible sur l'espace des opérateurs nous avons

$$\sum_{i=1}^{\infty} \texttt{proj}_{u_i} = \mathrm{Id}\,. \tag{23.124}$$

Démonstration. Pour chaque $N \in \mathbb{N}$ et $x \in \mathscr{H}$, en vertu de la décomposition (23.88) nous avons

$$\sum_{i=1}^{N} \texttt{proj}_{u_i}(x) - x = \sim_{i=N+1}^{\infty} \langle x, u_i \rangle u_i. \tag{23.125}$$

Par l'orthonormalité de la base nous avons

$$\sum_{i=N+1}^{\infty} \| \langle x, u_i \rangle u_i \| = \sum_{i=N+1}^{\infty} |\langle x, u_i \rangle|, \tag{23.126}$$

dont la limite $N \to \infty$ est zéro étant donné que la suite $i \mapsto |\langle x, u_i \rangle|$ est dans $\ell^2(\mathbb{R})$ par le théorème 23.45. □

23.3.7 Applications linéaires et continuité

Nous avons déjà vu dans l'exemple 10.20 que la fonction

$$\begin{aligned} f\colon H &\to H \\ e_k &\mapsto k e_k \end{aligned} \tag{23.127}$$

n'était pas continue en zéro alors qu'elle est linéaire. Nous allons maintenant voir qu'elle est un contre-exemple à la proposition 12.204. Calculons les dérivées partielles :

$$\frac{\partial f}{\partial e_j}(x) = \frac{d}{dt}\Big[f(x + te_j) \Big]_{t=0} = \frac{d}{dt}\Big[\sum_k k(x_k + t\delta_{jk}) e_k \Big]_{t=0} = j e_j. \tag{23.128}$$

où nous avons permuté la somme et la dérivée en considérant la suite de fonctions $f_k(t) = k(x_k + tu_k)e_k$. Donc $\frac{\partial f}{\partial e_j}(x)$ existe et est continue sur un voisinage de $x = 0$ (c'est même constant). Nous savons pourtant que la fonction f n'est pas différentiable en zéro parce que non continue.

L'endroit qui coince dans la preuve de la proposition 12.204 est l'introduction des «contres-termes» dans l'équation (12.495). En effet les contre-termes à ajouter seraient

$$l(x) = \frac{d}{dt}\Big[f\big(a + s(x-a)\big) \Big]_{t=0} \tag{23.129}$$

qui ici serait

$$l\big(\sum_k x_k e_k\big) = \frac{d}{dt}\Big[f\big(s \sum_k x_k e_k\big) \Big]_{t=0} = \sum_k x_k e_k, \tag{23.130}$$

dont la convergence est plus que douteuse.

Notons que les dérivées directionnelles n'existent pas toutes, loin s'en faut : si $u \in H$ nous avons

$$\frac{\partial f}{\partial u}(x) = \frac{d}{dt}\Big[f(x + tu) \Big]_{t=0} = \frac{d}{dt}\Big[\sum_k k(x_k + tu_k) e_k \Big]_{t=0} = \sum_k k u_k e_k \tag{23.131}$$

La convergence de la dernière somme n'est pas garantie pour tout u.

23.4 Théorème de Kochen-Specker

Le théorème suivant est central en mécanique quantique. La démonstration provient de [271] et de Wikipédia. Nous allons démontrer complètement le théorème seulement pour les espaces de Hilbert de dimension plus grande ou égale à 4.

Théorème 23.54 (Kochen-Specker[271])**.**
Soit $\mathscr{H}$ un espace de Hilbert de dimension plus grande ou égale à 3. Une fonction v sur l'ensemble des opérateurs de $\mathscr{H}$ ne peut pas satisfaire aux conditions suivantes :

(1) Si A et B sont compatibles, alors $v(A+B)=v(A)+v(B)$,

(2) Si A et B sont compatibles, alors $v(AB)=v(A)v(B)$.

*Ici nous disons que deux opérateurs sont **compatibles** lorsqu'ils possèdent une base hilbertienne commune de vecteurs propres.*

Démonstration. Soit $\{u_n\}_{n\in\mathbb{N}}$ une base hilbertienne de $\mathscr{H}$. Nous notons $\mathtt{proj}_i$ l'opérateur de projection sur l'espace (fermé) engendré par u_i. Ce sont des opérateurs compatibles deux à deux parce que la base $\{u_n\}_{n\in\mathbb{N}}$ est une base commune de vecteurs propres [14].

D'abord nous devons avoir $v(\mathbb{1})=1$. En effet pour tout opérateur A, nous avons

$$v(A)=v(A\mathbb{1})=v(A)v(\mathbb{1}). \tag{23.132}$$

Pour peu que $v(A)\neq 0$, cela nous fait $v(\mathbb{1})=1$.

En vertu du théorème 23.45, un vecteur $x\in\mathscr{H}$ se décompose en $x=\sum_n\langle x,u_n\rangle u_n$, et nous avons

$$\mathtt{proj}_i\, x=\langle x,u_i\rangle u_i. \tag{23.133}$$

En effet le théorème de la projection orthogonale 23.8 nous enseigne que $\mathtt{proj}_i\, x$ serait l'unique vecteur de la forme λu_i tel que $\lambda u_i - x\perp u_i$. Il est facile de vérifier que le vecteur proposé par (23.133) vérifie cette propriété.

Une conséquence est que

$$\left(\sum_{i=0}^{\infty}\mathtt{proj}_i\right)(x)=x. \tag{23.134}$$

Par conséquent, par hypothèse du théorème nous devons avoir

$$\sum_i v(\mathtt{proj}_i)=v(\mathbb{1})=1. \tag{23.135}$$

Étant donné que les projections sont idempotentes,

$$v(\mathtt{proj}_i)=v(\mathtt{proj}_i^2)=v(\mathtt{proj}_i)^2 \tag{23.136}$$

et donc $v(\mathtt{proj}_i)$ doit valoir zéro ou un. Mais la relation (23.135) donne une forte contrainte sur le choix de 0 et de 1. En effet, parmi les $v(\mathtt{proj}_i)$, un et un seul doit valoir 1, les autres doivent valoir 0.

Refaisant le raisonnement pour une autre base orthonormale hilbertienne, nous trouvons que les valeurs de v sur les opérateurs de projection sur les différentes directions doivent être choisies de telle façon que tout choix de base hilbertienne orthogonale contienne exactement un 1 et les reste de zéros.

Nous voudrions maintenant insister sur un point. Le problème de déterminer de façon cohérente les valeurs 0 ou 1 pour tous les $v(\mathtt{proj}_i)$ revient à attacher 0 ou 1 à tous les *rayons* de $\mathscr{H}$ de façon que toute base orthogonale de $\mathscr{H}$ contienne exactement un 1. Un rayon est une direction, c'est à dire une classe d'équivalence $x\sim\lambda x$. Si nous décidons de nommer «blanc» les rayons attachés à

14. Pour les besoins de la physique, nous remarquons que ces opérateurs sont des opérateurs hermitiens qui commutent, mais ça ne joue pas ici.

la valeur 0 et «noirs» ceux attachés à la valeur 1, le problème se réduit à colorer la boule unité de façon compatible.

Soit $\{u_i\}_{i\in\mathbb{N}}$ une base orthogonale de $\mathscr{H}$ numérotée de telle sorte que $v(u_0) = 1$ et $v(u_k) = 0$ pour $k \neq 0$. Nous allons maintenant nous particulariser au cas de dimension supérieure ou égale à 4. Si R est une rotation dans le plan $\text{Span}\{u_0, u_1, u_2, u_3\}$, alors l'ensemble

$$\{Ru_0, Ru_1, Ru_2, Ru_3, u_k\}_{k\geqslant 3} \tag{23.137}$$

est encore une base orthogonale de $\mathscr{H}$ et nous avons encore $v(u_k) = 0$ pour $k \geqslant 4$. Par conséquent un et un seul des vecteurs Ru_0, Ru_1, Ru_2 ou Ru_3 est colorié en noir ; les trois autres étant blancs. Le problème est maintenant complètement réduit à la dimension 4. Note : pour réduire à la dimension 3, on procède de même, mais pour conclure, il faut travailler plus.

Nous allons construire 9 base orthogonales de $\mathbb{R}^4$ à partir de 18 vecteurs, chacun arrivant dans exactement deux des bases. Ils sont donnés dans le tableau suivant :

0	1	0	0	0	0	1	1	1	−1	1	−1	−1	1	1	1	1	1
0	0	1	0	0	1	1	1	1	−1	−1	1	1	1	−1	1	1	−1
0	0	0	1	0	0	1	−1	1	−1	1	1	1	1	−1	1	1	1
0	1	0	0	1	0	1	−1	−1	1	1	1	1	−1	1	1	−1	1
1	0	1	0	1	1	1	0	1	1	1	0	1	0	1	0	1	−1
1	0	0	1	0	0	−1	0	0	0	0	−1	0	1	1	0	0	0
1	0	1	0	1	−1	0	1	0	0	0	1	0	1	0	1	0	0
−1	0	0	−1	0	0	0	−1	1	1	−1	0	−1	0	0	−1	1	1

Chaque case de ce table représente un rayon de $\mathbb{R}^4$; il y en a 18 différents, chacun écris deux fois. Une simple vérification montre que chaque colonne est un système orthogonal. La preuve du théorème de Kochen-Specker revient à montrer que nous ne pouvons pas colorier ce tableau de façon cohérente. En effet, étant donné que chaque vecteur est écrit deux fois, le tableau doit contenir un nombre pair de cases blanches et un nombre pair de cases noires.

Par ailleurs chaque colonne étant un système orthogonal, chaque colonne contient exactement une case noire ; il y a donc exactement neuf cases noires dans le tableau, ce qui est impossible.

□

23.5 Théorème de Lax-Milgram

Définition 23.55.
Une forme bilinéaire $a\colon V \times V \to \mathbb{R}$ sur un espace vectoriel normé V est ***coercitive*** *s'il existe $\alpha > 0$ tel que $a(u, u) \geqslant \alpha\|u\|^2$ pour tout $u \in V$.*

Théorème 23.56 (Lax-Milgram[272])**.**
Soit un espace de Hilbert réel V muni de différentes choses.

1. *L'application linéaire $L\colon V \to \mathbb{R}$ qui est bornée sur V. Nous notons C sa norme.*
2. *La forme bilinéaire continue a sur $V \times V$. Nous considérons $M > 0$ tel que $|a(u, v)| \leqslant M\|u\|\|v\|$ pour tout $u, v \in V$.*
3. *La forme a est coercitive*[15]*.*

Alors le problème qui consiste à chercher $u \in V$ tel que $a(u, v) = L(v)$ pour tout $v \in V$ admet une unique solution. De plus cette solution vérifie l'inégalité

$$\|u\| \leqslant \frac{M}{\alpha}C \tag{23.138}$$

Problèmes et choses à faire

Personnellement, je ne parviens pas à montrer l'inégalité (23.138), mais seulement

$$\|u\| \leqslant \frac{C}{\alpha}. \tag{23.139}$$

15. Définition 23.55.

Démonstration. Avant de commencer, certaines précisions. D'abord nous nous souvenons de la proposition 10.18 qui donne la continuité de L. Ensuite, a est continue et linéaire en tant qu'application $a\colon V\times V\to\mathbb{R}$. Nous aurions donc envie d'écrire qu'il existe $M>0$ tel que

$$|a(u,v)|\leqslant M\|(u,v)\|,\tag{23.140}$$

mais la norme produit [16] sur $V\times V$ ne donne pas $\|u\|\|v\|$. Le fait pour a d'être *bi*linéaire donne en réalité plus que la linéarité, et la proposition 10.76 nous assure l'existence du M.

Reformulation en équation linéaire La forme L est continue et donc dans le dual V' ; le théorème de Riesz 23.18 nous donne donc $f\in V$ tel que

$$L(v)=\langle f,v\rangle\tag{23.141}$$

pour tout $v\in V$. De plus si $w\in V$ est fixé, l'application $b_w\colon v\mapsto a(w,v)$ est linéaire et bornée parce que

$$\|b_w\|=\sup_{\|v\|=1}|b_w(v)|=\sup_{\|v\|=1}|a(w,v)|\leqslant M\|w\|\|v\|=M\|w\|.\tag{23.142}$$

Encore une fois, b_w étant continue et linéaire, elle est dans V' et Riesz nous fournit un élément $A(w)\in V$ tel que

$$b_w(v)=\langle A(w),v\rangle\tag{23.143}$$

pour tout $v\in V$.

Le problème variationnel $a(u,v)=L(v)$ est équivalent à $\langle A(u),v\rangle=\langle f,v\rangle$. L'ensemble des solution de cette dernière est égal à l'ensemble des solution de l'équation

$$A(u)=f.\tag{23.144}$$

A est linéaire Soient $\alpha,\beta\in\mathbb{R}$ et $w,z\in V$. Nous avons pour tout $v\in V$:

$$\begin{aligned}\langle A(\alpha w+\beta z),v\rangle&=a(\alpha w+\beta z,v)&(23.145a)\\&=\alpha a(w,v)+\beta a(z,v)&(23.145b)\\&=\alpha\langle A(w),v\rangle+\beta\langle A(z),v\rangle&(23.145c)\\&=\langle\alpha A(w)+\beta A(z),v\rangle.&(23.145d)\end{aligned}$$

Étant donné que nous avons égalité pour tout $v\in V$ nous en déduisons que $A(\alpha w+\beta z)=\alpha A(w)+\beta A(z)$, ce qui signifie que A est linéaire.

Une autre propriété de A Nous déduisons une majoration de $\|A(v)\|^2$ lorsque ce n'est pas nul. Pour ce faire,

$$\begin{aligned}\|A(v)\|^2a&=\langle A(v),A(v)\rangle&(23.146a)\\&=a(v,A(v))&(23.146b)\\&\leqslant M\|v\|\|A(v)\|.&(23.146c)\end{aligned}$$

En simplifiant, $\|A(v)\|\leqslant M\|v\|$. Et donc

$$\|A(v)\|^2\leqslant M^2\|v\|^2.\tag{23.147}$$

Une contraction Nous allons choisir une valeur de $\rho>0$ telle que l'application

$$T\colon w\mapsto w-\rho\big(A(w)-f\big)\tag{23.148}$$

16. Définition 10.73.

soit une contraction[17]. Nous avons $T(w)-T(w')=w-w'-\rho\big(A(w-w')\big)$ et donc

$$\|T(w)-T(w')\|^2=\|w-w'\|^2+\rho^2\|A(w-w')\|^2-2\rho\langle A(w-w'),w-w'\rangle \tag{23.149a}$$
$$=\|w-w'\|^2+\rho^2\|A(w-w')\|^2-2\rho a(w-w',w-w'). \tag{23.149b}$$

Vu que le dernier terme arrive avec un signe moins, pour majorer l'expression, il faut minorer ce terme, c'est à dire utiliser $a(w-w',w-w")\geqslant\alpha\|w-w-w'\|$. Et en même temps nous utilisons (23.147) pour le second terme. Au final pour pouvons factoriser $\|w-w'\|$ et

$$\|T(w)-T(w')\|\leqslant\|w-w'\|\big(1+\rho^2M^2-2\rho\alpha\big). \tag{23.150}$$

Pout que T soit contractante, il faut $0<P(x)<1$ avec $P(x)=M^2x^2-2\alpha x+1$. Le minimum de ce polynôme est obtenu en $x=\frac{\alpha}{M^2}$ (la formule du $x_{min}=-b/2a$) et vaut $1-\frac{\alpha^2}{M^2}<1$. Vu que par ailleurs $\lim_{x\to\infty}P(x)=+\infty$, et que ce polynôme passe par au moins une valeur strictement inférieure à 1, nous savons qu'il existe un x tel que $0<P(x)<1$. En donnant à ρ cette valeur, l'application T est une contraction.

Point fixe et conclusion L'ensemble des solutions du problème (23.144) est égal à l'ensemble des points fixes de $T(v)=v-\rho\big(A(v)-f\big)$.

L'application $T\colon V\to V$ est contractante et V est métrique et complet. Ergo le théorème de point fixe de Picard 18.23 s'applique et il existe un unique point fixe $u\in V$ pour l'application T. Ce point fixe est l'unique solution de notre problème initial.

La majoration Nous savons que pour tout $v\in V$, la relation $a(u,v)=L(v)$ est vérifiée. En particulier pour $v=u$ nous avons

$$a(u,u)=L(u). \tag{23.151}$$

D'un côté nous utilisons $a(u,u)\geqslant\alpha\|u\|^2$ et de l'autre, $L(u)\leqslant C\|u\|$:

$$\alpha\|u\|^2\leqslant C\|u\| \tag{23.152}$$

et donc

$$\|u\|\leqslant\frac{C}{\alpha}. \tag{23.153}$$

Notons que $L(u)$ et $a(u,u)$ sont positifs.

□

Théorème 23.57 (Lax-Milgram version symétrique[272])**.**
Nous considérons les mêmes hypothèses que celles du théorème de Lax-Milgram, c'est à dire un espace de Hilbert réel V muni de différentes choses.

1. *L'application linéaire $L\colon V\to\mathbb{R}$ qui est bornée sur V. Nous notons C sa norme.*
2. *La forme bilinéaire continue a sur $V\times V$. Nous considérons $M>0$ tel que $|a(u,v)|\leqslant M\|u\|\|v\|$ pour tout $u,v\in V$.*
3. *La forme a est coercitive.*

Nous supposons de plus que a est symétrique.

Alors un élément $u\in V$ est tel que $a(u,w)=l(w)$ pour tout $w\in V$ si et seulement si elle minimise la ***fonctionnelle d'énergie***

$$J(w)=\frac{1}{2}a(w,w)-l(w). \tag{23.154}$$

Démonstration. Nous séparons les deux sens.

17. Définition 18.22.

$\Rightarrow$ Soit u, un élément vérifiant $a(u,w)=l(w)$ pour tout $w\in V$. Soit aussi un élément quelconque $w\in V$, et montrons que $J(u+w)\geqslant J(u)$ (tout élément de V peut être écrit sous la forme $u+w$). Nous avons :

$$J(u+w)=\frac{1}{2}\big(a(u,u)+a(u,w)+a(w,u)+a(w,w)\big)-l(u)-l(w) \tag{23.155a}$$

$$=J(u)+\underbrace{a(u,w)-l(w)}_{=0}+\frac{1}{2}a(w,w) \tag{23.155b}$$

$$=J(u)+\frac{1}{2}a(w,w) \tag{23.155c}$$

Vu que a est coercive, le second terme est positif et nous avons

$$J(u+w)\geqslant J(u), \tag{23.156}$$

ce qu'il fallait.

$\Leftarrow$ Soit $u\in V$, un élément minimisant la fonctionnelle J. Nous fixons $w\in V$ et considérons la fonction $g\colon\mathbb{R}\to\mathbb{R}$ définie par

$$g(\epsilon)=J(u+\epsilon w). \tag{23.157}$$

En développant un peu et en regroupant les termes,

$$g(\epsilon)=J(u)+\epsilon\big(a(u,w)-l(w)\big)+\frac{\epsilon^2}{2}a(w,w). \tag{23.158}$$

Cela est une fonction éminemment continue et dérivable ; en réalité c'est un bête polynôme de degré deux. Vu que u minimise J, pour tout $\epsilon\neq 0$ nous avons $g(\epsilon)\geqslant g(0)$ ou encore : $\epsilon=0$ est minimum local (et même global) de g. Le polynôme (23.158) prend son minimum en $\epsilon=0$ si et seulement si

$$a(u,w)-l(w)=0. \tag{23.159}$$

Vous ne me croyez pas ? Faites $g'(\epsilon)=0$ ou bien reprenez la formule du $-b/2a$ pour le sommet d'une parabole, en tenant compte que $a(w,w)\neq 0$. Notons qu'ici encore le fait que a soit coercive joue parce que c'est cela qui nous permet d'affirmer que la parabole a un minimum et non un maximum.

□

Chapitre 24

Analyse fonctionnelle

Proposition 24.1.
Soit $f \in L^1(\mathbb{R})$ une fonction telle que

$$\int_{\mathbb{R}} f(t)\chi(t)dt = 0 \tag{24.1}$$

pour toute fonction $\chi \in \mathscr{D}(\mathbb{R})$. Alors $f = 0$ presque partout.

Théorème 24.2 (Théorème d'isomorphisme de Banach)**.**
Une application linéaire continue et bijective entre deux espaces de Banach est un homéomorphisme.

24.1 Théorème d'Ascoli

Définition 24.3.
*Une partie A d'un espace topologique X est **relativement compacte** dans X si sa fermeture est compacte.*

Proposition 24.4 ([148])**.**
Soient E et F deux espaces vectoriels normés sur $\mathbb{R}$ ou $\mathbb{C}$ et une application $f \in \mathcal{L}(E, F)$. Les propriétés suivantes sont équivalentes.

(1) L'image d'un borné de E par f est relativement compact dans F.

(2) L'image par f de la boule unité fermée est relativement compacte dans F.

(3) Si (x_n) est une suite bornée dans E, alors nous pouvons en extraire une sous-suite $(x_{\varphi(n)})$ telle que $\left(fx_{\varphi(n)}\right)$ converge dans F.

Définition 24.5.
*Une application vérifiant les conditions équivalentes de la proposition 24.4 est dite **compacte**.*

Définition 24.6.
*Soit $(f_i)_{i\in I}$ une famille de fonctions $f_i \colon X \to Y$ entre espaces métriques. Cette famille est **équicontinue** si pour tout $\epsilon > 0$ et pour tout $x \in X$, il existe un $\delta(x, \epsilon) > 0$ tel que*

$$\|x - y\|_X \leqslant \delta \Rightarrow \|f_i(x) - f_i(y)\|_Y \leqslant \epsilon \tag{24.2}$$

pour tout $i \in I$.

Théorème 24.7 (Théorème d'Ascoli[273])**.**
Soit K un espace topologique compact et un espace métrique (E, d). Nous considérons la topologie uniforme sur $C(K, E)$. Une partie A de $C(K, E)$ est relativement compacte si et seulement si les deux conditions suivantes sont remplies :

(1) A est équicontinu,

(2) $\forall x \in K$, l'ensemble $\{f(x)$ tel que $f \in A\}$ est relativement compact dans E.

24.2 Théorème de Banach-Steinhaus

Théorème 24.8 (Théorème de Banach-Steinhaus[2, 274])**.**
Soit E un espace de Banach[1] *et F un espace vectoriel normé. Nous considérons une partie $H \subset \mathcal{L}_c(E,F)$ (espace des fonctions linéaires continues). Alors H est uniformément borné si et seulement s'il est simplement borné.*

Démonstration. Si H est uniformément borné, il est borné ; pas besoin de rester longtemps sur ce sens de l'équivalence. Supposons donc que H soit borné. Pour chaque $k \in \mathbb{N}^*$ nous considérons l'ensemble

$$\Omega_k = \{x \in E \text{ tel que } \sup_{f\in H} \|f(x)\| > k\}. \tag{24.3}$$

Les Ω_k sont ouverts Soit $x_0 \in \Omega_k$; nous avons alors une fonction $f \in H$ telle que $\|f(x_0)\| > k$, et par continuité de f il existe $\rho > 0$ tel que $\|f(x)\| > k$ pour tout $x \in B(x_0,\rho)$. Par conséquent $B(x_0,\rho) \subset \Omega_k$ et Ω_k est ouvert par le théorème 6.3.

Les Ω_k ne sont pas tous denses dans E Nous supposons que les ensembles Ω_k soient tous dense dans E. Le théorème de Baire 6.202 nous indique que E est un espace de Baire (parce que de Banach) et donc que

$$\overline{\bigcap_{k\in\mathbb{N}} \Omega_k} = E. \tag{24.4}$$

En particulier l'intersection des Ω_k n'est pas vide. Soit $x_0 \in \bigcap_{k\in\mathbb{N}} \Omega_k$. Nous avons alors

$$\sup_{f\in H} \|f(x)\| = \infty, \tag{24.5}$$

ce qui est contraire à l'hypothèse. Donc les ouverts Ω_k ne sont pas tous denses dans E.

La majoration Il existe $k \geqslant 0$ tel que Ω_k ne soit pas dense dans E, et nous voulons prouver que $\{\|f\| \text{ tel que } f \in H\}$ est un ensemble borné. Soit donc $k \geqslant 0$ tel que Ω_k ne soit pas dense dans E ; il existe un $x_0 \in E$ et $\rho > 0$ tels que

$$B(x_0,\rho) \cap \Omega_k = \varnothing. \tag{24.6}$$

Si $x \in B(x_0,\rho)$ alors x n'est pas dans Ω_k et donc

$$\sup_{f\in H} \|f(x)\| \leqslant k. \tag{24.7}$$

Afin d'évaluer $\|f\|$ nous devons savoir ce qu'il se passe avec les vecteurs sur une boule autour de 0. Pour tout $x \in B(0,\rho)$ et pour tout $f \in H$, la linéarité de f donne

$$\|f(x)\| = \|f(x+x_0) - f(x_0)\| \leqslant \|f(x+x_0) + f(x_0)\| \leqslant 2k. \tag{24.8}$$

Par continuité nous avons alors $\|f(x)\| \leqslant 2k$ pour tout $x \in \overline{B(0,\rho)}$. Si maintenant $x \in F$ vérifie $\|x\| = 1$ nous avons

$$\|f(x)\| = \frac{1}{\rho}\|f(\rho x)\| \leqslant \frac{2k}{\rho}, \tag{24.9}$$

et donc $\|f\| \leqslant \frac{2k}{\rho}$, ce qui montre que $2k/\rho$ est un majorant de l'ensemble $\{\|f\| \text{ tel que } f \in H\}$.

□

Une application du théorème de Banach-Steinhaus est l'existence de fonctions continues et périodiques dont la série de Fourier ne converge pas. Ce sera l'objet de la proposition 26.18.

1. Définition 23.1.

Corollaire 24.9.
Soient $\big(E,(p_l)\big)$ et $\big(F,(q_k)\big)$ deux espaces localement compacts munis de semi-normes. Nous supposons que E est métrisable et complet. Soit $(T_j)_{j\in\mathbb{N}}$ une suite d'application linéaires $E\to F$ telles que pour tout $x\in E$ il existe $\alpha_x\in F$ tel que

$$T_j x \xrightarrow{F} \alpha_x. \tag{24.10}$$

Si nous posons $Tx=\alpha_x$ alors

(1) l'application T est linéaire et continue,

(2) pour tout compact K dans E et pour tout k nous avons

$$\lim_{j\to\infty}\sup_{x\in K} q_k(T_jx-Tx)=0, \tag{24.11}$$

(3) si $x_j\to x$ dans E alors

$$T_jx_j\to Tx \tag{24.12}$$

dans F.

La version suivante du théorème de Banach-Steinhaus est énoncée de façon *ad hoc* pour fonctionner avec l'espace $\mathscr{D}(K)$ des fonctions de classe C^∞ à support dans le compact K. Un énoncé un peu plus fort est donné dans le cadre des espaces de Fréchet dans [62].

Théorème 24.10 (Banach-Steinhaus avec des semi-normes)**.**
Soit (E,d) un espace vectoriel métrique complet dont la topologie est également[2] *donnée par une famille $\mathcal{P}$ de semi-normes. Soit $\{T_\alpha\}_{\alpha\in A}$ une famille d'applications linéaires continues $T_\alpha\colon E\to\mathbb{R}$ telles que pour tout $x\in E$ nous ayons*

$$\sup_{\alpha\in A}\big|T_\alpha(x)\big|<\infty. \tag{24.13}$$

Alors il existe une constante $C>0$ et un sous-ensemble fini $J\subset\mathcal{P}$ tels que pour tout $x\in E$ nous ayons

$$\big|T_\alpha(x)\big|\leqslant C\max_{j\in J}p_j(x). \tag{24.14}$$

Démonstration. Pour chaque $k\in\mathbb{N}^*$ nous posons

$$\Omega_k=\{x\in E \text{ tel que } \sup_{\alpha\in A}\big|T_\alpha(x)\big|>k\}. \tag{24.15}$$

Ces ensembles sont des ouverts (pour la même raison que dans la preuve du théorème 24.8) et leur union est E en entier parce que par hypothèse $\sup_{\alpha\in A}\big|T_\alpha(x)\big|<\infty$.

Si les Ω_k étaient tous dense, le théorème de Baire 6.202 nous dit que leur intersection est dense également ; elle est donc non vide et si $x_0\in\bigcap_{k\in\mathbb{N}}\Omega_k$ nous avons

$$\sup_{\alpha\in A}\big|T_\alpha(x_0)\big|=\infty, \tag{24.16}$$

ce qui contredirait l'hypothèse. Donc les Ω_k ne sont pas tous denses. Soit $k_0\in\mathbb{N}^*$ tel que Ω_{k_0} n'est pas dense dans E. Il existe donc $x_0\in E$ et un ouvert autour de x_0 n'intersectant pas Ω_{k_0}.

Nous jouons à présent sur la topologie de E. L'ouvert dont il est question est un d-ouvert et donc un $\mathcal{P}$-ouvert, lequel contient une $\mathcal{P}$-boule ouverte. Cette dernière boule n'est pas spécialement une d-boule, mais c'est un d-ouvert.

Il existe dont J fini dans $\mathcal{P}$ et $\rho>0$ tels que $B_J(x_0,\rho)\cap\Omega_{k_0}=\varnothing$. Donc pour tout $y\in B_J(x_0,\rho)$ nous avons

$$\sup_{\alpha\in A}\big|T_\alpha(y)\big|\leqslant k_0. \tag{24.17}$$

2. Au sens où les ouverts sont les mêmes.

Si maintenant $y \in B(0,\rho)$, nous avons $y = (y+x_0) - x_0$ et donc

$$\begin{aligned} \big|T_\alpha(y)\big| &= \big|T_\alpha(y+x_0) - T_\alpha(x_0)\big| && (24.18a)\\ &\leqslant \big|T_\alpha(y+x_0)\big| + \big|T_\alpha(x_0)\big| && (24.18b)\\ &\leqslant k_0 + C && (24.18c)\end{aligned}$$

où nous avons posé $C = \sup_{\alpha\in A}\big|T_\alpha(x_0)\big|$. En normalisant, sur la boule $B_J(0,1)$ nous avons

$$\big|T_\alpha(y)\big| \leqslant \rho^{-1}(k_0+C). \tag{24.19}$$

Enfin su $x \in E$ nous avons

$$\frac{x}{\max_{j\in J} p_j(x)} \in B_J(0,1) \tag{24.20}$$

et donc

$$\left|T_\alpha\left(\frac{x}{\max_{j\in J}}p_j(x)\right)\right| \leqslant \rho^{-1}(k_0+C). \tag{24.21}$$

Utilisant encore la linéarité de T_α nous trouvons ce que nous devions trouver :

$$\big|T_\alpha(x)\big| \leqslant \max_{j\in J} p_j(x)\rho^{-1}(k_0+C) \tag{24.22}$$

à redéfinition près de $\rho^{-1}(k_0+C)$ en C. □

24.3 Espaces L^p

24.3.1 Généralités

Soit $(\Omega,\mathcal{F},\mu)$ un espace mesuré. Deux fonctions à valeurs complexes f et g sur cet espaces sont dites **équivalentes** et nous notons $f\sim g$ si elles sont μ-presque partout égales. Nous notons $[f]$ la classe de f pour cette relation.

Lemme 24.11.
Une classe contient au maximum une seule fonction continue.

Démonstration. Soient deux fonctions continues f_1 et f_2 avec $f_1(a)\neq f_2(a)$. Si $|f_1(a)-f_2(a)|=\delta$ alors il existe un ϵ tel que $|f_1(x)-f_1(a)|<\delta$ pour tout $x\in B(a,\epsilon)$. En particulier $f_1\neq f_2$ sur $B(a,\epsilon)$. Cette dernière boule est de mesure de Lebesgue non nulle ; ergo f_1 et f_2 ne sont pas dans la même classe. □

Nous introduisons l'opération

$$\|f\|_p = \left(\int_\Omega |f(x)|^p d\mu(x)\right)^{1/p} \tag{24.23}$$

et nous notons $\mathcal{L}^p(\Omega,\mu)$ l'ensemble des fonctions mesurables sur Ω telles que $\|f\|_p<\infty$.

Lemme 24.12.
L'ensemble $\mathcal{L}^p$ est un espace vectoriel.

Démonstration. Le fait que si $f\in L^p$, alors $\lambda f\in L^p$ est évident. Ce qui est moins immédiat, c'est le fait que $f+g\in L^p$ lorsque f et g sont dans L^p. Cela découle du fait que la fonction $\varphi\colon x\mapsto x^p$ est convexe, de telle sorte que

$$\varphi\left(\frac{a+b}{2}\right) \leqslant \frac{\varphi(a)+\varphi(b)}{2}, \tag{24.24}$$

ou encore

$$(a+b)^p \leqslant 2^{p-1}(a^p+b^p) \tag{24.25}$$

□

L'opération $f \mapsto \|f\|_p$ n'est pas une norme sur $\mathcal{L}^p$ parce que pour f presque partout nulle, nous avons $|f|_p = 0$. Il y a donc des fonctions non nulles sur lesquelles $\|.\|_p$ s'annule.

Lemme 24.13.
Si $f \in \mathcal{L}^p(\Omega)$ et $f \sim g$, alors $g \in \mathcal{L}^p(\Omega)$ et $\|f\|_p = \|g\|_p$.

Démonstration. Soit $h(x) = |g(x)|^p - |f(x)|^p$; c'est une fonction par hypothèse presque partout nulle et donc intégrable sur Ω ; son intégrale y vaut zéro. Nous avons

$$\int_\Omega |f(x)|^p d\mu(x) = \int_\Omega \Big(|f(x)|^p + h(x)\Big) d\mu(x) = \int_\omega |g(x)|^p d\mu(x). \tag{24.26}$$

Cela prouve que la dernière intégrale existe et vaut la même chose que la première. □

Nous pouvons donc considérer la norme $|.|_p$ comme une norme sur l'ensemble des classes plutôt que sur l'ensemble des fonctions. Nous notons L^p l'ensemble des classes des fonctions de $\mathcal{L}^p$. Cet espace est muni de la norme

$$\|[f]\|_p = \|f\|_p, \tag{24.27}$$

formule qui ne dépend pas du représentant par le lemme 24.13.

Maintenant la formule

$$\|[f]\|_p = \left(\int_\Omega |f(x)|^p d\mu(x) \right)^{1/p} \tag{24.28}$$

défini une norme sur $L^p(\Omega, \mu)$. En effet si $\|[f]\|_p = 0$, nous avons

$$\int_\Omega |f(x)|^p d\mu(x) = 0, \tag{24.29}$$

ce qui par le lemme 13.123 implique que $|f(x)|^p = 0$ pour presque tout x. Ou encore $f \sim 0$, c'est à dire $[f] = [0]$ au niveau des classes. À partir de maintenant $\big(L^p(\Omega, \mu), \|.\|_p\big)$ est un espace métrique avec toute la topologie qui va avec.

Dans la suite nous n'allons pas toujours écrire $[f]$ pour la classe de f. Par abus de notations nous allons souvent parler de $f \in L^p$ comme si c'était une fonction.

Proposition 24.14 ([275])**.**
Soit $1 \leqslant p \leqslant \infty$ et supposons que la suite $[f_n]$ dans $L^p(\Omega, \mathcal{F}, \mu)$ converge vers $[f]$ au sens L^p. Alors il existe une sous-suite (h_n) qui converge ponctuellement μ-presque partout vers f.

Démonstration. Si $p = \infty$ nous sommes en train de parler de la convergence uniforme et il ne faut même pas prendre ni de sous-suite ni de « presque partout ».

Supposons que $1 \leqslant p < \infty$. Nous considérons une sous-suite $[h_n]$ de $[f_n]$ telle que

$$\|[h_j] - [f]\|_p < 2^{-j}, \tag{24.30}$$

puis nous posons $u_k(x) = |h_k(x) - f(x)|^p$. Notons que ce u_k est une vraie fonction, pas une classe. Et en plus c'est une fonction positive. Nous avons

$$\int_\Omega u_k d\mu = \int_\omega |h_k(x) - f(x)|^p d\mu(x) = \|h_k - f\|_p^p \leqslant 2^{-kp}. \tag{24.31}$$

Vu que u_k est une fonction positive la suite des sommes partielles de $\sum_k u_k$ est croissante et vérifie donc le théorème de la convergence monotone 13.129 :

$$\int_\Omega \left(\sum_{k=0}^{\infty} u_k(x) \right) d\mu(x) = \sum_{k=0}^{\infty} \int_\Omega u_k(x) d\mu(x) \leqslant \sum_{k=0}^{\infty} 2^{-kp} < \infty. \tag{24.32}$$

Le fait que l'intégrale de la fonction $\sum_k u_k$ est finie implique que cette fonction est finie μ-presque partout. Donc le terme général tend vers zéro presque partout, c'est à dire

$$|h_k(x) - f(x)|^p \to 0. \tag{24.33}$$

Cela signifie que $h_k \to f$ presque partout ponctuellement. □

Est-ce qu'on peut faire mieux que la convergence ponctuelle presque partout d'une sous-suite ? En tout cas on ne peut pas espérer grand chose comme convergence pour la suite elle-même, comme le montre l'exemple suivant.

Exemple 24.15
Nous allons montrer une suite de fonctions qui converge vers zéro dans $L^p[0,1]$ (avec $p<\infty$) mais qui ne converge ponctuellement pour *aucun* point. Cet exemple provient de bibmath.net.

Nous construisons la suite de fonctions par paquets. Le premier paquet est formé de la fonction constante 1.

Le second paquet est formé de deux fonctions. La première est $\mathbb{1}_{[0,1/2]}$ et la seconde $\mathbb{1}_{[1/2,1]}$.

Plus généralement le paquet numéro k est constitué des k fonctions $\mathbb{1}_{[i/k,(i+1)/k]}$ avec $i=0,\ldots,k-1$.

Vu que les fonctions du paquet numéro k ont pour norme $\|f\|_p=\frac{1}{k}$, nous avons évidemment $f_n\to 0$ dans L^p. Il est par contre visible que chaque paquet passe en revue tous les points de $[0,1]$. Donc pour tout x et pour tout N, il existe (même une infinité) $n>N$ tel que $f_n(x)=1$. Il n'y a donc convergence ponctuelle nulle part. △

La proposition suivante est une espèce de convergence dominée de Lebesgue pour L^p.

Proposition 24.16.
Soit $f\in L^p(\Omega)$ avec $1\leqslant p<\infty$ et (f_n) une suite de fonctions convergeant ponctuellement vers f et telle que $|f_n|\leqslant|f|$. Alors $f_n\xrightarrow{L^p}f$.

Démonstration. Nous avons immédiatement $|f_n(x)|^p\leqslant|f(x)|^p$, de telle sorte que le théorème de la convergence dominée implique que $f_n\in L^p$. La convergence dominée donne aussi que $\|f_n\|_p\to\|f\|_p$, mais cela ne nous intéresse pas ici.

Nous posons $h_n(x)=|f_n(x)-f(x)|$. En reprenant la formule de majoration (24.25) et en tenant compte du fait que $|f_n(x)|\leqslant|f(x)|$, nous avons

$$h_n(x)\leqslant 2^{p-1}\big(|f_n(x)|^p+|f(x)|^p\big)\leqslant 2^p|f(x)|^p, \tag{24.34}$$

ce qui prouve que $|h_n|$ est uniformément (en n) majorée par une fonction intégrable, donc h_n est intégrable et on peut permuter la limite et l'intégrale (théorème de la convergence dominée 13.134) :

$$\lim_{n\to\infty}\|f_n-f\|_p^p=\lim_{n\to\infty}\int_{\mathbb{R}^d}|f_n(x)-f(x)|^p dx=\int_{\mathbb{R}^d}\lim_{n\to\infty}h_n(x)dx=0. \tag{24.35}$$

□

Proposition 24.17.
Soit un espace mesuré σ-fini $(\Omega,\mathcal{A},\mu)$ et une fonction $f\in L^p(\Omega)$ telle que

$$\int_\Omega f\varphi=0 \tag{24.36}$$

pour tout $\varphi\in L^q(\Omega)$. Alors $f=0$.

Démonstration. Vu que Ω est σ-finie, nous pouvons considérer des parties K_n de Ω telles que $\mu(K_n)<\infty$ et $\bigcup_n K_n=\Omega$. Nous posons

$$A_n=\{x\in\Omega\text{ tel que }\operatorname{Re}\big(f(x)\big)>0\}\operatorname{cat}K_n. \tag{24.37}$$

Cela est une partie mesurable de mesure finie de Ω, donc $\varphi_n=\mathbb{1}_{A_n}\in L^q(\Omega)$ et nous avons

$$\int_{A_n}f=0. \tag{24.38}$$

Mais $\operatorname{Re}(f)>0$ sur A_n, donc $\mu(A_n)=0$. Ensuite nous passons à l'union :

$$A=\{x\in\Omega\text{ tel que }\operatorname{Re}(f(x))>0\}=\bigcup_n A_n. \tag{24.39}$$

Ce ensemble est union dénombrable d'ensembles de mesure nulle ; donc $\mu(A) = 0$. Nous faisons de même pour les ensembles $\operatorname{Re}(f) < 0$, $\operatorname{Im}(f) > 0$ et $\operatorname{Im}(f) < 0$. Au final, l'ensemble $\{x \in \Omega \text{ tel que } f(x) \neq 0\}$ est de mesure nulle, c'est à dire que $f = 0$ au sens des classes de L^p. □

24.3.2 L'espace L^∞

Il n'est pas possible de définir le supremum d'une fonction définie à ensemble de mesure nulle près parce que toute classe contient des fonctions qui peuvent être arbitrairement grandes en n'importe que point. Nous cherchons alors à définir une notion de supremum qui ne tient pas compte des ensembles de mesure nulle.

Définition 24.18.
*Soit $f : \Omega \to \mathbb{C}$. Un nombre M est un **majorant essentiel** de f si*

$$\mu\big(|f(x)| \leqslant M\big) = 0. \tag{24.40}$$

Nous posons alors $N_\infty(f) = \inf\{M \text{ tel que } |f(x)| \leqslant M \text{ presque partout}\}$. Cela revient à prendre le supremum à ensemble de mesure nulle près. Nous définissons alors les espaces de Lebesgue correspondants :

$$\mathcal{L}^\infty(\Omega) = \{f : \Omega \to \mathbb{C} \text{ tel que } N_\infty(f) < \infty\}, \tag{24.41}$$

et L^∞ en est le quotient usuel. Dans ce contexte nous notons $\|f\|_\infty$ le supremum essentiel de f, qui est indépendant de la classe.

24.3.3 Quelque identifications

Il est intuitivement clair que ce qui peut arriver à une fonction en un seul point ne va pas influencer la fonction lorsqu'elle est vue dans L^p. En tout cas lorsqu'on considère des mesures pour lesquelles les singletons sont de mesure nulle, et c'est bien le cas de la mesure de Lebesgue. Il est peut-être intuitivement moins clair que l'on peut non seulement modifier le comportement d'une fonction en un point, mais également modifier l'ensemble de base. En voici un exemple.

Proposition 24.19.
Nous avons les égalités suivantes d'espaces

$$L^p\big(]0, 2\pi[\big) = L^p\big([0, 2\pi]\big) = L^p(S^1) \tag{24.42}$$

au sens où il existe des bijections isométriques de l'un à l'autre. Ici nous sous-entendons la mesure de Lebesgue partout[3].

Démonstration. Voici une application bien définie où le crochet dénote la prise de classe :

$$\begin{aligned} \psi : L^p\big(]0, 2\pi[\big) &\to L^p\big([0, 2\pi]\big) \\ [f] &\mapsto \text{la classe de } f_e(x) = \begin{cases} f(x) & \text{si } x \in]0, 2\pi[\\ 0 & \text{si } x = 0 \text{ ou } x = 2\pi. \end{cases} \end{aligned} \tag{24.43}$$

Injective Si $[f] = [g]$ dans $L^p\big(]0, 2\pi[\big)$ alors $f_e(x) = g_e(x)$ pour tout $x \in]0, 2\pi[$ sauf une partie de mesure nulle. L'union de cette partie avec $\{0, 2\pi\}$ est encore de mesure nulle dans $[0, 2\pi]$. Les images par ψ sont donc égales dans $L^p\big([0, 2\pi]\big)$.

Surjective Un élément de $L^p\big([0, 2\pi]\big)$ est l'image de sa restriction ... ou plutôt l'image de la classe de la restriction d'un quelconque de ses représentants.

Isométrie L'intégrale qui donne la norme sur L^p ne change pas selon que nous ajoutions ou non les bornes au domaine d'intégration.

3. Vu que la mesure de Lebesgue est définie pour $\mathbb{R}^d$ munie de sa tribu des boréliens (complétée), vous êtes en droit de vous demander quelle est la tribu et la mesure que nous considérons sur le cercle S^1.

De la même manière nous avons

$$L^p\big([0,2\pi[\big) = L^p\big([0,2\pi]\big). \tag{24.44}$$

En ce qui concerne l'identification avec $L^p(S^1)$, il faut passer par l'isométrie $\varphi\colon [0,2\pi[\to S^1$ donnée par $\varphi(t) = e^{it}$, et être heureux que ce soit bien une isométrie parce qu'il faudra l'utiliser pour un changement de variables pour montrer que

$$\int_0^{2\pi} f(t)dt = \int_{S^1} (f\circ\varphi^{-1})(z)dz. \tag{24.45}$$

□

24.3.4 Inégalité de Jensen, Hölder et de Minkowski

Proposition 24.20 (Inégalité de Jensen[179])**.**
Soit un espace mesuré de probabilité[4] $(\Omega,\mathcal{A},\mu)$ *ainsi qu'une fonction convexe* $f\colon \mathbb{R}\to\mathbb{R}$ *et une application* $\alpha\colon \Omega\to\mathbb{R}$ *tels que* α *et* $f\circ\alpha$ *soient intégrables sur* Ω*. Alors*

$$f\Big(\int_\Omega \alpha\, d\mu\Big) \leqslant \int_\Omega (f\circ\alpha)d\mu. \tag{24.46}$$

Démonstration. Soit $a\in\mathbb{R}$ et le nombre c_a donné par la proposition 18.85 : pour tout $\omega\in\Omega$ nous avons

$$f\big(\alpha(\omega)\big) - f(a) \geqslant c_a\big(\alpha(\omega)-a\big). \tag{24.47}$$

Cela est en particulier vrai pour $a = \int_\Omega \alpha\, d\mu$. Nous intégrons l'inégalité (24.47) sur Ω en nous souvenant que $\int d\mu = 1$:

$$\int_\Omega (f\circ\alpha)d\mu - \int_\Omega f(a)d\mu \geqslant c_a\Big(\int_\Omega \alpha - \int_\Omega a\Big) \tag{24.48a}$$

$$\int_\Omega (f\circ\alpha)d\mu - f(a) \geqslant c_a(a-a) \tag{24.48b}$$

$$f(a) \leqslant \int_\Omega (f\circ\alpha)d\mu. \tag{24.48c}$$

Cette dernière inégalité est celle que nous devions prouver. □

Corollaire 24.21.
Soit un espace mesuré de probabilité $(\Omega,\mathcal{A},\mu)$ *et une application* $\alpha\in L^1(\Omega,\mu)$ *et* $\alpha\in L^p(\Omega)$ *avec* $1\leqslant p<+\infty$*. Alors*

$$\Big|\int_\Omega \alpha(s)d\mu(s)\Big| \leqslant \|\alpha\|_p. \tag{24.49}$$

Démonstration. Il suffi d'utiliser l'inégalité de Jensen sur la fonction convexe $f(x) = |x|^p$. Nous avons alors

$$\Big|\int_\Omega \alpha(s)d\mu(s)\Big|^p \leqslant \int_\Omega |\alpha(s)|^p d\mu(s), \tag{24.50}$$

c'est à dire

$$\Big|\int_\Omega \alpha(s)d\mu(s)\Big| \leqslant \left[\int_\Omega |\alpha(s)|^p d\mu(s)\right]^{1/p} = \|\alpha\|_p \tag{24.51}$$

où ma norme $\|.\|_p$ est prise au sens de la mesure μ. □

4. C'est à dire que $\int_\Omega d\mu = 1$.

Proposition 24.22 (Inégalité de Hölder)**.**
Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et $1 \leqslant p$, $q \leqslant \infty$ satisfaisant $\frac{1}{p} + \frac{1}{q} = 1$. Si $f \in L^p(\Omega)$, $g \in L^q(\Omega)$, alors le produit fg est dans $L^1(\Omega)$ et nous avons

$$\|fg\|_1 \leqslant \|f\|_p \|g\|_q. \tag{24.52}$$

Plus généralement si $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$ alors

$$\|fg\|_r \leqslant \|f\|_p \|g\|_q \tag{24.53}$$

Ce qui est appelé « inégalité de Hölder » est généralement l'inéquation (24.52).

Démonstration. Nous allons prouver l'inégalité (24.53). D'abord nous supposons $\|g\|_q = 1$ et nous posons

$$A = \{x \in \Omega \text{ tel que } |g(x)| > 0\}. \tag{24.54}$$

Hors de A, les intégrales que nous allons écrire sont nulles. Nous avons

$$\|fg\|_r^p = \Big| \int_A |f|^r |g|^{r-q} |g|^q \Big|^{p/r}, \tag{24.55}$$

et le coup tordu est de considérer cette intégrale comme étant une intégrale par rapport à la mesure $\nu = |g|^q d\mu$ qui a la propriété d'être une mesure de probabilité par hypothèse sur g. Nous pouvons alors utiliser l'inégalité de Jensen [5] parce que $p/r > 1$, ce qui fait de $x \mapsto |x|^{p/r}$ une fonction convexe. Nous avons alors

$$\|fg\|_r^p \leqslant \int_A \big(|f|^r |g|^{r-q}\big)^{p/r} |g|^q d\mu \tag{24.56a}$$

$$= \int_A |f|^p |g|^{p(r-q)/r} |g|^q d\mu \tag{24.56b}$$

La puissance de $|g|$ dans cette expression est : $q + \frac{p(r-q)}{r} = 0$ parce que $p(q-r) = rq$. Nous avons alors montré que

$$\|fg\|_r^p \leqslant \int_A |f|^p d\mu \leqslant \|f\|_p^p. \tag{24.57}$$

La dernière inégalité est le fait que le domaine A n'est pas tout le domaine Ω.

Si maintenant $\|g\|_q \neq 1$ alors nous calculons

$$\|fg\|_r = \|g\|_q \|f \frac{g}{\|g\|_q}\|_r \leqslant \|g\|_q \|f\|_p \tag{24.58}$$

en appliquant la première partie à la fonction $\frac{g}{\|g\|_q}$ qui est de norme 1. □

Remarque 24.23.
Dans le cas d'un espace de probabilité, la fonction constante $g = 1$ appartient à $L^p(\Omega)$. En prenant $p = q = 2$ nous obtenons

$$\|f\|_1 \leqslant \|f\|_2. \tag{24.59}$$

Lemme 24.24.
Lorsque I est borné nous avons $L^2(I) \subset L^1(I)$. Si I n'est pas borné alors $L^2(I) \subset L^1_{loc}(I)$.

Démonstration. En effet si I est borné, alors la fonction constante 1 est dans $L^2(I)$ et l'inégalité de Hölder 24.22 nous dit que le produit $1u$ est dans $L^1(I)$.

Si I n'est pas borné, nous refaisons le même raisonnement sur un compact K de I. □

5. Proposition 24.20.

Corollaire 24.25 ([276]).
Soit l'espace $L^2(I)$ avec $I =]0,1[$ avec la mesure de Lebesgue. Si $u_n \in L^2$ converge vers u dans L^2 alors nous pouvons permuter l'intégrale et la limite :

$$\lim_{n\to\infty} \int_I u_n = \int_I u. \tag{24.60}$$

Démonstration. Nous considérons la forme linéaire

$$\begin{aligned} T\colon L^2(I) &\to \mathbb{R} \\ u &\mapsto \int_I u. \end{aligned} \tag{24.61}$$

Elle est bien définie par l'inégalité de Hölder $\|fg\|_1 \leqslant \|f\|_2\|g\|_2$ appliqué à $g(x) = 1$ qui vérifie $\|g\|_2 = 1$. Nous avons aussi

$$T(u) \leqslant \int_I |u| \leqslant \|u\|_1 \leqslant \|u\|_2 \tag{24.62}$$

où la dernière inégalité est celle de Hölder 24.22. Bref, T est continue. Cela signifie que si $u_n \xrightarrow{L^2(I)} u$ alors $T(u_n) = T(u)$. Cela est l'égalité demandée. □

Proposition 24.26 (Inégalité de Minkowski[277, 278]).
Si $1 \leqslant p < \infty$ et si $f, g \in L^p(\Omega, \mathcal{A}, \mu)$ alors

(1) $\|f + g\|_p \leqslant \|f\|_p + \|g\|_p$

(2) Il y a égalité si et seulement si les vecteurs $f(x)$ et $g(x)$ sont presque partout colinéaires : il existe α, β tels que $\alpha f + \beta g = 0$ presque partout.

(3) Si $f(x,y)$ est mesurable sur l'espace produit $\big(X \times Y, \mu \otimes \nu\big)$ et si $p \geqslant 1$, alors

$$\left\| x \mapsto \int_Y f(x,y) d\nu(y) \right\|_p \leqslant \int_Y \|f_y\|_p d\nu(y) \tag{24.63}$$

où $f_y(x) = f(x,y)$.

La partie (3) est une généralisation de l'inégalité triangulaire (c'est à dire du point (1)) dans le cas où nous n'avons pas une somme de deux fonctions mais d'une infinité paramétrée par $y \in Y$. Elle sera le plus souvent utilisée sous la forme déballée :

$$\left[\int_X \Big(\int_Y |f(x,y)| d\nu(y) \Big)^p d\mu(x) \right]^{1/p} \leqslant \int_Y \Big(\int_X |f(x,y)|^p d\mu(x) \Big)^{1/p} d\nu(y). \tag{24.64}$$

24.3.5 Ni inclusions ni inégalités

Aucun espace $L^p(\mathbb{R})$ n'est inclus dans aucun autre ni aucune norme n'est plus grande qu'une autre (sur les intersections). Nous verrons cependant en la proposition 24.27 que de telles inclusions et inégalités sont possibles pour $L^p\big([0,1]\big)$.

Nous allons donner des exemples de tout ça en supposant $p < q$ et en nous appuyant lourdement sur les intégrales de $\frac{1}{x^\alpha}$ étudiées par la proposition 13.220.

<u>$L^p \not\subseteq L^q$</u> La fonction

$$f(x) = \begin{cases} \frac{1}{x^{1/q}} & \text{si } 0 < x < 1 \\ 0 & \text{sinon} \end{cases} \tag{24.65}$$

est dans L^p mais pas dans L^q. En effet

$$\|f\|_p^p = \int_0^1 \frac{1}{x^{p/q}} dx < \infty \tag{24.66}$$

parce que $p < q$ et $p/q < 1$. Par contre

$$\|f\|_q^q = \int_0^1 \frac{1}{x} dx = \infty. \tag{24.67}$$

$\underline{L^q \not\subseteq L^p}$ La fonction

$$f(x) = \begin{cases} \frac{1}{x^{1/p}} & \text{si } x > 1 \\ 0 & \text{sinon} \end{cases} \tag{24.68}$$

est dans L^q mais pas dans L^p. En effet

$$\|f\|_p^p = \int_1^\infty \frac{1}{x} = \infty \tag{24.69}$$

alors que

$$\|f\|_q^q = \int_1^\infty \frac{1}{x^{q/p}} dx < \infty. \tag{24.70}$$

Exemple de $\|f\|_p > \|f\|_q$ La fonction

$$f(x) = \begin{cases} 1 & \text{si } x \in [0,2] \\ 0 & \text{sinon.} \end{cases} \tag{24.71}$$

Nous avons

$$\|f\|_p = 2^{1/p} \tag{24.72a}$$
$$\|f\|_q = 2^{1/q}. \tag{24.72b}$$

Mais comme $p < q$ donc $\|f\|_p > \|f\|_q$.

Exemple de $\|f\|_p < \|f\|_q$ La fonction

$$f(x) = \begin{cases} 1 & \text{si } x \in [0, \frac{1}{2}] \\ 0 & \text{sinon.} \end{cases} \tag{24.73}$$

Alors

$$\|f\|_p = \frac{1}{2^{1/p}} \tag{24.74a}$$
$$\|f\|_q = \frac{1}{2^{1/q}} \tag{24.74b}$$

et donc $\|f\|_p < \|f\|_q$.

Ces exemples donnent un exemple de fonction f telle que $\|f\|_p < \|f\|_q$ pour tout espace $L^p(I)$ et $L^q(I)$ avec $I \subset \mathbb{R}$. Par contre l'exemple $\|f\|_p > \|f\|_q$ ne fonctionne que si la taille de I est plus grande que 1. Et pour cause : il y a des inclusions si I est borné.

Proposition 24.27 ([7])**.**
Inclusions et inégalités dans le cas d'un ensemble de mesure finie.

(1) Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré fini et $1 \leqslant p \leqslant +\infty$. Alors $L^q(\Omega) \subset L^p(\Omega)$ dès que $p \leqslant q$.

(2) Si $1 < p < 2$ et si $f \in L^2\big([0,1]\big)$ alors $\|f\|_p \leqslant \|f\|_2$.

Démonstration. Pour la simplicité des notations nous allons noter L^p pour $L^p(\Omega)$, et pareillement pour L^q. Soit $f \in L^q$. Nous posons

$$A = \{x \in [0,1] \text{ tel que } |f(x)| \geqslant 1\}. \tag{24.75}$$

Étant donné que $p \leqslant q$ nous avons $|f|^p \leqslant |f|^q$ sur A ; par conséquent $\int_A |f|^p$ converge parce que $\int_A |f|^q$ converge.

L'ensemble A^c est évidemment borné (complémentaire dans Ω) et sur A^c nous avons $|f(x)| \leqslant 1$ et donc $|f|^p \leqslant 1$. L'intégrale $\int_{A^c} |f|^p$ converge donc également.

Au final $\int_\Omega |f|^p$ converge et $f \in L^p$.

Soit à présent $f \in L^2$; par le premier point nous avons immédiatement $f \in L^2 \cap L^p$. Soit aussi $r \in \mathbb{R}$ tel que $\frac{1}{2/p} + \frac{1}{r} = 1$. Nous avons $|f|^p \in L^{2/p}$, et vu que nous sommes sur un domaine borné, $1 \in L^r$. Nous écrivons l'inégalité de Hölder (24.52) avec ces fonctions. D'une part

$$\|f\|_1 = \||f|^p\|_1 = \|f\|_p^p. \tag{24.76}$$

D'autre part

$$\||f|^p\|_{2/p} = \left(\int |f|^2\right)^{p/2} = \|f\|_2^p. \tag{24.77}$$

Donc $\|f\|_p^p \leqslant \|f\|_2^p$, ce qui prouve l'assertion (2) parce que $p > 1$. □

Remarque 24.28.
Nous n'avons cependant pas $L^2\big([0,1]\big) = L^p\big([0,1]\big)$ parce que l'exemple (24.65) fonctionne encore :

$$f(x) = \frac{1}{\sqrt{x}} \tag{24.78}$$

pour $x \in [0,1]$ donne bien

$$\|f\|_2 = \int_0^1 \frac{1}{x} = \infty \tag{24.79}$$

et $\|f\|_p = \int_0^1 \frac{1}{x^{p/2}} < \infty$ parce que $1 < p < 2$.

24.3.6 Complétude

Théorème 24.29 ([279, 280])**.**
Pour $1 \leqslant p < \infty$, l'espace $L^p(\Omega, \mathcal{A}, \mu)$ est complet.

Démonstration. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans L^p. Pour tout i, il existe $N_i \in \mathbb{N}$ tel que $\|f_p - f_q\|_p \leqslant 2^{-i}$ pour tout $p, q \geqslant N_i$. Nous considérons la sous suite $g_i = f_{N_i}$, de telle sorte qu'en particulier

$$\|g_i - g_{i-1}\|_p \leqslant 2^{-i}. \tag{24.80}$$

Pour chaque j nous considérons la somme télescopique

$$g_j = g_0 + \sum_{i=1}^{j} (g_i - g_{i-1}) \tag{24.81}$$

et l'inégalité

$$|g_j| \leqslant |g_0| + \sum_{i=1}^{j} |g_i - g_{i-1}|. \tag{24.82}$$

Nous allons noter

$$h_j = |g_0| + \sum_{i=1}^{j} |g_i - g_{i-1}|. \tag{24.83}$$

La suite de fonctions (h_j) ainsi définie est une suite croissante de fonctions positive qui converge donc (ponctuellement) vers une fonction h qui peut éventuellement valoir l'infini en certains points. Par continuité de la fonction $x \mapsto x^p$ nous avons

$$\lim_{j\to\infty} h_j^p = h^p, \tag{24.84}$$

puis par le théorème de la convergence monotone (théorème 13.129) nous avons

$$\lim_{j\to\infty} \int_\Omega h_j^p d\mu = \int_\Omega h^p d\mu. \tag{24.85}$$

Utilisant à présent la continuité de la fonction $x \mapsto x^{1/p}$ nous trouvons

$$\lim_{j\to\infty} \left(\int h_j^p \right)^{1/p} = \left(\int |h|^p \right)^{1/p}. \tag{24.86}$$

Nous avons donc déjà montré que

$$\lim_{j\to\infty} \|h_j\|_p = \left(\int |h|^p \right)^{1/p} \tag{24.87}$$

où, encore une fois, rien ne garantit à ce stade que l'intégrale à droite soit un nombre fini. En utilisant l'inégalité de Minkowski (proposition 24.26) et l'inégalité (24.80) nous trouvons

$$\|h_j\|_p \leqslant \|g_0\|_p + \sum_{i=1}^{j} \|g_i - g_{i-1}\|_p \leqslant \|g_0\|_p + 1. \tag{24.88}$$

En passant à la limite,

$$\left(\int |h|^p \right)^{1/p} = \lim_{j\to\infty} \|h_j\|_p \leqslant \|g_0\|_p + 1 < \infty. \tag{24.89}$$

Par conséquent $\int |h|^p$ est finie et

$$h \in L^p(\Omega, \mathcal{A}, \mu). \tag{24.90}$$

En particulier, l'intégrale $\int h$ est finie (parce que $p \geqslant 1$) et donc que $h(x) < \infty$ pour presque tout $x \in \Omega$.

Nous savons que $h(x)$ est la limite des sommes partielles (24.83), en particulier la série

$$\sum_{j=1}^{\infty} |g_i - g_{i-1}| \tag{24.91}$$

converge ponctuellement. En vertu du corollaire 13.136, la série de terme général $g_i - g_{i-1}$ converge ponctuellement. La suite g_i converge donc vers une fonction que nous notons g. Par ailleurs la suite g_i est dominée par $h \in L^p$, le théorème de la convergence dominée (théorème 13.134) implique que

$$\lim_{j\to\infty} \|g_j - g\|_p = 0. \tag{24.92}$$

Nous allons maintenant prouver que $\lim_{n\to\infty \|f_n - g\|_p} = 0$. Soit $\epsilon > 0$. Pour tout n et i nous avons

$$\|f_n - g\|_p = \|f_n - f_{N_i} + f_{N_i} - g\|_p \leqslant \|f_n - f_{N_i}\|_p + \|f_{N_i} - g\|_p. \tag{24.93}$$

Pour rappel, $f_{N_i} = g_i$. Si i et n sont suffisamment grands nous pouvons obtenir que chacun des deux termes est plus petit que $\epsilon/2$.

Il nous reste à prouver que $g \in L^p(\Omega, \mathcal{A}, \mu)$. Nous avons déjà vu (équation (24.90)) que $h \in L^p$, mais $|g_i| \leqslant h^p$, par conséquent $g \in L^p$.

Nous avons donc montré que la suite de Cauchy (f_n) converge vers une fonction de L^p, ce qui signifie que L^p est complet. □

Théorème 24.30 (Fischer-Riesz[2])**.**
Soit un ouvert Ω de $\mathbb{R}^n$ et $p \in [1, \infty]$. Alors

1. *Toute suite convergente dans $L^p(\Omega)$ admet une sous-suite convergente presque partout sur Ω.*
2. *La sous-suite donnée en (1) est dominée par un élément de $L^p(\Omega)$.*
3. *L'espace $L^p(\Omega)$ est de Banach.*

Démonstration. Le cas $p = \infty$ est à séparer des autres valeurs de p parce qu'on y parle de norme uniforme, et aucune sous-suite n'est à considérer.

Cas $p = \infty$. Nous commençons par prouver dans le cas $p = \infty$. Soit (f_n) une suite de Cauchy dans $L^\infty(\Omega)$, ou plus précisément une suite de représentants d'éléments de L^p. Pour tout $k \geqslant 1$, il existe $N_k \geqslant 0$ tel que si $m, n \geqslant N_k$, on a

$$\|f_m - f_n\|_\infty \leqslant \frac{1}{k}. \tag{24.94}$$

En particulier, il existe un ensemble de mesure nulle E_k sur lequel

$$|f_m(x) - f_n(x)| \leqslant \frac{1}{k}, \tag{24.95}$$

et si nous posons $E = \bigcup_{k\in\mathbb{N}} E_k$, nous avons encore un ensemble de mesure nulle (lemme 13.34). En résumé, nous avons un N_k tel que si $m, n \geqslant N_k$, alors

$$|f_n(x) - f_m(x)| \leqslant \frac{1}{k} \tag{24.96}$$

pour tout x hors de E. Donc pour chaque $x \in \Omega\backslash E$, la suite $n \mapsto f_n(x)$ est de Cauchy dans $\mathbb{R}$ et converge donc. Cela défini donc une fonction

$$\begin{aligned} f\colon \Omega\backslash E &\to \mathbb{R} \\ x &\mapsto \lim_{n\to\infty} f_n(x). \end{aligned} \tag{24.97}$$

Cela prouve le point (1) : la convergence ponctuelle.

En passant à la limite $n \to \infty$ dans l'équation 24.96 et tenant compte que cette majoration tient pour presque tout x dans Ω, nous trouvons

$$\|f - f_n\|_\infty \leqslant \frac{1}{k}. \tag{24.98}$$

Donc non seulement f est dans L^∞, mais en plus la suite (f_n) converge vers f au sens L^∞, c'est à dire uniformément. Cela prouve le point (3). En ce qui concerne le point (2), la suite f_n est entièrement (à partir d'un certain point) dominée par la fonction $1 + |f|$ qui est dans L^p.

Cas $p < \infty$. Toute suite convergente étant de Cauchy, nous considérons une suite de Cauchy (f_n) dans $L^p(\Omega)$ et ce sera suffisant pour travailler sur le premier point. Pour montrer qu'une suite de Cauchy converge, il est suffisant de montrer qu'une sous-suite converge. Soit $\varphi\colon \mathbb{N} \to \mathbb{N}$ une fonction strictement croissante telle que pour tout $n \geqslant 1$ nous ayons

$$\|f_{\varphi(n+1)} - f_{\varphi(n)}\|_p \leqslant \frac{1}{2^n}. \tag{24.99}$$

Pour créer la fonction φ, il est suffisant de prendre le N_k donné par la condition de Cauchy pour $\epsilon = 1/2^k$ et de considérer la fonction définie par récurrence par $\varphi(1) = N_1$ et $\varphi(n+1) > \max\{N_n, \varphi(n-1)\}$. Ensuite nous considérons la fonction

$$g_n(x) = \sum_{k=1}^{n} |f_{\varphi(k+1)}(x) - f_{\varphi(k)}(x)|. \tag{24.100}$$

Notons que pour écrire cela nous avons considéré des représentants f_k qui sont alors des fonctions à l'ancienne. Étant donné que g_n est une somme de fonctions dans L^p, c'est une fonction L^p, comme nous pouvons le constater en calculant sa norme :

$$\|g_n\|_p \leqslant \sum_{k=1}^{n} \|f_{\varphi(k+1)} - f_{\varphi(k)}\|_p \leqslant \sum_{k=1}^{n} \frac{1}{2^k} \leqslant \sum_{k=1}^{\infty} \frac{1}{2^k} = 1. \tag{24.101}$$

Étant donné que tous les termes de la somme définissant g_n sont positifs, la suite (g_n) est croissante. Mais elle est bornée en norme L^p et donc sujette à obéir au théorème de Beppo-Levi 13.129 sur la convergence monotone. Il existe donc une fonction $g \in L^p(\Omega)$ telle que $g_n \to g$ presque partout.

Soit un $x \in \Omega$ pour lequel $g_n(x) \to g(x)$; alors pour tout $n \geqslant 2$ et $\forall q \geqslant 0$,

$$|f_{\varphi(n+q)}(x) - f_{\varphi(n)}(x)| = \left| f_{\varphi(n+q)}(x) - \sum_{k=1}^{q-1} f_{\varphi(n+k)}(x) - \sum_{k=1}^{q-1} f_{\varphi(n+k)}(x) - f_{\varphi(n)}(x) \right| \tag{24.102a}$$

$$= \left| \sum_{k=1}^{q} f_{\varphi(n+k)} - \sum_{k=1}^{q} f_{\varphi(n+k-1)}(x) \right| \tag{24.102b}$$

$$\leqslant \sum_{k=1}^{q} \left| f_{\varphi(n+k)}(x) - f_{\varphi(n+k-1)}(x) \right| \tag{24.102c}$$

$$= g_{n+q+1}(x) - g_{n+1}(x) \tag{24.102d}$$

$$\leqslant g(x) - g_{n-1}(x). \tag{24.102e}$$

Nous prenons la limite $n \to \infty$; la dernière expression tend vers zéro et donc

$$|f_{\varphi(n+q)}(x) - f_{\varphi(n)}(x)| \to 0 \tag{24.103}$$

pour tout q. Donc pour presque tout $x \in \Omega$, la suite $n \mapsto f_{\varphi(n)}(x)$ est de Cauchy dans $\mathbb{R}$ et donc y converge vers un nombre que nous nommons $f(x)$. Cela définit une fonction

$$\begin{aligned} f \colon \Omega \backslash E &\to \mathbb{R} \\ x &\mapsto \lim_{n\to\infty} f_{\varphi(n)}(x) \end{aligned} \tag{24.104}$$

où E est de mesure nulle. Montrons que f est bien dans $L^p(\Omega)$; pour cela nous complétons la série d'inégalités (24.102) en

$$\big|f_{\varphi(n+q)}(x) - f_{\varphi(n)}(x)\big| \leqslant g(x) - g_{n-1}(x) \leqslant g(x). \tag{24.105}$$

En prenant la limite $q \to \infty$ nous avons l'inégalité

$$|f(x) - f_{\varphi(n)}(x)| \leqslant g(x) \tag{24.106}$$

pour presque tout $x \in \Omega$, c'est à dire pour tout $x \in \Omega \backslash E$. Cette inégalité implique deux choses valables pour presque tout x dans Ω :

$$f(x) \in B\big(g(x), f_{\varphi(n)}(x)\big) \tag{24.107a}$$

$$f_{\varphi(n)}(x) \leqslant |f(x)| + |g(x)|. \tag{24.107b}$$

La première inégalité assure que $|f|^p$ est intégrable sur $\Omega \backslash E$ parce que $|f|$ est majorée par $|g| + |f_{\varphi(n)}|$. Elle prouve par conséquent le point (1) parce que $n \mapsto f_{\varphi(n)}$ est une sous-suite convergente presque partout. La seconde montre le point (2).

Attention : à ce point nous avons prouvé que $n \mapsto f_{\varphi(n)}$ est une suite de fonctions qui converge *ponctuellement presque partout* vers une fonction f qui s'avère être dans L^p. Nous n'avons pas montré que cette suite convergeait au sens de L^p vers f. Ce que nous devons montrer est que

$$\|f - f_{\varphi(n)}\|_p \to 0. \tag{24.108}$$

L'inégalité (24.106) nous donne aussi, toujours pour presque tout $x \in \Omega$:

$$\big|f(x) - f_{\varphi(n)}(x)\big|^p \leqslant g(x)^p \tag{24.109}$$

ce qui signifie que la suite[6] $|f - f_{\varphi(n)}|^p$ est dominée par la fonction $|g|^p$ qui est intégrable sur $\Omega \backslash E$ et tout autant sur Ω parce que E est négligeable ; cela prouve au passage le point (2), et le théorème de la convergence dominée de Lebesgue (13.134) nous dit que

$$\lim_{n\to\infty} \int_\Omega |f(x) - f_{\varphi(n)}(x)|^p dx = \int_\Omega \lim_{n\to\infty} \big|f(x) - f_{\varphi(n)}(x)\big| dx = 0. \tag{24.110}$$

6. À ce point, [2] se contente de majorer $|f_{\varphi(n)}(x)|$ par $|f(x)| + |g(x)$, mais je ne comprends pas comment cette majoration nous permet d'utiliser la convergence dominée de Lebesgue pour montrer (24.108).

Cette dernière suite d'égalités se lit de la façon suivante :

$$\lim_{n\to\infty} \|f - f_{\varphi(n)}\|_p = \Big\| \lim_{n\to\infty} |f - f_{\varphi(n)}| \Big\|_p = 0. \tag{24.111}$$

Nous en déduisons que la suite $n \mapsto f_{\varphi(n)}$ est convergente vers f au sens de la norme $L^p(\Omega)$. Or la suite de départ (f_n) était de Cauchy (pour la norme L^p) ; donc l'existence d'une sous-suite convergente implique la convergence de la suite entière vers f, ce qu'il fallait démontrer.

□

24.3.7 Densité des fonctions infiniment dérivables à support compact

Définition 24.31.
*Une fonction est **étagée par rapport à** L^p si elle est de la forme*

$$f = \sum_{k=1}^{N} c_k \mathbb{1}_{B_k} \tag{24.112}$$

où les B_k sont des mesurables disjoints et $\mathbb{1}_{B_k} \in L^p$ pour tout k.

Lemme 24.32.
Si f est une fonction étagée en même temps qu'être dans L^p, alors elle est étagée par rapport à L^p.

Démonstration. Nous pouvons écrire

$$f = \sum_{k=1}^{N} c_k \mathbb{1}_{B_k} \tag{24.113}$$

où les B_k sont disjoints. Par hypothèse $\|f\|_p$ existe. Donc chacune des intégrales $\int_\Omega |\mathbb{1}_{B_k}|^p$ doit exister parce que les B_k étant disjoints, nous pouvons inverser la norme et la somme ainsi que la somme et l'intégrale :

$$\int_\Omega |f|^p = \int_\Omega \sum_{k=1}^{N} |c_k \mathbb{1}_{B_k}(x)|^p dx = \sum_{k=1}^{N} \int |c_k \mathbb{1}_{B_k}(x)|^p dx = \sum_{k=1}^{N} |c_k|^p \int_\Omega |\mathbb{1}_{B_k}(x)|^p dx. \tag{24.114}$$

□

Le contraire n'est pas vrai : la fonction étagée sur $\mathbb{R}$ qui vaut n sur $B(n, \frac{1}{4})$ est étagée par rapport à L^p, mais n'est pas dans L^p.

L'ensemble $C_c^\infty(\mathbb{R}^d)$ des fonction de classe C^∞ et à support compact sur $\mathbb{R}^d$ est souvent également noté $\mathscr{D}(\mathbb{R}^d)$.

Théorème 24.33 ([277])**.**
Nous avons des densités emboitées. Ici D est un borélien borné de $\mathbb{R}^d$ contenu dans $B(0,r)$ et K est un compact contenant $B(0, r+2)$.

(1) Les fonctions étagées par rapport à L^p sur $\mathbb{R}^d$ sont denses dans $L^p(\mathbb{R}^d)$. A fortiori les fonctions étagées sont denses dans L^p, mais nous n'en aurons pas besoin ici.

(2) Il existe une suite f_n dans $C(K, \mathbb{C})$ telle que

$$f_n \xrightarrow{L^p} \mathbb{1}_D. \tag{24.115}$$

(3) Si A est un borélien tel que $\mathbb{1}_A \in L^p(\mathbb{R}^d)$ [7] et si $\epsilon > 0$, alors il existe une suite de boréliens bornée $(D_n)_{n\in\mathbb{N}}$ tels que

$$\mathbb{1}_{D_n} \xrightarrow{L^p} \mathbb{1}_A. \tag{24.116}$$

7. Je pense que cette hypothèse manque dans [277]. En tout cas je vois mal comment je pourrais justifier les différentes étapes de la preuve en prenant par exemple $A = \mathbb{R}^d$.

(4) Il existe une suite φ_n dans $\mathscr{D}(\mathbb{R}^d) = C_c^\infty(\mathbb{R}^d)$ telle que

$$\varphi_n \xrightarrow{L^p} \mathbb{1}_D. \tag{24.117}$$

(5) L'ensemble $\mathscr{D}(\mathbb{R}^d) = C_c^\infty(\mathbb{R}^d)$ est dense dans $L^p(\mathbb{R}^d)$ pour tout $1 \leqslant p < \infty$.

Démonstration. Nous allons montrer les choses point par point.

(1) Si $f \in L^1(\mathbb{R}^d)$, nous savons par la proposition 13.106 qu'il existe une suite f_n de fonctions étagées convergeant ponctuellement vers f telle que $|f_n| \leqslant |f|$. La proposition 24.16 nous dit qu'alors $f_n \xrightarrow{L^p} f$.
La fonction f_n étant étagée et dans L^p en même temps, elle est automatiquement étagée par rapport à L^p par le lemme 24.32.

(2) C'est le théorème d'approximation 13.150 appliqué au borélien D contenu dans l'espace mesuré K.

(3) En vertu du point (2), il existe $f \in C^0(K, \mathbb{R})$ telle que

$$\|f - \mathbb{1}_D\|_p \leqslant \epsilon. \tag{24.118}$$

Ensuite, par le théorème de Weierstrass, il existe $\varphi \in C^\infty(K, \mathbb{R})$ telle que $\|f - \varphi\|_\infty \leqslant \epsilon$. Nous avons aussi

$$\|\varphi - f\|_p^p = \int_K |\varphi(x) - f(x)|^p dx \leqslant \mu(X)\|\varphi - f\|_\infty^p \leqslant \epsilon^p \mu(K). \tag{24.119}$$

Quitte à prendre un φ correspondant à un ϵ plus petit, nous avons

$$\|\varphi - f\| \leqslant \epsilon. \tag{24.120}$$

En combinant et en passant à $\epsilon/2$ nous avons trouvé une fonction $\varphi \in C^\infty(K, \mathbb{R})$ telle que

$$\|\varphi - \mathbb{1}_D\| \leqslant \epsilon. \tag{24.121}$$

(4) Nous considérons les boréliens fermés $D_n = A \cap B(0, n)$. Alors $\mathbb{1}_{D_n} \in L^p$ et nous avons pour n assez grand :

$$\int_{\mathbb{R}^d} |\mathbb{1}_{D_n}(x) - \mathbb{1}_A(x)|^p dx = \int_{\mathbb{R}^d \setminus B(0,n)} |\mathbb{1}_A(x)|^p < \epsilon, \tag{24.122}$$

c'est à dire que $\mathbb{1}_{D_n} \xrightarrow{L^p} \mathbb{1}_A$.

(5) Il suffit de remettre tout ensemble. Si $f \in L^p(\mathbb{R}^d)$, par le point (2) nous commençons par prendre σ étagée par rapport à L^p telle que

$$\|\sigma - f\|_p \leqslant \epsilon. \tag{24.123}$$

Ensuite nous écrivons σ sous la forme

$$\sigma = \sum_{k=1}^N c_k \mathbb{1}_{B_k} \tag{24.124}$$

et nous appliquons le point (3) à chacune des $\mathbb{1}_{B_k}$ pour trouver des boréliens bornés D_k tels que

$$\|\mathbb{1}_{D_k} - \mathbb{1}_{B_k}\|_p \leqslant \epsilon. \tag{24.125}$$

Enfin nous appliquons le point (4) pour trouver des fonctions $\varphi_k \in C_c^\infty(\mathbb{R}^d)$ telles que

$$\|\varphi_k - \mathbb{1}_{D_k}\|_p \leqslant \epsilon. \tag{24.126}$$

Il n'est pas compliqué de calculer que

$$\Big\| \sum_{k=1}^N c_k \varphi_k - f \Big\|_p \leqslant 2\epsilon \sum_k c_k + \epsilon. \tag{24.127}$$

□

Corollaire 24.34.
Si $1 < p < \infty$ alors l'ensemble[8] $L^2\big([0,1]\big) \cap L^p\big([0,1]\big)$ *est dense dans* $L^p\big([0,1]\big)$.

Démonstration. Nous savons du théorème 24.33(5) que $C_c^\infty\big([0,1]\big)$ est dense dans L^p. Mais nous avons évidemment $C_c^\infty \subset L^2 \cap L^p$, donc $L^2 \cap L^p$ est dense dans L^p. □

Lemme 24.35 ([277, 270])**.**
Soit $1 \leqslant p < \infty$ et $f \in L^p(\Omega)$. Nous notons τ_v l'opérateur de translation par v :

$$\begin{aligned}\tau_v\colon L^p(\Omega) &\to L^p(\Omega)\\ f &\mapsto \Big[x \mapsto f(x-v)\Big].\end{aligned} \tag{24.128}$$

Cet opérateur est continu en $v = 0$, c'est à dire

$$\lim_{v\to 0} \|\tau_v - f\|_p = 0. \tag{24.129}$$

Démonstration. Nous commençons par supposer que f est dans $\mathscr{D}(\Omega)$, et nous verrons ensuite comment généraliser.

Si $f \in \mathscr{D}(\Omega)$ Soit une suite $v_i \xrightarrow{\mathbb{R}^d} 0$, et posons $f_i = \tau_{v_i}(f)$; le but est de montrer que $f_i \xrightarrow{L^p} f$. Pour cela, la fonction $f - f_i$ est également à support compact, et qui plus est, si $\text{supp}(f) \subset B(0,r)$, alors $\text{supp}(f - f_i) \subset B(0, r + |v_i|)$, et l'ensemble

$$S = \overline{B\big(0, r + \max_i |v_i|\big)} \tag{24.130}$$

est un compact contenant les supports de tous les $f - f_i$. Le maximum existe parce que $v_i \to 0$. Voila qui «majore» le domaine de $f - f_i$ uniformément en i.

Majorons maintenant $|f - f_i|^p$ de façon uniforme en i. Soit le nombre

$$M = 2 \max_{x\in\mathbb{R}^d}\{f(x)\}. \tag{24.131}$$

La fonction qui vaut M^p sur S et zéro ailleurs est une fonction intégrable qui majore $|f-f_i|^p$. Nous pouvons donc utiliser la convergence dominée de Lebesgue (théorème 13.134) pour écrire

$$\lim_{i\to\infty} \|f - f_i\|_p^p = \lim_{i\to\infty} \int_\Omega |f(x) - f(x - v_i)|^p dx = \int_\Omega \lim_{i\to\infty} |f(x) - f(x - v_i)| dx = 0. \tag{24.132}$$

Pour $f \in L^p(\Omega)$ Soit $\epsilon > 0$, $f \in L^p(\Omega)$ et $\varphi \in \mathscr{D}(\Omega)$ tel que $\|f - \varphi\|_p \leqslant \epsilon$. Cela est possible par la densité de $\mathscr{D}(\Omega)$ dans $L^p(\Omega)$ vue en 24.33(5). Nous choisissons de plus $|v|$ assez petit pour avoir $\|\tau_v(\varphi) - \varphi\|_p < \epsilon$, qui est possible en vertu de ce que nous venons de démontrer à propos des fonctions à support compact. De plus τ_v étant une isométrie de L^p nous avons $\|\tau_v(\varphi) - \tau_v(f)\| = \|\varphi - f\| < \epsilon$. Nous avons tout pour majorer :

$$\|f - \tau_v(f)\| \leqslant \|f - \varphi\| + \|\varphi - \tau_v(\varphi)\| + \|\tau_v(\varphi) - \tau_v(f)\| \leqslant 3\epsilon. \tag{24.133}$$

Nous avons donc bien $\lim_{v\to 0} \|f - \tau_v(f)\| = 0$.

□

8. Nous parlons bien ici de l'*ensemble* L^2 parce que nous le considérons sans norme ou topologie particulière. La densité dont nous parlons ici est celle pour la topologique de L^p.

24.3.8 Approximation

Lemme 24.36 (Théorème fondamental d'approximation [166])**.**
Soit Ω un espace mesurable et $f\colon \Omega \to [0,\infty]$ une application mesurable. Alors il existe une suite croissante d'applications étagées $\varphi_n\colon \Omega \to \mathbb{R}^+$ dont la limite est f.

De plus si f est bornée, la convergence est uniforme.

Théorème 24.37 ([281])**.**
Soit I un intervalle de $\mathbb{R}$. L'espace $C_c(I)$ des fonctions continues à support compact sur I est dense dans $L^2(I)$.

Ce théorème sera généralisé à tous les $L^p(\mathbb{R}^d)$ par le théorème 24.33. Cependant L^p n'étant pas un Hilbert, il faudra travailler sans produit scalaire.

Démonstration. Soit $g \in L^2(I)$ une fonction telle que $g \perp f$ pour toute fonction $f \in C_c(I)$. Nous avons donc

$$\langle f, g\rangle = \int_I f\bar{g} = 0. \tag{24.134}$$

En passant éventuellement aux composantes réelles et imaginaires nous pouvons supposer que les fonctions sont toutes réelles. Nous décomposons g en parties positives et négatives : $g = g^+ - g^-$. Notre but est de montrer que $g^+ = g^-$, c'est à dire que g est nulle. La proposition 23.15 conclura que $C_c(I)$ est dense dans $L^2(I)$.

Soit un intervalle $[a,b] \subset I$ et une suite croissante de fonctions $f_n \in C_c(I)$ qui converge vers $\mathbb{1}_{[a,b]}$. Par hypothèse pour chaque n nous avons

$$\int_I f_n g^+ = \int_I f_n g^-. \tag{24.135}$$

La suite étant croissante, le théorème de la convergence monotone (théorème 13.129) s'applique et nous avons

$$\lim_{n\to\infty} \int_I f_n g^+ = \int_a^b g^+, \tag{24.136}$$

de telle sorte que nous ayons, pour tout intervalle $[a,b] \subset I$ l'égalité

$$\int_a^b g^+ = \int_a^b g^-. \tag{24.137}$$

De plus ces intégrales sont finies parce que

$$\int_a^b g^+ \leqslant \int_a^b |g| = \int_I |g|\mathbb{1}_{[a,b]} = \langle |g|, \mathbb{1}_{[a,b]}\rangle \leqslant \|g\|_{L^2}\sqrt{b-a} < \infty \tag{24.138}$$

par l'inégalité de Cauchy-Schwarz.

Soit maintenant un ensemble mesurable $A \subset I$. La fonction caractéristique $\mathbb{1}_A$ est mesurable et il existe une suite croissante de fonctions étagées (φ_n) convergente vers a par le lemme 24.36. À multiples près, les fonctions φ_n sont des sommes de fonctions caractéristiques du type $\mathbb{1}_{[a,b]}$, par conséquent, en vertu de (24.137) nous avons

$$\int_I \varphi_n g^+ = \int_I \varphi_n g^-. \tag{24.139}$$

Une fois de plus nous pouvons utiliser le théorème de la convergence monotone et obtenir

$$\int_A g^+ = \int_A g^- \tag{24.140}$$

pour tout ensemble mesurable $A \subset I$. Si nous notons dx la mesure de Lebesgue, les mesures g^+dx et g^-dx sont par conséquent égales et dominées par dx. Par le corollaire 13.145 du théorème de Radon Nikodym, les fonctions g^+ et g^- sont égales. □

24.4 Convolution

Le théorème qui permet de définir le produit de convolution est la suivant.

Théorème-définition 24.38 ([179]).
Soient $f, g \in L^1(\mathbb{R}^n)$.

(1) Pour presque tout $x \in \mathbb{R}^n$, la fonction

$$y \mapsto g(x-y)f(y) \tag{24.141}$$

*est dans $L^1(\mathbb{R}^n)$, et nous définissons le **produit de convolution** de f et g par*

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y)g(x-y)dy. \tag{24.142}$$

(2) $f * g \in L^1(\mathbb{R}^n)$.

(3) $\|f * g\|_1 \leqslant \|f\|_1\|g\|_1$.

L'ensemble $L^1(\mathbb{R}^n)$ devient alors une algèbre de Banach.

Lemme 24.39.
Le produit de convolution est commutatif : $f * g = g * f$.

Démonstration. Le théorème de Fubini (théorème 13.230) permet d'écrire

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y)g(x-y)dy = \int_{-\infty}^{\infty} dy_1 \dots \int_{-\infty}^{\infty} dy_n f(y)g(x-y). \tag{24.143}$$

En effectuant le changement de variable $z_i = x_i - y_i$ dans chacune des intégrales nous obtenons

$$(f * g)(x) = \int_{\mathbb{R}^n} g(z)f(x-z)dz = (g * f)(x). \tag{24.144}$$

Attention : on pourrait croire qu'un signe apparaît du fait que $z = x - y$ donne $dz = -dy$. Mais en réalité, l'intégrale $\int_{-\infty}^{+\infty}$ devient par le même changement de variables $\int_{+\infty}^{-\infty}$ qui redonne un nouveau signe au moment de remettre dans l'ordre. □

Proposition 24.40.
Si $1 \leqslant p \leqslant \infty$ et si $f \in L^p(\mathbb{R}^d)$ et $g \in L^1(\mathbb{R}^d)$ alors

(1) $f * g \in L^p$

(2) $\|f * g\|_p \leqslant \|f\|_p\|g\|_1$.

Cette proposition est une conséquence de l'inégalité de Minkowski sous forme intégrale : 24.26(3).

Proposition 24.41 ([282]).
Si $f \in L^1(\mathbb{R})$ et si g est dérivable avec $g' \in L^\infty$, alors $f * g$ est dérivable et $(f * g)' = f * g'$.

Démonstration. La fonction qu'il faut intégrer pour obtenir $f * g$ est $f(t)g(x-t)$, dont la dérivée par rapport à x est $f(t)g'(x-t)$. La norme de cette dernière est majorée (uniformément en x) par $G(t) = |f(t)|\|g'\|_\infty$. La fonction f étant dans $L^1(\mathbb{R})$, la fonction G est intégrable et le théorème de dérivation sous l'intégrale (théorème 14.65) nous dit que $f * g$ est dérivable et

$$(f * g)'(x) = \frac{d}{dx}\int_{\mathbb{R}} f(t)g(x-t)dt = \int_{\mathbb{R}} f(t)g'(x-t)dt = (f * g')(x). \tag{24.145}$$

□

Corollaire 24.42.
Si $f \in L^1(\mathbb{R}^d)$ et si g est de classe C^∞, alors $f * g$ est de classe C^∞.

Démonstration. Il s'agit d'itérer la proposition 24.41. □

Lemme 24.43.
Soit $f \in L^2(I)$ telle que

$$\int_I f\varphi = 0 \tag{24.146}$$

pour toute fonction $\varphi \in C_c^\infty(I)$. Alors $f = 0$ presque partout sur I.

Démonstration. Nous considérons la forme linéaire

$$\begin{aligned} \phi\colon L^2(I) &\to \mathbb{C} \\ g &\mapsto \langle f, g\rangle = \int_I f\bar{g}. \end{aligned} \tag{24.147}$$

Par densité [9] nous pouvons aussi considérer une suite (φ_n) dans $C_c^\infty(I)$ convergeant dans L^2 vers f. Alors nous avons pour tout n :

$$\langle f, \varphi_n\rangle = 0. \tag{24.148}$$

En passant à la limite, $\langle f, f\rangle = 0$, ce qui implique $f = 0$ dans L^2 et donc $f = 0$ presque partout en tant que bonne fonction. □

Ce résultat est encore valable dans les espaces L^p (proposition 24.65), mais il demande le théorème de représentation de Riesz [10].

24.4.1 Approximation de l'unité

Nous considérons $\Omega = \mathbb{R}^d$ ou $(S^1)^d$.

Définition 24.44.
*Une **approximation de l'unité** sur Ω est une suite (φ_n) dans $L^1(\Omega)$ telle que*

$$\sup_k \|\varphi_k\|_1 < \infty \tag{24.149a}$$

$$\int_\Omega \varphi_n = 1 \tag{24.149b}$$

$$\lim_{k\to\infty} \int_{\Omega\setminus B(0,\alpha)} |\varphi_k| = 0 \tag{24.149c}$$

pour tout n et pour tout $\alpha > 0$.

Ce sont des fonctions dont la masse vient s'accumuler autour de zéro. En effet quel que soit le voisinage $B(0,\alpha)$, si k est assez grand, il n'y a presque plus rien en dehors.

Pour le point (24.149b), si Ω est S^1, la mesure que nous considérons est $\frac{dx}{2\pi}$.

Exemple 24.45
Une façon de construire une approximation de l'unité sur $\mathbb{R}$ est de considérer une fonction $\varphi \in L^1(\Omega)$ telle que $\int\varphi = 1$ puis de poser

$$\varphi_k(x) = k^d\varphi(kx). \tag{24.150}$$

Ici, Ω peut être $\mathbb{R}$ ou S^1. △

Le lemme suivant permet de construire des approximations de l'unité intéressantes.

9. Théorème 24.33(5).
10. Théorème 24.64.

Lemme 24.46 ([277]).
Si nous posons

$$\varphi_n(x) = \left(\int \varphi(y)^n \right)^{-1} \varphi(x)^n, \tag{24.151}$$

alors nous obtenons une approximation de l'unité dans les deux cas suivants :

(1) Soit φ une fonction continue et positive sur S^1 telle que $\varphi(x) < \varphi(0)$ pour tout $x \notin 2\pi\mathbb{Z}$. Dans ce cas la mesure à prendre pour l'intégrale est $\frac{dy}{2\pi}$.

(2) Soit φ est une fonction continue et positive à support compact sur $\mathbb{R}^d$ telle que $\varphi(x) > \varphi(0)$ pour tout $x \neq 0$.

Théorème 24.47 ([277]).
Soit (φ_k) une approximation de l'unité sur $\Omega = \mathbb{R}^d$ ou $(S^1)^d$.

(1) Si g est mesurable et bornée sur Ω et si g est continue en x_0 alors

$$(\varphi_k * g)(x_0) \to g(x_0). \tag{24.152}$$

(2) Si $g \in L^p(\Omega)$ $(1 \leqslant p < \infty)$ alors

$$\varphi_k * g \xrightarrow{L^p} g. \tag{24.153}$$

(3) Si g est uniformément continue et bornée, alors

$$\varphi_k * g \xrightarrow{L^\infty} g \tag{24.154}$$

Démonstration. Les trois points vont se ressembler.

(1) Nous notons $d_k = (\varphi_k * g)(x_0) - g(x_0)$ et nous devons prouver que $d_k \to 0$. Vu que φ_k est d'intégrale 1 sur Ω nous pouvons écrire

$$|d_k| = \Big| \int_\Omega \big(g(x_0 - y) - g(x_0)\big)\varphi_k(y)dy \Big| \leqslant \int_\Omega \big|g(x_0 - y) - g(x_0)\big||\varphi_k(y)|dy. \tag{24.155}$$

Nous notons $M = \sup_k \|\varphi_k\|_1$, et nous considérons $\alpha > 0$ tel que

$$\big|g(x_0 - y) - g(x_0)\big| \leqslant \epsilon. \tag{24.156}$$

Nous nous restreignons maintenant aux k suffisamment grands pour que $\int_{\complement B(0,\alpha)} |\varphi_k(y)|dy \leqslant \epsilon$. Alors en découpant l'intégrale en $B(0,\alpha)$ et son complémentaire dans Ω,

$$|d_k| \leqslant \epsilon M + \int_{\complement B(0,\alpha)} 2\|g\|_\infty |\varphi_k(y)|dy \leqslant \epsilon M + 2\|g\|_\infty \epsilon \leqslant \epsilon C. \tag{24.157}$$

Donc oui, nous avons $|d_k| \to 0$, et donc le premier point du théorème.

(2) Nous posons $d_k(x) = (\varphi_k * g)(x) - g(x)$ et nous voulons prouver que $\|d_k\|_\infty \to 0$, c'est à dire que $d_k(x)$ converge vers zéro uniformément en x. Nous posons aussi

$$\tau_y(g) \colon x \mapsto g(x - y). \tag{24.158}$$

En récrivant le produit de convolution, une petite majoration donne

$$|d_k(x)| \leqslant \int_\Omega \|\tau_y(g) - g\|_\infty |\varphi_k(y)|dy. \tag{24.159}$$

L'uniforme continuité de g signifie que pour tout ϵ, il existe un α tel que pour tout $y \in B(0,\alpha)$,

$$\|\tau_y(g) - g\|_\infty \leqslant \epsilon. \tag{24.160}$$

Encore une fois nous découpons le domaine d'intégration en $B = B(0, \alpha)$ et son complémentaire :

$$\|d_k\|_\infty \leqslant \int_B \underbrace{\|\tau_y(g) - g\|_\infty}_{\leqslant \epsilon} |\varphi_k(y)|dy + \int_{\complement B} \underbrace{\|\tau_y(g) - g\|_\infty}_{\leqslant 2\|g\|_\infty} |\varphi_k(y)| \tag{24.161a}$$

$$\leqslant \epsilon M + 2\|g\|_\infty \epsilon \tag{24.161b}$$

où la seconde ligne est justifiée par le choix d'un k assez grand pour que $\int_{\complement B} |\varphi_k(y)|dy \leqslant \epsilon$. Nous avons donc bien $\|d_k\|_\infty \to 0$.

(3) Cette fois $g \in L^p(\Omega)$ et nous cherchons à montrer que $\|d_k\|_p \to 0$. Encore qu'ici d_k soit défini à partir d'un représentant dans la classe de g et que d'ailleurs, nous allons travailler avec ce représentant.
D'abord nous développons un peu ce d_k :

$$\|d_k\|_p = \left[\int_\Omega \left|\int_\Omega \big(g(x-y) - g(x)\big)\varphi_k(y)dy\right|^p dx\right]^{1/p} \tag{24.162a}$$

$$\leqslant \left[\int_\Omega \Big(\int_\Omega |g(x-y) - g(x)| \cdot |\varphi_k(y)|dy\Big)^p dx\right]^{1/p}. \tag{24.162b}$$

À cette dernière expression nous appliquons l'inégalité de Minkowski (théorème 24.26) sous la forme (24.64) pour la norme $d\nu(y) = |\varphi_k(y)|dy$ et $f(x, y) = g(x - y) - g(x)$:

$$\|d_k\|_p \leqslant \int_\Omega \Big(\int_\Omega \big|g(x-y) - g(x)\big|^p dx\Big)^{1/p} |\varphi_k(y)|dy = \int_\Omega \|\tau_y g - g\|_p |\varphi_k(y)|dy. \tag{24.163}$$

Par le lemme 24.35 nous pouvons trouver $\alpha > 0$ tel que $\|\tau_y g - g\|_p \leqslant \epsilon$ pour tout $y \in B(0, \alpha)$. Avec cela nous découpons encore le domaine d'intégration :

$$\|d_k\|_p \leqslant \int_{B(0,\alpha)} \underbrace{\|\tau_y g - g\|_p}_{\leqslant \epsilon} |\varphi_k(y)|dy + \int_{\complement B(0,\alpha)} \underbrace{\|\tau_y g - g\|_p}_{\leqslant 2\|g\|_p} |\varphi_k(y)|dy \leqslant \epsilon M + 2\epsilon\|g\|_p. \tag{24.164}$$

□

Une petite remarque en passant : aussi triste que cela en ait l'air, la convergence uniforme n'implique pas la convergence $L^p(\Omega)$ si Ω n'est pas borné. En effet si $f \in L^p$, la suite donnée par

$$f_n(x) = f(x) + \frac{1}{n} \tag{24.165}$$

converge uniformément vers f, mais

$$\|f_n - f\|_p = \int_\Omega \frac{1}{n} \tag{24.166}$$

n'existe même pas si le domaine Ω n'est pas borné.

24.4.2 Densité des polynôme trigonométrique

Nous allons beaucoup travailler avec le **système trigonométrique** donné par $\{e_n\}_{n\in\mathbb{Z}}$ et

$$e_n(t) = e^{int}. \tag{24.167}$$

Une bonne partie de la douleur qu'évoque mot « densité » consiste à montrer que ce système est total dans $L^2(S^1) = L^2([0, 2\pi])$, et donc en est une base hilbertienne.

Définition 24.48.
*Un **polynôme trigonométrique** est une fonction de la forme*

$$P(t) = \sum_{n=-N}^{N} c_n e^{int}. \tag{24.168}$$

Lemme 24.49.
Deux petits résultats simples mais utiles à propos des polynômes trigonométriques.

(1) Si $f \in L^1(S^1)$, alors nous avons la formule

$$f * e_n = c_n(f)e_n. \tag{24.169}$$

(2) Si P est un polynôme trigonométrique et si $f \in L^1(S^1)$ alors $f * P$ est encore un polynôme trigonométrique.

Démonstration. Le premier point est un simple calcul :

$$(f * e_n)(x) = \frac{1}{2\pi}\int_0^{2\pi} f(x-t)e_n(t)dt \tag{24.170a}$$

$$= e^{inx}\int_0^{2\pi} f(x-t)e^{-in(x-t)}dt \tag{24.170b}$$

$$= c_n(f)e_n(x). \tag{24.170c}$$

En ce qui concerne le second point, nous notons $P = \sum_{k=-N}^{N} P_k e_k$, et par linéarité de la convolution,

$$f * P = \sum_{k=-N}^{N} P_k f * e_k = \sum_{k=-N}^{n} c_k(f)e_k, \tag{24.171}$$

qui est encore un polynôme trigonométrique. □

Exemple 24.50
Sur S^1 nous construisons alors l'approximation de l'unité basée sur la fonction $1 + \cos(x)$ et le lemme 24.46. Cette fonction est évidemment un polynôme trigonométrique parce que

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2}. \tag{24.172}$$

Ensuite les puissances le sont aussi à cause de la formule du binôme :

$$\big(1 + \cos(x)\big)^n = \sum_{k=0}^{n} \binom{n}{k} \cos^n(x), \tag{24.173}$$

dans laquelle nous pouvons remettre $\cos(x)$ comme un polynôme trigonométrique et développer à nouveau la puissance avec (encore) la formule du binôme. La chose importante est qu'il existe une approximation de l'unité (φ_n) formée de polynômes trigonométrique.

Ce qui fait la spécificité des polynômes trigonométriques est qu'ils sont à la fois stables par convolution (lemme 24.49) et qu'ils permettent de créer une approximation de l'unité sur $[0, 2\pi]$. Ce sont ces deux choses qui permettent de prouver l'important théorème suivant. △

Théorème 24.51.
Les polynôme trigonométriques sont dense dans $L^p(S^1)$ pour $1 \leqslant p < \infty$.

Démonstration. Soit $f \in L^p(S^1) = L^p\big([0, 2\pi]\big)$, et (φ_n) une approximation de l'unité sur S^1 formée de polynômes trigonométriques, par exemple ceux de l'exemple 24.50. D'abord $\varphi_k * f$ est un polynôme trigonométrique par le lemme 24.49, et ensuite nous avons convergence

$$\varphi_k * f \xrightarrow{L^p} f \tag{24.174}$$

par le théorème 24.47. Nous avons donc convergence L^p d'une suite de polynômes trigonométrique, ce qui prouve que l'espace de polynômes trigonométriques est dense dans $L^p(S^1)$. □

Remarque 24.52.
Deux remarques.

— Il n'est pas possible que les polynômes trigonométriques soient dense dans L^∞ parce qu'une limite uniforme de fonctions continues est continue (c'est le théorème 12.297).
— Nous donnerons au théorème 26.6 une démonstration indépendante de la densité des polynômes trigonométriques dans $L^p(S^1)$.

24.5 L'espace L^2

L'espace L^2 est l'espace L^p définit à la section 24.3 avec $p = 2$. Cependant il possède une propriété extraordinaire par rapport aux autres L^p, c'est que la norme $|.|_2$ dérive d'un produit scalaire. Il sera donc un espace de Hilbert.

Nous en rappelons la construction. Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. Nous considérons l'opération

$$\langle f, g \rangle = \int_\Omega f(\omega)\overline{g(\omega)}d\mu(\omega) \tag{24.175}$$

et la norme associée

$$\|f\|_2 = \sqrt{\langle f, f \rangle}. \tag{24.176}$$

Nous considérons l'ensemble

$$\mathcal{L}^2(\Omega, \mu) = \{f : \Omega \to \mathbb{R} \text{ tel que } \|f\|_2 < \infty\} \tag{24.177}$$

et la relation d'équivalence $f \sim g$ si et seulement si $f(x) = g(x)$ pour μ-presque tout x.

Définition 24.53.
Nous définissons le quotient

$$L^2(\Omega, \mu) = \mathcal{L}^2(\Omega, \mu)/\sim . \tag{24.178}$$

Problèmes et choses à faire

Le lemme suivant dit que L^2 est un Hilbert. Il faudrait utiliser le théorème de Riesz pour prouver que les autres L^p n'en sont pas.

Lemme 24.54.
La formule (24.175) *définit un produit scalaire sur L^2, et ce dernier est un espace de Hilbert.*

Démonstration. D'abord si f et g sont dans L^2, alors l'inégalité de Hölder (proposition 24.22) nous indique que le produit fg est un élément de L^1. Par conséquent la formule a un sens.

Ensuite nous montrons que la formule passe au quotient. Pour cela, nous considérons des fonctions α et β nulles presque partout et nous regardons le produit de $f_1 = f + \alpha$ par $g_1 = g + \beta$:

$$\langle f_1, g_1 \rangle = \int fg + \beta f + \alpha g + \alpha\beta. \tag{24.179}$$

Les fonction βf, αg et $\alpha\beta$ étant nulles presque partout, leur intégrale est nulle et nous avons bien $\langle f_1, g_1 \rangle = \langle f, g \rangle$. Nous pouvons donc considérer le produit sur l'ensemble des classes.

Pour vérifier que la formule est un produit scalaire, le seul point non évidement est de prouver que $\langle f, f \rangle = 0$ implique $f = 0$. Cela découle du fait que

$$\langle f, f \rangle = \int_\Omega |f|^2. \tag{24.180}$$

La fonction $x \mapsto |f(x)|^2$ vérifie les hypothèses du lemme 13.123. Par conséquent $|f(x)|^2$ est presque partout nulle.

En ce qui concerne le fait que $L^2(\Omega)$ soit un espace de Hilbert, il s'agit simplement de se remémorer que c'est un espace complet (théorème 24.29) et dont la norme dérive d'un produit scalaire. Nous sommes donc bien dans la définition 23.2. □

Attention : dans 24.5, le produit scalaire a été défini avec un coefficient $1/2\pi$. Voir l'équation (24.183).

24.5.1 Coefficients et série de Fourier

Nous notons ici une conséquence du théorème 24.30 dans le cas de l'espace L^2. La proposition suivante est une petite partie du corollaire 23.50, qui sera d'ailleurs démontré de façon indépendante.

Proposition 24.55.
Si nous avons une suite de réels (a_k) telle que $\sum_{k=0}^{\infty} |a_k|^2 < \infty$ alors la suite

$$f_n(x) = \sum_{k=0}^{n} a_k e^{ikx} \tag{24.181}$$

converge dans $L^2\big(]0, 2\pi[\big)$.

Démonstration. Quitte à séparer les parties réelles et imaginaires, nous pouvons faire abstraction du fait que nous parlons d'une série de fonctions à valeurs dans $\mathbb{C}$ au lieu de $\mathbb{R}$.

Un simple calcul est :

$$\|f_n - f_m\|^2 \leqslant \int_0^{2\pi} \sum_{k=n}^{m} |a_k|^2 dx \leqslant 2\pi \sum_{k=n}^{m} |a_k|^2. \tag{24.182}$$

Par hypothèse le membre de droite est $|s_m - s_n|$ où s_k dénote la suite des somme partielle de la série des $|a_k|^2$. Cette dernière est de Cauchy (parce que convergente dans $\mathbb{R}$) et donc la limite $n \to \infty$ (en gardant $m > n$) est zéro. Donc la suite des f_n est de Cauchy dans L^2 et donc converge dans L^2. □

Nous utilisons ici des résultats de bases hilbertiennes autour du théorème 23.45. Nous considérons l'espace de Hilbert $L^2[-T/2, T/2]$ muni du produit scalaire [11]

$$\langle f, g \rangle = \frac{1}{T} \int_{-T/2}^{T/2} f(t)\overline{g(t)} dt. \tag{24.183}$$

Pour toute fonction pour laquelle ça a un sens (que ce soit des fonctions L^2 ou non), nous posons

$$c_n(f) = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-2i\pi nt/T} dt. \tag{24.184}$$

Ces nombres sont les **coefficients de Fourier** de f. Leur importance dans le cadre de L^2 provient du fait que la famille de fonctions

$$e_k(t) = e^{2i\pi kt/T} \tag{24.185}$$

est une base hilbertienne de $L^2[-T/2, T/2]$ et que

$$c_n(f) = \langle f, e_n \rangle. \tag{24.186}$$

Adaptons tout cela pour l'espace $L^2\big([0, 2\pi]\big)$. Nous posons

$$\langle f, g \rangle = \int_0^{2\pi} f(t)\overline{g(t)} dt \tag{24.187}$$

et

$$e_n(t) = \frac{1}{\sqrt{2\pi}} e^{int}. \tag{24.188}$$

11. Attention à la convention : nous mettons un coefficient qui n'est pas dans la définition générale (24.175).

Lemme 24.56.
Le système trigonométrique $\{e_n\}_{n\in\mathbb{Z}}$ est une base hilbertienne de $L^2\big([0,2\pi]\big)$.

Démonstration. Pour rappel, une base hilbertienne est la définition 23.27. Nous prouvons d'abord que le système est orthogonal. Nous avons

$$\langle e_n, e_m\rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{i(n-m)t}dt. \tag{24.189}$$

Si $n=m$, alors cela est égal à 1. Sinon c'est une intégrale simple :

$$\langle e_n, e_m\rangle = \left[\frac{i}{n-m}e^{i(n-m)t}\right]_0^{2\pi} = 0. \tag{24.190}$$

Cela est pour l'orthogonalité.

Pour que le système soit total, il faut que son espace vectoriel engendré soit dense. Cela est le théorème 24.51. □

Note : le théorème 26.6 donné aussi la densité, mais sera démontré plus tard, indépendamment. Voir aussi les thèmes 11 et 12.

Pour un élément donné $f\in L^2\big([0,2\pi]\big)$, nous définissons

$$S_nf = \sum_{k=-n}^{n}\langle f, e_k\rangle e_k \tag{24.191}$$

et nous avons le théorème suivant, qui récompense les efforts consentis à propos de la densité des polynômes trigonométriques dans L^2.

Théorème 24.57.
Soit $f\in L^2\big([0,2\pi]\big)$. Nous avons égalité[12]

$$f = \sum_{n\in\mathbb{Z}} c_n(f)e_n \tag{24.192}$$

dans L^2.

Nous avons aussi la convergence

$$S_nf \xrightarrow{L^2} f. \tag{24.193}$$

Démonstration. Le système trigonométrique $\{e_n\}_{n\in\mathbb{Z}}$ est total pour l'espace de Hilbert $L^2\big([0,2\pi]\big)$ (sans périodicité particulière). Donc le point (1) du théorème 23.45 nous donne l'égalité demandée.

La convergence (24.193) est une reformulation de l'égalité (24.192). □

24.58.
Obtenir la convergence L^2 ne demande pas d'hypothèse de périodicité : la convergence (24.193) est automatique du fait que le système trigonométrique soit total. Ce n'est cependant pas plus qu'une convergence L^2 et elle ne demande pas $f(0)=f(2\pi)$, même si pour chacun des e_k nous avons $e_k(0)=e_k(2\pi)$.

Si $f(2\pi)\neq f(0)$, alors il existe tout de même une suite (f_n) convergente vers f au sens L^2 telle que $f_n(0)=f_n(2\pi)$. Cela ne contredit en rien le fait que $e_k(0)=e_k(2\pi)$ parce que dans L^2, la valeur d'un point seul n'a pas d'importance.

Si nous voulons une vraie convergence ponctuelle voir uniforme $(S_nf)(x)\to f(x)$, alors il faut ajouter des hypothèses sur la continuité de f, sa périodicité ou le comportement des coefficients c_n. Voir aussi le thème 4.

12. Notons que la somme sur $\mathbb{Z}$ dans (24.192) est commutative ; il n'est donc pas besoin d'être plus précis.

Exemple 24.59
Si $f \in L^2\big([0,2\pi]\big)$ est (la classe de) une fonction à valeurs réelles, alors on peut la développer avec nettement moins de termes. D'abord nous savons que $e_{-n} = \overline{e_n}$, et donc

$$\langle f, e_n\rangle = \overline{\langle f, e_{-n}\rangle}, \tag{24.194}$$

ce qui donne

$$f = \sum_{n\in\mathbb{Z}} \langle f, e_n\rangle e_n = \sum_{n\in\mathbb{N}} \langle f, e_n\rangle e_n + \overline{\langle f, e_n\rangle e_n} = \sum_{n\in\mathbb{N}} \Re\big(\langle f, e_n\rangle e_n\big). \tag{24.195}$$

Or

$$\Re\big(\langle f, e_n\rangle e_n\big) = \frac{1}{(2\pi)^{3/2}}\cos(nx)\int_0^{2\pi} f(t)\cos(nt)dt - \frac{1}{(2\pi)^{3/2}}\sin(nx)\int_0^{2\pi} f(t)\sin(nt)dt. \tag{24.196}$$

Considérons la fonction impaire $\tilde{f} \in Ł^2\big([-2\pi, 2\pi]\big)$ créée à partir de f. Elle se développe de même et nous avons la même formule (24.196) à part quelque coefficients et le fait que les intégrales sont entre -2π et 2π. Vu que $\tilde{f}$ est impaire, l'intégrale avec $\cos(nt)$ s'annule et

$$\tilde{f}(x) = \sum_{n\in\mathbb{N}} c_n \sin(nx) \tag{24.197}$$

pour certains coefficients réels c_n. Cette égalité est à considérer dans L^2, c'est à dire presque partout et en particulier presque partout sur $[0,2\pi]$.

Donc les fonctions réelles sur $[0,2\pi]$ peuvent être écrites sous la forme d'une série de seulement des sinus.

Note : en choisissant $\tilde{f}$ paire, nous aurions eu une série de cosinus. △

24.5.2 Dualité et théorème de représentation de Riesz

Dans la suite E' est le dual topologique, c'est à dire l'espace des formes linéaires et continues sur E.

Proposition 24.60 ([117])**.**
Soit $1 < p < 2$ et q tel que $\frac{1}{p} + \frac{1}{q} = 1$. L'application

$$\begin{aligned}\Phi\colon L^q\big([0,1]\big) &\to L^p\big([0,1]\big)'\\ \Phi_g(f) &= \int_{[0,1]} f\bar{g}.\end{aligned} \tag{24.198}$$

est une isométrie linéaire surjective.

Démonstration. Pour la simplicité des notations nous allons noter L^2 pour $L^2\big([0,1]\big)$, et pareillement pour L^p.

<u>Φ_g **est un élément de** $(L^p)'$</u> Si $f \in L^p$ et $g \in L^q$ nous devons prouver que $\Phi_q(f)$ est bien définie. Pour cela nous utilisons l'inégalité de Hölder [13] qui dit que $fg \in L^1$; par conséquent la fonction $f\bar{g}$ est également dans L^1 et nous avons

$$|\Phi_g(f)| \leqslant \int_{[0,1]} |f\bar{g}| = \|fg\|_1 \leqslant \|f\|_p\|g\|_q. \tag{24.199}$$

En ce qui concerne la norme de l'application Φ_g nous avons tout de suite

$$\|\Phi_g\| = \sup_{\|f\|_p=1} \big|\Phi_g(f)\big| \leqslant \|g\|_q. \tag{24.200}$$

Cela signifie que l'application Φ_g est bornée et donc continue par la proposition 10.18. Nous avons donc bien $\Phi_g \in (L^p)'$.

13. Proposition 24.22.

Isométrie Afin de prouver que $\|\Phi_g\| = \|g\|_q$ nous allons trouver une fonction $f \in L^p$ telle que $\frac{|\Phi_g(f)|}{\|f\|_p} = \|g\|_q$. De cette façon nous aurons prouvé que $|\Phi_g| \geqslant \|g\|_q$, ce qui conclurait que $|\Phi_g| = \|g\|_q$.

Nous posons $f = g|g|^{q-2}$, de telle sorte que $|f| = |g|^{q-1}$ et

$$\|f\|_p = \left(\int |g|^{p(q-1)} \right)^{1/p} = \left(\int |g|^q \right)^{1/p} = \|g\|_q^{q/p} \tag{24.201}$$

où nous avons utilisé le fait que $p(q-1) = q$. La fonction f est donc bien dans L^p. D'autre part,

$$\Phi_g(f) = \int f\bar{g} = \int g|g|^{q-2}\bar{g} = \int |g|^q = \|g\|_q^q. \tag{24.202}$$

Donc

$$\frac{|\Phi_g(f)|}{\|f\|_p} = \|g\|_q^{q-\frac{q}{p}} = \|g\|_q \tag{24.203}$$

où nous avons encore utilisé le fait que $q - \frac{q}{p} = \frac{q(p-1)}{p} = 1$.

Surjectif Soit $\ell \in (L^p)'$; c'est une application $\ell\colon L^p \to \mathbb{C}$ sont nous pouvons prendre la restriction à L^2 parce que la proposition 24.27 nous indique que $L^2 \subset L^p$. Nous nommons $\phi\colon L^2 \to \mathbb{C}$ cette restriction.

$\phi \in (L^2)'$ Nous devons montrer que ϕ est continue pour la norme sur L^2. Pour cela nous montrons que sa norme opérateur (subordonnée à la norme de L^2 et non de L^p) est finie :

$$\sup_{f\in L^2} \frac{|\phi(f)|}{\|f\|_2} \leqslant \sup_{f\in L^2} \frac{|\ell(f)|}{\|f\|_p} < \infty. \tag{24.204}$$

Nous avons utilisé l'inégalité de norme $\|f\|_p \leqslant \|f\|_2$ de la proposition 24.27(2).

Utilisation du dual de L^2 Étant donné que L^2 est un espace de Hilbert (lemme 24.54) et que $\phi \in (L^2)'$, le théorème 23.18 nous donne un élément $g \in L^2$ tel que $\phi(f) = \Phi_g(f)$ pour tout $f \in L^2$.

Nous devons prouver que $g \in L^q$ et que pour tout $f \in L^p$ nous avons $\ell(f) = \Phi_g(f)$.

$g \in L^q$ Nous posons $f_n = g|g|^{q-2}\mathbb{1}_{|g|<n}$. Nous avons d'une part

$$\Phi_g(f_n) = \int_0^1 f_n\bar{g} = \int_{|g|<n} |g|^q. \tag{24.205}$$

Et d'autre part comme $f_n \in L^2$ nous avons aussi $\phi(f_n) = \Phi_g(f_n)$ et donc

$$0 \leqslant \Phi(f_n) = \phi(f_n) \leqslant \|\ell\|\|f_n\|_p \tag{24.206a}$$

$$= \|\ell\| \left(\int_{|g|<n} |g|^{(q-1)p} \right)^{1/p} \tag{24.206b}$$

$$= \|\ell\| \left(\int_{|g|<n} |g|^q \right)^{1/p}. \tag{24.206c}$$

où nous avons à nouveau tenu compte du fait que $p(q-1) = q$. En combinant avec (24.205) nous trouvons

$$\int_{|g|<n} |g|^q \leqslant \|\ell\| \left(\int_{|g|<n} |g|^q \right)^{1/p}, \tag{24.207}$$

et donc

$$\left(\int_{|g|<n} |g|^q \right)^{1-\frac{1}{p}} \leqslant \|\ell\|, \tag{24.208}$$

c'est à dire

$$\Big(\int_{|g|<n} |g|^q \Big)^{1/q} \leqslant \|\ell\|. \tag{24.209}$$

Si ce n'était pas encore fait nous nous fixons un représentant de la classe g (qui est dans L^2), et nous nommons également g ce représentant. Nous posons alors

$$g_n = |g|^q \mathbb{1}_{|g|<n} \tag{24.210}$$

qui est une suite croissante de fonctions convergeant ponctuellement vers $|g|^q$. Le théorème de Beppo-Levi 13.129 nous permet alors d'écrire

$$\lim_{n\to\infty} \int_{|q|<n} |g|^q = \int_0^1 |g|^q. \tag{24.211}$$

Mais comme pour chaque n nous avons $\int_{|g|<n} |q|^q \leqslant \|\ell\|^q$, nous conservons l'inégalité à la limite et

$$\int_0^1 |g|^q \leqslant \|\ell\|^q. \tag{24.212}$$

Cela prouve que $g \in L^p$.

$\underline{\ell(f) = \Phi_g(f)}$ Soit $f \in L^p$. En vertu de la densité de L^2 dans L^p prouvée dans le corollaire 24.34 nous pouvons considérer une suite (f_n) dans L^2 telle que $f_n \xrightarrow{L^p} f$. Pour tout n nous avons

$$\ell(f_n) = \Phi_g(f_n). \tag{24.213}$$

Mais ℓ et Φ_g étant continues sur L^p nous pouvons prendre la limite et obtenir

$$\ell(f) = \Phi_g(f). \tag{24.214}$$

□

Lemme 24.61 ([7])**.**
Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et $X \subset \Omega$ une partie de mesure $\mu(X) < \infty$. Soit $g \in L^1(\Omega)$ et S fermé dans $\mathbb{C}$. Si pour tout $E \in \mathcal{A}$ nous avons

$$\frac{1}{\mu(E)} \int_e g d\mu \in S, \tag{24.215}$$

alors $g(x) \in S$ pour presque tout $x \in X$.

Démonstration. Soit $D = \overline{B(a,r)}$ un disque fermé dans le complémentaire de S (ce dernier étant fermé, le complémentaire est ouvert). Posons $E = g^{-1}(D)$. Prouvons que $\mu(E) = 0$ parce que cela prouverait que $g(x) \in D$ pour seulement un ensemble de mesure nulle. Mais S^c pouvant être écrit comme une union dénombrable de disques fermés [14], nous aurions $g(x) \in S^c$ presque nulle part.

Vu que $\frac{1}{\mu(E)} \int_E a = a$ nous avons

$$\Big| \frac{1}{\mu(E)} g d\mu - a \Big| = \Big| \frac{1}{\mu(E)} \int_E (g-a) \Big| \leqslant \frac{1}{\mu(E)} \int_E |g-a| \leqslant \frac{1}{\mu(E)} \mu(E) r = r. \tag{24.216a}$$

Donc

$$\frac{1}{\mu(E)} \int_E g d\mu \in D, \tag{24.217}$$

ce qui est une contradiction avec le fait que $D \subset S^c$. □

14. Tout ouvert peut être écrit comme union dénombrable d'éléments d'une base de topologie par la proposition 6.17 et $\mathbb{C}$ a une base dénombrable de topologie par la proposition 6.60.

Théorème 24.62 ([7, 176])**.**
Soit $1 \leqslant p < \infty$ et un espace mesuré fini $(\Omega, \mathcal{A}, \mu)$. Soit q tel que $\frac{1}{p} + \frac{1}{q} = 1$. Alors l'application

$$\begin{aligned} \Phi\colon L^q &\to (L^p)' \\ \Phi_g(f) &= \int_\Omega f\bar{g}d\mu \end{aligned} \tag{24.218}$$

est une bijection isométrique.

Note : ici nous considérons dans L^p des fonctions à valeurs complexes, et donc les éléments du dual sont des applications linéaires continues à valeurs dans $\mathbb{C}$.

Démonstration. Nous commençons par prouver que Φ est injectif. Soient $g, g' \in L^q$ tels que $\Phi_g = \Phi_{g'}$. Alors pour tout $f \in L^p$ nous avons

$$\int_\Omega f(g - g')d\mu = 0. \tag{24.219}$$

En particulier pour tout ensemble mesurable A dans Ω, $\int_A (g - g')d\mu = 0$ parce que $\mathbb{1}_A \in L^p(\Omega)$ parce que nous avons supposé que l'espace était fini. La proposition 24.17 nous dit alors que $g - g' = 0$ dans L^q.

La partie difficile est de montrer que Φ est surjective.

Soit $\phi \in L^p(\Omega)'$. Si $\phi = 0$, c'est bien dans l'image de Φ ; nous supposons donc que non. Nous allons commencer par prouver qu'il existe une (classe de) fonction $g \in L^1(\Omega)$ telle que $\Phi_g(f) = \phi(f)$ pour tout $f \in L^\infty(\Omega, \mu)$; nous montrerons ensuite que $g \in L^q$ et que le tout est une isométrie.

Une mesure complexe Si $E \in \mathcal{A}$ nous notons $\nu(E) = \phi(\mathbb{1}_E)$. Nous prouvons maintenant que ν est une mesure complexe[15] sur $(\Omega, \mathcal{A})$. a seule condition pas facile est la condition de dénombrable additive. Il est déjà facile de voir que A et B sont disjoints, $\nu(A \cup B) = \nu(A) + \nu(B)$. Soient ensuite des ensembles A_n deux à deux disjoints et posons $E_k = \bigcup_{i \leqslant k} A_i$ pour avoir $\bigcup_k A_k = \bigcup_k E_k$ avec l'avantage que les E_k soient emboîtés. Cela donne

$$\|\mathbb{1}_E - \mathbb{1}_{E_k}\|_p = \mu(E \backslash E_k)^{1/p}, \tag{24.220}$$

mais vu que $1 \leqslant p < \infty$, avoir $x_k \to 0$ implique d'avoir $x_k^{1/p} \to 0$. Prouvons que $\mu(E\backslash E_k) \to 0$. En vertu du lemme 13.26 nous avons pour chaque k :

$$\mu(E\backslash E_k) = \mu(E) - \mu(E_k), \tag{24.221}$$

et vu que $E_k \to E$ est une suite croissante, le lemme 13.27(1), sachant que μ est une mesure « normale », donne

$$\lim_{n\to\infty} \mu(E_k) = \mu\Big(\bigcup_k E_k\Big). \tag{24.222}$$

Donc effectivement $\mu(E_k) \to \mu(E)$ et donc oui : $\mu(E\backslash E_k) \to 0$. Jusqu'à présent nous avons

$$\lim_{k\to\infty} \|\mathbb{1}_E - \mathbb{1}_{E_k}\|_p = 0, \tag{24.223}$$

c'est à dire $\mathbb{1}_{E_k} \xrightarrow{L^p} \mathbb{1}_E$. La continuité de ϕ sur L^p donne alors

$$\lim_{k\to\infty} \nu(E_k) = \lim_{k\to\infty} \phi(\mathbb{1}_{E_k}) = \phi(\lim_{k\to\infty} \mathbb{1}_{E_k}) = \phi(\mathbb{1}_E) = \nu(E). \tag{24.224}$$

Par additivité finie de ν nous avons

$$\nu(E_k) = \sum_{i \leqslant k} \nu(A_i) \tag{24.225}$$

et en passant à la limite, $\sum_{i=1}^\infty \nu(A_i) = \nu(\bigcup_i A_i)$. L'application ν est donc une mesure complexe.

15. Définition 13.147.

Mesure absolument continue En prime, si $\mu(E) = 0$ alors $\nu(E) = 0$ parce que

$$\mu(E) = 0 \Rightarrow \|\mathbb{1}_E\|_p = 0 \Rightarrow \mathbb{1}_E = 0 \text{ (dans } L^p) \Rightarrow \phi(\mathbb{1}_E) = 0 \tag{24.226}$$

Utilisation de Radon-Nikodym Nous sommes donc dans un cas où $\nu \ll \mu$ et nous utilisons le théorème de Radon-Nikodym 13.148 sous la forme de la remarque 13.149 : il existe une fonction intégrable $g\colon \Omega \to \mathbb{C}$ [16] telle que pour tout $A \in \mathcal{A}$,

$$\nu(A) = \int_A \bar{g} d\mu. \tag{24.227}$$

C'est à dire que

$$\phi(\mathbb{1}_A) = \int_A \bar{g} d\mu = \int_\Omega \bar{g} \mathbb{1}_A d\mu. \tag{24.228}$$

Nous avons donc exprimé ϕ comme une intégrale pour les fonctions caractéristiques d'ensembles.

Pour les fonctions étagées Par linéarité si f est mesurable et étagée nous avons aussi

$$\phi(f) = \int f\bar{g} d\mu = \Phi_g(f). \tag{24.229}$$

Pour $f \in L^\infty(\Omega)$ Une fonction $f \in L^\infty$ est une fonction presque partout bornée. Nous supposons que f est presque partout bornée par M. Par ailleurs cette f est limite uniforme de fonctions étagées : $\|f_k - f\|_\infty \to 0$ en posant $f_k = f\mathbb{1}_{|f|\leqslant k}$. Pour chaque k nous avons l'égalité

$$\Phi_g(f_k) = \phi(f_k). \tag{24.230}$$

Par ailleurs la fonction $f_k\bar{g}$ est majorée par la fonction intégrable $M\bar{g}$ et le théorème de la convergence dominée 13.134 nous donne

$$\lim_{k\to\infty} \Phi_g(f_i) = \lim_{k\to\infty} \int f_k\bar{g} = \int f\bar{g} = \Phi_g(f). \tag{24.231}$$

Et la continuité de ϕ sur L^p couplée à la convergence $f_k \xrightarrow{L^p} f$ donne $\lim_{k\to\infty} \phi(f_k) = (f)$. Bref prendre la limite dans (24.230) donne

$$\Phi_g(f) = \phi(f) \tag{24.232}$$

pour tout $f \in L^\infty(\Omega)$.

La suite ... Voici les prochaines étapes.

— Nous avons $\int f\bar{g} = \phi(f)$ tant que $f \in L^\infty$. Nous allons étendre cette formule à $f \in L^p$ par densité. Cela terminera de prouver que notre application est une bijection.

— Ensuite nous allons prouver que $\|\phi\| = \|\Phi_g\|$, c'est à dire que la bijection est une isométrie.

De L^∞ à L^p Soit $f \in L^p$. Si nous avions une suite (f_n) dans L^∞ telle que $f_n \xrightarrow{L^p} f$ alors $\lim \phi(f_n) = \phi(f)$ par continuité de ϕ. La difficulté est de trouver une telle suite de façon à pouvoir permuter l'intégrale et la limite :

$$\lim_{n\to\infty} \int_\Omega f_n\bar{g} = \int_\Omega \lim_{n\to\infty} f_n\bar{g} = \int_\Omega f\bar{g} = \Phi_g(f). \tag{24.233}$$

Nous allons donc maintenant nous atteler à la tâche de trouver $f_n \in L^\infty$ avec $f_n \xrightarrow{L^p} f$ et telle que (24.233) soit valide.

Nous allons d'abord supposer que $f \in L^p$ est positive à valeurs réelles. Nous avons alors par la proposition 13.106 qu'il existe une suite croissante de fonction étagées (et donc L^∞)

16. On peut écrire, pour utiliser de la notation compacte que $g \in L^1(\Omega, \mathbb{C})$.

telles que $f_n \to f$ ponctuellement. De plus étant donné que $|f_n| \leqslant |f|$, la proposition 24.16 nous dit que $f_n \xrightarrow{L^p} f$. Pour chaque n nous avons

$$\int_\Omega f_n \bar{g} = \phi(f_n). \tag{24.234}$$

Soit g^+ la partie réelle positive de $\bar{g}$. Alors nous avons la limite croissante ponctuelle $f_n g^+ \to fg^+$ et le théorème de la convergence monotone 13.129 nous permet d'écrire

$$\lim_{n\to\infty} \int f_n g^+ = \int fg^+. \tag{24.235}$$

Faisant cela pour les trois autres parties de $\bar{g}$ nous avons prouvé que si $f \in L^p$ est réelle et positive,

$$\int f\bar{g} = \phi(f), \tag{24.236}$$

c'est à dire que $\Phi_g(f) = \phi(f)$.
Refaisant le tout pour les trois autres parties de f nous montrons que

$$\Phi_g(f) = \phi(f) \tag{24.237}$$

pour tout $f \in L^p(\Omega)$. Nous avons donc égalité de ϕ et Φ_g dans $(L^p)'$ et donc bijection entre $(L^p)'$ et L^q.

Isométrie : mise en place Nous devons prouver que cette bijection est isométrique. Soit $\phi \in (L^p)'$ et $g \in L^q$ telle que $\Phi_g = \phi$. Il faut prouver que

$$\|g\|_q = \|\phi\|_{(L^p)'}. \tag{24.238}$$

$\|\phi\| \leqslant \|g\|_q$ dans $(L^p)'$: Nous savons que $\phi(f) = \int f\bar{g}$, et nous allons écrire la définition de la norme dans

$$\|\phi\|_{(L^p)'} = \sup_{\|f\|_p=1} \big|\phi(f)\big| \tag{24.239a}$$

$$= \sup |\int f\bar{g}| \tag{24.239b}$$

$$\leqslant \sup \underbrace{\int |f\bar{g}|}_{=\|f\bar{g}\|_1}. \tag{24.239c}$$

Il s'agit maintenant d'utiliser l'inégalité de Hölder 24.22 :

$$\|\phi\| \leqslant \sup_{\|f\|_p=1} \|f\|_p \|\bar{g}\|_q = \|g\|_q. \tag{24.240}$$

L'inégalité dans l'autre sens sera démontrée en séparant les cas $p = 1$ et $1 < p < \infty$.

Si $p = 1$ Si E est un ensemble mesurable, alors

$$|\int_E g d\mu| = \big|\phi(\mathbb{1}_E)\big|. \tag{24.241}$$

Mais le fait que $\mu(\Omega) < \infty$ donne que $\mathbb{1}_E \in L^1(\Omega)$. Donc $\mathbb{1}_E \in L^\infty \cap L^1$; nous pouvons alors écrire $\phi(\mathbb{1}_E) = \int_\Omega \mathbb{1}_E \bar{g} d\mu$ et donc

$$|\int_\Omega \mathbb{1}_E \bar{g} d\mu| = |\int_E g d\mu| = \big|\phi(\mathbb{1}_E)\big| \leqslant \|\phi\|_{(L^1)'} \|\mathbb{1}_E\|_1 = \|\phi\|\mu(E). \tag{24.242}$$

Pour utiliser le lemme 24.61, nous écrivons cela dans l'autre sens :

$$\|\phi\| \geqslant \frac{1}{\mu(E)} |\int_\Omega \mathbb{1}_E \bar{g} d\mu| = |\frac{1}{\mu(E)} \int_E \bar{g} d\mu|. \tag{24.243}$$

Si nous prenons $S = \{t \in \mathbb{C} \text{ tel que } |t| \leqslant \|\phi\|\}$, c'est un fermé vérifiant que

$$\frac{1}{\mu(E)} \int_E \bar{g} d\mu \in S, \tag{24.244}$$

donc $\bar{g} \in S$ presque partout, ce qui signifie que $\|g\|_\infty \in S$. Nous en concluons que

$$\|g\|_\infty \leqslant \|\phi\| \tag{24.245}$$

et donc que $g \in L^\infty$.

Notons que cet argument ne tient pas avec $p \neq 1$ parce que l'équation (24.242) terminerait sur $\|\phi\|\mu(E)^{1/p}$. Du coup l'ensemble S à prendre serait $S = \{t \in \mathbb{C} \text{ tel que } |t| \leqslant \|\phi\|\mu(E)^{1/p-1}\}$ et nous sommes en dehors des hypothèses du lemme parce qu'il n'y a pas d'ensemble *indépendant* de E dans lequel l'intégrale $\frac{1}{\mu(E)} \int_E \bar{g} d\mu$ prend ses valeurs.

$\underline{1 < p < \infty}$ La fonction

$$\alpha(x) = \begin{cases} \frac{g(x)}{|g(x)|} & \text{si } g(x) \neq 0 \\ 1 & \text{si } g(x) = 0 \end{cases} \tag{24.246}$$

a la propriété de faire $\alpha g = |g|$ en même temps que $|\alpha(x)| = 1$ pour tout x. Nous définissons

$$E_n = \{x \text{ tel que } |g(x)| \leqslant n\} \tag{24.247}$$

et

$$f_n = \mathbb{1}_{E_n} |g^{q-1}| \alpha. \tag{24.248}$$

Ce qui est bien avec ces fonctions c'est que [17]

$$|f_n|^p = |g^{p(q-1)}| |\alpha|^p = |g|^q \tag{24.249}$$

sur E_n. Dans E_n nous avons $|f_n| = |g^{q-1}| \leqslant n^{q-1}$ et dans E_n nous avons $f_n = 0$. Au final, $f_n \in L^\infty$. Par ce que nous avons vu plus haut, nous avons alors

$$\phi(f_n) = \Phi_g(f_n). \tag{24.250}$$

Par ailleurs,

$$f_n \bar{g} = \mathbb{1}_{E_n} |g^{q-1}| \frac{g}{|g|} \bar{g}, \tag{24.251}$$

donc [18]

$$\left| \int_{E_n} |g|^q d\mu \right| = | \int_\Omega f_n \bar{g} d\mu | = |\phi(f_n)| \leqslant \|\phi\| \|f_n\|_p = \|\phi\| \left(\int_{E_n} |f_n|^p \right)^{1/p} = \|\phi\| \left(\int_{E_n} |g|^q \right)^{1/p}. \tag{24.252}$$

Nous avons de ce fait une inégalité de la forme $A \leqslant \|\phi\| A^{1/p}$ et donc aussi $A^{1/p} \leqslant \|\phi\|^{1/p} A^{1/p^2}$, et donc $A \leqslant \|\phi\| \|\phi\|^{1/p} A^{1/p^2}$. Continuant ainsi à injecter l'inégalité dans elle-même,

$$\left| \int_{E_n} |g|^q d\mu \right| \leqslant \|\phi\|^{1 + \frac{1}{p} + \cdots + \frac{1}{p^k}} \left(\int_{E_n} |g|^q d\mu \right)^{1/p^k}. \tag{24.253}$$

Nous pouvons passer à la limite $k \to \infty$. Sachant que $p > 1$ nous savons $A^{1/k} \to 1$ et

$$1 + \frac{1}{p} + \cdots + \frac{1}{p^k} \to \frac{p}{p-1} = q. \tag{24.254}$$

Nous avons alors

$$\int_{E_n} |g|^q d\mu \leqslant \|\phi\|^q. \tag{24.255}$$

17. C'est ici que nous utilisons le lien entre p et q. En l'occurrence, de $1/p + 1/q = 1$ nous déduisons $q(p-1) = p$.

18. Dans [7], cette équation arrive sans modules, ce qui me laisse entendre que $\phi(f_n)$ est réel et positif pour pouvoir écrire que $\phi(f_n) \leqslant \|\phi\| \|f_n\|_p$, mais je ne comprends pas pourquoi.

L'intégrale s'écrit tout aussi bien sous la forme $\int_\Omega |g|^q \mathbb{1}_{E_n}$. La fonction dans l'intégrale est une suite croissante de fonctions mesurables à valeurs dans $[0,\infty]$. Nous pouvons alors permuter l'intégrale et la limite $n \to \infty$ en utilisant la convergence monotone (théorème 13.129) qui donne alors $\int_\Omega |g|^q \leqslant \|\phi\|^q$ ou encore

$$\|g\|_q \leqslant \|\phi\|. \tag{24.256}$$

Ceci achève de prouver que l'application $\phi \mapsto \Phi_g$ est une isométrie, et donc le théorème.

□

Définition 24.63.
*Un espace V est **réflexif** si l'injection naturelle $V \to V'$ est surjective.*

Théorème 24.64 (Théorème de représentation de Riesz[283, 181])**.**
Nous considérons pour ce théorème des fonctions à valeurs réelles.

Soit l'espace mesuré $(\Omega, \mathcal{A}, \mu)$ et $1 < p < \infty$.

(1) L'espace $L^p(\Omega)$ est réflexif.

(2) Si q est le nombre tel que $\frac{1}{p} + \frac{1}{q} = 1$ alors l'application

$$\begin{aligned} \Phi\colon L^q &\to (L^p)' \\ u &\mapsto \Big(\Phi_u\colon f \to \int_\Omega fu\, d\mu\Big) \end{aligned} \tag{24.257}$$

est une bijection isométrique.

Si μ est une mesure σ-finie, alors

(1) $(L^1)' = L^\infty$

(2) $L^1 \subset (L^\infty)'$ avec une inclusion stricte sauf dans les cas triviaux.

Proposition 24.65.
Soit $f \in L^p(\Omega)$ telle que

$$\int_\Omega f\varphi = 0 \tag{24.258}$$

pour tout $\varphi \in C_c^\infty(\Omega)$. Alors $f = 0$ presque partout.

Démonstration. Nous considérons la forme linéaire $\Phi_f \in (L^q)'$ donnée par

$$\begin{aligned} \Phi_f\colon L^p &\to \mathbb{C} \\ u &\mapsto \int_\Omega fu \end{aligned} \tag{24.259}$$

Par hypothèse cette forme est nulle sur la partie dense $C_c^\infty(\Omega)$. Si (φ_n) est une suite dans $C_c^\infty(\Omega)$ convergente vers u dans L^p, nous avons pour tout n que

$$0 = \Phi_f(\varphi_n) \tag{24.260}$$

En passant à la limite, nous voyons que Φ_f est la forme nulle. Elle est donc égale à Φ_0. La partie « unicité » du théorème de représentation de Riesz 24.64 nous indique alors que $f = 0$ dans L^p et donc $f = 0$ presque partout. □

Proposition 24.66.
Si $f \in L^1_{loc}(I)$ est telle que

$$\int_I f\varphi' = 0 \tag{24.261}$$

pour tout $\varphi \in C_c^\infty(I)$, alors il existe une constante C telle que $f = C$ presque partout.

Démonstration. Soit $\psi \in C_c^\infty(I)$ une fonction d'intégrale 1 sur I. Si $w \in C_c^\infty(I)$ alors nous considérons la fonction

$$h = w - \psi \int_I w, \tag{24.262}$$

qui est dans $C_c^\infty(I)$ et dont l'intégrale sur I est nulle. Par la proposition 18.11, la fonction h admet une primitive dans $C_c^\infty(I)$; et nous notons φ cette primitive. L'hypothèse appliquée à φ donne

$$0 = \int_I f\varphi' = \int_I f\left(w - \psi\int_I w\right) = \int_I fw - \underbrace{\left(\int_I f(x)\psi(x)dx\right)}_{C}\left(\int_I w(y)dy\right) = \int_I w(f-C). \tag{24.263}$$

L'annulation de la dernière intégrale implique par la proposition 24.65 que $f - C = 0$ dans L^2, c'est à dire $f = C$ presque partout. □

24.6 Théorèmes de Hahn-Banach

Théorème 24.67 (Hahn-Banach[284, 62])**.**
Soit E, un espace vectoriel réel et une application $p\colon E \to \mathbb{R}$ satisfaisant

(1) $p(\lambda x) = \lambda p(x)$ pour tout $x \in E$ et pour tout $\lambda > 0$,

(2) $p(x+y) \leqslant p(x) + p(y)$ pour tout $x, y \in E$.

Soit de plus $G \subset E$ un sous-espace vectoriel muni d'une application $g\colon G \to \mathbb{R}$ vérifiant $g(x) \leqslant p(x)$ pour tout $x \in G$. Alors il existe $f \in \mathcal{L}(E, \mathbb{R})$ telle que $f(x) = g(x)$ pour tout $x \in G$ et $f(x) \leqslant p(x)$ pour tout $x \in E$.

Démonstration. Si h une application linéaire définie sur un sous-espace de E, nous notons D_h ledit sous-espace.

Un ensemble inductif Nous considérons P, l'ensemble des fonctions linéaires suivant

$$P = \left\{ h\colon D_h \to \mathbb{R} \text{ tel que } \begin{cases} G \subset D_h \\ h(x) = g(x) \quad \forall x \in G \\ h(x) \leqslant p(x) \quad \forall x \in D_h \end{cases} \right\} \tag{24.264}$$

Cet ensemble est non vide parce que g est dedans. Nous le munissons de la relation d'ordre $h_1 \leqslant h_2$ si et seulement si $D_{h_1} \subset D_{h_2}$ et h_2 prolonge h_1. Nous montrons à présent que P est un ensemble inductif. Soit un sous-ensemble totalement ordonné $Q \subset P$; nous définissons une fonction h de la façon suivante. D'abord $D_h = \sup_{l\in Q} D_l$ et ensuite

$$\begin{aligned} h\colon D_h &\to \mathbb{R} \\ x &\mapsto l(x) \quad \text{si } x \in D_l \end{aligned} \tag{24.265}$$

Cela est bien définit parce que si $x \in D_l \cap D_{l'}$ alors, vu que Q est totalement ordonné (i.e. $l \leqslant l'$ ou $l' \leqslant l$), on a obligatoirement $D_l \subset D_{l'}$ et l' qui prolonge l (ou le contraire). Donc h est un majorant de Q dans P parce que $h \geqslant l$ pour tout $l \in Q$. Cela montre que P est inductif (définition 2.9). Le lemme de Zorn 2.10 nous dit alors que P possède un maximum f qui va être la réponse à notre théorème.

Le support de f La fonction f est dans P ; donc $f(x) \leqslant p(x)$ pour tout $x \in D_h$ et $f(x) = g(x)$ pour tout $x \in G$. Pour terminer nous devons montrer que $D_f = E$. Supposons donc que $D_f \neq E$ et prenons $x_0 \notin D_f$. Nous allons contredire la maximalité de f en considérant la fonction h donnée par $D_h = D_f + \mathbb{R}x_0$ et

$$h(x + tx_0) = f(x) + t\alpha \tag{24.266}$$

où α est une constante que nous allons fixer plus tard.

Nous commençons par prouver que f est dans P. Nous devons prouver que

$$h(x+tx_0)=f(x)+t\alpha\leqslant p(x+tx_0) \tag{24.267}$$

Pour cela nous allons commencer par fixer α pour avoir les relations suivantes :

$$\begin{cases} f(x)+\alpha\leqslant p(x+x_0) & (24.268a)\\ f(x)-\alpha\leqslant p(x-x_0) & (24.268b)\end{cases}$$

pour tout $x\in D_f$. Ces relations sont équivalentes à demander α tel que

$$\begin{cases} \alpha\leqslant p(x+x_0)-f(x) & (24.269a)\\ \alpha\geqslant f(x)-p(x-x_0) & (24.269b)\end{cases}$$

Nous nous demandons donc s'il existe un α qui satisfasse

$$\sup_{y\in D_f}\big(f(y)-p(y-x_0)\big)\leqslant\alpha\leqslant\inf_{z\in D_f}\big(p(z+x_0)-f(z)\big). \tag{24.270}$$

Ou encore nous devons prouver que pour tout $y,z\in D_f$,

$$p(z+x_0)-f(x)\geqslant f(y)-p(y-x_0)\geqslant 0. \tag{24.271}$$

Par les propriétés de p et de f,

$$p(z+x_0)+p(y-x_0)-f(z)-f(y)\geqslant p(z+y)-f(z+y)\geqslant 0. \tag{24.272}$$

La dernière inégalité est le fait que $f\in P$. Un choix de α donnant les inéquations (24.268) est donc possible.

À partir des inéquations (24.268) nous obtenons la relation (24.267) de la façon suivante. Si $t>0$ nous multiplions l'équation (24.268a) par t :

$$tf(x)+t\alpha\leqslant tp(x+x_0). \tag{24.273}$$

Et nous écrivons cette relation avec x/t au lieu de x en tenant compte de la linéarité de f :

$$f(x)+t\alpha\leqslant tp\big(\frac{x}{t}+x_0\big)=p(x+tx_0). \tag{24.274}$$

Avec $t<0$, c'est similaire, en faisant attention au sens des inégalités.

Nous avons donc construit $h\colon D_h\to\mathbb{R}$ avec $h\in P$, $D_f\subset D_h$ et $h(x)=f(x)$ pour tout $x\in D_f$. Cela pour dire que $h>f$, ce qui contredit la maximalité de f. Le domaine de f est donc E tout entier.

La fonction f est donc une fonction qui remplit les conditions.

□

Définition 24.68.
Un espace topologique est ***localement convexe*** *si tout point possède un système fondamental de voisinages formé de convexes.*

Définition 24.69 (Hyperplan qui sépare)**.**
Soit E un espace vectoriel topologique ainsi que A, B des sous-ensembles de E. Nous disons que l'hyperplan d'équation $f=\alpha$ ***sépare au sens large*** *les parties A et B si $f(x)\leqslant\alpha$ pour tout $x\in A$ et $f(x)\geqslant\alpha$ pour tout $x\in B$.*

La séparation est ***au sens strict*** *s'il existe $\epsilon>0$ tel que*

$$\begin{aligned} f(x)&\leqslant\alpha-\epsilon && \text{pour tout } x\in A & (24.275a)\\ f(x)&\geqslant\alpha+\epsilon && \text{pour tout } x\in B. & (24.275b)\end{aligned}$$

Théorème 24.70 (Hahn-Banach, première forme géométrique[62])**.**
Soit E un espace vectoriel topologique et A, B deux convexes non vides disjoints de E. Si A est ouvert, il existe un hyperplan fermé qui sépare A et B au sens large.

Théorème 24.71 (Hahn-Banach, seconde forme géométrique)**.**
Soit un espace vectoriel topologique localement convexe[19] *ainsi que des convexes non vides disjoints A et B tels que A soit compact et B soit fermé. Alors il existe un hyperplan fermé qui sépare strictement A et B.*

Démonstration. Vu que B est fermé, A est dans l'ouvert $E\backslash B$. Donc si $a \in A$, il existe un voisinage ouvert convexe de a inclus à A. Soit U_a un voisinage ouvert et convexe de 0 tel que $(a+U_a)\cap B = \varnothing$.

Vu que la fonction $(x, y) \mapsto x + y$ est continue, nous pouvons trouver un ouvert convexe V_a tel que $V_a + V_a \subset U_a$. L'ensemble $a + V_a$ est alors un voisinage ouvert de a et bien entendu $\bigcup_a (a + V_a)$ recouvre A qui est compact. Nous en extrayons un sous-recouvrement fini, c'est à dire que nous considérons $a_1, \dots, a_n \in A$ tels que

$$A \subset \bigcup_{i=1}^{n} (a_i + V_{a_i}). \tag{24.276}$$

Nous posons alors

$$V = \bigcap_{i=1}^{n} V_{a_i}. \tag{24.277}$$

Cet ensemble est non vide parce et il contient un voisinage de zéro parce que c'est une intersection finie de voisinages de zéro. Soit $x \in A + V$. Il existe i tel que

$$x \in a_i + U_{a_i} + V \subset a_i + V_{a_i} + V_{a_i} \subset a_i + U_{a_i} \subset E\backslash B. \tag{24.278}$$

Donc $(A+V)\cap B = \varnothing$. L'ensemble $A+V$ est alors un ouvert convexe disjoint de B. Par la première forme géométrique du théorème de Hahn-Banach 24.70 nous avons un hyperplan qui sépare $A + V$ de B au sens large : il existe $f \in E'\backslash\{0\}$ tel que $f(a) + f(v) \leqslant f(b)$ pour tout $a \in A$, $v \in V$ et $b \in B$.

Il suffit donc de trouver un $v \in V$ tel que $f(v) \neq 0$ pour avoir la séparation au sens strict. Cela est facile : V étant un voisinage de zéro et f étant linéaire, si elle était nulle sur V, elle serait nulle sur E. □

24.7 Théorème de Tietze

Définition 24.72.
*Si E et F sont des espaces normés, une application $f : E \to F$ est **presque surjective** s'il existe $\alpha \in]0, 1[$ et $C > 0$ tels que pour tout $y \in \overline{B_F(0, 1)}$, il existe $x \in \overline{B_E(0, C)}$ tel que $\|y - f(x)\| \leqslant \alpha$.*

Lemme 24.73 ([2])**.**
Soient E et F des espaces de Banach et $f \in L(E, F)$[20]*. Si f est presque surjective, alors*

(1) f est surjective

(2) pour tout $y \in \overline{B_F(0, 1)}$, il existe $x \in \overline{B_E(0, \frac{C}{1-\alpha})}$ tel que $y = f(x)$.

Le point (2) est une précision du point (1) : il dit quelle est la taille de la boule de E nécessaire à obtenir la boule unité dans F.

Démonstration. Soit $y \in \overline{B_F(0, 1)}$. Nous allons construire $x \in B\big(0, \frac{C}{1-\alpha}\big)$ qui donne $f(x) = y$. Ce x sera la limite d'une série que nous allons construire par récurrence. Pour $n = 1$ nous utilisons la presque surjectivité pour considérer $x_1 \in \overline{B_E(0, C)}$ tel que $\|y - f(x_1)\| \leqslant \alpha$. Ensuite nous considérons la récurrence

$$x_n \in \overline{B_E(0, C)} \tag{24.279}$$

19. Définition 24.68.
20. L'ensemble des applications linéaires continues

tel que

$$\|y - \sum_{i=1}^{n} \alpha^{i-1} f(x_i)\| \leqslant \alpha^n \tag{24.280}$$

Pour montrer que cela existe nous supposons que la série est déjà construire jusqu'à $n > 1$:

$$\frac{1}{\alpha^n}\Big(y - \sum_{i=1}^{n} \alpha^{i-1} f(x_i)\Big) \in \overline{B_F(0,1)} \tag{24.281}$$

À partir de là, par presque surjectivité il existe un $x_{n+1} \in \overline{B_E(0,C)}$ tel que

$$\|\frac{y - \sum_{i=1}^{n} \alpha^{i-1} f(x_i)}{\alpha^n} - f(x_{n+1})\| \leqslant \alpha. \tag{24.282}$$

En multipliant par α^n, le terme $\alpha^n f(x_{n+1})$ s'intègre bien dans la somme :

$$\|y = \sum_{i=1}^{n+1} \alpha^{i-1} f(x_i)\| \leqslant \alpha^{n+1}. \tag{24.283}$$

Nous nous intéressons à une éventuelle limite à la somme des $\alpha^{n-1}x_n$. D'abord nous avons la majoration $\|\alpha^{n-1}x_n\| \leqslant \alpha^{n-1}C$, et vu que par la définition de la presque surjectivité $0 < \alpha < 1$, la série

$$\sum_{n=1}^{\infty} \alpha^{n-1} x_n \tag{24.284}$$

converge absolument [21] parce que la suite des normes est une suite géométrique de raison α. Vu que E est de Banach, la convergence absolue implique la convergence simple (la suite des sommes partielles est de Cauchy et Banach est complet). Nous posons

$$x = \sum_{n=1}^{\infty} \alpha^{n-1} x_n \in E, \tag{24.285}$$

et en termes de normes, ça vérifie

$$\|x\| \leqslant \sum_{n=1}^{\infty} \alpha^{n-1}\|x_n\| \leqslant C \sum_{n=1}^{\infty} \alpha^{n-1} = \frac{C}{1-\alpha}. \tag{24.286}$$

Donc c'est bon pour avoir $x \in B\big(0, \frac{C}{1-\alpha}\big)$. Nous devons encore vérifier que $y = f(x)$. Pour cela nous remarquons que

$$\|y - f\Big(\sum_{n=1}^{N} \alpha^{n-1} x_n\Big)\| \leqslant \alpha^N. \tag{24.287}$$

Nous pouvons prendre la limite $N \to \infty$ et permuter f avec la limite (par continuité de f). Vu que $0 < \alpha < 1$ nous avons

$$\|y - f(x)\| = 0. \tag{24.288}$$

□

Théorème 24.74 (Tietze[2, 209])**.**
Soit un espace métrique (X, d) et un fermé $Y \subset X$. Soit $g_0 \in C^0(Y, \mathbb{R})$. Alors g_0 admet un prolongement continu sur X.

Démonstration. Soit l'opération de restriction

$$\begin{aligned} T\colon (C_b^0(X,\mathbb{R}), \|.\|_\infty) &\to (C_b^0(Y,\mathbb{R}), \|.\|_\infty) \\ f &\mapsto f|_Y. \end{aligned} \tag{24.289}$$

L'application T est évidemment linéaire. Elle est de plus borné pour la norme opérateur usuelle donnée par la proposition 10.5 parce que $\|T(f)\| \leqslant \|f\| < \infty$. L'application T est alors continue par la proposition 10.18.

21. Définition 10.45.

Presque surjection Soit $g \in C_b^0(Y,\mathbb{R})$ avec $\|g\|_\infty \leqslant 1$. Nous posons

$$Y^+ = \{x \in Y \text{ tel que } \frac{1}{3} \leqslant g(x) \leqslant 1\} \tag{24.290a}$$

$$Y^- = \{x \in Y \text{ tel que } -1 \leqslant g(x) \leqslant -\frac{1}{3}\}. \tag{24.290b}$$

Nous considérons alors

$$\begin{aligned} f\colon X &\to \mathbb{R} \\ x &\mapsto \frac{1}{3}\frac{d(x,Y^-) - d(x,Y^+)}{d(x,Y^-) + d(x,Y^+)} \end{aligned} \tag{24.291}$$

Vu qu'en valeur absolue le dénominateur est plus grand que le numérateur nous avons $\|f\|_\infty \leqslant \frac{1}{3}$. Notons que

— Si $x \in Y^+$ alors $f(x) = \frac{1}{3}$ et $g(x) \in [\frac{1}{3}, 1]$;
— Si $x \in Y^-$ alors $f(x) = -\frac{1}{3}$ et $g(x) \in [-1, -\frac{1}{3}]$;
— Si x n'est ni dans Y^+ ni dans Y^- alors nous avons [22] $g(x) \in [-\frac{1}{3}, \frac{1}{3}]$ et donc $\big|f(x)-g(x)\big| \leqslant \big|f(x)\big| + \big|g(x)\big| \leqslant \frac{2}{3}$.

Dans les deux cas nous avons $\big|f(x) - g(x)\big| \in [0, \frac{2}{3}]$ pour tout $x \in X$. Cela prouve que

$$\|T(f) - g\|_{Y,\infty} \leqslant \frac{2}{3}. \tag{24.292}$$

En résumé nous avons pris g dans la boule $\overline{B(0,1)}$ de $\big(C_b^0(Y,\mathbb{R}), \|.\|_\infty\big)$ et nous avons construit une fonction f dans la boule $\overline{B(0,\frac{1}{3})}$ de $\big(C_b^0(X,\mathbb{R}), \|.\|_\infty\big)$ telle que $\|T(f)-g\|_\infty \leqslant \frac{2}{3}$. L'application T est donc une presque surjection avec $\alpha = \frac{1}{3}$ et $C = \frac{2}{3}$.

Prolongement dans les boules unité fermées La proposition 12.299 nous assure que les espaces $C_b^0(X,\mathbb{R})$ et $C_b^0(Y,\mathbb{R})$ sont de Banach (complets), et le lemme 24.73 nous dit alors que T est surjective et que pour tout $g \in \overline{B(0,1)}$, il existe

$$f \in \overline{B\left(0, \frac{1/3}{1-\frac{2}{3}}\right)} = \overline{B(0,1)}. \tag{24.293}$$

telle que $g = T(f)$.

Prolongement pour les boules ouvertes Jusqu'à présent nous avons montré qu'une fonction $g \in \overline{B(0,1)}$ admet une prolongement continu dans $\overline{B(0,1)}$. Nous allons montrer que si g est dans la boule ouverte $B(0,1)$ de $\big(C_b^0(Y,\mathbb{R}), \|.\|_\infty\big)$ alors g admet un prolongement dans la boule ouverte $B(0,1)$ de $\big(C_b^0(X,\mathbb{R}), \|.\|_\infty\big)$.

Soit $g \in B_{C_b^0(Y)}(0,1)$ et son prolongement $h \in \overline{B_{C_b^0(X)}(0,1)}$. Si $\|h\|_\infty < 1$ alors le résultat est vrai. Sinon nous considérons l'ensemble

$$Z = \{x \in X \text{ tel que } |h(x)| = 1\}. \tag{24.294}$$

Nous avons $Y \cap Z = \varnothing$ parce que nous avons $h = g$ sur Y et nous avons choisi $\|g\|_\infty < \infty$. Par ailleurs Y est fermé par hypothèse et Z est fermé parce que h est continue ; par conséquent $Y \cap Z$ est fermé, donc [23]

$$\bar{Y} \cap \bar{Z} = Y \cap Z = \varnothing. \tag{24.295}$$

Nous posons

$$\begin{aligned} u\colon X &\to \mathbb{R}^+ \\ x &\mapsto \frac{d(x,Z)}{d(x,Y) + d(x,Z)} \end{aligned} \tag{24.296}$$

22. Nous rappelons que $\|g\| = 1$, donc $g(x)$ est forcément ente -1 et 1.

23. Si vous avez l'intention de dire que $\overline{Y \cap Z} = \bar{Y} \cap \bar{Z} = Y \cap Z = \varnothing$, allez d'abord voir l'exemple 6.180. Ici c'est correct parce que Y et Z sont fermés.

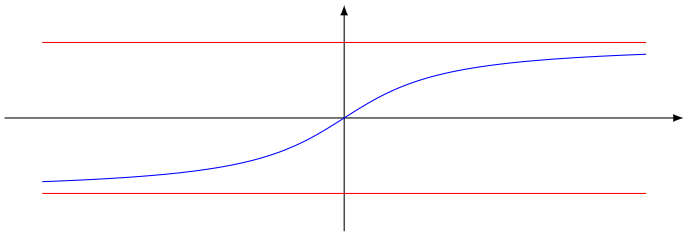
Le dénominateur n'est pas nul parce qu'il faudrait $d(x,Y) = d(x,Z) = 0$, ce qui demanderait $x \in \bar{Y} \cap \bar{Z}$, ce qui n'est pas possible. Nous posons $f = uh$. Si $x \in Y$ alors $u(x) = 1$, donc f est encore un prolongement de g. De plus f est encore continue, et donc encore un bon candidat. Enfin si x est hors de Y alors $d(x,Y) > 0$ (strictement parce que Y est fermé) et donc $0 < u(x) < 1$, ce qui donne $|f(x)| < |h(x)| \leqslant 1$. Donc $\|f\|_\infty < 1$.

Nous avons donc trouvé qu'une fonction dans la boule ouverte $B_{C_b^0(Y)}(0,1)$ se prolonge en une fonction dans la boule ouverte $B_{C_b^0(X)}(0,1)$.

Le cas non borné Soit enfin $g_0 \in C^0(Y,\mathbb{R})$. Nous allons nous ramener au cas de la boule unité ouverte en utilisant un homéomorphisme $\phi\colon \mathbb{R} \to]-1,1[$. L'application $g = \phi \circ g_0$ est dans la boule unité ouvert de $C^0(Y,\mathbb{R})$ et donc admet un prolongement f dans la boule unité ouverte de $C^0(X)$. L'application $f_0 = \phi^{-1} \circ f$ est un prolongement continu de g_0.

□

Un homéomorphisme $\phi\colon \mathbb{R} \to]-1,1[$ est donné par exemple par la fonction $\phi(t) = \frac{2}{\pi}\arctan(t)$ dont le graphique est donné ci-dessous :



24.8 Espace de Schwartz

Pour un multiindice $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$, nous notons

$$\partial^\alpha \varphi = \partial_{x_1}^{\alpha_1} \dots \partial_{x_d}^{\alpha_d} \varphi \tag{24.297}$$

pour peu que la fonction φ soit $|\alpha| = \alpha_1 + \cdots + \alpha_d$ fois dérivable.

Définition 24.75.
*Soit $\Omega \subset \mathbb{R}^d$. L'**espace de Schwartz** $\mathscr{S}(\Omega)$ est le sous-ensemble de $C^\infty(\Omega)$ des fonctions dont toutes les dérivées décroissent plus vite que tout polynôme :*

$$\mathscr{S}(\Omega) = \big\{\varphi \in C^\infty(\Omega) \text{ tel que } \forall \alpha, \beta \in \mathbb{N}^d, p_{\alpha,\beta}(\varphi) < \infty\big\} \tag{24.298}$$

où nous avons considéré

$$p_{\alpha,\beta}(\varphi) = \sup_{x\in\Omega} |x^\beta(\partial^\alpha\varphi)(x)| = \|x^\beta \partial^\alpha \varphi\|_\infty. \tag{24.299}$$

Pour simplifier les notations (surtout du côté de Fourier), nous allons parfois écrire $M_i\varphi$ pour la fonction $x \mapsto x_i\varphi(x)$.

Exemple 24.76
La fonction e^{-x^2} est une fonction à décroissance rapide sur $\mathbb{R}$. △

Définition 24.77.
*Une fonction $f\colon \mathbb{R}^d \to \mathbb{C}$ est dite à **décroissance rapide** si elle décroît plus vite que n'importe quel polynôme. Plus précisément, si pour tout polynôme Q, il existe un $r > 0$ tel que $|f(x)| < \frac{1}{|Q(x)|}$ pour tout $\|x\| \geqslant r$.*

Proposition 24.78.
Une fonction Schwartz est à décroissance rapide.

Démonstration. Nous commençons par considérer un polynôme P donné par

$$P(x) = \sum_k c_k x^{\beta_k} \tag{24.300}$$

où les β_k sont des multiindices, les c_k sont des constantes et la somme est finie. Nous avons la majoration

$$\sup_{x\in\mathbb{R}^d} |\varphi(x)P(x)| \leqslant \sum_k \sup_x \left|c_k\varphi(x)x^{\beta_k}\right| \leqslant \sum_k |c_k| p_{0,\beta_k}(\varphi) < \infty. \tag{24.301}$$

Nous allons noter M_P la constante $\sum_k |c_k| p_{0,\beta_k}(\varphi)$, de sorte que pour tout $x \in \mathbb{R}^d$ nous ayons $|\varphi(x)P(x)| \leqslant M_P$ et donc

$$|\varphi(x)| \leqslant \frac{M_P}{|P(x)|} = \frac{1}{|\frac{1}{M_P}P(x)|}. \tag{24.302}$$

Notons que cette inégalité est a fortiori correcte pour les x sur lesquels P s'annule.

Soit maintenant un polynôme Q. Nous considérons le polynôme $P(x) = \|x\|Q(x)$. Étant de plus haut degré, pour toute constante C il existe un rayon r_C tel que $|P(x)| \geqslant C|Q(x)|$ pour tout $|x| \geqslant r_C$. En particulier pour $|x| \geqslant r_{M_P}$ nous avons

$$|P(x)| \geqslant M_P|Q(x)| \tag{24.303}$$

et donc, pour ces x,

$$|\varphi(x)| \leqslant \frac{1}{|\frac{1}{M_P}P(x)|} \leqslant \frac{1}{|Q(x)|}. \tag{24.304}$$

La première inégalité est valable pour tout x, et la seconde pour $\|x\| \geqslant r_{M_P}$. □

Corollaire 24.79 ([1])**.**
Soit φ une fonction Schwartz sur $\mathbb{R}^m \times \mathbb{R}^n$. Alors la fonction

$$y \mapsto \sup_{x\in\mathbb{R}^n} |\varphi(x,y)| \tag{24.305}$$

est intégrable.

Démonstration. Soit un polynôme Q en la variable y. Par la proposition 24.78, il existe $r > 0$ tel que

$$|\varphi(x,t)| < \frac{1}{Q(y)} \tag{24.306}$$

pour tout $\|(x,y)\| > r$. A fortiori l'inégalité tient pour tout $|y| > r$. Donc

$$\int_{\mathbb{R}^m} \sup_{x\in\mathbb{R}^n} |\varphi(x,y)|dy = \int_{\|y\|\leqslant r} \sup_x |\varphi(x,y)|dy + \int_{\|y\|>r} \sup_x |\varphi(x,y)|dy. \tag{24.307}$$

La première intégrale est bornée par $\text{Vol}\big(B(0,r)\big)\|f\|_\infty$ tandis que la seconde est bornée par l'intégrale de $\frac{1}{Q(y)}$. En prenant Q de degré suffisamment élevé en toutes les composantes de y nous avons intégrabilité. □

24.8.1 Topologie

Lemme-définition 24.80.
Les $p_{\alpha,\beta}$ donnés par l'équation (24.299) *ci-dessus sont des semi-normes*[24]*. La topologie considérée sur $\mathscr{S}(\mathbb{R}^d)$ est celle des semi-normes $p_{\alpha,\beta}$.*

24.81.
Nous avons un enchaînement de résultats qui nous aident à prouver la continuité d'une application $T\colon \mathscr{S}(\mathbb{R}^d) \to X$.

24. Définition 6.190.

(1) La topologie de $\mathscr{S}(\mathbb{R}^d)$ est donnée par une famille dénombrable de semi-normes. Donc la proposition 6.133 nous dit que $\mathscr{S}(\mathbb{R}^d)$ est métrisable.

(2) La proposition 6.141 nous dit alors que si X est métrique, toute application séquentiellement continue $T\colon \mathscr{S}(\mathbb{R}^d) \to X$ est continue.

(3) Donc si X est métrique, il suffit de prouver que pour $f_n \xrightarrow{\mathscr{S}(\mathbb{R}^d)} 0$ nous avons $T(f_n) \xrightarrow{X} 0$ où $f_n\colon \mathscr{S}(\mathbb{R}^d) \to X$. Dans les cas usuels, T sera une distribution et $X = \mathbb{C}$.

(4) En vertu de la proposition 6.193, la convergence $f_n \xrightarrow{\mathscr{S}(\mathbb{R}^d)} 0$ signifie que pour tout choix de multiindice α et β, $p_{\alpha,\beta}(f_n) \to 0$, c'est à dire

$$\|x^\beta \partial^\alpha f_n\|_\infty \to 0. \tag{24.308}$$

(5) Et enfin, la technique pour montrer que $T\colon \mathscr{S}(\mathbb{R}^d) \to \mathbb{C}$ est continue est de montrer que sous l'hypothèse d'avoir (24.308) pour tout choix de α et β, nous avons $T(f_n) \to 0$ dans $\mathbb{C}$.

Lemme 24.82 ([4])**.**
La topologie sur $\mathscr{S}(\mathbb{R}^d)$ est donnée aussi par les semi-normes

$$q_{n,m} = \max_{|\alpha|\leqslant n} \sup_{x\in\mathbb{R}^d} \big(1 + \|x\|\big)^m \big|\partial^\alpha \varphi(x)\big|. \tag{24.309}$$

Autrement dit, une suite $\varphi_n \xrightarrow{\mathscr{S}(\mathbb{R}^d)} 0$ si et seulement si $q_{n,m(\varphi)} \to 0$ pour tout n et m.

Le fait que les $q_{n,m}(\varphi)$ restent bornés est la proposition 24.78. Cependant ce lemme est plus précis parce qu'en disant seulement que φ est majoré par des polynôme, nous ne disons pas que les polynômes correspondants aux φ_n tendent vers zéro si $\varphi_n \xrightarrow{\mathscr{S}} 0$. Et d'ailleurs on ne sait pas très bien ce que signifierait $P_n \to 0$ pour une suite de polynômes.

Proposition 24.83.
Pour $p \in [1, \infty]$, l'espace $\mathscr{S}(\mathbb{R}^d)$ s'injecte continument dans $L^p(\mathbb{R}^d)$.

Démonstration. L'injection dont nous parlons est l'identité ou plus précisément l'identité suivie de la prise de classe. Il faut vérifier que cela est correct et continu, c'est à dire d'abord qu'une fonction à décroissance rapide est bien dans L^p et ensuite que si $f_n \xrightarrow{\mathscr{S}} 0$, alors $f_n \xrightarrow{L^p} 0$.

Commençons par $p = \infty$. Alors $\|f_n\|_\infty = p_{0,0}(f_n) \to 0$ parce que si $f_n \xrightarrow{\mathscr{S}} 0$, alors en particulier $p_{0,0}(f_n) \to 0$.

Au tour de $p < \infty$ maintenant. Nous savons qu'en dimension d, la fonction

$$x \mapsto \frac{1}{(1 + \|x\|)^s} \tag{24.310}$$

est intégrable dès que $s > d$. Pour toute valeur de m nous avons

$$\|\varphi\|_p^p = \int_{\mathbb{R}^d} |\varphi(x)|^p dx = \int_{\mathbb{R}^d} \frac{\big|(1 + \|x\|)^m \varphi(x)\big|^p}{\big(1 + \|x\|\big)^{mp}} \leqslant \int_{\mathbb{R}^d} \frac{q_{0,m}(\varphi)^p}{\big(1 + \|x\|\big)^{mp}}. \tag{24.311}$$

En choisissant m de telle sorte que $mp > d$, nous avons convergence de l'intégrale et donc $\|\varphi\|_p < \infty$. Nous retenons que

$$\|\varphi\|_p^p \leqslant C q_{0,m}(\varphi)^p \tag{24.312}$$

pour une certaine constance C et un bon choix de m.

Ceci prouve que $\mathscr{S}(\mathbb{R}^d) \subset L^p(\mathbb{R}^d)$. Nous devons encore vérifier que l'inclusion est continue. Si $\varphi_n \xrightarrow{\mathscr{S}} 0$, alors en particulier nous avons $q_{0,m}(\varphi_n) \to 0$ par le lemme 24.82. Par conséquent la majoration (24.312) nous dit que $\|\varphi_n\|_p \to 0$ également.

□

En résumé, si $\varphi_n \xrightarrow{\mathscr{S}(\mathbb{R}^d)} \varphi$ alors $\varphi_n \xrightarrow{L^p} \varphi$.

Théorème 24.84 ([179]).
Soit μ une mesure sur les boréliens de $\mathbb{R}^n$ finie sur les compacts. Alors $\mathscr{D}(\mathbb{R}^n)$ est dense dans $L^1(\mathbb{R}^n, \mathcal{B}or(\mathbb{R}^n), \mu)$.

Proposition 24.85 ([285]).
La partie $\mathscr{D}(\mathbb{R}^d)$ est dense dans $\mathscr{S}(\mathbb{R}^d)$.

Démonstration. Soit $f \in \mathscr{S}(\mathbb{R}^d)$, et ϕ, une fonction de $\mathscr{D}(\mathbb{R}^d)$ telle que $\phi(x) = 1$ pour $|x| \leqslant 1$ (l'existence de telles fonctions est discutée en 15.13.1). Soit aussi $\phi_k(x) = \phi(x/k)$. Nous posons

$$f_k(x) = \phi_k(x)f(x), \tag{24.313}$$

et nous allons prouver que pour tout multiindices α et γ,

$$p_{\alpha,\gamma}(f_k - f) = \|x^\gamma \partial^\alpha (f_k - f)\|_\infty \to 0. \tag{24.314}$$

Pour cela nous allons noter $\beta \leqslant \alpha$ lorsque β est un multiindice contenu dans α. En utilisant la dérivée du produit nous avons

$$(\partial^\alpha f_k)(x) = \sum_{\beta\leqslant\alpha} (\partial^{\alpha-\beta}\phi_k)(x)\partial^\beta f(x) \tag{24.315a}$$

$$= \sum_{\beta\leqslant\alpha} k^{-|\alpha-\beta|}(\partial^{\alpha-\beta}\phi)(x/k)(\partial^\beta f)(x) \tag{24.315b}$$

$$= \sum_{\beta<\alpha} k^{-|\alpha-\beta|}(\partial^{\alpha-\beta}\phi)(x/k)(\partial^\beta f)(x) + \phi(x/k)(\partial^\alpha f)(x). \tag{24.315c}$$

Nous devons donc étudier et majorer

$$\begin{aligned}\sup_{x\in\mathbb{R}^d} |x^\gamma \partial^\alpha (f_k - f)| \leqslant \sup \Big|x^\gamma \sum_{\beta<\alpha} k^{-|\alpha-\beta|}(\partial^{\alpha-\beta}\phi)(x/k)(\partial^\beta f)(x)\Big| \\ + \sup \big|x^\gamma \big(\phi(x/k) - 1\big)(\partial^\alpha f)(x)\big|\end{aligned} \tag{24.316}$$

En ce qui concerne le second terme, soit $\epsilon > 0$, vu que f est Schwartz, il existe R tel que

$$|x^\gamma(\partial^\alpha f)(x)| < \epsilon \tag{24.317}$$

dès que $\|x\| > R$. En prenant $k > R$,

$$|x^\gamma(\partial^\alpha f)(x)| \begin{cases} = 0 & \text{si } \|x\| < R \\ \leqslant \epsilon & \text{si } \|x\| > R. \end{cases} \tag{24.318}$$

En ce qui concerne le premier terme,

$$\sup_{x\in\mathbb{R}^d} \Big|x^\gamma \sum_{\beta<\alpha} k^{-|\alpha-\beta|}(\partial^{\alpha-\beta}\phi)(x/k)(\partial^\beta f)(x)\Big| \tag{24.319a}$$

$$\leqslant \frac{1}{k} \sup_x \Big| \sum_{\beta<\alpha} (\partial^{\alpha-\beta}\phi)(x/k)(x^\gamma \partial^\beta f)(x)\Big| \tag{24.319b}$$

$$= \frac{1}{k} \sup_x \Big| \sum_{\beta<\alpha} (\partial^{\alpha-\beta}\phi)(x/k) p_{\beta,\gamma}(f)\Big| \tag{24.319c}$$

La somme ne contient qu'un nombre fini de β différents, donc nous pouvons considérer un nombre K qui majore tous les $p_{\beta,\gamma}(f)$ en même temps. La partie avec ϕ peut être majorée par $\|\partial^{\alpha-\beta}\phi\|_\infty$ (qui est fini) dont nous pouvons prendre le maximum sur $\beta < \alpha$. Toute l'expression dans la somme est donc majorée par un nombre qui ne dépend ni de x ni de β. Vu que la somme est finie, elle est majorée par ce nombre multiplié par le nombre de termes dans la somme et au final

$$\sup_{x\in\mathbb{R}^d} \Big|x^\gamma \sum_{\beta<\alpha} k^{-|\alpha-\beta|}(\partial^{\alpha-\beta}\phi)(x/k)(\partial^\beta f)(x)\Big| \leqslant \frac{K'}{k}. \tag{24.320}$$

La limite $k \to \infty$ ne fait alors plus de doutes. □

Remarque 24.86.

Vu la topologie de $\mathscr{S}(\mathbb{R}^d)$ (définition 24.80), la convergence $f_k \xrightarrow{\mathscr{S}(\mathbb{R}^d)} f$ peut être exprimée par le fait que pour tout k, l,

$$t^k f_n^{(l)} \xrightarrow{unif} t^k f^{(l)}. \tag{24.321}$$

C'est à dire convergence uniforme de toutes les dérivées multipliées par n'importe quel polynôme.

24.8.2 Produit de convolution

Proposition 24.87 (Stabilité de Schwartz par convolution[25] [282])**.**
Si $\varphi \in L^1(\mathbb{R}^d)$ et $\psi \in \mathscr{S}(\mathbb{R}^d)$, alors $\varphi * \psi \in \mathscr{S}(\mathbb{R}^d)$.

Démonstration. Nous devons prouver que

$$p_{\alpha,\beta}(\varphi * \psi) = \sup_{x\in\mathbb{R}^d} |x^\beta(\partial^\alpha(\varphi * \psi))(x)| \tag{24.322}$$

est borné pour tout multiindices α et β. En appliquant $|\alpha|$ fois la proposition 24.41, nous mettons toutes les dérivées sur ψ : $\partial^\alpha(\varphi * \psi) = (\varphi * \partial^\alpha\psi)$. Cela étant fait, nous majorons

$$\begin{aligned}
\big|x^\beta(\varphi * \partial^\alpha\psi)(x)\big| &\leqslant |x^\beta| \int_{\mathbb{R}^d} |\varphi(y)| \underbrace{\big|(\partial^\alpha\psi)(x-y)\big|}_{\leqslant\|\partial^\alpha\psi\|_\infty} dy\Big| && \text{(24.323a)}\\
&\leqslant |x^\beta|\|\partial^\alpha\psi\|_\infty \int_{\mathbb{R}^d} |\varphi(y)|dy && \text{(24.323b)}\\
&\leqslant p_{\alpha,\beta}(\psi)\|\varphi\|_{L^1}. && \text{(24.323c)}
\end{aligned}$$

Par conséquent, $p_{\alpha,\beta}(\varphi * \psi) \leqslant \|\varphi\|_{L^1} p_{\alpha,\beta}(\psi) < \infty$. □

25. Définition 24.38.

Chapitre 25

Analyse complexe

25.1 Fonctions holomorphes

La dérivée complexe est discutée à la section 15.1.

25.1.1 Équations de Cauchy-Riemann

Notons que la formule (15.19) donne un **développement limité** pour les fonctions holomorphes. Si f est holomorphe en z_0 alors si z est dans un voisinage de z_0, il existe une fonction $s\colon \mathbb{R} \to \mathbb{C}$ telle que $\lim_{t\to 0} s(t)/t = 0$ et

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + s(|z - z_0|). \tag{25.1}$$

Nous introduisons les opérateurs

$$\frac{\partial}{\partial z} = \partial = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) \tag{25.2a}$$

$$\frac{\partial}{\partial \bar{z}} = \bar{\partial} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right) \tag{25.2b}$$

Si f est une fonction $\mathbb{C}$-dérivable représentée par la fonction $F = P + iQ$, les équations de Cauchy-Riemann signifient que $\Delta P = \Delta Q = 0$, c'est à dire que la fonction f a des composantes harmoniques.

Théorème 25.1.
Si $f \in C^1(\Omega)$ alors nous avons équivalence des faits suivants :

(1) f est holomorphe sur Ω,

(2) f vérifie $\partial_{\bar{z}} f = 0$.

Proposition 25.2.
Une application $f\colon \Omega \to \mathbb{C}$ est $\mathbb{C}$-dérivable sur Ω si et seulement si elle est différentiable et

$$\begin{cases} \dfrac{\partial u}{\partial x} = \dfrac{\partial v}{\partial y} & \text{(25.3a)} \\ \dfrac{\partial u}{\partial y} = -\dfrac{\partial v}{\partial x} & \text{(25.3b)} \end{cases}$$

où $f(x + iy) = u(x, y) + iv(x, y)$. Ces équations se notent de façon plus compacte

$$\frac{\partial f}{\partial \bar{z}} = 0. \tag{25.4}$$

Les équations (25.3) sont les équations de **Cauchy-Riemann**.

Démonstration. La différentielle de $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ est donnée par la matrice

$$T = \begin{pmatrix} \partial_x u(a) & \partial_y u(a) \\ \partial_x v(a) & \partial_y v(a) \end{pmatrix}. \tag{25.5}$$

Cette matrice est une similitude si et seulement si les équations de Cauchy-Riemann sont satisfaites. En effet si $1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $i = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, la matrice T est une similitude (écrivons $\alpha + i\beta$ son coefficient) si

$$\begin{cases} T(1) = \alpha + i\beta & \text{(25.6a)} \\ T(i) = -\beta + i\alpha, & \text{(25.6b)} \end{cases}$$

c'est à dire

$$T = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}. \tag{25.7}$$

Identifier cette matrice à (25.5) fournit le résultat annoncé. □

Proposition 25.3.
Si $f\colon \mathbb{C} \to \mathbb{C}$ est holomorphe, alors nous avons

$$df_{z_0} = (\partial_z f)(z_0) \tag{25.8}$$

au sens où l'opérateur linéaire $df_{z_0}\colon \mathbb{C} \to \mathbb{C}$ est l'opération de multiplication par le nombre complexe $(\partial_z f)(z_0)$.

Démonstration. Soit $f(x+iy) = f_1(x,y) + if_2(x,y)$ une fonction holomorphe. Les fonctions réelles f_1 et f_2 sont assujetties aux équations de Cauchy-Riemann de la proposition 25.2 :

$$\begin{cases} \partial_x f_1 = \partial_y f_2 & \text{(25.9a)} \\ \partial_x f_2 = -\partial_y f_1. & \text{(25.9b)} \end{cases}$$

Nous avons, en recourant à un petit abus de notation entre $f_i\colon \mathbb{R}^2 \to \mathbb{R}$ et $f_i\colon \mathbb{C} \to \mathbb{R}$:

$$\begin{aligned} df_{z_0}(u) &= \frac{d}{dt}\Big[f(z_0 + tu)\Big]_{t=0} && \text{(25.10a)} \\ &= \frac{d}{dt}\Big[f_1(z_0 + tu) + if_2(z_0 + tu)\Big]_{t=0} && \text{(25.10b)} \\ &= \partial_x f_1 u_1 + \partial_y f_1 u_2 + i\big(\partial_x f_2 u_1 + \partial_y f_2 u_2\big) && \text{(25.10c)} \\ &= \begin{pmatrix} \partial_x f_1 & \partial_y f_1 \\ \partial_x f_2 & \partial_y f_2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} && \text{(25.10d)} \\ &= \begin{pmatrix} \partial_x f_1 & \partial_y f_1 \\ -\partial_y f_1 & \partial_y f_2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. && \text{(25.10e)} \end{aligned}$$

En utilisant le lemme 15.4 nous reconnaissons la matrice de multiplication par le nombre $\partial_x f_1 - i\partial_y f_1$. Or justement,

$$\partial_z f = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) f = \frac{1}{2}\big(\partial_x f_1 + i\partial_x f_2 - i\partial_y f_1 + \partial_y f_2\big), \tag{25.11}$$

qui se réduit à $\partial_x f_1 - i\partial_y f_1$ lorsque nous y appliquons les équations de Cauchy-Riemann. □

25.1.2 Intégrales sur des chemins fermés

Lemme 25.4.
Si g est une fonction continue dans un ouvert $\Omega \subset \mathbb{C}$ et si g admet une primitive complexe sur Ω alors

$$\int_\gamma g(z)dz = 0 \tag{25.12}$$

pour tout chemin fermé γ de classe C^1 contenu dans Ω.

Démonstration. Nommons G une primitive de g. Par définition,

$$\int_\gamma g = \int_\gamma G' \tag{25.13a}$$
$$= \int_0^1 G'\big(\gamma(t)\big)\gamma'(t)dt \tag{25.13b}$$
$$= \int_0^1 (G\circ g\gamma)'(t)dt \tag{25.13c}$$
$$= G(\gamma(1)) - G\big(\gamma(0)\big) \tag{25.13d}$$
$$= 0 \tag{25.13e}$$

parce que le chemin est fermé : $\gamma(0)=\gamma(1)$. □

Lemme 25.5 (Goursat[286])**.**
Soit Ω un ouvert dans $\mathbb{C}$ et f une fonction continue sur Ω, holomorphe sur Ω moins éventuellement un point (nommé $z_1\in\Omega$). Soit T, un triangle[1] *fermé inclus à Ω. Alors nous avons*

$$\int_{\partial T} f(z)dz = 0. \tag{25.14}$$

Démonstration. Nous notons $\gamma=\partial T$. Dans la suite nous allons définir une suite de triangles $T^{(n)}$ et nous noterons $\gamma_n=\partial T^{(n)}$ avec une orientation que nous allons expliquer. Pour commencer nous posons $T^{(0)}=T$ et $\gamma_0=\partial T^{(0)}$.

Nous considérons le cas $z_1\notin T$, et nous posons

$$c = l(\gamma)^{-2}|\int_\gamma f|. \tag{25.15}$$

Notre objectif est de montrer que $c=0$. Soit A,B,C les trois sommes du triangle ; nous divisons le triangle de la façons suivante. D'abord nous considérons les points A',B,C' respectivement milieux de BC, AC et AB. En traçant le triangle $A'B'C'$, nous construisons quatre triangles que nous nommons $T_i^{(0)}$. Le théorème de Thalès assure que le périmètre de chacun des quatre triangles est la moitié du périmètre du grand triangle T.

Sur T nous choisissons l'orientation ABC. De façon à être «compatible», nous choisissons les orientations $AC'B'$, $BA'C'$ et $A'CB'$. La somme de ces trois triangles donne T plus le triangle $A'C'B'$. Par conséquent nous choisissons sur le triangle central l'orientation (inverse) $AB'C'$ de façon à avoir

$$\int_\gamma f = \sum_{i=1}^4 \int_{\partial T_i^{(0)}} f. \tag{25.16}$$

Cela implique que pour au moins un des quatre triangles (disons $T_k^{(0)}$ pour fixer les idées) nous ayons

$$\int_{\partial T_k^{(0)}} f \geqslant \frac{1}{4}\int_{\partial T^{(0)}} f \tag{25.17}$$

Nous notons $T^{(1)}$ ce triangle. Comme noté précédemment nous avons

$$l(\partial T^{(1)}) = \frac{1}{2}l(\partial T^{(0)}), \tag{25.18}$$

et donc

$$l(\gamma_1)^{-2}|\int_{\gamma_1}|f = 4l(\gamma_0)^{-2}|\int_{\gamma_1} f| \geqslant 4l(\gamma_0)^{-2}\frac{1}{4}|\int_{\gamma_0} f| = c. \tag{25.19}$$

En répétant le procédé nous construisons une suite de triangles $T^{(n)}$ qui satisfont toujours

$$l(\partial T^{(n)}) = \frac{1}{2^n}l(\partial T^{(0)}). \tag{25.20}$$

1. Nous considérons ici le triangle «plein».

Ces triangles forment une suite de fermés emboités dont le diamètre tend vers zéro. Leur intersection contient donc exactement un point (lemme 18.110) que nous nommons z_0 (et qui appartient évidemment à Ω). Étant donné que f est holomorphe nous utilisons le développement limité (25.1) autour de z_0 :

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + s(|z - z_0|)(z - z_0) \tag{25.21}$$

avec $\lim_{t\to 0} s(t) = 0$. Nous posons $g(z) = f(z_0) + f'(z_0)(z - z_0)$ et nous considérons $\epsilon > 0$. Soit $\alpha > 0$ tel que

$$|f(z) - g(z)| < \epsilon|z - z_0| \tag{25.22}$$

pour tout $|z - z_0| < \alpha$. Le α à choisir pour obtenir cet effet est celui qui donne $s(|z - z_0|) < \epsilon$. Soit $N \in \mathbb{N}$ tel que $l(\gamma_n) < \alpha$ pour tout $n > N$. D'autre part, deux points dans un triangle sont toujours à distance moindre que la longueur d'un côté, donc pour tout $z \in T^{(n)}$ nous avons $|z - z_0| < \alpha$ et par conséquent pour tout z dans $T^{(n)}$ nous avons

$$|f(z) - g(z)| < \epsilon|z - z_0|. \tag{25.23}$$

Notons que la fonction g est une dérivée : c'est la dérivée de la fonction

$$G(z) = zf(z_0) + \frac{1}{2}f'(z_0)(z - z_0)^2. \tag{25.24}$$

Par conséquent nous avons

$$\int_{\gamma_n} g = 0 \tag{25.25}$$

par le lemme 25.4. Nous avons donc

$$\left|\int_{\gamma_n} f\right| = \left|\int_{\gamma_n} (f - g)\right| \tag{25.26a}$$

$$\leqslant l(\gamma_n) \max\{|f(z) - g(z)| \text{ tel que } z \in T^{(n)}\} \tag{25.26b}$$

$$\leqslant \epsilon l(\gamma_n)^2, \tag{25.26c}$$

et par conséquent

$$c \leqslant l(\gamma_n)^{-2} \left|\int_{\gamma_n} f\right| \leqslant \epsilon, \tag{25.27}$$

ce qui signifie que $c = 0$ parce que ϵ est arbitraire. Nous avons donc prouvé le lemme de Goursat dans le cas où le point de non holomorphie z_1 est en dehors de T.

Si z_1 est sur un côté, disons sur le côté AB, alors nous considérons un vecteur $v \in \mathbb{C}$ tel que $T_\epsilon = T + \epsilon v$ ne contienne z_1 pour aucun ϵ. Le vecteur $v = z_1 - C$ fait par exemple l'affaire. En vertu du point précédent nous avons

$$\int_{\partial T_\epsilon} f = 0 \tag{25.28}$$

pour tout $\epsilon > 0$. Étant donné que la fonction f est continue (y compris en z_1), l'intégrale sur ∂T est également nulle.

Si maintenant le point z_1 est à l'intérieur de T nous décomposons T en trois triangles ayant z_1 comme sommet commun. Si nous considérons les orientations Az_1C, ABz_1 et BCz_1, alors nous avons

$$\int_T f = \int_{Az_1C} f + \int_{ABz_1} f + \int_{BCz_1} f, \tag{25.29}$$

alors que par le point précédent les trois intégrales du membre de droite sont nulles. □

Proposition 25.6 ([286])**.**
Soit Ω un ouvert étoilé et f une fonction holomorphe sur Ω sauf éventuellement en un point z_1 où f est seulement continue. Alors si γ est un chemin fermé dans Ω, nous avons

$$\int_\gamma f = 0. \tag{25.30}$$

Proposition 25.7.
Si $f(z) = \sum_n a_n z^n$ a pour rayon de convergence R, alors f est $\mathbb{C}$-dérivable et nous pouvons dériver terme à terme dans la boule ouverte $B(0,R)$.

Démonstration. Cela est exactement la proposition 15.41. □

25.1.3 Lacets, indice et homotopie

Proposition-définition 25.8.
Soit γ un chemin fermé[2] *dans $\mathbb{C}$. L'**indice** de la courbe γ est la fonction*

$$\begin{aligned}\mathrm{Ind}_\gamma\colon \mathbb{C}\backslash\gamma &\to \mathbb{Z}\\ z &\mapsto \frac{1}{2\pi i}\int_\gamma \frac{d\omega}{\omega - z}.\end{aligned} \tag{25.31}$$

*Un chemin continu et fermé (au sens $\gamma(1) = \gamma(0)$) est un **lacet**.*

(1) La fonction Ind_γ est continue et prend effectivement des valeurs entières.

(2) La fonction indice est constante sur chaque composante connexe de $\mathbb{C}\backslash\gamma$ et est nulle sur la composante non bornée.

Le second point est en partie la proposition 6.85.

Définition 25.9.
*Si γ_1 et γ_2 sont deux lacets en $x_0 \in X$ (un espace topologique), une **équivalence d'homotopie** est une application $f\colon [0,1]\times[0,1] \to X$ telle que*

(1) $f(0,t) = \gamma_1(t)$ pour tout t ;

(2) $f(1,t) = \gamma_1(t)$ pour tout t ;

(3) pour chaque $t \in [0,1]$, l'application $s \mapsto f(s,t)$ est continue ;

(4) pour chaque $s \in [0,1]$, l'application $t \mapsto f(s,t)$ est un lacet basé en x_0.

Exemple 25.10
Si γ est un cercle de centre $z_0 \in \mathbb{C}$ et de rayon r, alors

$$\mathrm{Ind}_\gamma(z) = \begin{cases} 2\pi i & \text{si } z \in B(z_0, r)\\ 0 & \text{sinon.}\end{cases} \tag{25.32}$$

La seconde ligne provient directement de la proposition 25.8. Pour la première, le cercle γ se paramètre par

$$\gamma(\theta) = z_0 + re^{i\theta}, \tag{25.33}$$

et l'intégrale vaut

$$\int_\gamma \frac{d\omega}{\omega - z_0} = \int_0^{2\pi} \frac{1}{re^{i\theta}} ire^{i\theta} d\theta = 2\pi i. \tag{25.34}$$

L'indice de ce chemin va évidemment jouer un rôle particulier dans la suite. △

Théorème 25.11 (Cauchy, version homotopique[287])**.**
Soit Ω un ouvert de $\mathbb{C}$ et f une fonction holomorphe sur Ω. Si γ_1 et γ_2 sont deux lacets homotopes de classe C^1 dans Ω, alors

$$\int_{\gamma_1} f(z)dz = \int_{\gamma_2} f(z)dz. \tag{25.35}$$

Corollaire 25.12 ([287])**.**
Soit $a \in \mathbb{C}$ ainsi que deux chemins γ_1 et γ_2 homotopes dans $\mathbb{C}\backslash\{a\}$. Alors $\mathrm{Int}(\gamma_1, a) = \mathrm{Ind}(\gamma_2, a)$.

Il y a aussi des choses sur l'indice dans [286].

2. Par abus de langage, nous désignerons par γ à la fois le chemin et son image.

25.1.4 Théorème de Cauchy et analycité

Cette sous-section veut prouver le théorème de Cauchy. Comme d'habitude, une référence qui ne peut pas rater est [286].

Théorème 25.13 (Formule de Cauchy)**.**
Soit Ω ouvert dans $\mathbb{C}$, $z_0 \in \Omega$ et f, une fonction holomorphe sur Ω. Soit $r > 0$ tel que $B(z_0, r) \subset \Omega$. Alors pour tout $z \in B(z_0, r)$ nous avons

$$f(z) = \frac{1}{2\pi i} \int_{\partial B(z_0,r)} \frac{f(\omega)}{\omega - z} d\omega. \tag{25.36}$$

Démonstration. Soit $z \in B(z_0, r)$ et considérons la fonction

$$g(\omega) = \begin{cases} \frac{f(\omega)-f(z)}{\omega-z} & \text{si } \omega \neq z \\ f'(z) & \text{si } \omega = z. \end{cases} \tag{25.37}$$

Cette fonction est holomorphe sur $B(z_0, r)\backslash\{z\}$ et continue en z. Elle vérifie donc la proposition 25.6 et nous avons

$$\int_\gamma g = 0 \tag{25.38}$$

où γ est le cercle de centre z_0 et de rayon r. Nous avons donc

$$0 = \int_\gamma \frac{f(\omega)}{\omega - z} - \int_\gamma \frac{f(z)}{\omega - z}, \tag{25.39}$$

et ayant déjà calculé la seconde intégrale dans l'exemple 25.10 nous en déduisons

$$\int_\gamma \frac{f(\omega)}{\omega - z} d\omega = 2\pi i f(z), \tag{25.40}$$

ce qu'il fallait. □

Définition 25.14.
*Une fonction $f : \Omega \to \mathbb{C}$ est $\mathbb{C}$-**analytique** sur Ω si pour tout $z_0 \in \Omega$, il existe une suite complexe (c_n) et $r > 0$ tels que*

$$f(z) = \sum_{n=0}^{\infty} c_n (z - z_0)^n \tag{25.41}$$

pour tout $z \in B(z_0, r)$.

Théorème 25.15.
Soit Ω ouvert dans $\mathbb{C}$ et f, holomorphe sur Ω. Soient encore $z_0 \in \Omega$ et r_0 tel que $B(z_0, r_0) \subset \Omega$. Alors :

(1) Sur $B(z_0, r_0)$, la fonction f s'écrit

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n. \tag{25.42}$$

(2) Nous avons

$$a_n = \frac{f^{(n)}(z_0)}{n!} = \frac{1}{2\pi i} \int_\gamma \frac{f(\omega)}{(\omega - z_0)^{n+1}} d\omega \tag{25.43}$$

où $\gamma = \partial B(z_0, r)$ avec $|z - z_0| < r < r_0$.

(3) En particulier f est infiniment dérivable.

Démonstration. Soit $r > 0$ tel que $|z - z_0| < r < r_0$. La formule de Cauchy (théorème 25.13) nous dit que

$$f(z) = \frac{1}{2\pi i} \int_\gamma \frac{f(\omega)}{\omega - z} d\omega \tag{25.44}$$

où $\gamma = \partial B(z_0, r)$. Nous pouvons paramétrer ce chemin par $\omega = z_0 + re^{i\theta}$ et $\theta \in [0, 2\pi]$. Nous avons

$$f(z) = \frac{1}{2\pi i} \int_0^{2\pi} \frac{f(z_0 + re^{i\theta})}{z_0 + re^{i\theta} - z} rie^{i\theta} d\theta \tag{25.45a}$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \frac{f(z_0 + re^{i\theta})}{1 - e^{-i\theta}(z - z_0)/r} d\theta. \tag{25.45b}$$

Nous pouvons développer l'intégrante en puissance de $(z - z_0)$ en utilisant la formule 15.252. Ici le rôle de x est tenu par

$$e^{-i\theta}(z - z_0)/r \tag{25.46}$$

dont le module est bien plus petit que 1, par hypothèse sur r. Nous avons donc

$$f(z) = \frac{1}{2\pi} \int_0^{2\pi} \sum_{n=0}^{\infty} f(z_0 + re^{i\theta}) e^{-in\theta} r^{-n} (z - z_0)^n d\theta. \tag{25.47}$$

L'art est maintenant de permuter la somme et l'intégrale. Pour cela nous remarquons que ce qui se trouve dans la somme est majoré en module par

$$M \left| \frac{z - z_0}{r} \right|^n \tag{25.48}$$

où M est le maximum de $|f|$ sur γ. La borne (25.48) ne dépend pas de θ ; par conséquent la convergence de la somme est uniforme en θ par le critère de Weierstrass (théorème 15.16). Le théorème 15.18 s'applique [3] et nous pouvons permuter la somme avec l'intégrale.

Ce que nous trouvons est que

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \tag{25.49}$$

où

$$a_n = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + re^{i\theta}) e^{-in\theta} r^{-n} d\theta = \frac{1}{2\pi i} \int_\gamma \frac{f(\omega)}{(\omega - z_0)^{n+1}}. \tag{25.50}$$

Cette formule est valable pour $|z - z_0| < r$. Sur cette boule, la fonction est donc une série entière. Le théorème de Taylor 15.77 nous permet donc d'affirmer que f est partout infiniment continument dérivable (parce que en chaque point on a un voisinage sur lequel c'est vrai), et d'identifier les coefficients (qui, eux, ne sont valables que localement) sous la forme

$$a_n = \frac{f^{(n)}(z_0)}{n!}. \tag{25.51}$$

□

Corollaire 25.16.
Soit f une fonction continue sur un ouvert Ω telle que pour toute boule $B(a, r)$ contenue dans Ω, nous ayons

$$f(a) = \frac{1}{2\pi i} \int_{\partial B(a,r)} \frac{f(\xi)}{\xi - a} d\xi. \tag{25.52}$$

Alors f est holomorphe.

3. Étant donné que nous savions déjà que la somme était une fonction intégrable, nous sommes loin d'avoir utilisé toute la puissance du théorème.

Démonstration. Il suffit de recopier la démonstration du théorème 25.15 pour savoir que f se développe en série de puissances et est donc en particulier dérivable. □

Le fait qu'une fonction holomorphe soit C^∞ comme dit dans la proposition 25.15 permet de démonter un résultat de dérivation sous l'intégrale, qui dépend de pouvoir majorer la différentielle.

Proposition 25.17.
Soit une fonction continue $g\colon \mathbb{R}\times\mathbb{C}\to\mathbb{C}$. Nous supposons que pour tout t, la fonction $z\mapsto g(t,z)$ est $\mathbb{C}$-dérivable (définition 15.1) et différentiable. Soit B compact dans $\mathbb{R}$ et la fonction

$$G(z)=\int_B g(t,z)dt. \tag{25.53}$$

que nous supposons exister pour tout z.

Alors

$$G'(z)=\int_B g'(t,z)dt \tag{25.54}$$

où le prime réfère à la $\mathbb{C}$-dérivée par rapport à la variable z à t fixé.

Démonstration. Nous fixons $z\in\mathbb{C}$ et nous considérons la suite de fonctions

$$g_i(t)=\frac{g(t,z+\epsilon_i)-g(t,z)}{\epsilon_i} \tag{25.55}$$

où ϵ_i est une suite de nombres complexes tendant vers zéro ($\epsilon_i \xrightarrow{\mathbb{C}} 0$). Si la limite existe et ne dépend pas de la suite choisie, alors $\lim_{i\to\infty} g_i(t)=g'(t,z)$. Et vu que g est supposée dérivable, c'est le cas.

Nous avons aussi, par linéarité de l'intégrale :

$$G'(z)=\lim_{i\to\infty}\int_B g_i(t)dt. \tag{25.56}$$

La difficulté est de permuter la limite et l'intégrale. Pour cela nous allons utiliser la convergence dominée de Lebesgue (théorème 13.134). Afin de majorer $|g_i(t)|$ par une fonction intégrable en t (uniformément en i), nous exploitons le théorème des accroissements finis, théorème 12.229. En notant dg la différentielle de g par rapport à z à t fixé, pour chaque t et chaque i nous avons

$$|g(t,z+\epsilon_i)-g(t,z)|\leqslant \sup_{\xi\in[z,z+\epsilon_i]}\|dg_\xi\|\|\epsilon_i\|. \tag{25.57}$$

Vu que z est fixé et que ξ est dans le compact $[z,z+\epsilon_i]$ et que dg est continue (parce que la $\mathbb{C}$-dérivabilité implique la continuité de la différentielle parce que nous avons l'analycité par le théorème 25.15), nous pouvons majorer $\|dg_\xi\|$ par une constante $M_i(z)$ qui dépend a priori de i et de z.

Heureusement, nous pouvons prendre a fortiori le supremum sur $\overline{B(z,|\epsilon_i|)}$ (qui est tout autant compact) et supposer que $|\epsilon_i|$ est strictement décroissante ; de toutes façons, il y a un maximum parce que $|\epsilon_i|\to 0$. Dans ce cas, il suffit de prendre le supremum de $\|dg_\xi\|$ pour $\xi\in\overline{B(z,|\epsilon_1|)}$ et ça contente tout le monde.

Quoi qu'il en soit nous avons une constante $M(z)$ telle que

$$|g(t,z+\epsilon_i)-g(t,z)|\leqslant M(z)\|\epsilon_i\| \tag{25.58}$$

et donc $|g_i(t)|\leqslant M(z)$. La constante (par rapport à t) $M(z)$ est évidemment intégrable sur le compact B et nous pouvons permuter la limite avec l'intégrale :

$$G'(z)=\lim_{i\to\infty}\int_B g_i(t)dt=\int_B\lim_{i\to\infty}g_i(t)dt=\int_B g'(t,z)dt. \tag{25.59}$$

□

Proposition 25.18.
Une fonction continue f est holomorphe si et seulement si la 1-forme différentielle $f(z)dz$ est localement exacte.

Démonstration. Si f est holomorphe, alors nous avons vu que f était différentiable et que $df_z = f(z)dz$ par la formule 15.25.

Dans le sens inverse, supposons que $f(z)dz$ est localement exacte, et soit F telle que $dF = f(z)dz$. Ce que nous allons faire est montrer que la dérivée de F existe et vaut f. En effet, la définition de la différentielle nous dit que

$$\lim_{h\to 0}\left|\frac{F(z+h)-F(z)-dF_z(h)}{h}\right| = 0. \tag{25.60}$$

La limite vaut évidemment encore zéro si nous enlevons les modules :

$$0 = \lim_{h\to 0}\frac{F(z+h)-F(z)-f(z)h}{h} \tag{25.61a}$$

$$= \lim_{h\to 0}\frac{F(z+h)-F(z)}{h} - f(z). \tag{25.61b}$$

Donc $F' = f$. Cela montre que F est $\mathbb{C}$-dérivable et donc holomorphe. En conséquence du théorème 25.15, la fonction F est infiniment dérivable et f l'est alors aussi. La fonction f est donc holomorphe [4]. □

25.1.5 Théorème de Brouwer en dimension 2

Pour d'autres versions du théorème de Brouwer, voir la sous-section 18.3.3.

Théorème 25.19 (Brouwer en dimension 2[2])**.**
Soit $\mathcal{B}$ la boule unité fermée de $\mathbb{R}^2$. Alors toute application continue de $\mathcal{B}$ dans elle-même admet un point fixe.

Démonstration. Supposons que la fonction $f \in C^0(\mathcal{B},\mathcal{B})$ n'admette pas de points fixes sur $\mathcal{B} = \overline{B(0,1)}$. Pour $x \in \mathcal{B}$ nous notons $g(x)$ l'intersection entre $\partial\mathcal{B}$ et la demi-droite allant de $f(x)$ vers x. C'est bien parce que f n'a pas de points fixes que g est bien définie.

En reprenant le même début de la preuve de la proposition 18.31 nous savons que la fonction

$$\begin{aligned} g\colon \overline{B(0,1)} &\to \partial B(0,1)\\ x &\mapsto \lambda(x)\big(x - f(x)\big) + f(x) \end{aligned} \tag{25.62}$$

est continue. De plus $g(x) = x$ sur $\partial B(0,1)$. Nous allons montrer qu'une telle fonction [5] ne peut pas exister.

Pour $s \in [0,1]$ nous paramétrons le cercle $\partial B(0,s)$ par

$$\begin{aligned} x_s\colon [0,1] &\to \partial B(0,1)\\ t &\mapsto \big(s\cos(2\pi t), s\sin(2\pi t)\big). \end{aligned} \tag{25.63}$$

Ensuite nous considérons les chemins

$$\begin{aligned} \gamma_s\colon [0,1] &\to \partial B(0,s)\\ t &\mapsto g\circ x_s. \end{aligned} \tag{25.64}$$

L'application γ_s est continue et $\gamma_s(0) = \gamma_s(1)$. Les chemins γ_s sont des lacets ; nous nous intéressons maintenant à l'indice au point 0 de γ_0 et γ_1. D'une part $\gamma_0(t) = g(0)$ (lacet constant) et $\gamma_1(t) = e^{2i\pi t}$ (parce que $g(x) = x$ sur le bord). Nous avons donc, en utilisant l'indice de la définition 25.8,

$$\mathrm{Ind}_{\gamma_0}(0) = \frac{1}{2\pi i}\mathrm{Ind}_{\gamma_0}\frac{d\omega}{\omega} = \frac{1}{2\pi i}\int_0^1 \frac{\gamma_0'(t)}{\gamma_0(t)}dt = 0, \tag{25.65}$$

4. Dire que la dérivée d'une fonction holomorphe est holomorphe est un raisonnement classique.
5. Qui est nommée *rétraction* de la sphère sur elle-même.

alors que

$$\mathrm{Ind}_{\gamma_1}(0) = \frac{1}{2\pi i}\int_0^1 \frac{2i\pi e^{2i\pi t}}{e^{2i\pi t}}dt = 1. \tag{25.66}$$

Nous considérons l'homotopie

$$\begin{aligned}\gamma\colon [0,1]\times[0,1] &\to \overline{B(0,1)}\\ (s,t) &\mapsto \gamma_s(t) = (g\circ x_s)(t).\end{aligned} \tag{25.67}$$

Nous avons $g(0) \neq 0$ parce que g prend ses valeurs sur le bord. Vu que c'est une équivalence d'homotopie[6] entre γ_1 et γ_2, les indices devraient être égaux par le corollaire 25.12. □

25.1.6 Principe des zéros isolés

Théorème 25.20 (Principe des zéros isolés [286])**.**
Soit f une fonction holomorphe et a, une zéro non isolé de f. Alors f est nulle sur un voisinage de a.

Démonstration. Nous écrivons f sous la forme d'une série entière autour de a :

$$f(z) = \sum_{n=0}^{\infty} c_n(z-a)^n \tag{25.68}$$

valable sur une boule $B(a,r)$. Soit c_m le premier coefficient non nul (si il n'existe pas c'est que f est nulle sur tout $B(a,r)$ et alors le théorème est prouvé). Nous avons alors

$$f(z) = c_m(z-a)^m\big(1 + \sum_{k=1}^{\infty} d_k(z-a)^k\big) \tag{25.69}$$

avec $d_k = c_{m-k}$. Le rayon de convergence de la série $\sum_k d_k(z-a)^k$ est le même que celui de (25.68) parce que la suite $d_k r^{m+k}$ reste bornée (critère d'Abel, lemme 15.26). Si nous posons

$$g(z) = 1 + \sum_{k=1}^{\infty} d_k(z-a)^k, \tag{25.70}$$

alors g est une fonction continue et $g(a) = 1$. De plus

$$f(z) = c_m(z-a)^m g(z). \tag{25.71}$$

Soit une suite (z_n) de zéros de f convergent vers a. Étant donné que g est continue, nous devrions avoir $\lim_{k\to\infty} g(z_k) = g(a) = 1$, mais si $f(z_k) = 0$ avec $z_k \neq a$, alors $g(z_k) = 0$. Cela est un paradoxe qui nous permet de conclure que si la suite z_n existe bien, alors f est identiquement nulle sur un voisinage, c'est à dire que tous les c_n sont nuls. □

Corollaire 25.21.
Soit f une fonction holomorphe sur un ouvert connexe Ω. Si f s'annule sur un un ouvert (non vide) de Ω, alors f s'annule sur tout Ω.

Démonstration. soit

$$N = \{z \in \Omega \text{ tel que } f = 0 \text{ sur un ouvert autour de } z\}. \tag{25.72}$$

Le fait que N soit ouvert est évident à partir de sa définition. Nous allons montrer que N est également fermé dans Ω, et donc conclure que $N = \Omega$. Soit (z_n) une suite dans N convergente vers $z \in \Omega$. Étant donné que $f(z_n) = 0$ et que f est continue, nous avons

$$f(z) = \lim_{n\to\infty} f(z_n) = 0, \tag{25.73}$$

ce qui fait de z un zéro non isolé de f. Par conséquent le principe des zéros isolés (théorème 25.20) nous enseigne que f s'annule dans un voisinage autour de z, c'est à dire que $z \in N$. L'ensemble N est donc fermé. □

6. Définition 25.9

25.1.7 Prolongement de fonctions holomorphes

Proposition 25.22.
Soit Ω, un ouvert de $\mathbb{C}$ et $f : \Omega \to \mathbb{C}$ une fonction holomorphe sur $\Omega\backslash\{a\}$ ($a \in \Omega$). Nous supposons qu'il existe $r > 0$ tel que f est bornée sur $B(a,r) \cap \Omega$. Alors f se prolonge en une fonction holomorphe sur Ω.

Le théorème de prolongement de Riemann 25.36 donnera plus d'informations.

Démonstration. Nous définissons la fonction $g : \Omega \to \mathbb{C}$ par

$$g(z) = \begin{cases} (z-a)f(z) & \text{si } z \neq a \\ 0 & \text{si } z = a. \end{cases} \tag{25.74}$$

Sur $\Omega\backslash\{a\}$, la fonction g est holomorphe (produit de fonctions holomorphes), et elle est continue en a. Par conséquent elle est holomorphe sur Ω. Nous la développons en série entière sur une boule $B(a,r)$:

$$g(z) = \sum_{n=0}^{\infty} c_n (z-a)^n. \tag{25.75}$$

Nous avons $g(a) = c_0 = 0$. Nous posons

$$\varphi(z) = \sum_{n=0}^{\infty} c_{n+1}(z-a)^n. \tag{25.76}$$

Si $z \neq a$, alors $\varphi(z) = f(a)$ parce que $\varphi(z) = g(z)/(z-a)$. Mais φ est continue en a, et donc holomorphe en a.

La fonction φ est par conséquent un prolongement holomorphe de f en a. □

25.1.8 Théorème de Runge

Le théorème que nous allons prouver n'est en réalité qu'une partie de ce qui est usuellement appelle le théorème de Runge.

Théorème 25.23 (Théorème de Runge)**.**
Soit K, un compact de $\mathbb{C}$ tel que $\complement K$ soit connexe. Si $a \in \complement K$ alors la fonction

$$\varphi_a(z) = \frac{1}{z-a} \tag{25.77}$$

est limite uniforme de polynômes sur K.

Démonstration. Nous considérons $P(K)$, l'adhérence des polynômes sur K pour la norme uniforme (sur K). Nous devons montrer que pour tout $a \in \complement K$, la fonction φ_a est dans $P(K)$. Pour cela nous considérons l'ensemble

$$A = \{a \in \complement K \text{ tel que } \varphi_a \in P(K)\} \tag{25.78}$$

et nous allons montrer qu'il est à la fois non vide, ouvert et fermé dans le connexe $\complement K$.

Je répète : nous allons prouver l'ouverture et la fermeture *pour la topologie de* $\complement K$. Nous n'allons pas prouver que A est un ouvert de $\mathbb{C}$. Ce qui sera par conséquent prouvé est que $A = \complement K$.

Non vide Soit $R = \sup_{z\in K} |z|$ et $a \in \complement K$ tel que $|a| > R$. Nous avons

$$\varphi_a(z) = \frac{1}{a}\frac{1}{\frac{z}{a}-1} = -\frac{1}{a}\frac{1}{1-\frac{z}{a}} = -\frac{1}{a}\sum_{k=0}^{\infty}\left(\frac{z}{a}\right)^k = \sum_{k=0}^{\infty}\frac{z^k}{a^{k+1}}. \tag{25.79}$$

Ici la convergence de la série et sa limite sont assurées par le fait que $|z/a| < 1$ par choix de R et a. La suite de polynômes

$$P_n(z) = \sum_{k=0}^{n} \frac{z^k}{a^{k+1}} \tag{25.80}$$

converge uniformément sur $B(0, R)$ et en particulier sur K. Donc $P_n \to \varphi_a$.

Fermé Nous allons montrer que la fermeture de A (dans $\complement K$) est inclue dans A, et donc qu'elle est égale à A et donc que A est fermé. Par le lemme 6.79, la fermeture de A dans $\complement K$ est l'ensemble $\bar{A} \cap \complement K$ où $\bar{A}$ est la fermeture de A au sens usuel.

Bref, soit $a \in \bar{A} \cap \complement K$, et montrons que $\varphi_a \in \overline{P(K)}$. Vu que $P(K)$ est déjà une fermeture, nous aurons en fait $\varphi_a \in P(K)$ et donc $a \in A$, ce qui signifierait que $\bar{A} \cap \complement A = A$ et donc que A est fermé.

Au travail.

Soit $(a_n) \in A$ une suite convergente vers a. Soit aussi $d = d(a, K)$; on a $d > 0$ parce que K est compact et a est hors de a alors le complémentaire de K est ouvert. Nous choisissons en plus la suite a_n pour avoir $|a_n - a| < \frac{d}{2}$; au pire on prend la queue de suite. Soit $z \in K$; nous avons

$$|\varphi_{a_n}(z) - \varphi_a(z)| = \left| \frac{1}{z - a_n} - \frac{1}{z - a} \right| = \left| \frac{a_n - a}{(z - a_n)(z - a)} \right|. \tag{25.81}$$

Vu que $a_n \in B(a, \frac{d}{2})$ et que $z \in K$ et $d = d(a, K)$ nous avons $|a_n - z| \geqslant \frac{d}{2}$; et aussi $|a - z| \geqslant \frac{d}{2}$. Nous pouvons donc majorer (25.81) par

$$|\varphi_{a_n}(z) - \varphi_a(z)| \leqslant 2\frac{|a_n - a|}{d^2}. \tag{25.82}$$

Donc nous avons

$$\|\varphi_a - \varphi_{a_n}\|_K \leqslant 2\frac{|a_n - a|}{d^2} \to 0 \tag{25.83}$$

où la norme $\|.\|_K$ est la norme supremum sur K. Donc $a \in \overline{P(K)} = P(K)$ et A est fermé.

Ouvert Vu que K est compact, il est fermé et donc $\complement K$ est ouvert. Par conséquent, ainsi que précisé dans l'exemple 6.80, les ouverts de $\complement K$ sont les ouverts de $\mathbb{C}$ contenus dans $\complement K$. Afin de prouver que A est ouvert, nous prenons $a \in A$ et nous cherchons une boule (au sens de $\mathbb{C}$) autour de a qui serait incluse dans A.

Soit donc $h \in \mathbb{C}$ «petit» dans un sens que nous allons préciser plus tard. Encore une fois nous posons $d = d(a, K)$. Nous avons

$$\varphi_{a+h}(z) = \frac{1}{z - a - h} = \frac{1}{z - a} \frac{1}{1 - \frac{h}{z-a}} = \sum_{k=0}^{\infty} \frac{h^k}{(z - a)^{k+1}}. \tag{25.84}$$

Déjà ici nous demandons $h < \sup_{z \in K} |z - a|$. Puisque $|z - a| > d$, nous avons alors

$$|\varphi_{a+h}(z)| \leqslant \sum_{k=0}^{\infty} \frac{h^k}{d^{k+1}} < \infty. \tag{25.85}$$

Cela pour dire que la somme à droite de (25.84) converge bien pourvu que h soit bien petit. Nous pouvons donc poursuivre :

$$\varphi_{a+h}(z) = \sum_{k=0}^{\infty} \frac{h^k}{(z - a)^{k+1}} = \sum_{k=0}^{\infty} h^k \varphi_a(z)^{k+1}. \tag{25.86}$$

Nous montrons maintenant que la convergence de la somme (25.86) est en réalité uniforme en z. En effet

$$\left|\varphi_{a+h}(z) - \sum_{k=0}^{N} h^k \varphi_a(z)^{k+1}\right| = \left|\sum_{k=N+1}^{\infty} h^k \varphi_a(z)^{k+1}\right| \tag{25.87a}$$

$$\leqslant \sum_{k=N+1}^{\infty} |h|^k |\varphi_a(z)|^{k+1}. \tag{25.87b}$$

Étant donné que φ_a est continue sur le compact K, elle y est majorée en module ; on peut même être plus précis :

$$|\varphi_a(z)| = \frac{1}{|z-a|} \leqslant \frac{1}{d}. \tag{25.88}$$

Nous pouvons donc écrire

$$\left|\varphi_{a+h}(z) - \sum_{k=0}^{N} h^k \varphi_a(z)^{k+1}\right| \leqslant \frac{1}{d} \sum_{k=N+1}^{\infty} \left|\frac{h}{d}\right|^k. \tag{25.89}$$

Étant donné que la somme $\sum_{k=0}^{\infty} |h/d|^k$ converge, la limite $N \to \infty$ est nulle et nous avons

$$\lim_{N\to\infty} \|\varphi_{a+h} - \sum_{k=0}^{N} h^k \varphi_a^{k+1}\|_K = 0. \tag{25.90}$$

Pour avoir $\varphi_{a+h} \in P(K)$, il faut encore savoir si les fonctions φ_a^k sont dans $P(K)$ pour tout k. Dans ce cas pour chaque N la somme sera encore dans $P(K)$ et φ_{a+h} sera limite uniforme d'éléments de $P(K)$.

Par hypothèse, $\varphi_a \in P(K)$; soit P_n une suite de polynômes qui converge uniformément vers φ_a. Nous allons montrer qu'alors la suite de polynômes P_n^k converge uniformément vers φ_a^k. Soit n tel que $\|P_n - \varphi_a\|_K \leqslant \epsilon$ et utilisons le produit remarquable

$$a^k - b^k = (a-b) \sum_{i=0}^{k-1} a^i b^{k-1-i} \tag{25.91}$$

pour obtenir

$$|P_n(z)^k - \varphi_a(z)^k| \leqslant |P_n(z) - \varphi_a(z)| \sum_{i=0}^{k-1} |P_n(z)^i \varphi_a(z)^{k-1-i}|. \tag{25.92}$$

Vu que P_n et φ_a sont continues sur le compact K, on peut majorer la somme par une constante M, et il restera

$$|P_n(z)^k - \varphi_a(z)^k| \leqslant M|P_n(z) - \varphi_a(z)|, \tag{25.93}$$

ou encore

$$\|P_n^k - \varphi_a^k\| \leqslant M\epsilon. \tag{25.94}$$

Cela prouve que $\varphi_a^k \in P(K)$ et donc que φ_{a+h} est limite uniforme (sur K) d'éléments de $P(K)$ et donc fait partie de $P(K)$ lui aussi.

Ceci achève de prouver que A est ouvert dans $\complement K$.

Conclusion L'ensemble A est non vide, ouvert et fermé dans $\complement K$, donc il est égal à $\complement K$. Le théorème est ainsi démontré.

□

25.2 Intégrales de fonctions holomorphes

Nous commençons par le lemme technique.

Lemme 25.24 ([286])**.**
Soit f une fonction holomorphe sur $B(z_0, r_0)$. Pour tout $z \in B(z_0, r)$ (avec $r < r_0$) nous avons

$$|f'(z)| \leqslant \frac{r}{\big(r - |z-z_0|\big)^2} \max\big\{f(z_0 + re^{i\theta})\big\}_{\theta\in\mathbb{R}}. \tag{25.95}$$

Démonstration. Par translation nous pouvons supposer que $z_0 = 0$. Étant donné que f est holomorphe, elle admet un développement en séries entières

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \tag{25.96}$$

et nous notons $M = \max\{f(z) \text{ tel que } z \in \overline{B(0,r)}\}$. Nous avons $r^n|a_n| \leqslant M$. Par conséquent

$$|f'(z)| = \left|\sum_{n=1}^{\infty} n a_n z^{n-1}\right| \tag{25.97a}$$

$$\leqslant \frac{1}{r}\sum r^n |a_n| n \left(\frac{|z|}{r}\right)^{n-1} \tag{25.97b}$$

$$< \frac{M}{r}\sum n \left(\frac{|z|}{r}\right)^{n-1}. \tag{25.97c}$$

À ce point nous devons utiliser la série de l'exemple 15.85. Nous avons alors

$$|f'(z)| \leqslant \frac{M}{r}\frac{1}{\left(1-\frac{|z|}{r}\right)^2} = \frac{Mr}{(r-|z|)^2}. \tag{25.98}$$

□

Théorème 25.25 (Holomorphie sous l'intégrale[286])**.**
Soit un espace mesuré (Ω, μ), un ouvert A dans $\mathbb{C}$ et une fonction $f\colon A \times \Omega \to \mathbb{C}$. Nous voulons étudier la fonction

$$F(z) = \int_\Omega f(z,\omega) d\mu(\omega) \tag{25.99}$$

pour tout $z \in A$. Nous supposons que

1. *la fonction $f(.,\omega)$ est holomorphe sur A pour chaque ω.*
2. *La fonction $f(z,.)$ est mesurable sur (Ω,μ).*
3. *Pour tout compact $K \subset A$, il existe une fonction $g_K\colon \Omega \to \mathbb{R}$ telle que $|f(z,\omega)| \leqslant g_K(\omega)$ et telle que*

$$\int_\Omega g_K(\omega) d\mu(\omega) \tag{25.100}$$

existe.

Alors la fonction F est holomorphe et

$$F'(z) = \int_\Omega \frac{\partial f}{\partial z}(z,\omega) d\mu(\omega). \tag{25.101}$$

Démonstration. Soient $z_0 \in A$ et $r > 0$ tels que $K = \overline{B(z_0,r)} \subset A$. Pour chaque $\omega \in \Omega$ nous considérons la fonction

$$\begin{aligned} f_\omega\colon \overline{B(z_0,r)} &\to \mathbb{C} \\ z &\mapsto f(z,\omega). \end{aligned} \tag{25.102}$$

Étant donné que $\overline{B(z_0,r)}$ est compacte, la fonction $|f_\omega|$ est majorée par un nombre que nous notons $f_K(\omega)$ qui est indépendant de z (pour autant que $z \in K$). Nous désignons par $S(z_0,r)$ la frontière de la boule $B(z_0,r)$. Étant donné que la majoration est valable sur $\overline{B(z_0,r)}$, nous avons en particulier

$$|f_\omega(z)| \leqslant f_K(\omega) \tag{25.103}$$

pour tout $z \in S$. En utilisant la lemme 25.24 nous avons

$$|f'_\omega(z)| \leqslant \frac{r}{(r-|z-z_0|)^2} \max\{f(z_0+re^{i\theta})\}_{\theta\in\mathbb{R}} \tag{25.104a}$$

$$\leqslant \frac{rf_K(\omega)}{(r-|z-z_0|)^2}. \tag{25.104b}$$

Cette majoration est valable pour tout $z \in B(z_0, r)$. Si nous supposons de plus que $z \in B(z_0, r/2)$ nous avons

$$|f'(z)| \leqslant \frac{rf_K(\omega)}{\left(r-\frac{r}{2}\right)^2} = \frac{4}{r}f_K(\omega). \tag{25.105}$$

Étant donné que la boule $B(z_0, r/2)$ est convexe, la fonction f_ω est Lipschitz et pour tout $h \in \mathbb{C}$ tel que $|h| < r/2$ nous avons

$$\left|\frac{f_\omega(z_0+h)-f_\omega(z_0)}{h}\right| \leqslant \frac{4f_K(\omega)}{r}. \tag{25.106}$$

Soit maintenant une suite (h_n) qui converge vers 0 dans $\mathbb{C}$. Nous considérons la suite de fonctions correspondantes

$$g_n(\omega) = \frac{f(z_0+h_n,\omega)-f(z_0,\omega)}{h_n}. \tag{25.107}$$

Cette suite de fonction vérifie la convergence ponctuelle

$$g_n(\omega) \to \frac{\partial f}{\partial z}(z_0,\omega). \tag{25.108}$$

De plus g_n est une fonction (de ω) dominée par $\frac{4f_K}{r}$ qui est intégrable. Par conséquent le théorème de la convergence dominée nous indique que

$$\int_\Omega g_n(\omega)d\mu(\omega) \to \int_\Omega \frac{\partial f}{\partial z}(z_0,\omega)d\mu(\omega), \tag{25.109}$$

tandis que

$$F'(z) = \lim_{n\to\infty}\frac{F(z_0+h_n)-F(z_0)}{h_n} = \lim_{n\to\infty}\int_\Omega g_N(\omega)d\mu(\omega). \tag{25.110}$$

□

Corollaire 25.26.
Si f est une fonction holomorphe sur l'ouvert ξ contenant la fermeture de la boule $B(z_0, r)$, alors pour tout z dans $B(z_0, \rho)$ ($\rho < r$) les dérivées de f s'expriment pas la formule suivante :

$$f^{(k)}(z) = \frac{1}{2\pi i}\int_{\partial B(z_0,r)}\frac{f(\xi)}{(\xi-z)^{k+1}}d\xi. \tag{25.111}$$

Démonstration. Nous faisons par récurrence.

Pour la dérivée première Nous appliquons le théorème 25.25 à la fonction

$$g(z,\xi) = \frac{f(\xi)}{\xi-z} \tag{25.112}$$

avec $\xi = \partial B(z_0, r)$ et $A = B(z_0, \rho)$ avec $\rho < r$. Étant donné que f est holomorphe, elle est continue et donc bornée sur tout compact $K \subset A$ par une constante M (qui dépend du compact choisi). D'autre part, nous avons toujours $|\xi - z| > r - \rho$ et donc

$$|g(z,\xi)| \leqslant \frac{M}{r-\rho}. \tag{25.113}$$

La fonction constante $g_K = \frac{M}{r-\rho}$ est évidemment intégrable. Le théorème conclut que f est holomorphe (cela, nous le savions déjà [7]), et

$$f'(z) = \frac{1}{2i\pi} \int_{\partial B} \frac{f(\xi)}{(\xi - z)^2} d\xi. \tag{25.114}$$

Les dérivées suivantes Pour la récurrence[288] nous supposons que

$$f^{(k)}(z) = \frac{k!}{2i\pi} \int_{\partial B(z_0,r)} \frac{f(\xi)}{(\xi - z)^{k+1}} d\xi, \tag{25.115}$$

et nous tentons de calculer $f^{(k+1)}(z)$. Pour cela nous paramétrons l'intégrale de façon très usuelle :

$$f^{(k)}(z) = \frac{k!}{2\pi i} \int_0^{2\pi} \frac{f(re^{it})}{(re^{it} - z)} ire^{it} dt. \tag{25.116}$$

Nous permettons de permuter la $\mathbb{C}$-dérivation (par rapport à z) et l'intégrale en vertu de la proposition 25.17 appliquée à la fonction

$$g(t, z) = \frac{f(re^{it})}{(re^{it} - z)^k + 1} ire^{it}. \tag{25.117}$$

Cela donne

$$f^{(k+1)}(z) = \frac{k!}{2i\pi} \int_0^{2\pi} f(re^{it}) ire^{it} \frac{k+1}{(re^{it} - z)^{k+1}} dt = \frac{(k+1)!}{2i\pi} \int_{\partial B} \frac{f(\xi)}{(\xi - z)^{k+2}} d\xi. \tag{25.118}$$

□

Théorème 25.27.
Si f est une fonction holomorphe sur le disque ouvert $B(z_0, R)$ alors

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n \tag{25.119}$$

et cette série converge uniformément sur tout compact.

Démonstration. Sans perte de généralité nous supposons que $z_0 = 0$. La formule de Cauchy (théorème 25.13) fournit, pour $z \in B(0, R)$,

$$f(z) = \frac{1}{2\pi i} \int_{\partial B} \frac{f(\xi)}{\xi - z} d\xi = \frac{1}{2\pi i} \int_{\partial B} \frac{f(\xi)}{1 - (z/\xi)} \frac{d\xi}{\xi}. \tag{25.120}$$

En particulier notons que $z \in B(0, R)$ alors que ξ est sur le bord de cette boule ouverte. Donc $|\xi| > |z|$ pour tous les ξ et z qui interviennent. Nous utilisons la série géométrique

$$\frac{1}{1 - (z/\xi)} = \sum_{n=0}^{\infty} \left(\frac{z}{\xi} \right)^n. \tag{25.121}$$

Permuter une intégrale et une somme En utilisant la mesure de comptage [8] sur $\mathbb{N}$ (qui est σ-finie), nous pouvons écrire

$$\int_{\partial B} \sum_{n=0}^{\infty} \frac{z^n f(\xi)}{\xi^{n+1}} d\xi = \int_{\partial B} \left(\int_{\mathbb{N}} g(\xi, n) dm(n) \right) d\xi \tag{25.122}$$

7. Et cela fournit une preuve alternative à la réciproque du théorème de Cauchy : une fonction continue qui vérifie la formule de Cauchy est holomorphe.

8. Définition 13.28.

où

$$\begin{aligned} g\colon \partial B \times \mathbb{N} &\to \mathbb{C} \\ (\xi, n) &\mapsto \frac{z^n f(\xi)}{\xi^{n+1}}. \end{aligned} \tag{25.123}$$

Nous allons permuter les intégrales en utilisant le théorème de Fubini, selon la procédure décrite en 13.231. Nous commençons par l'intégrale sur $\mathbb{N}$:

$$\int_{\mathbb{N}} |g(n,\xi)| = |\frac{f(\xi)}{\xi}| \sum_{n\in\mathbb{N}} |\frac{z}{\xi}|^n = \frac{1}{R}|f(\xi)|\frac{1}{1-|z|/R}. \tag{25.124}$$

Ici nous avons utilisé $|\xi| = R$. Notons que z est fixé depuis longtemps à l'intérieur de la boule de rayon R de telle sorte que $|z/\xi|$ est une constante strictement inférieure à 1.

L'intégrale sur $\xi \in \partial B$ n'a pas à être effectuée explicitement : nous nous contentons de prouver qu'elle est finie. La fonction f est continue sur le compact ∂B. Cela parce que B est une boule fermée dans l'ouvert Ω sur lequel f est continue. Au final l'expression à droite de (25.124) est bornée sur le compact ∂B et son intégrale donne un nombre fini.

Tout ceci pour invoquer le corollaire 13.229 qui nous indique que $g \in L^1(\mathbb{N} \times \partial B)$.

Une fois g intégrable sur l'espace produit $\mathbb{N} \times \partial B$, nous pouvons utiliser Fubini 13.230 pour permuter les intégrales.

Une fois la somme et l'intégrale permutées, nous avons

$$f(z) = \frac{1}{2\pi i}\sum_{n=0}^{\infty}\int_{\partial B}\frac{z^n f(\xi)}{\xi^{n+1}} = \sum_{n=0}^{\infty}\left(\frac{1}{2\pi i}\int_{\partial B}\frac{f(\xi)}{\xi^{n+1}}\right)z^n. \tag{25.125}$$

Nous devons maintenant montrer que ce qui se trouve dans la grande parenthèse vaut $f^{(n)}(0)/n!$. Cela est immédiat en comparant avec la formule (25.111).

□

Proposition 25.28 (Morera [289])**.**
Soit Ω ouvert dans $\mathbb{C}$ et f continue. Si

$$\int_{\partial T} f = 0 \tag{25.126}$$

pour tout triangle (plein) T contenu dans Ω, alors f est holomorphe sur Ω.

Démonstration. Il est suffisant de prouver que f est holomorphe sur toute boule ouverte $B(a,r)$ inclue dans Ω. Nous posons, pour tout $z \in B(a,r)$,

$$F(z) = \int_{[p,z]} f, \tag{25.127}$$

et nous considérons le chemin triangulaire $a \to z \to z+h \to a$ où $h \in \mathbb{C}$ est choisit assez petit pour que $z+h \in B(a,r)$. L'intégrale sur le triangle étant nulle, nous avons

$$0 = \int_{a\to z} f + \int_{z\to z+h} f + \int_{z+h\to a} f, \tag{25.128}$$

c'est à dire

$$F(z+h) - F(z) = \int_{z\to z+h} f. \tag{25.129}$$

En paramétrant le chemin par $z+th$ avec $t \in [0,1]$, et en tenant compte de la remarque 14.21,

$$F'(z) = \lim_{h\to 0}\frac{F(z+h)-F(z)}{h} \tag{25.130a}$$

$$= \lim_{h\to 0}\frac{1}{h}\int_0^1 f(z+th)h dt, \tag{25.130b}$$

ce qui prouve que F est dérivable et $F' = f$. Par définition (15.2), F est holomorphe, et donc C^∞ par le théorème 25.15. Du coup f est également C^∞ et donc en particulier holomorphe. □

25.2.1 Mesure de Radon

Définition 25.29.
*Une **mesure de Radon** sur un compact K de $\mathbb{C}$ est une forme linéaire continue sur $C(K)$. Si μ est une mesure de Radon, on définit la **transformée de Cauchy** de μ par*

$$\begin{aligned}\hat{\mu}\colon \mathbb{C}\backslash K &\to \mathbb{C}\\ z &\mapsto -\frac{1}{\pi}\mu\left(\frac{1}{\xi - z}\right).\end{aligned} \tag{25.131}$$

Théorème 25.30.
Si μ est une mesure de Radon sur K alors $\hat{\mu}$ est infiniment $\mathbb{C}$-dérivable sur $\Omega = \mathbb{C}\backslash K$ et nous avons

$$\hat{\mu}^{(n)}(z) = -\frac{n!}{\pi}\mu\left(\frac{1}{(\xi - z)^{n+1}}\right). \tag{25.132}$$

Cette théorie permet de fournir une démonstration plus technologique du corollaire 25.26.

Lemme 25.31.
Si f est holomorphe sur Ω et si B est une boule fermée dans Ω alors pour tout $z \in \mathrm{Int}(B)$ nous avons

$$f^{(k)}(z) = \frac{k!}{2i\pi}\int_{\partial B}\frac{f(\xi)}{(\xi - z)^{k+1}}d\xi. \tag{25.133}$$

Démonstration. Appliquer le théorème 25.30 à la mesure de Radon

$$\mu(\phi) = \int_{\partial B}\phi(\xi)d\xi. \tag{25.134}$$

□

Tout ce petit monde à propos de la mesure de Radon permet également de redémontrer que

$$\left(\frac{1}{2\pi i}\int_{\partial B}\frac{f(\xi)}{\xi^{n+1}}\right) = f^{(n)}(0)/n!, \tag{25.135}$$

comme nous l'avons déjà fait autour de l'équation (25.125). Nous utilisons le théorème de Radon 25.30 à la mesure

$$\mu(\phi) = \int_{\partial B}\phi(\xi)d\xi. \tag{25.136}$$

La transformée de Cauchy est

$$\hat{\mu}(z) = -\frac{1}{\pi}\mu\left(\frac{1}{\xi - z}\right) = -\frac{1}{\pi}\int_{\partial B}\frac{1}{\xi - z}d\xi, \tag{25.137}$$

et le théorème assure que

$$\hat{\mu}^{(n)}(z) = -\frac{n!}{\pi}\mu\left(\frac{1}{(\xi - z)^{n+1}}\right) = -\frac{n!}{\pi}\int_{\partial B}\frac{1}{(\xi - z)^{n+1}}d\xi. \tag{25.138}$$

En comparant la formule (25.137) avec la formule de Cauchy nous voyons que $\hat{\mu}(z) = -2if(z)$. Par conséquent

$$f^{(n)}(z) = -\frac{1}{2i}\hat{\mu}^{(n)}(z) = \frac{n!}{2\pi i}\int_{\partial B}\frac{1}{(\xi - z)^{n+1}}d\xi, \tag{25.139}$$

et

$$f^{(n)}(0) = \frac{n!}{2\pi i}\int_{\partial B}\frac{1}{\xi^{n+1}}d\xi. \tag{25.140}$$

25.3 Conditions équivalentes à l'holomorphie

Nous nous proposons de lister les conditions que nous avons vues être équivalentes à l'holomorphie.

Théorème 25.32.
Soit Ω un ouvert de $\mathbb{C}$ et $f\colon \Omega \to \mathbb{C}$ une fonction continue. Les conditions suivantes sont équivalentes.

(1) f est holomorphe.

(2) Pour tout triangle (plein) T contenu dans Ω, $\int_T f = 0$.

(3) f est $\mathbb{C}$-dérivable.

(4) f est C^∞

(5) $\frac{\partial f}{\partial \bar{z}} = 0$; ce sont les équations de Cauchy-Riemann.

(6) La 1-forme différentielle $f(z)dz$ est localement exacte.

(7) Pour toute boule $B(a,r)$ contenue dans Ω nous avons

$$f(a) = \frac{1}{2\pi i}\int_{\partial B(a,r)} \frac{f(z)}{z-a}dz. \tag{25.141}$$

Démonstration. (1) implique (2) est le lemme de Goursat 25.5. (2) implique (1) est le théorème de Morera 25.28.

(3) est la définition de l'holomorphie.

(4) implique (1) est un a fortiori sur la définition. (1) implique (4) est contenu dans le théorème de développement en série entière 25.15.

L'équivalence entre (5) et l'holomorphie est le théorème 25.2.

L'équivalence entre (6) et (1) est la proposition 25.18.

L'équivalence entre (1) et (7) est d'une part le théorème 25.15 et d'autre part le corollaire 25.16. □

25.4 Singularités, pôles et méromorphe

Définition 25.33.
*Si f est holomorphe sur un ouvert Ω, alors une **singularité** de f est un point isolé du bord de Ω.*

*(1) La singularité est **effaçable** si la fonction f s'y prolonge en une fonction holomorphe.*

*(2) La singularité Z est un **pôle** d'ordre k de f si elle n'est pas effaçable et si la fonction $z \mapsto (z-Z)^k f(z)$ se prolonge en une fonction holomorphe en Z.*

Exemple 25.34
La fonction

$$z \mapsto \frac{\sin(z)}{z} \tag{25.142}$$

n'est pas définie en $z = 0$, mais elle s'y prolonge en une fonction continue en posant $f(0) = 1$. △

Proposition 25.35.
Une singularité de f est un pôle si et seulement si

$$\lim_{z\to Z} f(z) = \infty. \tag{25.143}$$

Le théorème suivant compète la proposition 25.22.

Théorème 25.36 (Prolongement de Riemann)**.**
Soit $f\colon \Omega \to \mathbb{C}$ et une singularité Z de f. Nous avons équivalence de

(1) la singularité Z est effaçable ;

(2) f possède un prolongement continue en Z ;

(3) il existe un voisinage épointé de Z sur lequel f est bornée ;

(4) $\lim_{z\to Z}(z-Z)f(z)=0$.

Définition 25.37 (Fonction méromorphe[290])**.**
*Soit un ouvert $\mathcal{U}$ de $\mathbb{C}$ et une suite de points $\{p_i\}$ dans $\mathcal{U}$ sans points d'accumulation (éventuellement il y a un nombre fini de p_i). Si la fonction f est holomorphe sur $\mathcal{U}\backslash\{p_i\}$ et si chaque p_i est un point régulier ou un pôle de f, alors nous disons que f est **méromorphe** sur $\mathcal{U}$.*

Proposition 25.38.
Soit Ω un ouvert de $\mathbb{C}$ et une suite de fonctions $f_n\colon\Omega\to\mathbb{C}$ telles que pour tout compact K de Ω il existe $N_K\geqslant 0$ tel que

(1) f_n n'a pas de pôles dans K dès que $n\geqslant N_K$;

(2) la série $\sum_{n\geqslant N_K}f_n$ converge uniformément sur K.

Alors

(1) La fonction

$$f(z)=\sum_{n=0}^{\infty}f_n(z) \tag{25.144}$$

est méromorphe sur Ω et ses pôles sont l'union de ceux des f_n.

(2) Nous pouvons permuter la somme et la dérivée :

$$f'(z)=\sum_{n=0}^{\infty}f_n'(z). \tag{25.145}$$

Théorème 25.39 (Série de Laurent)**.**
*Soient C une couronne de rayons $r_1<r_2$ centrée en zéro et une fonction f holomorphe dans cette couronne. Alors nous avons la **série de Laurent***

$$f(z)=\sum_{n\in\mathbb{Z}}a_nz^n. \tag{25.146}$$

(1) Cette série converge uniformément sur tout compact de C.

(2) Les coefficients sont donnés par

$$a_n=\frac{1}{2\pi i}\int_\gamma\frac{f(z)}{z^{n+1}}dz \tag{25.147}$$

où γ est un cercle centré en zéro.

(3) Ce développement en série est unique.

(4) La valeur des a_n ne dépend pas du choix du rayon du cercle γ.

25.5 Fonctions d'Euler

Théorème 25.40 (Prolongement méromorphe de la fonction Γ d'Euler[2])**.**
Nous considérons la formule

$$\Gamma(z)=\int_0^\infty e^{-t}t^{z-1}dt. \tag{25.148}$$

Alors

(1) Cette formule définit une fonction holomorphe sur

$$\mathcal{P}=\{z\in\mathbb{C}\text{ tel que }\Re(z)>0\}. \tag{25.149}$$

(2) La fonction $\Gamma\colon \mathcal{P} \to \mathbb{C}$ admet un unique prolongement méromorphe sur $\mathbb{C}$, lequel a des pôles sur les entiers négatifs.

Démonstration. **Holomorphie sous l'intégrale** Pour étudier l'holomorphie de la fonction Γ sur $\mathcal{P}$ nous utilisons le théorème 25.25.

Nous considérons la fonction

$$\begin{aligned} g\colon \mathcal{P} \times \mathbb{R}^+ &\to \mathbb{C} \\ (z,t) &\mapsto e^{-t}z^{z-1} \end{aligned} \tag{25.150}$$

et nous commençons par montrer que c'est holomorphe en z pour chaque $t > 0$ fixé. Nous le vérifions par le critère de $\partial_{\bar{z}f=0}$ [9] et en nous souvenant que $t^i = e^{\ln(t^i)} = e^{i\ln(t)}$. Nous obtenons rapidement que

$$\frac{\partial g}{\partial \bar{z}} = 0. \tag{25.151}$$

Le fait que la fonction $t \mapsto g(z,t)$ soit mesurable pour tout z est d'accord.

Et enfin soit K compact dans $\mathcal{P}$. Il faut trouver une fonction $g_K(t)$ intégrable sur $[0,\infty[$ telle que pour tout $z \in K$ et $t \in [0,\infty[$ nous ayons $|f(z,t) \leqslant g(t)|$. Pour cela nous majorons séparément les parties $t \in]0,1[$ et $t \geqslant 1$.

Soit donc K compact dans $\mathcal{P}$; nous posons $M = \max_{z\in K} \Re(z)$ et $\epsilon = \min_{z\in K} \Re(z)$.

Si $t \in]0,1[$ alors nous avons

$$e^{-t}t^{z-1} = e^{-t}e^{(z-1)\ln(t)}, \tag{25.152}$$

de telle façon à que que

$$\begin{aligned} |e^{-t}t^{z-1}| &\leqslant |e^{(x-1+iy)\ln(t)}| && (25.153a)\\ &= |e^{(\Re(z)-1)\ln(t)}| && (25.153b)\\ &= |t^{\Re(z)-1}| && (25.153c)\\ &\leqslant |t^{\epsilon-1}| && (25.153d)\\ &= \frac{1}{t^{1-\epsilon}}. && (25.153e) \end{aligned}$$

Cette dernière fonction est intégrable sur $]0,1[$.

Nous considérons maintenant $t \geqslant 1$. Dans ce cas nous avons

$$|e^{-t}z^{z-1}| = e^{-t}t^{\Re(z)-1} \leqslant e^{-t}t^{M-1}. \tag{25.154}$$

Cette dernière fonction est un produit d'une exponentielle décroissante avec un polynôme. C'est donc intégrable entre 1 et l'infini.

La fonction g_K que nous considérons est donc

$$g_K(t) = \begin{cases} \frac{1}{t^{1-\epsilon}} & \text{si } t < 1 \\ \text{borné} & \text{si } 1 \leqslant t \leqslant b \\ e^{-t}t^{M-1} & \text{si } t > b. \end{cases} \tag{25.155}$$

Cela est une fonction intégrable sur $]0\infty,[$ et qui majore f uniformément en z sur le compact K de $\mathcal{P}$. Le théorème 25.25 nous permet donc de conclure que

$$\Gamma(z) = \int_0^\infty f(z,t)dt \tag{25.156}$$

est holomorphe en z sur $\mathcal{P}$ et que

$$\Gamma'(z) = \int_0^\infty \frac{\partial f}{\partial z}(z,t)dt. \tag{25.157}$$

9. Théorème 25.2.

En deux morceaux Nous passons maintenant à la seconde partie du théorème. Pour $z \in \mathcal{P}$ nous coupons l'intégrale en deux :

$$\Gamma(z) = \int_0^1 e^{-t}t^{z-1}dt + \int_1^\infty e^{-t}t^{z-1}dt \tag{25.158}$$

Première partie Nous commençons par parler de la première partie : $\int_0^1 e^{-t}t^{z-1}dt$ dans laquelle nous voulons utiliser le développement en série de l'exponentielle e^{-t}. Nous devons donc traiter

$$\int_0^1 \sum_{n=0}^\infty \frac{(-1)^n}{n!}t^{n+z-1}dt. \tag{25.159}$$

Nous allons permuter la somme avec l'intégrale à l'aide du théorème de Fubini 13.230 en posant la fonction

$$g(n,t) = \frac{(-1)^n}{n!}t^{n+z-1} \tag{25.160}$$

et en considérant le produit entre la mesure de Lebesgue sur $\mathbb{C}$ et la mesure de comptage sur $\mathbb{N}$, c'est à dire que nous étudions

$$\int_0^1 \int_{\mathbb{N}} g(n,t)dndt. \tag{25.161}$$

Pour permuter il suffit de prouver que $|g|$ est intégrable pour la mesure produit, c'est à dire que

$$\int_0^1 \int_{\mathbb{N}} \left|\frac{(-1)^n}{n!}t^{n+z-1}\right| < \infty. \tag{25.162}$$

Nous avons $|t^z = t^{\Re(z)}|$, donc

$$\sum_{n=0}^\infty \left|\frac{t^{n+z-1}}{n!}\right| = t^{\Re(z)-1}\sum_{n=0}^\infty \frac{t^n}{n!} = t^{\Re(z)-1}e^t. \tag{25.163}$$

Étant donné que nous avons fixé $z \in \mathcal{P}$, nous avons $\Re(z) - 1 > -1$ et donc $t^{\Re(z)-1}$ est intégrable entre 0 et 1. La partie e^t se majore sur $[0,1]$ par une constante quelconque. Nous avons donc payé le droit d'inverser la somme et l'intégrale :

$$\int_0^1 e^{-t}t^{z-1}dt = \sum_{n=0}^\infty \int_0^1 \frac{(-1)^n}{n!}t^{n+z-1}dt = \sum_{n=0}^\infty \frac{(-1)^n}{n!}[t^{n+z}]_0^1 = \sum_{n=0}^\infty \frac{(-1)^n}{n!(n+z)}. \tag{25.164}$$

Nous avons donc l'intéressante formule suivante, valable pour tout $z \in \mathcal{P}$:

$$\Gamma(z) = \sum_{n=0}^\infty \frac{(-1)^n}{n!(n-z)} + \int_1^\infty e^{-t}t^{z-1}dt. \tag{25.165}$$

Prolongation de la première partie Nous voudrions montrer maintenant que la fonction

$$\sum_{n=0}^\infty \frac{(-1)^n}{n!(n-z)} \tag{25.166}$$

est méromorphe sur $\mathbb{C}$ avec des pôles en les entiers négatifs. Pour cela nous considérons la suite de fonctions

$$f_n(z) = \frac{(-1)^n}{n!(z+n)} \tag{25.167}$$

et nous allons utiliser la proposition 25.38. Si $n \geqslant 0$, la fonction f_n est méromorphe sur $\mathbb{C}$ avec un pôle simple en $z = -n$. Soit K compact de $\mathbb{C}$ et N_K tel que $K \subset \overline{B(0,N_K)}$. Pour $n \geqslant N_K + 1$, la fonction f_n n'a pas de pôles dans K et de plus pour tout $z \in K$ nous avons

$$|z+n| = |z-(-z)| \geqslant \big|n - |z|\big| \geqslant n - |z| \geqslant n - N_K, \tag{25.168}$$

et par conséquent

$$|f_n(z)| \leqslant \frac{1}{n!(n-N)}, \tag{25.169}$$

ou pour le dire de façon plus snob :

$$\|f_n\|_{\infty,K} \leqslant \frac{1}{n!(n-N)}, \tag{25.170}$$

dont la série converge. Cela signifie que la série $\sum_{n\geqslant N} f_n$ converge normalement [10] sur K, donc la fonction

$$f(z) = \sum_{n=0}^{\infty} f_n(z) \tag{25.171}$$

est une fonction méromorphe dont les pôles sont ceux des f_n, c'est à dire les entiers négatifs (proposition 25.38).

La seconde partie Nous allons à présent prouver que la fonction

$$g(z) = \int_1^{\infty} e^{-t}t^{z-1}dt \tag{25.172}$$

est holomorphe sur $\mathbb{C}$. Pour cela nous considérons la fonction de deux variables $f(z,t) = e^{-t}t^{z-1}$ et nous utilisons le théorème d'holomorphie sous l'intégrale 25.25. D'abord pour z_0 fixé dans $\mathbb{C}$ nous avons

$$\int_1^{\infty} |e^{-t}t^{z_0-1}| \leqslant \int_1^{\infty} e^{-t}t^{\Re(z_0)-1}dt, \tag{25.173}$$

donc l'intégrale converge parce que c'est polynôme contre exponentielle. Par ailleurs pour chaque t_0 fixé sur $[0,\infty[$, la fonction $z \mapsto e^{-t_0}t_0^{z-1}$ est holomorphe sur $\mathbb{C}$ comme en témoigne le calcul suivant :

$$\frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right) t_0^{x+iy-1} = 0. \tag{25.174}$$

Et enfin si K est compact dans $\mathbb{C}$ nous avons

$$|f(z,t)| = |e^{-t}t^{z-1}| = e^{-t}|t^{\Re(z)-1}| \leqslant e^{-t}t^{M-1} \tag{25.175}$$

où $M = \max_{z\in K} \Re(z)$. Nous en déduisons que la fonction

$$z \mapsto \int_1^{\infty} e^{-t}t^{z-1}dt \tag{25.176}$$

est une fonction holomorphe sur $\mathbb{C}$.

Conclusion Au final nous avons prouvé que la fonction Γ d'Euler admet le prolongement méromorphe sur $\mathbb{C}$ donné par

$$\Gamma(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!(z+n)} + \int_1^{\infty} e^{-t}t^{z-1}dt. \tag{25.177}$$

□

25.5.1 Euler et factorielle

Proposition 25.41.
Nous avons la formule $\Gamma(n) = (n-1)!$ *pour tout* $n \in \mathbb{N}$.

10. Définition 10.46.

Démonstration. Nous partons de la formule

$$\Gamma(n) = \int_0^\infty e^{-t}t^{n-1}dt \tag{25.178}$$

que nous intégrons par partie en posant

$$\begin{aligned} u &= t^{n-1} & u' &= (n-1)t^{n-1} \\ v &= e^{-t} & v' &= -e^{-t}. \end{aligned} \tag{25.179}$$

Les termes au bord s'annulent (ici il y a un passage à la limite qui n'est pas écrit) et nous trouvons

$$\Gamma(n) = \int_0^\infty (n-1)e^{-t}t^{n-2}dt = (n-1)\Gamma(n-1). \tag{25.180}$$

Pour conclure il suffit de remarquer que

$$\Gamma(1) = \int_0^\infty = -[e^{-t}]_0^\infty = 1. \tag{25.181}$$

□

25.6 Partition d'un entier en parts fixées

Proposition 25.42 ([2])**.**
Soient $a_1, \dots, a_k \in \mathbb{N}^*$ des entiers premiers entre eux dans leur ensemble. Pour $n \geqslant 1$ nous posons

$$u_n = \text{Card}\left\{(x_1, \dots, x_k) \in \mathbb{N}^* \text{ tel que } \sum_{i=1}^k a_i x_i = n\right\}, \tag{25.182}$$

et $u_0 = 1$.

Alors nous avons l'équivalence de suite (pour $n \to \infty$) :

$$u_n \sim \frac{1}{a_1 \dots a_k} \frac{n^{k-1}}{(k-1)!}. \tag{25.183}$$

Démonstration. Pour chacun des $i \in \{1, \dots, k\}$ nous considérons la série entière

$$\sum_{x=0}^\infty z^{xa_i} = \sum_k (z^{a_i})^x. \tag{25.184}$$

Étant donné que $|z^{a_i}| < 1$ si et seulement si $|z < 1|$, cette série a un rayon de convergence égal à 1. Nous allons calculer le produit de Cauchy de ces k séries, en nous souvenant que le théorème 15.31 nous assure que la série résultante aura un rayon de convergence au moins égal à 1 et vaudra le produit des différentes séries.

Le coefficient de z^n dans cette série vaut

$$\sum_{\sum^{x\in\mathbb{N}^k} x_i a_i = n} 1 = u_n \tag{25.185}$$

parce que dans chacune des séries, le coefficient de tous les z^{xa_i} est 1. Nous définissons la fonction

$$f(z) = \sum_{n=0}^\infty u_n z^n = \prod_{i=1}^k \left(\sum_{x=0}^\infty z^{xa_i} \right) = \prod_{i=1}^k \frac{1}{1-z^{a_i}}. \tag{25.186}$$

La fonction f existe sur $|z| < 1$ parce que nous venons de voir qu'elle peut s'exprimer comme un produit de Cauchy ; et la dernière égalité est simplement la somme de la série harmonique. D'autre part la fonction f est la série génératrice de la suite (u_n).

Nous sommes en présence d'une fonction ayant des pôles aux racines $a_1,\dots,\ a_k$[e] de l'unité. Étant donné que 1 est une racine de l'unité de tous les ordres, le pôle en $z=1$ est de multiplicité k. Les autres pôles sont de multiplicité strictement inférieure ; en effet soit $\omega\in\mathbb{C}$ tel que $\omega^{a_i}=1$ pour tout i. Alors Bezout [11] nous donne des entiers $v_i\in\mathbb{Z}$ tels que $\sum_i v_i a_i=1$. Alors nous avons

$$\omega=\omega^{\sum_{v_i a_i}}=\prod_{i=1}^{k}(\omega^{a_i})^{v_i}=1. \tag{25.187}$$

Donc nous voyons que 1 est le seul à être racine de tous les ordres en même temps. Nous notons

$$P=\{\omega_1,\dots,\omega_p\} \tag{25.188}$$

l'ensemble des pôles avec $\omega_1=1$. Par ailleurs la fonction f est une fraction rationnelle dont nous connaissons les racines du dénominateur (ce sont les ω_i) et à peu près leurs ordres. Nous utilisons le truc de la décomposition en fractions simples en séparant le terme de puissance k qui n'existe que pour la racine $\omega_1=1$:

$$f(z)=\frac{\alpha}{(1-z)^k}+\sum_{i=1}^{p}\sum_{j=1}^{k-1}\frac{c_{ij}}{(\omega_i-z)^j}. \tag{25.189}$$

Ce développement est valable pour tout $|z|<1$. Nous considérons maintenant $\omega\in P$ et $j\in\mathbb{N}$ et nous étudions la fonction

$$g(z)=\frac{1}{\omega-z}. \tag{25.190}$$

Un rapide calcul (par exemple par récurrence) montre que

$$g^{(k)}(z)=\frac{k!}{(\omega-z)^{k+1}}, \tag{25.191}$$

et étant donné que $|\omega|=1$ nous pouvons écrire la série

$$\frac{1}{\omega-z}=\sum_{k=0}^{\infty}\frac{z^k}{\omega^{k+1}}, \tag{25.192}$$

valable pour $|z|<1$. Ce qui nous intéresse, c'est d'exprimer une série pour $1/(\omega-z)^j$; et voyant (25.191), nous voyons qu'il suffit de calculer les dérivées de la série de g. Nous dériver terme à terme à l'intérieur du rayon de convergence. Avec quelques abus d'écriture, et en utilisant la bête formule (15.238) nous avons [12]

$$\frac{1}{(\omega-z)^j}=\frac{g^{(j-1)}(z)}{(j-1)!} \tag{25.193a}$$

$$=\frac{1}{(j-1)!}\left(\frac{1}{\omega-z}\right)^{(j-1)} \tag{25.193b}$$

$$=\frac{1}{(j-1)!}\sum_{k=0}^{\infty}\frac{1}{\omega^{k+1}}(z^k)^{(j-1)} \tag{25.193c}$$

$$=\sum_{k=j-^{\infty}}\frac{1}{(j-1)!}\frac{1}{\omega^{k+1}}\frac{k!}{(k-j+1)!}z^{k-j+1} \tag{25.193d}$$

$$=\sum_{n=0}^{\infty}\frac{1}{\omega^{n+j}}\frac{(n+j-1)!}{n!(j-1)!}z^n \tag{25.193e}$$

$$=\sum_{n=0}^{\infty}\frac{1}{\omega^{n+j}}\binom{n+j-1}{n}z^n. \tag{25.193f}$$

11. Théorème 3.32.

12. À ce niveau j'ai pas exactement le même coefficient binomial que dans [2], mais je n'exclus absolument pas que ce soit moi qui me trompe. Écrivez-moi si vous pouvez infirmer ou confirmer l'erreur. Quoi qu'il en soit, cela ne change pas le résultat asymptotique que nous cherchons.

Nous pouvons utiliser cela pour récrire la formule (25.189) de façon considérablement plus compliquée :

$$f(z)=\alpha\sum_{n=0}^{\infty}\binom{n+j-1}{n}z^n+\sum_{i=1}^{p}\sum_{j=1}^{k-1}\sum_{n=0}^{\infty}c_{ij}\binom{n+j-1}{n}\frac{z^n}{\omega_i^{n+j}}. \tag{25.194}$$

Mais nous savons que ce f est la série génératrice de la suite (u_n) et que nous pouvons donc utiliser la formule (15.240) pour exprimer les nombres u_l : u_l est simplement le coefficient de z^l divisé par $l!$. C'est à dire

$$u_l=\alpha\binom{l+k-1}{l}+\sum_{i=1}^{p}\sum_{j=1}^{k-1}c_{ij}\binom{l+j-1}{l}\frac{1}{\omega_i^{l+j}}. \tag{25.195}$$

Notre boulot est d'examiner le comportement de cela lorsque $l\to\infty$, c'est à dire regarder quels sont les puissances de l en présence. Notons que

En ce qui concerne le premier terme, la puissance dominante dans le coefficient binomial est l^{k-1}. Dans les autres termes [13], c'est l^{j-1} qui est de degré moins grand. Donc le comportement de u_l en terme de l est

$$u_l\sim\alpha\frac{l^{k-1}}{(k-1)!}. \tag{25.196}$$

Il nous reste à voir ce que vaut α. Pour cela nous repartons de l'expression 25.186 que nous écrivons sous la forme

$$(1-z)^kf(z)=\prod_{i=1}^{k}\frac{1-z}{1-z^{a_i}}. \tag{25.197}$$

Nous reconnaissons l'inverse d'une somme harmonique partielle :

$$(1-z)^kf(z)=\prod_{i=1}^{k}\frac{1}{1+z+z^2+\cdots+z^{a_i-1}}. \tag{25.198}$$

Par ailleurs, nous ne savons pas si $f(1)$ existe parce que son rayon de convergence n'est que de 1 ; et nous savons même qu'elle n'existe pas (parce que ce serait la somme des u_n). Mais nous savons aussi que le pôle de plus grande multiplicité de f est en $z=1$ et est de multiplicité k. Donc $(1-z)^kf(z)$ devrait converger pour $z\to1$. Pour tout $|z<1|$ nous avons

$$(1-z)^kf(z)=\alpha+\sum_{i=1}^{p}\sum_{j=1}^{k-1}c_{ij}\frac{(1-z)^k}{(\omega_i-z)^j}. \tag{25.199}$$

Lorsque $z\to1$, tous les termes des sommes tendent vers zéro, y compris ceux avec $i=1$ parce que $j<k$. Il reste donc

$$\lim_{z\to0}(1-z)^kf(z)=\alpha. \tag{25.200}$$

En calculant la même limite avec (25.198) nous trouvons

$$\lim_{z\to1}(1-z)^kf(z)=\lim_{z\to1}\prod_{i=1}^{k}\frac{1}{1+z+z^2+\cdots+z^{a_i-1}}=\frac{1}{a_1\dots a_k}. \tag{25.201}$$

Donc

$$\alpha=\frac{1}{a_1\dots a_k}, \tag{25.202}$$

et le résultat est prouvé.

□

13. Attention : les termes $i=1$ ont $\omega_1=1$ et il n'est donc pas possible de conclure simplement en disant que $\omega_i^{l-j}\to0$ pour $l\to\infty$; bien que cela soit vrai pour tous les $i\neq1$.

25.7 Exponentielle et logarithme complexe

25.7.1 Définition et propriétés de l'exponentielle

Définition 25.43.
*Soit $z = x + iy \in \mathbb{C}$. Nous définissons l'**exponentielle** de z par*

$$\begin{aligned} \exp\colon \mathbb{C} &\to \mathbb{C} \\ z &\mapsto \sum_{n=0}^{\infty} \frac{z^n}{n!}. \end{aligned} \tag{25.203}$$

Le rayon de convergence de cette somme est infini.

Proposition 25.44 ([291])**.**
Quelques propriétés de l'exponentielle.

(1) Le fonction exp *est continue.*
(2) Nous avons la formule $e^{z+w} = e^z + e^w$ pour tout $z, w \in \mathbb{C}$.
(3) $(e^z)^{-1} = e^{-z}$
(4) $(\exp(z))^n = \exp(nz)$.

Démonstration. L'exponentielle est continue parce qu'elle est la somme d'une série entière de rayon de convergence infini (proposition 15.12).

Les séries $\exp(z)$ et $\exp(w)$ ayant un rayon de convergence infini nous pouvons utiliser le produit de Cauchy (théorème 15.31) :

$$e^z e^w = \sum_{n=0}^{\infty} \left(\sum_{i+j=n} \frac{z^i w^j}{i!j!} \right) \tag{25.204a}$$

$$= \sum_{n=0}^{\infty} \left(\sum_{i=0}^{n} \frac{z^i w^{n-i}}{i!(n-i)!} \right) \tag{25.204b}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{i=0}^{n} \binom{n}{i} z^i w^{n-i} \tag{25.204c}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} (z+w)^n \tag{25.204d}$$

$$= \exp(z+w). \tag{25.204e}$$

Nous avons utilisé la formule du binôme (proposition 4.9).

Les autres propriétés énoncées sont des corollaires :

$$e^z e^{-z} = e^0 = 1. \tag{25.205}$$

□

Proposition 25.45.
Si $z = x + iy \in \mathbb{C}$ alors

$$e^{x+iy} = e^x \big(\cos(y) + i \sin(y) \big). \tag{25.206}$$

Démonstration. Par la proposition 25.44 nous savons que $e^{x+iy} = e^x e^{iy}$. Nous devons donc seulement étudier e^{iy}. Nous avons

$$e^{iy} = \sum_{n=0}^{\infty} \frac{(iy)^n}{n!} \tag{25.207a}$$

$$= \sum_{n=0}^{\infty} (-1)^n \frac{y^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} (-1)^n \frac{y^{2n+1}}{(2n+1)!} \tag{25.207b}$$

$$= \cos(y) + i \sin(y). \tag{25.207c}$$

Nous avons utilisé le fait que $i^{2n} = (-1)^n$ et $i^{2n+1} = i(-1)^n$. □

Proposition 25.46.
Soit $z \in \mathbb{C}$ fixé. La fonction

$$\begin{aligned} E\colon \mathbb{R} &\to \mathbb{C} \\ t &\mapsto e^{tz} \end{aligned} \tag{25.208}$$

est C^∞, sa dérivée est

$$E'(t) = ze^{tz}. \tag{25.209}$$

La fonction E est développable en série entière (voir définition 15.75) sur $\mathbb{R}$ en $t = 0$ et

$$e^{tz} = \sum_{n=0}^{\infty} \frac{z^n}{n!} t^n. \tag{25.210}$$

Démonstration. Nous fixons $z \in \mathbb{C}$. Par définition 25.43, la série suivante est e^{tz} :

$$f(t) = \sum_{n=0}^{\infty} \frac{z^n}{n!} t^n. \tag{25.211}$$

Cette série a un rayon de convergence infini et la fonction f est donc C^∞ sur $\mathbb{R}$. Nous pouvons la dériver terme à terme :

$$f'(t) = \sum_{n=1}^{\infty} \frac{z^n}{n!} n t^{n-1} = z \sum_{n=1}^{\infty} \frac{z^{n-1}}{(n-1)!} t^{n-1} = ze^{tz}. \tag{25.212}$$

□

Théorème 25.47.
La fonction exponentielle vérifie les propriétés suivantes.

(1) exp *est holomorphe.*

(2) $(e^z)' = e^z$.

(3) L'exponentielle est développable en série entière,

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} \tag{25.213}$$

et la série converge normalement sur tout compact de $\mathbb{C}$.

Démonstration. En tant que application $E\colon \mathbb{R}^2 \to \mathbb{C}$, la fonction

$$E(x, y) = e^x(\cos y + i \sin y) \tag{25.214}$$

est C^∞. De plus nous avons

$$\frac{\partial E}{\partial x}(x, y) = e^{x+iy} = E(x, y) \tag{25.215a}$$

$$\frac{\partial E}{\partial y}(x, y) = iE(x, y), \tag{25.215b}$$

et par conséquent la fonction E vérifie les équations de Cauchy-Riemann.

Si r est fixé, par le critère d'Abel appliqué à la suite $r/n!$ nous savons que la série $\sum z^n/n!$ converge normalement sur le compact $B(0, r)$. □

25.7.2 Intégrale de Fresnel

Nous allons calculer l'**intégrale de Fresnel**

$$\int_0^\infty e^{-ix^2}dx = \frac{\sqrt{\pi}}{2}e^{-i\pi/4} \tag{25.216}$$

en suivant la démarche présentée par wikipédia. Nous commençons par prouver que l'intégrale est convergente en nous contentant de justifier la convergence de

$$\int_0^\infty \sin(x^2)dx. \tag{25.217}$$

Pour tout $a > 0$, l'intégrale $\int_0^a \sin(x^2)dx$ ne pose pas de problèmes. En tenant compte du lemme 14.188, nous devons donc seulement calculer

$$\lim_{b\to\infty}\int_a^b \sin(x^2)dx \tag{25.218}$$

où a est une constante strictement positive. Nous effectuons une intégration par partie en posant

$$u = \frac{1}{x} \qquad u' = -\frac{1}{x^2} \tag{25.219a}$$

$$v' = x\sin(x) \qquad v = \frac{1-\cos(x)}{2}. \tag{25.219b}$$

Notons que la primitive v a été choisie pour avoir $v(0) = 0$. Nous avons

$$\int_a^b \sin(x^2)dx = \left[\frac{1-\cos(x^2)}{2x}\right]_a^b - \int_a^b \frac{\cos(x^2)-1}{2x^2}dx \tag{25.220}$$

Pour le premier terme nous avons

$$\lim_{b\to\infty}\left[\frac{1-\cos(x^2)}{2x}\right]_a^b = \lim_{b\to\infty}\frac{1-\cos(b^2)}{2b} - \frac{1-\cos(a^2)}{2a} = -\frac{1-\cos(a^2)}{2a}. \tag{25.221}$$

C'est borné. Pour le second terme de (25.220), la fonction

$$\frac{\cos(x^2)-1}{2x^2} \tag{25.222}$$

est majorée par la fonction $1/x^2$ qui est intégrable entre a et ∞.

Nous allons calculer l'intégrale demandée en passant par la fonction

$$f(x) = e^{-z^2} \tag{25.223}$$

définie sur le plan complexe. Nous l'intégrons sur le chemin $\gamma = \gamma_1 + \gamma_2 - \gamma_3$ indiqué à la figure 25.1.

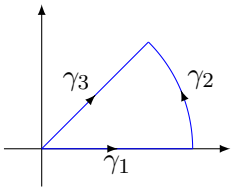


FIGURE 25.1 – Chemin d'intégration pour l'intégrale de Fresnel

Ces chemins sont donnés par

$$\begin{aligned}\gamma_1 : [0,R] &\to \mathbb{C}\\ t &\mapsto t,\end{aligned} \tag{25.224}$$

$$\begin{aligned}\gamma_2\colon [0,\frac{\pi}{4}] &\to \mathbb{C}\\ t &\mapsto Re^{it},\end{aligned} \tag{25.225}$$

$$\begin{aligned}\gamma_3\colon [0,R] &\to \mathbb{C}\\ t &\mapsto te^{i\pi/4}.\end{aligned} \tag{25.226}$$

Tout d'abord la fonction f est bien holomorphe par le critère du théorème 25.2. Le calcul de $\frac{\partial f}{\partial \bar{z}}$ se fait simplement en posant $f(x,y)=e^{-(x+iy)^2}$. Le calcul est usuel :

```
----------------------------------------------------------------------
| Sage Version 4.8, Release Date: 2012-01-20                         |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y)=exp(-(x+I*y)**2)
sage: A=f.diff(x)+I*f.diff(y)
sage: A.simplify_full()
(x, y) |--> 0
```

Nous avons donc

$$0=\int_\gamma f=\underbrace{\int_0^R e^{-t^2}dt}_{I_1(R)}+\underbrace{\int_0^{\pi/4} e^{-R^2e^{2it}}Rie^{it}dt}_{I_2(R)}+\underbrace{\int_0^R e^{-t^2e^{i\pi/2}}e^{i\pi/4}dt}_{I_3(R)}. \tag{25.227}$$

L'intégrale est nulle pour tout R en vertu de la proposition 25.6. L'intégrale I_1 est une gaussienne et nous avons

$$\lim_{R\to\infty} I_1(R)=\frac{\sqrt{\pi}}{2} \tag{25.228}$$

par l'exemple 13.233. Nous montrons maintenant que $\lim_{R\to\infty}|I_2(R)|=0$ [14]. D'abord nous majorons en prenant la norme puis nous effectuons le changement de variables $u=2t$:

$$|I_2(R)|\leqslant\int_0^{\pi/4} Re^{-R^2\cos(2t)}dt \tag{25.229a}$$

$$=\frac{R}{2}\int_0^{\pi/2} e^{-R^2\cos(u)}du. \tag{25.229b}$$

Nous savons que le graphe du cosinus est concave : il reste au dessus de la droite que joint $(0,1)$ à $(\frac{\pi}{2},0)$. Du coup $\cos(u)\geqslant 1-\frac{2}{\pi}u$ et par conséquent

$$e^{-R^2\cos(u)}\leqslant e^{-R^2(1-\frac{2}{\pi}u)}=e^{R^2(\frac{2}{\pi}u-1)}. \tag{25.230}$$

Nous effectuons l'intégrale

$$|I_2(R)|\leqslant\frac{R}{2}\int_0^{\pi/2} e^{-R^2}e^{\frac{2R^2}{\pi}u}du \tag{25.231a}$$

$$=\frac{R}{2}e^{-R^2}\Big[\frac{\pi}{2R^2}e^{2R^2u/\pi}\Big]_0^{\pi/2} \tag{25.231b}$$

$$=\frac{\pi}{4R}-\frac{\pi e^{-R^2}}{4R}, \tag{25.231c}$$

14. Il y a moyen de démontrer cela via le lemme de Jordan. Nous donnons ici une démonstration moins technologique.

et nous avons bien $\lim_{R\to\infty}|I_2(R)|=0$. Nous passons à la troisième intégrale. En tenant compte que $e^{i\pi/2}=i$, nous avons

$$I_3(R)=-\int_0^R e^{-\gamma_3(t)^2}e^{i\pi/4}dt \tag{25.232a}$$
$$=-\frac{1+i}{\sqrt{2}}\int_0^R e^{-t^2}e^{2i\pi/4} \tag{25.232b}$$
$$=-\frac{1+i}{\sqrt{2}}\int_0^R e^{-it^2}. \tag{25.232c}$$

En passant à la limite $R\to 0$, de l'équation (25.227) il ne reste que

$$0=\frac{\sqrt{2}}{2}-\frac{1+i}{\sqrt{2}}\int_0^\infty e^{-it^2}dt, \tag{25.233}$$

ce qui signifie que

$$\int_0^\infty e^{-it^2}dt=\frac{\sqrt{2\pi}}{2(1+i)}=\frac{\sqrt{\pi}}{2}e^{-i\pi/4}. \tag{25.234}$$

25.7.3 Logarithme complexe

25.7.3.1 La fonction argument

Nous savons la définition 25.43 de l'exponentielle complexe.

Définition 25.48.
*Un **logarithme** de $\alpha\in\mathbb{C}$ est une solution de l'équation $e^z=\alpha$.*

Notons bien que cela définit *un* logarithme, et non *le* logarithme.

Lemme 25.49.
Si z_1 et z_2 sont des logarithmes de α alors il existe $k\in\mathbb{Z}$ tel que $z_1=z_2+2ik\pi$.

Démonstration. Nous commençons par déterminer les logarithmes de $\alpha=1$. Nous avons besoin de $e^{a+bi}=1$ ($a,b\in\mathbb{R}$). Nous avons

$$e^ae^{bi}=1, \tag{25.235}$$

et en prenant la norme nous trouvons $|e^a|=1$, ce qui donne $a=0$. Ensuite $e^{bi}=1$, qui signifie $b=2k\pi$. Les logarithmes de 1 sont donc les nombres de la forme $2ik\pi$.

Soient maintenant z_1 et z_2 des logarithmes de α. Alors $e^{z_1}=e^{z_2}$, donc [15] $e^{z_1-z_2}=1$, ce qui signifie que z_1-z_2 est un logarithme de 1. Donc il existe un $k\in\mathbb{Z}$ tel que $z_1-z_2=2ik\pi$. □

Remarque 25.50.
Jusqu'ici nous n'avons pas donné de conditions donnant l'existence d'un logarithme. Nous avons seulement supposé des existences et donné des propriétés sur ces hypothétiques objets.

Définition 25.51 ([292])**.**
*Si $z\in\mathbb{C}^*$ nous définissons la **valeur principale** de son argument le nombre $\theta\in\left]-\pi,\pi\right]$ tel que*

$$z=|z|e^{i\theta} \tag{25.236}$$

Nous le notons $\arg(z)$.

25.52.
Il ne faut pas se ruer sur $\arg(x+iy)=\arctan(y/x)$.Pour rappel, la fonction arctan a été définie

15. C'est facile de dire «donc». Il faut surtout citer la proposition 25.44(2).

dans le théorème 16.38, et elle prend ses valeurs dans $]-\pi/2, \pi/2[$. La formule $v(x,y) = \arctan(y/x)$ n'est donc valable que pour $x > 0$. Les valeurs sont :

$$\arg(x+iy) = \begin{cases} \arctan(y/x) & \text{si } x > 0 \\ \pi + \arctan(y/x) & \text{si } x < 0 \text{ et } y \geqslant 0 \\ -\pi + \arctan(y/x) & \text{si } x < 0 \text{ et } y < 0 \\ \frac{\pi}{2} & \text{si } x = 0 \text{ et } y > 0 \\ \frac{-\pi}{2} & \text{si } x = 0 \text{ et } y < 0. \end{cases} \tag{25.237}$$

Pour $x > 0$ nous avons $\arg(x+iy) = \arctan(y/x)$ parce que justement la fonction arctan prend ses valeurs en particulier entre $-\pi$ et π. Pour $x < 0$ et $y > 0$ nous avons $\arg(x+iy) = \pi + \arctan(y/x)$ (dans ce cas, $\arctan(y/x) < 0$) et si $x < 0$, $y < 0$ nous avons $\arg(x+iy) = -\pi + \arctan(y/x)$.

25.53 (Les dérivées partielles de la fonction argument)**.**
Vu que nous en aurons besoin plusieurs fois, nous calculons maintenant les dérivées partielles de la fonction

$$\begin{aligned} \varphi\colon \mathbb{R}^2 &\to \mathbb{R} \\ (x,y) &\mapsto \arg(x+iy). \end{aligned} \tag{25.238}$$

Nous commençons par la dérivée $\partial_x\varphi(x,y)$. Et il y a de nombreux cas à séparer.

<u>$x > 0$</u> Nous avons

$$\frac{\partial\varphi}{\partial x}(x,y) = \lim_{\epsilon\to 0} \frac{\arctan(y/(x+\epsilon)) - \arctan(y/x)}{\epsilon}, \tag{25.239}$$

qui n'est autre que la dérivée de la fonction $x \mapsto \arctan(y/x)$. Nous pouvons la calculer facilement avec le théorème 16.38(2) :

$$\frac{\partial\varphi}{\partial x}(x,y) = -\frac{y}{x^2+y^2}. \tag{25.240}$$

<u>$x < 0$</u> Nous avons

$$\frac{\partial\varphi}{\partial x}(x,y) = \lim_{\epsilon\to 0} \frac{\pm\pi + \arctan(y/(x+\epsilon)) - \big(\pm\pi + \arctan(y/x)\big)}{\epsilon} \tag{25.241}$$

où les signes $\pm$ dépendent du signe de y. De toutes façons, les termes en π se simplifient et le calcul est le même que celui du cas $x > 0$. Encore une fois nous avons

$$\frac{\partial\varphi}{\partial x}(x,y) = -\frac{y}{x^2+y^2}. \tag{25.242}$$

<u>$x = 0$</u> Nous devons calculer

$$\frac{\partial\varphi}{\partial x}(0,y) = \lim_{\epsilon\to 0} \frac{\arg(\epsilon+iy) - \arg(iy)}{\epsilon}. \tag{25.243}$$

Il y a quatre cas d'après les signes de ϵ (séparer limite à gauche et à droite) et y.

Si $\epsilon > 0$ et $y > 0$ alors nous avons à faire le calcul

$$\lim_{\epsilon\to 0^+} \frac{\arctan(y/\epsilon) - \pi/2}{\epsilon} \tag{25.244}$$

qui se traite par la règle de l'Hospital. Cela donne $-1/y$.

Les trois autres cas ne se distinguent que par des constantes au numérateur, lesquelles disparaissent en appliquant la règle de l'Hospital [16]. Au final,

$$\frac{\partial\varphi}{\partial x}(0,y) = -\frac{1}{y}. \tag{25.245}$$

16. Nonobstant le fait que ces constantes se mettent bien pour avoir un vrai cas d'indétermination 0/0, sinon la règle de l'Hospital ne s'applique pas.

Nous avons calculé jusqu'ici :

$$\frac{\partial\varphi}{\partial x}(x,y) = \frac{-y}{x^2+y^2} \tag{25.246}$$

pour tout $(x,y) \in \mathbb{R}^2\backslash\{(0,0)\}$. En particulier vous avez noté que cette dérivée partielle est continue sur $\mathbb{R}^2\backslash\{(0,0)\}$.

Nous calculons à présent la dérivée partielle par rapport à y :

$$\frac{\partial\varphi}{\partial y}(x,y) = \lim_{\epsilon\to 0}\frac{\arg(x+iy+i\epsilon)-\arg(x+iy)}{\epsilon}. \tag{25.247}$$

$\underline{x>0}$ Nous avons à calculer

$$\lim_{\epsilon\to 0}\frac{\arctan\frac{y+\epsilon}{x}-\arctan\frac{y}{x}}{\epsilon}, \tag{25.248}$$

qui n'est autre que la dérivée de la fonction $t\mapsto \arctan\frac{t}{x}$ en $t=y$. Résultat :

$$\frac{\partial\varphi}{\partial y}(x,y) = \frac{x}{x^2+y^2}. \tag{25.249}$$

$\underline{x<0 \textbf{ et } y\neq 0}$ Le calcul à faire est :

$$\lim_{\epsilon\to 0}\frac{\pm\pi+\arctan\frac{y+\epsilon}{x}-\left(\pm\pi+\arctan\frac{y}{x}\right)}{\epsilon} \tag{25.250}$$

Une chose importante à remarquer est que dans le calcul de la limite nous pouvons supposer que y et $y+\epsilon$ aient le même signe, quelle que soit la valeur et le signe de ϵ (assez petit). C'est pour cela que les deux termes $\pm\pi$ arrivent avec le même signe des deux côtés de la différence, et se simplifient. Nous tombons sur une limite déjà faite et

$$\frac{\partial\varphi}{\partial y}(x,y) = \frac{x}{x^2+y^2} \tag{25.251}$$

$\underline{x<0 \textbf{ et } y=0}$ Vu que $x<0$ nous avons $\arg(x)=\pi$ et nous devons calculer

$$\lim_{\epsilon\to 0}\frac{\arg(x+i\epsilon)-\pi}{\epsilon}. \tag{25.252}$$

La limite $\epsilon\to 0^+$ est classique et donne $1/x$.

Mais la limite $\epsilon\to 0^-$ n'existe pas :

$$\lim_{\epsilon\to 0^-}\frac{-\pi+\arctan(\epsilon/x)-\pi}{\epsilon} \tag{25.253}$$

n'existe pas.

Donc

$$\frac{\partial\varphi}{\partial y}(x,0) \tag{25.254}$$

n'existe pas pour $x<0$.

$\underline{x=0 \textbf{ et } y\neq 0}$ Le calcul est immédiat

$$\lim_{\epsilon\to 0}\frac{\arg(iy+i\epsilon)-\arg(iy)}{\epsilon} = 0, \tag{25.255}$$

donc

$$\frac{\partial\varphi}{\partial y}(0,y) = 0. \tag{25.256}$$

En ce qui concerne la continuité, nous avons que $\partial_y\varphi$ est continue partout sauf sur la demi-droite $\{(x,0) \text{ tel que } x\leqslant 0\}$ où elle n'existe pas.

25.7.3.2 Une définition possible du logarithme

Définition 25.54.
Nous définissons la fonction ***logarithme*** *par*

$$\begin{aligned} \ln\colon \mathbb{C}^* &\to \mathbb{C} \\ z &\mapsto \ln\big(|z|\big) + i\arg(z) \end{aligned} \tag{25.257}$$

où le ln *à droite est le logarithme usuel sur* $\mathbb{R}^+$.

Remarque 25.55.
Cette fonction généralise le logarithme déjà vu sur $]0,\infty[\subset \mathbb{R}$. En effet pour des valeurs de z dans cette partie nous avons $\arg(z) = 0$ et $|z| = z$.

Lemme 25.56.
Le nombre $\ln(z)$ *est un logarithme de* z.

Démonstration. Nous avons

$$e^{\ln(z)} = e^{\ln|z|}e^{i\arg(z)} = |z|e^{i\arg(z)} = z. \tag{25.258}$$

Nous avons utilisé le fait que $e^{\ln(x)} = x$ pour $x \in \mathbb{R}^+$ et $|z|e^{i\arg(z)} = z$ par définition de la fonction arg. □

Notons que si on avait pris d'autres conventions pour définir arg, nous aurions eu d'autres définitions possibles de ln.

Exemple 25.57
Nous avons

$$\ln(-1) = \ln(1) + i\arg(-1). \tag{25.259}$$

Mais $\ln(1) = 0$ et $\arg(-1) = \pi$ (et non $-\pi$), donc

$$\ln(-1) = i\pi. \tag{25.260}$$

C'est cette définition du logarithme qui est prise par Sage, et c'est cela qui lui permet de donner la primitive de $1/x$ comme $\ln(x)$ et non $\ln(|x|)$, parce que Sage connait les logarithmes de nombres réels négatifs :

```
1
2  SageMath version 7.3, Release Date: 2016-08-04
3  Type "notebook()" for the browser-based notebook interface.
4  Type "help()" for help.
5
6 sage: ln(-1)
7 I*pi
8 sage: f(x)=1/x
9 sage: f.integrate(x)
10 x |--> log(x)
```

tex/sage/sageSnip010.sage

△

Nous avons jusqu'ici défini une fonction sur $\mathbb{C}^*$ qui fait correspondre à chaque nombre complexe un de ses logarithmes. Il reste quelque questions à régler :

— Est-ce que cette fonction est continue ? Holomorphe ? (réponses : non et non)

— Si non, est-ce qu'il y avait moyen de trouver une définition plus efficace ? (réponse : non)

Lemme 25.58.
La fonction ln *n'est pas continue sur* $]-\infty,0]$.

Démonstration. Attention à bien comprendre l'énoncé. La fonction

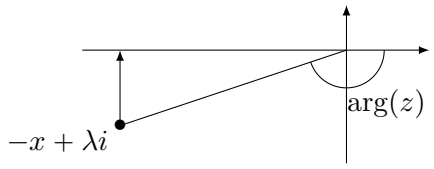
$$\begin{aligned} f\colon]-\infty,0[&\to \mathbb{C} \\ x &\mapsto \ln(x) \end{aligned} \tag{25.261}$$

est continue. D'ailleurs c'est $\ln(x) = \ln(|x|) + i\pi$. Ce dont il est question dans l'énoncé, c'est de la fonction ln vue comme fonction sur $\mathbb{C}^*$.

Soit $x > 0$ dans $\mathbb{R}$; nous avons

$$\ln(-x) = \ln(x) + i\pi. \tag{25.262}$$

Cependant $\lim_{\substack{\lambda\to 0^-\\ \lambda\in\mathbb{R}}} \ln(-x+\lambda i)$ va valoir $\ln(|x| - i\pi)$. En effet lorsque $\lambda < 0$ est petit, l'argument de $-x + \lambda i$ se rapproche de $-\pi$ (et non de π).



Donc

$$\lim_{\substack{\lambda\to 0^-\\ \lambda\in\mathbb{R}}} \ln(-x+\lambda i) = \lim \ln(|x+\lambda i|) + i\arg(-x+\lambda i) = \ln(|x|) - i\pi. \tag{25.263}$$

Nous n'avons donc pas continuité de la fonction logarithme comme fonction sur $\mathbb{C}^*$. □

Théorème 25.59.
La restriction

$$\ln\colon \mathbb{C}\backslash]-\infty,0] \to \mathbb{C} \tag{25.264}$$

est holomorphe.

Démonstration. Nous allons utiliser la proposition 15.7 et considérer la fonction

$$\begin{aligned} F\colon S &\to \mathbb{R}^2 \\ (x,y) &\mapsto \big(\ln(|x+iy|), \arg(x+iy)\big) \end{aligned} \tag{25.265}$$

où $S = \mathbb{R}^2\backslash\{(x,0) \text{ tel que } x \leqslant 0\}$. Nous devons vérifier que F est différentiable et que sa différentielle en un point de S est une similitude.

Nous posons

$$u(x,y) = \ln\big(\sqrt{x^2+y^2}\big) \tag{25.266}$$

et

$$v(x,y) = \arg(x+iy). \tag{25.267}$$

Les dérivées partielles de u ne sont pas très compliquées :

```
sage: var('x,y')
(x, y)
sage: u(x,y)=ln(sqrt(x**2+y**2))
sage: u.diff(x)
(x, y) |--> x/(x^2 + y^2)
```

tex/sage/sageSnip011.sage

c'est à dire

$$\frac{\partial u}{\partial x} = \frac{x}{x^2+y^2} \tag{25.268a}$$

$$\frac{\partial u}{\partial y} = \frac{y}{x^2+y^2}. \tag{25.268b}$$

Pour celles de v par contre, il faut se poser des questions, par exemples résister à la tentation d'écrire $v(x,y) = \arctan(y/x)$ et lire 25.52.

Nous avons déjà calculé les dérivées partielles de v dans 25.53, et nous avons vu qu'elles étaient continues sur $\mathbb{R}^2$ privé de la demi-droite.

Vu que les dérivées partielles sont continues, la proposition 12.204 nous dit que F est différentiable. La matrice de la différentielle est alors la matrice des dérivées partielles

$$\begin{pmatrix} \frac{x}{x^2+y^2} & \frac{y}{x^2+y^2} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{pmatrix}, \tag{25.269}$$

qui a la forme requise (15.12) pour que la proposition 15.7 nous assure que ln soit $\mathbb{C}$-dérivable, c'est à dire holomorphe. □

25.7.3.3 Pas plus de continuité

Bon. La fonction logarithme que nous avons définie est holomorphe sur $\mathbb{C}^*$ privé d'une demi-droite U. Et elle n'est pas continue sur U ; elle y est cependant continue «par le haut». Pouvons-nous faire mieux ? Nous allons maintenant prouver quelque résultats d'impossibilité de faire mieux que holomorphe partout sauf une partie pas si petite que ça.

Proposition 25.60.
Il n'existe pas de fonctions continues $f\colon \mathbb{C}^* \to \mathbb{C}$ telle que $e^{f(z)} = z$ pour tout $z \in \mathbb{C}^*$.

Démonstration. Pour tout z, le nombre $f(z)$ est un logarithme de z. Or $\ln(z)$ en est également un. Donc par le lemme 25.49

$$f(z) = \ln(z) + 2ik(z)\pi \tag{25.270}$$

pour une certaine fonction $k\colon \mathbb{C}^* \to \mathbb{Z}$. Sur le domaine d'holomorphie de ln, les fonction ln et f étant continues, la fonction k l'est aussi. Mais une fonction continue à valeurs dans $\mathbb{Z}$ est constante (son domaine est connexe).

Il existe donc $k \in \mathbb{Z}$ tel que

$$f(z) = \ln(z) + 2ik\pi \tag{25.271}$$

au moins pour tout $z \in \mathbb{C}^* \backslash U$. Une telle fonction ne peut pas être continue sur U parce que ln ne l'est pas. □

Ok. Pas continue sur tout $\mathbb{C}$. Mais continue sur un peu plus que $\mathbb{C}$ privé de toute une demi-droite ? La proposition suivante répond que bof.

Proposition 25.61.
Soit Ω un ouvert de $\mathbb{C}$ contenant $S(0,r)$ (le cercle centré en 0 et de rayon $r > 0$). Il n'existe pas de fonction continue $f\colon \Omega \to \mathbb{C}$ telle que $e^{f(z)} = z$ pour tout $z \in \Omega$.

Démonstration. Encore une fois, pour tout $z \in \Omega$ nous avons

$$f(z) = \ln(z) + 2i\pi k(z) \tag{25.272}$$

pour une certaine fonction $k\colon \Omega \to \mathbb{Z}$. Sur $\Omega \backslash U$, la fonction ln est continue et k doit également l'être. Donc k est constante sur les composantes connexes de $\Omega \backslash U$.

Vu que $S(0,r)$ est compact, on peut le recouvrir par un nombre fini de boules centrées en des points de $S(0,r)$. En prenant le minimum des rayons de ces boules, nous voyons que Ω contient une couronne

$$\{z \in \mathbb{C} \text{ tel que } r - \delta \leqslant |z| \leqslant r + \delta\}. \tag{25.273}$$

Soit le point $x_0 = -r$. C'est un point de Ω contenu dans U. Nous allons prouver que $B(x_0, \delta)\backslash U$ est dans une seule composante connexe de Ω.

Soit un point $z_1 \in B(x_0, \delta)$ situé au-dessus de U, et z_2 un point de $B(x_0, \delta)$ situé en dessous de U. Le cercle $S(0, r)$ coupe $B(x_0, \delta)$ en deux points : un au-dessus et un en-dessous de U. On peut lier z_1 au point de «sortie» supérieur de $S(0, r)$ en restant dans $B(x_0, \delta)$; ce point est ensuite relié en suivant le cercle au point d'entrée inférieur du cercle dans $B(x_0, \delta)$. Ce dernier point est lié à z_2 par un chemin restant dans la boule.

Tout cela pour dire que z_1 et z_2 sont dans la même composante connexe de Ω et que $k(z_1) = k(z_2)$. Il existe donc $k \in \mathbb{Z}$ tel que

$$f(z) = \ln(z) + 2ik\pi \tag{25.274}$$

sur $B(x_0, \delta)\backslash U$. Une telle fonction f ne peut pas être continue. □

25.7.3.4 Pas d'unicité : autres déterminations de l'argument

Nous avons pris la fonction d'argument $\arg\colon \mathbb{C} \to]-\pi, \pi]$. Il y en a évidemment beaucoup d'autres de possibles. Par exemple pour $\alpha \in \mathbb{R}$ nous pouvons considérer

$$\arg_{\alpha^+}\colon \mathbb{C} \to]\alpha, \alpha + 2\pi] \tag{25.275}$$

ou

$$\arg_{\alpha^-}\colon \mathbb{C} \to [\alpha, \alpha + 2\pi[. \tag{25.276}$$

En posant

$$\ln_{\alpha^\pm}(z) = \ln(|z|) + i \arg_{\alpha^\pm}(z) \tag{25.277}$$

nous avons une fonction réciproque de l'exponentielle définie sur $\mathbb{C}^*$ et holomorphe sur $\mathbb{C}^*$ privé d'une demi-droite D_α (dépendante de la valeur de ψ).

La différence entre $\ln_{\alpha^+}$ et $\ln_{\alpha^-}$ est seulement la valeur sur la demi-droite de non-holomorphie. L'une sera semi-continue d'un côté et l'autre, de l'autre côté.

Définition 25.62 ([293])**.**
*Soit un ouvert $\Omega \subset \mathbb{C}^*$. Nous disons que la fonction $f\colon \Omega \to \mathbb{C}$ est une **détermination** sur Ω si elle est continue et vérifie*

$$e^{f(z)} = z \tag{25.278}$$

pour tout $z \in \mathbb{C}$.

Les différents résultats vus jusqu'ici montrent qu'il n'existe pas de détermination du logarithme sur $\mathbb{C}^*$.

Définition 25.63.
*La **détermination principale** du logarithme est la restriction de notre logarithme 25.54*

$$\begin{aligned} \ln\colon \mathbb{C}^* &\to \mathbb{C} \\ z &\mapsto \ln(|z|) + i \arg(z) \end{aligned} \tag{25.279}$$

à l'ouvert $\mathbb{C}^*\backslash U$ où U est la partie $\Re(z) \leqslant 0$ de $\mathbb{C}$.

Remarque 25.64.
Beaucoup de sources[294] ne définissent pas $\ln_{\alpha^\pm}$ sur la droite D_α. C'est à dire qu'ils notent $\ln_\alpha$ notre fonction $\ln_{\alpha^+}$ restreinte à $\mathbb{C}^*\backslash D_\alpha$. Dans ce cas, les fonctions $\ln_{\alpha^+}$ et $\ln_{\alpha^-}$ sont identiques[17].

Cette remarque est importante parce que certains vont vous dire «le logarithme n'est pas définit sur la demi-droite» ; de leur point de vue, la fonction que nous avons définie est une prolongation (non continue) à U du logarithme, qui est continu.

17. Cela n'est pas tout à fait évident ; vous devriez y penser.

(1) Certaines personnes pourraient vous dire que notre logarithme «n'est pas bien définit parce que si on fait le tour dans un sens ou dans l'autre nous n'obtenons pas la même valeur pour $\ln(z)$ lorsque z est sur U». Et cela avec des arguments aussi forts que «2π et 0, c'est le même point».

Nous préférons être bien clairs sur ce point : notre fonction ln est parfaitement définie sur $\mathbb{C}^*$ et 2π n'est pas la même chose que zéro. En particulier $\arg(e^{2i\pi}) = 0$ et $\arg(e^{-i\pi}) = \pi$ et non π.

(2) Il n'en reste pas moins que Sage donne $\ln(-1) = I\pi$ et que nous avons choisi de faire de même, parce que le Frido n'est pas un cours d'agrégation, mais un texte qui donne quelque éléments de mathématique dans le but d'utiliser Sage efficacement.

(3) Tout ceci pour dire que si vous utilisez ce livre pour l'agrégation, vous devriez sérieusement considérer l'option de ne pas donner du logarithme la définition donnée ici, mais bien sa restriction.

En fait notre logarithme est maximum pour la propriété «être une réciproque de l'exponentielle» alors que beaucoup de monde préfère avoir une fonction maximale pour la propriété «être réciproque de l'exponentielle tout en étant continue».

De toutes les fonctions ayant le droit de vouloir être appelée «logarithme», celle que nous avons choisie (un peu arbitrairement) pour s'appeler «logarithme» et accaparer de la notation «ln» est $\ln_{\pi^+}$. Elle est d'une certaine manière celle qui arrive le plus naturellement.

En effet si nous pensons au logarithme népérien $\ln\colon \mathbb{R}^+ \to \mathbb{R}$ que nous voulons prolonger sur $\mathbb{R}$, nous devons poser

$$\ln(-x) = \ln(-1) + \ln(x) \tag{25.280}$$

pour $x > 0$. Que peut valoir $\ln(-1)$? Il doit vérifier $e^{\ln(-1)} = -1$. La première valeur qui nous tombe sous la main est $\ln(-1) = \pi$. Bien entendu, d'autres possibilités étaient possibles, comme $\ln(-1) = 2017\pi$ par exemple.

25.7.3.5 Pas d'unicité : développement en série

Pour $z_0 \in \mathbb{C}^*$ nous pouvons écrire un développement en série de la réciproque de l'exponentielle autour de z_0. La fonction ainsi définie est holomorphe sur la boule $B(z_0, |z_0|)$ et diverge en dehors de cette boule.

Voila encore une fonction «logarithme» pour chaque point de $\mathbb{C}^*$. Nous nommons $\ln_{z_0}$ la fonction

$$\ln_{z_0}\colon B(z_0, |z_0|) \to \mathbb{C} \tag{25.281}$$

donnée par la série.

En général nous n'avons pas $\ln_{z_1} = \ln_{z_2}$ sur l'intersection des disques de convergence. Si c'était le cas, de proche en proche nous pourrions construire une fonction continue réciproque du logarithme sur $\mathbb{C}^*$, ce qui est impossible.

25.7.3.6 Pas d'unicité : laquelle choisir ?

Bon. Pour chaque demi-droite D nous avons une détermination du logarithme sur $\mathbb{C}^*\backslash D$. Et pour tout $z_0 \in \mathbb{C}^*$ nous en avons une sur $B(z_0, |z_0|)$.

En pratique, quel logarithme choisir ? Cela dépend du problème.

Si vous avez besoin ou envie de travailler avec des série entières, le mieux est de choisir une détermination donnée par un développement autour d'un point bien choisit par rapport à votre problème.

Si vous avez surtout besoin d'holomorphie, et que vous en avez besoin sur un grand domaine, vous devriez choisir une détermination sur un des ensembles $\mathbb{C}^*\backslash D_\alpha$ en choisissant α de telle sorte que la demi-droite maudite ne passe pas par la zone sur laquelle vous travaillez.

Dans tous les cas, vous devez préciser très explicitement la détermination choisie. Dans ce texte, sauf mention du contraire, nous utiliserons la détermination principale, et même son extension (non

continue) à $\mathbb{C}^*$. Lorsque nous aurions besoin d'holomorphie, nous préciserons que nous considérons la restriction.

25.7.3.7 Logarithme comme primitive

Tout le monde sait que le logarithme $\ln\colon \mathbb{R}^+ \to \mathbb{R}$ est une primitive de la fonction $x \mapsto 1/x$. Qu'en est-il dans le cas complexe ? Tout d'abord précisons que nous ne comptons pas encore parler d'intégrale sur $\mathbb{C}$, mais seulement d'intégrales sur $\mathbb{R}$ d'une fonction à valeur complexes.

Proposition 25.65.
Si $z \in \mathbb{C}$ alors

$$\int \frac{1}{x+z}dx = \ln(x+z) \tag{25.282}$$

Démonstration. Il est important de comprendre que la formule (25.282) est un abus de notation pour dire que si nous considérons la fonction

$$\begin{aligned}\varphi\colon \mathbb{R} &\to \mathbb{C}\\ x &\mapsto \ln(x+z)\end{aligned} \tag{25.283}$$

alors nous avons $\varphi'(x) = \frac{1}{x+z}$. Ici la dérivation est une dérivation sur $\mathbb{R}$ et l'intégrale est une intégrale sur $\mathbb{R}$, c'est à dire «composante par composantes». La fonction φ se décompose en partie réelle et imaginaire qui sont à dériver séparément :

$$\varphi(x) = \ln(|x+z|) + i\arg(x+z). \tag{25.284}$$

Si z est imaginaire pur Nous posons $z = \lambda i$ avec $\lambda \in \mathbb{R}^*$. D'abord nous avons

$$\frac{1}{x+\lambda i} = \frac{x}{x^2+\lambda^2} - i\frac{\lambda}{x^2+\lambda^2}. \tag{25.285}$$

La partie réelle de $\varphi(x)$ est

$$\varphi_1(x) = \ln\left(\sqrt{x^2+\lambda^2}\right), \tag{25.286}$$

dont la dérivée est

$$\varphi_1'(x) = \frac{x}{x^2+\lambda^2}, \tag{25.287}$$

qui correspond bien à la partie réelle de $\frac{1}{x+\lambda i}$.
En ce qui concerne la partie imaginaire, $\varphi_2(x) = \arg(x+\lambda i)$, et sa dérivée n'est rien d'autre que la dérivée partielle par rapport à x de la fonction argument, déjà calculée en (25.246) :

$$\varphi_2'(x) = \frac{-\lambda}{x^2+\lambda}. \tag{25.288}$$

Cela est bien la partie imaginaire de $\frac{1}{x+\lambda i}$.
Notons que nous n'avons pas de problèmes sur la demi-droite des réels négatifs parce que nous ne considérons au final que la dérivée partielle par rapport à x de la fonction argument, laquelle existe et est continue, même sur cette partie.

Pour z quelconque Soit $z = s + \lambda i$ avec $s, \lambda \in \mathbb{R}$. En posant $\varphi_0(x) = \ln(x+\lambda i)$ nous avons $\varphi(x) = \varphi_0(x+s)$ et donc

$$\varphi'(x) = \varphi_0'(x+s) = \frac{1}{x+s+\lambda i} = \frac{1}{x+z}. \tag{25.289}$$

Tout va bien.

□

Exemple 25.66
Un petit calcul d'intégrale, que nous avions déjà faite dans l'exemple 14.93 (avec la méthode de Rothstein-Trager). En passant par une décomposition en fractions simples :

$$\int \frac{1}{x^3+x} = \int \left(\frac{1}{x} - \frac{1/2}{x-i} - \frac{1/2}{x+i} \right) \tag{25.290a}$$

$$= \ln(x) - \frac{1}{2}\ln(x-i) - \frac{1}{2}\ln(x+i) \tag{25.290b}$$

$$= \ln(x) - \frac{1}{2}\ln(x^2+1). \tag{25.290c}$$

Attention aux justifications. Il n'est pas vrai en général dans le cas de nombres complexes a et b que $\ln(ab) = \ln(a) + \ln(b)$. En effet, pour la partie réelle, ça passe parce que $|ab| = |a||b|$. Mais en ce qui concerne la partie imaginaire,

$$\arg(ab) \neq \arg(a) + \arg(b) \tag{25.291}$$

lorsque la somme dépasse les bornes de $]-\pi, \pi]$. Le passage à (25.290c) fonctionne parce que dans le cas particulier des nombres $x+i$ et $x-i$, les arguments se somment à zéro : $\arg(x+i)+\arg(x-i) = 0$.
△

25.8 Théorème de Weierstrass

Théorème 25.67 (Théorème de Weierstrass[295])**.**
Soit (f_n) une suite de fonctions holomorphes sur un ouvert Ω de $\mathbb{C}$ que nous supposons converger uniformément sur tout compact vers f. Alors f est holomorphe sur Ω et pour tout k nous avons

$$f_n^{(k)} \to f^{(k)} \tag{25.292}$$

uniformément sur tout compact.

Dit en peu de mots, la limite uniforme d'une suite de fonctions holomorphes est holomorphe, et on peut permuter la limite avec la dérivation.

Démonstration. Chacune des fonctions f_n étant holomorphes, si $a \in \Omega$ et r est tel que $B(a,r) \subset \Omega$, nous avons par la formule de Cauchy 25.13 :

$$f_n(z) = \frac{1}{2\pi i} \int_{\partial B(a,r)} \frac{f_n(\xi)}{\xi - z} d\xi \tag{25.293}$$

pour tout z dans un boule $B(a,\rho)$ incluse dans $B(a,r)$. Étant donné que le cercle ∂B est compact, elle y est majorée par une constante M. Montrons que de plus nous pouvons choisir M de telle façon à avoir $|f_n(\xi)| \leqslant M$ pour tout n et tout ξ en même temps. D'abord nous utilisons la continuité de la limite f sur le compact ∂B pour poser $A = \max_{z\in\partial B} |f(z)|$. Ensuite nous considérons un $\epsilon > 0$ et N tel que $\| f_n - f\|_{\partial B} \leqslant \epsilon$ pour tout $n \geqslant N$. Nous savons maintenant que

$$\{|f_n(\xi)| \text{ tel que } n \geqslant N, \xi \in \partial B\} \tag{25.294}$$

est majoré par $A + \epsilon$. Nous posons enfin

$$B = \max_{n\leqslant N} \max_{\xi\in\partial B} |f_n(z)|, \tag{25.295}$$

et alors le nombre $M = \max\{A+\epsilon, B\}$ majore $|f_n(\xi)|$ pour tout n et tout $\xi \in \partial B$.

De plus pour tout $\xi \in \partial B$ et pour tout z dans la petite boule, nous avons $|\xi - z| > r - \rho$, donc la fonction dans l'intégrale est majorée par une constante ne dépendant ni de n ni de ξ. Nous pouvons donc permuter l'intégrale et la limite sur n :

$$f(z) = \frac{1}{2i\pi} \int_{\partial B} \frac{f(\xi)}{\xi - z}. \tag{25.296}$$

Cela implique que la fonction f est holomorphe par le corollaire 25.16.

Nous voudrions maintenant parler des dérivées des f_n et de f. Pour cela nous voulons permuter l'intégrale et les dérivées, ce qui est fait au corollaire 25.26 :

$$f_n^{(k)} = \frac{1}{2\pi i}\int_{\partial B(z_0,r)} \frac{f(\omega)}{(\omega - z)^{k+1}} d\omega. \tag{25.297}$$

Nous voulons la convergence sur tout compact contenu dans l'ouvert Ω. Pour ce faire, nous allons considérer un compact $K \subset \Omega$ et prouver la convergence uniforme dans toute boule de la forme $B(z_0, r)$ avec $z_0 \in K$ et $B(z_0, r) \subset \Omega$. Pour chaque tel couple (z_0, r), nous aurons un $N_{(z_0,r)} \in \mathbb{N}$ tel que si $n \geqslant N_{(z_0,r)}$,

$$\| f_n^{(k)} - f^{(k)} \|_{B(z_0,r)} \leqslant \epsilon. \tag{25.298}$$

Vu que ces boules $B(z_0, r)$ forment un recouvrement de K par des ouverts, nous pouvons en retirer un sous-recouvrement fini et prendre, comme N, le maximum des $N_{(z_0,r)}$ correspondants. Pour ce N nous aurons

$$\| f_n^{(k)} - f^{(k)} \|_K \leqslant \epsilon. \tag{25.299}$$

Au travail !

Pour $z \in B(z_0, r)$ nous considérons $r' > r$ tel que $B(z_0, r') \subset \Omega$ et nous avons

$$|f_n^{(k)}(z) - f^{(k)}(z)| = \left| \frac{1}{2\pi i}\int_{\partial B(z_0,r')} \frac{f_n(\xi) - f(\xi)}{(\xi - z)^{k+1}} d\xi \right| \tag{25.300a}$$

$$\leqslant \frac{1}{2\pi}\int_{\partial B(z_0,r')} \frac{|f_n(\xi) - f(\xi)|}{|r - r'|^{k+1}} d\xi. \tag{25.300b}$$

Nous avons pris ce r' de telle manière que $|\xi - z|$ soit borné par le bas par $|r - r'|$; sinon la majoration que nous venons de faire ne marche pas. Étant donné que $f_n \to f$ uniformément, nous pouvons considérer n assez grand pour que le numérateur soit plus petit que ϵ indépendamment de ξ et de z. Donc pour un n assez grand,

$$|f_n^{(k)}(z) - f^{(k)}(z)| \leqslant \frac{\epsilon}{2\pi} \frac{2\pi r'}{|r - r'|^{k+1}} \tag{25.301}$$

pour tout $z \in B(z_0, r)$. Donc nous avons convergence uniforme $f_n^{(k)} \to f^{(k)}$ sur cette boule. Par l'argument de compacité donné plus haut, nous avons la convergence uniforme sur tout compact. □

25.9 Théorème de Montel

Théorème 25.68 (Montel[2])**.**
Soit Ω un ouvert de $\mathbb{C}$ et $\mathcal{F}$ une famille de fonctions holomorphes sur Ω, uniformément bornée sur tout compact de Ω. Alors de toute suite dans $\mathcal{F}$ nous pouvons extraire une sous-suite convergeant uniformément sur tout compact de Ω.

Démonstration. **Un ensemble équicontinu** Nous commençons par prendre une suite de compacts dans Ω comme dans le lemme 6.156, et une suite δ_n de réels strictement positifs tels que

$$B(z, 2\delta_n) \subset K_{n+1} \tag{25.302}$$

pour tout $z \in K_n$. Soient $x, y \in K_n$ tels que $|x - y| < \delta_n$; nous notons $\partial B(x, 2\delta_n)$ le cercle de rayon $2\delta_n$ autour de x, parcouru dans le sens positif. La formule de Cauchy 25.36 nous donne

$$f(x) - f(y) = \frac{1}{2\pi i}\int_{\partial B} \left(\frac{f(\xi)}{\xi - x} - \frac{f(\xi)}{\xi - y} \right) d\xi = \frac{x - y}{2\pi i}\int_{\partial B} \frac{f(\xi)}{(\xi - x)(\xi - y)} d\xi \tag{25.303}$$

Nous majorons ça par

$$\big|f(x)-f(y)\big| \leqslant \frac{|x-y|}{2\pi}\int_{\partial B}\frac{|f(\xi)|}{2\delta_n^2}d\xi \leqslant \frac{|x-y|}{\delta_n}M_n. \tag{25.304}$$

Justifications :

— $|\xi - x| = 2\delta_n$ et $|\xi - y| \geqslant \delta_n$ parce que ξ est au mieux sur le rayon passant par x et y.
— $|f(\xi)| \leqslant M_n$ où M_n est la borne uniforme de $\mathcal{F}$ sur le compact K_n.
— Nous avons aussi fini par calculer l'intégrale dans laquelle il ne restait plus rien, ça a donné la circonférence du cercle de rayon $2\delta_n$.

Jusqu'à présent nous avons prouvé que l'ensemble

$$\mathcal{F}_n = \{f|_{K_n} \text{ tel que } f \in \mathcal{F}\} \tag{25.305}$$

est équicontinu. Il est aussi équiborné par hypothèse.

Application du théorème d'Ascoli L'ensemble $\mathcal{F}_n$ vérifie les hypothèses du théorème d'Ascoli 24.7. Donc l'ensemble $\mathcal{F}_n$ est relativement compact dans $C(K_n, \mathbb{C})$ pour la norme uniforme. Autrement dit l'ensemble $\bar{\mathcal{F}}$ est compact et si nous avons une suite de fonctions dans $\mathcal{F}_n$, il existe une sous-suite convergeant dans $\bar{\mathcal{F}}_n$, c'est à dire uniformément. Autrement dit il existe une fonction strictement croissante $\varphi\colon \mathbb{N} \to \mathbb{N}$ telle que la suite $k \mapsto f_{\varphi(k)}$ converge uniformément sur K_n. La limite n'est cependant pas spécialement dans $\mathcal{F}_n$.

L'argument diagonal La suite $k \mapsto f_{\varphi_1\circ\dots\varphi_k(k)}$ converge uniformément sur tous les K_n. Si K est un compact de Ω, alors les petites propriétés sympas du lemme 6.156 nous disent que $K \subset \text{Int}(K_m)$ pour un certain m. Ladite suite convergeant uniformément sur K_m, elle converge uniformément sur K et nous avons montré la convergence uniforme sur tout compact de Ω.

□

Corollaire 25.69 ([2])**.**
Soit Ω un ouvert connexe borné de $\mathbb{C}$ et $a \in \Omega$. Soit f holomorphe sur Ω telle que $f(a) = a$ et $|f'(a)| < 1$.

Alors de (f^n) on peut extraire une sous-suite convergeant uniformément sur tout compact de Ω vers la fonction constante a.

Démonstration. Nous considérons un voisinage de a inclus à Ω ; sachant que $|f(a)| < 1$, nous trouvons un voisinage encore plus petit de a sur lequel $|f'(z)| < 1$. Soit donc r tel que $\overline{B(a,r)} \subset \Omega$ et tel que $|f'(z)| < 1$ sur $\overline{B(a,r)}$. Étant donné que $f'(z)$ est continue sur le compact $\overline{B(a,r)}$, nous en prenons le maximum λ (qui est strictement inférieur à 1) et nous avons au final

$$|f'(z)| \leqslant \lambda < 1 \tag{25.306}$$

pour tout $z \in \overline{B(a,r)}$. Le théorème des accroissements finis 12.241 nous dit que

$$\big|f(z) - a\big| \leqslant \lambda|z-a| \tag{25.307}$$

pour tout $z \in \overline{B(a,r)}$. C'est ici que nous utilisons l'hypothèse de convexité de Ω. Nous montrons alors par récurrence que

$$\big|f^n(z) - a\big| \leqslant \lambda^n|z-a| \leqslant \lambda^n r \leqslant r. \tag{25.308}$$

L'ensemble $A = \{f^n \text{ tel que } n \geqslant 1\}$ est donc uniformément borné sur $\overline{B(a,r)}$ par $a + r$. Autre manière de le dire : pour tout $z \in \overline{B(a,r)}$ nous avons

$$f^n(z) \in \overline{B(a,r)}. \tag{25.309}$$

La suite (f^n) est donc uniformément bornée sur tout compact de $B(a,r)$. Le théorème de Montel 25.68 nous indique que l'on peut extraire une sous-suite convergente uniformément sur tout compact. Au vu de (25.308) cette convergence ne peut avoir lieu que vers une fonction g qui vaut la constante a sur $B(a,r)$.

D'autre par la fonction g est holomorphe en tant que limite uniforme de fonctions holomorphes, théorème 25.67. Or une fonction holomorphe constante sur un ouvert est constante sur tout son domaine d'holomorphie (principe d'extension analytique, théorème 18.120). □

25.10 Espaces de Bergman

Source : [209].
Soit Ω un borné dans $\mathbb{C}$ et D le disque unité ouvert de $\mathbb{C}$.

Définition 25.70.
*L'**espace de Bergman** sur Ω, noté $A^2(\Omega)$ est l'espace des fonctions holomorphes sur Ω qui sont en même temps dans $L^2(\Omega)$.*

Nous mettons sur $A^2(\Omega)$ le produit scalaire usuel hérité de L^2 :

$$\langle f, g\rangle = \int_\Omega f(z)\overline{g(z)}dz. \tag{25.310}$$

Lemme 25.71.
Soit $K \subset \Omega$ un compact et $f \in A^2(\Omega)$. Alors

$$\max_{z\in K} |f(z)| \leqslant \frac{1}{\sqrt{\pi}} \frac{1}{d(K, \partial\Omega)} \|f\|_2. \tag{25.311}$$

Démonstration. Soient $a \in \Omega$ et $r > 0$ tels que $B(a, r) \subset \Omega$. Nous considérons aussi $\rho \leqslant r$. La formule de Cauchy (25.36) nous donne

$$f(a) = \frac{1}{2\pi i} \int_{B(a,\rho)} \frac{f(\xi)}{\xi - a} f\xi = \frac{1}{2\pi} \int_0^{2\pi} f(a + \rho e^{i\theta})d\theta \tag{25.312}$$

où nous avons utilisé le chemin $\gamma(\theta) = a + \rho e^{i\theta}$, $\gamma'(\theta) = i\rho e^{i\theta}$ et $\rho = |\xi - a|$. Maintenant une astuce est d'écrire

$$\frac{r^2}{2} f(a) = \int_0^r f(a)\rho d\rho, \tag{25.313}$$

et d'y substituer la valeur de $f(a)$ que nous venons de calculer :

$$\begin{aligned}
\frac{r^2}{2} f(a) &= \int_0^r \frac{1}{2\pi} \int_0^{2\pi} f(a + \rho e^{i\theta})d\theta\rho d\rho && (25.314a)\\
&= \frac{1}{2\pi} \int_{B(a,r)} f(z)dz && \text{passage aux polaires} \quad (25.314b)\\
&= \frac{1}{2\pi} \langle 1, f\rangle_B && \text{produit scalaire sur } B(a, r) \quad (25.314c)\\
&\leqslant \frac{1}{2\pi} \sqrt{\langle 1, 1\rangle_B \langle f, f\rangle_B} && (25.314d)
\end{aligned}$$

Nous avons donc

$$r^2 f(a) \leqslant \frac{1}{\pi} \sqrt{\langle 1, 1\rangle_B \langle f, f\rangle_B}, \tag{25.315}$$

et donc

$$\pi r^2 f(a) \leqslant \sqrt{\pi r^2} \|f\|_2, \tag{25.316}$$

parce que $\langle f, f\rangle_B \leqslant \|f\|_2^2$. En effet le produit scalaire $\|.\|_2$ est donné par une intégrale sur Ω alors que $B(a, r) \subset \Omega$ et que la fonction qu'on y intègre est positive (c'est $|f(z)|^2$). En simplifiant,

$$f(a) \leqslant \frac{1}{\sqrt{\pi} r} \|f\|_2. \tag{25.317}$$

Mais r a été choisit pour avoir $B(a, r) \subset \Omega$, donc $r \leqslant d(a, \partial\Omega)$ et

$$|f(a)| \leqslant \frac{1}{d(a, \partial\Omega)\sqrt{\pi}} \|f\|_2. \tag{25.318}$$

Maintenant si nous prenons $a \in K$, nous avons encore la minoration $d(a, \partial K) \leqslant d(a, \partial\Omega)$ et donc

$$|f(a)| \leqslant \frac{1}{d(a, \partial K)\sqrt{\pi}} \|f\|_2. \tag{25.319}$$

□

Théorème 25.72.
Soit Ω un ouvert de $\mathbb{C}$.

(1) L'espace $A^2(\Omega)$ est un espace de Hilbert.

(2) Si D est la boule unité dans $\mathbb{C}$, une base hilbertienne de $A^2(D)$ est donnée par les fonctions

$$e_n(z) = \sqrt{\frac{n+1}{\pi}} z^n \tag{25.320}$$

pour $n \geqslant 0$.

Démonstration. Nous commençons par montrer que $A^2(\Omega)$ est complet. Pour cela nous considérons une suite de Cauchy (f_n) dans $A^2(\Omega)$ et un compact $K \subset \Omega$. Nous savons par le lemme 25.71 que

$$\max_{z\in K} \big|f_n(z) - f_m(z)\big| \leqslant \frac{1}{\sqrt{\pi} d(K, \partial\Omega)} \|f_n - f_m\|_2. \tag{25.321}$$

Donc f_n converge uniformément sur K. Par le théorème de Weierstrass 25.67, la fonction f est holomorphe. Il existe donc une fonction holomorphe f qui est limite uniforme sur tout compact de Ω de la suite (f_n).

Mais $L^2(\Omega)$ étant complet, la suite (f_n) a une limite $g \in L^2(\Omega)$. Ce que nous voudrions faire est prouver que $f = g$. Notons que tel quel, ce n'est pas vrai parce que f est une vraie fonction alors que g est une classe. Ce que nous enseigne la proposition 24.14 est qu'il existe une sous-suite (qu'on note (g_n)) qui converge vers g presque partout. Dans cette dernière phrase, g_n et g sont de vraies fonctions, des représentants des classes dans L^2.

Nous déduisons que $f = g$ presque partout (ici f et g sont les fonctions) parce que la sous-suite converge uniformément vers f en même temps que presque partout vers g. Donc $f = g$ dans $L^2(\Omega)$ (ici f et g sont les classes). Donc $f \in L^2(\Omega)$ et l'espace $A^2(\Omega)$ est de Hilbert.

Il nous faut encore prouver que $(e_n)_{n\geqslant 0}$ est une base orthonormale. En ce qui concerne les produits scalaires,

$$\begin{aligned}
\langle e_m, e_n\rangle &= \sqrt{\frac{(m+1)(n+1)}{\pi}} \int_D z^n \overline{z^m} dz && (25.322a)\\
&= \sqrt{\frac{(m+1)(n+1)}{\pi^2}} \int_0^1 \rho\, d\rho \int_0^{2\pi} d\theta \rho^{m+n} e^{i\theta(n-m)} && (25.322b)\\
&= \sqrt{\frac{(m+1)(n+1)}{\pi^2}} \frac{1}{m+n+2} \underbrace{\int_0^{2\pi} e^{i\theta(n-m)} d\theta}_{2\pi\delta_{mn}} && (25.322c)\\
&= \sqrt{\frac{(n+1)^2}{\pi^2}} \frac{1}{2n+2} 2\pi\delta_{nm} && (25.322d)\\
&= \delta_{nm}. && (25.322e)
\end{aligned}$$

Donc les fonctions données sont bien orthonormales. Nous devons montrer qu'elles sont denses dans $A^2(D)$. Soit $f \in A^2(D)$ et $c_n(f) = \langle f, e_n\rangle$; nous allons montrer que

$$\|f\|_2^2 = \sum_{n=0}^{\infty} |\langle f, e_n\rangle|^2, \tag{25.323}$$

parce que le point (5) du théorème 23.45 nous indique que ce sera suffisant pour avoir une base hilbertienne.

Étant donné que f est holomorphe sur D, le théorème 25.13 nous développe f en série entière :

$$f(z) = \sum_{k=0}^{\infty} a_k z^k. \tag{25.324}$$

En permutant la somme avec le produit scalaire,

$$c_n(f) = \int_D f(z)\bar{e}_n(z) = \sqrt{\frac{n+1}{\pi}} \int_D f(z)\bar{z}^n dz. \tag{25.325}$$

Afin de profiter de la convergence uniforme de la série (25.324) à l'intérieur de D, nous allons exprimer l'intégrale sur D comme une intégrale sur $|z| < r$ en faisant tendre r vers 1 (par le bas). Pour ce faire nous considérons les fonctions

$$g_k(z) = \begin{cases} f(z)\bar{z}^n & \text{si } |z| < 1 - 1/k \\ 0 & \text{sinon.} \end{cases} \tag{25.326}$$

Ces fonctions sont intégrables sur D et dominées par $f(z)\bar{z}^n$ qui est intégrable sans dépendre de k. Mais nous avons évidemment $g_k(z) \to f(z)\bar{z}^n$. Le théorème de la convergence dominée permet alors de permuter l'intégrale et la limite $k \to \infty$. Cela nous permet d'écrire

$$c_n(f) = \sqrt{\frac{n+1}{\pi}} \lim_{r\to 1^-} \int_{|z|<r} \bar{z}^n f(z) dz = \sqrt{\frac{n+1}{\pi}} \lim_{r\to 1^-} \int_{|z|<r} \sum_{k=0}^{\infty} a_k z^k \bar{z}^n. \tag{25.327}$$

Par la convergence uniforme de la série entière *à l'intérieur* du disque D nous pouvons permuter l'intégrale et la somme (proposition 15.46) :

$$c_n(f) = \sqrt{\frac{n+1}{\pi}} \lim_{r\to 1^-} \sum_{k=0}^{\infty} a_k \int_{|z|<r} z^k \bar{z}^n dz. \tag{25.328}$$

L'intégrale proprement dite est vite calculée et vaut

$$\int_{|z|<1} \bar{z}^n z^k dz = \frac{\pi r^{2n+2}}{n+1} \delta_{kn}. \tag{25.329}$$

Nous pouvons donc continuer le calcul de $c_n(f)$ en effectuant la somme sur k qui se réduit à changer k en n puis en effectuant la limite :

$$c_n(f) = \sqrt{\frac{n+1}{\pi}} \lim_{r\to 1^-} \sum_k a_k \frac{\pi r^{2n+2}}{n+1} \delta_{kn} = \sqrt{\frac{\pi}{n+1}} a_n. \tag{25.330}$$

Nous effectuons le même genre de calculs pour évaluer $\|f\|_2^2$:

$$\|f\|_2^2 = \int_D |f(z)|^2 dz \tag{25.331a}$$

$$= \lim_{r\to 1^-} \int_{|z|<r} f(z) \sum_{k=0}^{\infty} \bar{a}_k \bar{z}_k dz \tag{25.331b}$$

$$= \lim_{r\to 1^-} \sum_{k=0}^{\infty} \bar{a}_k \int_{|z|<r} f(z)\bar{z}^k dz \qquad \text{permuter } \sum \text{ et } \int \tag{25.331c}$$

$$= \lim_{r\to 1^-} \sum_{k=0}^{\infty} \bar{a}_k a_k \frac{\pi r^{2k+2}}{k+1} \qquad \text{intégrale déjà faite.} \tag{25.331d}$$

Mais nous savons déjà que $c_n(f) = \sqrt{\pi/(n+1)}$, donc ce qui est dans la somme est $\pi\bar{a}_k a_k/(n+1) = |c_k(f)|^2$. Nous avons donc

$$\|f\|_2^2 = \lim_{r\to 1^-} \sum_{k=0}^{\infty} |c_k(f)|^2 r^{2k+2}. \tag{25.332}$$

La fonction (de r) constante $|c_k(f)|^2$ domine $|c_k(f)r^{2k+2}|$ tout en ayant une somme (sur k) qui converge ; en effet la proposition 23.25 nous indique que $\sum_j |c_k(f)|^2 \leqslant \|f\|_2^2$. Le théorème de la convergence dominée nous permet d'inverser la limite et la somme pour obtenir le résultat attendu :

$$\|f\|_2^2 = \sum_{k=0}^{\infty} |c_k(f)|^2. \tag{25.333}$$

□

Chapitre 26

Série de Fourier

Soit f une fonction ; nous définissons ses **coefficients de Fourier** par

$$c_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-int} \tag{26.1}$$

avec plus de détails en 24.5.1.

26.1 Densité des polynômes trigonométriques

26.1.1 Convergence pour les fonctions continues (via Weierstrass)

Le résultat fondamental qui nous permet d'utiliser les polynômes trigonométriques comme base pour les fonctions *continues* périodiques est le suivant. Notons que pour les fonctions non continues, il y a encore du travail.

Lemme 26.1.
Si $f\colon \mathbb{R} \to \mathbb{C}$ est une fonction continue 2π-périodique et si $\epsilon > 0$, alors il existe un polynôme trigonométrique P tel que $\|f - P\|_\infty \leqslant \epsilon$.

Démonstration. Nous allons utiliser le théorème de Stone-Weierstrass 18.4. Soit le compact Hausdorff

$$S^1 = \{z \in \mathbb{C} \text{ tel que } |z| = 1\}, \tag{26.2}$$

et $C(S^1, \mathbb{C})$ l'algèbre des fonctions continues de S^1 vers $\mathbb{C}$. Il suffit de vérifier que les polynômes trigonométriques vérifient les hypothèse du théorème de Stone-Weierstrass. Un polynôme trigonométrique est un polynôme en z et $\bar{z}$ défini sur S^1.

(1) Le polynôme constant est dans l'algèbre, ok.

(2) Pour la séparation des points, le polynôme trigonométrique $x \mapsto e^{ix}$.

(3) Si P est un polynôme en z et $\bar{z}$, alors $\bar{P}$ l'est encore.

Donc si $\epsilon > 0$ et $\tilde{f} \in C(S^1, \mathbb{C})$ sont donnés, il existe un polynôme trigonométrique P tel que

$$\sum_t |\tilde{f}(e^{it}) - P(t)| < \epsilon. \tag{26.3}$$

Soit $f\colon \mathbb{R} \to \mathbb{C}$ une fonction continue 2π-périodique. Nous considérons $\tilde{f} \in C(S^1, \mathbb{C})$ donnée par $\tilde{f}(e^{it}) = f(t)$. Alors $\sup_t |f(t) - P(t)| \leqslant \epsilon$. □

26.1.2 Convergence pour les fonctions continues (via Fejér)

Si nous ne voulons pas passer par le gros théorème de Stone-Weierstrass pour prouver la densité des polynômes trigonométrique dans $\big(C^0_{2\pi}, \|.\|_\infty\big)$, nous pouvons passer par le gros théorème de Fejér. C'est ce que nous faisons maintenant.

Le **noyau de Dirichlet** est la fonction

$$D_n(t) = \sum_{k=-n}^{n} e^{int}. \tag{26.4}$$

Le **noyau de Fejér** est la moyenne de Cesaro des noyaux de Dirichlet :

$$F_n(t) = \frac{1}{n} \sum_{k=0}^{n-1} D_k(t). \tag{26.5}$$

Lemme 26.2.
Le noyau de Dirichlet s'exprime sous la forme

$$D_n(t) = \sum_{k=-n}^{n} e^{-ikt} = \frac{\sin\left(\frac{2n+1}{2}t\right)}{\sin(t/2)} \tag{26.6}$$

Note : ce noyau n'est pas positif.

Démonstration. Nous commençons par mettre en facteur le premier terme :

$$D_n(t) = \sum_{k=-n}^{n} e^{int} = e^{-int} \sum_{k=0}^{2n} e^{ikt}. \tag{26.7}$$

En utilisant la formule de la somme géométrique,

$$D_n(t) = e^{-int} \frac{1 - (e^{it})^{2n+1}}{1 - e^{it}} \tag{26.8a}$$

$$= e^{-int} \frac{1 - e^{(2n+1)it}}{1 - e^{it}} \tag{26.8b}$$

$$= e^{-int} \frac{e^{(2n+1)it/2}}{e^{i\frac{t}{2}}} \frac{e^{-(2n+1)it/2} - e^{(2n+1)it/2}}{e^{-it/2} - e^{it/2}} \tag{26.8c}$$

$$= \frac{(-2i)\sin\left(\frac{2n+1}{2}t\right)}{(-2i)\sin\left(\frac{t}{2}\right)}. \tag{26.8d}$$

□

Théorème 26.3 (Théorème de Dirichlet)**.**
Soit f une fonction 2π-périodique et C^1 par morceaux. Pour tout $x \in \mathbb{R}$ nous posons

$$s_n(x) = \sum_{k=-n}^{n} c_k(f) e^{ikx}. \tag{26.9}$$

Alors nous avons

$$\lim_{n\to\infty} s_n(x) = \frac{f(x^+) + f(x^-)}{2}. \tag{26.10}$$

Lemme 26.4.
Le noyau de Fejér s'exprime sous la forme

$$F_n(t) = \frac{1}{n} \left(\frac{\sin \frac{nt}{2}}{\sin \frac{t}{2}} \right)^2. \tag{26.11}$$

Note : ce noyau est positif. C'est important parce qu'on s'en sert dans la preuve du théorème de Fejér.

Démonstration. L'astuce est de noter $\sin(x) = \Im(e^{ix})$ et de repartir du résultat à propos du noyau de Dirichlet. En utilisant encore la formule de la série géométrique partielle,

$$F_n(t) = \frac{1}{n\sin(t/2)}\Im\sum_{k=0}^{n-1} e^{(2k+1)it/2} \tag{26.12a}$$

$$= \frac{1}{n\sin(t/2)}\Im e^{\frac{it}{2}}\sum_{k=0}^{n-1} \tag{26.12b}$$

$$= \frac{1}{n\sin(t/2)}\Im e^{\frac{it}{2}}\left(\frac{1-e^{nit}}{1-e^{it}}\right) \tag{26.12c}$$

$$= \frac{1}{n\sin(t/2)}\Im e^{it/2}\frac{e^{\frac{nit}{2}}\left(e^{-\frac{int}{2}}-e^{\frac{nit}{2}}\right)}{e^{\frac{it}{2}}\left(e^{-it/2}-e^{it/2}\right)} \tag{26.12d}$$

$$= \frac{1}{n\sin(t/2)}\underbrace{\Im e^{\frac{nit}{2}}}_{\sin(nt/2)}\frac{\sin\left(\frac{nt}{2}\right)}{\sin(\frac{t}{2})} \tag{26.12e}$$

$$= \frac{1}{n}\left(\frac{\sin\frac{nt}{2}}{\sin\frac{t}{2}}\right)^2. \tag{26.12f}$$

□

Théorème 26.5 (Fejèr)**.**
Soit une fonction continue et 2π-périodique $f\colon \mathbb{R}\to\mathbb{C}$. Pour tout $k\in\mathbb{Z}$ nous notons

$$\begin{aligned} e_k\colon \mathbb{R} &\to \mathbb{C} \\ x &\mapsto e^{ikx}. \end{aligned} \tag{26.13}$$

Pour chaque $n\in\mathbb{N}$ nous posons

$$D_n = \sum_{k=-n}^{n} e_k \qquad \tilde{S}_n(f) = \sum_{k=-n}^{n} c_k(f)e_k \tag{26.14a}$$

$$F_n = \frac{D_0+\cdots+D_{n-1}}{n} \qquad \tilde{F}_n = \sigma_n(f) = \frac{1}{n}\sum_{k=0}^{n-1} S_k(f). \tag{26.14b}$$

Alors

1. *(1)* $\frac{1}{2\pi}\int_{-\pi}^{\pi} F_n(t)dt = 1$.
2. *(2) Pour tout $\alpha\in\,]0,\pi[$, F_n converge uniformément vers 0 sur $[-\pi,\pi]\backslash[-\alpha,\alpha]$.*
3. *(3) La suite $\tilde{F}_n$ converge uniformément sur $\mathbb{R}$ vers f.*
4. *(4) Le système trigonométrique $\{e_k\}_{k\in\mathbb{Z}}$ est total pour l'espace $\left(C^0(S^1),\|.\|_\infty\right)$ des fonctions continues 2π-périodiques.*

Démonstration. Un calcul usuel montre que

$$\int_{-\pi}^{\pi} e_l(t)dt = \begin{cases} 0 & \text{si } l\neq 0 \\ 2\pi & \text{si } l=0 \end{cases} \tag{26.15}$$

Nous avons alors

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} F_n(t)dt = \frac{1}{2\pi}\frac{1}{n}\sum_{k=0}^{n-1}\sum_{l=-k}^{k}\underbrace{\int_{-\pi}^{\pi} e_l(t)dt}_{2\pi\delta_l} = \frac{1}{n}\sum_{k=0}^{n-1} 1 = 1. \tag{26.16}$$

Cela prouve déjà le premier point.

Pour le second point, en partant de l'expression (26.11) et en considérant $x \in [-\pi, \pi,]\backslash[-\alpha, \alpha]$ (ce qui nous évite l'annulation du dénominateur),

$$|F_n(x)| \leqslant \frac{1}{(n+1)\sin^2(\alpha/2)}, \tag{26.17}$$

et donc $F_n \to 0$ uniformément sur l'ensemble considéré.

Nous passons maintenant à cette histoire de convergence uniforme de la moyenne de Cesaro vers f. Pour tout $n \in \mathbb{N}$ nous avons

$$\tilde{D}_n(x) = \frac{1}{2\pi} \sum_{k=-n}^{n} \left(\int_{-\pi}^{\pi} f(t) e^{-ikt} dt \right) e^{ikx} \tag{26.18a}$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \sum_{k=-n}^{n} e_k(x-t) \tag{26.18b}$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) D_k(x-t). \tag{26.18c}$$

Par conséquent, en effectuant le changement de variable $u = x - t$ et la périodicité,

$$\tilde{F}_n(x) = \int_{-\pi}^{\pi} f(t) F_n(x-t) dt \tag{26.19a}$$

$$= -\int_{x+\pi}^{x-\pi} f(x-u) F_n(u) du \tag{26.19b}$$

$$= \int_{-\pi}^{\pi} f(x-u) F_n(u) du. \tag{26.19c}$$

Nous prouvons à présent l'uniforme continuité. Soit $\epsilon > 0$; étant donné que f est continue et 2π-périodique, elle est uniformément continue et nous considérons $\delta > 0$ tel que $|x-y| < \delta$ implique $\big|f(x) - f(y)\big| < \epsilon$. Soit M un majorant de $|f|$ sur $\mathbb{R}$. L'équation (26.19) nous donne

$$\big|f(x) - \tilde{F}_n(x)\big| = \big|\frac{1}{2\pi} \int_{-\pi}^{\pi} \big(f(x-t) - f(x)\big) F_n(t) dt\big| \tag{26.20a}$$

$$\leqslant \frac{1}{2\pi} \int_{\delta \leqslant |t| \leqslant \pi} |2M F_n(t)| dt + \frac{1}{2\pi} \int_{-\delta}^{\delta} \epsilon |F_n(t)| dt \tag{26.20b}$$

$$\leqslant \frac{2M}{2\pi} \int_{\delta \leqslant |t| \leqslant \pi} F_n(t) dt + \epsilon' \tag{26.20c}$$

Pour obtenir (26.20a) nous avons pu rentrer $f(x)$ dans l'intégrale en utilisant le premier point. Pour obtenir (26.20c) nous avons d'abord utilisé la positivité de F_n (lemme 26.4) pour enlever les valeurs absolues, et nous avons ensuite utilisé le fait que son intégrale valait 2π.

Étant donné que $F_n \to 0$ uniformément sur $[-\pi, \pi,]\backslash[-\alpha, \alpha]$, il existe un N tel que

$$\int_{\delta \leqslant |t| \leqslant \pi} F_n(t) dt \leqslant \epsilon \tag{26.21}$$

dès que $n > N$. Le résultat découle.

Pour le point (4), il suffit de remarquer que chacun des $\tilde{F}_n$ est une combinaison finie d'éléments du système trigonométrique. □

26.1.3 Densité dans L^p

Nous venons de voir (de deux façons différentes) que les polynômes trigonométriques étaient dense dans $\big(C^0_{2\pi}(\mathbb{R}), \|.\|_\infty\big)$. Nous avons aussi déjà vu par le théorème 24.51 que ces polynômes trigonométriques étaient denses dans $L^p(S^1)$. Nous présentons à présent une autre façon de prouver cette dernière densité.

Théorème 26.6.
Les polynômes trigonométriques sont denses dans $L^p(S^1)$ pour $1 \leqslant p < \infty$.

Démonstration. Par les théorèmes 26.1 ou 26.5 (au choix), nous savons que les polynômes trigonométriques sont denses dans $\big(C^0_{2\pi}(S^1), \|.\|_\infty\big)$. Vu que S^1 est compact, la densité est également au sens L^p. En effet si $\|f_n - f\|_\infty \leqslant \epsilon$, alors

$$\|f_n - f\|_\infty = \int_0^{2\pi} |f_n - f|^p \leqslant \int_0^{2\pi} \epsilon^p = 2\pi\epsilon^p. \tag{26.22}$$

Donc les polynômes trigonométriques sont denses dans $\big(C^0_{2\pi}(S^1), \|.\|_p\big)$. Mais nous savons par (un a fortiori sur) le théorème 24.33 que les fonction continues sont denses dans $L^p(S^1)$.

Par densité de la densité, les polynômes trigonométriques sont denses dans $L^p(S^1)$. □

26.1.4 Suite équirépartie, critère de Weyl

Définition 26.7.
Soit u une suite dans $[0,1]$. Pour $0 \leqslant a \leqslant b \leqslant 1$ nous posons

$$X_n(a,b) = \text{Card}\big\{k \in \{1,\dots,n\} \text{ tel que } u_k \in [a,b]\big\}. \tag{26.23}$$

*Nous disons que la suite u est **équirépartie** si pour tout $0 \leqslant a < b < 1$, on a*

$$\lim_{n\to\infty} \frac{X_n(a,b)}{n} = b - a. \tag{26.24}$$

Voir aussi la remarque 34.132 sur les nombres normaux.

Proposition 26.8 (Critère de Weyl[209, 2])**.**
Soit (x_n) une suite dans $[0,1[$. Les conditions suivantes sont équivalentes.

(1) La suite (x_n) est équirépartie.

(2) Pour toute fonction continue à valeurs réelles sur $[0,1]$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(x_k) = \int_0^1 f(x)dx. \tag{26.25}$$

(3) Pour tout $p \in \mathbb{N}^*$ nous avons

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} e^{2i\pi p x_k} = 0. \tag{26.26}$$

Démonstration. On pose

$$S_n(f) = \frac{1}{n} \sum_{k=1}^{n} f(x_k). \tag{26.27}$$

Une espèce de lemme Supposons connaitre un ensemble de fonctions A dense dans $C^0([0,1])$ pour toutes les fonctions duquel nous avons la limite (26.25). Alors la limite a lieu pour toute fonction de $C^0([0,1])$. En effet, soit $f \in C^0([0,1])$ et $g \in A$ tel que $\|f-g\|_\infty < \epsilon$. Alors

$$\left\| \frac{1}{n} \sum_{k=1}^{n} f(x_k) - \int_0^1 f(t)dt \right\| \leqslant \left\| \frac{1}{n} \sum_{k=1}^{n} \big(f(x_k) - g(x_k)\big) \right\| \tag{26.28a}$$

$$+ \left\| \frac{1}{n} \sum_{k=1}^{n} g(x_k) - \int_0^1 g(t)dt \right\| \tag{26.28b}$$

$$+ \left\| \int_0^1 g(t)dt - \int_0^1 f(t)dt \right\|. \tag{26.28c}$$

Le premier terme se majore par ϵ. Le troisième est la même majoration : $\int_0^1 \big(f(t)-g(t)\big)dt \leqslant \|f-g\|_\infty = \epsilon$. Par hypothèse sur l'espace A, le second terme se majore par ϵ lorsque n est grand.

(1)⇒(2) Nous supposons que la suite est équirépartie et nous commençons par montrer le résultat pour les fonctions en escalier. Soit donc la fonction en escalier $\eta(x) = c_j$ sur $a_{j-1} < x < a_j$. Sur le point a_j lui-même, la fonction η vaut soit c_j soit c_{j+1}. Nous avons

$$\frac{1}{n}\sum_{k=1}^{n}\eta(x_k) = \frac{1}{n}\left[\sum_{j=1}^{m} c_j X_n(a_j,a_{j+1}) - \sum_{j=1}^{m} c_j X_n(a_j,a_j) + \sum_{j=1}^{m}\eta(a_j)X_n(a_j,a_j)\right]. \tag{26.29}$$

À la limite $n\to\infty$, les deux derniers termes tombent [1] et il reste

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\eta(x_k) = \sum_{j=1}^{m} c_j(a_{j-1}-a_j). \tag{26.30}$$

Or par construction, pour une fonction en escalier,

$$\sum_{j=1}^{m} c_j(a_{j-1}-a_j) = \int_0^1 \eta. \tag{26.31}$$

Étant donné que les fonctions en escalier sont denses dans les fonctions continues, l'espèce de lemme plus haut conclut.

(2)⇒(1) Nous prouvons maintenant le sens inverse. C'est à dire que pour toute fonction continue sur $[0,1]$, nous avons

$$\int_0^1 f(x)dx = \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} f(x_k). \tag{26.32}$$

Nous devons en déduire que (x_n) est équirépartie. Pour ce faire, soit $x\in[0,1[$ et $\epsilon>0$ tel que $x+\epsilon<1$. Nous considérons $\varphi = \mathbb{1}_{[x,1[}$ et

$$\varphi_{\epsilon(t)} = \begin{cases} 0 & \text{si } t\in[0,x[\\ \frac{t-x}{\epsilon} & \text{si } t\in[x,x+\epsilon[\\ 1 & \text{si } t\geqslant x+\epsilon. \end{cases} \tag{26.33}$$

Cela est une fonction continue, donc

$$\lim_{n\to\infty} S_n\big(\varphi_\epsilon(t)\big) = \int_0^1 \varphi_\epsilon(t)dt = \int_x^{x+\epsilon}\frac{t-x}{\epsilon}dt + \int_{x+\epsilon}^1 1dt = 1-x-\frac{\epsilon}{2}. \tag{26.34}$$

Mais $\varphi_\epsilon \leqslant \varphi$, donc $S_n(\varphi_\epsilon)\leqslant S_n(\varphi)$ et donc

$$\liminf_{n\to\infty} S_n(\varphi) \geqslant 1-x. \tag{26.35}$$

Notons que nous ne savons pas si la *vraie* limite de gauche existe ; c'est pourquoi nous prenons la limite inférieure, qui existe toujours.

Nous définissons aussi

$$\psi_\epsilon(t) = \begin{cases} 0 & \text{si } t\in[0,x-\epsilon[\\ \frac{t-x+\epsilon}{\epsilon} & \text{si } t\in[x-\epsilon,x[\\ 1 & \text{si } t>x. \end{cases} \tag{26.36}$$

C'est encore une fonction continue et nous trouvons [2]

$$\int_0^1 \psi_\epsilon(t)dt = 1-x+\frac{\epsilon}{2}. \tag{26.37}$$

1. J'en profite pour mentionner que mon équation (26.29) n'est pas la même que celle de [209] dans laquelle il me semble voir une faute ; quoi qu'il en soit, les termes litigieux tombent.

2. Je recommande chaudement de dessiner les fonctions φ_ϵ et ψ_ϵ pour avoir une idée de la situation.

Vu que $\psi_\epsilon \geqslant \varphi$, nous avons $S_n(\psi_\epsilon) \geqslant S_n(\varphi)$ et donc

$$\limsup_n S_n(\varphi) \leqslant 1 - x. \tag{26.38}$$

Nous avons déjà obtenu que

$$1 - x \leqslant \liminf S_n(\varphi) \leqslant \limsup S_n(\varphi) \leqslant 1 - x, \tag{26.39}$$

donc la limite existe et vaut

$$\lim_{n\to\infty} S_n(\varphi) = 1 - x. \tag{26.40}$$

Cela est pour la fonction caractéristique $\varphi = \mathbb{1}_{[x,1[}$. Si nous prenons une fonction caractéristique $\mathbb{1}_{[a,b]}$, nous avons la même chose parce que $\mathbb{1}_{[a,b[}$ est une combinaisons linéaire de fonctions du type $\mathbb{1}_{[x,1[}$.
Nous avons donc

$$\lim_{n\to\infty} S_n\big(\mathbb{1}_{[a,b]}\big) = b - a, \tag{26.41}$$

alors que le membre de gauche n'est autre que

$$S_n\big(\mathbb{1}_{[a,b]}\big) = \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{[a,b]}(x_k) = \frac{1}{n}N(n,a,b). \tag{26.42}$$

(2)⇒(3) Vu que $e^{2i\pi p x_k} = \cos(2\pi p x_k) + \sin(2\pi i x_k)$ est une fonction périodique, c'est immédiat.

(3)⇒(2) Par linéarité, le point (2) montre que si f est un polynôme trigonométrique, alors

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(x_k) = \int_0^1 f(t)dt. \tag{26.43}$$

Densité des polynômes trigonométriques Il nous reste à prouver que les polynômes trigonométriques sont denses dans les fonction continues sur $[0,1]$. Soit une fonction continue sur $[0,1]$ avec $f(0) = f(1)$. Alors le théorème de Stone-Weierstrass dans sa version trigonométrique (lemme 26.1) nous donne la densité.

Si $f(1) \neq f(0)$ c'est pas très grave : on peut trouver une fonction g vérifiant $g(0) = g(1)$ et $\|f - g\|_\infty \leqslant \epsilon$. Ensuite un polynôme trigonométrique approxime très bien g. .

□

26.2 Fonctions de Dirichlet

Définition 26.9.
*Une fonction $f\colon \mathbb{R} \to \mathbb{C}$ est une **fonction de Dirichlet** si*

(1) elle est 2π-périodique,

(2) elle est continue par morceaux,

(3) pour tout $x \in \mathbb{R}$ nous avons

$$f(x) = \frac{f(x^+) + f(x^-)}{2}. \tag{26.44}$$

Nous notons $\mathcal{D}$ l'ensemble des fonctions de Dirichlet.

Lemme 26.10 ([296])**.**
L'ensemble $C^0(S^1)$ est dense dans $\big(\mathcal{D}, \|.\|_2\big)$.

Démonstration. Nous commençons par supposer que $f \in \mathcal{D}$ n'ait qu'un seul point de discontinuité, x_0. Alors nous considérons la fonction f_n qui est égale à f sur $S^1 \backslash B(x_0, \frac{1}{n})$ et qui sur $B(x_n, \frac{1}{n})$ est le segment de droite joignant $f(x_0 - \frac{1}{n})$ et $f(x_0 + \frac{1}{n})$. Cela est une fonction continue, et de plus nous avons

$$|f_n(x)| \leqslant \|f\|_\infty \tag{26.45}$$

pour tout x. En effet si x est en dehors de $B(x_0, \frac{1}{n})$ c'est évident, et si $x \in B(x_0, \frac{1}{n})$, alors $|f_n(x)|$ est majoré soit par $f(x_0 - \frac{1}{n})$ soit par $f(x_0 + \frac{1}{n})$ suivant que le raccord affin soit croissant ou décroissant. Avec ça nous avons

$$\|f_n - f\|_2^2 = \int_{x_0-1/n}^{x_0+1/n} |f(x) - f_n(x)|^2 dx \leqslant \int_{x_0-1/n}^{x_0+1/n} 4\|f\|_\infty = \frac{8\|f\|_\infty}{n}. \tag{26.46}$$

Et nous voyons que $\|f_n - f\|_2 \to 0$.

Si f contient plusieurs points de continuité, on fait le même coup autour de chaque point, en prenant n assez grand pour que si x_0 est un point de discontinuité, $B(x_0, \frac{1}{n})$ n'en contienne pas d'autres. □

Notons que la densité de $C^0(S^1)$ dans $\big(\mathcal{D}, \|.\|_\infty\big)$ est impossible parce qu'une limite uniforme de fonctions continue est continue.

Théorème 26.11.
Le système trigonométrique $\{e_n\}_{n\in\mathbb{Z}}$ est total dans $\big(\mathcal{D}, \|.\|_2\big)$.

Démonstration. Soit $f \in \mathcal{D}$. Si elle est continue, le théorème de Fejèr 26.5 nous donne convergence uniforme sur S^1 d'une suite de polynômes trigonométriques vers f. Cette convergence est également une convergence L^2 parce que S^1 est compact.

Prenons donc $f \in \mathcal{D}$ non continue et $\epsilon > 0$ [3]. Par le lemme 26.10, il existe une fonction $g \in C^0(S^1)$ telle que

$$\|g - f\|_2 \leqslant \epsilon. \tag{26.47}$$

Le théorème de Fejèr donne aussi un polynôme trigonométrique P tel que $\|P - g\|_2 < \epsilon$; nous avons alors

$$\|P - f\|_2 \leqslant \|P - g\|_2 + \|g - f\|_2 \leqslant 2\epsilon. \tag{26.48}$$

□

Notons que cette histoire de fonctions de Dirichlet n'a pas attaquée le vrai fond du problème de la densité des polynômes trigonométriques dans $L^2(S^1)$ parce que nous restons avec une hypothèse de continuité, alors que les représentants des éléments de $L^2(S^1)$ n'ont strictement aucune régularité a priori.

26.3 Coefficients et série de Fourier

Définition 26.12.
*La **série de Fourier** associée à f est*

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n(f) e^{2\pi i \frac{n}{T} x}. \tag{26.49}$$

Cette expression est pour l'instant purement formelle. Cela ne présume ni de la convergence de la série, ni, au cas où elle serait convergente, que la limite soit f.

Pour la suite nous allons considérer des fonctions périodiques de période 2π, et les coefficients de Fourier de f (quand ils existent) sont alors

$$c_n(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-int} dt \tag{26.50}$$

3. Par exemple $\epsilon = 0.4$, mais ce n'est qu'un exemple hein. Si vous en voulez un autre, prenez p, un nombre premier puis calculez $\epsilon = 1/p$.

Proposition 26.13 ([297])**.**
Soit f une fonction continue et périodique telle que sa série de Fourier converge uniformément. Alors la convergence est vers f.

Démonstration. Notons d'abord que f étant continue sur $[0,2\pi]$, elle y est bornée et L^2. Par conséquent Parseval nous enseigne que

$$\|S_N(f)-f\|_{L^2}\to 0. \tag{26.51}$$

Cela signifie que

$$\lim_{N\to\infty}\frac{1}{2\pi}\int_0^{2\pi}|f(t)-S_N(t)|^2dt=0. \tag{26.52}$$

L'hypothèse de convergence uniforme nous dit que la fonction $|f(t)-S_N(t)|^2$ converge uniformément vers la fonction $|f(t)-S(t)|^2$ où nous avons écrit S la limite de S_N. En permutant la limite et l'intégrale,

$$\frac{1}{2\pi}\int_0^{2\pi}|f(t)-S(t)|^2dt=0, \tag{26.53}$$

ce qui signifie que la fonction $t\mapsto|f(t)-S(t)|^2$ est la fonction nulle. Nous en déduisons que $f=S$. □

Proposition 26.14.
Soit f une fonction 2π-périodique. Si $\sum_{n\in\mathbb{Z}}|c_n(f)|<\infty$, alors pour tout $x\in\mathbb{R}$ nous avons

$$f(x)=\sum_{n\in\mathbb{Z}}c_n(f)e^{inx}. \tag{26.54}$$

De plus, la suite (S_nf) converge uniformément vers f.

Démonstration. Nous posons

$$g(x)=\sum_{n\in\mathbb{Z}}c_n(f)e^{inx}. \tag{26.55}$$

Étant donné les hypothèses, la série de droite converge absolument, la fonction g est continue sur $\mathbb{R}$. Nous avons

$$\big|g(x)-(S_nf)(x)\big|\leqslant\sum_{|k|>n}|c_k(f)|, \tag{26.56}$$

mais le terme de droite tend vers zéro lorsque $n\to\infty$ parce que c'est le reste d'une série convergente. Cela signifie que S_nf converge uniformément vers g.

Par ailleurs nous savons que dans L^2 nous avons la convergence $S_nf\to f$ (parce que f est continue sur le compact $[0,2\pi]$ et donc y est bornée et L^2), ce qui signifie que $g=f$ presque partout au sens L^2. Ces deux fonctions étant continues, elles sont égales partout. □

Théorème 26.15.
Soit f, une fonction C^1 et 2π-périodique. Nous notons $(c_n)_{n\in\mathbb{Z}}$ la suite de ses coefficients de Fourier. Alors $(c_n)\in\ell^1(\mathbb{Z})$ et pour tout $x\in\mathbb{R}$ nous avons

$$f(x)=\sum_{n\in\mathbb{Z}}c_n(f)e^{inx}. \tag{26.57}$$

Démonstration. Soit $n\in\mathbb{Z}$. Nous posons $g(t)=f(t)e^{-int}$. Nous avons

$$0=g(2\pi)-g(0)=\int_0^{2\pi}g'(t)dt=\int_0^{2\pi}\big[f'(t)e^{-int}-inf(t)e^{-int}\big]. \tag{26.58}$$

Du coup, $c_n(f')=inc_n(f)$. La fonction f' étant bornée (parce que continue sur $[0,2\pi]$), elle est de carré intégrable sur $[0,2\pi]$ et par les inégalités de Parseval (théorème 23.45) nous avons

$$\sum_{n\in\mathbb{Z}}|c_n(f')|^2<\infty. \tag{26.59}$$

Par conséquent $(c_n(f')) \in \ell^2(\mathbb{Z})$ et a fortiori $(c_n(f'))_{n\in\mathbb{N}} \in \ell^2(\mathbb{N})$. L'inégalité de Cauchy-Schwartz nous indique alors

$$\sum_{n\in\mathbb{N}} |c_n(f)| = \sum_{n\in\mathbb{N}} \frac{1}{n}|c_n(f')| \leqslant \left(\sum_n \frac{1}{n^2}\right)^{1/2} \left(\sum_n |c_n(f')|^2\right)^{1/2} < \infty. \tag{26.60}$$

Nous procédons de même pour $n < 0$. Cela prouve que

$$\sum_{n\in\mathbb{Z}} |c_n(f)| < \infty. \tag{26.61}$$

□

Corollaire 26.16.
Soient f, g deux fonctions continues et 2π-périodiques. Si $c_n(f) = c_n(g)$ alors $f = g$.

Démonstration. Dans le cas de fonctions continues, le théorème de Fejér nous enseigne que si nous posons

$$S_n(x) = \sum_{k=-n}^{n} c_k(f)e^{ikx} \tag{26.62}$$

alors nous avons la convergence

$$\frac{1}{N+1}\sum_{n=0}^{N} S_n(f)(x) \to f(x). \tag{26.63}$$

C'est à dire qu'une fonction continue est déterminée par ses coefficients de Fourier. □

Exemple 26.17
Considérons la fonction

$$f(x) = 1 - \frac{x^2}{\pi^2} \tag{26.64}$$

sur $[-\pi, \pi]$. Nous la développons en série trigonométrique, et étant paire il n'y a pas de sinus. Un calcul montre que

$$a_0 = \frac{4}{3} \tag{26.65}$$

et

$$a_n = (-1)^{n+1}\frac{4}{n^2\pi^2}, \tag{26.66}$$

de telle sorte que

$$f(x) = \frac{2}{3} - \frac{4}{\pi^2}\sum_{n=1}^{\infty}(-1)^n\frac{\cos(nx)}{n^2}. \tag{26.67}$$

Nous avons $f(\pi) = 0$, mais vu le développement,

$$f(\pi) = \frac{2}{3} - \frac{4}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2}, \tag{26.68}$$

donc

$$\sum_{n=1}^{\infty}\frac{1}{n^2} = \frac{\pi^2}{6}. \tag{26.69}$$

△

26.3.1 Le contre-exemple que nous attendions tous

Nous montrons maintenant que la continuité et la périodicité ne sont pas suffisantes pour avoir convergence de la série de Fourier.

Proposition 26.18 ([2])**.**
Soit $C^0_{2\pi}$ l'ensemble des fonctions continues muni de la norme uniforme. Nous définissons

$$S_n(f)(x) = \sum_{k=-n}^{n} c_k(f)e^{ikx}. \tag{26.70}$$

Alors il existe $f \in C^0_{2\pi}$ tel que la suite $n \mapsto S_n(f)(0)$ soit divergente. En particulier f n'est pas la somme de sa série de Fourier.

Démonstration. Nous considérons la forme linéaire

$$\begin{aligned} l_n\colon C^0_{2\pi} &\to \mathbb{C} \\ f &\mapsto S_n(f)(0) = \sum_{k=-n}^{n} c_k(f). \end{aligned} \tag{26.71}$$

La forme est continue Nous montrons d'abord que $\|l_n\|$ est continue en montrant que $\|l_n\| < \infty$ et en utilisant la proposition 10.18. Pour cela nous calculons un peu :

$$l_n(f) = \sum_{k=-n}^{n} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\sum_{k=-n}^{n} e^{-ikt}dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)D_n(t)dt \tag{26.72}$$

où $D_n(t)$ est le noyaux de Dirichlet dont nous savons une formule par le lemme 26.2. Nous avons donc

$$|l_n(f)| \leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_n(t)|\|f\|_\infty dt. \tag{26.73}$$

En prenant $\|f\|_\infty = 1$ nous avons la borne suivante pour la norme de l_n :

$$\|l_n\| \leqslant \frac{1}{2n}\int_{-\pi}^{\pi} |D_n(t)|dt < \infty. \tag{26.74}$$

Notons que la convergence de l'intégrale vient de la continuité de la fonction

$$t \mapsto \frac{\sin\left(\frac{2n+1}{2}t\right)}{\sin\left(\frac{t}{2}\right)} \tag{26.75}$$

qui, elle même, se prouve avec une règle de l'Hospital :

$$\lim_{t\to 0}\frac{\sin(at)}{\sin(t)} = \lim_{t\to 0}\frac{a\cos(at)}{\cos(t)} = a. \tag{26.76}$$

Donc $D_n(t)$ a une limite bien définie pour $t \to 0$ et est alors une fonction continue sur le compact $[-\pi, \pi]$.

La norme de l_n (début) Nous avons prouvé que $\|l_n\| \leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi}|D_n(t)|dt$. Nous allons à présent prouver que cela est effectivement la norme de l_n. Pour $\epsilon > 0$ nous considérons la fonction

$$\begin{aligned} f_\epsilon\colon \mathbb{R} &\to \mathbb{C} \\ x &\mapsto \frac{D_n(x)}{|D_n(x)| + \epsilon}. \end{aligned} \tag{26.77}$$

C'est une fonction continue et 2π-périodique satisfaisant $\|f_\epsilon\| \leqslant 1$ parce que le dénominateur est toujours plus grand que le numérateur. Nous nous proposons de calculer

$$l_n(f_\epsilon) = \sum_{k=-n}^{n} \frac{1}{2\pi}\int_{-\pi}^{\pi} f_\epsilon(t)e^{-ikt}dt. \tag{26.78}$$

Vu que $f_\epsilon(t)e^{-ikt}$ vaut en norme $|f_\epsilon(t)|$ qui est une fonction intégrable (ne dépendant pas de k) sur $[-\pi,\pi]$, le théorème de la convergence dominée 13.134 nous permet de permuter la somme et l'intégrale :

$$l_n(f_\epsilon)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{D_n(t)}{|D_n(t)|+\epsilon}\underbrace{\sum_{k=-n}^{n}e^{-ikt}}_{=D_n(t)}dt=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\big|D_n(t)\big|^2}{|D_n(t)|+\epsilon}dt. \tag{26.79}$$

Nous avons donc

$$\lim_{\epsilon\to 0}l_n(f_\epsilon)=\frac{1}{2\pi}\int_{-\pi}^{\pi}|D_n(t)|dt. \tag{26.80}$$

Mais vue l'inégalité (26.74) nous avons

$$\|l_n\|=\frac{1}{2\pi}\int_{-\pi}^{\pi}|D_n(t)|dt. \tag{26.81}$$

Notre tâche est maintenant de donner une valeur à cette intégrale.

Norme de l_n tend vers ∞ Nous utilisons la divergence de l'intégrale du sinus cardinal (14.567). D'abord nous écrivons

$$\|l_n\|=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\big|\sin\big(\frac{2n+1}{2}t\big)\big|}{\big|\sin(t/2)\big|}dt, \tag{26.82}$$

ensuite nous nous souvenons que $|\sin(x)|\leqslant|x|$ pour tout x, ce qui nous permet de changer le dénominateur :

$$\|l_n\|\geqslant\frac{2}{\pi}\int_{0}^{\pi}\frac{\big|\sin\big(\frac{2n+1}{2}t\big)\big|}{|t|}dt \tag{26.83}$$

Nous y effectuons le changement de variable $u=\frac{2n+1}{2}t$ qui donne

$$\|l_n\|\geqslant\frac{2}{\pi}\int_{0}^{(n+\frac{1}{2})\pi}\frac{\big|\sin(u)\big|}{|u|}. \tag{26.84}$$

Utilisant cette histoire de sinus cardinal nous avons

$$\lim_{n\to\infty}\|l_n\|=\infty. \tag{26.85}$$

La conclusion L'espace $\big(C^0_{2\pi},\|.\|_\infty\big)$ est complet [4], donc le théorème de Banach-Steinhaus 24.8 s'applique. Par rapport aux notations de l'énoncé de Banch-Steinhaus, nous posons

$$E=\big(C^0_{2\pi},\|.\|_\infty\big) \tag{26.86a}$$
$$F=\mathbb{R} \tag{26.86b}$$
$$H=\{l_n\}_{n\in\mathbb{N}}. \tag{26.86c}$$

Vu que la suite $(\|l_n\|)$ n'est pas bornée, il existe $f\in C^0_{2\pi}$ tel que

$$\sup_n\|l_n(f)\|=\infty. \tag{26.87}$$

Pour cette fonction nous avons

$$\sup_{n\geqslant 0}S_n(f)(0)=\infty, \tag{26.88}$$

et donc la série de Fourier de f ne converge pas en zéro.

□

4. Parce qu'une limite uniforme de fonctions continues est continue, théorème 12.297.

26.3.2 Inégalité isopérimétrique

Définition 26.19.
*Une **courbe de Jordan** dans le plan est une application $\gamma\colon S^1 \to \mathbb{R}^2$ qui est continue et injective.*

Une telle courbe peut évidemment être vue comme une application $\gamma\colon [0,2\pi] \to \mathbb{R}^2$ telle que $\gamma(0) = \gamma(2\pi)$. En particulier il n'est jamais mauvais de se rappeler qu'on peut choisir une paramétrisation normale par la proposition 19.43.

Théorème 26.20 (Théorème de Jordan[298])**.**
Si γ est une courbe de Jordan, alors l'ensemble $\mathbb{R}^2\backslash\gamma$ a exactement deux composantes connexes. L'une est bornée, l'autre non. Les deux ont γ comme frontière.

Le théorème suivant dit que parmi les courbes C^1, le cercle a la plus grande surface possible à périmètre donné.

Théorème 26.21 (Inégalité isopérimétrique[2])**.**
Soit $f\colon S^1 \to \mathbb{C}$ une courbe de Jordan de classe C^1. Nous notons L sa longueur et S l'aire contenue de la surface délimitée[5] *par f. Alors*

*(1) Nous avons l'**inégalité isopérimétrique** : $L^2 \geqslant 4\pi S$.*

(2) Nous avons l'égalité $L^2 = 4\pi S$ si et seulement si la courbe donnée par f est un cercle.

Démonstration. Nous commençons par considérer un chemin dont la longueur est 2π et nous en considérons sa paramétrisation normale. Nous allons exprimer l'aire S en utilisant le théorème de Green, et plus particulièrement la formule de surface (14.187).

Si $f(s) = x(s) + iy(s)$, nous devons intégrer $y'x - x'y$, qui n'est rien d'autre que la partie imaginaire de $f'(s)\overline{f(s)}$. Donc

$$S = \frac{1}{2}\operatorname{Im}\int_0^{2\pi} f'(s)\overline{f(s)}ds \tag{26.89}$$

Nous considérons les coefficients de Fourier de f donnés par la formule (26.50) :

$$c_n(f) = \frac{1}{2\pi}f(s)e^{-ins}. \tag{26.90}$$

Ceux de f' (qui est aussi continue sur le compact S^1 et donc tout autant L^2) sont donnés par

$$c_n(f') = inc_n(f). \tag{26.91}$$

D'autre part en vertu du théorème 19.9, la longueur de γ s'exprime en terme de l'intégrale de la norme de sa dérivée :

$$2\pi = L = \int_0^{2\pi} |f'(s)|ds = \int_0^{2\pi} |f'(s)|^2 ds \tag{26.92}$$

parce que nous avons choisit une paramétrisation normale qui vérifie automatiquement $|f'(s)| = 1$ pour tout s. L'identité de Parseval sous sa forme (23.106) appliquée à f' nous enseigne que

$$L = 2\pi = \int_0^{2\pi} |f'(s)|^2 ds = 2\pi\langle f', f'\rangle = 2\pi\sum_{n=-\infty}^{\infty} |c_n(f')|^2 = 2\pi\sum_n n^2|c_n(f)|^2. \tag{26.93}$$

Par ailleurs le système trigonométrique étant une base hilbertienne, et les fonctions f et f' étant dans $L^2\big([0,2\pi]\big)$ (parce que continues sur un compact), elles sont égales à leurs séries de Fourier

5. C'est la partie connexe bornée de $\mathbb{C}\backslash\gamma$ dont l'existence est donnée par le théorème de Jordan 26.20.

(au sens L^2), c'est à dire que nous avons l'égalité (24.192). Nous avons alors

$$\langle f', f\rangle_{L^2} = \langle \sum_{n\in\mathbb{Z}} c_n(f')e_n, \sum_{m\in\mathbb{Z}} c_m(f)e_m\rangle \tag{26.94a}$$
$$= \sum_m \sum_n c_n(f')\overline{c_m(f)}\underbrace{\langle e_n, e_m\rangle}_{\delta_{m,n}} \tag{26.94b}$$
$$= \sum_{n\in\mathbb{Z}} c_n(f')\overline{c_n(f)} \tag{26.94c}$$
$$= \sum_n in|c_n(f)|^2 \tag{26.94d}$$

où nous avons utilisé la continuité du produit scalaire pour sortir les sommes. Avec cela nous pouvons exprimer l'aire (26.89) en termes de coefficients de Fourier :

$$S = \frac{1}{2}\operatorname{Im} 2\pi\langle f', f\rangle = \pi\sum_{n\in\mathbb{Z}} n|c_n(f)|^2. \tag{26.95}$$

En utilisant les expressions (26.93) et (26.95) pour L et S, et en écrivant $L = 2\pi L$, nous avons

$$L^2 - 4\pi S = 4\pi^2\sum_{n\in\mathbb{Z}}(n^2 - n)|c_n(f)|^2 \geqslant 0. \tag{26.96}$$

Cela prouve l'inégalité demandée dans le cas où $L = 2\pi$.

Si γ n'est pas de longueur 2π mais L, alors nous considérons le chemin $\sigma(t) = \frac{2\pi\gamma(t)}{L}$. Sa longueur est 2π et son aire, au vu de la formule de Green (26.89), son aire est $4\pi^2\frac{S}{L^2}$. L'inégalité isopérimétrique appliquée au chemin σ donne alors $L^2 \geqslant 4\pi S$.

Le cas d'égalité s'obtient uniquement si $c_n = 0$ pour tout n différent de 0 ou 1. Dans ce cas nous avons

$$f(s) = c_0(f) + c_1(f)e^{is}, \tag{26.97}$$

qui est un cercle de centre $c_0(f)$ et de rayon $|c_1(f)|$. □

26.3.3 À propos des coefficients

Nous considérons l'application

$$\begin{aligned} c\colon \big(L^1_{2\pi}, \|.\|_1\big) &\to \big(C_0, \|.\|_\infty\big) \\ f &\mapsto (c_n(f))_{n\in\mathbb{Z}} \end{aligned} \tag{26.98}$$

qui à une fonction 2π-périodique fait correspondre la suite (bornée) de ses coefficients de Fourier. Nous rappelons la définition

$$c_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-int}. \tag{26.99}$$

Nous allons montrer que cette application est linéaire, continue, injective et non surjective. Pour la continuité, par la linéarité il suffit de la montrer en 0. Nous devons donc montrer que si nous avons une suite de fonctions f_k qui tend vers 0 au sens L^1, alors $c(f_k) \to 0$ au sens de la norme $\|.\|_\infty$ sur l'ensemble des suites.

Si nous posons $r_k = \int_0^{2\pi}|f_k(t)|dt$, alors $r_k = \|f_k\|_1$ et nous avons $r_k \to 0$. Mais par définition

$$|c_n(f_k)| \leqslant r_k, \tag{26.100}$$

et donc $\|c(f_k)\|_\infty \leqslant r_k$. L'application c est donc continue. L'injectivité est donnée par le corollaire 26.16.

Si nous supposons que l'application c est continue, alors le théorème d'isomorphisme de Banach (24.2) nous dit que cela devrait être un homéomorphisme, c'est à dire que c^{-1} serait également continue. Nous allons montrer qu'il n'en est rien.

Nous considérons la suite de suite

$$(c_n)_k = \begin{cases} 1 & \text{si } k < n \\ 0 & \text{sinon.} \end{cases} \tag{26.101}$$

Ici $(c_n)_k$ est le terme numéro k de la suite n. Par injectivité de l'application qui à une fonction fait correspondre la suite de ses coefficients de Fourier, la seule fonction qui possède ces coefficients est

$$f_n(t) = \sum_{k\in\mathbb{N}} c_{n,k} e^{ikt}. \tag{26.102}$$

Étant donné que $\|f_n\|_1 = n$, la suite $(\|f_n\|_1)$ n'est pas bornée alors que a suite de suites (26.101) est bornée dans l'ensemble des suites parce que $\|c_n\|_\infty = 1$.

Chapitre 27

Transformation de Fourier

Ici nous utilisons la convention de la transformée de Fourier de wikipedia, c'est à dire

$$\hat{f}(\xi) = \int_{\mathbb{R}} e^{-i\xi x} f(x) dx \tag{27.1a}$$

$$f(x) = 2\pi \int_{\mathbb{R}} e^{i\xi x} \hat{f}(\xi) d\xi. \tag{27.1b}$$

Nous allons par ailleurs utiliser indifféremment les notations $\mathcal{F}(f)$ ou $\hat{f}$ pour la transformée de Fourier de f. La notation $\mathcal{F}$ est pratique pour les transformées de looooooongues expressions ainsi que pour parler de l'application «transformée de Fourier» d'un espace de fonction vers un autre.

27.1.
Nous verrons dans le théorème 27.25 que la Transformée de Fourier n'est pas une isométrie de L^2. Pour avoir une isométrie, il aurait fallu choisir des coefficients moins simples dans (27.1a).

27.1 Transformée de Fourier dans $L^1(\mathbb{R}^d)$

Définition 27.2.
*Si $f \in L^1(\mathbb{R}^d)$ alors nous définissons sa **transformée de Fourier** est la fonction donnée par*

$$\mathcal{F}(f)(\xi) = \hat{f}(\xi) = \int_{\mathbb{R}} e^{-i\xi x} f(x) dx \tag{27.2}$$

Lemme 27.3.
Si $f \in L^1(\mathbb{R}^d)$ et si $g(x) = f(\lambda x)$ alors

$$\hat{g}(\xi) = \lambda^{-d} \hat{f}(\xi/\lambda). \tag{27.3}$$

Démonstration. Il s'agit de faire le changement de variable $y = \lambda x$ dans l'intégrale

$$\hat{g}(\xi) = \int_{\mathbb{R}^d} f(\lambda x) e^{-i\xi x} dx. \tag{27.4}$$

Dans le changement de variables, vient le coefficient $dx = \lambda^{-d} dy$. □

Proposition 27.4.
La transformée de Fourier est un morphisme vis-à-vis de la convolution sur $L^1(\mathbb{R}^n)$:

$$\widehat{f * g} = \hat{f}\hat{g}. \tag{27.5}$$

Démonstration. Nous devons étudier l'intégrale

$$\widehat{f * g}(\xi) = \int_{\mathbb{R}} \left[\int_{\mathbb{R}} f(y) g(t-y) \right] e^{-it\xi} dt. \tag{27.6}$$

Ici nous avons choisit des représentants f et g dans les classes de L^1. Montrons que f est borélienne. D'abord $f(x) = f_+(x) - f_-(x)$ où f_+ et f_- sont des fonctions positives. Afin d'alléger les notations nous supposons un instant que f est positive et nous posons

$$f_n(x) = \sum_{k=1}^{2^n} \frac{k}{n} \mathbb{1}_{f(x)\in[\frac{k}{n},\frac{k+1}{n}[}. \tag{27.7}$$

Le fait que f soit dans L^1 implique que chacune des fonctions f_n est borélienne [1] et donc que f l'est aussi en tant que limite ponctuelle de fonctions boréliennes [2].

Nous allons appliquer le théorème de Fubini 13.229 à la fonction

$$\phi(x, y) = f(x)g(y)e^{-i\xi(x+y)} \tag{27.8}$$

qui est borélienne en tant que produit et composé de fonctions boréliennes. Nous avons

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x)e^{-i\xi x}||g(y)e^{-i\xi y}|dy \right) dx = \int_{\mathbb{R}} \left(|f(x)| \int_{\mathbb{R}} |g(y)|dy \right) dx \tag{27.9a}$$

$$= \int_{\mathbb{R}} |f(x)|\|g\|_1 \tag{27.9b}$$

$$= \|f\|_1\|g\|_1 < \infty. \tag{27.9c}$$

Le théorème est donc applicable. D'abord nous avons :

$$\hat{f}(\xi)\hat{g}(\xi) = \left(\int_{\mathbb{R}} f(x)e^{-i\xi x}dx \right) \left(\int_{\mathbb{R}} g(y)e^{-i\xi y}dy \right) \tag{27.10a}$$

$$= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x)g(y)e^{-i\xi(x+y)}dy \right) dx \tag{27.10b}$$

$$= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x)g(t-x)e^{-i\xi t} \right) dx. \tag{27.10c}$$

Jusqu'ici nous n'avons pas utilisé Fubini. Nous avons seulement introduit le nombre $\int_{\mathbb{R}} g(y)e^{-i\xi y}dy$ dans l'intégrale par rapport à x et effectué le changement de variables $y \mapsto t = x + y$. Maintenant nous appliquons le théorème de Fubini pour inverser l'ordre des intégrales :

$$\hat{f}(\xi)\hat{g}(\xi) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x)g(t-x)e^{-it\xi}dx \right) dy \tag{27.11a}$$

$$= \int_{\mathbb{R}} e^{-it\xi} \left(\int_{\mathbb{R}} f(x)g(t-x)dx \right) dt \tag{27.11b}$$

$$= \int_{\mathbb{R}} e^{-it\xi}(f * g)(t)dt \tag{27.11c}$$

$$= \widehat{f * g}(\xi). \tag{27.11d}$$

□

Proposition 27.5.
Soit une fonction $f \in L^1(\mathbb{R}^d)$. Alors sa transformée de Fourier est continue.

Démonstration. Nous considérons une fonction f définie sur $\mathbb{R}^d$ et à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$. Sa transformée de Fourier est donnée par

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} e^{-i\xi x} f(x)dx. \tag{27.12}$$

1. Ceci demanderait plus de justification. Dites moi si vous savez comment justifier que les f_n soient boréliennes.
2. Le fait que f soit borélienne est une conséquence du théorème 24.84.

Pour montrer que cette fonction $\hat{f}$ est continue en ξ_0 nous considérons une suite $(\xi_n) \to \xi_0$ et nous voulons montrer que $\hat{f}(\xi_n) \to \hat{f}(\xi_0)$. Pour cela nous considérons les fonctions

$$g_n(x) = e^{-i\xi_n x} f(x) \tag{27.13}$$

qui convergent simplement vers $g(x) = e^{-i\xi x} f(x)$. Étant donné que

$$|g_n(x)| < |f(x)|, \tag{27.14}$$

le théorème de la convergence dominée donne alors

$$\lim_{n\to\infty} \int g_n(x) = \int \lim_{n\to\infty} g_n(x), \tag{27.15}$$

c'est à dire $\lim_{n\to\infty} \hat{f}(\xi_n) = \hat{f}(\xi)$. La fonction $\hat{f}$ est donc continue. □

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

Démonstration. Cela est une simple vérification :

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-ix\xi} dx, \tag{27.16}$$

nous avons, pour tout ξ,

$$|\hat{f}(\xi)| \leqslant \int_{\mathbb{R}} |f(x)| dx, \tag{27.17}$$

ce qui signifie exactement $\|\hat{f}\|_\infty \leqslant \|f\|_1$. □

Lemme 27.7 (Lemme de Riemann-Lebesgue[299])**.**
Si f est une fonction $L^1(\mathbb{R})$ alors $\lim_{\xi\to\pm\infty} \hat{f}(\xi) = 0$.

Démonstration. Nous commençons par prouver le résultat dans le cas d'une fonction g en escalier, et plus précisément par une fonction caractéristique d'un compact $K = [a, b]$. Au niveau de la transformée de Fourier nous avons

$$\hat{\mathbb{1}}_K(\xi) = \int_a^b e^{-i\xi x} dx = -\frac{1}{i\xi}(e^{-ib\xi} - e^{-ia\xi}). \tag{27.18}$$

Par conséquent

$$|\hat{\mathbb{1}}_K(\xi)| \leqslant \frac{2}{|\xi|}. \tag{27.19}$$

Plus généralement si $g = \sum_{i=1}^N c_i \mathbb{1}_{K_i}$, alors

$$|\hat{g}(\xi)| \leqslant \frac{2}{|\xi|} \sum_{i=1}^N |c_i|, \tag{27.20}$$

et donc nous avons effectivement $\lim_{\xi\to\pm\infty} |\hat{g}(\xi)| = 0$.

Nous passons maintenant au cas général $f \in L^1(\mathbb{R})$. Étant donné que les fonctions L^1 en escalier sont denses dans L^1, nous considérons une fonction $g \in L^1(\mathbb{R})$ en escalier telle que $\|f - g\|_1 < \epsilon$. Nous avons donc

$$\|\hat{f} - \hat{g}\|_\infty \leqslant \|f - g\|_1 < \epsilon. \tag{27.21}$$

Donc

$$\|\hat{f}(\xi)\| \leqslant \|\hat{f}(\xi) - \hat{g}(\xi)\|_|\hat{g}(\xi)|. \tag{27.22}$$

Le premier terme est plus petit que ϵ. Il nous reste à voir que

$$\lim_{\xi\to\infty} |\hat{g}(\xi)| = 0, \tag{27.23}$$

mais cela est le résultat de la première partie de la preuve. □

Corollaire 27.8.
La transformée de Fourier d'une fonction $L^1(\mathbb{R})$ est bornée.

Démonstration. Par le corollaire 27.5, la transformée de Fourier d'une fonction L^1 est continue. Le lemme de Riemann-Lebesgue 27.7 impliquant qu'elle tend vers zéro en $\pm\infty$, elle doit être bornée. □

27.1.1 Formule sommatoire de Poisson

Proposition 27.9 (Formule sommatoire de Poisson)**.**
Soit $f\colon \mathbb{R} \to \mathbb{C}$ une fonction continue et $L^1(\mathbb{R})$. Nous supposons que

(1) il existe $M > 0$ et $\alpha > 1$ tels que

$$|f(x)| \leqslant \frac{M}{(1+|x|)^\alpha}, \tag{27.24}$$

(2) $\sum_{n=-\infty}^{\infty} |\hat{f}(2\pi n)| < \infty$.

Alors nous avons

$$\sum_{n=-\infty}^{\infty} f(n) = \sum_{n=-\infty}^{\infty} \hat{f}(2\pi n). \tag{27.25}$$

Démonstration. **Convergence normale** Nous commençons par montrer qu'il y a convergence normale sur tout compact séparément des séries sur les $n \geqslant 0$ et sur les $n < 0$.
Soit K un compact de $\mathbb{R}$ contenu dans $[-A, A]$ et $n \in \mathbb{Z}$ tel que $|n| \geqslant 2A$. Pour $x \in K$ nous avons

$$|x+n| \geqslant |n| - |x| \geqslant |n| - A \geqslant \frac{|n|}{2}. \tag{27.26}$$

Du coup nous avons un $\alpha > 1$ tel que

$$|f(x+n)| \leqslant \frac{M}{\big(1+|x+n|\big)^\alpha} \leqslant \frac{M}{\Big(1+\frac{|n|}{2}\Big)^\alpha}. \tag{27.27}$$

Lorsque n est grand, cela a le comportement de $M/|n|^\alpha$ et donc la série

$$\sum_{n=0}^{\infty} f(x+n) \tag{27.28}$$

est une série convergent normalement. Les deux séries (usuelles)

$$a_- = \sum_{n\leqslant 0} f(x+n) \tag{27.29a}$$

$$a_- = \sum_{n>0} f(x+n) \tag{27.29b}$$

convergent normalement.

Convergence commutative Au sens de la définition 9.179 nous avons

$$\sum_{n\in\mathbb{Z}} f(x+n) = a_+ + a_-. \tag{27.30}$$

En effet si nous prenons $J_0' \subset \mathbb{N}$ fini tel que $|\sum_{\mathbb{N}\setminus J_0} f(x+n) - a_+| \leqslant \epsilon$ et $J_1' \in -\mathbb{N}$ tel que $|\sum_{n\in-\mathbb{N}\setminus J_1'} f(x+n)| - a_- < \epsilon$, et si nous posons $J_0 = J_0' \cup J_1'$ alors si K est un ensemble fini de $\mathbb{Z}$ contenant J_0 nous avons

$$|\sum_{n\in K} f(n+x) - (a_+ + a_-)| \leqslant |\sum_{n\in K^+} f(n+x) - a_+| + |\sum_{n\in K^-} f(n+x) - a_-| \leqslant 2\epsilon \tag{27.31}$$

où K^+ sont les éléments positifs de K et K^- sont les *strictement* négatifs. Maintenant que la famille $\{f(n+x)\}_{n\in\mathbb{Z}}$ est une famille sommable, nous savons qu'elle est commutativement sommable et que la proposition 9.184 nous permet de sommer dans l'ordre que l'on veut. Nous pouvons donc écrire sans ambigüité l'expression $\sum_{n\in\mathbb{Z}} f(x+n)$ ou $\sum_{n=-\infty}^{\infty} f(x+n)$.

re-convergence normale Nous posons donc sans complexes la série

$$F(x) = \sum_{n\in\mathbb{Z}} f(x+n) \tag{27.32}$$

qui converge tant commutativement que normalement. Notons que nous pouvons maintenant dire que la série sur $\mathbb{Z}$ converge normalement ; pas seulement les deux séries séparément.

Continuité, périodicité Étant donné que chacune des fonctions $f(x+n)$ est continue, la convergence normale nous assure que F est continue.

De plus F est périodique parce que

$$F(x+1) = \sum_{n=-\infty}^{\infty} f(x+1+n) = \sum_{p=-\infty}^{\infty} f(x+p) \tag{27.33}$$

où nous avons posé $p = 1+n$.

Coefficients de Fourier En vertu de la définition (24.184) et de la périodicité de F,

$$\begin{aligned}
c_n(F) &= \int_{-1/2}^{1/2} F(t)e^{-2\pi int}dt && (27.34a)\\
&= \int_0^1 F(t)e^{-2\pi int}dt && (27.34b)\\
&= \int_0^1 \sum_{n\in\mathbb{Z}} f(t+n)e^{-2i\pi nt}dt && (27.34c)\\
&= \sum_{n\in\mathbb{Z}} \int_n^{n+1} f(u)e^{-2\pi i(u-n)t}du && (27.34d)\\
&= \int_{-\infty}^{\infty} f(u)e^{-2\pi inu}du && (27.34e)\\
&= \hat{f}(2\pi n). && (27.34f)
\end{aligned}$$

où nous avons effectué le changement de variables $u = t+n$, et permuté l'intégrale et la somme en vertu du fait que la somme converge normalement.

Conclusion Étant donné l'hypothèse $\sum_{n\in\mathbb{Z}} |\hat{f}(n)| < \infty$ la proposition 26.14 nous dit que

$$F(x) = \sum_{n\in\mathbb{Z}} c_n(F)e^{2\pi inx}, \tag{27.35}$$

c'est à dire que

$$\sum_{n-\infty}^{\infty} f(x+n) = \sum_{n=-\infty}^{\infty} \hat{f}(2\pi n)e^{2\pi inx}. \tag{27.36}$$

En écrivant cette égalité en $x = 0$ nous trouvons le résultat :

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \hat{f}(2\pi n). \tag{27.37}$$

□

Exemple 27.10

La formule sommatoire de Poisson peut être utilisée pour calculer des sommes dans l'espace de Fourier plutôt que dans l'espace direct. Nous allons montrer dans cet exemple l'égalité

$$\sum_{n=-\infty}^{\infty} e^{-\alpha n^2} = \sum_{n=-\infty}^{\infty} \sqrt{\frac{\pi}{\alpha}} e^{-\pi^2 n^2/\alpha}. \tag{27.38}$$

Si α est grand, alors la somme de gauche est plus rapide, tandis que si α est petit, c'est le contraire.

Nous appliquons la formule sommatoire de Poisson à la fonction

$$f(x) = e^{-\alpha x^2}. \tag{27.39}$$

Nous avons

$$\hat{f}(x) = \int_{\mathbb{R}} e^{-\alpha t^2 - ixt} dt \tag{27.40a}$$

$$= e^{-x^2/4\alpha} \int_{\mathbb{R}} e^{-(\sqrt{\alpha}t + \frac{ix}{2\sqrt{\alpha}})^2} \tag{27.40b}$$

$$= e^{-x^2/4\alpha} \frac{1}{\sqrt{\alpha}} \int_{\mathbb{R}+\frac{ix}{2\sqrt{\alpha}}} e^{-u^2} du. \tag{27.40c}$$

Pour traiter cette intégrale nous utilisons la proposition 25.6 en considérant le chemin rectangulaire fermé qui joint les points $-R$, R, $R+ai$, $-R+ai$ et $f(z) = e^{-z^2}$. Calculons l'intégrale sur les deux côtés verticaux. Nous posons

$$\gamma_R(t) = R + tia \tag{27.41}$$

avec $t\colon 0 \to 1$. Nous avons

$$\int_{\gamma_R} f = \int_0^1 f\big(\gamma_R(t)\big) \|\gamma_R'(t)\| dt \tag{27.42a}$$

$$= ae^{-R^2} \int_0^1 e^{-2tRia + at^2} dt, \tag{27.42b}$$

donc en module nous avons

$$|\int_{\gamma_R} f| \leqslant ae^{-R^2} \int_0^1 e^{at^2} dt \leqslant Me^{-R^2}, \tag{27.43}$$

où M est une constante ne dépendant pas de R. Lorsque $R \to \infty$, la contribution des chemins verticaux s'annule et nous trouvons que

$$\int_{\mathbb{R}+ai} e^{-u^2} du = \int_{\mathbb{R}} e^{-u^2} du, \tag{27.44}$$

que nous pouvons utiliser pour continuer le calcul (27.40). Nous avons

$$\hat{f}(x) = \frac{e^{-x^2/4\alpha}}{\sqrt{\alpha}} \int_R e^{-u^2} du = \sqrt{\frac{\pi}{\alpha}} e^{-x^2/4\alpha} \tag{27.45}$$

où nous avons utilisé la formule (13.702). Par conséquent ce qui rentre dans la formule sommatoire de Poisson est

$$\hat{f}(2\pi n) = \sqrt{\frac{\pi}{\alpha}} e^{-\pi^2 n^2/\alpha}. \tag{27.46}$$

△

27.2 Transformée de Fourier dans l'espace de Schwartz

La définition de la transformée de Fourier de $\varphi \in \mathscr{S}(\mathbb{R}^d)$ est

$$\hat{\varphi}(\xi) = \int_{\mathbb{R}^n} \varphi(x) e^{-ix \cdot \xi} dx. \tag{27.47}$$

Si α est un multiindice de taille m, nous notons

$$(M_\alpha f)(x) = x_{\alpha_1} \dots x_{\alpha_m} f(x). \tag{27.48}$$

Lemme 27.11 (Lemme de transfert)**.**
Si $\varphi \in \mathscr{S}(\mathbb{R}^d)$ et si α est un multiindice, alors

$$\partial^\alpha \hat{\varphi} = (-i)^{|\alpha|}\widehat{M_\alpha \varphi}. \tag{27.49}$$

et

$$\widehat{\partial^\alpha \varphi}(\xi) = (-i)^{|\alpha|}\xi^\alpha \hat{\varphi}(\xi). \tag{27.50}$$

Démonstration. Nous considérons la fonction $h(x,\xi) = \varphi(x)e^{-ix\cdot\xi}$ dont la dérivée par rapport à ξ_i est donnée par $-i(M_i\varphi)(x)e^{x\cdot\xi}$. Cette fonction est majorée en norme par

$$G(x) = M_i\varphi(x), \tag{27.51}$$

qui est encore une fonction à décroissance rapide et donc parfaitement intégrable sur $\mathbb{R}^d$. Le théorème 14.65 nous dit donc que la dérivée de $\hat{\varphi}$ par rapport à ξ_i existe et vaut

$$\frac{\partial\hat{\varphi}}{\partial\xi_i}(\xi) = -i\int_{\mathbb{R}^n} x_i\varphi(x)e^{-i\xi\cdot x} = -i\widehat{M_i\varphi}(\xi). \tag{27.52}$$

En appliquant ce résultat en chaîne, nous trouvons la première formule annoncée.

Nous passons à la seconde formule annoncée. Étant donné que $\varphi \in \mathscr{S}$, ses dérivées le sont aussi et par conséquent, il n'y a pas de problèmes pour écrire

$$\widehat{\partial_{x_k}\varphi}(\xi) = \int_{\mathbb{R}^d} \frac{\partial\varphi}{\partial x_k}(x)e^{-ix\cdot\xi}dx. \tag{27.53}$$

Étant donné que

$$\frac{\partial}{\partial x_k}\left(\varphi(x)e^{-ix\cdot\xi}\right) = \frac{\partial\varphi}{\partial x_k}(x)e^{-ix\cdot\xi} - i\xi_k\varphi(x)e^{-ix\cdot\xi}, \tag{27.54}$$

notre tâche sera de prouver que

$$\int_{\mathbb{R}^d}\frac{\partial}{\partial x_k}\left(\varphi(x)e^{-ix\cdot\xi}\right)dx = 0. \tag{27.55}$$

Autrement dit, nous voulons montrer que le terme au bord d'une intégration par partie s'annule. D'abord le fait que φ soit à décroissance rapide nous assure que l'intégrale (27.55) converge. Pour chaque ξ, la fonction

$$f(x,\xi) = \frac{\partial}{\partial x_k}\left(\varphi(x)e^{-ix\cdot\xi}\right) \tag{27.56}$$

est intégrable par rapport à x. De plus, f est dans $\mathscr{S}(\mathbb{R})$ pour chacune de ses variables (les autres étant fixées). Le théorème de Fubini 13.230 nous permet alors de décomposer l'intégrale en

$$\int_{\mathbb{R}^d} f(x,\xi)dx = \int_{\mathbb{R}}\dots\int_{\mathbb{R}} f(x_1,\dots,x_d)dx_1\dots dx_d. \tag{27.57}$$

De plus nous pouvons intégrer dans l'ordre de notre choix et nous choisissons évidemment d'intégrer d'abord par rapport à x_k. Étudions donc l'intégrale

$$\int_{\mathbb{R}}\frac{\partial}{\partial x}\left(\varphi(x)e^{-ix\xi}\right)dx = \lim_{A\to\infty}\int_{-A}^{A}\frac{\partial}{\partial x}\left(\varphi(x)e^{-ix\xi}\right)dx \tag{27.58}$$

dans laquelle nous avons un peu allégé les notations. Une primitive de ce qui est intégré est toute trouvée : c'est $\varphi(x)e^{-ix\xi}$, et nous pouvons utiliser le théorème fondamental du calcul intégral pour écrire que

$$\int_{-A}^{A}\left(\varphi(x)e^{-ix\xi}\right)'dx = \left[\varphi(x)e^{-ix\xi}\right]_{x=-A}^{x=A}. \tag{27.59}$$

Vu que φ est dans $\mathscr{S}$, la limite $A\to\infty$ donne zéro.

En substituant maintenant (27.54) dans (27.53) et en tenant compte du terme que nous venons de montrer s'annuler, nous avons

$$\widehat{\partial_k\varphi}(\xi) = -i\xi_k \int_{\mathbb{R}^d} \varphi(x)e^{-ix\cdot\xi} = -i\xi_k\hat{\varphi}(\xi). \tag{27.60}$$

En recommençant la procédure $|\alpha|$ fois nous trouvons la seconde formule annoncée. □

Proposition 27.12 ([179])**.**
L'espace de Schwartz est stable par transformée de Fourier. De plus l'application

$$\mathcal{F}\colon \mathscr{S}(\mathbb{R}^d) \to \mathscr{S}(\mathbb{R}^d) \tag{27.61}$$

est une bijection linéaire et continue.

Démonstration. La linéarité découle de celle de l'intégrale. La difficulté est de prouver que pour $\varphi \in \mathscr{S}(\mathbb{R}^d)$ nous avons bien que $\hat{\varphi} \in \mathscr{S}(\mathbb{R}^d)$ et que cette association est continue[3].

Stabilité Nous devons prouver que pour tout multiindices α et β, nous avons $p_{\alpha,\beta}(\hat{\varphi}) < \infty$. Nous avons

$$\xi^\beta\partial^\alpha\hat{\varphi}(\xi) = \xi^\beta(-i)^{|\alpha|}\widehat{M_\alpha\varphi}(\xi) = (-i)^{|\alpha|+|\beta|}\widehat{\partial^\beta M_\alpha\varphi}(\xi). \tag{27.62}$$

Ensuite nous nous souvenons que $\|\hat{f}\|_\infty \leqslant \|f\|_1$ parce que

$$|\hat{f}(\xi)| \leqslant \int_{\mathbb{R}^d} \left|f(x)e^{-ix\cdot\xi}\right| = \int_{\mathbb{R}^d} |f(x)|dx = \|f\|_1. \tag{27.63}$$

Donc

$$p_{\alpha,\beta}(\hat{\varphi}) = \|\widehat{\partial^\beta M_\alpha\varphi}\|_\infty \leqslant \|\partial^\beta M_\alpha\varphi\|_1. \tag{27.64}$$

Du fait que φ soit dans $\mathscr{S}$, la dernière expression est finie. Cela prouve déjà que

$$\mathcal{F}\big(\mathscr{S}(\mathbb{R}^d)\big) \subset \mathscr{S}(\mathbb{R}^d). \tag{27.65}$$

Continuité Nous supposons avoir une suite $\varphi_n \xrightarrow{\mathscr{S}} \varphi$, et nous devons prouver que $\hat{\varphi}_n \xrightarrow{\mathscr{S}} \hat{\varphi}$. Pour alléger les notations, nous posons $f_n = \varphi_n - \varphi$. Nous avons

$$\|\hat{f}\|_{\alpha,\beta} = \|\xi^\beta\partial^\alpha\hat{f}\|_\infty \tag{27.66a}$$

$$= \|\widehat{\partial^\beta M_\alpha f}\|_\infty \text{ lemme 27.11.} \tag{27.66b}$$

$$\leqslant \|\partial^\beta M_\alpha f\|_1 \tag{27.66c}$$

La convergence $f_n \xrightarrow{\mathscr{S}} 0$ nous dit ente autres que $\partial^\beta M_\alpha f_n \xrightarrow{\mathscr{S}} 0$; en particulier la proposition 24.83 nous dit que $\partial^\beta M_\alpha f_n \xrightarrow{L^1} 0$, ce qui signifie, par les majorations (27.66) que

$$\|\hat{f}_n\|_{\alpha,\beta} \leqslant \|\partial^\beta M_\alpha f_n\|_1 \to 0, \tag{27.67}$$

ce qui prouve la continuité de transformée de Fourier dans $\mathscr{S}(\mathbb{R}^d)$.

Bijection Une preuve peut être trouvée dans [300].

□

Proposition 27.13 ([1])**.**
Soit $\varphi \in \mathscr{S}(\mathbb{R}^n \times \mathbb{R}^m)$ et la transformée de Fourier partielle

$$\tilde{\varphi}(x,k) = \int_{\mathbb{R}^m} e^{-iky}\varphi(x,y)dy. \tag{27.68}$$

Alors $\tilde{\varphi} \in \mathscr{S}(\mathbb{R}^n \times \mathbb{R}^m)$.

3. Pour rappel, en dimension infinie, il n'est pas garanti qu'une application linéaire soit continue.

Démonstration. Il s'agit de reprendre les étapes de la partie correspondante de la preuve de la proposition 27.12. Soient des multiindices α, α', β et β' où α et β se réfèrent à la variable x tandis que α' et β' se réfèrent à la variable k.

Vu que la multiplication par $k^{\beta'}$ commute avec ∂^α nous avons

$$x^\beta k^{\beta'}\partial^\alpha\partial^{\alpha'}\tilde{\varphi}(x,k) = x^\beta k^{\beta'}\partial^\alpha(-i)^{|\alpha'|}\widetilde{M_{\alpha'}\varphi}(x,k) = (-i)^{|\alpha'|+|\beta'|}x^\beta\partial^\alpha\partial^{\beta'}\widetilde{M_{\alpha'}}\varphi(x,k). \tag{27.69}$$

D'autre part nous avons $\partial^\alpha\tilde{\varphi} = \widetilde{\partial^\alpha\varphi}$ parce que la fonction $\partial_x\varphi$ étant Schwartz, la fonction

$$G(y) = \sup_{x\in\mathbb{R}^n}|(\partial_x\varphi)(x,y)| \tag{27.70}$$

est dans $L^1(\mathbb{R}^m)$ par le corollaire 24.79. Par conséquent le théorème 14.65 permet de permuter la dérivée et l'intégrale dans

$$\frac{\partial}{\partial x}\tilde{\varphi}(x,k) = \frac{\partial}{\partial x}\int_{\mathbb{R}^m} e^{-iky}\varphi(x,y)dy. \tag{27.71}$$

Dans le même ordre d'esprit mais dans difficultés de permutation de limites nous avons $M_\beta\tilde{\varphi} = \widetilde{M_\beta\varphi}$.

D'autre part nous avons encore $\|\tilde{\varphi}\|_\alpha < \infty$ parce que

$$|\tilde{\varphi}(x,k)| \leqslant \int_{\mathbb{R}^m}|\varphi(x,y)|dy \leqslant \sup_x\int_{\mathbb{R}^m}|\varphi(x,y)|dy \leqslant \int_{\mathbb{R}^m}|\sup_x\varphi(x,y)|dy < \infty \tag{27.72}$$

parce que φ est Schwartz et le corollaire 24.79 donne l'intégrabilité.

Donc nous avons

$$p_{(\alpha\alpha'),(\beta\beta')(\tilde{\varphi})} = \|\partial^{\beta'}\widetilde{M_{\alpha'}M_\beta}\partial^\alpha\varphi\|_\infty < \infty. \tag{27.73}$$

Cela prouve que $\tilde{\varphi}$ est Schwartz. □

27.2.1 Quelque transformées de Fourier

Exemple 27.14([2])
Soit la fonction $g_\epsilon(x) = e^{-\epsilon x^2}$. Sa transformée de Fourier sera déduite dans le lemma 27.15 en utilisant le lemme de transfert 27.11. Nous nous proposons ici de déduire de façon directe l'équation différentielle vérifiée par la transformée de Fourier de g_ϵ.

Nous posons

$$I(k) = \int_{\mathbb{R}} e^{-ikx}e^{-\epsilon x^2}dx. \tag{27.74}$$

et nous considérons la fonction

$$f(k,x) = e^{-ikx}e^{-\epsilon x^2}. \tag{27.75}$$

Elle est de classe C^1 par rapport à k, et intégrable en x pour chaque k. De plus sa dérivée

$$(\partial_k f)(k,x) = -ixe^{-ikx}e^{-\epsilon x^2} \tag{27.76}$$

vérifie $|\partial_k f| \leqslant xe^{-\epsilon x^2}$. La dérivée est donc majorée (uniformément en k) par une fonction intégrable. Le théorème 14.65 permet de permuter la dérivée et l'intégrale :

$$\begin{aligned}
I'(k) &= \int_{\mathbb{R}} -ixe^{-ikx}e^{-\epsilon x^2}dx && (27.77a)\\
&= i\int_{\mathbb{R}} e^{-ikx}\frac{1}{2\epsilon}\frac{d}{dx}\left(e^{-\epsilon x^2}\right)dx && (27.77b)\\
&= \frac{-i}{2\epsilon}\int_{\mathbb{R}}\frac{d}{dx}\left(e^{-ikx}\right)e^{-\epsilon x^2}dx \qquad \text{par partie} && (27.77c)\\
&= \frac{-k}{2\epsilon}\int_{\mathbb{R}} e^{-ikx}e^{-\epsilon x^2}dx && (27.77d)\\
&= \frac{-k}{2\epsilon}I(k). && (27.77e)
\end{aligned}$$

D'où l'équation différentielle $I'(k) = -\frac{k}{2\epsilon}I(k)$. △

Lemme 27.15 (Transformée de Fourier de la Gausienne [301])**.**
La transformée de Fourier de

$$\begin{aligned} g_\epsilon\colon \mathbb{R}^d &\to \mathbb{R} \\ x &\mapsto e^{-\epsilon\|x\|^2} \end{aligned} \tag{27.78}$$

est donnée par

$$\hat{g}_\epsilon(\xi) = \Big(\frac{\pi}{\epsilon}\Big)^{d/2} e^{-\|\xi\|^2/4\epsilon} \tag{27.79}$$

Démonstration. Nous commençons par la fonction $g(x) = e^{-\|x\|^2/2}$ et nous prouvons que sa transformée de Fourier est $\hat{g}(\xi) = (2\pi)^{d/2} g(\xi)$.

Réduction à la dimension 1 La fonction g est dans l'espace de Schwartz. Par le théorème de Fubini,

$$\hat{g}(\xi) = \int_{\mathbb{R}^d} \prod_{k=1}^{d} e^{-x_k^2} e^{-i\xi_k x_k} dx = \prod_{k=1}^{d} \int_{\mathbb{R}} e^{-t^2/2} e^{-\xi_k x} dt = \prod_{k=1}^{d} \hat{f}(\xi_k) \tag{27.80}$$

où f est la fonction d'une variable

$$f(x) = e^{-x^2/2}. \tag{27.81}$$

Notons que $f \in \mathscr{D}(\mathbb{R})$.

Une équation différentielle Voyons l'équation différentielle satisfaite par la transformée de Fourier $\hat{f}$ de la fonction (27.81). Grâce au lemme 27.11 nous trouvons l'équation différentielle [4]

$$\xi \hat{f}(\xi) + (\hat{f})'(\xi) = 0. \tag{27.82}$$

C'est le moment d'utiliser le théorème de Cauchy-Lipschitz (18.37), appliqué à la fonction $f(t, y) = -ty$ qui est Lipschitz et continue au au problème

$$\begin{cases} y' + ty = 0 & \text{(27.83a)} \\ y(0) = y_0 & \text{(27.83b)} \end{cases}$$

possède une unique solution maximale, en l'occurrence $y(x) = y_0 e^{-x^2/2}$. En ce qui concerne la condition initiale nous avons

$$\hat{f}(0) = \int_{\mathbb{R}} e^{-x^2/2} dx = \sqrt{2\pi}. \tag{27.84}$$

par l'exemple 13.233. Donc

$$\hat{f}(\xi) = \sqrt{2\pi} e^{-\xi^2/2}. \tag{27.85}$$

En reformant le produit (27.80) nous concluons.

Nous passons maintenant à la fonction g_ϵ. Nous pouvons écrire g_ϵ sous la forme

$$g_\epsilon(x) = g(\sqrt{2\epsilon}x). \tag{27.86}$$

Utilisant successivement la transformée de Fourier de g que nous venons de calculer et 27.3 (facteur d'échelle) nous trouvons

$$\hat{g}(\xi) = (2\pi)^{d/2} g(\xi) \tag{27.87a}$$

$$\hat{g}_\epsilon(\xi) = (2\epsilon)^{-d/2} \hat{g}\big(\xi/\sqrt{2\epsilon}\big) \tag{27.87b}$$

$$= \Big(\frac{\pi}{\epsilon}\Big)^{d/2} e^{-|\xi|^2/4\epsilon}.. \tag{27.87c}$$

Nous voyons que $\hat{g}_\epsilon \in \mathscr{S}(\mathbb{R}^d)$ (c'était gagné d'avance par la proposition 27.12). □

4. Une façon directe de déduire cette équation différentielle est donnée dans l'exemple 27.14.

27.3 Suite régularisante

Définition 27.16.
*Une **suite régularisante** est une suite (ρ_n) dans $L^1(\mathbb{R}^d)$ telle que*

(1) pour tout n, $\rho_n \geqslant 0$ et $\int_{\mathbb{R}^d} \rho_n = 1$;

(2) pour tout $\alpha > 0$,

$$\lim_{n\to\infty} \int_{|t|>\alpha} \rho_n = 0. \tag{27.88}$$

Une telle suite est régularisante parce que souvent $\rho_n \in \mathscr{D}(\mathbb{R}^d)$, ce qui donne $f * \rho_n \in C^\infty$ par le corollaire 24.42.

Proposition 27.17 ([302, 303])**.**
Soit une suite régularisante $\rho_n \in L^1(\mathbb{R}^d)$. Alors :

(1) Si f est continue à support compact, nous avons la convergence uniforme sur $\mathbb{R}^d$:

$$f * \rho_n \xrightarrow{unif} f. \tag{27.89}$$

(2) Si $g \in L^p$ $(1 \leqslant p < \infty)$ alors

$$g * \rho_n \xrightarrow{L^p} g. \tag{27.90}$$

Démonstration. Si f est continue à support compact, elle est uniformément continue[5], et elle est bornée. Soit $\epsilon > 0$ et $\alpha > 0$ tel que pour tout x, y tels que $\|x-y\| < \alpha$ nous ayons $|f(x)-f(y)| < \epsilon$. Nous prenons de plus n suffisamment grand pour avoir $\int_{B(0,\alpha)^c} \rho_n < \epsilon$. Nous avons alors

$$\begin{aligned}
|f(x) - (f*\rho_n)(x)| &= |\int_{\mathbb{R}^d} \big(f(x)-f(y)\big)\rho_n(x-y)dy| && \text{(27.91a)}\\
&\leqslant \int_{B(x,\alpha)} \underbrace{|f(x)-f(y)|}_{\leqslant\epsilon} \rho_n(x-y)dy + \int_{B(x,\alpha)^c} \underbrace{|f(x)-f(y)|}_{\leqslant 2\|f\|_\infty} \rho_n(x-y)dy && \text{(27.91b)}\\
&\leqslant \epsilon(1+2\|f\|_\infty). && \text{(27.91c)}
\end{aligned}$$

Nous avons prouvé que pour tout $\epsilon > 0$, il existe N tel que $n > N$ implique $\big|f(x)-(f*\rho_n)(x)\big| \leqslant \epsilon$. Cela prouve l'uniforme convergence sur $\mathbb{R}^d$ de $f * \rho_n$ vers f.

Pour le point (2) nous considérons $g \in L^1(\mathbb{R}^d)$ et $\phi \in \mathscr{D}(\mathbb{R}^d)$. Nous avons la majoration

$$\|g*\rho_n - g\|_p \leqslant \|g*\rho_n - \phi*\rho_n\|_p + \|\phi*\rho_n - \phi\|_p + \|\phi - g\|_p \tag{27.92}$$

En ce qui concerne le premier terme ;

$$\|(g-\phi)*\rho_n\|_p \leqslant \|g-\phi\|_p \tag{27.93}$$

par la proposition 24.40. Donc

$$\|g*\rho_n - g\|_p \leqslant 2\|g-\phi\|_p + \|\phi*\rho_n - \phi\|_p. \tag{27.94}$$

Par la densité de $\mathscr{D}$ dans L^p (théorème 24.33(5)) nous pouvons considérer une suite $\phi_i \xrightarrow{L^p} g$ dans $\mathscr{D}(\mathbb{R}^d)$. Pour tout i nous avons

$$\|g*\rho_n - g\|_p \leqslant 2\|g-\phi_i\|_p + \|\phi_i*\rho_n - \phi\|_p. \tag{27.95}$$

Nous effectuons la limite sur $n \to \infty$:

$$\lim_{n\to\infty} \|g*\rho_n - g\|_p \leqslant 2\|g-\phi_i\| + \underbrace{\lim_{n\to\infty} \|\phi_i*\rho_n - \phi_i\|_p}_{=0} \tag{27.96}$$

5. Théorème de Heine 12.62.

parce que le point (1) s'applique à ϕ_i. Nous effectuons ensuite la limite sur $i \to \infty$ dans

$$\lim_{n\to\infty} \|g * \rho_n - g\| \leqslant 2\|g - \phi_i\| \to 0. \tag{27.97}$$

□

Lemme 27.18.
Si $g_\epsilon(x) = e^{-\epsilon\|x\|^2}$ alors la suite

$$\rho_n = \frac{1}{(2\pi)^d}\hat{g}_{1/n} \tag{27.98}$$

est une suite régularisante (définition 27.16).

Démonstration. Nous savons déjà la transformée de Fourier de g_ϵ par le lemme 27.15. Nous montrons que la suite ρ_n est régularisante. Nous avons $\hat{g}_\epsilon \in L^1(\mathbb{R}^d)$ et $\hat{g}_\epsilon \geqslant 0$ ainsi que $\lim_{\epsilon\to 0}\int_{B(0,\alpha)} \hat{g}_\epsilon = 0$ pour tout α. Il y a seulement un couac avec la norme. Nous calculons $\int_{\mathbb{R}^d} \hat{g}_\epsilon(\xi)d\xi$ avec la forme (27.87c). En utilisant sauvagement Fubini [6] pour séparer les intégrales et en effectuant le changement de variable $u = t/(2\sqrt{\epsilon})$ nous calculons :

$$\int_{\mathbb{R}^d} e^{-|\xi|^2/4\epsilon}d\xi = \prod_{k=1}^{d}\int_{\mathbb{R}} e^{-t^2/4\epsilon}dt \tag{27.99a}$$

$$= 2\sqrt{\epsilon}\prod_{k=1}^{d}\int_{\mathbb{R}} e^{-u^2}du \tag{27.99b}$$

$$= \prod_{k=1}^{d} 2\sqrt{\epsilon}\sqrt{\pi} \tag{27.99c}$$

$$= 2^d(\pi\epsilon)^{d/2}. \tag{27.99d}$$

Nous avons utilisé l'exemple 13.233 pour le calcul de l'intégrale gaussienne. Avec tout cela nous avons

$$\int_{\mathbb{R}^d} \hat{g}_\epsilon = (2\pi)^d. \tag{27.100}$$

Donc $\frac{1}{(2\pi)^d}\hat{g}_{1/n}$ est une suite régularisante. □

Le corollaire suivant regroupe les résultats à propos des suites régularisantes, leur utilité et leur existence.

Corollaire 27.19.
Si la suite régularisante ρ_n est dans $L^1(\mathbb{R}^d) \cap C^\infty(\mathbb{R}^d)$ alors pour $f \in L^p(\mathbb{R}^d)$ en posant $f_n = \rho_n * f$ nous avons

(1) $f_n \in C^\infty(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$

(2) $f_n \xrightarrow{L^p} f$

De plus, de telles suites existent.

Démonstration. Le fait que f_n soit de classe C^∞ est le corollaire 24.42, et la convergence est la proposition 27.17(2).

De telles suites existent, par exemple celle donnée par le lemme 27.18. □

6. Le pauvre !

27.3.1 Formule d'inversion

Proposition 27.20 (Formule d'inversion de Fourier[2])**.**
Si $f \in \mathscr{S}(\mathbb{R})$, alors nous avons la formule d'inversion

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ikx} \hat{f}(k) dk. \tag{27.101}$$

Cette formule peut d'écrire de plusieurs autres façons :

$$\mathcal{F}\big(\mathcal{F}(f)\big)(x) = 2\pi f(-x), \tag{27.102a}$$
$$\mathcal{F}^{-1}(f)(x) = \frac{1}{2\pi} \hat{f}(-x), \tag{27.102b}$$
$$f(x) = \frac{1}{2\pi} \mathcal{F}(\hat{f})(-x). \tag{27.102c}$$

Démonstration. Pour $\epsilon > 0$ nous posons

$$f_\epsilon(k) = e^{-\epsilon k^2} e^{ikx} \hat{f}(k). \tag{27.103}$$

Nous allons calculer

$$\lim_{\epsilon \to 0} \int_{\mathbb{R}} e^{-\epsilon k^2} e^{ikx} \hat{f}(k) dk \tag{27.104}$$

de deux façons.

D'abord en utilisant directement le théorème de la convergence dominée 13.134. La fonction $\hat{f}$ est dans $\mathscr{S}(\mathbb{R})$ (théorème 27.12) et par conséquent $f_\epsilon \in L^1(\mathbb{R})$ parce que le facteur $e^{-\epsilon k^2}$ ne va certainement pas empêcher de converger. De plus $|f_\epsilon| \leqslant |\hat{f}|$ et $\hat{f} \in L^1$. Le théorème est de la convergence dominée est applicable et

$$\lim_{\epsilon \to 0} \int_{\mathbb{R}} e^{-\epsilon k^2} e^{ikx} \hat{f}(k) dk = \int_{\mathbb{R}} e^{ikx} \hat{f}(k) dk. \tag{27.105}$$

Pour le deuxième calcul nous allons faire appel à Fubini [7] pour la fonction

$$\begin{aligned} u\colon \mathbb{R} \times \mathbb{R} &\to \mathbb{R} \\ (k, y) &\mapsto e^{ik(x-y)} e^{-\epsilon k^2} f(y). \end{aligned} \tag{27.106}$$

D'abord nous nous assurons que $u \in L^1(\mathbb{R} \times \mathbb{R})$ par le corollaire 13.229, et ensuite nous utilisons le théorème de Fubini 13.230 pour manipuler les intégrales (et en particulier les inverser). Dans un premier temps nous avons :

$$\int_{\mathbb{R}} \int_{\mathbb{R}} |e^{ik(x-y)} e^{-\epsilon k^2} f(y)| dy\, dk \leqslant \int_{\mathbb{R}} e^{-\epsilon k^2} \Big[\int_{\mathbb{R}} |f(y)| dy \Big] dk < \infty \tag{27.107}$$

parce que f étant dans $\mathscr{S}(\mathbb{R})$, l'intégrale intérieure se réduit à un nombre. Nous savons maintenant

7. Parce qu'il est toujours plus simple de refiler le boulot aux autres que de le faire soi-même... pauvre Fubini !

que $u \in L^1(\mathbb{R} \times \mathbb{R})$. Nous pouvons alors calculer un peu ...

$$\int_{\mathbb{R}} e^{ikx} e^{-\epsilon k^2} \hat{f}(k) dk = \int_{\mathbb{R}} \int_{\mathbb{R}} e^{ikx} e^{-\epsilon k^2} e^{-iky} f(y) dy\, dk \tag{27.108a}$$

$$= \int_{\mathbb{R}} \Big[\int_{\mathbb{R}} e^{ik(x-y)} e^{-\epsilon k^2} f(y) dk \Big] dy \tag{27.108b}$$

$$= \int_{\mathbb{R}} f(y) \Big[\int_{\mathbb{R}} e^{ik(x-y)} e^{-\epsilon k^2} dk \Big] dy \tag{27.108c}$$

$$= \int_{\mathbb{R}} f(y) \hat{g}_\epsilon(y - x) dy \tag{27.108d}$$

$$= \sqrt{\frac{\pi}{\epsilon}} \int_{\mathbb{R}} f(y) e^{-(y-x)^2/4\epsilon} dy \tag{27.108e}$$

$$= 2\sqrt{\epsilon} \sqrt{\frac{\pi}{\epsilon}} \int_{\mathbb{R}} f(x + 2\sqrt{\epsilon}t) e^{-t^2} dt \tag{27.108f}$$

$$= 2\sqrt{\pi} \int_{\mathbb{R}} f(x + 2\sqrt{\epsilon}t) e^{-t^2} dt \tag{27.108g}$$

$$\tag{27.108h}$$

Justifications :

— La fonction g_ϵ est la gaussienne dont la transformée de Fourier est calculée dans le lemme 27.15.

— Nous avons effectué le changement de variables $t = (y - x)/(2\sqrt{\epsilon})$ qui donne $dt = dy/2\sqrt{\epsilon}$.

La fonction f étant Schwartz (en particulier bornée), dans la dernière intégrale, nous pouvons effectuer la majoration

$$f(x + 2\sqrt{\epsilon}t) e^{-t^2} \leqslant \|f\|_\infty e^{-t^2}, \tag{27.109}$$

qui est une fonction intégrable. Nous pouvons donc permuter la limite et l'intégrale. Dans l'égalité

$$\lim_{\epsilon \to 0} \int_{\mathbb{R}} e^{ikx} e^{-\epsilon k^2} \hat{f}(k) dk = \lim_{\epsilon \to 0} 2\sqrt{\pi} \int_{\mathbb{R}} f(x + 2\sqrt{\epsilon}t) e^{-t^2} dt \tag{27.110}$$

À gauche nous avons déjà la limite depuis (27.105), et à droite nous obtenons

$$\lim_{\epsilon \to 0} 2\sqrt{\pi} \int_{\mathbb{R}} f(x + 2\sqrt{\epsilon}t) e^{-t^2} dt = \int_{\mathbb{R}} f(x) e^{-t^2} dt = 2\sqrt{\pi} f(x) \sqrt{\pi} = 2\pi f(x) \tag{27.111}$$

où nous avons utilisé l'intégrale gaussienne faite dans l'exemple 13.233.

En remettant tout ensemble,

$$2\pi f(x) = \lim_{\epsilon \to 0} \int_{\mathbb{R}} e^{-\epsilon k^2} e^{ikx} \hat{f}(k) dk = \int_{\mathbb{R}} e^{ikx} \hat{f}(k) dk, \tag{27.112}$$

ce qu'il fallait prouver. □

Corollaire 27.21.
Nous avons la formule

$$\int_{\mathbb{R}} \int_{\mathbb{R}} e^{-ikx} f(x) dx\, dk = 2\pi f(0). \tag{27.113}$$

Démonstration. Poser $x = 0$ dans l'équation (27.101). □

27.22.
Les physiciens qui n'ont que rarement peur écrivent souvent la formule (27.113) sous la forme

$$\int_{\mathbb{R}} e^{-ikx} dk = \delta(x) \tag{27.114}$$

où δ serait la fonction de Dirac qui vaut zéro partout sauf en $x = 0$ où elle vaudrait l'infini, mais pas n'importe quel infini ; juste celui qu'il faut pour que sont intégrale valle 1.

Lemme 27.23.
Si $\phi \in \mathscr{S}(\mathbb{R} \times \mathbb{R}^n)$, alors

$$\partial_t \hat{\phi} = \widehat{\partial_t \phi} \tag{27.115}$$

où le chapeau dénote la transformée de Fourier par rapport à la variable dans $\mathbb{R}^n$ et non par rapport à celle dans $\mathbb{R}$. Le t par contre est la variable dans $\mathbb{R}$.

Démonstration. Par définition de la transformée de Fourier nous avons

$$(\partial_t \hat{\phi})(t, \xi) = \frac{\partial}{\partial t} \int_{\mathbb{R}^n} \phi(t, x) e^{-ix\xi} dx. \tag{27.116}$$

Notre but est de permuter l'intégrale et la dérivée en utilisant le théorème 14.65. Il nous faut une fonction $G\colon \mathbb{R}^n \to \mathbb{R}$ qui soit intégrable sur $\mathbb{R}^n$ et telle que

$$\left| \frac{\partial \phi}{\partial t} \phi(t, x) \right| \leqslant G(x) \tag{27.117}$$

pour tout $t \in B(t_0, \delta)$. Étant donné que la fonction $\partial_t \phi$ est tout autant Schwartz que ϕ elle-même nous pouvons alléger les notations et chercher une fonction G qui convient pour toute fonction $\varphi \in \mathscr{S}(\mathbb{R} \times \mathbb{R}^n)$. Soit la fonction

$$G(x) = \sup_{t \in B(t_0, \delta)} |\varphi(t, x)|. \tag{27.118}$$

Pour tout multiindice α nous avons alors

$$\sup_{x \in \mathbb{R}^n} \left| x^\alpha G(x) \right| \leqslant \sup_{(t,x) \in \mathbb{R} \times \mathbb{R}^n} \left| x^\alpha \varphi(t, x) \right| \leqslant M_\alpha \in \mathbb{R}. \tag{27.119}$$

Grâce à la proposition 24.78, cela signifie que φ décroît plus vite que n'importe quel polynôme ; G est donc intégrable sur $\mathbb{R}^n$. □

27.4 Transformée de Fourier sur $L^2(\mathbb{R}^d)$

La théorie des transformées de Fourier est intéressante sur $L^2(\mathbb{R}^d)$ parce qu'elle y donne une isométrie. Nous allons la donner avec des fonctions à valeurs dans $\mathbb{C}$.

Remarque 27.24.
Une remarque qui vaut ce qu'elle vaut, mais si u est une classe de fonction pour la relation $u \sim v$ si et seulement si $u(x) = v(x)$ pour presque tout v alors l'intégrale

$$\hat{u}(\xi) = \int_{\mathbb{R}^d} u(x) e^{ix\xi} dx \tag{27.120}$$

ne dépend pas du choix du représentant. Nous pouvons donc parfaitement parler de transformée de Fourier d'une classe de fonctions.

27.4.1 Extension de $L^1 \cap L^2$ vers L^2

Théorème 27.25 (Extention de la transformée de Fourier vers $L^2(\mathbb{R}^d)$[301])**.**
Soit $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. Alors

(1) Nous avons $\mathcal{F}(f) \in L^2$ et $\|\hat{f}\|_{L^2} = (2\pi)^d \|f\|_{L^2}$.

(2) L'application $\mathcal{F}\colon L^1 \cap L^2 \to L^2$ peut être étendue en une application $\mathcal{F}\colon L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ vérifiant

$$\|\hat{f}\|_{L^2} = (2\pi)^d \|f\|_{L^2} \tag{27.121}$$

pour tout $f \in L^2(\mathbb{R}^d)$.

Démonstration. Le fait que $f \in L^1$ implique $\|\mathcal{F}(f)\|_\infty \leqslant \|f\|_1$ (c'est le lemme 27.6). En particulier, $|\mathcal{F}(f)(\xi)|^2$ est majoré et l'intégrale

$$\clubsuit = \int_{\mathbb{R}^d} |\hat{f}|^2 e^{-\epsilon\xi^2} d\xi \tag{27.122}$$

existe et est finie.

Découper l'intégrale Dans un premier temps nous développons les intégrales. Dans les égalités suivantes, $x\xi$ est le produit scalaire $x \cdot \xi$ dans $\mathbb{R}^d$.

$$\clubsuit = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \overline{f(x)} e^{ix\xi} dx \right) \left(\int_{\mathbb{R}^d} f(y) e^{-y\xi} \right) e^{-\epsilon|\xi|^2} d\xi \tag{27.123a}$$

$$= \int_{\mathbb{R}^d} \left[\int_{\mathbb{R}^d \times \mathbb{R}^d} \overline{f(x)} f(y) e^{i\xi(x-y)} dxdy \right] e^{-\epsilon|\xi|^2} d\xi. \tag{27.123b}$$

Nous avons utilisé le théorème de Fubini pour regrouper les intégrales [8]. Vu que $f \in L^1(\mathbb{R}^d)$, la fonction $(x, y, \xi) \mapsto f(x)f(y)e^{-\epsilon|\xi|^2}$ est dans $L^1(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d)$ et le théorème de Fubini 13.230 avec $\Omega_1 = \mathbb{R}^d \times \mathbb{R}^d$ et $\Omega_2 = \mathbb{R}^d$ nous permet de permuter les intégrales pour avoir

$$\clubsuit = \int_{\mathbb{R}^d \times \mathbb{R}^d} \overline{f(x)} f(y) \left[\int_{\mathbb{R}^d} e^{i\xi(x-y)} e^{-\epsilon|\xi|^2} d\xi \right] dxdy. \tag{27.124}$$

Discuter de cette gaussienne En posant

$$g(x) = e^{-|x|^2/2} \tag{27.125a}$$

$$g_\epsilon(x) = g(\sqrt{2\epsilon}x) = e^{-\epsilon|x|^2} \tag{27.125b}$$

nous avons $g_\epsilon \in \mathscr{S}(\mathbb{R}^d)$ et le lemme 27.18 nous autorise à écrire

$$\hat{g}(\xi) = (2\pi)^{d/2} g(\xi) \tag{27.126a}$$

$$\hat{g}_\epsilon(\xi) = \left(\frac{\pi}{\epsilon}\right)^{d/2} e^{-|\xi|^2/4\epsilon} \tag{27.126b}$$

Nous voyons que $\hat{g}_\epsilon \in \mathscr{S}(\mathbb{R}^d)$ (c'était gagné d'avance par la proposition 27.12) et que $\hat{g}_\epsilon$ est une fonction paire (encore une fois, c'était gagné d'avance parce que la transformée de Fourier d'une fonction paire est paire).

Tout cela pour dire que l'intégrale entre crochet dans (27.124) est $\hat{g}_\epsilon(y - x) = \hat{g}_\epsilon(x - y)$, et donc

$$\clubsuit = \int_{\mathbb{R}^d \times \mathbb{R}^d} \overline{f(x)} f(y) \hat{g}_\epsilon(x - y) dxdy. \tag{27.127}$$

Encore une fois le théorème de Fubini permet de séparer les intégrales et de calculer l'intégrale sur y en premier. Vu que $f \in L^1$ et que $\hat{g}_\epsilon \in \mathscr{S}(\mathbb{R}^d)$, le produit de convolution $f * \hat{g}_\epsilon$ est un élément de $\mathscr{S}(\mathbb{R}^d)$ par la proposition 24.87. Nous avons donc

$$\clubsuit = \int_{\mathbb{R}^d} \overline{f(x)} (f * \hat{g}_\epsilon)(x) dx. \tag{27.128}$$

Là, nous reconnaissons un produit scalaire dans $L^2(\mathbb{R}^d)$, et donc

$$\int_{\mathbb{R}^d} |\hat{f}|^2 e^{-\epsilon\xi^2} d\xi = \langle f, f * \hat{g}_\epsilon \rangle_{L^2(\mathbb{R}^d)}. \tag{27.129}$$

Notons que tout a un sens : $f \in L^2(\mathbb{R}^d)$ et $f * \hat{g}_\epsilon \in \mathscr{S}(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$.

8. Dans la suite nous allons encore utiliser Fubini quelque fois pour regrouper et dégrouper des intégrales.

Suite régularisante Nous prenons la suite régularisante du lemme 27.18 donnée par

$$\rho_n = \frac{1}{(2\pi)^d}\hat{g}_{1/n}. \tag{27.130}$$

Première conclusion Nous reprenons (27.129)

$$\int_{\mathbb{R}^d} |\hat{f}|^2 e^{-|\xi^2|/n} d\xi = \langle f, f * \hat{g}_{1/n}\rangle_{L^2(\mathbb{R}^d)} = (2\pi)^d \langle f, f * \rho_n\rangle. \tag{27.131}$$

En prenant la limite $n \to \infty$ nous trouvons

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} |\hat{f}|^2 e^{-\epsilon\xi^2} d\xi = (2\pi)^d \|f\|^2. \tag{27.132}$$

Pour effectuer la limite du membre de gauche nous devons remarquer qu'en posant

$$g_n(\xi) = |\hat{f}(\xi)| e^{-|\xi|^2/n}, \tag{27.133}$$

nous avons une suite décroissante de fonction (c'est à dire que à ξ fixé, c'est décroissant en n). Par ailleurs ces fonctions sont toujours à valeurs dans $[0,\infty]$ et nous pouvons utiliser le théorème de la convergence monotone 13.129 pour permuter la limite et l'intégrale. Au final :

$$\|\hat{f}\|_{L^2} = (2\pi)^d \|f\|_{L^2}. \tag{27.134}$$

En ce qui concerne l'extension, soit $f \in L^2(\mathbb{R}^d)$ et une suite (f_n) dans $L^1 \cap L^2$ telle que $f_n \xrightarrow{L^2} f$.

Existence d'une telle suite Si $f \in L^2(\mathbb{R}^d)$, alors nous pouvons poser

$$f_n(x) = f(x) e^{-|x|^2/n^2}. \tag{27.135}$$

Par l'inégalité de Hölder (24.52) nous avons $f_n \in L^1(\mathbb{R}^d)$; de plus $f_n \in L^2(\mathbb{R}^d)$ parce que pour tout x nous avons $|f_n(x)| \leqslant |f(x)|$. Montrons que $f_n \xrightarrow{L^2} f$. Nous avons

$$\|f_n - f\|_{L^2}^2 = \int_{\mathbb{R}^d} |f(x)(1 - e^{-|x^2|/n^2})|^2 dx. \tag{27.136}$$

Nous voulons prendre la limite $n \to \infty$. Pour ce faire à à droite nous remarquons que $e^{-|x|^2/n^2}$ est majoré par 1 ; ce qui se trouve dans l'intégrale est donc majoré (uniformément en n) par $|f(x)|^2$, qui est une fonction L^1 parce que f est L^2. Le théorème de la convergence dominée 13.134 nous permet alors de permuter la limite et l'intégrale, ce qui donne

$$\lim_{n\to\infty} \|f_n - f\|_2^2 = \int_{\mathbb{R}^d} \lim_{n\to\infty} |f(x)(1 - e^{-|x|^2/n^2})|^2 dx = 0. \tag{27.137}$$

Définition de $\mathcal{F}: L^2 \to L^2$ La suite (f_n) est une suite convergence dans L^2, et elle est donc de Cauchy. De plus pour chaque n, m nous avons

$$\|\hat{f}_n - \hat{f}_m\| = (2\pi)^d \|f_n - f_m\|. \tag{27.138}$$

La suite $(\hat{f}_n)$ est donc elle aussi de Cauchy, dans l'espace $L^2(\mathbb{R}^d)$ qui est complet (lemme 24.54). Nous posons

$$\hat{f} = \lim_{n\to\infty} \hat{f}_n. \tag{27.139}$$

Indépendance aux choix Nous devons montrer que la définition de $\hat{f}$ ne dépend pas de la suite approximant f dans $L^1 \cap L^2$. Soient dans deux suites $f_n \xrightarrow{L^2} f$ et $g_n \xrightarrow{L^2} f$ telles que $\hat{f}_n \xrightarrow{L^2} F$ et $\hat{g}_n \xrightarrow{L^2} G$. Alors

$$\|\hat{f}_n - \hat{g}_n\| = (2\pi)^d \|f_n - g_n\| \leqslant (2\pi)^d \|f_n - f\| + (2\pi)^d \|g_n - f\| \to 0. \tag{27.140}$$

Par conséquent $(\hat{f}_n - \hat{g}_n)_n$ est une suite qui converge vers zéro. Par unicité de la limite, $F = G$.

□

Remarque 27.26.
Une autre suite possible, à la place de (27.135), est

$$f_n(x) = f(x)\mathbb{1}_{|x|<n}. \tag{27.141}$$

C'est à dire la fonction f limitée à une boule de rayon n autour de 0.