

Analyse

Notes de cours et exercices corrigés

Arnaud Durand

Département de Mathématiques d'Orsay

arnaud.durand@universite-paris-saclay.fr

Avant-propos

Ce document rassemble des notes de cours et des exercices d'analyse que j'ai donnés dans le cadre de la préparation à l'agrégation de mathématiques à Orsay, entre les années 2015 et 2023. Les intitulés et numéros de leçons sont ceux utilisés à l'époque de ces enseignements.

Conventions et notations

On note $\mathbb{N} = \{0, 1, 2, \dots\}$ et $\mathbb{N}^* = \{1, 2, \dots\}$. Le corps $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$, selon le contexte, et i est l'unité imaginaire. « Positif », « négatif », « croissant » et « décroissant » s'entendent au sens large. Le symbole $\subset$ autorise l'égalité et $\sqcup$ désigne une union disjointe. Une somme vide vaut 0 et un produit vide vaut 1.

Pour $A \subset X$, on note $\overline{A}$ son adhérence, $\overset{\circ}{A}$ son intérieur, $\partial A = \overline{A} \setminus \overset{\circ}{A}$ sa frontière et $A^c = X \setminus A$ son complémentaire. Un espace *séparable* possède une partie dénombrable dense (« dénombrable » inclut le cas fini). Dans un espace métrique (X, d) , $B(x, r) = \{y : d(x, y) < r\}$ et $\overline{B}(x, r) = \{y : d(x, y) \leq r\}$ désignent les boules ouverte et fermée. Attention : $\overline{B}(x, r)$ peut différer de $\overline{B(x, r)}$. Pour $A \neq \emptyset$, on pose $d(x, A) = \inf_{a \in A} d(x, a)$, et $\text{diam } A = \sup_{a, b \in A} d(a, b)$; on convient que $d(x, \emptyset) = +\infty$ et $\text{diam } \emptyset = 0$.

Pour des espaces normés E, F , $\mathcal{L}(E, F)$ ou $\mathcal{L}(E, F)$ désigne les applications linéaires continues de E dans F , avec $\mathcal{L}(E) = \mathcal{L}(E, E)$. La norme d'opérateur est $\|u\| = \sup_{\|x\| \leq 1} \|u(x)\|$; elle est aussi notée $\|u\|$ dans le chapitre de calcul différentiel. Le dual topologique est $E' = \mathcal{L}(E, \mathbb{K})$. Un espace de Banach est un espace normé complet. On note $\text{Vect } A$ le sous-espace engendré par A , Id l'identité, $\ker u$ le noyau et $\text{Im } u$ (ou $\text{im } u$) l'image d'une application linéaire. Une base sans autre précision est algébrique : les décompositions sont finies, contrairement aux développements sur une base hilbertienne.

On note $\mathcal{M}_n(\mathbb{K})$ l'espace des matrices carrées d'ordre n et $\text{GL}_n(\mathbb{K})$ le groupe des matrices inversibles. La notation ${}^t A$ désigne la transposée de A . Dans les espaces de Hilbert complexes, le produit scalaire $\langle u, v \rangle$ est linéaire en u et antilinéaire en v ; ainsi, sur L^2 , $\langle f, g \rangle = \int f \overline{g}$. Pour $\mathbb{R}^n$, $x \cdot y$ est le produit scalaire euclidien.

Les espaces $\mathcal{C}(X, F)$ et $\mathcal{C}^k(U, F)$ sont ceux des fonctions continues et de classe C^k . Sans espace d'arrivée précisé, les fonctions sont scalaires, réelles ou complexes suivant le contexte. Pour une fonction bornée, $\|f\|_\infty = \sup_x \|f(x)\|$. Une série $\sum f_n$ converge *normalement* sur A si $\sum_n \sup_{x \in A} \|f_n(x)\| < \infty$; pour une série de vecteurs d'un espace normé, cela signifie $\sum_n \|u_n\| < \infty$.

Sur un espace mesuré $(X, \mathcal{A}, \mu)$, « presque partout » (p.p.) signifie hors d'un ensemble de mesure nulle. Pour $1 \leq p < \infty$, $L^p(X, \mu)$ est l'espace des fonctions mesurables telles que $\int_X |f|^p d\mu < \infty$, identifiées lorsqu'elles sont égales presque partout, avec $\|f\|_p = (\int_X |f|^p d\mu)^{1/p}$. De même, L^∞ est l'espace des fonctions essentiellement bornées, avec $\|f\|_\infty = \inf\{M \geq 0 : |f| \leq M \text{ p.p.}\}$. Sur une partie de $\mathbb{R}^d$, la mesure est celle de Lebesgue sauf indication contraire. Les espaces $\ell^p(I)$ sont les espaces analogues de suites indexées par I , pour la mesure de comptage. La fonction indicatrice $\mathbf{1}_A$ vaut 1 sur A et 0 ailleurs.

En probabilités, $\mathbb{P}$ désigne la probabilité, $\mathbb{E}[X]$ l'espérance d'une variable intégrable et $\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$ la variance d'une variable réelle de carré intégrable. « Presque sûrement » (p.s.) signifie avec probabilité 1. La loi de X est la mesure $A \mapsto \mathbb{P}(X \in A)$; sa fonction de répartition, lorsque X est réelle, est $F_X(t) = \mathbb{P}(X \leq t)$.

Sauf convention locale explicitement annoncée, la transformée de Fourier est $\widehat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-ix \cdot \xi} dx$ pour $f \in L^1(\mathbb{R}^d)$. Son prolongement à L^2 est celui du théorème de Plancherel et vérifie $\|\widehat{f}\|_2 = (2\pi)^{d/2} \|f\|_2$. L'exercice 19.2 emploie le noyau $e^{-2\pi i x \cdot \xi}$; l'isométrie utilisée dans le corrigé de l'exercice 12.4 est $f \mapsto \widehat{f}/\sqrt{2\pi}$.

On note $\log = \ln$ le logarithme népérien réel. Pour $x > 0$ et $z \in \mathbb{C}$, on pose $x^z = \exp(z \log x)$. La branche du logarithme complexe est précisée localement. Au voisinage considéré, $f \sim g$ signifie $f/g \rightarrow 1$ et $f = o(g)$ signifie $f/g \rightarrow 0$, avec $g \neq 0$; pour f vectorielle et $r > 0$, $f = o(r)$ signifie $\|f\|/r \rightarrow 0$.

Table des matières

Avant-propos	i
I Notes de cours	1
1 Compacité	2
1.1 Généralités	2
1.2 Le théorème de Tychonoff et ses applications	6
1.3 Le théorème d'Ascoli	9
2 Espaces normés de dimension finie	11
2.1 Normes et compacité dans $\mathbb{K}^n$	11
2.2 Normes et compacité en dimension finie	12
2.3 Applications	12
3 Calcul différentiel	14
3.1 Différentiabilité et dérivées partielles	14
3.2 Dérivées partielles d'ordre supérieur	23
3.3 Points critiques et extrema	26
3.4 Inversion locale et fonctions implicites	29
4 Équations différentielles	35
4.1 Généralités	35
4.2 Équations différentielles linéaires	38
4.3 Équations différentielles non linéaires	41
4.4 Compléments	44
5 Espaces de Hilbert	48
5.1 Définitions et propriétés élémentaires	48
5.2 Projection sur un convexe fermé	49
5.3 Dualité et théorème de représentation de Riesz	52
5.4 Théorème de Lax-Milgram	53
5.5 Application au problème de Dirichlet	55
5.6 Bases hilbertiennes	56
5.7 Orthonormalisation de Gram-Schmidt	57
5.8 Inégalités de Wirtinger et isopérimétrique	58
6 Théorème de Baire et applications	60
6.1 Le théorème de Baire	60
6.2 Applications aux espaces vectoriels normés	61

6.3 Applications en analyse réelle	64
7 Théorèmes de Hahn–Banach et applications	68
7.1 Forme analytique	68
7.2 Formes géométriques	69
7.3 Applications	71
II Exercices corrigés	75
8 Compacité et connexité	76
9 Espaces vectoriels normés	82
10 Calcul différentiel et extrema	84
11 Équations différentielles	101
12 Espaces de Hilbert	119
13 Théorème de Baire	131
14 Suites numériques et récurrences	133
15 Espaces de fonctions, suites et séries de fonctions	141
16 Intégration et intégrales à paramètre	146
17 Convexité	154
18 Séries entières et analyse complexe	159
19 Transformation de Fourier	171
A Leçons couvertes par les exercices	175
Bibliographie	177

Première partie

Notes de cours

1.1 Généralités

1.1.1 Propriété de Borel-Lebesgue

Définition 1.1. Un espace métrique (E, d) est dit *compact* si de tout recouvrement de E par des ouverts de E , on peut extraire un sous-recouvrement fini. Autrement dit, si $E = \bigcup_{i \in I} O_i$, avec O_i ouvert pour tout $i \in I$, alors il existe $J \subset I$ fini tel que $E = \bigcup_{i \in J} O_i$.

Exemples. • Tout espace métrique fini est compact.

- $\mathbb{R}$ n'est pas compact car on ne peut pas extraire de sous-recouvrement fini du recouvrement $\mathbb{R} = \bigcup_{n \in \mathbb{N}^*}]-n, n[$.

Proposition 1.1. *Un espace métrique compact est borné.*

Démonstration. On peut supposer E non vide. Pour $x_0 \in E$ arbitraire, on extrait un recouvrement fini du recouvrement formé de boules ouvertes $E = \bigcup_{n \in \mathbb{N}^*} B(x_0, n)$, ce qui est possible quand E est compact. Cela implique que $E = B(x_0, n_0)$ pour un certain entier n_0 , et donc que E est borné. $\square$

Proposition 1.2 (Aspect dual). *Un espace métrique (E, d) est compact si et seulement si de toute famille de fermés de E d'intersection vide, on peut extraire une sous-famille finie d'intersection vide.*

Démonstration. Il suffit de passer aux complémentaires pour transformer les recouvrements ouverts en familles de fermés d'intersection vide. $\square$

Proposition 1.3. *Si $(F_n)_{n \in \mathbb{N}}$ est une suite décroissante de fermés non vides dans un espace compact E , alors $\bigcap_{n \in \mathbb{N}} F_n \neq \emptyset$.*

Remarque. Le résultat est encore vrai si E est complet et $\text{diam}(F_n) \xrightarrow{n \rightarrow \infty} 0$. Si cette dernière condition n'est pas vérifiée, c'est faux puisque $\bigcap_{n \in \mathbb{N}} [n, +\infty[= \emptyset$.

Parties compactes Pour caractériser les parties compactes, on peut utiliser directement les ouverts de l'espace ambiant, grâce à la définition de la topologie induite.

Proposition 1.4. *Soit (E, d) un espace métrique. Une partie A de E est compacte si et seulement si de tout recouvrement de A par des ouverts de E , on peut extraire un sous-recouvrement fini, i.e. si $A \subset \bigcup_{i \in I} O_i$, avec O_i ouvert de E pour tout $i \in I$, il existe $J \subset I$ fini tel que $A \subset \bigcup_{i \in J} O_i$.*

Cette caractérisation conduit aux deux propriétés de stabilité suivantes.

Proposition 1.5. *Une réunion finie de parties compactes est compacte.*

Proposition 1.6. *L'intersection d'une famille non vide de parties compactes d'un espace métrique est compacte.*

1.1.2 Propriété de Bolzano-Weierstrass

Théorème 1.7 (Bolzano-Weierstrass). *Un espace métrique (E, d) est compact si et seulement si de toute suite de points de E , on peut extraire une sous-suite convergente dans E .*

La propriété de Borel-Lebesgue implique celle de Bolzano-Weierstrass. Voici deux démonstrations de cette implication.

Première preuve. Soit $(x_n)_{n \in \mathbb{N}}$ une suite dans E . Pour tout $p \in \mathbb{N}$, on note $A_p = \{x_n, n \geq p\}$. La suite $(\overline{A_p})_{p \in \mathbb{N}}$ est une suite décroissante de fermés non vides, donc par compacité $\bigcap_{p \in \mathbb{N}} \overline{A_p} \neq \emptyset$. On note x un élément de cette dernière intersection, puis on construit une sous-suite $(x_{\varphi(n)})_{n \in \mathbb{N}}$ ainsi :

- on choisit $x_{\varphi(0)} \in A_0$ quelconque ;
- si $x_{\varphi(n)}$ est construit, on prend $x_{\varphi(n+1)} \in A_{\varphi(n)+1}$ tel que $\varphi(n+1) > \varphi(n)$ et $d(x_{\varphi(n+1)}, x) < 2^{-(n+1)}$, ce qui est possible puisque $x \in \overline{A_{\varphi(n)+1}}$.

On obtient bien une sous-suite $(x_{\varphi(n)})$ de (x_n) qui converge vers x .

Deuxième preuve. Soit $(x_n)_{n \in \mathbb{N}}$ une suite dans E . Si les valeurs de cette suite forment un ensemble fini, l'une d'elles apparaît une infinité de fois ; elle fournit une sous-suite constante, donc convergente. Sinon, l'ensemble $A = \{x_n, n \in \mathbb{N}\}$ est infini. On va montrer que A possède au moins un point d'accumulation. Si ce n'est pas le cas, alors pour tout $x \in E$, il existe $r_x > 0$ tel que $B(x, r_x) \cap A$ est fini. Par compacité, on extrait du recouvrement de E par les boules $B(x, r_x)$ un sous-recouvrement fini. Donc, il existe $J \subset E$ fini tel que $E = \bigcup_{x \in J} B(x, r_x)$. On en déduit

que l'ensemble $A = \bigcup_{x \in J} (B(x, r_x) \cap A)$ est fini, ce qui donne une contradiction. Ainsi, A admet au moins un point d'accumulation $x \in E$, donc il existe une sous-suite de (x_n) qui converge vers x .

La propriété de Bolzano-Weierstrass implique celle de Borel-Lebesgue. La démonstration repose sur les deux lemmes suivants.

Définition 1.2. Un espace métrique (E, d) est dit *précompact* si pour tout $\varepsilon > 0$, il existe un recouvrement fini de E par des boules ouvertes de rayon ε .

Lemme 1.8. *Tout espace métrique (E, d) vérifiant la propriété de Bolzano-Weierstrass est précompact.*

Démonstration. Supposons par l'absurde qu'il existe $\varepsilon > 0$ tel que E ne puisse pas être recouvert par un nombre fini de boules ouvertes de rayon ε . Soit $x_0 \in E$. On a nécessairement $B(x_0, \varepsilon) \neq E$, donc il existe $x_1 \in E$ tel que $d(x_1, x_0) \geq \varepsilon$. De même, $B(x_0, \varepsilon) \cup B(x_1, \varepsilon) \neq E$, donc il existe $x_2 \in E$ tel que $d(x_2, x_0) \geq \varepsilon$ et $d(x_2, x_1) \geq \varepsilon$. Ainsi, si les points $x_0, \dots, x_n$ sont construits, on a $B(x_0, \varepsilon) \cup \dots \cup B(x_n, \varepsilon) \neq E$, donc il existe $x_{n+1} \in E$ tel que $d(x_{n+1}, x_0), \dots, d(x_{n+1}, x_n) \geq \varepsilon$. On obtient finalement une suite $(x_n)_{n \in \mathbb{N}}$ de points de E telle que $d(x_i, x_j) \geq \varepsilon$ pour tout $i < j$. La séparation uniforme des termes empêche toute sous-suite d'être de Cauchy, donc de converger. On obtient la contradiction recherchée. $\square$

Lemme 1.9. Soit (E, d) un espace métrique vérifiant la propriété de Bolzano-Weierstrass, et soit $(O_i)_{i \in I}$ un recouvrement de E par des ouverts. Alors,

$$\exists \alpha > 0 \quad \forall x \in E \quad \exists i \in I \quad B(x, \alpha) \subset O_i.$$

Démonstration. Par l'absurde, supposons que pour tout $\alpha > 0$, il existe $x \in E$ tel que pour tout $i \in I$, on a $B(x, \alpha) \not\subset O_i$. En particulier, pour tout entier $n \geq 1$, il existe un point $x_n \in E$ tel que $B(x_n, 1/n) \not\subset O_i$ pour tout $i \in I$. On extrait de la suite $(x_n)_{n \in \mathbb{N}^*}$ une sous-suite convergente $(x_{\varphi(n)})$, de limite notée x . On note un indice $i \in I$ tel que $x \in O_i$. Comme O_i est ouvert, on a $B(x, 2r) \subset O_i$ pour un certain $r > 0$. De plus, comme $x_{\varphi(n)} \xrightarrow[n \rightarrow \infty]{} x$, il existe $N \in \mathbb{N}$ tel que $d(x_{\varphi(n)}, x) < r$ et $\varphi(n) > 1/r$ pour tout entier $n \geq N$. Ainsi, pour tout $y \in B(x_{\varphi(n)}, 1/\varphi(n))$, on a

$$d(x, y) \leq d(x, x_{\varphi(n)}) + d(x_{\varphi(n)}, y) < r + \frac{1}{\varphi(n)} < 2r.$$

On en déduit que $B(x_{\varphi(n)}, 1/\varphi(n)) \subset B(x, 2r) \subset O_i$, ce qui donne une contradiction. $\square$

Réunissons maintenant les conclusions de ces deux lemmes. Soit $(O_i)_{i \in I}$ un recouvrement de E par des ouverts. D'après le lemme 1.9, on a

$$\exists \alpha > 0 \quad \forall x \in E \quad \exists i \in I \quad B(x, \alpha) \subset O_i.$$

D'après le lemme 1.8, on peut recouvrir E par un nombre fini de boules ouvertes de rayon α , donc il existe des points $x_1, \dots, x_n \in E$ tels que $E = \bigcup_{j=1}^n B(x_j, \alpha)$. Pour tout indice j , il existe un indice $i_j \in I$ tel que $B(x_j, \alpha) \subset O_{i_j}$.

Finalement, $E = \bigcup_{j=1}^n B(x_j, \alpha) \subset \bigcup_{j=1}^n O_{i_j}$.

Corollaire 1.10 (Autres formulations de Bolzano-Weierstrass). Un espace métrique (E, d) est compact si et seulement si l'une des assertions suivantes est vérifiée :

- toute suite de E admet au moins une valeur d'adhérence dans E ;
- toute partie infinie de E admet au moins un point d'accumulation dans E .

1.1.3 Propriétés générales des compacts

Proposition 1.11. Soit (E, d) un espace métrique.

- Si E est compact et si A est une partie fermée de E , alors A est compact.
- Si A est une partie compacte de E , alors A est fermée et bornée.

Proposition 1.12. Un espace compact est complet.

Proposition 1.13. Soit (E, d) un espace métrique compact et (x_n) une suite de E admettant une et une seule valeur d'adhérence x . Alors, (x_n) converge vers x .

Démonstration. Par l'absurde, supposons que (x_n) ne converge pas vers x . Alors, il existe $\varepsilon > 0$ tel que pour tout entier $N \in \mathbb{N}$, il existe un entier $n \geq N$ tel que $d(x_n, x) \geq \varepsilon$. On peut donc construire une sous-suite $(x_{\varphi(n)})$ de (x_n) telle que $d(x_{\varphi(n)}, x) \geq \varepsilon$ pour tout n . Comme E est compact, on peut en extraire une sous-suite $(x_{\varphi \circ \psi(n)})$ qui converge vers une limite notée y . On a donc $d(y, x) \geq \varepsilon$, puis $x \neq y$, ce qui contredit le fait que (x_n) a une unique valeur d'adhérence. $\square$

Proposition 1.14. Soit $E_1, \dots, E_n$ des espaces métriques non vides. L'espace produit $E = E_1 \times \dots \times E_n$ est compact si et seulement si l'espace E_i est compact pour tout $i \in \{1, \dots, n\}$.

Remarques. • Si (x_n) et (y_n) sont à valeurs dans E , un espace métrique compact, alors (x_n, y_n) est à valeurs dans le compact $E \times E$, donc on peut extraire une sous-suite $(x_{\varphi(n)}, y_{\varphi(n)})$ qui converge, de sorte que les sous-suites $(x_{\varphi(n)})$ et $(y_{\varphi(n)})$ convergent toutes les deux. Ainsi, la sous-suite d'indices le long desquels il y a convergence est la même.

- Le théorème de Tychonoff ci-après donne une extension de ce résultat à un nombre infini dénombrable de compacts.

Une application $f : (X, d) \rightarrow (Y, \delta)$ est dite *L-lipschitzienne* si $\delta(f(x), f(y)) \leq Ld(x, y)$ pour tous $x, y \in X$. Une bijection est *bilipschitzienne* si elle et son inverse sont lipschitziens.

Proposition 1.15. Les parties compactes d'un $\mathbb{R}$ -espace vectoriel normé de dimension finie sont les fermés bornés.

Démonstration. Le sens direct résulte des propriétés déjà établies : toute partie compacte est fermée et bornée.

Pour la réciproque, le cas $E = \{0\}$ est immédiat ; on suppose donc $\dim E \geq 1$. On commence par observer qu'un segment de $\mathbb{R}$ est compact. En effet, soit $(O_i)_{i \in I}$ un recouvrement ouvert d'un segment $[a, b]$. On considère l'ensemble F des points $x \in [a, b]$ tels que l'intervalle $[a, x]$ peut être recouvert par un nombre fini d'ouverts O_i . Comme F est non vide (il contient a , puisque $\{a\}$ peut être recouvert par un seul ouvert) et inclus dans $[a, b]$, il admet une borne supérieure $m \in [a, b]$. Cette borne m appartient à un ouvert O_i . Il existe alors $\varepsilon > 0$ tel que le segment $V = [m - \varepsilon, m + \varepsilon] \cap [a, b] = [c, d]$ soit inclus dans O_i . Puisque m est adhérent à F , alors V rencontre F en un point noté x . En ajoutant O_i au recouvrement fini de $[a, x]$, on obtient un recouvrement fini de $[a, d]$, donc $d \in F$, puis $m \geq d = \min\{m + \varepsilon, b\}$, ce qui implique que $b = d \in F$.

La proposition 1.14 implique alors qu'un produit de n segments est compact dans $\mathbb{R}^n$. Ainsi, si on munit $\mathbb{R}^n$ de la norme $\|\cdot\|_\infty$, tout fermé borné est un fermé inclus dans un produit de segments, donc est compact.

Enfin, si E est un $\mathbb{R}$ -espace vectoriel de dimension n muni d'une norme $\|\cdot\|$, et si $(e_1, \dots, e_n)$ désigne une base de E , l'application $x = (x_1, \dots, x_n) \mapsto \|x_1 e_1 + \dots + x_n e_n\|$ est lipschitzienne, avec

$$\|x_1 e_1 + \dots + x_n e_n\| \leq (\|e_1\| + \dots + \|e_n\|) \|x\|_\infty.$$

Elle est donc continue sur le compact $\{x \in \mathbb{R}^n \mid \|x\|_\infty = 1\}$. Son minimum a est atteint (voir ci-après) en un point non nul, donc est strictement positif. Pour tout $x \in \mathbb{R}^n$ non nul, on a $\|(x_1/\|x\|_\infty)e_1 + \dots + (x_n/\|x\|_\infty)e_n\| \geq a$, donc

$$\|x_1 e_1 + \dots + x_n e_n\| \geq a \|x\|_\infty.$$

L'application linéaire $x = (x_1, \dots, x_n) \mapsto x_1 e_1 + \dots + x_n e_n$ est une bijection bilipschitzienne, de sorte qu'elle fait correspondre les parties fermées bornées et les parties compactes de $\mathbb{R}^n$ et de E . Les parties fermées bornées de E sont donc compactes. $\square$

Remarque. Le résultat précédent est faux en dimension infinie (voir le théorème de Riesz 2.6).

Proposition 1.16. Soit (x_n) une suite convergente d'un espace métrique (E, d) , dont la limite est notée ℓ . Alors, l'ensemble $\Gamma = \{x_n, n \in \mathbb{N}\} \cup \{\ell\}$ est compact.

Démonstration. Soit $(O_i)_{i \in I}$ un recouvrement de Γ par des ouverts de E . Comme $\ell \in \Gamma$, il existe $i_0 \in I$ tel que $\ell \in O_{i_0}$. De plus, comme O_{i_0} est ouvert et $x_n \xrightarrow{n \rightarrow \infty} \ell$, on a $x_n \in O_{i_0}$ pour tout $n > N$, pour un certain entier $N \in \mathbb{N}$. Pour tout entier $n \leq N$, il existe $j_n \in I$ tel que $x_n \in O_{j_n}$. On pose $J = \{j_n, n \leq N\} \cup \{i_0\}$, de sorte que $\Gamma \subset \bigcup_{j \in J} O_j$, avec J fini. $\square$

1.1.4 Compacts et applications continues

Proposition 1.17. Soit (E, d) un espace métrique compact, (F, δ) un espace métrique, et $f : E \rightarrow F$ une application continue. Alors, $f(E)$ est compact.

Remarque. Cela implique que l'application f est fermée, i.e. l'image de tout fermé est un fermé. C'est faux en général si E n'est pas compact.

Proposition 1.18. Soit $f : (E, d) \rightarrow (F, \delta)$ une application continue et bijective. Si (E, d) est compact, alors $f^{-1} : F \rightarrow E$ est continue, i.e. f est un homéomorphisme.

Remarque. Si $f : I \rightarrow J$ est continue et bijective, avec I et J des intervalles, alors f^{-1} est continue. On applique en effet la proposition à toute restriction de f à des segments, qui sont compacts.

Proposition 1.19. Soit $f : (E, d) \rightarrow \mathbb{R}$ une application continue, où (E, d) est compact et non vide. Alors, f est bornée et atteint ses bornes. Ainsi, il existe $a, b \in E$ tels que

$$f(a) = \inf_{x \in E} f(x) \quad \text{et} \quad f(b) = \sup_{x \in E} f(x).$$

Remarque. D'après le théorème des valeurs intermédiaires, si $f : \mathbb{R} \rightarrow \mathbb{R}$ est continue, elle envoie tout intervalle sur un intervalle. Donc, l'image par f d'un intervalle fermé borné est un intervalle fermé borné (i.e. un segment).

Théorème 1.20 (Heine). Soit (E, d) et (F, δ) deux espaces métriques, avec E compact. Soit $f : E \rightarrow F$ une application continue. Alors, f est uniformément continue.

Démonstration. Pour tout $\varepsilon > 0$, l'ensemble

$$K = \{(x, y) \in E \times E \mid \delta(f(x), f(y)) \geq \varepsilon\}$$

est compact (car fermé dans un compact). Si K est non vide, l'application distance d atteint son minimum sur K , noté η . On a nécessairement $\eta > 0$. Si K est vide, on prend $\eta > 0$ arbitraire. Dans les deux cas, on a bien que si $d(x, y) < \eta$, alors $(x, y) \notin K$, et donc $\delta(f(x), f(y)) < \varepsilon$. $\square$

1.2 Le théorème de Tychonoff et ses applications

1.2.1 Le théorème de Tychonoff

Théorème 1.21 (Tychonoff). Soit (E_n, d_n) , pour $n \in \mathbb{N}^*$, des espaces métriques. On munit l'espace $E = \prod_{n \in \mathbb{N}^*} E_n$ de la topologie produit (la topologie la moins fine qui rend chaque projection $E \rightarrow E_n$ continue). Alors, E est métrique et si chaque E_n est compact, alors E est compact.

Démonstration. Quitte à remplacer d_n par une distance équivalente, par exemple $\frac{d_n}{1 + d_n}$, on peut supposer que $d_n \leq 1$ uniformément. En désignant par x_n et y_n les composantes de x et y , respectivement, on munit alors E de la distance $d(x, y) = \sum_{n=1}^{\infty} 2^{-n} d_n(x_n, y_n)$. Clairement, si $d(x^{(k)}, x) \xrightarrow{k \rightarrow \infty} 0$, alors pour tout $n \in \mathbb{N}^*$ fixé, on a $2^{-n} d_n(x_n^{(k)}, x_n) \xrightarrow{k \rightarrow \infty} 0$, et donc $x_n^{(k)} \xrightarrow{k \rightarrow \infty} x_n$. Réciproquement, supposons que $x_n^{(k)} \xrightarrow{k \rightarrow \infty} x_n$ pour tout n . Alors, pour tout $\varepsilon > 0$, on choisit $N \in \mathbb{N}^*$ tel que $\sum_{n > N} 2^{-n} \leq \varepsilon$. Pour tout $n \leq N$, il existe K_n tel que $d_n(x_n^{(k)}, x_n) \leq \varepsilon$ pour tout $k \geq K_n$. Pour $k \geq \max\{K_1, \dots, K_N\}$, on a donc

$$d(x^{(k)}, x) = \sum_{n \leq N} 2^{-n} d_n(x_n^{(k)}, x_n) + \sum_{n > N} 2^{-n} d_n(x_n^{(k)}, x_n) \leq 2\varepsilon,$$

ce qui montre bien que $x^{(k)} \xrightarrow{k\infty} x$.

On utilise alors le théorème de Bolzano-Weierstrass et une extraction diagonale. On considère une suite $(x^{(k)})_{k \in \mathbb{N}^*}$ dans E . Alors, $(x_1^{(k)})_{k \in \mathbb{N}^*}$ est une suite dans E_1 , et par compacité on en extrait une sous-suite $(x_1^{(\varphi_1(k))})_{k \in \mathbb{N}^*}$ qui converge dans E_1 vers x_1 . On extrait ensuite de $(x_2^{(\varphi_1(k))})_{k \in \mathbb{N}^*}$ une sous-suite $(x_2^{(\varphi_1 \circ \varphi_2(k))})_{k \in \mathbb{N}^*}$ qui converge dans E_2 vers x_2 . On construit ainsi une suite d'injections $\varphi_n : \mathbb{N}^* \rightarrow \mathbb{N}^*$ strictement croissantes telles que pour tout $n \in \mathbb{N}^*$, la suite $(x_n^{(\varphi_1 \circ \dots \circ \varphi_n(k))})_{k \in \mathbb{N}^*}$ converge vers x_n dans E_n . On pose alors $\psi(k) = \varphi_1 \circ \dots \circ \varphi_k(k)$ pour tout $k \in \mathbb{N}^*$. Pour tout $n \in \mathbb{N}^*$, on a

$$d_n(x_n^{(\psi(k))}, x_n) = d_n(x_n^{(\varphi_1 \circ \dots \circ \varphi_n(\varphi_{n+1} \circ \dots \circ \varphi_k(k)))}, x_n)$$

dès que $k > n$. Soit $\varepsilon > 0$. On a $d_n(x_n^{(\varphi_1 \circ \dots \circ \varphi_n(k))}, x_n) \leq \varepsilon$ pour tout entier $k \geq K_n$, pour un certain $K_n \in \mathbb{N}^*$. Pour $k \geq K_n$, on a aussi par injectivité des extractions $\varphi_{n+1} \circ \dots \circ \varphi_k(k) \geq k \geq K_n$, ce qui donne $d_n(x_n^{(\psi(k))}, x_n) \leq \varepsilon$. On en déduit que $x_n^{(\psi(k))} \xrightarrow{k\infty} x_n$ pour tout $n \in \mathbb{N}^*$, ce qui implique que $x^{(\psi(k))} \xrightarrow{k\infty} x = (x_1, x_2, \dots)$ pour la topologie produit. On a utilisé ici le procédé diagonal de Cantor. $\square$

1.2.2 Le théorème de Banach-Alaoglu

On rappelle que E' est le dual *topologique*, constitué des formes linéaires continues. La topologie de la convergence simple sur E' est appelée topologie *faible étoile*, notée $\sigma(E', E)$: c'est la topologie la moins fine qui rend continues les évaluations $\varphi \mapsto \varphi(x)$ pour tous $x \in E$.

Théorème 1.22 (Banach-Alaoglu). *Pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, soit E un $\mathbb{K}$ -espace vectoriel normé séparable, de dual topologique E' . On note B et B' la boule unité fermée de E et celle de E' , respectivement. Soit $(x_n)_{n \in \mathbb{N}^*}$ une suite de B dense dans B . On pose*

$$d(\varphi, \psi) = \sum_{n=1}^{\infty} 2^{-n} \frac{|\varphi(x_n) - \psi(x_n)|}{1 + |\varphi(x_n) - \psi(x_n)|}$$

pour tous $\varphi, \psi \in B'$. Alors :

- (i). *d est une distance sur B' et l'espace (B', d) est compact;*
- (ii). *si $(\varphi_k)_{k \in \mathbb{N}^*}$ est dans B' et $\varphi \in B'$, on a $d(\varphi_k, \varphi) \xrightarrow{k\infty} 0$ si et seulement si $\varphi_k(x) \xrightarrow{k\infty} \varphi(x)$ pour tout $x \in E$.*

Ainsi, la topologie définie par d sur B' est, séquentiellement parlant, celle de la convergence simple.

Démonstration. Si $\mathbb{K} = \mathbb{R}$, on note $X_n = [-1, 1]$, et si $\mathbb{K} = \mathbb{C}$, on note $X_n = \{z \in \mathbb{C} \mid |z| \leq 1\}$, que l'on munit de la métrique usuelle δ_n . On considère ensuite $X = \prod_{n \in \mathbb{N}^*} X_n$, muni de la métrique $\delta = \sum_{n=1}^{\infty} 2^{-n} \frac{\delta_n}{1 + \delta_n}$, c'est-à-dire que

$$\delta(u, v) = \sum_{n=1}^{\infty} 2^{-n} \frac{|u_n - v_n|}{1 + |u_n - v_n|}$$

pour $u = (u_n)_{n \geq 1}$ et $v = (v_n)_{n \geq 1}$ dans X .

On définit une injection $j : B' \rightarrow X$ par $j(\varphi) = (\varphi(x_n))_{n \geq 1}$. On a bien $j(\varphi) \in X$ car $|\varphi(x_n)| \leq \|\varphi\| \|x_n\| \leq 1$. On observe que $j(B')$ est fermé dans (X, δ) . En effet, si $y = (y_n)_{n \geq 1} \in \overline{j(B')}$, il existe une suite $(\varphi_k)_{k \geq 1}$ dans B' telle que $\delta(j(\varphi_k), y) \xrightarrow{k\infty} 0$, ce qui donne $\varphi_k(x_n) \xrightarrow{k\infty} y_n$ pour tout n . Prenons maintenant $x \in B$ et $\varepsilon > 0$. Par densité, il existe $n \geq 1$ tel que $\|x - x_n\| \leq \varepsilon$. Comme la suite $(\varphi_k(x_n))_{k \geq 1}$ converge, il existe aussi $k_0 \geq 1$ tel que $|\varphi_k(x_n) - \varphi_\ell(x_n)| \leq \varepsilon$ pour $k, \ell \geq k_0$. On en déduit que

$$\begin{aligned} |\varphi_k(x) - \varphi_\ell(x)| &\leq |\varphi_k(x) - \varphi_k(x_n)| + |\varphi_k(x_n) - \varphi_\ell(x_n)| + |\varphi_\ell(x_n) - \varphi_\ell(x)| \\ &\leq \|\varphi_k\| \|x - x_n\| + |\varphi_k(x_n) - \varphi_\ell(x_n)| + \|\varphi_\ell\| \|x - x_n\| \leq 3\varepsilon, \end{aligned}$$

de sorte que la suite $(\varphi_k(x))_{k \geq 1}$ est de Cauchy dans $\mathbb{K}$, qui est complet. Cette suite converge donc vers une limite notée $\varphi(x)$. Par homogénéité, cette convergence a lieu pour tout $x \in E$, et la limite définit une forme linéaire φ , comme limite simple de formes linéaires. De plus, si $x \in B$, on a $|\varphi_k(x)| \leq 1$, et donc $|\varphi(x)| \leq 1$, ce qui donne finalement $\varphi \in B'$ et $y = j(\varphi)$. Finalement, $j(B')$ est bien fermé dans (X, δ) .

Comme (X, δ) est compact d'après le théorème de Tychonoff, on obtient que $j(B')$ est compact. De plus, j est injective par densité de $(x_n)_{n \geq 1}$. En posant $d(\varphi, \psi) = \delta(j(\varphi), j(\psi))$, on obtient alors une distance d sur B' qui rend (B', d) isométrique, donc homéomorphe, à $(j(B'), \delta)$, ce qui implique la compacité de (B', d) .

La deuxième assertion du théorème vient du fait que $d(\varphi_k, \varphi) \xrightarrow{k \rightarrow \infty} 0$ est équivalent à $\varphi_k(x_n) \xrightarrow{k \rightarrow \infty} \varphi(x_n)$ pour tout $n \in \mathbb{N}^*$, qui est équivalent à $\varphi_k(x) \xrightarrow{k \rightarrow \infty} \varphi(x)$ pour tout $x \in E$. En effet, pour $x \in B$ et $n \in \mathbb{N}^*$, on a

$$\begin{aligned} |\varphi_k(x) - \varphi(x)| &\leq |\varphi_k(x) - \varphi_k(x_n)| + |\varphi_k(x_n) - \varphi(x_n)| + |\varphi(x_n) - \varphi(x)| \\ &\leq \|\varphi_k\| \|x - x_n\| + |\varphi_k(x_n) - \varphi(x_n)| + \|\varphi\| \|x - x_n\|. \end{aligned}$$

Si la convergence a lieu en tout point x_n , alors on a

$$\limsup_{k \rightarrow \infty} |\varphi_k(x) - \varphi(x)| \leq 2\|x - x_n\|,$$

et on conclut pour $x \in B$ par densité de (x_n) dans B , puis pour tout $x \in E$ par homogénéité. $\square$

1.2.3 Un théorème de point fixe pour les systèmes dynamiques

La tribu borélienne $\mathcal{B}(X)$ est la tribu engendrée par les ouverts de X . Si $T : X \rightarrow X$ est mesurable, la *mesure image* $T(\mu)$ est définie par $T(\mu)(A) = \mu(T^{-1}(A))$ pour tout borélien A . Une mesure de probabilité est une mesure positive de masse totale 1. La mesure de Dirac δ_x est caractérisée par $\delta_x(A) = \mathbf{1}_A(x)$, ou encore $\int f d\delta_x = f(x)$.

Proposition 1.23. *Soit (X, d) un espace métrique compact non vide, $\mathcal{B}$ sa tribu borélienne, et $T : X \rightarrow X$ une application continue. Alors, T admet au moins une mesure de probabilité invariante, i.e. il existe au moins une mesure de probabilité μ sur $(X, \mathcal{B})$ telle que $T(\mu) = \mu$, i.e. pour toute fonction $f \in \mathcal{C}(X)$,*

$$\int_X f(T(x)) \mu(dx) = \int_X f(x) \mu(dx).$$

Démonstration. Comme l'espace (X, d) est métrique et compact, l'ensemble $\mathcal{C}(X)$ des fonctions continues sur X , muni de la norme sup, est séparable. D'après le théorème de représentation de Riesz, son dual est l'espace $\mathcal{M}(X)$ des mesures complexes sur $(X, \mathcal{B})$ muni de la norme en variation totale. Ici, $\mathcal{M}(X)$ désigne les mesures boréliennes complexes de variation totale finie ; leur norme est $\|\mu\|_{TV} = \sup \sum_{j=1}^m |\mu(A_j)|$, où le supremum porte sur toutes les partitions boréliennes finies $(A_1, \dots, A_m)$ de X .

On note B' la boule unité fermée de $\mathcal{M}(X)$. Pour $x_0 \in X$, on pose

$$\mu_k = \frac{1}{k} \sum_{\ell=0}^{k-1} \delta_{T^\ell(x_0)},$$

où T^ℓ est la ℓ -ième itérée de T , de sorte que $\mu_k \in B'$. Comme B' est compact d'après le théorème de Banach-Alaoglu, le théorème de Bolzano-Weierstrass implique que, quitte à extraire une sous-suite, il existe une mesure $\mu \in B'$ telle que pour toute fonction $f \in \mathcal{C}(X)$,

$$\int_X f(x) \mu_k(dx) \xrightarrow{k \rightarrow \infty} \int_X f(x) \mu(dx).$$

En prenant f positive, on obtient $\int_X f(x) \mu(dx) \geq 0$ par passage à la limite, et en prenant $f = 1$, on trouve $\mu(X) = 1$. Ainsi, la mesure limite μ est une mesure de probabilité.

Enfin, pour toute fonction $f \in \mathcal{C}(X)$, on a

$$\begin{aligned} \left| \int_X f(T(x)) \mu_k(dx) - \int_X f(x) \mu_k(dx) \right| &= \left| \frac{1}{k} \sum_{\ell=1}^k f(T^\ell(x_0)) - \frac{1}{k} \sum_{\ell=0}^{k-1} f(T^\ell(x_0)) \right| \\ &= \frac{1}{k} |f(T^k(x_0)) - f(x_0)| \leq \frac{2\|f\|_\infty}{k}. \end{aligned}$$

Quand $k \rightarrow \infty$, il y a convergence des intégrales pour f et $f \circ T$ qui sont deux fonctions continues. On obtient donc en passant à la limite

$$\int_X f(T(x)) \mu(dx) = \int_X f(x) \mu(dx),$$

ce qui montre bien que la mesure μ est T -invariante. □

1.3 Le théorème d'Ascoli

On considère deux espaces métriques compacts non vides X et Y , de distance notée d pour les deux. On note $\mathcal{C}(X, Y)$ l'espace des fonctions continues de X dans Y , muni de la distance de la convergence uniforme

$$d(f, g) = \sup_{x \in X} d(f(x), g(x)).$$

Cette distance est toujours finie car X est compact. L'espace $(\mathcal{C}(X, Y), d)$ est complet et il est constitué d'applications uniformément continues d'après le théorème de Heine.

Définition 1.3. Une partie A de $\mathcal{C}(X, Y)$ est dite *équicontinue* si

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x, y \in X \forall f \in A \quad d(x, y) \leq \delta \implies d(f(x), f(y)) \leq \varepsilon.$$

Il s'agit d'une continuité uniforme commune à toutes les fonctions de A : pour chaque $\varepsilon > 0$, le même $\delta > 0$ convient à toute la famille.

Définition 1.4. Une partie A d'un espace métrique est dite *relativement compacte* si son adhérence $\overline{A}$ est compacte.

Théorème 1.24 (Ascoli). Soit A une partie de $\mathcal{C}(X, Y)$. Il y a équivalence entre :

- (i). A est équicontinue ;
- (ii). A est relativement compacte.

Démonstration. Commençons par supposer A relativement compacte. Soit $\varepsilon > 0$. La partie $\overline{A}$ est compacte, et recouverte par les boules ouvertes de rayon ε et de centre f , pour $f \in A$. Il existe donc des fonctions $f_1, \dots, f_p$ dans A telles que

$$A \subset \overline{A} \subset B(f_1, \varepsilon) \cup \dots \cup B(f_p, \varepsilon).$$

Les fonctions $f_1, \dots, f_p$ sont continues sur le compact X , donc uniformément continues d'après le théorème de Heine. Il existe donc $\delta > 0$ tel que pour tous $x, y \in X$,

$$d(x, y) \leq \delta \implies d(f_j(x), f_j(y)) \leq \varepsilon,$$

et ce pour tout $j \in \{1, \dots, p\}$, avec δ indépendant de j . (Puisqu'il n'y a qu'un nombre fini de valeurs de j , on peut trouver un δ qui convient pour toutes les fonctions f_j .) Soit $f \in A$ et $j \in \{1, \dots, p\}$ tel que $d(f, f_j) \leq \varepsilon$. Alors, pour $x, y \in X$ tels que $d(x, y) \leq \delta$, on a

$$d(f(x), f(y)) \leq d(f(x), f_j(x)) + d(f_j(x), f_j(y)) + d(f_j(y), f(y)) \leq 2\varepsilon + \varepsilon = 3\varepsilon.$$

Ainsi, la partie A est bien équicontinue.

Réciproquement, supposons que A est équicontinue et commençons par montrer que l'adhérence $\overline{A}$ est précompacte (c'est-à-dire recouverte par un nombre fini de boules fermées de rayon donné). Soit $\varepsilon > 0$ et $f, g \in A$. Choisissons $\delta > 0$ fourni par l'équicontinuité pour ε . Comme X est compact, il est recouvert par p boules ouvertes de rayon δ , de centres $x_1, \dots, x_p$. Supposons que $d(f(x_j), g(x_j)) \leq \varepsilon$ pour tout $j \in \{1, \dots, p\}$. Pour $x \in X$, il existe $j \in \{1, \dots, p\}$ tel que $d(x, x_j) < \delta$, ce qui donne $d(f(x), f(x_j)) \leq \varepsilon$ et $d(g(x), g(x_j)) \leq \varepsilon$. On a donc

$$d(f(x), g(x)) \leq d(f(x), f(x_j)) + d(f(x_j), g(x_j)) + d(g(x_j), g(x)) \leq 3\varepsilon.$$

Il en ressort que

$$[\forall j \in \{1, \dots, p\} \quad d(f(x_j), g(x_j)) \leq \varepsilon] \implies d(f, g) \leq 3\varepsilon.$$

De plus, comme Y est compact, il est recouvert par q boules ouvertes de rayon $\varepsilon/2$, de centres $y_1, \dots, y_q$. Pour toute application φ de $\{x_1, \dots, x_p\}$ dans $\{y_1, \dots, y_q\}$, on considère la partie

$$A_\varphi = \{f \in A \mid \forall j \in \{1, \dots, p\}, d(f(x_j), \varphi(x_j)) < \varepsilon/2\}.$$

On a donc $d(f, g) \leq 3\varepsilon$ pour toutes fonctions $f, g \in A_\varphi$. On a aussi

$$A = \bigcup_{\varphi} A_\varphi \subset \bigcup_{\varphi} \overline{B}(f_\varphi, 3\varepsilon),$$

où f_φ désigne un élément arbitraire de A_φ , et les unions sont prises sur les applications φ pour lesquelles la partie A_φ est non vide. Il y a au plus q^p applications de la sorte. Finalement, pour tout $\varepsilon > 0$, la partie A est recouverte par un nombre fini de boules fermées de rayon 3ε . Il en va de même de l'adhérence $\overline{A}$, ce qui montre que $\overline{A}$ est précompacte.

De plus, $\overline{A}$ est fermée dans $\mathcal{C}(X, Y)$ qui est complet, donc $\overline{A}$ est complète. On conclut à l'aide de la proposition ci-dessous. $\square$

Proposition 1.25. *Tout espace métrique précompact et complet est compact.*

Démonstration. Soit (E, d) un espace métrique précompact et complet, et soit $(O_i)_{i \in I}$ un recouvrement ouvert de E . On dit qu'une partie de E est non finiment recouverte si aucun sous-recouvrement fini extrait de $(O_i)_{i \in I}$ ne la recouvre. On suppose par l'absurde que E lui-même est non finiment recouvert. On va construire par récurrence une suite $(B_n)_{n \geq 1}$ de boules fermées non finiment recouvertes de la forme $B_n = \overline{B}(x_n, 2^{-n})$, dont les centres vérifient $d(x_{n+1}, x_n) \leq 2^{-n}$.

Comme E est précompact, il est recouvert par un nombre fini de boules fermées de rayon $1/2$ et l'une de ces boules $B_1 = \overline{B}(x_1, 2^{-1})$ est non finiment recouverte, puisque sinon E serait finiment recouvert. Comme B_1 est une partie de E , elle est aussi précompacte, donc on peut lui appliquer le même raisonnement avec des boules de rayon $1/4$. On obtient donc une boule $B_2 = \overline{B}(x_2, 2^{-2})$ non finiment recouverte de centre $x_2 \in B_1$, ce qui donne $d(x_2, x_1) \leq 2^{-1}$. On continue de même la construction de la suite $(B_n)_{n \geq 1}$.

La suite $(x_n)_{n \geq 1}$ est de Cauchy dans E complet, donc converge vers un certain $x \in E$. Le point x appartient à un ouvert O_i , et pour $r > 0$ suffisamment petit, on a $\overline{B}(x, r) \subset O_i$. Ensuite, pour n assez grand tel que $d(x_n, x) \leq r/2$ et $2^{-n} \leq r/2$, on a $B_n \subset \overline{B}(x, r) \subset O_i$. Cela contredit le fait que B_n est non finiment recouverte. On en déduit la compacité de E . $\square$

Le théorème d'Ascoli peut s'appliquer par exemple en analyse complexe et permet d'obtenir le théorème des familles normales, ou théorème de Montel. Si Ω est un ouvert non vide de $\mathbb{C}$, on désigne par $\mathcal{H}(\Omega)$ l'ensemble des fonctions holomorphes (dérivables au sens complexe en tout point) de Ω dans $\mathbb{C}$ muni de la topologie de la convergence uniforme sur tout compact : $f_j \rightarrow f$ signifie que $\sup_{z \in K} |f_j(z) - f(z)| \rightarrow 0$ pour tout compact $K \subset \Omega$. On peut la définir à l'aide d'une suite *exhaustive* de compacts non vides (K_n) , c'est-à-dire $K_n \subset \overset{\circ}{K}_{n+1}$ et $\bigcup_n K_n = \Omega$, par la distance

$$d_{\mathcal{H}}(f, g) = \sum_{n=1}^{\infty} 2^{-n} \min \left(1, \sup_{z \in K_n} |f(z) - g(z)| \right).$$

Cette topologie fait de $\mathcal{H}(\Omega)$ un espace de Fréchet : un espace vectoriel topologique localement convexe, métrisable par une distance complète invariante par translation. Ici, la convexité locale vient des voisinages définis par un nombre fini d'inégalités $\sup_{K_n} |f - g| < \varepsilon_n$. Une partie A de $\mathcal{H}(\Omega)$ est dite *localement bornée* si pour tout compact $K \subset \Omega$,

$$\exists M(K) > 0 \quad \forall z \in K \quad \forall f \in A \quad |f(z)| \leq M(K).$$

Le théorème de Montel dit qu'une partie A de $\mathcal{H}(\Omega)$ est relativement compacte si et seulement si elle est localement bornée.

Espaces normés de dimension finie

Référence : [6].

2.1 Normes et compacité dans $\mathbb{K}^n$

On note $\mathbb{K}$ le corps $\mathbb{R}$ ou le corps $\mathbb{C}$, munis respectivement de la valeur absolue ou du module usuels. Pour un entier $n \geq 1$, l'espace produit $\mathbb{K}^n$ est muni de la topologie produit, que l'on peut définir par la norme

$$\|x\|_\infty = \max_{1 \leq j \leq n} |x_j|, \quad \text{pour } x = (x_1, \dots, x_n).$$

On a la caractérisation suivante des compacts de $\mathbb{K}^n$.

Proposition 2.1. *Une partie de $\mathbb{K}^n$ est compacte si et seulement si elle est fermée et bornée.*

Démonstration. On a vu précédemment qu'une partie compacte d'un espace métrique est fermée et bornée. Pour la réciproque, on observe que si A désigne une partie bornée de $\mathbb{K}^n$, alors elle est incluse dans un ensemble produit de la forme $\{x \in \mathbb{K} \mid |x| \leq M\}^n$. Ce dernier ensemble est précompact, comme produit d'ensembles précompacts. Il est aussi complet, car fermé dans $\mathbb{K}^n$, qui est complet. (On peut travailler coordonnée par coordonnée pour montrer que $\mathbb{K}^n$ est complet quand $\mathbb{K}$ est complet.) Ainsi, l'ensemble produit est compact. Finalement, la partie A est compacte aussi car fermée et incluse dans un compact. $\square$

On rappelle que deux normes p et q sur $\mathbb{K}^n$ sont dites *équivalentes* s'il existe deux constantes $a, b > 0$ telles que $a p(x) \leq q(x) \leq b p(x)$ pour tout $x \in \mathbb{K}^n$. On définit ainsi une relation d'équivalence.

Proposition 2.2. *Toute norme $\|\cdot\|$ sur $\mathbb{K}^n$ est équivalente à la norme $\|\cdot\|_\infty$. Par conséquent, toutes les normes sur $\mathbb{K}^n$ sont équivalentes.*

Démonstration. On note $(e_1, \dots, e_n)$ la base canonique sur $\mathbb{K}^n$. Comme $\|\cdot\|$ est une norme, on a pour tout $x \in \mathbb{K}^n$,

$$\|x\| = \left\| \sum_{j=1}^n x_j e_j \right\| \leq \sum_{j=1}^n |x_j| \|e_j\| \leq \|x\|_\infty \underbrace{\sum_{j=1}^n \|e_j\|}_{=b}.$$

Cela montre en particulier que la norme $\|\cdot\|$ est continue sur $\mathbb{K}^n$. En effet, l'inégalité triangulaire implique que pour tous $x, y \in \mathbb{K}^n$,

$$\| \|x\| - \|y\| \| \leq \|x - y\| \leq b \|x - y\|_\infty.$$

De plus, la sphère unité $S = \{x \in \mathbb{K}^n \mid \|x\|_\infty = 1\}$ est fermée et bornée, donc compacte dans $\mathbb{K}^n$. La norme $\|\cdot\|$ atteint alors son infimum sur S , de sorte qu'il existe $x_0 \in S$ tel que $\|x\| \geq \|x_0\| = a$ pour tout $x \in S$. On a nécessairement $a > 0$, puisque x_0 est non nul. Finalement, pour $x \in \mathbb{K}^n$ non nul, on a $\left\| \frac{x}{\|x\|_\infty} \right\| \geq a$, et donc $\|x\| \geq a \|x\|_\infty$. $\square$

2.2 Normes et compacité en dimension finie

Le cas de l'espace nul est immédiat pour les résultats de cette section et pour la continuité automatique démontrée plus loin ; dans leurs preuves, on suppose donc $\dim E \geq 1$.

Comme un espace vectoriel de dimension finie n est isomorphe à $\mathbb{K}^n$, on peut transposer comme suit les résultats ci-dessus.

Proposition 2.3. *Soit E un espace vectoriel sur $\mathbb{K}$ de dimension finie.*

- (i). *Toutes les normes sur E sont équivalentes.*
- (ii). *Une norme étant fixée, les parties compactes de E sont exactement les fermés bornés.*

Démonstration. On note f un isomorphisme de $\mathbb{K}^n$ vers E . Si p désigne une norme sur E , alors $p \circ f$ est une norme sur $\mathbb{K}^n$. Par conséquent, si p et q sont deux normes sur E , alors les normes $p \circ f$ et $q \circ f$ sont équivalentes sur $\mathbb{K}^n$. Comme f est bijective, les normes p et q sont donc équivalentes sur E .

Pour le deuxième point, on note $\|\cdot\|$ la norme considérée sur E , et on munit $\mathbb{K}^n$ de la norme $x \mapsto \|f(x)\|$. L'isomorphisme f est donc bicontinu. Il fait correspondre les parties compactes, fermées, et bornées de E avec celles de $\mathbb{K}^n$. On déduit donc la caractérisation des compacts de E de celle des compacts de $\mathbb{K}^n$. $\square$

Proposition 2.4. *Tout espace vectoriel normé de dimension finie est complet. (C'est un espace de Banach.)*

Démonstration. On sait que $\mathbb{K}^n$ muni de la norme $\|\cdot\|_\infty$ est complet puisque $\mathbb{K}$ l'est. Par équivalence des normes, $\mathbb{K}^n$ est aussi complet si muni de la norme $x \mapsto \|f(x)\|$, où $\|\cdot\|$ est la norme considérée sur E , et f est un isomorphe de $\mathbb{K}^n$ vers E . On en déduit que E est complet. $\square$

2.3 Applications

2.3.1 Théorème de Riesz

On sait qu'en dimension finie, la boule unité fermée d'un espace vectoriel normé est compacte, puisqu'elle est fermée et bornée. Cette propriété est en fait caractéristique de la dimension finie.

Lemme 2.5 (Riesz). *Soit E un espace vectoriel normé et soit F un sous-espace vectoriel fermé de E tel que $F \neq E$. Alors,*

$$\forall \varepsilon > 0 \quad \exists u \in E \quad \begin{cases} \|u\| = 1 \\ d(u, F) \geq 1 - \varepsilon. \end{cases}$$

Démonstration. On considère un point quelconque $v \in E \setminus F$. Comme F est fermé, $\{v\}$ est compact et $F \cap \{v\} = \emptyset$, de sorte que $d = d(v, F) > 0$. Pour $\varepsilon \geq 1$, tout vecteur unitaire convient. Supposons donc $0 < \varepsilon < 1$. Il existe un vecteur $f_0 \in F$ tel que $d \leq \|v - f_0\| \leq \frac{d}{1 - \varepsilon}$. Alors, le vecteur défini par $u = \frac{v - f_0}{\|v - f_0\|}$ convient. En effet, ce vecteur est de norme 1 et pour tout $f \in F$, on a

$$\|u - f\| = \left\| \frac{v - f_0}{\|v - f_0\|} - f \right\| = \frac{\|v - \overbrace{(f_0 + \|v - f_0\| f)}^{\in F}\|}{\|v - f_0\|} \geq \frac{d}{d/(1 - \varepsilon)} = 1 - \varepsilon.$$

On conclut en prenant l'infimum sur tous les vecteurs $f \in F$. $\square$

Théorème 2.6 (Riesz). *Soit E un espace vectoriel normé. Si la boule unité fermée de E est compacte, alors E est de dimension finie.*

Démonstration. On suppose par l'absurde que E est de dimension infinie. Alors, il existe une suite strictement croissante $(E_n)_{n \geq 0}$ de sous-espaces vectoriels de dimension finie de E . Le lemme précédent permet de construire une suite $(u_n)_{n \geq 1}$ de vecteurs de E , de norme égale à 1 (donc appartenant à la boule unité fermée), vérifiant $u_n \in E_n$ et $d(u_n, E_{n-1}) \geq 1/2$ pour tout $n \geq 1$. En particulier, on a $\|u_n - u_m\| \geq 1/2$ si $m < n$, de sorte que la suite $(u_n)_{n \geq 1}$ ne peut admettre de sous-suite convergente. Cela contredit la propriété de Bolzano-Weierstrass, et donc la compacité de la boule unité de E . Ainsi, l'espace E est de dimension finie. $\square$

2.3.2 Continuité automatique

Proposition 2.7. *Soit E et F deux espaces vectoriels normés sur $\mathbb{K}$, avec E de dimension finie. Alors, toute application linéaire de E dans F est continue.*

Démonstration. On note $(e_1, \dots, e_n)$ une base de E , et on munit E de la norme vérifiant $\|x\| = \max\{|x_1|, \dots, |x_n|\}$, avec $x = x_1 e_1 + \dots + x_n e_n$. En notant p la norme considérée sur F et u une application linéaire de E dans F , on peut écrire

$$p(u(x)) = p\left(\sum_{j=1}^n x_j u(e_j)\right) \leq \sum_{j=1}^n |x_j| p(u(e_j)) \leq \|x\| \sum_{j=1}^n p(u(e_j)).$$

Ainsi, u est continue quand E est muni de la norme $\|\cdot\|$. Par équivalence des normes en dimension finie, on en déduit que u reste continue quand E est muni de la norme initiale. $\square$

Calcul différentiel

Références : [1, 2, 3, 4, 5].

3.1 Différentiabilité et dérivées partielles

3.1.1 Définitions

Pour décrire les variations d'une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$, nous introduisons la notion de dérivée partielle. Dans tout ce cours, on note $(e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$ et $(\eta_1, \dots, \eta_m)$ la base canonique de $\mathbb{R}^m$. Ces espaces sont munis de leur structure euclidienne.

Définition 3.1. Soit U un ouvert de $\mathbb{R}^n$ et $f : U \rightarrow \mathbb{R}$. On dit que la fonction f admet une *dérivée partielle* en $a = (a_1, \dots, a_n) \in U$ par rapport à sa i -ème variable si la limite suivante existe :

$$\frac{\partial f}{\partial x_i}(a) := \lim_{h \rightarrow 0} \frac{f(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_n) - f(a)}{h} = \lim_{h \rightarrow 0} \frac{f(a + he_i) - f(a)}{h}.$$

Le calcul de $\frac{\partial f}{\partial x_i}(a)$ consiste donc à ne dériver l'expression de f que par rapport à la variable x_i . La notion de dérivée partielle suppose *a priori* le choix d'une base dans $\mathbb{R}^n$. La différentielle donne une description intrinsèque de ces variations au voisinage d'un point. Sa définition s'applique aussi aux fonctions à valeurs vectorielles, que nous considérons dès maintenant.

Définition 3.2. Soient U un ouvert de $\mathbb{R}^n$ et $f : U \rightarrow \mathbb{R}^m$. On dit que f est *différentiable* en $a \in U$ s'il existe une application linéaire $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ telle que

$$\lim_{\|h\| \rightarrow 0} \frac{\|f(a+h) - f(a) - L(h)\|}{\|h\|} = 0.$$

Autrement dit, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que $B(a, \delta) \subset U$ et, pour tout $h \in \mathbb{R}^n$,

$$0 < \|h\| < \delta \quad \implies \quad \|f(a+h) - f(a) - L(h)\| < \varepsilon \|h\|.$$

Dans ce cas, l'application linéaire continue L est unique. Elle est notée $df(a)$ et appelée *différentielle* de f en a .

Remarque. Une autre façon de vérifier la différentiabilité d'une fonction f en un point a est de montrer l'existence d'une fonction ε , définie au voisinage de 0 dans $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^m$, vérifiant $\varepsilon(0) = 0$ et $\varepsilon(h) \rightarrow 0$ quand $h \rightarrow 0$, et telle que pour tout $h \in \mathbb{R}^n$ de norme assez petite,

$$f(a+h) = f(a) + df(a)(h) + \|h\|\varepsilon(h),$$

c'est-à-dire de montrer que

$$f(a+h) = f(a) + df(a)(h) + o(\|h\|) \quad \text{quand} \quad \|h\| \rightarrow 0.$$

On reconnaît ici le développement limité de la fonction f à l'ordre 1 au point a .

Point de vigilance. Si f est différentiable sur un ouvert U , l'application $a \mapsto df(a)$ est définie sur U et prend ses valeurs dans $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Elle n'est en général ni constante ni linéaire. La linéarité de la différentielle concerne l'application $h \mapsto df(a)(h)$, le point a étant fixé.

Remarque. Dans le cas d'une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$, la différentielle est une application linéaire de $\mathbb{R}$ dans $\mathbb{R}$. Elle peut donc être identifiée à la multiplication par un réel noté $f'(a)$, et on a pour tout $h \in \mathbb{R}$,

$$df(a)(h) = f'(a)h.$$

Autrement dit, l'application linéaire $h \mapsto df(a)(h)$ n'est autre que la multiplication du réel h par le réel $f'(a)$.

Proposition 3.1. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est différentiable en $a \in \mathbb{R}^n$, alors la fonction f est continue au point a .

Démonstration. On a $f(a+h) - f(a) = df(a)(h) + o(\|h\|) \rightarrow 0$ quand $\|h\| \rightarrow 0$. □

Remarque. L'existence des seules dérivées partielles n'implique pas la continuité. Pour s'en convaincre, on peut considérer la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par

$$f(x, y) = \begin{cases} \left(\frac{x^2 y}{x^4 + y^2} \right)^2 & \text{si } y \neq 0 \\ 0 & \text{si } y = 0. \end{cases}$$

En effet, on a $\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$, mais la fonction f n'est pas continue en $(0, 0)$, car par exemple $f(x, x^2) = \frac{1}{4}$ pour tout $x \neq 0$.

La proposition suivante exprime les dérivées partielles à partir de la différentielle.

Proposition 3.2. Soient U un ouvert de $\mathbb{R}^n$ et $f : U \rightarrow \mathbb{R}^m$. Si la fonction f est différentiable en un point $a \in U$, alors pour tout entier $1 \leq i \leq m$, les fonctions f_i admettent des dérivées partielles, et on a pour tout $h \in \mathbb{R}^n$,

$$df(a)(h) = \sum_{i=1}^m \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(a) h_j \eta_i.$$

Démonstration. Pour tout $1 \leq i \leq m$ et tout $h \in \mathbb{R}^n$, on a

$$|f_i(a+h) - f_i(a) - (df(a)(h))_i| \leq \|f(a+h) - f(a) - df(a)(h)\|.$$

On en déduit que

$$\lim_{\|h\| \rightarrow 0} \frac{|f_i(a+h) - f_i(a) - (df(a)(h))_i|}{\|h\|} = 0.$$

En particulier, en prenant $h = te_j$ pour $1 \leq j \leq n$, il vient

$$\lim_{t \rightarrow 0} \frac{|f_i(a+te_j) - f_i(a) - t(df(a)(e_j))_i|}{|t|} = 0,$$

et donc

$$(df(a)(e_j))_i = \lim_{t \rightarrow 0} \frac{f_i(a+te_j) - f_i(a)}{t}.$$

Ainsi, pour tout $1 \leq i \leq m$, les fonctions f_i admettent pour dérivées partielles

$$\frac{\partial f_i}{\partial x_j}(a) = (df(a)(e_j))_i.$$

En outre, pour tout vecteur $h = \sum_{j=1}^n h_j e_j$ dans $\mathbb{R}^n$, on peut écrire le vecteur $df(a)(h)$ de $\mathbb{R}^m$ sous la forme

$$df(a)(h) = \sum_{i=1}^m (df(a)(h))_i \eta_i.$$

Par linéarité de la différentielle $h \mapsto df(a)(h)$, il vient

$$df(a)(h) = \sum_{i=1}^m \sum_{j=1}^n h_j (df(a)(e_j))_i \eta_i = \sum_{i=1}^m \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(a) h_j \eta_i,$$

ce qui conclut la preuve. $\square$

Ainsi, une fonction différentiable en un point a admet des dérivées partielles. La réciproque est fautive en général comme le montre l'exemple suivant.

Exemple 3.1. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie pour tout $(x, y) \in \mathbb{R}^2$ par

$$f(x, y) = \begin{cases} \frac{y^2}{x} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0. \end{cases}$$

On calcule alors

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{t \rightarrow 0} \frac{f(t, 0) - f(0, 0)}{t} = 0 \quad \text{et} \quad \frac{\partial f}{\partial y}(0, 0) = \lim_{t \rightarrow 0} \frac{f(0, t) - f(0, 0)}{t} = 0,$$

ce qui montre que f admet des dérivées partielles en $(0, 0)$. Si f était différentiable en ce point, la proposition 3.2 montrerait que sa différentielle $df(0, 0)$ est l'application linéaire nulle. Cependant,

$$\frac{f(h, h) - f(0, 0) - df(0, 0)(h, h)}{h} = 1,$$

qui ne tend pas vers 0 lorsque $h \rightarrow 0$. Donc, f n'est pas différentiable en $(0, 0)$.

Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est une fonction différentiable en un point $a \in \mathbb{R}^n$, sa différentielle $df(a)$ est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$. Les bases $(e_1, \dots, e_n)$ de $\mathbb{R}^n$ et $(\eta_1, \dots, \eta_m)$ de $\mathbb{R}^m$ étant fixées, l'application linéaire $df(a)$ peut être représentée par une matrice de taille $m \times n$ notée $J_f(a)$. D'après la proposition 3.2, la composante (i, j) de cette matrice est donnée par la j -ème dérivée partielle de la i -ème composante de f .

Définition 3.3. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ une fonction qui admet des dérivées partielles au point $a \in \mathbb{R}^n$. On appelle *matrice jacobienne* de f en a la matrice

$$J_f(a) = \left(\frac{\partial f_i}{\partial x_j}(a) \right)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}.$$

Lorsque $m = n$, la matrice jacobienne $J_f(a)$ est carrée et son déterminant

$$\det J_f(a)$$

s'appelle le *jacobien* (ou le *déterminant jacobien*) de f en a .

Le cas des fonctions $f : \mathbb{R}^n \rightarrow \mathbb{R}$ à valeurs scalaires ($m = 1$) représente un cas particulier qu'il convient de distinguer. En effet, dans ce cas, la différentielle $df(a)$ est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}$, i.e. une forme linéaire sur $\mathbb{R}^n$, et la matrice jacobienne est un vecteur ligne à n composantes qui sont les dérivées partielles de f , que l'on peut assimiler à un vecteur de $\mathbb{R}^n$.

Définition 3.4. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction qui admet des dérivées partielles au point $a \in \mathbb{R}^n$. Le vecteur de $\mathbb{R}^n$ défini par

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right)$$

est appelé *gradient* de f en a .

Remarque. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est différentiable en $a \in \mathbb{R}^n$, la proposition 3.2 montre que pour tout $h \in \mathbb{R}^n$, on a

$$df(a)(h) = J_f(a)^t h = \nabla f(a) \cdot h,$$

On utilise ici des notations usuelles pour la transposée et le produit scalaire. De plus, on dispose de la formule de Taylor suivante :

$$f(a+h) = f(a) + \nabla f(a) \cdot h + o(\|h\|) \quad \text{quand } h \rightarrow 0.$$

3.1.2 Propriétés

Pour calculer une dérivée partielle, on fixe toutes les variables sauf une. Les règles de dérivation usuelles s'appliquent donc ; nous les rappelons sans démonstration.

Propriété 3.3. Soient $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ deux fonctions, et soient $\lambda, \mu \in \mathbb{R}$.

(i). Si la fonction f ne dépend pas de la i -ème variable, alors sa dérivée partielle par rapport à cette variable existe en tout point $a \in \mathbb{R}^n$ et vérifie $\frac{\partial f}{\partial x_i}(a) = 0$.

(ii). Si les fonctions f et g admettent une dérivée partielle par rapport à la i -ème variable en un point $a \in \mathbb{R}^n$, alors la combinaison linéaire $\lambda f + \mu g$ admet une dérivée partielle en a par rapport à la i -ème variable, et on a

$$\frac{\partial(\lambda f + \mu g)}{\partial x_i}(a) = \lambda \frac{\partial f}{\partial x_i}(a) + \mu \frac{\partial g}{\partial x_i}(a).$$

(iii). Si les fonctions f et g admettent une dérivée partielle par rapport à la i -ème variable en un point $a \in \mathbb{R}^n$, alors le produit fg admet une dérivée partielle en a par rapport à la i -ème variable, et on a

$$\frac{\partial(fg)}{\partial x_i}(a) = f(a) \frac{\partial g}{\partial x_i}(a) + g(a) \frac{\partial f}{\partial x_i}(a).$$

Passons aux règles de calcul pour la différentielle.

Propriété 3.4. Soient $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ deux fonctions, et soient $\lambda, \mu \in \mathbb{R}$.

(i). Si la fonction f est constante, alors f est différentiable en tout point $a \in \mathbb{R}^n$, de différentielle $df(a) = 0$.

(ii). Si $f \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ est une application linéaire, alors f est différentiable en tout point $a \in \mathbb{R}^n$, de différentielle $df(a) = f$.

(iii). Si les fonctions f et g sont différentiables en un point $a \in \mathbb{R}^n$, alors la combinaison linéaire $\lambda f + \mu g$ est différentiable en a , de différentielle

$$d(\lambda f + \mu g)(a) = \lambda df(a) + \mu dg(a).$$

(iv). Si $m = 1$, et si les fonctions f et g sont différentiables en un point $a \in \mathbb{R}^n$, alors le produit fg est différentiable en a , de différentielle

$$d(fg)(a) = f(a)dg(a) + g(a)df(a).$$

Démonstration. (i). Si la fonction f est constante, alors il existe $c \in \mathbb{R}^m$ tel que $f(x) = c$ pour tout $x \in \mathbb{R}^n$. On a donc $f(a+h) - f(a) = 0$ pour tous $a, h \in \mathbb{R}^n$. D'où $df(a)(h) = 0$.

(ii). Si $f \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ est une application linéaire, alors $f(a+h) - f(a) = f(h)$ pour tous $a, h \in \mathbb{R}^n$. Comme f est linéaire, cela montre que $df(a)(h) = f(h)$.

(iii). Si $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ sont deux fonctions différentiables en $a \in \mathbb{R}^n$, alors il existe des fonctions $\varepsilon_1, \varepsilon_2 : \mathbb{R}^n \rightarrow \mathbb{R}^m$ telles que $\varepsilon_1(h) \rightarrow 0$ et $\varepsilon_2(h) \rightarrow 0$ quand $h \rightarrow 0$ telles que pour tout $h \in \mathbb{R}^n$ assez petit,

$$\begin{cases} f(a+h) - f(a) = df(a)(h) + \|h\| \varepsilon_1(h) \\ g(a+h) - g(a) = dg(a)(h) + \|h\| \varepsilon_2(h). \end{cases}$$

Par conséquent,

$$\lambda f(a+h) + \mu g(a+h) = \lambda f(a) + \mu g(a) + \lambda df(a)(h) + \mu dg(a)(h) + \|h\|(\lambda \varepsilon_1(h) + \mu \varepsilon_2(h)).$$

Comme la fonction $h \mapsto \lambda df(a)(h) + \mu dg(a)(h)$ est linéaire et comme $\varepsilon(h) = \lambda \varepsilon_1(h) + \mu \varepsilon_2(h) \rightarrow 0$ quand $h \rightarrow 0$, on en déduit que la combinaison linéaire $\lambda f + \mu g$ est différentiable en a avec

$$d(\lambda f + \mu g)(a) = \lambda df(a) + \mu dg(a).$$

(iv). On suppose maintenant $m = 1$. On écrit alors

$$\begin{aligned} f(a+h)g(a+h) - f(a)g(a) &= f(a+h)(g(a+h) - g(a)) + g(a)(f(a+h) - f(a)) \\ &= f(a+h)(dg(a)(h) + \|h\|\varepsilon_2(h)) + g(a)(df(a)(h) + \|h\|\varepsilon_1(h)). \end{aligned}$$

Par conséquent,

$$\begin{aligned} & \frac{|f(a+h)g(a+h) - f(a)g(a) - f(a)dg(a)(h) - g(a)df(a)(h)|}{\|h\|} \\ &= \frac{|(f(a+h) - f(a))dg(a)(h) + f(a+h)\|h\|\varepsilon_2(h) + g(a)\|h\|\varepsilon_1(h)|}{\|h\|} \\ &\leq |f(a+h) - f(a)| \|dg(a)\| + |f(a+h)| |\varepsilon_2(h)| + |g(a)| |\varepsilon_1(h)|. \end{aligned}$$

Cette borne tend vers zéro quand $\|h\| \rightarrow 0$ car la fonction f est continue en a , d'après la proposition 3.1. Ainsi, le produit fg est différentiable en a , et on a

$$d(fg)(a) = f(a)dg(a) + g(a)df(a),$$

qui est la formule attendue. □

Examinons maintenant comment la différentielle se comporte par composition.

Proposition 3.5. Soient $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ une fonction différentiable en un point $a \in \mathbb{R}^n$, et soit $g : \mathbb{R}^m \rightarrow \mathbb{R}^p$ une fonction différentiable en $f(a) \in \mathbb{R}^m$. Alors, la composée $g \circ f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ est différentiable en $a \in \mathbb{R}^n$, de différentielle

$$d(g \circ f)(a) = dg(f(a)) \circ df(a). \quad (3.1)$$

Démonstration. Notons $b = f(a)$. Il existe des fonctions $\varepsilon_1 : \mathbb{R}^n \rightarrow \mathbb{R}^m$ et $\varepsilon_2 : \mathbb{R}^m \rightarrow \mathbb{R}^p$, nulles en 0 et vérifiant $\varepsilon_1(h) \rightarrow 0$ quand $h \rightarrow 0$ et $\varepsilon_2(k) \rightarrow 0$ quand $k \rightarrow 0$, et telles que pour tous $h \in \mathbb{R}^n$ et $k \in \mathbb{R}^m$ assez petits,

$$f(a+h) - f(a) = df(a)(h) + \|h\|\varepsilon_1(h) \quad (3.2)$$

et

$$g(b+k) - g(b) = dg(b)(k) + \|k\|\varepsilon_2(k). \quad (3.3)$$

Posons $k_h = f(a+h) - f(a)$. D'après (3.2), il existe $c > 0$ tel que $\|k_h\| \leq c\|h\|$ pour tout h assez petit. En reportant dans (3.3), on obtient

$$\begin{aligned} g(f(a+h)) &= g(f(a) + k_h) = g(f(a)) + dg(f(a))(k_h) + \|k_h\|\varepsilon_2(k_h) \\ &= g(f(a)) + dg(f(a))(df(a)(h) + \|h\|\varepsilon_1(h)) + \|k_h\|\varepsilon_2(k_h). \end{aligned}$$

Par conséquent,

$$\begin{aligned} & \frac{\|g(f(a+h)) - g(f(a)) - (dg(f(a)) \circ df(a))(h)\|}{\|h\|} \\ &= \frac{\|\|h\|dg(f(a))\varepsilon_1(h) + \|k_h\|\varepsilon_2(k_h)\|}{\|h\|} \\ &\leq \|dg(f(a))\| \|\varepsilon_1(h)\| + c\|\varepsilon_2(k_h)\|, \end{aligned}$$

avec une borne qui tend vers zéro quand $\|h\| \rightarrow 0$. Cela conclut la preuve. □

Point de vigilance. Contrairement aux fonctions d'une seule variable, l'ordre d'apparition des différentielles dans la formule (3.1) est extrêmement important. En effet, comme $g \circ f : \mathbb{R}^n \rightarrow \mathbb{R}^p$, sa différentielle en $a \in \mathbb{R}^n$ (quand elle existe) doit être une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$. Comme $df(a) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, on en déduit que $df(a)(h) \in \mathbb{R}^m$ pour tout $h \in \mathbb{R}^n$. De plus, comme $dg(f(a)) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^p)$, alors $dg(f(a))(df(a)(h)) \in \mathbb{R}^p$. Cela montre bien que $dg(f(a)) \circ df(a) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$. On peut voir cela également avec les matrices jacobiniennes. En effet, on a

$$J_{g \circ f}(a) = J_g(f(a))J_f(a).$$

Comme $J_f(a)$ est une matrice $m \times n$ et comme $J_g(f(a))$ est une matrice $p \times m$, la multiplication de matrices $J_g(f(a))J_f(a)$ a un sens car $J_g(f(a))$ possède m colonnes et $J_f(a)$ possède m lignes. En revanche, le produit de ces matrices dans l'ordre inverse n'est pas toujours bien défini. Enfin, dans ce qui suit, on omet parfois le symbole $\circ$ quand on considère la composition dans $\mathcal{L}(\mathbb{R}^n)$.

Remarque. Il est intéressant d'écrire la formule (3.1) avec les dérivées partielles. Si l'on note $u \in \mathbb{R}^m$ la variable de g , alors pour tous entiers $1 \leq i \leq p$ et $1 \leq j \leq n$,

$$\frac{\partial (g \circ f)_i}{\partial x_j}(a) = \sum_{k=1}^m \frac{\partial g_i}{\partial u_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a).$$

Le calcul consiste donc à dériver g_i par rapport à la variable u_k et d'évaluer le résultat en $f(a)$, de multiplier par la dérivée de f_k par rapport à la variable x_j évaluée en a , puis de sommer par rapport à k .

En plusieurs variables, le théorème des accroissements finis prend la forme de l'inégalité suivante.

Définition 3.5. Un sous-ensemble C de $\mathbb{R}^n$ est dit *convexe* si pour tous $x, y \in C$, le segment entre x et y est inclus dans C , soit

$$[x, y] = \{ty + (1-t)x, t \in [0, 1]\} \subset C.$$

Théorème 3.6 (des accroissements finis). Soit $U \subset \mathbb{R}^n$ un ouvert convexe, et soit $f : U \rightarrow \mathbb{R}^m$ une fonction différentiable. Alors, pour tous $x, y \in U$,

$$\|f(x) - f(y)\| \leq \|x - y\| \sup_{z \in [x, y]} \|df(z)\|.$$

Démonstration. Soient $x, y \in U$. Comme l'ensemble U est convexe, le segment $[x, y]$ est bien inclus dans U . Soit $g : [0, 1] \rightarrow \mathbb{R}$ la fonction définie pour tout $t \in [0, 1]$ par le produit scalaire

$$g(t) = f(ty + (1-t)x) \cdot (f(y) - f(x)) = \sum_{i=1}^m f_i(ty + (1-t)x)(f_i(y) - f_i(x)).$$

Notons que la fonction g est bien définie sur $[0, 1]$ puisque $ty + (1-t)x \in U$, et elle est par ailleurs continue sur cet intervalle. De plus, la fonction f est différentiable sur U , et la fonction $t \mapsto ty + (1-t)x$ est dérivable sur $]0, 1[$. La formule de différentiation des fonctions composées assure que g est dérivable sur $]0, 1[$, de dérivée donnée par

$$\forall t \in]0, 1[\quad g'(t) = (df(ty + (1-t)x)(y - x)) \cdot (f(y) - f(x)).$$

D'après le théorème des accroissements finis unidimensionnel, il existe un $t_0 \in]0, 1[$ tel que $g(1) - g(0) = g'(t_0)$. On en déduit que

$$\|f(y) - f(x)\|^2 = (df(t_0y + (1-t_0)x)(y - x)) \cdot (f(y) - f(x)).$$

Par conséquent, d'après l'inégalité de Cauchy-Schwarz,

$$\begin{aligned} \|f(y) - f(x)\|^2 &\leq \|f(y) - f(x)\| \|df(t_0y + (1-t_0)x)(y - x)\| \\ &\leq \|f(y) - f(x)\| \|y - x\| \|df(t_0y + (1-t_0)x)\| \\ &\leq \|f(y) - f(x)\| \|y - x\| \sup_{z \in [x, y]} \|df(z)\|, \end{aligned}$$

ce qui conduit au résultat. $\square$

Définition 3.6. Un sous-ensemble C de $\mathbb{R}^n$ est dit *connexe* s'il n'existe aucune paire d'ouverts (U_1, U_2) tels que $C \cap U_1 \neq \emptyset$, $C \cap U_2 \neq \emptyset$ et

$$C \subset U_1 \cup U_2 \quad \text{et} \quad (C \cap U_1) \cap (C \cap U_2) = \emptyset.$$

Autrement dit, un ensemble est connexe s'il n'est pas la réunion de deux ouverts relatifs disjoints non vides. Dans le cas d'un ensemble lui-même ouvert, cela se traduit par la propriété plus intuitive suivante.

Théorème 3.7. *Un ensemble ouvert et non vide U de $\mathbb{R}^n$ est connexe si et seulement si pour tous $x, y \in U$, il existe une ligne brisée contenue dans U qui relie x et y , i.e. il existe un nombre fini de points $x_1, \dots, x_p \in U$ tels que $x = x_1$ et $y = x_p$, et les segments $[x_i, x_{i+1}]$ sont inclus dans U pour tout $1 \leq i \leq p-1$.*

Démonstration. Nous ne montrons que la condition nécessaire qui sera la seule utilisée par la suite. Fixons $x_0 \in U$, et notons V l'ensemble des points de U que l'on peut relier à x_0 par une ligne brisée dans U . Cet ensemble vérifie les propriétés suivantes.

- *V est ouvert :* Soit $a \in V$. Il existe une ligne brisée $L \subset U$ qui relie x_0 et $a \in U$. L'ensemble U étant ouvert, il existe un réel $r > 0$ tel que $B(a, r) \subset U$. Donc, pour tout $x \in B(a, r)$, on a $[a, x] \subset B(a, r) \subset U$, et donc $L' = L \cup [a, x] \subset U$ est une ligne brisée reliant x_0 et x . Cela montre que $x \in V$. Ainsi, $B(a, r) \subset V$.
- *V est fermé dans U :* Soit $b \in \bar{V} \cap U$. Il existe un réel $r > 0$ tel que $B(b, r) \subset U$, car U est ouvert, et il existe un point $a \in V$ tel que $\|a - b\| < r$, car b est adhérent à V . Soit $L \subset U$ une ligne brisée reliant x_0 et a . Alors, $L' = L \cup [a, b]$ est une ligne brisée dans U reliant x_0 et b , ce qui montre que $b \in V$, et donc $V = \bar{V} \cap U$.

On écrit alors que $U = (U \setminus V) \cup V$, avec V et $U \setminus V = U \setminus \bar{V}$ ouverts et disjoints. Comme l'ensemble U est connexe, on a soit $V = \emptyset$, soit $U \setminus V = \emptyset$. Cependant, on sait que $x_0 \in V$, donc V est non vide, ce qui impose $V = U$. Enfin, si $x, y \in U$, il existe une ligne brisée $L_x \subset U$ reliant x_0 et x , et une ligne brisée $L_y \subset U$ reliant x_0 et y . Finalement $L = L_x \cup L_y$ est une ligne brisée dans U qui relie x et y . $\square$

Théorème 3.8. *Soit $U \subset \mathbb{R}^n$ un ouvert connexe, et soit $f : U \rightarrow \mathbb{R}^m$ une fonction telle que $df(x) = 0$ pour tout $x \in U$. Alors, la fonction f est constante sur U , i.e. il existe $c \in \mathbb{R}^m$ tel que $f(x) = c$ pour tout $x \in U$.*

Démonstration. Soit $x, y \in U$. L'ensemble U est connexe, donc le théorème 3.7 assure l'existence d'une ligne brisée contenue dans U qui relie x et y . Ainsi, il existe $x_1, \dots, x_p \in U$ tels que $x = x_1$ et $y = x_p$, et les segments $[x_i, x_{i+1}]$ sont inclus dans U pour tout $1 \leq i \leq p-1$. D'après le théorème des accroissements finis, on a

$$\|f(x_{i+1}) - f(x_i)\| \leq \|x_{i+1} - x_i\| \sup_{z \in [x_i, x_{i+1}]} \|df(z)\| = 0,$$

puisque $df(z) = 0$ pour tout $z \in [x_{i+1}, x_i] \subset U$. Enfin, d'après l'inégalité triangulaire,

$$\|f(x) - f(y)\| \leq \sum_{i=1}^{p-1} \|f(x_{i+1}) - f(x_i)\| = 0,$$

ce qui montre que f est constante dans U . $\square$

L'exemple suivant montre pourquoi on ne peut pas supprimer l'hypothèse de connexité.

Exemple 3.2. On considère dans $\mathbb{R}^2$ les points $x_0 = (-1, 0)$ et $y_0 = (1, 0)$. On définit l'ensemble $U = B(x_0, 1/2) \cup B(y_0, 1/2)$. Il s'agit d'un ensemble ouvert, car c'est la réunion de deux boules ouvertes, mais non connexe puisque U est lui-même l'union de deux ouverts disjoints. Si l'on définit la fonction $f : U \rightarrow \mathbb{R}$ par $f(x, y) = -1$ si $(x, y) \in B(x_0, 1/2)$, et $f(x, y) = 1$ si $(x, y) \in B(y_0, 1/2)$, alors $df(x, y) = 0$ pour tout $(x, y) \in U$. Cependant, la fonction f n'est pas constante sur U .

3.1.3 Fonctions de classe $\mathcal{C}^1$

L'existence des dérivées partielles ne suffit pas, à elle seule, à assurer la différentiabilité. Leur continuité fournit en revanche un critère que nous allons établir.

Définition 3.7. Soit $U \subset \mathbb{R}^n$ un ouvert, et soit $f : U \rightarrow \mathbb{R}^m$. On dit que la fonction f est de classe $\mathcal{C}^1$ sur U si f est continue et différentiable sur U , et si sa différentielle $df : U \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ est continue. On note $\mathcal{C}^1(U; \mathbb{R}^m)$ l'ensemble des fonctions de classe $\mathcal{C}^1$ sur U à valeurs dans $\mathbb{R}^m$ et, plus simplement, $\mathcal{C}^1(U)$ quand $m = 1$.

Il est clair que si f est de classe $\mathcal{C}^1$ sur U , alors toutes les dérivées partielles de f sont continues sur U . Le résultat suivant montre que la réciproque est vraie.

Théorème 3.9. Soit $U \subset \mathbb{R}^n$ un ouvert, et soit $f : U \rightarrow \mathbb{R}^m$. On suppose que pour tous entiers $1 \leq i \leq m$ et $1 \leq j \leq n$, la fonction f admet des dérivées partielles $\frac{\partial f_i}{\partial x_j}$ qui sont continues sur U . Alors, f est différentiable sur U et $f \in \mathcal{C}^1(U; \mathbb{R}^m)$.

Démonstration. Pour simplifier, nous présentons la démonstration pour $m = 1$. Soit $a \in U$. Montrons que la fonction f est différentiable en a . Par continuité des dérivées partielles, on sait que pour tout $\varepsilon > 0$, il existe $r > 0$ tel que $\bar{B}(a, r) \subset U$, et pour tous $1 \leq j \leq n$ et $x \in \bar{B}(a, r)$,

$$\left| \frac{\partial f}{\partial x_j}(x) - \frac{\partial f}{\partial x_j}(a) \right| < \frac{\varepsilon}{\sqrt{n}}. \quad (3.4)$$

Pour $h \in \mathbb{R}^n$ fixé tel que $\|h\| \leq r$, on note $v^{(0)} = 0$, et pour tout $1 \leq j \leq n$,

$$v^{(j)} = h_1 e_1 + \dots + h_j e_j.$$

On peut alors écrire

$$f(a + h) - f(a) = \sum_{j=1}^n (f(a + v^{(j)}) - f(a + v^{(j-1)})). \quad (3.5)$$

Pour tout $t \in [0, 1]$, posons ensuite

$$g_j(t) = f(a + v^{(j-1)} + t h_j e_j).$$

Chaque fonction g_j est bien définie car le point $a + v^{(j-1)} + t h_j e_j$ est dans la boule $\bar{B}(a, r)$. De plus, g_j est continue sur le fermé $[0, 1]$. Par définition des dérivées partielles, elle est aussi dérivable sur l'ouvert $]0, 1[$, avec pour dérivée sur cet intervalle

$$g'_j(t) = \frac{\partial f}{\partial x_j}(a + v^{(j-1)} + t h_j e_j) h_j.$$

En vertu du théorème des accroissements finis unidimensionnel, on en déduit l'existence d'un réel $\theta_j \in [0, 1]$ tel que $g_j(1) - g_j(0) = g'_j(\theta_j)$, ce qui donne

$$f(a + v^{(j)}) - f(a + v^{(j-1)}) = \frac{\partial f}{\partial x_j}(a + v^{(j-1)} + \theta_j h_j e_j) h_j.$$

Comme le point $a + v^{(j-1)} + \theta_j h_j e_j$ est dans la boule $\bar{B}(a, r)$, la borne (3.4) donne

$$\left| \frac{\partial f}{\partial x_j}(a + v^{(j-1)} + \theta_j h_j e_j) - \frac{\partial f}{\partial x_j}(a) \right| < \frac{\varepsilon}{\sqrt{n}}.$$

On déduit alors de (3.5) que pour tout $h \in \bar{B}(0, r)$,

$$\left| f(a + h) - f(a) - \sum_{j=1}^n h_j \frac{\partial f}{\partial x_j}(a) \right| \leq \sum_{j=1}^n |h_j| \left| \frac{\partial f}{\partial x_j}(a + v^{(j-1)} + \theta_j h_j e_j) - \frac{\partial f}{\partial x_j}(a) \right| \leq \frac{\varepsilon}{\sqrt{n}} \sum_{j=1}^n |h_j| \leq \varepsilon \|h\|.$$

On a ainsi montré que la fonction f est différentiable en a et que sa différentielle vérifie pour tout $h \in \mathbb{R}^n$,

$$df(a)(h) = \sum_{j=1}^n h_j \frac{\partial f}{\partial x_j}(a).$$

Il reste à montrer que la différentielle $df : U \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ est continue. Dans ce but, on utilise (3.4) afin d'écrire que pour tous $v \in \mathbb{R}^n$ et $x \in \bar{B}(a, r)$,

$$\|df(x)(v) - df(a)(v)\| \leq \frac{\varepsilon}{\sqrt{n}} \sum_{j=1}^n |v_j| \leq \varepsilon \|v\|.$$

Par conséquent, $\|df(x) - df(a)\| \leq \varepsilon$, et donc la différentielle df est continue en a . $\square$

Le théorème 3.9 est très utile en pratique. En effet, il est généralement difficile de montrer directement qu'une fonction f est différentiable sur un ouvert U . Une possibilité consiste donc à montrer que f admet des dérivées partielles sur U , et que celles-ci sont continues sur U . Le résultat précédent assure alors la différentiabilité de f sur U .

Exemple 3.3. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction définie par

$$f(x, y) = \begin{cases} xy \frac{x^2 - y^2}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0). \end{cases}$$

La fonction f est correctement définie, continue, et elle admet des dérivées partielles continues sur $\mathbb{R}^2 \setminus \{(0, 0)\}$. Pour tout $(x, y) \neq (0, 0)$, on calcule

$$\frac{\partial f}{\partial x}(x, y) = y \frac{x^4 + 4x^2y^2 - y^4}{(x^2 + y^2)^2} \quad \text{et} \quad \frac{\partial f}{\partial y}(x, y) = x \frac{x^4 - 4x^2y^2 - y^4}{(x^2 + y^2)^2}.$$

En outre, au point $(0, 0)$, les dérivées partielles sont données par

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{x \rightarrow 0} \frac{f(x, 0) - f(0, 0)}{x - 0} = 0 \quad \text{et} \quad \frac{\partial f}{\partial y}(0, 0) = \lim_{y \rightarrow 0} \frac{f(0, y) - f(0, 0)}{y - 0} = 0.$$

Comme annoncé précédemment, les dérivées partielles $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont continues sur l'ensemble $\mathbb{R}^2 \setminus \{(0, 0)\}$.

Pour étudier leur continuité en $(0, 0)$, passons en coordonnées polaires en posant $x = r \cos \theta$ et $y = r \sin \theta$, avec $r = \sqrt{x^2 + y^2} > 0$ et $\theta \in [0, 2\pi[$. On observe que $(x, y) \rightarrow 0$ si et seulement si $r \rightarrow 0$. Ainsi,

$$\left| \frac{\partial f}{\partial x}(x, y) \right| \leq r |\sin \theta| |\cos^4 \theta + 4 \cos^2 \theta \sin^2 \theta - \sin^4 \theta| \leq 6r \xrightarrow{(x, y) \rightarrow 0} 0 = \frac{\partial f}{\partial x}(0, 0),$$

ce qui prouve que la dérivée partielle $\frac{\partial f}{\partial x}$ est continue en $(0, 0)$. On montre de même que $\frac{\partial f}{\partial y}$ est continue en $(0, 0)$.

Finalement, les deux dérivées partielles de f sont continues sur $\mathbb{R}^2$ tout entier. On en conclut que la fonction f est différentiable sur $\mathbb{R}^2$.

3.2 Dérivées partielles d'ordre supérieur

Tout comme pour les fonctions d'une seule variable, on dispose aussi de notions de dérivées d'ordre supérieur pour les fonctions de plusieurs variables. Nous nous limitons ici à la notion de dérivée partielle d'ordre deux, même s'il est possible de définir la différentielle d'ordre quelconque $k \geq 2$.

Définition 3.8. Soit $U \subset \mathbb{R}^n$ un ouvert, et soit $f : U \rightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^1$. On dit que la fonction f admet des dérivées partielles d'ordre deux en un point $a \in U$ si les fonctions $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$ admettent des dérivées partielles en a . Pour tous $1 \leq i, j \leq n$, la i -ème dérivée partielle de la j -ème dérivée partielle de f en a est notée

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right)(a).$$

De plus, on appelle *matrice hessienne* de la fonction f au point a la matrice formée des dérivées partielles d'ordre deux en a , c'est-à-dire

$$D^2 f(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)_{1 \leq i, j \leq n}.$$

Le théorème de Schwarz permet de permuter l'ordre des dérivations partielles lorsque les dérivées partielles secondes sont continues.

Théorème 3.10 (Schwarz). Soit $U \subset \mathbb{R}^n$ un ouvert, et soit $f : U \rightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ qui admet des dérivées partielles secondes sur U , lesquelles sont continues en un point $a \in U$. Alors, pour tous $1 \leq i, j \leq n$,

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

En particulier, la matrice hessienne $D^2 f(a)$ est symétrique.

Démonstration. Comme U est ouvert, on peut fixer un $r > 0$ tel que $\overline{B}(a, 2r) \subset U$.

Cas $n = 2$. Pour tout $h \in \mathbb{R}^2$ avec $\|h\| \leq r$, on pose

$$\Delta_h = (f(a+h) - f(a_1+h_1, a_2)) - (f(a_1, a_2+h_2) - f(a)),$$

ainsi que pour tout $x \in [a_1 - r, a_1 + r]$,

$$\varphi(x) = f(x, a_2 + h_2) - f(x, a_2).$$

La fonction φ ainsi définie est continue sur le fermé $[a_1 - r, a_1 + r]$, et dérivable sur l'ouvert $]a_1 - r, a_1 + r[$, avec pour dérivée sur cet intervalle

$$\varphi'(x) = \frac{\partial f}{\partial x}(x, a_2 + h_2) - \frac{\partial f}{\partial x}(x, a_2).$$

Le théorème des accroissements finis unidimensionnel montre alors l'existence d'un réel $\theta_1 \in [0, 1]$ tel que

$$\varphi(a_1 + h_1) - \varphi(a_1) = h_1 \varphi'(a_1 + \theta_1 h_1),$$

ce qui donne encore

$$\Delta_h = h_1 \left(\frac{\partial f}{\partial x}(a_1 + \theta_1 h_1, a_2 + h_2) - \frac{\partial f}{\partial x}(a_1 + \theta_1 h_1, a_2) \right).$$

On définit ensuite pour tout $y \in [a_2 - r, a_2 + r]$,

$$\psi(y) = \frac{\partial f}{\partial x}(a_1 + \theta_1 h_1, y).$$

La fonction ψ ainsi définie est continue sur le fermé $[a_2 - r, a_2 + r]$, et dérivable sur l'ouvert $]a_2 - r, a_2 + r[$, avec pour dérivée sur cet intervalle

$$\psi'(y) = \frac{\partial^2 f}{\partial y \partial x}(a_1 + \theta_1 h_1, y).$$

Le théorème des accroissements finis montre l'existence d'un réel $\theta_2 \in [0, 1]$ tel que

$$\psi(a_2 + h_2) - \psi(a_2) = h_2 \psi'(a_2 + \theta_2 h_2),$$

c'est-à-dire

$$\Delta_h = h_1 h_2 \frac{\partial^2 f}{\partial y \partial x}(a_1 + \theta_1 h_1, a_2 + \theta_2 h_2).$$

En échangeant le rôle des variables x et y , le même argument montre l'existence de deux réels $\eta_1, \eta_2 \in [0, 1]$ tels que

$$\Delta_h = h_1 h_2 \frac{\partial^2 f}{\partial x \partial y}(a_1 + \eta_1 h_1, a_2 + \eta_2 h_2)$$

Pour tout h avec $h_1 h_2 \neq 0$ et $\|h\| \leq r$, il existe donc des réels $\theta_1, \theta_2, \eta_1, \eta_2 \in [0, 1]$ tels que

$$\frac{\partial^2 f}{\partial y \partial x}(a_1 + \theta_1 h_1, a_2 + \theta_2 h_2) = \frac{\partial^2 f}{\partial x \partial y}(a_1 + \eta_1 h_1, a_2 + \eta_2 h_2).$$

Cependant, on observe que les deux vecteurs $(\theta_1 h_1, \theta_2 h_2)$ et $(\eta_1 h_1, \eta_2 h_2)$ sont de norme au plus celle de h . De plus, les dérivées partielles d'ordre deux $\frac{\partial^2 f}{\partial y \partial x}$ et $\frac{\partial^2 f}{\partial x \partial y}$ sont continues au point a . On en conclut par passage à la limite quand $h \rightarrow (0, 0)$ que

$$\frac{\partial^2 f}{\partial y \partial x}(a) = \frac{\partial^2 f}{\partial x \partial y}(a).$$

Cas $n \geq 3$. Le cas $i = j$ est immédiat. Fixons des indices distincts $1 \leq i, j \leq n$, et pour tout point $(x, y) \in \mathbb{R}^2$ avec $\|(x, y)\| \leq r$, considérons

$$g(x, y) := f(a + x e_i + y e_j).$$

La fonction g est de classe $\mathcal{C}^1$ sur $B(0, r)$, et elle admet dans cette boule des dérivées partielles d'ordres un et deux qui vérifient

$$\frac{\partial g}{\partial x}(x, y) = \frac{\partial f}{\partial x_i}(a + x e_i + y e_j) \quad \text{et} \quad \frac{\partial g}{\partial y}(x, y) = \frac{\partial f}{\partial x_j}(a + x e_i + y e_j),$$

ainsi que

$$\frac{\partial^2 g}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a + x e_i + y e_j) \quad \text{et} \quad \frac{\partial^2 g}{\partial y \partial x}(x, y) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a + x e_i + y e_j).$$

Comme les dérivées partielles $\frac{\partial^2 f}{\partial x_i \partial x_j}$ et $\frac{\partial^2 f}{\partial x_j \partial x_i}$ sont continues en a , les dérivées partielles $\frac{\partial^2 g}{\partial x \partial y}$ et $\frac{\partial^2 g}{\partial y \partial x}$ sont continues en $(0, 0)$. On déduit alors du cas précédent que

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 g}{\partial x \partial y}(0, 0) = \frac{\partial^2 g}{\partial y \partial x}(0, 0) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a),$$

ce qui conclut la preuve du théorème. □

Définition 3.9. Soit $U \subset \mathbb{R}^n$ un ouvert, et soit $f : U \rightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^1$. On dit que f est de classe $\mathcal{C}^2$ sur U si f admet des dérivées partielles d'ordre deux en tout point de U , qui sont continues sur U . Pour les fonctions à valeurs dans $\mathbb{R}^m$, cette définition s'applique composante par composante. On note $\mathcal{C}^2(U; \mathbb{R}^m)$ l'ensemble des fonctions de classe $\mathcal{C}^2$ sur U à valeurs dans $\mathbb{R}^m$ et, plus simplement $\mathcal{C}^2(U)$ quand $m = 1$.

Cette notation s'étend aux ordres supérieurs de la façon suivante :

Définition 3.10. Soit $U \subset \mathbb{R}^n$ un ouvert, et soit $f : U \rightarrow \mathbb{R}$ une fonction.

- On dit que f est de classe $\mathcal{C}^p$, pour un entier $p \in \mathbb{N}$, si f admet des dérivées partielles sur U d'ordre inférieur ou égal à p , qui sont continues sur U . On note $\mathcal{C}^p(U; \mathbb{R}^m)$ l'ensemble des fonctions de classe $\mathcal{C}^p$ sur U à valeurs dans $\mathbb{R}^m$ et, plus simplement $\mathcal{C}^p(U)$ quand $m = 1$.
- On dit que f est de classe $\mathcal{C}^\infty$ sur U si f est de classe $\mathcal{C}^p$ sur U pour tout entier $p \in \mathbb{N}$. Les ensembles de fonctions correspondants sont notés $\mathcal{C}^\infty(U; \mathbb{R}^m)$ et $\mathcal{C}^\infty(U)$, selon le modèle ci-dessus.

Les dérivées partielles secondes interviennent dans le terme quadratique de la formule de Taylor en plusieurs variables.

Théorème 3.11 (Formule de Taylor). Soit $U \subset \mathbb{R}^n$ un ouvert, et soit $f : U \rightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^2$. Pour tout $a \in U$, on a

$$f(a+h) = f(a) + \nabla f(a) \cdot h + \frac{1}{2} (D^2 f(a)h) \cdot h + o(\|h\|^2) \quad \text{quand } h \rightarrow 0.$$

Point de vigilance. Le terme $(D^2 f(a)h) \cdot h$ est le produit scalaire du vecteur $D^2 f(a)h$ avec le vecteur h . En effet, comme la matrice hessienne $D^2 f(a)$ est une matrice de taille $n \times n$, si on identifie h à un vecteur colonne de taille n , alors le produit $D^2 f(a)h$ est un vecteur colonne de taille n , identifié à un élément de $\mathbb{R}^n$. En terme de dérivées partielles secondes, on peut écrire

$$(D^2 f(a)h) \cdot h = \sum_{i=1}^n (D^2 f(a)h)_i h_i = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j.$$

Démonstration. Choisissons $r > 0$ tel que $\bar{B}(a, r) \subset U$. Pour $\|h\| < r$, la fonction $g_h(t) = f(a+th)$ est de classe $\mathcal{C}^2$ au voisinage de $[0, 1]$. La formule de Taylor avec reste intégral en une variable donne

$$f(a+h) = f(a) + \nabla f(a) \cdot h + \int_0^1 (1-t)(D^2 f(a+th)h) \cdot h dt.$$

En soustrayant le terme quadratique $\frac{1}{2}(D^2 f(a)h) \cdot h$, le reste $R(h)$ vérifie

$$|R(h)| \leq \frac{\|h\|^2}{2} \sup_{0 \leq t \leq 1} \|D^2 f(a+th) - D^2 f(a)\|.$$

La continuité de la hessienne en a assure que ce supremum tend vers zéro lorsque $h \rightarrow 0$, uniformément par rapport à la direction de h . Ainsi, $R(h) = o(\|h\|^2)$. $\square$

3.3 Points critiques et extrema

Définition 3.11. Soit $U \subset \mathbb{R}^n$ un ouvert, et soit $f : U \rightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ sur U . On dit qu'un point $a \in U$ est un *point critique* de f si $df(a) = 0$.

Définition 3.12. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction, et soit E un sous-ensemble de $\mathbb{R}^n$.

- On dit que f admet un *minimum global* sur E en un point $a \in E$ si

$$\forall y \in E \quad f(a) \leq f(y).$$

- On dit que f admet un *maximum global* sur E en un point $a \in E$ si

$$\forall y \in E \quad f(a) \geq f(y).$$

- On dit que f admet un *minimum local* sur E en un point $a \in E$ s'il existe un ouvert $V \subset \mathbb{R}^n$ contenant a tel que

$$\forall y \in E \cap V \quad f(a) \leq f(y).$$

- On dit que f admet un *maximum local* sur E en un point $a \in E$ s'il existe un ouvert $V \subset \mathbb{R}^n$ contenant a tel que

$$\forall y \in E \cap V \quad f(a) \geq f(y).$$

Nous utilisons la dénomination d'*extremum* pour désigner sans distinction un maximum ou un minimum. Nous montrons à présent une condition nécessaire qui assure que tout extremum local est forcément un point critique.

Théorème 3.12. Soit $U \subset \mathbb{R}^n$ un ouvert, et soit $f : U \rightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ sur U . Si f admet un *extremum local* sur U en un point $a \in U$, alors il s'agit d'un point critique de f , c'est-à-dire que $df(a) = 0$.

Démonstration. Supposons que f admet un minimum local sur U en $a \in U$. Alors, il existe un ouvert V de $\mathbb{R}^n$ contenant a tel que

$$\forall y \in U \cap V \quad f(a) \leq f(y).$$

Comme l'ensemble $U \cap V$ est ouvert, il existe un réel $r > 0$ tel que $\bar{B}(a, r) \subset U \cap V$. Pour tout $h \in \mathbb{R}^n$ avec $\|h\| \leq r$, et pour tout $t \in [-1, 1]$, on pose

$$g(t) = f(a + th).$$

La fonction g ainsi définie est de classe $\mathcal{C}^1$ sur l'intervalle ouvert $] -1, 1[$. De plus, elle admet un minimum global sur cet intervalle au point 0, car

$$\forall t \in] -1, 1[\quad g(0) = f(a) \leq f(a + th) = g(t).$$

Cela implique alors que $g'(0) = 0$. En effet, on a clairement

$$\frac{g(t) - g(0)}{t} \begin{cases} \geq 0 & \text{si } 0 < t < 1 \\ \leq 0 & \text{si } -1 < t < 0. \end{cases}$$

En prenant la limite à droite ou la limite à gauche quand $t \rightarrow 0$, on obtient respectivement que la dérivée $g'(0)$ est positive et négative. Elle est donc nulle. La formule de différentiation des fonctions composées montre que

$$\forall t \in] -1, 1[\quad g'(t) = df(a + th)(h).$$

En prenant $t = 0$, on en déduit que

$$\forall h \in \bar{B}(0, r) \quad df(a)(h) = 0.$$

Finalement, pour $v \in \mathbb{R}^n \setminus \{0\}$ quelconque, on applique cela au vecteur $h = r \frac{v}{\|v\|} \in \bar{B}(0, r)$, et on conclut que $df(a)(v) = 0$. Le cas $v = 0$ résulte de la linéarité. Ainsi, la différentielle $df(a)$ est nulle. $\square$

Un extremum intérieur est donc nécessairement un point critique. L'exemple suivant montre que la réciproque est fausse.

Exemple 3.4. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction définie par

$$\forall (x, y) \in \mathbb{R}^2 \quad f(x, y) = x^2 - y^2.$$

Il s'agit d'une fonction de classe $\mathcal{C}^\infty$ puisque c'est une fonction polynomiale.

Cherchons d'abord les points critiques de f . Pour ce faire, on résout l'équation $\nabla f(x, y) = 0$, c'est-à-dire

$$\frac{\partial f}{\partial x}(x, y) = 2x = 0 \quad \text{et} \quad \frac{\partial f}{\partial y}(x, y) = -2y = 0,$$

et on trouve que $(x, y) = (0, 0)$. La fonction f admet $(0, 0)$ pour unique point critique.

Par ailleurs, on constate que f n'admet ni un minimum local, ni un maximum local, sur $\mathbb{R}^2$ en $(0, 0)$. En effet, si U désigne un ouvert quelconque de $\mathbb{R}^2$ qui contient $(0, 0)$, alors on peut trouver $\varepsilon > 0$ assez petit tel que $\bar{B}((0, 0), \varepsilon) \subset U$. Alors, les deux points $(\varepsilon, 0)$ et $(0, \varepsilon)$ appartiennent à U et vérifient

$$f(\varepsilon, 0) = \varepsilon^2 > 0 = f(0, 0) \quad \text{et} \quad f(0, \varepsilon) = -\varepsilon^2 < 0 = f(0, 0).$$

Cela montre respectivement que $(0, 0)$ ne conduit ni à un maximum local, ni à un minimum local pour la fonction f .

Afin d'identifier plus précisément les extrema parmi les points critiques, il convient d'utiliser le développement de Taylor à l'ordre deux pour les fonctions de classe $\mathcal{C}^2$.

Théorème 3.13. Soit $U \subset \mathbb{R}^n$ un ouvert, et soit $f : U \rightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^2$.

- (i). Si f admet un minimum local sur U en un point $a \in U$, alors les valeurs propres de la matrice hessienne $D^2 f(a)$ sont toutes positives ou nulles.
- (ii). Si f admet un maximum local sur U en un point $a \in U$, alors les valeurs propres de la matrice hessienne $D^2 f(a)$ sont toutes négatives ou nulles.

Démonstration. Le théorème de Schwarz montre que la matrice hessienne $D^2 f(a)$ est symétrique. Elle est donc diagonalisable dans une base orthonormée de vecteurs propres. Ainsi, il existe des valeurs propres $\lambda_1 \leq \dots \leq \lambda_n$ et des vecteurs propres $v_1, \dots, v_n \in \mathbb{R}^n$ tels que la famille $(v_1, \dots, v_n)$ forme une base orthonormée de $\mathbb{R}^n$ et

$$\forall 1 \leq i \leq n \quad D^2 f(a)v_i = \lambda_i v_i.$$

D'après la formule de Taylor, on a pour tout $1 \leq i \leq n$,

$$f(a + tv_i) = f(a) + t \nabla f(a) \cdot v_i + \frac{t^2}{2} (D^2 f(a)v_i) \cdot v_i + o(t^2) \quad \text{quand } t \rightarrow 0.$$

Comme f admet un extremum local sur U en $a \in U$, le théorème 3.12 montre que $\nabla f(a) = 0$. De plus, comme les vecteurs v_i sont unitaires, on a

$$(D^2 f(a)v_i) \cdot v_i = \lambda_i v_i \cdot v_i = \lambda_i \|v_i\|^2 = \lambda_i.$$

On en déduit que pour tout $1 \leq i \leq n$,

$$f(a + tv_i) - f(a) = \frac{t^2}{2} \lambda_i + o(t^2) \quad \text{quand } t \rightarrow 0.$$

Finalement, si f admet un minimum local sur U en $a \in U$, alors il existe un ouvert V de $\mathbb{R}^n$ tel que

$$\forall x \in U \cap V \quad f(a) \leq f(x).$$

L'ensemble $U \cap V$ étant ouvert, on peut trouver un réel $r > 0$ tel que $B(a, r) \subset U \cap V$. Pour tous $1 \leq i \leq n$ et $t \in]-r, r[$, le point $a + tv_i$ est dans la boule $B(a, r)$, donc

$$0 \leq \frac{f(a + tv_i) - f(a)}{t^2} = \frac{\lambda_i}{2} + o(1) \quad \text{quand } t \rightarrow 0.$$

Par passage à la limite quand $t \rightarrow 0$, on obtient que $\lambda_i \geq 0$ pour tout $1 \leq i \leq n$.

De façon totalement similaire, si f admet un maximum local sur U en $a \in U$, les valeurs propres vérifient $\lambda_i \leq 0$ pour tout $1 \leq i \leq n$. $\square$

Lorsque la matrice hessienne est non dégénérée, le critère suivant permet de préciser la nature d'un point critique.

Théorème 3.14. Soient $U \subset \mathbb{R}^n$ un ouvert, $f : U \rightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^2$ sur U , et $a \in U$ un point critique de f .

- (i). Si les valeurs propres de la matrice hessienne $D^2 f(a)$ sont toutes strictement positives, alors f admet un minimum local sur U en $a \in U$.
- (ii). Si les valeurs propres de la matrice hessienne $D^2 f(a)$ sont toutes strictement négatives, alors f admet un maximum local sur U en $a \in U$.

Démonstration. Comme l'ensemble U est ouvert et contient a , il existe un réel $r > 0$ tel que $B(a, r) \subset U$. Notons $\lambda_1 \leq \dots \leq \lambda_n$ les valeurs propres de la matrice hessienne $D^2 f(a)$, ainsi que $v_1, \dots, v_n \in \mathbb{R}^n$ les vecteurs propres associés tels que la famille $(v_1, \dots, v_n)$ forme une base orthonormée de $\mathbb{R}^n$ et vérifie

$$\forall 1 \leq i \leq n \quad D^2 f(a)v_i = \lambda_i v_i.$$

On décompose un vecteur $h \in B(0, r)$ arbitraire sur cette base, ce qui donne des réels $h_1, \dots, h_n$ tels que $h = \sum_{i=1}^n h_i v_i$.

On a alors

$$(D^2 f(a)h) \cdot h = \sum_{i=1}^n h_i (D^2 f(a)v_i) \cdot h = \sum_{i=1}^n \lambda_i h_i v_i \cdot h = \sum_{i=1}^n \lambda_i |h_i|^2.$$

Comme le point a est un point critique de la fonction f , la formule de Taylor conduit à

$$\begin{aligned} f(a + h) - f(a) &= \nabla f(a) \cdot h + \frac{1}{2} (D^2 f(a)h) \cdot h + o(\|h\|^2) \\ &= \frac{1}{2} \sum_{i=1}^n \lambda_i |h_i|^2 + o(\|h\|^2) \quad \text{quand } h \rightarrow 0. \end{aligned}$$

Si les valeurs propres sont toutes strictement positives, on obtient

$$f(a + h) - f(a) \geq \|h\|^2 \left(\frac{\lambda_1}{2} + \varepsilon(h) \right),$$

pour une fonction ε définie au voisinage de 0 dans $\mathbb{R}^n$ qui vérifie $\varepsilon(h) \rightarrow 0$ quand $h \rightarrow 0$. Il existe donc un réel $0 < \delta \leq r$ tel que $|\varepsilon(h)| \leq \frac{\lambda_1}{2}$ pour tout $\|h\| < \delta$, ce qui implique

$$\forall h \in B(0, \delta) \quad f(a + h) - f(a) \geq 0.$$

En notant $V = B(a, \delta)$, on a donc montré que $f(a) \leq f(y)$ pour tout $y \in V$, ce qui assure que f admet un minimum local sur U en $a \in U$.

Si les valeurs propres sont toutes strictement négatives, un argument similaire montre que f admet un maximum local sur U en $a \in U$. $\square$

Remarque. En dimension $n = 2$, la matrice hessienne est une matrice symétrique de taille 2×2 , et il est facile de connaître le signe des valeurs propres λ_1 et λ_2 . Il suffit de remarquer que

$$\det D^2 f(a) = \lambda_1 \lambda_2 \quad \text{et} \quad \text{Tr } D^2 f(a) = \lambda_1 + \lambda_2.$$

On constate alors que si a est un point critique et si le déterminant de la hessienne $D^2 f(a)$ est strictement positif, alors les deux valeurs propres sont non nulles et de même signe, ce qui implique que a est un extremum local. De plus, si la trace de la hessienne est strictement positive, alors les deux valeurs propres sont strictement positives, et a est un point de minimum local. De même, si la trace de la hessienne est strictement négative, alors les deux valeurs propres sont strictement négatives et a est un point de maximum local.

3.4 Inversion locale et fonctions implicites

3.4.1 Théorème d'inversion locale

Motivation. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ telle que $f'(x_0) \neq 0$. Pour fixer les idées, supposons que $f'(x_0) > 0$. Par continuité de f' , il existe $\delta > 0$ tel que

$$\forall x \in I =]x_0 - \delta, x_0 + \delta[\quad f'(x) > 0,$$

ce qui montre que f est strictement croissante et continue sur I . La fonction f réalise donc une bijection de I sur $J = f(I)$ qui est un intervalle ouvert de $\mathbb{R}$. De plus, la fonction réciproque f^{-1} est de classe $\mathcal{C}^1$ sur J . En effet, la continuité de la réciproque et le passage à la limite dans le quotient $(x' - x)/(f(x') - f(x))$ lorsque $x' \rightarrow x$ donnent

$$\forall x \in I \quad (f^{-1})'(f(x)) = \frac{1}{f'(x)}.$$

Par conséquent, sur l'intervalle J , on a l'identité

$$(f^{-1})' = \frac{1}{f' \circ f^{-1}}.$$

Le théorème d'inversion locale étend ce résultat aux dimensions supérieures. Introduisons d'abord la notion de difféomorphisme.

Définition 3.13. Soient U et V des ouverts de $\mathbb{R}^n$, et $f : U \rightarrow V$ une fonction. On dit que f réalise un $\mathcal{C}^p$ -difféomorphisme de U sur V , pour un entier $p \in \mathbb{N}$, si

- (i). la fonction f est de classe $\mathcal{C}^p$ sur U ;
- (ii). la fonction $f : U \rightarrow V$ est bijective;
- (iii). la fonction f^{-1} est de classe $\mathcal{C}^p$ sur V .

Remarque. (1). Pour $p = 0$, on dit plus habituellement que f réalise un *homéomorphisme* de U sur V .

(2). Si $p \geq 1$, on a $f^{-1} \circ f(x) = x$ pour tout $x \in U$. Le théorème de différentiation des fonctions composées montre que

$$d(f^{-1})(f(x)) \circ df(x) = \text{Id}.$$

Ainsi, la différentielle $df(x)$ est inversible dans $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, et son inverse vérifie

$$\forall x \in U \quad (df(x))^{-1} = d(f^{-1})(f(x)),$$

ou de façon équivalente

$$\forall y \in V \quad d(f^{-1})(y) = (df(f^{-1}(y)))^{-1}.$$

Théorème 3.15 (d'inversion locale). Soient $U \subset \mathbb{R}^n$ un ouvert, et $f : U \rightarrow \mathbb{R}^n$ une fonction de classe $\mathcal{C}^p$, avec $p \in \mathbb{N}^*$. On suppose qu'il existe un point $a \in U$ tel que $df(a) \in \text{GL}_n(\mathbb{R})$. Alors, il existe un ouvert $V \subset U$ contenant a , et un ouvert $W \subset \mathbb{R}^n$ contenant $f(a)$, tels que f réalise un $\mathcal{C}^p$ -difféomorphisme de V sur W .

La preuve du théorème d'inversion locale s'appuiera sur les deux résultats préliminaires suivants.

Proposition 3.16. Soit $u \in \mathcal{L}(\mathbb{R}^n)$ une application linéaire telle que

$$\|u\| = \sup_{x \in \mathbb{R}^n \setminus \{0\}} \frac{\|u(x)\|}{\|x\|} < 1.$$

Alors, l'application $\text{Id} - u$ est inversible et son inverse vérifie

$$(\text{Id} - u)^{-1} = \sum_{k=0}^{\infty} u^k.$$

Démonstration. Comme $\|u^k\| \leq \|u\|^k$ pour tout entier $k \in \mathbb{N}$, la série $\sum u^k$ converge normalement dans $\mathcal{L}(\mathbb{R}^n)$. De plus, on calcule

$$(\text{Id} - u) \sum_{k=0}^m u^k = \sum_{k=0}^m u^k - \sum_{k=1}^{m+1} u^k = \text{Id} - u^{m+1}.$$

Comme $\|u\| < 1$, il vient $\|u^{m+1}\| \leq \|u\|^{m+1} \xrightarrow{m \rightarrow \infty} 0$, ce qui conduit à

$$(\text{Id} - u) \sum_{k=0}^{\infty} u^k = \text{Id}.$$

De la même manière, on montre que

$$\left(\sum_{k=0}^{\infty} u^k \right) (\text{Id} - u) = \text{Id}.$$

On en déduit comme annoncé que $\text{Id} - u \in \text{GL}_n(\mathbb{R})$ et $(\text{Id} - u)^{-1} = \sum_{k=0}^{\infty} u^k$. □

Proposition 3.17. L'ensemble $\text{GL}_n(\mathbb{R})$ est ouvert dans $\mathcal{L}(\mathbb{R}^n)$. De plus, l'application de passage à l'inverse $\text{inv} : u \mapsto u^{-1}$ est de classe $\mathcal{C}^\infty$ sur $\text{GL}_n(\mathbb{R})$, et vérifie

$$\forall u \in \text{GL}_n(\mathbb{R}) \quad \forall h \in \mathcal{L}(\mathbb{R}^n) \quad d \text{inv}(u)(h) = -u^{-1}hu^{-1}.$$

Démonstration. Soient $u \in \text{GL}_n(\mathbb{R})$ et $h \in \mathcal{L}(\mathbb{R}^n)$ deux applications linéaires telles que $\|h\| < \frac{1}{\|u^{-1}\|} = r$. Comme

$$u + h = u(\text{Id} + u^{-1}h) \quad \text{et} \quad \|u^{-1}h\| \leq \|u^{-1}\| \|h\| < 1,$$

on observe que l'application $\text{Id} + u^{-1}h$ est inversible, de même que $u + h$. Il en ressort que $B(u, r) \subset \text{GL}_n(\mathbb{R})$, et donc que $\text{GL}_n(\mathbb{R})$ est ouvert dans $\mathcal{L}(\mathbb{R}^n)$.

De plus,

$$(u + h)^{-1} = (\text{Id} + u^{-1}h)^{-1}u^{-1} = \sum_{k=0}^{\infty} (-1)^k (u^{-1}h)^k u^{-1} = u^{-1} - u^{-1}hu^{-1} + \sum_{k=2}^{\infty} (-1)^k (u^{-1}h)^k u^{-1}.$$

L'application qui à $h \in \mathcal{L}(\mathbb{R}^n)$ associe $-u^{-1}hu^{-1} \in \mathcal{L}(\mathbb{R}^n)$ est linéaire. De plus,

$$\left\| \sum_{k=2}^{\infty} (-1)^k (u^{-1}h)^k u^{-1} \right\| \leq \sum_{k=2}^{\infty} \|u^{-1}\|^{k+1} \|h\|^k = \|u^{-1}\|^3 \|h\|^2 \sum_{l=0}^{\infty} \|u^{-1}\|^l \|h\|^l = \frac{\|u^{-1}\|^3 \|h\|^2}{1 - \|u^{-1}\| \|h\|} = o(\|h\|),$$

quand $h \rightarrow 0$. On en déduit que l'application inv est différentiable sur $\text{GL}_n(\mathbb{R})$, de différentielle donnée par la formule annoncée.

Comme en outre $d \text{inv}(u)(h) = -\text{inv}(u) \circ h \circ \text{inv}(u)$, l'application $u \mapsto d \text{inv}(u)$ est continue. Ainsi, l'application inv est de classe $\mathcal{C}^1$ sur $\text{GL}_n(\mathbb{R})$. Cette observation permet en fait de montrer par récurrence que inv est de classe $\mathcal{C}^\infty$ sur $\text{GL}_n(\mathbb{R})$. $\square$

Démonstration du théorème d'inversion locale. Quitte à remplacer f par la fonction $x \mapsto df(a)^{-1}(f(a+x) - f(a))$ et U par l'ouvert $-a + U$, on peut faire sans perte de généralité les hypothèses suivantes : $a = 0$, $f(a) = 0$ et $df(a) = \text{Id}$. La preuve se déroule en quatre étapes.

Étape 1 : Il existe un réel $r > 0$ tel que pour tout $x \in \bar{B}(0, r)$, la différentielle $df(x)$ est inversible et vérifie $\|df(x)^{-1}\| \leq 2$. En effet, comme f est de classe $\mathcal{C}^1$ sur U et comme $df(0) = \text{Id}$, il existe un réel $r > 0$ tel que

$$\forall x \in \bar{B}(0, r) \quad \|\text{Id} - df(x)\| \leq \frac{1}{2}. \quad (3.6)$$

Pour $x \in \bar{B}(0, r)$, en appliquant la proposition 3.16 avec $u = \text{Id} - df(x)$, on obtient

$$\|df(x)^{-1}\| \leq \sum_{k=0}^{\infty} \|u\|^k \leq \sum_{k=0}^{\infty} \frac{1}{2^k} = 2. \quad (3.7)$$

Étape 2 : La fonction f réalise une bijection de l'ouvert $V = B(0, r) \cap f^{-1}(B(0, r/2))$ sur l'ouvert $W = B(0, r/2)$. Soit $y \in B(0, r/2)$. Pour montrer qu'il existe un unique $x \in B(0, r)$ tel que $y = f(x)$, considérons l'application $h_y : U \rightarrow \mathbb{R}^n$ définie par

$$\forall x \in U \quad h_y(x) = y + x - f(x).$$

Il s'agit d'une fonction de classe $\mathcal{C}^1$ sur U , de différentielle $dh_y(x) = \text{Id} - df(x)$. D'après (3.6), cette différentielle est de norme au plus $1/2$ pour tout $x \in \bar{B}(0, r)$. En vertu du théorème des accroissements finis, il vient

$$\forall x, x' \in \bar{B}(0, r) \quad \|h_y(x) - h_y(x')\| \leq \frac{1}{2} \|x - x'\|, \quad (3.8)$$

ce qui montre que h_y est contractante sur $\bar{B}(0, r)$. De plus, pour x dans cette boule,

$$\|h_y(x)\| \leq \|y\| + \|x - f(x)\| \leq \|y\| + \|h_y(x) - h_y(0)\| < \frac{r}{2} + \frac{r}{2} = r.$$

La boule compacte $\bar{B}(0, r)$ est donc stable par h_y . Le théorème du point fixe de Banach assure qu'il existe un unique $x \in \bar{B}(0, r)$ tel que $h_y(x) = x$, c'est-à-dire $y = f(x)$. Comme l'image par l'application h_y de la boule $\bar{B}(0, r)$ est en fait incluse dans $B(0, r)$, on a $x \in B(0, r)$. Aussi, comme $y \in B(0, r/2)$, on a également $x \in f^{-1}(B(0, r/2))$.

Étape 3 : L'application $g = f^{-1} : W \rightarrow V$ est lipschitzienne. Pour tous $x, x' \in B(0, r)$, en utilisant l'inégalité triangulaire et la borne (3.8), on a

$$\|x - x'\| \leq \|h_0(x) - h_0(x')\| + \|f(x) - f(x')\| \leq \frac{1}{2} \|x - x'\| + \|f(x) - f(x')\|,$$

ce qui implique

$$\|x - x'\| \leq 2\|f(x) - f(x')\|.$$

On en conclut finalement que

$$\forall y, y' \in B(0, r/2) \quad \|g(y) - g(y')\| \leq 2\|f(g(y)) - f(g(y'))\| = 2\|y - y'\|. \quad (3.9)$$

Étape 4 : L'application g est de classe $\mathcal{C}^p$ sur W . Soient un point $y \in W$ et un vecteur $k \in \mathbb{R}^n$ tels que $y + k \in W$. Il existe un unique point $x \in V$ tel que $y = f(x)$, et un unique vecteur $h \in \mathbb{R}^n$ tel que $x + h \in V$ et $f(x + h) = y + k$. On observe que $g(y) = x$ et $g(y + k) = x + h$. Posons

$$\Delta(k) = g(y + k) - g(y) - df(x)^{-1}(k) = -df(x)^{-1}(f(x + h) - f(x) - df(x)(h)).$$

Comme l'application f est différentiable en x , on déduit de (3.7) que

$$\|\Delta(k)\| \leq 2\|f(x+h) - f(x) - df(x)(h)\| = \|h\|\varepsilon(h),$$

où ε est une fonction positive ou nulle définie au voisinage de 0 dans $\mathbb{R}^n$, avec $\varepsilon(0) = 0$ et $\varepsilon(h) \rightarrow 0$ quand $h \rightarrow 0$. Cependant, la borne (3.9) assure que

$$\|h\| = \|x+h-x\| = \|g(y+k) - g(y)\| \leq 2\|k\|.$$

On en déduit que

$$\frac{\|\Delta(k)\|}{\|k\|} \leq 2\varepsilon(h) \rightarrow 0$$

quand $\|k\| \rightarrow 0$, car cela implique $\|h\| \rightarrow 0$. Comme le terme $df(x)^{-1}(k)$ est linéaire en k , on en déduit que l'application g est différentiable au point y , de différentielle

$$dg(y) = df(x)^{-1} = df(f^{-1}(y))^{-1}.$$

Enfin, comme $dg = \text{inv} \circ df \circ f^{-1}$, on en déduit que la différentielle dg est continue, comme composée de fonctions continues (voir la proposition 3.17). L'application g est donc de classe $\mathcal{C}^1$ sur W . En outre, comme df est de classe $\mathcal{C}^{p-1}$, et comme inv est de classe $\mathcal{C}^\infty$, et $g = f^{-1}$ de classe $\mathcal{C}^1$, la relation précédente implique que g est de classe $\mathcal{C}^2$ sur W lorsque $p \geq 2$. Plus généralement, si g est de classe $\mathcal{C}^k$ avec $1 \leq k < p$, alors dg est de classe $\mathcal{C}^k$, donc g est de classe $\mathcal{C}^{k+1}$. On obtient ainsi par récurrence que g est de classe $\mathcal{C}^p$ sur W . $\square$

Le théorème précédent fournit des inverses au voisinage de chaque point. Sous les hypothèses du résultat suivant, ces inverses définissent un difféomorphisme global.

Théorème 3.18 (d'inversion globale). Soient $U \subset \mathbb{R}^n$ un ouvert, et $f : U \rightarrow \mathbb{R}^n$ une fonction de classe $\mathcal{C}^1$. On suppose f injective sur U , et $df(x)$ inversible pour tout $x \in U$. Alors, $f(U)$ est un ouvert, et f réalise un $\mathcal{C}^1$ -difféomorphisme de U sur $f(U)$.

Démonstration. D'après le théorème d'inversion locale, pour tout $x \in U$, il existe un voisinage ouvert $V_x \subset U$ de x et un voisinage ouvert W_x de $f(x)$ tels que f réalise un $\mathcal{C}^1$ -difféomorphisme de V_x sur W_x . On a alors

$$U = \bigcup_{x \in U} V_x \quad \text{et} \quad f(U) = f\left(\bigcup_{x \in U} V_x\right) = \bigcup_{x \in U} f(V_x) = \bigcup_{x \in U} W_x,$$

ce qui montre que $f(U)$ est ouvert. Comme f est injective sur U , on en déduit que f est bijective de U sur $f(U)$. Par ailleurs, comme $f^{-1}|_{W_x} = (f|_{V_x})^{-1}$ est de classe $\mathcal{C}^1$ sur W_x pour tout $x \in U$, on en déduit que f^{-1} est de classe $\mathcal{C}^1$ sur $f(U)$. $\square$

3.4.2 Théorème des fonctions implicites

Le théorème des fonctions implicites permet de résoudre localement une équation cartésienne $f(x, y) = 0$ sous la forme $y = \varphi(x)$. Autrement dit, il permet (localement) de montrer qu'un ensemble de niveau peut s'écrire comme le graphe d'une fonction.

Pour toute fonction $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ et tout point $(a_1, a_2) \in \mathbb{R}^n \times \mathbb{R}^m$, on peut considérer les fonctions partielles

$$x_1 \mapsto f(x_1, a_2) \quad \text{et} \quad x_2 \mapsto f(a_1, x_2),$$

définies respectivement sur $\mathbb{R}^n$ et sur $\mathbb{R}^m$. Quand elles existent, on peut alors considérer les différentielles correspondantes

$$d_{x_1}f(a_1, a_2) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \quad \text{et} \quad d_{x_2}f(a_1, a_2) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m),$$

appelées *différentielles partielles* de f en (a_1, a_2) . Si la fonction f est différentiable au point (a_1, a_2) , alors pour tout $h = (h_1, h_2) \in \mathbb{R}^n \times \mathbb{R}^m$,

$$df(a_1, a_2)(h_1, h_2) = d_{x_1}f(a_1, a_2)(h_1) + d_{x_2}f(a_1, a_2)(h_2).$$

Théorème 3.19 (des fonctions implicites). Soient $U \subset \mathbb{R}^n$ et $V \subset \mathbb{R}^m$ des ouverts, et soit $f : U \times V \rightarrow \mathbb{R}^m$ une fonction de classe $\mathcal{C}^p$, pour un entier $p \in \mathbb{N}^*$. Soit un point $a = (a_1, a_2) \in U \times V$ tel que

$$f(a_1, a_2) = 0 \quad \text{et} \quad d_{x_2} f(a_1, a_2) \in \text{GL}_m(\mathbb{R}).$$

Alors, il existe un ouvert $U' \subset U$ contenant a_1 , un ouvert $V' \subset V$ contenant a_2 , et une fonction $\varphi : U' \rightarrow V'$ de classe $\mathcal{C}^p$ tels que

$$\begin{cases} (x_1, x_2) \in U' \times V' \\ f(x_1, x_2) = 0 \end{cases} \iff \begin{cases} x_1 \in U' \\ x_2 = \varphi(x_1). \end{cases}$$

De plus, pour tout $x_1 \in U'$, on a

$$d\varphi(x_1) = -(d_{x_2} f(x_1, \varphi(x_1)))^{-1} d_{x_1} f(x_1, \varphi(x_1)).$$

Démonstration. Nous identifions systématiquement dans cette preuve les espaces $\mathbb{R}^n \times \mathbb{R}^m$ et $\mathbb{R}^{n+m}$. Soit $F : U \times V \rightarrow \mathbb{R}^n \times \mathbb{R}^m$ la fonction définie par

$$\forall (x_1, x_2) \in U \times V \quad F(x_1, x_2) = (x_1, f(x_1, x_2)).$$

La fonction F ainsi définie est de classe $\mathcal{C}^p$ sur l'ouvert $U \times V$, et vérifie pour tout point $x = (x_1, x_2) \in U \times V$ et tout vecteur $h = (h_1, h_2) \in \mathbb{R}^n \times \mathbb{R}^m$,

$$dF(x_1, x_2)(h_1, h_2) = (h_1, df(x_1, x_2)(h_1, h_2)).$$

En appliquant cette relation au point a , on en déduit que $dF(a) \in \text{GL}_{n+m}(\mathbb{R})$. En effet, en écrivant la différentielle de f en a avec ses différentielles partielles, on observe que

$$dF(a_1, a_2)(h_1, h_2) = k \iff \begin{cases} h_1 = k_1 \\ h_2 = d_{x_2} f(a_1, a_2)^{-1}(k_2 - d_{x_1} f(a_1, a_2)(k_1)). \end{cases}$$

Comme en outre $F(a_1, a_2) = (a_1, 0)$, le théorème d'inversion locale montre l'existence d'un ouvert $\Omega \subset \mathbb{R}^n \times \mathbb{R}^m$ contenant a et d'un ouvert $\Omega' \subset \mathbb{R}^n \times \mathbb{R}^m$ contenant $(a_1, 0)$ tels que F réalise un $\mathcal{C}^p$ -difféomorphisme de Ω sur Ω' . Quitte à restreindre Ω , on peut toujours supposer que $\Omega = \tilde{U} \times V'$, avec $\tilde{U} \subset U$ un voisinage ouvert de a_1 et $V' \subset V$ un voisinage ouvert de a_2 , et que $\Omega' = F(\tilde{U} \times V')$. De plus, comme $(a_1, 0) \in \Omega'$, il existe un ouvert $U' \subset \tilde{U}$ de $\mathbb{R}^n$ contenant a_1 tel que $U' \times \{0\} \subset \Omega'$.

Comme $F(x_1, x_2) = (x_1, f(x_1, x_2))$ pour tout $(x_1, x_2) \in \Omega$, il existe une fonction $g : \Omega' \rightarrow \mathbb{R}^m$ de classe $\mathcal{C}^p$ telle que

$$\forall (y_1, y_2) \in \Omega' \quad F^{-1}(y_1, y_2) = (y_1, g(y_1, y_2)).$$

En particulier, pour tout $y_1 \in U'$, il existe un unique $(x_1, x_2) \in \tilde{U} \times V'$ tel que

$$\begin{aligned} (y_1, 0) &= F(x_1, x_2) = (x_1, f(x_1, x_2)) \\ \iff (x_1, x_2) &= F^{-1}(y_1, 0) = (y_1, g(y_1, 0)) \\ \iff y_1 &= x_1 \quad \text{et} \quad g(x_1, 0) = x_2. \end{aligned}$$

Notons $\varphi : U' \rightarrow V'$ la fonction définie par $\varphi(x_1) = g(x_1, 0)$ pour tout $x_1 \in U'$. Il s'agit d'une fonction de classe $\mathcal{C}^p$. On a alors

$$\begin{aligned} x_1 &\in U' \quad \text{et} \quad x_2 = \varphi(x_1) \\ \iff (x_1, x_2) &= (x_1, g(x_1, 0)) = F^{-1}(x_1, 0) \in U' \times V' \\ \iff (x_1, x_2) &\in U' \times V' \quad \text{et} \quad (x_1, 0) = F(x_1, x_2) = (x_1, f(x_1, x_2)) \\ \iff (x_1, x_2) &\in U' \times V' \quad \text{et} \quad f(x_1, x_2) = 0. \end{aligned}$$

Le théorème d'inversion locale montre que $dF(x) \in GL_{n+m}(\mathbb{R})$ pour tout $x \in \Omega$, ce qui implique $d_{x_2}f(x_1, x_2) \in GL_m(\mathbb{R})$. En outre, comme $f(x_1, \varphi(x_1)) = 0$ pour tout $x_1 \in U'$, on en déduit par le théorème de différentiation des fonctions composées, pour tout $x_1 \in U'$, les relations

$$\begin{aligned} d_{x_1}f(x_1, \varphi(x_1)) + d_{x_2}f(x_1, \varphi(x_1))d\varphi(x_1) &= 0, \\ \text{d'où } d\varphi(x_1) &= -(d_{x_2}f(x_1, \varphi(x_1)))^{-1}d_{x_1}f(x_1, \varphi(x_1)). \end{aligned}$$

□

Équations différentielles

Référence : [6].

4.1 Généralités

4.1.1 Définitions

Définition 4.1. Soit $n \in \mathbb{N}^*$ et E un espace de Banach. Soit une application $F : \Omega \rightarrow E$, où Ω est un ouvert de $\mathbb{R} \times E^n$. On appelle *solution* de l'équation différentielle d'ordre n

$$y^{(n)} = F(t, y, y', \dots, y^{(n-1)}) \quad (*)$$

toute application $\varphi : I \rightarrow E$, où I est un intervalle de $\mathbb{R}$, qui est n fois dérivable et telle que

$$\forall t \in I \quad \begin{cases} (t, \varphi(t), \varphi'(t), \dots, \varphi^{(n-1)}(t)) \in \Omega \\ F(t, \varphi(t), \varphi'(t), \dots, \varphi^{(n-1)}(t)) = \varphi^{(n)}(t). \end{cases}$$

Pour ramener l'équation précédente à l'ordre 1, on regroupe la fonction inconnue et ses dérivées dans un même vecteur. On définit les applications

$$\begin{aligned} \Phi &: I \rightarrow E^n \\ t &\mapsto (\varphi(t), \varphi'(t), \dots, \varphi^{(n-1)}(t)) \end{aligned}$$

ainsi que, en notant $y = (x_0, x_1, \dots, x_{n-1})$,

$$\begin{aligned} G &: \Omega \rightarrow E^n \\ (t, y) &\mapsto (x_1, \dots, x_{n-1}, F(t, y)). \end{aligned}$$

Alors φ est solution de (*) si et seulement si Φ est solution de l'équation différentielle $\Phi' = G(t, \Phi)$. L'étude des équations du premier ordre permet ainsi de traiter les ordres supérieurs.

Définition 4.2. Une fonction $\varphi : I \rightarrow E$, où I est un intervalle de $\mathbb{R}$, est dite *solution maximale* de l'équation différentielle $y^{(n)} = F(t, y, y', \dots, y^{(n-1)})$ si elle est solution et s'il n'existe pas de solution $\psi : J \rightarrow E$ de cette équation différentielle telle que $I \subsetneq J$ et $\psi|_I = \varphi$, où J est un autre intervalle de $\mathbb{R}$.

4.1.2 Problème de Cauchy, théorème de Cauchy-Lipschitz

Problème de Cauchy

Définition 4.3. Soit une équation différentielle $y^{(n)} = F(t, y, y', \dots, y^{(n-1)})$ notée (*).

Soit aussi $(t_0, x_0, \dots, x_{n-1}) \in \Omega$. On appelle *problème de Cauchy* de (*) en $(t_0, x_0, \dots, x_{n-1})$ la recherche d'une fonction $\varphi : I \rightarrow E$, où I est un intervalle de $\mathbb{R}$, contenant t_0 , solution de (*) et vérifiant

$$\varphi(t_0) = x_0, \quad \varphi'(t_0) = x_1, \quad \dots \quad \varphi^{(n-1)}(t_0) = x_{n-1}.$$

On dit qu'il y a unicité au problème de Cauchy de (*) en $(t_0, x_0, \dots, x_{n-1})$ s'il existe au moins une solution à ce problème de Cauchy et si pour toutes solutions $\varphi : I \rightarrow E$ et $\psi : J \rightarrow E$, les fonctions φ et ψ coïncident sur $I \cap J$.

Existence et unicité locale d'une solution : théorème de Cauchy-Lipschitz

Définition 4.4. On dit qu'une application $F : \Omega \rightarrow \mathbb{R}^n$, avec Ω un ouvert de $\mathbb{R} \times \mathbb{R}^n$, est *localement lipschitzienne en la seconde variable* si pour tout $(t_0, x_0) \in \Omega$, il existe un voisinage V de (t_0, x_0) et un réel $K > 0$ tels que pour tous t, x, x' avec (t, x) et (t, x') dans V ,

$$\|F(t, x) - F(t, x')\| \leq K\|x - x'\|.$$

Théorème 4.1 (Cauchy-Lipschitz). Soit $F : \Omega \rightarrow \mathbb{R}^n$ une fonction continue et localement lipschitzienne en la seconde variable. Alors, pour tout $(t_0, x_0) \in \Omega$, il existe un intervalle I voisinage de t_0 dans $\mathbb{R}$ et une application $\varphi : I \rightarrow \mathbb{R}^n$ solution de $y' = F(t, y)$ tels que $\varphi(t_0) = x_0$. De plus, il y a unicité au problème de Cauchy de cette équation différentielle en (t_0, x_0) .

Démonstration. • Existence : On utilise la notion de *cylindre de sécurité*. Soit $(t_0, x_0) \in \Omega$. On note $B_r = \overline{B}(x_0, r)$ la boule fermée de centre x_0 et de rayon $r > 0$, ainsi que $I_\alpha =]t_0 - \alpha, t_0 + \alpha[$ pour $\alpha > 0$. Notons V un voisinage de (t_0, x_0) dans Ω sur lequel F est K -lipschitzienne en la seconde variable. Notons M un majorant de $\|F\|$ sur V . (On peut choisir V relativement compact et $M = \sup_V \|F\| < \infty$.) On peut ensuite choisir $r, \alpha > 0$ tels que $I_\alpha \times B_r \subset V$ et $\alpha M < r$. On dit alors que $I_\alpha \times B_r$ est un cylindre de sécurité pour F en (t_0, x_0) . Comme F est continue, toute solution est de classe C^1 . Donc $\varphi : I_\alpha \rightarrow \mathbb{R}^n$ est solution si et seulement si pour tout $t \in I_\alpha$,

$$(t, \varphi(t)) \in \Omega \quad \text{et} \quad \varphi(t) = x_0 + \int_{t_0}^t F(u, \varphi(u)) \, du.$$

On note Γ l'ensemble des fonctions continues $\psi : I_\alpha \rightarrow \mathbb{R}^n$ telles que $\psi(I_\alpha) \subset B_r$. Pour $\psi \in \Gamma$, l'application

$$\begin{aligned} \tilde{\psi} : I_\alpha &\rightarrow \mathbb{R}^n \\ t &\mapsto x_0 + \int_{t_0}^t F(u, \psi(u)) \, du \end{aligned}$$

est telle que pour tout $t \in I_\alpha$,

$$\|\tilde{\psi}(t) - x_0\| \leq \left| \int_{t_0}^t \|F(u, \psi(u))\| \, du \right| \leq \alpha M < r,$$

de sorte que $\tilde{\psi} \in \Gamma$. On peut donc définir la suite de fonctions $(\psi_n)_{n \geq 0}$ telle que $\psi_0(t) = x_0$ pour tout $t \in I_\alpha$, puis $\psi_{n+1} = \tilde{\psi}_n$ pour tout $n \geq 0$. On remarque que pour tout $t \in I_\alpha$,

$$\|\psi_{n+1}(t) - \psi_n(t)\| \leq \left| \int_{t_0}^t \|F(u, \psi_n(u)) - F(u, \psi_{n-1}(u))\| \, du \right| \leq K \left| \int_{t_0}^t \|\psi_n(u) - \psi_{n-1}(u)\| \, du \right|.$$

On en déduit par récurrence sur $n \in \mathbb{N}$ que pour tout $t \in I_\alpha$,

$$\|\psi_{n+1}(t) - \psi_n(t)\| \leq r K^n \frac{|t - t_0|^n}{n!}.$$

En effet, si on suppose que pour tout $u \in I_\alpha$,

$$\|\psi_n(u) - \psi_{n-1}(u)\| \leq r K^{n-1} \frac{|u - t_0|^{n-1}}{(n-1)!},$$

alors on peut écrire que pour tout $t \in I_\alpha$,

$$\|\psi_{n+1}(t) - \psi_n(t)\| \leq K \left| \int_{t_0}^t r K^{n-1} \frac{|u - t_0|^{n-1}}{(n-1)!} \, du \right| = r K^n \frac{|t - t_0|^n}{n!}.$$

De plus, pour $n = 0$, on a

$$\|\psi_1(t) - \psi_0(t)\| = \left\| x_0 + \int_{t_0}^t F(u, x_0) du - x_0 \right\| \leq \alpha M < r.$$

On en déduit que la série de fonctions $\sum(\psi_{n+1} - \psi_n)$ converge normalement sur l'intervalle I_α . Ainsi, la suite (ψ_n) converge uniformément sur I_α vers une fonction limite notée ψ . On a pour tout $t \in I_\alpha$,

$$\|\psi(t) - x_0\| = \lim_{n \rightarrow \infty} \|\psi_n(t) - x_0\| \leq r,$$

de sorte que ψ appartient à Γ . De plus, pour tout $n \in \mathbb{N}$, on a

$$\psi_{n+1}(t) = x_0 + \int_{t_0}^t F(u, \psi_n(u)) du,$$

et en faisant tendre n vers l'infini, on obtient

$$\psi(t) = x_0 + \int_{t_0}^t F(u, \psi(u)) du.$$

On en conclut que ψ est solution du problème de Cauchy en (t_0, x_0) sur l'intervalle I_α .

- **Unicité :** Soient $\varphi : I \rightarrow \mathbb{R}^n$ et $\psi : J \rightarrow \mathbb{R}^n$ deux solutions du même problème de Cauchy. L'ensemble $T = \{t \in I \cap J : \varphi(t) = \psi(t)\}$ est non vide et fermé dans $I \cap J$. Montrons qu'il est aussi ouvert dans cet intervalle. Pour $t_1 \in T$, la continuité des solutions et le caractère localement lipschitzien de F donnent $\delta > 0$ et $K > 0$ tels que les deux graphes restent dans un même voisinage où la constante de Lipschitz est K , pour $t \in I \cap J$ avec $|t - t_1| < \delta$. On choisit δ assez petit pour que $K\delta < 1$. Pour un tel t , notons S_t le segment d'extrémités t_1 et t , et posons $M_t = \max_{u \in S_t} \|\varphi(u) - \psi(u)\|$. Les équations intégrales donnent, pour tout $u \in S_t$,

$$\|\varphi(u) - \psi(u)\| \leq K|u - t_1|M_t \leq K\delta M_t.$$

En prenant le maximum, on obtient $M_t \leq K\delta M_t$, donc $M_t = 0$. Ainsi, T est ouvert dans $I \cap J$. Par connexité, $T = I \cap J$. $\square$

Remarques. • L'existence reste vraie si la fonction F est seulement continue (théorème de Peano), mais il n'y a pas forcément unicité. En effet, pour l'équation différentielle

$$y' = \begin{cases} 0 & \text{si } y < 0 \\ \sqrt{y} & \text{si } y \geq 0 \end{cases}$$

sur $\mathbb{R}$, il n'y a pas unicité au problème de Cauchy puisque toutes les fonctions

$$\varphi_c(t) = \frac{(t-c)^2}{4} \mathbf{1}_{\{t > c\}}$$

sont solutions, coïncident sur $\mathbb{R}^-$ si $c > 0$, et ne sont pas identiques.

- Si F est de classe C^1 , elle est localement lipschitzienne en la seconde variable (accroissements finis), et donc le théorème de Cauchy-Lipschitz s'applique.
- Pour les équations différentielles d'ordre n , le théorème de Cauchy-Lipschitz s'exprime ainsi : si $y^{(n)} = F(t, y, y', \dots, y^{(n-1)})$ est une équation différentielle d'ordre n et si F est continue et localement lipschitzienne par rapport à la seconde variable $Y = (y_0, \dots, y_{n-1})$, alors il y a unicité au problème de Cauchy.

Corollaire 4.2. Soit $F : \Omega \rightarrow \mathbb{R}^n$, où Ω est un ouvert de $\mathbb{R} \times \mathbb{R}^n$, une fonction continue et localement lipschitzienne en la seconde variable. Alors, pour tout $(t_0, x_0) \in \Omega$, il existe une unique solution maximale φ de l'équation différentielle $y' = F(t, y)$ prenant la valeur x_0 en t_0 . Cette solution est définie sur un intervalle ouvert de $\mathbb{R}$.

4.2 Équations différentielles linéaires

On se limite au cas de la dimension finie. On désigne par I un intervalle de $\mathbb{R}$, par $\mathbb{K}$ le corps $\mathbb{R}$ ou $\mathbb{C}$, et par n un entier strictement positif. On identifie un élément de $\mathbb{K}^n$ à un vecteur colonne.

4.2.1 Généralités

Définition 4.5. Soit $p \in \mathbb{N}^*$. Toute équation différentielle sur $\mathbb{K}^n$ d'ordre p du type

$$Y^{(p)} = A_{p-1}(t)Y^{(p-1)} + \dots + A_0(t)Y + B(t), \quad (\text{L})$$

où $A_{p-1}, \dots, A_0$ sont des fonctions continues de I dans $\mathcal{M}_n(\mathbb{K})$ et B est une fonction continue de I dans $\mathbb{K}^n$, est appelée *équation différentielle linéaire* d'ordre p . Lorsque la fonction B est identiquement nulle sur I , l'équation différentielle linéaire (L) est dite *homogène*.

On peut ramener toute équation différentielle d'ordre p à une équation différentielle d'ordre 1. En effet, (L) est équivalente à

$$\frac{d}{dt} \begin{pmatrix} Y \\ Y' \\ \vdots \\ Y^{(p-1)} \end{pmatrix} = \begin{pmatrix} 0 & I_n & & (0) \\ \vdots & & \ddots & \\ 0 & (0) & & I_n \\ A_0(t) & \dots & \dots & A_{p-1}(t) \end{pmatrix} \begin{pmatrix} Y \\ Y' \\ \vdots \\ Y^{(p-1)} \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ B(t) \end{pmatrix},$$

qui est une équation d'ordre 1 où la fonction inconnue est maintenant à valeurs dans $\mathbb{K}^{np}$.

Remarque. Si l'équation différentielle porte sur des fonctions à valeurs dans $\mathbb{K}^n$, pour $n \geq 2$, on parle de système différentiel linéaire. Si $n = 1$, on parle d'équation différentielle scalaire.

Solutions des équations différentielles linéaires

Leurs solutions maximales sont définies sur tout l'intervalle I où les fonctions de l'équations sont définies. Ce résultat sera prouvé ultérieurement.

Théorème 4.3. Soit une équation différentielle linéaire $Y' = A(t)Y + B(t)$, où $A : I \rightarrow \mathcal{M}_n(\mathbb{K})$ et $B : I \rightarrow \mathbb{K}^n$ sont des fonctions continues. Alors, pour tout $t_0 \in I$ et tout $X_0 \in \mathbb{K}^n$, il existe une unique solution V de l'équation différentielle définie sur I tout entier, telle que $V(t_0) = X_0$.

Remarque. Pour une équation différentielle linéaire d'ordre p , on a : pour tout $t_0 \in I$ et pour tous $X_0, \dots, X_{p-1} \in \mathbb{K}^n$, il existe une unique solution Y de (L) définie sur I tout entier telle que $Y(t_0) = X_0, \dots, Y^{(p-1)}(t_0) = X_{p-1}$.

Théorème 4.4. Soit $A : I \rightarrow \mathcal{M}_n(\mathbb{K})$ une fonction continue. L'ensemble $\mathcal{S}_H$ des solutions de l'équation linéaire homogène $\frac{dY}{dt} = A(t)Y$, notée (H), est un sous-espace vectoriel de dimension n du $\mathbb{K}$ -espace vectoriel $C^1(I, \mathbb{K}^n)$.

Démonstration. La forme de (H) montre que $\mathcal{S}_H$ est un sous-espace vectoriel de $C^1(I, \mathbb{K}^n)$. De plus, pour $t_0 \in I$ fixé, l'application $\Phi : \mathcal{S}_H \rightarrow \mathbb{K}^n : V \mapsto V(t_0)$ est linéaire. D'après le théorème précédent, c'est un isomorphisme. D'où le résultat. $\square$

Remarques. • L'ensemble des solutions de l'équation différentielle $Y' = A(t)Y + B(t)$, notée (L), est un espace affine de dimension n . En effet, si V_0 est une solution particulière de (L), alors les solutions de (L) sont de la forme $V_0 + V$, où V est solution de (H).

• Pour une équation différentielle linéaire homogène d'ordre p , les solutions forment un espace vectoriel de dimension np . Dans le cas non homogène, elles forment un espace affine de même dimension.

Wronskien

On considère une équation différentielle linéaire homogène d'ordre 1 sur $\mathbb{K}^n$ de la forme $Y' = A(t)Y$, où $A : I \rightarrow \mathcal{M}_n(\mathbb{K})$ est continue.

Définition 4.6. Soit $V_1, \dots, V_n$ des solutions de (H). On appelle *wronskien* de $V_1, \dots, V_n$ l'application wronskien : $I \rightarrow \mathbb{K} : t \mapsto \det(V_1(t), \dots, V_n(t))$.

Remarque. Si (H_p) désigne l'équation différentielle $y^{(p)} = a_{p-1}(t)y^{(p-1)} + \dots + a_0(t)y$, où $a_i : I \rightarrow \mathbb{K}$ est continue, alors v est solution de (H_p) si et seulement si $\begin{pmatrix} v \\ \vdots \\ v^{(p-1)} \end{pmatrix}$ est solution du système associé (H_1) , une équation d'ordre 1 sur $\mathbb{K}^p$. On appelle wronskien de p solutions $v_1, \dots, v_p$ de (H_p) le wronskien de (H_1) des solutions $\begin{pmatrix} v_i \\ \vdots \\ v_i^{(p-1)} \end{pmatrix}$, pour $1 \leq i \leq p$, c'est-à-dire

$$\text{wronskien}(v_1, \dots, v_p)(t) = \begin{vmatrix} v_1(t) & \dots & v_p(t) \\ \vdots & & \vdots \\ v_1^{(p-1)}(t) & \dots & v_p^{(p-1)}(t) \end{vmatrix}.$$

Par exemple, le wronskien de deux solutions u et v d'une équation différentielle linéaire homogène d'ordre 2, de la forme $y'' = p(t)y' + q(t)y$ est

$$\begin{vmatrix} u & v \\ u' & v' \end{vmatrix} = u v' - u' v.$$

Proposition 4.5. Soient $V_1, \dots, V_n$ des solutions de (H). Le rang des vecteurs $V_1(t), \dots, V_n(t)$ est indépendant de $t \in I$.

Démonstration. Cela vient du fait que $\Phi : \mathcal{S}_H \rightarrow \mathbb{K}^n : V \mapsto V(t_0)$ est un isomorphisme. □

Corollaire 4.6. Des solutions $V_1, \dots, V_n$ de (H) forment une base de solutions de (H) si et seulement s'il existe $t_0 \in I$ tel que $\text{wronskien}(V_1, \dots, V_n)(t_0) \neq 0$, et dans ce cas, on a $\text{wronskien}(V_1, \dots, V_n)(t) \neq 0$ pour tout $t \in I$.

Résolution de l'équation différentielle linéaire d'ordre 1 sur $\mathbb{K}$

- (H) : $y' = a(t)y$, où $a : I \rightarrow \mathbb{K}$ est continue. Les solutions forment un $\mathbb{K}$ -espace vectoriel de dimension 1 et sont proportionnelles à $t \mapsto e^{\psi(t)}$, où ψ est une primitive de a sur l'intervalle I .
- (L) : $y' = a(t)y + b(t)$, où $a, b : I \rightarrow \mathbb{K}$ sont continues. On applique la méthode de variation de la constante. On cherche les solutions de la forme $t \mapsto \lambda(t)e^{\psi(t)}$, où λ est une fonction inconnue de classe C^1 . On voit que $\lambda(t)e^{\psi(t)}$ est solution si et seulement si $\lambda'(t)e^{\psi(t)} = b(t)$ pour tout $t \in I$, c'est-à-dire que λ est une primitive de $e^{-\psi(t)}b(t)$ sur I . Enfin, les solutions générales de (L) sont sommes d'une solution particulière de (L) et des solutions de (H).

Résolution d'un système linéaire d'ordre 1

Pour (H) : $Y' = A(t)Y$, avec $n \geq 2$, il n'y a pas de méthode générale. Cependant, si on connaît n solutions $V_1, \dots, V_n$ de l'équation (H) linéairement indépendantes, alors la solution générale de (H) est $\sum \lambda_i V_i$, avec $\lambda_i \in \mathbb{K}$. On peut ensuite résoudre le système différentiel inhomogène (L) : $Y' = A(t)Y + B(t)$ par la méthode de variation des

constantes. On cherche les solutions de (L) sous la forme $V(t) = \sum_i \lambda_i(t) V_i(t)$, où les $\lambda_i : I \rightarrow \mathbb{K}$ sont des fonctions inconnues de classe C^1 . Par dérivation, V est solution de (L) si et seulement si

$$\sum_{i=1}^n \lambda_i'(t) V_i(t) + \sum_{i=1}^n \lambda_i(t) V_i'(t) = \sum_{i=1}^n \lambda_i(t) A(t) V_i(t) + B(t).$$

Comme $V_i'(t) = A(t) V_i(t)$, on a $\sum_i \lambda_i'(t) V_i(t) = B(t)$. Comme les V_i sont linéairement indépendants, on peut résoudre le système puis intégrer pour trouver les λ_i .

Cas particulier des équations différentielles linéaires d'ordre 2 sur $\mathbb{K}$

Supposons connues deux solutions linéairement indépendantes u et v de (H) : $y'' = a(t) y'(t) + b(t) y$ et cherchons la solution générale de (L) : $y'' = a(t) y'(t) + b(t) y + c(t)$. On se ramène à un système d'ordre 1 en posant $Y = \begin{pmatrix} y \\ y' \end{pmatrix}$, donc (H) et (L) deviennent respectivement $(H_1) : Y' = A(t) Y$ et $(L_1) : Y' = A(t) Y + C(t)$, avec

$$A(t) = \begin{pmatrix} 0 & 1 \\ b(t) & a(t) \end{pmatrix} \quad \text{et} \quad C(t) = \begin{pmatrix} 0 \\ c(t) \end{pmatrix}.$$

Pour trouver les solutions de (L_1) , on pose $W(t) = \lambda(t) U(t) + \mu(t) V(t)$, avec $U(t) = \begin{pmatrix} u(t) \\ u'(t) \end{pmatrix}$ et $V(t) = \begin{pmatrix} v(t) \\ v'(t) \end{pmatrix}$. On a vu que W est solution de (L_1) si et seulement si $\lambda'(t) U(t) + \mu'(t) V(t) = C(t)$ pour tout $t \in I$. Pour que $W = (y, y')$ avec $y = \lambda u + \mu v$, on impose la condition auxiliaire $\lambda' u + \mu' v = 0$. Sous cette condition, y est solution de (L) si et seulement si $\lambda' u' + \mu' v' = c$. On est donc conduit au système

$$\begin{cases} \lambda' u + \mu' v = 0 \\ \lambda' u' + \mu' v' = c(t). \end{cases}$$

Toutes les solutions de (L) s'obtiennent ainsi : pour une solution y , les identités $y = \lambda u + \mu v$ et $y' = \lambda u' + \mu v'$ déterminent des fonctions λ et μ de classe C^1 , puisque le wronskien de u et v ne s'annule pas. En dérivant ces deux identités, on retrouve les deux équations du système précédent.

Abaissement de l'ordre dans une équation linéaire sur $\mathbb{K}$

La méthode d'abaissement de l'ordre s'applique à une équation de la forme

$$y^{(n)} = a_{n-1}(t) y^{(n-1)} + \dots + a_1(t) y' + a_0(t) y, \quad (L)$$

où les fonctions $a_i : I \rightarrow \mathbb{K}$ sont continues. Si on connaît une solution φ de (L) qui ne s'annule pas sur un sous-intervalle ouvert $J \subset I$, on peut y abaisser l'ordre de (L) : toute solution s'écrit $f = g \varphi$, où $g = f/\varphi$ est de classe C^n sur J . Alors f est solution de (L) si et seulement si

$$\sum_{k=0}^n \binom{n}{k} g^{(k)} \varphi^{(n-k)} = a_{n-1}(t) \sum_{k=0}^{n-1} \binom{n-1}{k} g^{(k)} \varphi^{(n-1-k)} + \dots + a_1(t) (g \varphi' + g' \varphi) + a_0(t) g \varphi,$$

et comme φ est solution de (L), tous les termes contenant g s'éliminent, ce qui donne

$$\sum_{k=1}^n \binom{n}{k} g^{(k)} \varphi^{(n-k)} = a_{n-1}(t) \sum_{k=1}^{n-1} \binom{n-1}{k} g^{(k)} \varphi^{(n-1-k)} + \dots + a_1(t) g' \varphi,$$

c'est-à-dire que g' est solution d'une équation différentielle d'ordre $n-1$ (abaissement de l'ordre).

Par exemple, à l'ordre 2, on considère $(L_2) : y'' = a(t) y' + b(t) y$. Si on connaît une solution φ , alors $f = g \varphi$ est solution de (L_2) si et seulement si $2g' \varphi' + g'' \varphi = a(t) g' \varphi$, qui est une équation linéaire d'ordre 1 en g' , que l'on sait intégrer, ce qui conduit à toutes les solutions de (L_2) sur chaque intervalle où φ ne s'annule pas.

4.2.2 Équations linéaires à coefficients constants

Si $A(t)$ est une matrice constante A , il est possible d'intégrer $Y' = AY$ avec l'exponentielle de matrice $\exp(A) = e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}$.

Proposition 4.7. Soit $A \in \mathcal{M}_n(\mathbb{K})$. L'équation différentielle (H) : $Y' = AY$ a ses solutions maximales définies sur $\mathbb{R}$, et la solution prenant une valeur donnée $V_0 \in \mathbb{K}^n$ en $t = 0$ est $V(t) = e^{tA}V_0$.

Démonstration. On a $V(0) = e^{0A}V_0 = V_0$. De plus, $V(t) = \sum_{n=0}^{\infty} \frac{t^n A^n V_0}{n!}$. La fonction V est de classe C^1 sur $\mathbb{R}$, par convergence normale de la série des dérivées sur tout segment, et pour tout $t \in \mathbb{R}$, on a

$$V'(t) = \sum_{n=1}^{\infty} \frac{t^{n-1}}{(n-1)!} A^n V_0 = A \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n V_0 = A V(t). \quad \square$$

Remarques. • En réduisant la matrice A , on peut avoir des renseignements sur les solutions de (H).

- Si on travaille sur $\mathbb{K} = \mathbb{R}$, on peut réduire A dans $\mathbb{C}$ puis écrire les solutions de (H) sous la forme $\varphi(t) + \overline{\varphi(t)}$, où φ est une solution complexe de (H).

Un cas particulièrement intéressant est celui des équations différentielles linéaires homogènes d'ordre p sur $\mathbb{K}$, à coefficients constants :

$$y^{(p)} + a_1 y^{(p-1)} + \dots + a_p y = 0 \quad (\text{E})$$

On se ramène au cas matriciel, avec $\mathbb{K} = \mathbb{C}$, et on a le résultat suivant. On considère $P = X^p + a_1 X^{p-1} + \dots + a_p$, le polynôme caractéristique de (E), que l'on factorise sous la forme $\prod_{j=1}^r (X - \lambda_j)^{m_j}$. Les solutions de (E) sont les applications $t \mapsto \sum_{j=1}^r e^{\lambda_j t} P_j(t)$, où les P_j sont des polynômes de degré $< m_j$. Dans le cas particulier de l'ordre 2, on a l'équation (H) : $y'' + a y' + b y = 0$ et le polynôme $P = X^2 + a X + b$. Alors :

- si P a deux racines distinctes r_1 et r_2 , les solutions de (H) sont les fonctions $t \mapsto \lambda_1 e^{r_1 t} + \lambda_2 e^{r_2 t}$, avec $\lambda_1, \lambda_2 \in \mathbb{C}$ (si on est sur $\mathbb{R}$, avec r_1 et r_2 complexes, on conjugue les expressions) ;
- si P a une racine double r , les solutions sont les fonctions de la forme $t \mapsto (\lambda t + \mu) e^{rt}$, avec $\lambda, \mu \in \mathbb{K}$.

4.3 Équations différentielles non linéaires

Les méthodes qui suivent donnent des solutions sur des intervalles où les changements de variables utilisés sont admissibles. Il faut ensuite examiner leurs prolongements et les raccords possibles pour déterminer les solutions maximales.

4.3.1 Équations incomplètes

Il s'agit des équations dans lesquelles x ou y ne figure pas.

Équations de la forme $F(x, y') = 0$

On note $\Gamma = \{(X, Y) \in \mathbb{R}^2 \mid F(X, Y) = 0\}$. On suppose avoir un paramétrage de Γ de la forme $t \mapsto (\varphi(t), \psi(t))$, avec φ un C^1 -difféomorphisme et ψ continue. Alors $F(x, y') = 0$ s'écrit aussi

$$\begin{cases} x = \varphi(t) \\ y' = \psi(t), \end{cases}$$

donc $dy = \psi(t) dx = \psi(t) \varphi'(t) dt$, et le graphe des solutions peut être paramétré par

$$x = \varphi(t) \quad \text{et} \quad y = \int_{t_0}^t \psi(u) \varphi'(u) du + y_0.$$

Équations de la forme $F(y, y') = 0$

En conservant les mêmes notations, l'équation s'écrit désormais

$$\begin{cases} y = \varphi(t) \\ y' = \psi(t) \end{cases} \implies \begin{cases} dy = \psi(t) dx \\ \varphi'(t) dt = \psi(t) dx \end{cases} \implies dx = \frac{\varphi'(t)}{\psi(t)} dt,$$

et ensuite on intègre pour obtenir un paramétrage en fonction de t du graphe des solutions. On remarque en outre que si la fonction ψ s'annule en un point t_0 , alors la fonction constante $t \mapsto \varphi(t_0)$ est solution. Elle peut se raccorder à des solutions dont la dérivée s'annule en t_0 .

Équations de la forme $F(y, y', y'') = 0$

On cherche des solutions $x \mapsto \varphi(x)$ telles que φ' ne s'annule pas, de sorte que φ soit un C^2 -difféomorphisme. On cherche ensuite à paramétrer le graphe des solutions avec la variable $y = \varphi(x)$. On pose $p = y' = \frac{dy}{dx}$ et on regarde p comme une fonction de y . On a

$$y' = p \quad \text{et} \quad y'' = \frac{dp}{dx} = \frac{dy}{dx} \cdot \frac{dp}{dy} = y' \frac{dp}{dy},$$

de sorte que la fonction p dépendante de y vérifie $F\left(y, p, p \frac{dp}{dy}\right) = 0$, qui est du premier ordre en p . Si on sait la résoudre, on écrit $dx = \frac{dy}{p(y)}$ puis, en intégrant, on trouve x en fonction de y .

Exemple. On considère l'équation $y'' + (y y')^3 = 0$. On pose $p = y'$ et on regarde p comme fonction de y . En observant comme ci-dessus que $y'' = p \frac{dp}{dy}$, l'équation devient $p \frac{dp}{dy} + (yp)^3 = 0$, donc $\frac{dp}{p^2} = -y^3 dy$. En intégrant, on trouve $\frac{1}{p} = \frac{y^4}{4} + \alpha$, avec $\alpha \in \mathbb{R}$, puis $p = \frac{dy}{dx} = \frac{1}{y^4/4 + \alpha}$, soit encore $\left(\frac{y^4}{4} + \alpha\right) dy = dx$, puis après intégration $\frac{y^5}{20} + \alpha y + \beta = x$, avec $\alpha, \beta \in \mathbb{R}$.

Remarques. • S'il existe y_0 tel que $F(y_0, 0, 0) = 0$, alors la fonction constante $y = y_0$ est solution et peut se raccorder avec les solutions trouvées précédemment.

• On peut généraliser cette méthode aux équations incomplètes d'ordre n du type $F(y, y', \dots, y^{(n)}) = 0$ pour abaisser d'une unité l'ordre de l'équation.

4.3.2 Équations homogènes

Il s'agit d'équations différentielles qui peuvent s'exprimer en fonction de y' et de y/x .

Équations homogènes résolues

Ce sont les équations du type $y' = f(y/x)$. On pose $t = y/x$, donc $y = tx$, puis $f(t) dx = dy = x dt + t dx$. D'où $(f(t) - t) dx = x dt$, que l'on sait intégrer.

Remarque. S'il existe t_0 tel que $f(t_0) = t_0$, alors $y = t_0 x$ est solution.

Équations homogènes non résolues

Ce sont les équations du type $F(y/x, y') = 0$. On suppose connu un paramétrage $t \mapsto (u(t), v(t))$ de la courbe $\Gamma = \{(U, V) \mid F(U, V) = 0\}$, où u est de classe C^1 et v est continue. On écrit $y/x = u(t)$ et $y' = v(t)$, donc

$$dy = v(t) dx \quad \text{et} \quad \frac{dy}{dt} = \frac{d}{dt}(u(t)x) = u'(t)x + u(t) \frac{dx}{dt}.$$

On en déduit que $dy = u'(t)x dt + u(t) dx$, ce qui donne

$$(v(t) - u(t)) dx = u'(t)x dt.$$

- Sur un intervalle où la fonction $v - u$ ne s'annule pas, on intègre

$$\frac{dx}{x} = \frac{u'(t)}{v(t) - u(t)} dt.$$

- S'il existe t_0 tel que $u(t_0) = v(t_0)$, alors $x \mapsto u(t_0)x$ est solution.

4.3.3 Étude d'équations particulières

Équations de Bernoulli

Il s'agit des équations de la forme $y' = a(t)y + b(t)y^\alpha$, avec $\alpha \in \mathbb{R} \setminus \{0, 1\}$, et a et b deux fonctions continues d'un intervalle $I \subset \mathbb{R}$ dans $\mathbb{R}$. (Si α vaut 0 ou 1, on a des équations différentielles linéaires.) On cherche des solutions qui ne s'annulent pas. Si α n'est pas entier, on cherche des solutions strictement positives. On peut écrire

$$\frac{y'}{y^\alpha} = a(t)\frac{1}{y^{\alpha-1}} + b(t),$$

c'est-à-dire, en posant $z = y^{1-\alpha}$,

$$\frac{1}{1-\alpha} z' = a(t)z + b(t),$$

qui est une équation différentielle linéaire d'ordre 1 que l'on sait résoudre.

Remarque. Si $\alpha > 0$, la fonction nulle est solution. Si de plus $\alpha \geq 1$, le théorème de Cauchy-Lipschitz s'applique et montre qu'aucune autre solution ne s'annule. Si $0 < \alpha < 1$, on peut faire des raccords de classe C^1 avec la solution nulle.

Équations de Riccati

Il s'agit des équations de la forme $y' = a(t)y^2 + b(t)y + c(t)$, où a, b, c sont des fonctions continues d'un intervalle J dans $\mathbb{R}$. Si on connaît une solution particulière φ_0 , on pose $y = \varphi_0 + z$ et en remplaçant dans l'équation, on obtient une équation de Bernoulli :

$$z' = (2a(t)\varphi_0(t) + b(t))z + a(t)z^2.$$

D'après le théorème de Cauchy-Lipschitz, aucune solution de cette nouvelle équation ne s'annule, en dehors de la solution nulle. On pose $z = 1/u$ et on se ramène à une équation différentielle du premier ordre. On remarque qu'il n'y a pas de raccords de solutions à faire ici.

Équations de Lagrange

Ces équations sont de la forme $y = a(y')t + b(y')$, avec a et b des fonctions de classe C^1 sur un intervalle J de $\mathbb{R}$. On cherche les solutions affines. Elles sont de la forme $t \mapsto mt + b(m)$, où $a(m) = m$. On cherche ensuite les solutions C^2 telles que y'' ne s'annule pas. On cherche à paramétrer le graphe d'une telle solution avec la variable admissible $p = y'$. On dérive par rapport à t l'égalité $y = a(p)t + b(p)$, et donc

$$p = \frac{dy}{dt} = a'(p)\frac{dp}{dt}t + a(p) + b'(p)\frac{dp}{dt},$$

puis $(p - a(p))\frac{dt}{dp} = a'(p)t + b'(p)$. La dérivation donne $p - a(p) = (a'(p)t + b'(p))y''$. Ainsi, $y'' = 0$ implique $p = a(p)$, mais la réciproque est fautive en général. Sur les intervalles où $p - a(p) \neq 0$, l'équation en $t(p)$ est linéaire résolue du premier ordre. On sait la résoudre, ce qui donne $p \mapsto t(p)$, puis $y(p)$ grâce à l'équation initiale. Sur ces intervalles, on obtient les branches de solutions pour lesquelles $t'(p) \neq 0$ en utilisant le paramétrage $p \mapsto (t(p), a(p)t(p) + b(p))$. Les valeurs où $p = a(p)$ doivent être examinées séparément ; on vérifie ensuite les raccords possibles dans l'équation initiale.

Équations de Clairaut

Il s'agit des équations de la forme $y = y' t + b(y')$, où b est une fonction de classe C^1 sur un intervalle J de $\mathbb{R}$. C'est un cas particulier d'équation de Lagrange dégénérée car $a(m) = m$ pour tout m . Les solutions affines sont de la forme $t \mapsto m t + b(m)$ pour tout $m \in J$. L'étude des équations de Lagrange montre que les solutions de classe C^2 dont la dérivée seconde ne s'annule pas peuvent se paramétrer sous la forme $p \mapsto (t(p), y(t(p)))$, où $t(p) = -b'(p)$ et $y(t(p)) = -p b'(p) + b(p)$. L'arc ainsi paramétré possède la propriété suivante : les tangentes à cet arc sont aussi solutions de l'équation différentielle (i.e. l'arc est l'enveloppe des solutions affines). Ces formules décrivent les branches où le paramétrage est admissible ; les raccords doivent être vérifiés dans l'équation initiale.

4.4 Compléments

4.4.1 Sur les équations différentielles linéaires d'ordre 2

On considère sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$, l'équation

$$y'' + p(t) y' + q(t) y = 0, \quad (\text{L})$$

où p et q sont deux fonctions continues d'un intervalle ouvert $I \subset \mathbb{R}$ dans $\mathbb{K}$. On étudie les solutions à valeurs dans $\mathbb{K}$.

Rappels

C'est un cas particulier d'équation différentielle linéaire homogène donc toute solution maximale est définie sur I tout entier. L'ensemble $\mathcal{S}$ des solutions maximales est un $\mathbb{K}$ -espace vectoriel de dimension 2 et pour tout $t_0 \in I$, l'application $f \mapsto (f(t_0), f'(t_0))$ est un isomorphisme de $\mathcal{S}$ dans $\mathbb{K}^2$. Tous les éléments de $f \in \mathcal{S}$ sont de classe C^2 . La connaissance de deux solutions linéairement indépendantes de (L) permet de résoudre l'équation

$$y'' + p(t) y' + q(t) y = r(t) \quad (\text{L}')$$

par variation des constantes. Lorsqu'une seule solution particulière de (L) est connue, l'abaissement de l'ordre permet d'en rechercher une seconde, indépendante de la première.

Wronskien

Si u et v sont des solutions maximales, leur wronskien est $\text{wronskien}(u, v) = u v' - u' v$ et vérifie pour tous $t_0, t \in I$,

$$\text{wronskien}(u, v)(t) = \text{wronskien}(u, v)(t_0) \exp\left(-\int_{t_0}^t p(s) ds\right).$$

En particulier, le wronskien est nul partout si et seulement s'il s'annule en un point. De plus, si p est la fonction nulle, le wronskien est constant.

Remarque. Si p est de classe C^1 , on peut supprimer le terme en dérivée première en procédant comme suit. On pose $y = zx$, où z et x sont deux fonctions de classe C^2 de I dans $\mathbb{K}$ inconnues. En remplaçant dans (L), on obtient

$$zx'' + (2z' + pz)x' + (z'' + pz' + qz)x = 0.$$

On choisit $z = \exp(-\frac{1}{2} \int_{t_0}^t p(s) ds)$, qui est de classe C^2 et ne s'annule pas. Alors $2z' + pz = 0$ et l'équation devient $x'' + (q - p'/2 - p^2/4)x = 0$. Ensuite, si p est constante nulle, l'équation (L) devient $y'' + qy = 0$, et donc $y''y' + qyy' = 0$, puis

$$\frac{d}{dt}((y')^2) + q(t) \frac{d}{dt}(y^2) = 0,$$

ce qui peut être utile dans la résolution.

4.4.2 Lemme de Gronwall et applications

Le lemme de Gronwall fournit des majorations à partir d'inégalités intégrales ; il permet aussi d'exploiter une inégalité différentielle après intégration.

Théorème 4.8 (Lemme de Gronwall). Soit φ, ψ et y trois fonctions continues sur un intervalle $[a, b]$, à valeurs positives et telles que pour tout $t \in [a, b]$,

$$y(t) \leq \varphi(t) + \int_a^t \psi(s) y(s) ds.$$

Alors, pour tout $t \in [a, b]$,

$$y(t) \leq \varphi(t) + \int_a^t \varphi(s) \psi(s) \exp\left(\int_s^t \psi(u) du\right) ds.$$

Démonstration. On pose $F(t) = \int_a^t \psi(s) y(s) ds$. En multipliant la condition par $\psi(t)$, on obtient

$$F'(t) - \psi(t) F(t) \leq \varphi(t) \psi(t),$$

et donc

$$G'(t) \leq \varphi(t) \psi(t) \exp\left(-\int_a^t \psi(s) ds\right) \text{ avec } G(t) = F(t) \exp\left(-\int_a^t \psi(s) ds\right).$$

En intégrant, on a

$$G(t) = G(t) - G(a) \leq \int_a^t \varphi(s) \psi(s) \exp\left(-\int_a^s \psi(u) du\right) ds.$$

Cependant, la condition de l'énoncé montre que

$$y(t) \leq \varphi(t) + G(t) \exp\left(\int_a^t \psi(u) du\right),$$

ce qui conduit au résultat. □

Corollaire 4.9. Soit ψ et y deux fonctions continues d'un intervalle $[a, b]$ dans $\mathbb{R}^+$ et vérifiant

$$\exists c \geq 0 \quad \forall t \in [a, b] \quad y(t) \leq c + \int_a^t \psi(s) y(s) ds.$$

Alors, pour tout $t \in [a, b]$,

$$y(t) \leq c \exp\left(\int_a^t \psi(s) ds\right).$$

Démonstration. On applique le lemme de Gronwall dans le cas où la fonction φ est constante égale à c . On a donc pour tout $t \in [a, b]$,

$$y(t) \leq c + \int_a^t c \psi(s) \exp\left(\int_s^t \psi(u) du\right) ds \leq c + c \left[-\exp\left(\int_s^t \psi(u) du\right) \right]_{s=a}^{s=t} = c \exp\left(\int_a^t \psi(u) du\right). \quad \square$$

L'application suivante sert notamment à contrôler les erreurs associées aux solutions approchées d'équations différentielles.

Corollaire 4.10. Soit $y : [a, b] \rightarrow \mathbb{R}^n$ une fonction de classe C^1 vérifiant

$$\exists \alpha > 0 \quad \exists \beta \geq 0 \quad \forall t \in [a, b] \quad \|y'(t)\| \leq \beta + \alpha \|y(t)\|.$$

Alors, pour tout $t \in [a, b]$, on a

$$\|y(t)\| \leq \|y(a)\| e^{\alpha(t-a)} + \frac{\beta}{\alpha} (e^{\alpha(t-a)} - 1).$$

Démonstration. On écrit pour tout $t \in [a, b]$,

$$\|y(t)\| \leq \|y(a)\| + \|y(t) - y(a)\| \leq \|y(a)\| + \int_a^t \|y'(s)\| ds \leq \|y(a)\| + \beta(t-a) + \alpha \int_a^t \|y(s)\| ds.$$

On applique le lemme de Gronwall avec $\varphi(t) = \|y(a)\| + \beta(t-a)$ et $\psi(s) = \alpha$, ainsi qu'à $\|y\|$. On en déduit que pour tout $t \in [a, b]$,

$$\begin{aligned} \|y(t)\| &\leq \|y(a)\| + \beta(t-a) + \int_a^t (\|y(a)\| + \beta(s-a)) \alpha \exp\left(\int_s^t \alpha du\right) ds \\ &= \|y(a)\| + \beta(t-a) + \alpha \int_a^t (\|y(a)\| + \beta(s-a)) e^{\alpha(t-s)} ds \\ &= \|y(a)\| + \beta(t-a) + \alpha \left(\left[(\|y(a)\| + \beta(s-a)) \frac{e^{\alpha(t-s)}}{-\alpha} \right]_a^t + \int_a^t \frac{\beta}{\alpha} e^{\alpha(t-s)} ds \right) \\ &= \|y(a)\| + \beta(t-a) + \alpha \left((\|y(a)\| + \beta(t-a)) \frac{1}{-\alpha} - \|y(a)\| \frac{e^{\alpha(t-a)}}{-\alpha} + \frac{\beta}{\alpha} \left[\frac{e^{\alpha(t-s)}}{-\alpha} \right]_a^t \right) \\ &= \|y(a)\| + \beta(t-a) - \|y(a)\| - \beta(t-a) + \|y(a)\| e^{\alpha(t-a)} + \beta \left(\frac{1 - e^{\alpha(t-a)}}{-\alpha} \right) \\ &= \|y(a)\| e^{\alpha(t-a)} + \frac{\beta}{\alpha} (e^{\alpha(t-a)} - 1). \end{aligned}$$

□

4.4.3 Principe de majoration a priori

Ce principe est aussi appelé théorème de sortie de tout compact et concerne le comportement des solutions maximales d'équations différentielles aux extrémités de leur intervalle de définition. Soit $F : \Omega \rightarrow \mathbb{R}^n$, avec $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ ouvert, une fonction continue et localement lipschitzienne en la seconde variable. Le théorème de Cauchy-Lipschitz indique que pour tout $(t_0, x_0) \in \Omega$, il existe un intervalle I centré en t_0 et une fonction φ définie sur I solution de $y' = F(t, y)$ vérifiant $\varphi(t_0) = x_0$. Ainsi, le problème de Cauchy a toujours une solution, définie sur un intervalle centré en t_0 . On s'intéresse au problème de l'existence d'un voisinage de V de (t_0, x_0) dans Ω tel que toutes les solutions au problème de Cauchy en $(t, x) \in V$ soient définies sur un même intervalle I .

Lemme 4.11. Soit $(t_0, x_0) \in \Omega$. Il existe un voisinage V de (t_0, x_0) dans Ω et un intervalle I ouvert centré en t_0 tel que pour tout $(t, x) \in V$, on a $t \in I$ et il existe une solution φ de $y' = F(t, y)$ définie sur I et vérifiant $\varphi(t) = x$.

Démonstration. Pour $t \in I$ et $\alpha > 0$, on note $I_\alpha(t) =]t - \alpha, t + \alpha[$, et pour $x \in \mathbb{R}^n$ et $r > 0$, on note $B_r(x) = \{y \in \mathbb{R}^n \mid \|x - y\| < r\}$. Si $U = I_\alpha(t) \times B_r(x)$ est un cylindre de sécurité pour F en (t, x) , c'est-à-dire que F est lipschitzienne en la seconde variable sur U et il existe un majorant $M > 0$ de F sur U tel que $\alpha M < r$, alors il existe une solution φ de $y' = F(t, y)$ définie sur $I_\alpha(t)$ et vérifiant $\varphi(t) = x$. Soit $U \subset \Omega$ un voisinage compact de (t_0, x_0) tel que F soit lipschitzienne en la seconde variable sur U . On note M un majorant de F sur U , ainsi que $\alpha, r > 0$ tels que $\alpha M < r$ et $I_\alpha(t_0) \times B_r(x_0) \subset U$, de sorte que $I_\alpha(t_0) \times B_r(x_0)$ est un cylindre de sécurité pour F en (t_0, x_0) . On pose $I = I_{\alpha/3}(t_0)$ et $B = B_{r/3}(x_0)$, puis $V = I \times B$. Pour $(t, x) \in V$, on observe que $I_{2\alpha/3}(t) \subset I_\alpha(t_0)$ et $B_{2r/3}(x) \subset B_r(x_0)$, et on a $(2\alpha/3)M < 2r/3 < r$, donc $I_{2\alpha/3}(t) \times B_{2r/3}(x)$ est un cylindre de sécurité pour F en (t, x) , et donc il existe une solution φ de $y' = F(t, y)$ définie sur $I_{2\alpha/3}(t)$ et vérifiant $\varphi(t) = x$. Pour tout $(t, x) \in V$, on a $I \subset I_{2\alpha/3}(t)$, ce qui conduit au résultat. □

Théorème 4.12 (Principe de majoration a priori). Soit $]a, b[$ un intervalle ouvert de $\mathbb{R}$, soit O un ouvert de $\mathbb{R}^n$, et soit $F :]a, b[\times O \rightarrow \mathbb{R}^n$ une fonction continue et localement lipschitzienne en la seconde variable. Soit $\varphi :]\alpha, \beta[\rightarrow \mathbb{R}^n$ une solution maximale de $y' = F(t, y)$. Alors, si $\beta < b$ (resp. si $a < \alpha$), pour tout compact $K \subset O$, il existe un voisinage V de β (resp. de α) dans $]a, b[$ tel que $\varphi(t) \notin K$ pour tout $t \in V$.

Démonstration. On se place au voisinage de β . Par l'absurde, supposons qu'il existe un compact $K \subset O$ et des réels $t_n \in]\alpha, \beta[$ tels que $t_n \rightarrow \beta$ quand $n \rightarrow \infty$, et $\varphi(t_n) \in K$ pour tout n . Par compacité, on extrait une sous-suite convergente, encore notée $\varphi(t_n)$ pour simplifier. On a donc $\varphi(t_n) \rightarrow x_0 \in K \subset O$. Comme $(\beta, x_0) \in]a, b[\times O$, il existe un voisinage $]\beta - \varepsilon, \beta + \varepsilon[\times V$ de (β, x_0) dans $]a, b[\times O$ tel que pour tous $t \in]\beta - \varepsilon, \beta + \varepsilon[$ et $x \in V$, il existe ψ solution de $y' = F(t, y)$ définie sur $]\beta - \varepsilon, \beta + \varepsilon[$ et vérifiant $\psi(t) = x$. On choisit n assez grand pour que $t_n \in]\beta - \varepsilon, \beta + \varepsilon[$ et $\varphi(t_n) \in V$. On note ψ la solution définie sur $]\beta - \varepsilon, \beta + \varepsilon[$ telle que $\psi(t_n) = \varphi(t_n)$. L'unicité au problème de Cauchy donnée par le théorème de Cauchy-Lipschitz montre que $\psi = \varphi$ sur $]\beta - \varepsilon, \beta[$. La fonction définie sur $]\alpha, \beta + \varepsilon[$ par $\varphi(t)$ si $\alpha < t < \beta$, et par $\psi(t)$ si $\beta \leq t < \beta + \varepsilon$, est donc solution sur $]\alpha, \beta + \varepsilon[$, ce qui contredit la maximalité de φ . $\square$

Remarques. • Si $O = \mathbb{R}^n$, le résultat indique que si $\beta < b$, alors $\|\varphi(t)\| \rightarrow +\infty$ quand $t \rightarrow \beta$.

- Si $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ est localement lipschitzienne, le théorème précédent implique que si $\varphi :]a, b[\rightarrow \mathbb{R}^n$ est solution maximale de $y' = F(y)$, et si $b < +\infty$, alors $\|\varphi(t)\| \rightarrow +\infty$ quand $t \rightarrow b$.

Corollaire 4.13. On considère un équation différentielle linéaire sur $\mathbb{R}^n$ de la forme $X' = A(t)X + B(t)$, avec deux fonctions continues $A :]a, b[\rightarrow \mathcal{M}_n(\mathbb{R})$ et $B :]a, b[\rightarrow \mathbb{R}^n$. Alors, les solutions maximales sont définies sur $]a, b[$.

Démonstration. La fonction F définie de $]a, b[\times \mathbb{R}^n$ dans $\mathbb{R}^n$ par la formule $F(t, X) = A(t)X + B(t)$ est continue et localement lipschitzienne en la seconde variable. En effet, on a $F(t, X) - F(t, Y) = A(t)(X - Y)$, et la fonction A est continue. Notons $X :]\alpha, \beta[\rightarrow \mathbb{R}^n$ une solution maximale, et supposons $\beta < b$ et choisissons $t_0 \in]\alpha, \beta[$. On munit $\mathcal{M}_n(\mathbb{R})$ de la norme d'algèbre usuelle, et on note

$$M = 1 + \sup_{t \in [t_0, \beta]} \|A(t)\| \quad \text{et} \quad N = \sup_{t \in [t_0, \beta]} \|B(t)\|.$$

On a donc pour tout $t \in [t_0, \beta]$, la majoration $\|X'(t)\| \leq M \|X(t)\| + N$. Le lemme de Gronwall implique que pour tout $t \in [t_0, \beta]$,

$$\|X(t)\| \leq \|X(t_0)\| e^{M(t-t_0)} + \frac{N}{M} (e^{M(t-t_0)} - 1).$$

Ainsi, la solution X est bornée au voisinage de β^- . Cela contredit le principe de majoration a priori, et donc $\beta = b$. De même, on a $\alpha = a$. $\square$

4.4.4 Dépendance des solutions vis-à-vis des conditions initiales

Soit $F : \Omega \rightarrow \mathbb{R}^n$, où $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ est un ouvert, une application continue et localement lipschitzienne en la seconde variable. Le théorème de Cauchy-Lipschitz indique que pour tout $(t_0, x_0) \in \Omega$, il existe une solution maximale de $y' = F(t, y)$ notée φ_{t_0, x_0} et définie sur un intervalle ouvert I_{t_0, x_0} contenant t_0 , et telle que $\varphi(t_0) = x_0$. On note

$$W = \bigcup_{(t_0, x_0) \in \Omega} \{(t_0, x_0)\} \times I_{t_0, x_0} \subset \Omega \times \mathbb{R},$$

et on définit une application $\Phi : W \rightarrow \mathbb{R}^n$ par $\Phi(t_0, x_0; t) = \varphi_{t_0, x_0}(t)$. Ainsi, $\Phi(t_0, x_0; \cdot)$ est la solution maximale de $y' = F(t, y)$ passant en x_0 à l'instant t_0 . On peut montrer en général que W est ouvert et que Φ est continue sur W .

Espaces de Hilbert

Références : [7, 8, 13].

5.1 Définitions et propriétés élémentaires

Définition 5.1. • Soit H un espace vectoriel sur $\mathbb{R}$. Un *produit scalaire* $\langle u, v \rangle$ est une forme bilinéaire de $H \times H$ dans $\mathbb{R}$ qui est symétrique (i.e. $\langle u, v \rangle = \langle v, u \rangle$ pour tous $u, v \in H$) et définie positive (i.e. $\langle u, u \rangle \geq 0$ pour tout $u \in H$, avec $\langle u, u \rangle = 0$ si et seulement si $u = 0$).

- Dans le cas où H est un espace vectoriel sur $\mathbb{C}$, on dit que $\langle u, v \rangle$ est un *produit scalaire hermitien* si c'est une forme sesquilinéaire (i.e. linéaire à gauche : $\langle \alpha u + v, w \rangle = \alpha \langle u, w \rangle + \langle v, w \rangle$; et semi-linéaire à droite : $\langle u, \alpha v + w \rangle = \bar{\alpha} \langle u, v \rangle + \langle u, w \rangle$), hermitienne (i.e. $\langle u, v \rangle = \overline{\langle v, u \rangle}$), définie positive.

Remarque. Par commodité, et sauf mention contraire, on se restreindra souvent dans ce cours au cas réel.

Proposition 5.1 (Inégalité de Cauchy-Schwarz). Soit H un espace vectoriel muni d'un produit scalaire noté $\langle \cdot, \cdot \rangle$. Alors, pour tous $u, v \in H$, on a l'inégalité $|\langle u, v \rangle|^2 \leq \langle u, u \rangle \langle v, v \rangle$.

Démonstration. Pour tout $t \in \mathbb{R}$, on a

$$0 \leq \langle u + tv, u + tv \rangle = \langle u, u \rangle + 2t\Re\langle u, v \rangle + t^2\langle v, v \rangle.$$

Il s'agit d'un polynôme en t qui ne prend que des valeurs positives. Ainsi, on a deux cas possibles :

- Si $\langle v, v \rangle = 0$, alors le polynôme est de degré 1 au plus et il est donc nécessairement constant. Il en ressort que

$$0 = (\Re\langle u, v \rangle)^2 \leq \langle u, u \rangle \langle v, v \rangle = 0.$$

- Sinon, le polynôme est de degré 2, avec un discriminant négatif ou nul, c'est-à-dire que

$$4(\Re\langle u, v \rangle)^2 - 4\langle u, u \rangle \langle v, v \rangle \leq 0.$$

Dans les deux cas, on a donc $(\Re\langle u, v \rangle)^2 \leq \langle u, u \rangle \langle v, v \rangle$. Pour conclure, on note α un nombre complexe de module 1 tel que $|\langle u, v \rangle| = \alpha \langle u, v \rangle$. On a

$$|\langle u, v \rangle| = \alpha \langle u, v \rangle = \langle \alpha u, v \rangle = \Re\langle \alpha u, v \rangle.$$

Avec ce qui précède, on en déduit que

$$(\Re\langle \alpha u, v \rangle)^2 \leq \langle \alpha u, \alpha u \rangle \langle v, v \rangle = \underbrace{\alpha \bar{\alpha}}_{=1} \langle u, u \rangle \langle v, v \rangle.$$

(On remarque dans cette preuve que la propriété selon laquelle $\langle u, u \rangle > 0$ pour $u \neq 0$ n'est jamais utilisée.) $\square$

Corollaire 5.2. Soit H un espace vectoriel muni d'un produit scalaire $\langle \cdot, \cdot \rangle$. Alors, l'expression $\|u\| = \langle u, u \rangle^{1/2}$ définit une norme sur H .

Démonstration. Pour $u, v \in H$, on observe que

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2 + 2\Re\langle u, v \rangle \leq \|u\|^2 + \|v\|^2 + 2\|u\| \|v\| = (\|u\| + \|v\|)^2.$$

Cela prouve l'inégalité triangulaire. Les autres propriétés de la norme sont trivialement vérifiées. $\square$

Proposition 5.3 (Cas d'égalité dans Cauchy-Schwarz). *Pour tous $u, v \in H$, on a $|\langle u, v \rangle| = \|u\| \|v\|$ si et seulement si u et v sont linéairement dépendants.*

Démonstration. On a clairement égalité quand u et v sont en dépendance linéaire. Supposons avoir l'égalité, et notons α un nombre complexe de module 1 tel que $|\langle u, v \rangle| = \Re \langle \alpha u, v \rangle$. Alors, on a

$$\|u\|v - \alpha\|v\|u\|^2 = 2\|u\|^2\|v\|^2 - 2\|u\| \|v\| \Re \langle \alpha u, v \rangle = 0,$$

ce qui donne $\|u\|v - \alpha\|v\|u = 0$, et donc la dépendance linéaire. $\square$

Une conséquence de la définition de la norme associée à un produit scalaire est l'identité du parallélogramme : pour tous $u, v \in H$,

$$\left\| \frac{u+v}{2} \right\|^2 + \left\| \frac{u-v}{2} \right\|^2 = \frac{1}{2} (\|u\|^2 + \|v\|^2),$$

c'est-à-dire aussi

$$\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

Définition 5.2. Un espace de Hilbert est un espace vectoriel H muni d'un produit scalaire $\langle \cdot, \cdot \rangle$ pour lequel la norme $\| \cdot \|$ définie précédemment le rend complet.

Exemples. • Les espaces vectoriels $\mathbb{R}^d$ et $\mathbb{C}^d$, munis respectivement des produits scalaires donnés par $\sum_{i=1}^d a_i u_i v_i$

et $\sum_{i=1}^d a_i u_i \overline{v_i}$ pour des vecteurs $u = (u_1, \dots, u_d)$ et $v = (v_1, \dots, v_d)$ dans $\mathbb{R}^d$ ou $\mathbb{C}^d$, avec des coefficients $a_i > 0$ fixés.

Si tous les coefficients a_i sont égaux à 1, on dit que les espaces $\mathbb{R}^d$ et $\mathbb{C}^d$ sont munis respectivement de leur structure euclidienne ou hermitienne canonique.

- Sur un espace mesuré $(\Omega, \mathcal{F}, \mu)$, l'espace $L^2(\Omega, \mathcal{F}, \mu)$ des classes de fonctions de carré intégrable modulo l'égalité presque partout est un espace de Hilbert, pour le produit scalaire $\langle f, g \rangle = \int_{\Omega} f \overline{g} d\mu$ dans le cas complexe, et $\int_{\Omega} f g d\mu$ dans le cas réel. Sur l'espace $\mathcal{L}^2$ des fonctions avant passage au quotient, cette forme est seulement positive : une fonction non nulle mais nulle presque partout a une norme associée nulle.
- Comme cas particulier du précédent, il y a l'ensemble $\mathbb{N}$ muni de la tribu discrète de ses parties $\mathcal{P}(\mathbb{N})$ et de la mesure de comptage : on obtient l'ensemble $\ell^2(\mathbb{N})$ muni du produit scalaire $\sum_{n=0}^{\infty} u_n v_n$ ou $\sum_{n=0}^{\infty} u_n \overline{v_n}$. Autre cas particulier, l'espace $L^2([0, 1])$ obtenu pour l'ensemble Ω l'intervalle $[0, 1]$, pour $\mathcal{F}$ la tribu des boréliens et pour μ la mesure de Lebesgue.
- L'espace de Sobolev $H^1([0, 1])$ est constitué des classes $f \in L^2([0, 1])$ admettant un représentant absolument continu de la forme $f(x) = c + \int_0^x g(t) dt$, avec $g \in L^2([0, 1])$. On note $f' = g$. Le produit scalaire $\langle f, h \rangle_{H^1} = \int_0^1 (f \overline{h} + f' \overline{h'}) dt$ en fait un espace de Hilbert. Les valeurs ponctuelles de f désignent celles de son unique représentant continu.

5.2 Projection sur un convexe fermé

On se restreint désormais au cas réel. Dans ce qui suit, H désigne un espace de Hilbert et on reprend les notations précédentes pour la norme et le produit scalaire.

Théorème 5.4. Soit $K \subset H$ un convexe fermé non vide. Alors, pour tout $f \in H$, il existe un unique $u \in K$ tel que

$$\|f - u\| = \min_{v \in K} \|f - v\|.$$

De plus, u est caractérisé par la propriété de l'angle obtus : $u \in K$ et

$$\forall v \in K \quad \langle f - u, v - u \rangle \leq 0.$$

On note $u = P_K f$ la projection de f sur K . Alors, l'application P_K est 1-lipschitzienne.

Démonstration. Existence. On note (v_n) une suite minimisante, i.e. $v_n \in K$ pour tout entier n et de plus $d_n = \|f - v_n\| \xrightarrow{n \rightarrow \infty} d = \inf_{v \in K} \|f - v\|$. Montrons que la suite (v_n) est de Cauchy. D'après l'identité du parallélogramme appliquée aux vecteurs $f - v_n$ et $f - v_m$, on a

$$\left\| f - \frac{v_n + v_m}{2} \right\|^2 + \left\| \frac{v_n - v_m}{2} \right\|^2 = \frac{1}{2} (d_n^2 + d_m^2).$$

Le premier terme est supérieur ou égal à d^2 car $(v_n + v_m)/2$ appartient à K par convexité. Par conséquent,

$$\|v_n - v_m\|^2 \leq 2(d_n^2 + d_m^2) - 4d^2 \xrightarrow{m, n \rightarrow \infty} 0.$$

La suite (v_n) est donc de Cauchy. On note u sa limite. Comme K est fermé, on a $u \in K$, et de plus $d = \|f - u\|$.

Caractérisation par l'angle obtus. Supposons qu'un vecteur $u \in K$ vérifie $\|f - u\| = \min_{v \in K} \|f - v\|$ et considérons $w \in K$ et $t \in]0, 1]$. Par convexité, le vecteur $v = (1 - t)u + tw$ appartient à K , donc

$$\|f - u\| \leq \|f - ((1 - t)u + tw)\| = \|(f - u) - t(w - u)\|,$$

puis

$$\|f - u\|^2 \leq \|f - u\|^2 - 2t\langle f - u, w - u \rangle + t^2\|w - u\|^2,$$

et finalement $2\langle f - u, w - u \rangle \leq t\|w - u\|^2$. En faisant tendre t vers 0, on obtient $\langle f - u, w - u \rangle \leq 0$. Réciproquement, si cette inégalité est vérifiée, on a pour tout $v \in K$,

$$\|f - u\|^2 - \|f - v\|^2 = 2 \underbrace{\langle f - u, v - u \rangle}_{\leq 0} - \|u - v\|^2 \leq 0,$$

ce qui implique que $\|f - u\| \leq \|f - v\|$ pour tout $v \in K$.

Unicité. Si u_1 et u_2 sont deux minimiseurs, on a pour tout $v \in K$, les inégalités $\langle f - u_1, v - u_1 \rangle \leq 0$ et $\langle f - u_2, v - u_2 \rangle \leq 0$. En prenant $v = u_2$ dans la première et $v = u_1$ dans la seconde, on obtient

$$\langle f - u_1, u_2 - u_1 \rangle \leq 0 \quad \text{et} \quad \langle f - u_2, u_1 - u_2 \rangle \leq 0.$$

En ajoutant ces deux inégalités, on a $\langle (f - u_1) + (u_2 - f), u_2 - u_1 \rangle \leq 0$, ce qui donne finalement $\|u_2 - u_1\|^2 \leq 0$, et donc $u_2 = u_1$.

Propriété de Lipschitz. Pour $f_1, f_2 \in H$, on note respectivement $u_1 = P_K f_1$ et $u_2 = P_K f_2$ les projections de f_1 et f_2 sur K . Alors, pour tout vecteur $v \in K$, on a

$$\langle f_1 - u_1, v - u_1 \rangle \leq 0 \quad \text{et} \quad \langle f_2 - u_2, v - u_2 \rangle \leq 0,$$

et donc en prenant $v = u_2$ dans la première inégalité et $v = u_1$ dans la seconde, on a

$$\langle f_1 - u_1, u_2 - u_1 \rangle \leq 0 \quad \text{et} \quad \langle f_2 - u_2, u_1 - u_2 \rangle \leq 0.$$

En ajoutant ces inégalités, on a $\langle (f_1 - u_1) + (u_2 - f_2), u_2 - u_1 \rangle \leq 0$, ce qui donne $\|u_2 - u_1\|^2 \leq \langle f_2 - f_1, u_2 - u_1 \rangle$. Avec l'inégalité de Cauchy-Schwarz, on majore ce dernier terme par $\|f_2 - f_1\| \|u_2 - u_1\|$. On conclut finalement que $\|u_2 - u_1\| \leq \|f_2 - f_1\|$. $\square$

Remarque. Dans un espace de Banach, la projection n'existe pas toujours. En effet, dans l'espace $\mathcal{C}^0([0, 1])$ des fonctions continues définies sur l'intervalle $[0, 1]$ muni de sa norme usuelle $\|\cdot\|_\infty$, considérons l'ensemble convexe

$$K = \left\{ f \in \mathcal{C}^0([0, 1]) \text{ positive} \mid f(0) = 0 \text{ et } \int_0^1 f(t) dt = 1 \right\}.$$

On observe que $d(0, K) = \inf_{f \in K} \|f\|_\infty = 1$. En effet, d'une part, toute fonction $f \in K$ vérifie

$$1 = \int_0^1 f(t) dt \leq \|f\|_\infty.$$

D'autre part, pour tout entier $n \geq 2$, en désignant par g_n la fonction définie sur $[0, 1]$, affine entre l'origine et le point de coordonnées $\left(\frac{2}{n}, \frac{n}{n-1}\right)$, puis constante, on observe que $g_n \in K$ et que $\|g_n\|_\infty \xrightarrow{n \rightarrow \infty} 1$. En effet, on a clairement g_n continue et positive, avec $g_n(0) = 0$ et $\|g_n\|_\infty = \frac{n}{n-1}$, et

$$\int_0^1 g_n(t) dt = \frac{1}{2} \frac{2}{n} \frac{n}{n-1} + \frac{n}{n-1} \left(1 - \frac{2}{n}\right) = 1.$$

On a donc comme annoncé $d(0, K) = 1$. Cependant, cette distance n'est pas atteinte. En effet, cela impliquerait l'existence d'une fonction $f \in K$ telle que $\|f\|_\infty = 1$. Ainsi, la fonction $1 - f$ serait continue, positive et d'intégrale nulle sur $[0, 1]$. La fonction f serait donc constante égale à 1 sur cet intervalle, ce qui contredirait sa nullité en 0.

Corollaire 5.5 (Cas d'un sous-espace vectoriel fermé). Soit $M \subset H$ un sous-espace vectoriel fermé et soit $f \in H$. Alors, $u = P_M f$ est caractérisé par les conditions $u \in M$ et

$$\forall v \in M \quad \langle f - u, v \rangle = 0.$$

De plus, P_M est un opérateur linéaire. C'est un projecteur linéaire de noyau $M^\perp$, et de norme 1 si $M \neq \{0\}$.

Démonstration. La caractérisation par l'angle obtus donne $\langle f - u, v - u \rangle \leq 0$ pour tout $v \in M$. Par linéarité, on a même $\langle f - u, tv - u \rangle \leq 0$ pour tout $v \in M$ et tout $t \in \mathbb{R}$, c'est-à-dire

$$t\langle f - u, v \rangle - \langle f - u, u \rangle \leq 0.$$

Nécessairement, le coefficient $\langle f - u, v \rangle$ est nul. Réciproquement, si on a $\langle f - u, v \rangle = 0$ pour tout $v \in M$, alors on a aussi en particulier $\langle f - u, u \rangle = 0$, et donc par soustraction $\langle f - u, v - u \rangle = 0$.

Pour la linéarité, on considère $g \in H$ et $v = P_M g$. Alors,

$$\langle (\lambda f + g) - (\lambda u + v), w \rangle = \lambda \langle f - u, w \rangle + \langle g - v, w \rangle = 0$$

pour tout $w \in M$, ce qui donne $P_M(\lambda f + g) = \lambda u + v = \lambda P_M f + P_M g$. □

Corollaire 5.6. Soit M un sous-espace vectoriel fermé de H . Alors, on a la décomposition en somme directe orthogonale $H = M \oplus M^\perp$: chaque vecteur de H s'écrit de façon unique $u + v$, avec $u \in M$ et $v \in M^\perp$, et ces deux composantes sont orthogonales. Le projecteur associé à cette décomposition est P_M .

Démonstration. Pour $f \in H$, on pose $u = P_M f$. Alors, on écrit $f = u + (f - u)$ et le résultat précédent montre que $\langle f - u, v \rangle = 0$ pour tout $v \in M$, ce qui montre que $f - u \in M^\perp$, en plus d'avoir $u \in M$. Enfin, si w appartient à l'intersection de M et $M^\perp$, on a $\langle w, w \rangle = 0$, et donc $w = 0$. □

Si M désigne un sous-espace vectoriel quelconque de H , alors l'adhérence $\overline{M}$ est un sous-espace vectoriel fermé. Le corollaire précédent implique que

$$H = \overline{M} \oplus \overline{M}^\perp = \overline{M} \oplus M^\perp,$$

où la dernière égalité vient du fait élémentaire que $\overline{M}^\perp = M^\perp$. On en déduit le critère utile de densité suivant.

Corollaire 5.7. Un sous-espace vectoriel M est dense dans H si et seulement si son orthogonal $M^\perp$ est réduit à $\{0\}$.

5.3 Dualité et théorème de représentation de Riesz

On note H' le dual topologique de H , i.e. l'espace vectoriel des formes linéaires continues sur H , muni de la norme duale notée aussi $\| \cdot \|$.

Théorème 5.8 (Riesz). *Pour toute forme linéaire $\varphi \in H'$, il existe un unique vecteur $f \in H$ tel que $\varphi(v) = \langle f, v \rangle$ pour tout $v \in H$. De plus, on a $\|f\| = \|\varphi\|$. Autrement dit, l'application suivante est une isométrie surjective :*

$$\begin{aligned} H &\longrightarrow H' \\ f &\longmapsto \langle f, \cdot \rangle. \end{aligned}$$

Démonstration. Existence. Notons $M = \ker \varphi = \varphi^{-1}(\{0\})$ le noyau de la forme linéaire φ . Il s'agit d'un sous-espace vectoriel fermé de H . Si $M = H$, on a $\varphi = 0$ et on peut prendre $f = 0$. Sinon, il existe un vecteur $g \in H \setminus M$ tel que $\|g\| = 1$ et $g \in M^\perp$. En effet, on peut partir de $g_0 \in H \setminus M$ quelconque, puis poser $g = \frac{g_0 - P_M g_0}{\|g_0 - P_M g_0\|}$. Ensuite, pour tout vecteur $v \in H$, on peut écrire $v = \lambda g + w$, avec $\lambda \in \mathbb{R}$ et $w \in M$ en posant $\lambda = \frac{\varphi(v)}{\varphi(g)}$ et $w = v - \lambda g$. On a donc $0 = \langle g, w \rangle = \langle g, v - \lambda g \rangle$, c'est-à-dire $\langle g, v \rangle = \lambda \|g\|^2 = \frac{\varphi(v)}{\varphi(g)}$, puis finalement $\varphi(v) = \langle \varphi(g)g, v \rangle$. On conclut en posant $f = \varphi(g)g$.

Isométrie. On a d'abord $\|f\| = \|\varphi(g)g\| = |\varphi(g)| \leq \|\varphi\|$ car $\|g\| = 1$. Réciproquement, l'inégalité de Cauchy-Schwarz implique que pour tout vecteur $v \in H$, on a $|\varphi(v)| = |\langle f, v \rangle| \leq \|f\| \|v\|$, et donc $\|\varphi\| \leq \|f\|$.

Unicité. Si f_1 et f_2 conviennent, alors $\langle f_1, v \rangle = \varphi(v) = \langle f_2, v \rangle$ pour tout vecteur $v \in H$. Donc, $\langle f_1 - f_2, v \rangle = 0$ pour tout $v \in H$, d'où $f_1 = f_2$. $\square$

Corollaire 5.9 (Existence et unicité de l'adjoint). *Soit $T \in \mathcal{L}(H)$ un opérateur linéaire continu de H dans H . Alors, il existe un unique opérateur linéaire continu $T^* \in \mathcal{L}(H)$ tel que*

$$\forall u, v \in H \quad \langle Tu, v \rangle = \langle u, T^*v \rangle.$$

De plus, il y a égalité des normes d'opérateur : $\|T^*\| = \|T\|$. On appelle T^* l'adjoint de T . Enfin, on a clairement $(T^*)^* = T$ et $(ST)^* = T^*S^*$ pour tous opérateurs $S, T \in \mathcal{L}(H)$.

Démonstration. Pour tout vecteur $v \in H$, l'application $u \mapsto \langle Tu, v \rangle$ est une forme linéaire continue sur H . Le théorème de Riesz implique l'existence d'un unique vecteur $T^*v \in H$ tel que $\langle Tu, v \rangle = \langle u, T^*v \rangle$ quel que soit $u \in H$. Il est facile de voir qu'on définit ainsi une application linéaire T^* sur H .

De plus, on sait que le vecteur T^*v a même norme que la forme linéaire qu'il représente, i.e. $\|T^*v\| = \sup_{\|u\| \leq 1} |\langle Tu, v \rangle|$.

L'inégalité de Cauchy-Schwarz implique en outre que pour tous vecteurs $u, v \in H$,

$$|\langle Tu, v \rangle| \leq \|Tu\| \|v\| \leq \|T\| \|u\| \|v\|.$$

On en déduit que $\|T^*v\| \leq \|T\| \|v\|$, donc que l'application linéaire T^* est continue, avec une norme d'opérateur telle que $\|T^*\| \leq \|T\|$. En appliquant cela à T^* , on obtient $\|T^*\| \geq \|(T^*)^*\| = \|T\|$. D'où $\|T^*\| = \|T\|$.

Pour établir l'unicité de l'opérateur T^* , supposons que deux opérateurs S_1 et S_2 conviennent, i.e. vérifient $\langle u, S_1v \rangle = \langle Tu, v \rangle = \langle u, S_2v \rangle$ pour tous $u, v \in H$. On a en particulier $\langle u, S_1v - S_2v \rangle = 0$ pour tout $u \in H$, ce qui implique que $S_1v = S_2v$ pour tout $v \in H$, et finalement $S_1 = S_2$. $\square$

Proposition 5.10. *Pour tout opérateur $T \in \mathcal{L}(H)$, on a*

$$\ker(T^*) = (\operatorname{im} T)^\perp \quad \text{et} \quad \overline{\operatorname{im} T^*} = (\ker T)^\perp.$$

Démonstration. La première identité est triviale. On l'applique à T^* , ce qui donne $\ker((T^*)^*) = (\operatorname{im} T^*)^\perp$. En prenant l'orthogonal, et en se rappelant que prendre deux fois l'orthogonal d'un sous-espace vectoriel revient à prendre son adhérence, on a $(\ker(T))^\perp = ((\operatorname{im} T^*)^\perp)^\perp = \overline{\operatorname{im} T^*}$. $\square$

Définition 5.3. Un opérateur $T \in \mathcal{L}(H)$ est dit *auto-adjoint* si $T = T^*$.

Remarques. Dans ce cas, on a clairement $\|T^2\| = \|T\|^2$, puisqu'en général $\|TT^*\| = \|T^*T\| = \|T\|^2$. De plus, les opérateurs TT^* et T^*T sont toujours auto-adjoints quel que soit $T \in \mathcal{L}(H)$.

Lemme 5.11. Supposons $H \neq \{0\}$. La norme d'un opérateur auto-adjoint quelconque $T \in \mathcal{L}(H)$ vérifie

$$\|T\| = \sup_{\|u\|=1} |\langle Tu, u \rangle|.$$

Démonstration. Notons s le supremum figurant dans l'énoncé. Pour tous vecteurs $u, v \in H$, on a par homogénéité puis grâce à l'identité du parallélogramme,

$$4|\langle Tu, v \rangle| = |\langle T(u+v), u+v \rangle - \langle T(u-v), u-v \rangle| \leq (\|u+v\|^2 + \|u-v\|^2) s = 2(\|u\|^2 + \|v\|^2) s.$$

Par conséquent, on a

$$\|T\| = \|T^*\| = \sup_{\|u\|=\|v\|=1} |\langle Tu, v \rangle| \leq s.$$

Enfin, on a bien sûr l'inégalité réciproque puisque le supremum de $|\langle Tu, v \rangle|$ sur tous les vecteurs u et v de norme 1 majore le supremum s , qui correspond au cas particulier où $v = u$. Finalement, $\|T\| = s$. $\square$

Remarque. En dimension finie, le supremum figurant dans l'énoncé du lemme est atteint par continuité de l'application $u \mapsto |\langle Tu, u \rangle|$ sur la sphère unité, qui est compacte. On dispose donc d'un vecteur unitaire $u \in H$ tel que $\|T\| = |\langle Tu, u \rangle|$. Ainsi, par définition de la norme d'opérateur, on a

$$|\langle Tu, u \rangle| = \|T\| = \sup_{\|v\|=1} \|Tv\| \geq \|Tu\| = \|Tu\| \|u\|.$$

Réciproquement, l'inégalité de Cauchy-Schwarz donne $|\langle Tu, u \rangle| \leq \|Tu\| \|u\|$. On est donc dans le cas d'égalité, ce qui implique l'existence d'une dépendance linéaire entre Tu et u . Autrement dit, u est un vecteur propre pour T . Notons λ la valeur propre associée. On a $\|T\| = |\langle Tu, u \rangle| = |\langle \lambda u, u \rangle| = |\lambda| \|u\|^2 = |\lambda|$, de sorte que $\lambda = \pm \|T\|$. Cela permet de démarrer la diagonalisation de l'opérateur T car ensuite l'espace vectoriel $\{u\}^\perp$ est stable par T , et on peut considérer la restriction correspondante.

5.4 Théorème de Lax-Milgram

Définition 5.4. On dit qu'une forme bilinéaire a de $H \times H$ dans $\mathbb{R}$ est :

- (i). *continue* s'il existe une constante $C > 0$ telle que $|a(u, v)| \leq C \|u\| \|v\|$ pour tous vecteurs $u, v \in H$;
- (ii). *coercive* s'il existe une constante $\alpha > 0$ telle que $a(v, v) \geq \alpha \|v\|^2$ pour tout vecteur $v \in H$.

Théorème 5.12 (Lax-Milgram). Soit H un espace de Hilbert et soit a une forme bilinéaire continue et coercive sur H . Alors, pour toute forme linéaire $\varphi \in H'$, il existe un unique vecteur $u \in H$ tel que $a(u, v) = \varphi(v)$ pour tout vecteur $v \in H$. De plus, si a est symétrique, alors u est caractérisé par

$$\frac{1}{2}a(u, u) - \varphi(u) = \min_{v \in H} \left(\frac{1}{2}a(v, v) - \varphi(v) \right).$$

Démonstration. Pour tout vecteur $u \in H$, l'application $v \mapsto a(u, v)$ est une forme linéaire sur H . Le théorème de Riesz implique l'existence d'un unique vecteur $Tu \in H$ tel que $a(u, v) = \langle Tu, v \rangle$ pour tout $v \in H$. Il est clair qu'on définit ainsi un unique opérateur linéaire continu $T \in \mathcal{L}(H)$.

La coercivité de a et l'inégalité de Cauchy-Schwarz impliquent la minoration $\|Tv\| \geq \alpha\|v\|$ pour tout vecteur $v \in H$, puisqu'en effet

$$\alpha\|v\|^2 \leq a(v, v) = \langle Tv, v \rangle \leq \|Tv\| \|v\|.$$

Cela impose en particulier l'injectivité de T .

De plus, l'image de l'opérateur T est fermée dans H . En effet, supposons que $w = \lim_{n \rightarrow \infty} w_n$ est limite d'une suite d'éléments $w_n \in \text{im } T$. On peut écrire $w_n = Tv_n$ pour tout n , et il résulte comme ci-dessus de la coercivité et de l'inégalité de Cauchy-Schwarz que

$$\alpha\|v_n - v_m\| \leq \|T(v_n - v_m)\| = \|w_n - w_m\| \xrightarrow{m, n \rightarrow \infty} 0.$$

La suite (v_n) est donc une suite de Cauchy dans H , qui est complet. Elle admet donc une limite, notée v . On a clairement $w = Tv \in \text{im } T$.

L'image de T est aussi dense dans H . En effet, si un vecteur v appartient à $(\text{im } T)^\perp$, alors on a $0 = \langle Tu, v \rangle = a(u, v)$ pour tout $u \in H$. En particulier, toujours par coercivité, on a

$$\alpha\|v\|^2 \leq a(v, v) = \langle Tv, v \rangle = 0.$$

Ainsi $v = 0$, ce qui prouve que l'orthogonal de $\text{im } T$ se réduit à $\{0\}$. Le critère de densité vu précédemment permet de conclure.

On déduit des arguments précédents que l'opérateur T est une bijection de H dans lui-même. Par conséquent, pour $\varphi \in H'$, le théorème de Riesz fournit un unique représentant $w \in H$. Si on pose $u = T^{-1}w$, alors pour tout vecteur $v \in H$, on a

$$\varphi(v) = \langle w, v \rangle = \langle Tu, v \rangle = a(u, v).$$

Cela prouve l'existence de u . Son unicité est claire puisque si u_1 et u_2 conviennent, on a $a(u_1, v) = \varphi(v) = a(u_2, v)$ pour tout $v \in H$, et donc $a(u_1 - u_2, v) = 0$ par bilinéarité. À nouveau par coercivité, on a

$$\alpha\|u_1 - u_2\|^2 \leq a(u_1 - u_2, u_1 - u_2) = 0,$$

ce qui donne $u_1 = u_2$.

Finalement, pour établir la caractérisation, supposons que la forme bilinéaire a est symétrique, et écrivons alors que pour $u, h \in H$,

$$\frac{1}{2}a(u + h, u + h) - \varphi(u + h) = \frac{1}{2}a(u, u) - \varphi(u) + \frac{1}{2}a(h, h) - \varphi(h) + a(u, h).$$

Si le vecteur u représente la forme linéaire φ pour la forme bilinéaire a , on a $a(u, h) = \varphi(h)$. De plus, par coercivité, on a clairement $a(h, h) \geq 0$. Ainsi, pour tout $h \in H$,

$$\frac{1}{2}a(u + h, u + h) - \varphi(u + h) \geq \frac{1}{2}a(u, u) - \varphi(u).$$

Comme le membre de gauche coïncide avec celui de droite pour $h = 0$, on en déduit que u est solution du problème de minimisation. Réciproquement, si u est solution du problème de minimisation, on a pour tout $h \in H$,

$$\frac{1}{2}a(h, h) - \varphi(h) + a(u, h) \geq 0.$$

En remplaçant h par son opposé, on obtient aussi

$$\frac{1}{2}a(-h, -h) - \varphi(-h) + a(u, -h) \geq 0.$$

En utilisant la linéarité et en regroupant ces deux inégalités, on en déduit que pour tout $h \in H$,

$$|a(u, h) - \varphi(h)| \leq \frac{1}{2}a(h, h).$$

Enfin, en remplaçant h par εh , avec $\varepsilon > 0$, cela donne

$$|a(u, h) - \varphi(h)| = \frac{1}{\varepsilon} |a(u, \varepsilon h) - \varphi(\varepsilon h)| \leq \frac{1}{2\varepsilon} a(\varepsilon h, \varepsilon h) = \frac{\varepsilon}{2} a(h, h).$$

En faisant $\varepsilon \rightarrow 0$, on conclut que $a(u, h) = \varphi(h)$ pour tout $h \in H$. □

5.5 Application au problème de Dirichlet

Une fonction $f \in \mathcal{C}^0([0, 1])$ étant donnée, on considère le problème différentiel suivant, appelé *problème de Dirichlet* :

$$\begin{cases} -u'' + u = f & \text{sur } [0, 1] \\ u(0) = u(1) = 0. \end{cases}$$

Il s'agit d'une équation différentielle linéaire d'ordre 2 avec des conditions aux bords de l'intervalle. Une solution classique, dite *forte*, est une fonction $u \in \mathcal{C}^2([0, 1])$ qui vérifie l'équation différentielle et les conditions aux bords.

On va adopter une approche variationnelle du problème. Pour cela, on commence par intégrer l'équation contre une fonction test ψ , en effectuant une intégration par parties. Ainsi, pour toute fonction $\psi \in \mathcal{C}^1([0, 1])$ telle que $\psi(0) = \psi(1) = 0$, on a

$$\int_0^1 u'(x)\psi'(x) dx + \int_0^1 u(x)\psi(x) dx = \int_0^1 f(x)\psi(x) dx.$$

Cette nouvelle équation a un sens si $u \in \mathcal{C}^1([0, 1])$, ou même si u appartient à l'espace de Sobolev $H_0^1([0, 1])$ constitué des fonctions $v \in L^2([0, 1])$ vérifiant $v(0) = v(1) = 0$ et telles qu'il existe une fonction $g \in L^2([0, 1])$ pour laquelle

$$v(x) = \int_0^x g(t) dt \quad \text{p.p.}$$

La primitive $x \mapsto \int_0^x g(t) dt$ est l'unique représentant continu de la classe v ; les conditions $v(0) = v(1) = 0$ portent sur ce représentant. La fonction g , définie presque partout, est alors notée v' . La formulation intégrale revient au fait que pour toute fonction $\psi \in \mathcal{C}^1([0, 1])$ telle que $\psi(0) = \psi(1) = 0$,

$$\langle u, \psi \rangle = \ell(\psi),$$

où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire qui fait de $H_0^1([0, 1])$ un espace de Hilbert et qui est défini par

$$\langle u, v \rangle = \int_0^1 u'(x)v'(x) dx + \int_0^1 u(x)v(x) dx,$$

et ℓ est la forme linéaire sur $H_0^1([0, 1])$ donnée par

$$\ell(v) = \int_0^1 f(x)v(x) dx.$$

Le théorème de représentation de Riesz prouve alors l'existence d'une unique fonction $u \in H_0^1([0, 1])$ telle que $\langle u, v \rangle = \ell(v)$ quelle que soit la fonction v dans $H_0^1([0, 1])$. En particulier, u apporte une solution à la formulation intégrale du problème de Dirichlet, appelée solution *faible*.

La fonction u peut également être vue comme solution d'un problème variationnel grâce au théorème de Lax-Milgram. En effet, ce théorème indique que u minimise la fonctionnelle

$$\frac{1}{2} \int_0^1 ((v'(x))^2 + (v(x))^2) dx - \int_0^1 f(x)v(x) dx.$$

Il s'agit là du *principe de Dirichlet*.

Pour revenir à la solution classique, on montre que la solution faible u est assez régulière en commençant par observer que $u'' = u - f$ est continue sur l'intervalle $[0, 1]$, où l'égalité est à prendre au sens faible. Plus précisément, si on note $g = u'$, alors la dérivée g' au sens faible est continue sur $[0, 1]$, c'est-à-dire qu'il existe une fonction h de $\mathcal{C}^0([0, 1])$ telle que pour toute fonction $\psi \in \mathcal{C}^1([0, 1])$ avec $\psi(0) = \psi(1) = 0$,

$$\int_0^1 g(x)\psi'(x) dx = - \int_0^1 h(x)\psi(x) dx.$$

En effet, on peut partir de la formulation intégrale précédente, et poser $h = u - f$ et $g = u'$. En choisissant une primitive H de h dont la constante est ajustée pour que $\int_0^1 (g - H) = 0$, on obtient après intégration par parties

$$\int_0^1 (g(x) - H(x))\psi'(x) dx = 0$$

pour toute fonction $\psi \in \mathcal{C}^1([0, 1])$ avec $\psi(0) = \psi(1) = 0$. Enfin, par densité, il existe une suite $(\psi_n)_{n \geq 1}$ de fonctions dans $\mathcal{C}^1([0, 1])$ vérifiant toutes $\psi_n(0) = \psi_n(1) = 0$ et telles que $\psi'_n \xrightarrow[n \rightarrow \infty]{L^2} g - H$. Pour le voir, on approche $g - H$ dans L^2 par des fonctions continues q_n , puis on remplace q_n par $q_n - \int_0^1 q_n$ et on pose $\psi_n(x) = \int_0^x q_n(t) dt$. L'intégrale nulle de $g - H$ garantit que la soustraction des moyennes conserve la convergence. On a finalement

$$\int_0^1 (g(x) - H(x))^2 dx = 0,$$

et donc $g = H$ presque partout, avec H de classe $\mathcal{C}^1$. Il en résulte que u est de classe $\mathcal{C}^2$ sur l'intervalle $[0, 1]$. On retrouve donc une solution forte.

5.6 Bases hilbertiennes

Définition 5.5. Une famille $(u_i)_{i \in I}$ d'éléments d'un espace de Hilbert H est :

- (i). *orthogonale* si $\langle u_i, u_j \rangle = 0$ pour tous indices $i, j \in I$ tels que $i \neq j$;
- (ii). *orthonormale* si elle est orthogonale et vérifie $\|u_i\| = 1$ pour tout indice $i \in I$;
- (iii). *une base hilbertienne* si elle est orthonormale et *totale*, c'est-à-dire vérifie la condition $\overline{\text{Vect}\{u_i, i \in I\}} = H$.

Remarque. D'après le théorème de Pythagore, pour toute famille orthogonale $(u_i)_{i \in I}$ et tout sous-ensemble fini $J \subset I$, on a

$$\left\| \sum_{i \in J} u_i \right\|^2 = \sum_{i \in J} \|u_i\|^2.$$

Ainsi, la famille $(u_i)_{i \in I}$ est libre si elle ne contient pas le vecteur nul.

Lemme 5.13 (Riesz-Fischer). Soit $(u_n)_{n \geq 0}$ une famille orthogonale d'un espace de Hilbert H . Alors, la série $\sum u_n$ converge dans H si et seulement si la série numérique $\sum \|u_n\|^2$ converge. De plus, $\left\| \sum_{n=0}^{\infty} u_n \right\|^2 = \sum_{n=0}^{\infty} \|u_n\|^2$.

Démonstration. On suppose que $S_n = \sum_{k=0}^n u_k$ converge dans H , vers une limite notée S . D'après le théorème de Pythagore, on observe que

$$\|S_n\|^2 = \sum_{k=0}^n \|u_k\|^2 \xrightarrow{n \rightarrow \infty} \|S\|^2.$$

Réciproquement, supposons que la série numérique $\sum \|u_n\|^2$ converge. Alors, à nouveau d'après le théorème de Pythagore, on a pour $m < n$,

$$\|S_n - S_m\|^2 = \sum_{k=m+1}^n \|u_k\|^2 \xrightarrow{m, n \rightarrow \infty} 0.$$

La suite $(S_n)_{n \geq 0}$ est donc de Cauchy, et par conséquent converge dans l'espace H qui est complet. $\square$

5.7 Orthonormalisation de Gram-Schmidt

Soit $(u_n)_{n \geq 0}$ une famille libre dans un espace de Hilbert H . On définit par récurrence une famille de vecteurs $(e_n)_{n \geq 0}$ comme suit. On commence par poser $e_0 = \frac{u_0}{\|u_0\|}$. Puis, les vecteurs $e_0, \dots, e_{n-1}$ étant construits, on pose

$$f_n = u_n - \sum_{k=0}^{n-1} \langle u_n, e_k \rangle e_k \quad \text{et} \quad e_n = \frac{f_n}{\|f_n\|}.$$

Alors, la famille $(e_n)_{n \geq 0}$ est orthonormale et pour tout entier $n \geq 0$, on observe que $\text{Vect}\{e_0, \dots, e_n\} = \text{Vect}\{u_0, \dots, u_n\}$.

Théorème 5.14. *Tout espace de Hilbert séparable admet une base hilbertienne (dénombrable).*

Démonstration. On peut supposer que H est de dimension infinie. Sinon, le résultat est clair puisqu'il suffit d'orthonormaliser une base quelconque de H .

Soit $(v_k)_{k \geq 0}$ une suite dense d'éléments de H . Pour tout entier $k \geq 0$, on note $F_k = \text{Vect}\{v_0, \dots, v_k\}$. Les F_k forment une suite croissante d'espaces vectoriels de dimension finie dont l'union est dense dans H . On pose ensuite $E_0 = F_{k_0}$, où k_0 est le plus petit entier k tel que $\dim F_k = 1$. Puis par récurrence, des espaces vectoriels $E_0, \dots, E_{n-1}$ étant construits, on pose $E_n = F_{k_n}$, où k_n est le plus petit entier k tel que $\dim F_k = \dim E_{n-1} + 1$. Les espaces vectoriels E_n ainsi définis forment donc une suite croissante de sous-espaces vectoriels de dimension finie dont l'union est dense dans H , et qui vérifient $\dim E_n = n + 1$ pour tout $n \in \mathbb{N}$. On remarque que pour tout entier $n \geq 1$, le vecteur v_{k_n} n'appartient pas à $E_{n-1} = \text{Vect}\{v_{k_0}, \dots, v_{k_{n-1}}\}$. La famille $(v_{k_n})_{n \geq 0}$ est donc libre. Comme cette famille est totale dans H , on peut en tirer une base hilbertienne en lui appliquant le procédé d'orthonormalisation de Gram-Schmidt. $\square$

Théorème 5.15. *Soit $(e_n)_{n \geq 0}$ une base hilbertienne d'un espace de Hilbert séparable H . Alors :*

- (i). *pour tout vecteur $u \in H$, on a $u = \sum_{n=0}^{\infty} \langle u, e_n \rangle e_n$;*
- (ii). *pour tout vecteur $u \in H$, on a $\|u\|^2 = \sum_{n=0}^{\infty} |\langle u, e_n \rangle|^2$, qui constitue l'identité de Bessel-Parseval ;*
- (iii). *pour toute suite $(c_n)_{n \geq 0}$ de réels tels que la série $\sum c_n^2$ converge, la série $\sum_{n=0}^{\infty} c_n e_n$ converge dans H .*

Démonstration. (i). On commence par observer que la projection du vecteur u sur l'espace vectoriel $E_n = \text{Vect}\{e_0, \dots, e_n\}$ s'écrit

$$u_n = p_{E_n}(u) = \sum_{k=0}^n \langle u, e_k \rangle e_k.$$

Le théorème de Pythagore implique alors que

$$\|u_n\|^2 = \sum_{k=0}^n \|\langle u, e_k \rangle e_k\|^2 = \sum_{k=0}^n |\langle u, e_k \rangle|^2.$$

Comme la projection p_{E_n} est 1-lipschitzienne, on a $\|u_n\|^2 \leq \|u\|^2$, de sorte que la série à termes positifs $\sum |\langle u, e_k \rangle|^2$ est convergente. Ainsi, la série $\sum \|\langle u, e_k \rangle e_k\|^2$ converge également, et le lemme de Riesz-Fischer implique que la série $\sum \langle u, e_k \rangle e_k$ converge dans H vers un certain vecteur v . Cependant, on a pour $m \geq n$,

$$\langle u_m - u, e_n \rangle = \langle u_m, e_n \rangle - \langle u, e_n \rangle = \sum_{k=0}^m \langle u, e_k \rangle \langle e_k, e_n \rangle - \langle u, e_n \rangle = 0,$$

puisque le terme $\langle e_k, e_n \rangle$ est non nul seulement si $k = n$, et ce terme vaut alors 1. Comme la suite (u_m) converge vers v , on en déduit par passage à la limite quand $m \rightarrow \infty$ que $\langle v - u, e_n \rangle = 0$ pour tout $n \in \mathbb{N}$. On a finalement $v = u$, ce qui donne le résultat.

- (ii). D'après le point précédent, on peut écrire le vecteur u comme somme de la série $\sum \langle u, e_n \rangle e_n$. Le lemme de Riesz-Fischer implique alors l'identité de Bessel-Parseval puisque

$$\|u\|^2 = \left\| \sum_{n=0}^{\infty} \langle u, e_n \rangle e_n \right\|^2 = \sum_{n=0}^{\infty} \|\langle u, e_n \rangle e_n\|^2 = \sum_{n=0}^{\infty} |\langle u, e_n \rangle|^2.$$

- (iii). On observe que $\|c_n e_n\|^2 = c_n^2$ pour tout n . Comme la série $\sum c_n^2$ converge, il en va de même de $\sum \|c_n e_n\|^2$, et le lemme de Riesz-Fischer permet de conclure à la convergence dans H de la série $\sum c_n e_n$. $\square$

Remarque. Ainsi, tout espace de Hilbert séparable de dimension infinie H est isométriquement isomorphe à l'espace de suites $\ell^2(\mathbb{N})$, par l'intermédiaire de l'application

$$\begin{aligned} H &\longrightarrow \ell^2(\mathbb{N}) \\ u &\longmapsto (\langle u, e_n \rangle)_{n \geq 0}, \end{aligned}$$

où $(e_n)_{n \geq 0}$ désigne une base hilbertienne de H . De plus, dans l'espace $\ell^2(\mathbb{N})$, un exemple de base hilbertienne est la base canonique formée des suites de la forme $w_n = (0, \dots, 0, 1, 0, \dots)$, où le coefficient égal à 1 a pour indice n .

5.8 Inégalités de Wirtinger et isopérimétrique

On munit l'ensemble quotient $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$ de la tribu borélienne et de la mesure de Lebesgue normalisée $\frac{dt}{2\pi}$. Alors, l'espace $L^2(\mathbb{T})$ s'identifie à l'espace des fonctions mesurables 2π -périodiques de $\mathbb{R}$ dans $\mathbb{C}$ dont le module au carré est intégrable sur une période, identifiées lorsqu'elles sont égales presque partout. Il s'agit d'un espace de Hilbert pour le produit scalaire

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t) \overline{g(t)} dt.$$

De plus, la famille formée des fonctions $e_n : t \mapsto e^{int}$, pour $n \in \mathbb{Z}$, est une famille totale d'après le théorème de Fejér. Il s'agit même d'une base hilbertienne. Pour toute fonction f de l'espace $L^2(\mathbb{T})$, on rappelle que le coefficient de Fourier d'indice n est défini par

$$c_n(f) = \langle f, e_n \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-int} dt.$$

Dans ce contexte, l'identité de Bessel-Parseval devient

$$\frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 dt = \sum_{n=-\infty}^{+\infty} |c_n(f)|^2.$$

On peut en donner comme application le résultat suivant.

Proposition 5.16 (Inégalité de Wirtinger). Soit f une fonction à valeurs complexes, 2π -périodique, de classe $\mathcal{C}^1$ et d'intégrale nulle sur une période. Alors, on a

$$\int_0^{2\pi} |f(t)|^2 dt \leq \int_0^{2\pi} |f'(t)|^2 dt.$$

De plus, il y a égalité si et seulement si f est de la forme $t \mapsto a e^{it} + b e^{-it}$, avec des coefficients quelconques $a, b \in \mathbb{C}$.

Démonstration. On peut bien sûr identifier la fonction f avec un élément de l'espace $L^2(\mathbb{T})$. Pour tout entier $n \in \mathbb{Z}$, une simple intégration par parties montre clairement que $c_n(f') = in c_n(f)$. En appliquant l'identité de Bessel-Parseval, et en observant que le coefficient de Fourier $c_0(f)$ est nul puisque la fonction f est d'intégrale nulle, on obtient

$$\int_0^{2\pi} |f(t)|^2 dt = 2\pi \sum_{n \in \mathbb{Z} \setminus \{0\}} |c_n(f)|^2 \leq 2\pi \sum_{n \in \mathbb{Z}} |in c_n(f)|^2 = 2\pi \sum_{n \in \mathbb{Z} \setminus \{0\}} |c_n(f')|^2 = \int_0^{2\pi} |f'(t)|^2 dt,$$

ce qui donne l'inégalité annoncée. De plus, on observe que la différence des deux intégrales s'écrit simplement

$$2\pi \sum_{n \in \mathbb{Z} \setminus \{0\}} (n^2 - 1) |c_n(f)|^2.$$

Le cas d'égalité correspond donc au fait que $c_n(f) = 0$ dès que $|n| \geq 2$. Les seules fonctions possibles sont donc $t \mapsto a e^{it} + b e^{-it}$, avec $a, b \in \mathbb{C}$. $\square$

Ce résultat permet lui-même d'établir l'inégalité isopérimétrique.

Proposition 5.17 (Inégalité isopérimétrique). On note γ une courbe fermée simple du plan, régulière de classe $\mathcal{C}^1$, de longueur L et entourant une aire A . Alors,

$$4\pi A \leq L^2,$$

avec égalité si et seulement si γ est un cercle.

Démonstration. Par homothétie, on peut supposer $L = 2\pi$. On peut aussi supposer que la courbe γ est parcourue dans le sens trigonométrique à vitesse 1 par une fonction $f : [0, 2\pi] \rightarrow \mathbb{C}$. Par translation, on peut aussi supposer que la fonction f est d'intégrale nulle. On décompose f sous la forme $x + iy$, où x et y sont deux fonctions à valeurs réelles. D'après la formule de Green-Riemann, l'aire délimitée par la courbe γ vaut

$$A = \frac{1}{2} \int_{\gamma} (x dy - y dx) = \frac{1}{2} \int_0^{2\pi} (x(t)y'(t) - y(t)x'(t)) dt.$$

Dans cette dernière expression, on observe que l'intégrande est la partie imaginaire du produit $f'(t)\overline{f(t)}$. Ainsi, l'aire peut s'écrire

$$A = \pi \Im \langle f', f \rangle.$$

En appliquant successivement l'inégalité de Cauchy-Schwarz et l'inégalité de Wirtinger, on obtient

$$A \leq \pi |\langle f', f \rangle| \leq \pi \|f'\| \|f\| \leq \pi \|f'\|^2.$$

Cependant, comme la courbe γ est supposée parcourue à vitesse 1 par la fonction f , on a en fait $\|f'\| = 1$. On en conclut que $A \leq \pi$, c'est-à-dire comme annoncé $4\pi A \leq L^2$, puisque $L = 2\pi$.

Enfin, s'il y a égalité dans l'inégalité isopérimétrique, il y a aussi égalité dans les inégalités de Wirtinger et de Cauchy-Schwarz. Par conséquent, la fonction f est de la forme $t \mapsto a e^{it} + b e^{-it}$, avec $a, b \in \mathbb{C}$. De plus, les fonctions f et f' sont colinéaires, donc f est soit de la forme $t \mapsto a e^{it}$, soit de la forme $t \mapsto b e^{-it}$. À nouveau comme la courbe γ est parcourue dans le sens trigonométrique et à vitesse 1, on a nécessairement $|a| = 1$, et donc $f : t \mapsto e^{i(t+\theta)}$ pour un $\theta \in \mathbb{R}$. La courbe γ est donc un cercle. $\square$

Théorème de Baire et applications

Référence : [6].

6.1 Le théorème de Baire

Définition 6.1. On dit qu'un espace métrique (X, d) est un *espace de Baire* si toute intersection dénombrable d'ouverts denses dans X est dense dans X , c'est-à-dire si pour toute suite $(O_n)_{n \geq 1}$ d'ouverts de X ,

$$[\forall n \geq 1 \quad \overline{O_n} = X] \quad \implies \quad \overline{\bigcap_{n=1}^{\infty} O_n} = X.$$

Remarques. • *Formulation équivalente* : (X, d) est un espace de Baire si toute union dénombrable de fermés d'intérieurs vides est d'intérieur vide, c'est-à-dire si pour toute suite $(F_n)_{n \geq 1}$ de fermés de X ,

$$[\forall n \geq 1 \quad \overset{\circ}{F_n} = \emptyset] \quad \implies \quad \bigcup_{n=1}^{\infty} \overset{\circ}{F_n} = \emptyset.$$

- *Terminologie* : Une intersection dénombrable d'ouverts est appelée un G_δ . On remarque qu'une intersection dénombrable de G_δ est un G_δ . Dans un espace de Baire, toute intersection dénombrable de G_δ denses est un G_δ dense. Une union dénombrable de fermés est appelée un F_σ . Dans un espace de Baire, toute union dénombrable de F_σ d'intérieurs vides est un F_σ d'intérieur vide.
- Si (X, d) est un espace de Baire non vide et si $(F_n)_{n \geq 1}$ est une suite de fermés de X telle que $\bigcup_{n=1}^{\infty} F_n = X$, alors $\overset{\circ}{F_{n_0}} \neq \emptyset$ pour un entier $n_0 \geq 1$.

Définition 6.2. Soit (X, d) un espace de Baire et soit A une partie de X .

- On dit que A est un *résiduel* si A contient un G_δ dense dans X .
- On dit que A est *maigre* si A est contenu dans un F_σ d'intérieur vide de X .

Proposition 6.1. Soit (X, d) un espace de Baire et soit A une partie de X . Alors, A est maigre si et seulement si $X \setminus A$ est un résiduel.

Théorème 6.2 (Baire). Tout espace métrique complet est un espace de Baire.

Démonstration. Soit (X, d) un espace métrique complet et soit $(O_n)_{n \geq 1}$ une suite d'ouverts denses dans X . Il suffit de montrer que pour tout ouvert non vide V de X , on a $V \cap \bigcap_{n=1}^{\infty} O_n \neq \emptyset$. Soit V un tel ouvert. On construit par récurrence une suite $(B_n)_{n \geq 1}$ de boules fermées de X de rayons strictement positifs telles que :

- (i). $B_1 \subset O_1 \cap V$;
- (ii). $B_{n+1} \subset O_{n+1} \cap \overset{\circ}{B_n}$, pour tout $n \geq 1$;
- (iii). B_n est de rayon inférieur à 2^{-n} , pour tout $n \geq 1$.

Pour commencer, l'ouvert O_1 est dense dans X , donc $O_1 \cap V \neq \emptyset$. Comme $O_1 \cap V$ est un ouvert non vide, il contient une boule fermée de rayon inférieur à $1/2$, notée B_1 . Supposons ensuite que les boules $B_1, \dots, B_n$ sont données. Comme O_{n+1} est un ouvert dense dans X , l'ouvert $O_{n+1} \cap \overset{\circ}{B_n}$ est non vide, donc contient une boule fermée de rayon au plus $2^{-(n+1)}$, notée B_{n+1} et qui est convenable.

Par construction, les boules $(B_n)_{n \geq 1}$ forment une suite décroissante de fermés non vides de X dont le diamètre tend vers 0. Comme X est complet, on en déduit l'existence d'un élément $x \in X$ tel que $\bigcap_{n=1}^{\infty} B_n = \{x\}$. Ainsi, $x \in B_1 \subset V$ et $x \in B_n \subset O_n$ pour tout $n \geq 1$, de sorte que $x \in V \cap \bigcap_{n=1}^{\infty} O_n$. Ce dernier ensemble est donc non vide. $\square$

Remarque. La preuve ci-dessus utilise le lemme des fermés emboîtés : si (X, d) est un espace métrique complet et si $(F_n)_{n \geq 1}$ est une suite décroissante de fermés non vides de X avec $\text{diam } F_n \xrightarrow{n \rightarrow \infty} 0$, alors on a $\bigcap_{n=1}^{\infty} F_n = \{x\}$, pour un $x \in X$. En effet, dans ce cas, pour tout entier $n \geq 1$, il existe $x_n \in F_n$. Pour $n \geq m$, on a $x_n, x_m \in F_m$, donc $d(x_n, x_m) \leq \text{diam } F_m$, qui est arbitrairement petit dès que m est assez grand. La suite $(x_n)_{n \geq 1}$ vérifie donc le critère de Cauchy dans X complet, donc converge vers un élément x . On a $x_n \in F_m$ pour tout $n \geq m$, et F_m est fermé, donc $x \in F_m$. Cela est valable pour tout $m \geq 1$, si bien que $x \in \bigcap_{m=1}^{\infty} F_m$.

Lemme 6.3. Soit (X, d) un espace de Baire et $(F_n)_{n \geq 1}$ une suite de fermés de X telle que $\bigcup_{n=1}^{\infty} F_n = X$. Alors, $\bigcup_{n=1}^{\infty} \overset{\circ}{F}_n$ est un ouvert dense dans X .

Démonstration. Il est clair que $G = X \setminus \bigcup_{n=1}^{\infty} \overset{\circ}{F}_n$ est un fermé de X . Il reste à montrer que G est d'intérieur vide. Cependant, pour tout $n \geq 1$, l'ensemble $G \cap F_n$ est fermé et d'intérieur vide, car $G \cap \overset{\circ}{F}_n \subset G \cap F_n = \emptyset$. Comme (X, d) est un espace de Baire, on en déduit que l'union $\bigcup_{n=1}^{\infty} (G \cap F_n)$ est d'intérieur vide également, mais il s'agit de $G \cap \bigcup_{n=1}^{\infty} F_n = G \cap X = G$. $\square$

Exemples. • L'ensemble $\mathbb{R} \setminus \mathbb{Q}$ est un G_δ dense de $\mathbb{R}$; l'ensemble $\mathbb{Q}$ est un F_σ d'intérieur vide.

• L'ensemble $\mathbb{Q}$ est dense mais ce n'est pas un G_δ . En effet, notons $(x_n)_{n \geq 1}$ une énumération de $\mathbb{Q}$ et supposons que $\mathbb{Q} = \bigcap_{n=1}^{\infty} O_n$, avec O_n ouvert. Alors, $\bigcap_{n=1}^{\infty} (O_n \setminus \{x_n\}) \subset \bigcap_{n=1}^{\infty} (\mathbb{Q} \setminus \{x_n\}) = \mathbb{Q} \setminus \bigcup_{n=1}^{\infty} \{x_n\} = \emptyset$. Cependant, chaque $O_n \setminus \{x_n\}$ est un ouvert dense, ce qui implique que $\bigcap_{n=1}^{\infty} (O_n \setminus \{x_n\})$ devrait être dense, et donc non vide.

6.2 Applications aux espaces vectoriels normés

6.2.1 Un espace vectoriel normé muni d'une base infinie dénombrable ne peut être complet

Théorème 6.4. Soit E un espace vectoriel normé admettant une base infinie dénombrable. Alors, E n'est pas complet.

Démonstration. Soit $(e_n)_{n \geq 1}$ une base de E . Pour tout $n \geq 1$, on considère $F_n = \text{Vect}\{e_1, \dots, e_n\}$. Il s'agit d'un sous-espace vectoriel de dimension finie, donc fermé. De plus, il est d'intérieur vide. (En effet, s'il contenait une boule, il contiendrait par translation une boule centrée en 0, puis par homothétie tout l'espace E .) Supposons E complet. Le théorème de Baire implique que $E = \bigcup_{n=1}^{\infty} F_n$ est d'intérieur vide, ce qui est absurde. $\square$

6.2.2 Théorème de l'application ouverte et théorème de Banach

Théorème 6.5 (application ouverte). Soit E et F deux espaces de Banach, et soit $T : E \rightarrow F$ une application linéaire continue surjective. Alors, T est une application ouverte, c'est-à-dire que pour tout ouvert Ω de E , l'ensemble $T(\Omega)$ est un ouvert de F .

Démonstration. On note $B(r)$ la boule ouverte de E centrée en 0 et de rayon $r > 0$. Alors, l'adhérence $\overline{T(B(r))}$ est un voisinage de 0 dans F . En effet, pour tout $n \geq 1$, l'ensemble $F_n = \overline{T(B(nr))}$ est un fermé de F . Comme T est surjective, on a $F = \bigcup_{n=1}^{\infty} F_n$. Comme F est complet, on déduit du théorème de Baire que $F_n \neq \emptyset$ pour un certain $n \geq 1$. Ainsi, F_n contient une boule ouverte $B(x, \rho)$. Il existe une suite $(x_p)_{p \geq 1}$ dans $B(nr)$ telle que $T(x_p) \xrightarrow[p \rightarrow \infty]{} x$. De même, si $y \in B(x, \rho)$, il existe une suite $(y_p)_{p \geq 1}$ dans $B(nr)$ telle que $T(y_p) \xrightarrow[p \rightarrow \infty]{} y$. On a donc $y_p - x_p \in B(2nr)$ pour tout $p \geq 1$, ainsi que $T(y_p - x_p) \xrightarrow[p \rightarrow \infty]{} y - x$, de sorte que $y - x \in F_{2n}$. Finalement, $B(x, \rho) - x \subset F_{2n}$, c'est-à-dire que F_{2n} contient la boule ouverte $B(0, \rho)$. On en conclut que

$$B\left(0, \frac{\rho}{2n}\right) = \frac{1}{2n}B(0, \rho) \subset \frac{1}{2n}F_{2n} = \overline{T(B(r))},$$

ce qui prouve que $\overline{T(B(r))}$ est bien un voisinage de 0 dans F .

De plus, pour tout $r > 0$, on a l'inclusion $\overline{T(B(r/2))} \subset T(B(r))$. En effet, posons $V_n = B(2^{-n}r)$ pour tout $n \geq 0$. Soit $y_1 \in \overline{T(V_1)}$. Comme $\overline{T(V_2)}$ est un voisinage de 0 d'après ce qui précède, on a $\overline{T(V_1)} \subset T(V_1) + \overline{T(V_2)}$, de sorte qu'il existe $x_1 \in V_1$ et $y_2 \in \overline{T(V_2)}$ tels que $y_1 - y_2 = T(x_1)$. De même, pour tout $n \geq 2$, on peut obtenir un point $x_n \in V_n$ et un point $y_n \in \overline{T(V_n)}$ tels que $y_n - y_{n+1} = T(x_n)$. Pour tout $n \geq 1$, on a $\|x_n\| < 2^{-n}r$, donc la série $\sum x_n$ converge normalement. En notant $x = \sum_{n=1}^{\infty} x_n$, on obtient $T(x) = y_1$, puisque pour tout $N \geq 1$,

$$T\left(\sum_{n=1}^N x_n\right) = \sum_{n=1}^N T(x_n) = \sum_{n=1}^N (y_n - y_{n+1}) = y_1 - y_{N+1},$$

avec $y_{N+1} \xrightarrow[N \rightarrow \infty]{} 0$, car $y_{N+1} \in \overline{T(V_{N+1})}$ et T est continue. On en déduit finalement que $y_1 \in T(V_0)$ puisque

$$\|x\| \leq \sum_{n=1}^{\infty} \|x_n\| < r \sum_{n=1}^{\infty} 2^{-n} = r.$$

Ce qui précède montre que $T(V_0)$ contient $\overline{T(V_1)}$, qui est un voisinage de 0. Donc, $T(B(r))$ est un voisinage de 0 pour tout $r > 0$. Ainsi, si U est un voisinage de 0, alors $T(U)$ est aussi un voisinage de 0. Finalement, si Ω est un ouvert qui contient un point x , alors $\Omega - x$ est un voisinage de 0, et donc $T(\Omega - x) = T(\Omega) - T(x)$ est un voisinage de 0 également. Ainsi, $T(\Omega)$ est un voisinage de $T(x)$. Cela étant valable pour tout $x \in \Omega$, on en déduit que $T(\Omega)$ est ouvert. $\square$

Corollaire 6.6 (théorème de Banach). Soit E et F deux espaces de Banach, et soit $T : E \rightarrow F$ une application linéaire continue et bijective. Alors, l'application T^{-1} est continue.

Démonstration. On pose $U = T^{-1}$. Le théorème de l'application ouverte assure que pour tout ouvert Ω de E , l'ensemble $U^{-1}(\Omega)$ est un ouvert de F . L'image réciproque de tout ouvert de E étant un ouvert de F , on en conclut que U est continue. $\square$

6.2.3 Théorème de Banach-Steinhaus

Soit E un espace de Banach et soit F un espace vectoriel normé. On note $\mathcal{L}(E, F)$ l'espace vectoriel des applications linéaires continues de E dans F , muni de la norme $\|f\| = \sup_{\|x\| \leq 1} \|f(x)\|$.

Théorème 6.7 (Banach-Steinhaus). Soit H une partie de $\mathcal{L}(E, F)$. Alors, l'une des deux propriétés suivantes est vérifiée :

$$\sup_{f \in H} \|f\| < \infty \quad \text{ou} \quad \exists x \in E \quad \sup_{f \in H} \|f(x)\| = \infty.$$

Démonstration. Pour tout entier $K \geq 0$, on considère l'ouvert Ω_K formé des points $x \in E$ tels que $\sup_{f \in H} \|f(x)\| > K$. (Il s'agit bien d'un ouvert car toutes les applications f sont continues.) On a deux situations possibles :

- Si Ω_K est dense dans E pour tout $K \geq 0$, alors le théorème de Baire implique que $\bigcap_{K=1}^{\infty} \Omega_K$ est dense dans E (qui est complet), et donc non vide. Chaque point x de cette intersection vérifie $\sup_{f \in H} \|f(x)\| = \infty$.
- Sinon, il existe un ouvert Ω_K qui n'est pas dense dans E . Ainsi, il existe une boule ouverte $B(x_0, \rho)$ de E qui ne rencontre pas Ω_K . Pour tout point $y \in B(x_0, \rho)$ et toute application $f \in H$, on a $\|f(y)\| \leq K$, donc pour $\|x\| < \rho$ et $f \in H$,

$$\|f(x)\| = \|f(x + x_0) - f(x_0)\| \leq \|f(x + x_0)\| + \|f(x_0)\| \leq 2K.$$

Par continuité de f , on a en fait $\|f(x)\| \leq 2K$ dès que $\|x\| \leq \rho$. On en déduit que $\|f(x)\| = \frac{1}{\rho} \|f(\rho x)\| \leq \frac{2K}{\rho}$ pour $\|x\| \leq 1$ et $f \in H$, et par conséquent que $\sup_{f \in H} \|f\| \leq \frac{2K}{\rho} < \infty$. $\square$

Corollaire 6.8. Soit $(f_n)_{n \geq 1}$ une suite d'applications linéaires continues de E dans F qui converge simplement vers f . Alors, f est linéaire et continue.

Démonstration. La linéarité de f est claire. Pour montrer la continuité, on applique le théorème de Banach-Steinhaus à la partie $H = \{f_n, n \geq 1\}$. Comme $(f_n)_{n \geq 1}$ converge simplement, on a $\sup_{n \geq 1} \|f_n(x)\| < \infty$ pour tout $x \in E$. Alors, on a nécessairement $\sup_{n \geq 1} \|f_n\| < \infty$ d'après le théorème de Banach-Steinhaus. Ainsi, il existe $M > 0$ tel que $\|f_n\| \leq M$ pour tout $n \geq 1$, et on a $\|f_n(x)\| \leq M\|x\|$ pour tout $x \in E$. En passant à la limite, on obtient $\|f(x)\| \leq M\|x\|$ pour tout $x \in E$, ce qui montre la continuité de f . $\square$

Corollaire 6.9. Soit E_1 un espace de Banach et E_2 un espace vectoriel normé. Soit $B : E_1 \times E_2 \rightarrow F$ une application bilinéaire dont les applications partielles $B(x, \cdot)$ et $B(\cdot, y)$ sont continues quels que soient $x \in E_1$ et $y \in E_2$. Alors, B est continue sur $E_1 \times E_2$.

Démonstration. Soit la partie $H = \{B(\cdot, y), \|y\| \leq 1\}$ de $\mathcal{L}(E_1, F)$. Pour tout $x \in E_1$, l'application $y \mapsto B(x, y)$ est continue donc $\sup_{\|y\| \leq 1} \|B(x, y)\|$ est fini. D'après le théorème de Banach-Steinhaus, il existe $M > 0$ tel que pour tous $x \in E_1$ et $\|y\| \leq 1$, on ait $\|B(x, y)\| \leq M\|x\|$. Finalement, on a $\|B(x, y)\| \leq M\|x\|\|y\|$ pour tous $x \in E_1$ et $y \in E_2$, ce qui garantit la continuité de B sur $E_1 \times E_2$. $\square$

6.3 Applications en analyse réelle

6.3.1 Une dérivée est continue sur un ensemble dense

Lemme 6.10. Soit (X, d) et (Y, δ) deux espaces métriques, avec (X, d) complet. Soit $(f_n)_{n \geq 1}$ une suite de fonctions continues de X dans Y qui converge simplement vers une fonction f . Alors, l'ensemble des points de continuité de f est un résiduel.

Démonstration. Pour tout $\varepsilon > 0$ et tout $n \geq 1$, on note $F_{n,\varepsilon}$ l'ensemble fermé constitué des $x \in X$ tels que $\delta(f_n(x), f_p(x)) \leq \varepsilon$ pour tout $p \geq n$, puis on considère l'ouvert de X défini par $\Omega_\varepsilon = \bigcup_{n=1}^{\infty} F_{n,\varepsilon}^\circ$. Pour tout $x \in X$, la suite $(f_n(x))_{n \geq 1}$ converge, donc est de Cauchy, ce qui implique que $x \in F_{n,\varepsilon}$ pour n assez grand. Ainsi, $X = \bigcup_{n=1}^{\infty} F_{n,\varepsilon}$ et le lemme précédent montre que Ω_ε est dense dans X , puisque X est un espace de Baire.

De plus, si $x_0 \in \Omega_\varepsilon$, il existe $n \geq 1$ tel que $x_0 \in F_{n,\varepsilon}^\circ$. Comme f_n est continue, il existe un voisinage V de x_0 tel que $x_0 \in V \subset F_{n,\varepsilon}^\circ$, ainsi que $\delta(f_n(x), f_n(x_0)) \leq \varepsilon$ pour tout $x \in V$. Donc, pour $x \in V$ et $p \geq n$, on a

$$\delta(f_p(x_0), f_p(x)) \leq \delta(f_p(x_0), f_n(x_0)) + \delta(f_n(x_0), f_n(x)) + \delta(f_p(x), f_n(x)) \leq 3\varepsilon.$$

En faisant tendre p vers l'infini, on en déduit que $\delta(f(x_0), f(x)) \leq 3\varepsilon$ pour tout $x \in V$. Autrement dit, l'ensemble $\Omega = \bigcap_{\varepsilon > 0} \Omega_\varepsilon$ est formé de points où f est continue. Cependant, les ensembles Ω_ε sont décroissants quand $\varepsilon \downarrow 0$. On peut donc réécrire Ω comme un G_δ . D'après le théorème de Baire, l'espace (X, d) est de Baire, donc Ω est dense dans X . L'ensemble des points de continuité de f est donc un résiduel. $\square$

Théorème 6.11. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une application dérivable sur $\mathbb{R}$. Alors, l'ensemble des points de continuité de f' est dense.

Démonstration. Pour tout entier $n \geq 1$, on considère la fonction $f_n : x \mapsto (f(x + 1/n) - f(x))n$. La suite $(f_n)_{n \geq 1}$ est formée de fonctions continues sur $\mathbb{R}$, qui est complet, et converge simplement vers f' . Le lemme précédent implique que l'ensemble des points de continuité de f' est un résiduel, donc est dense dans $\mathbb{R}$. $\square$

6.3.2 Fonctions nulles à l'infini

Proposition 6.12. Soit $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ une fonction continue pour laquelle on a $f(nx) \xrightarrow{n \rightarrow \infty} 0$ pour tout $x > 0$. Alors, on a $f(x) \xrightarrow{x \rightarrow +\infty} 0$.

Démonstration. Soit $\varepsilon > 0$. Pour tout $n \geq 1$, on considère l'ensemble F_n des $x \geq 1$ tels que $|f(px)| \leq \varepsilon$ pour tout $p \geq n$. L'hypothèse faite sur f implique que $[1, +\infty[\subset \bigcup_{n=1}^{\infty} F_n$. Les ensembles F_n sont donc des fermés, par continuité de f , dont l'union coïncide avec $[1, +\infty[$, qui est complet, donc un espace de Baire. On en déduit que $F_{n_0} \neq \emptyset$ pour un certain $n_0 \geq 1$. Ainsi, il y a un intervalle $]\alpha, \beta[\subset F_{n_0}$. On choisit un entier $n_1 \geq 1$ tel que $(n_1 + 1)\alpha < n_1\beta$, ce qui revient à $n_1(\beta - \alpha) > \alpha$. Pour $x > \alpha \max\{n_0, n_1\}$, on a $p\alpha < x \leq (p + 1)\alpha$ avec $p \geq \max\{n_0, n_1\}$, de sorte que $\alpha < x/p < \beta$ et $|f(p x/p)| \leq \varepsilon$. Finalement, $|f(x)| \leq \varepsilon$ pour x assez grand, donc f tend vers 0 à l'infini. $\square$

6.3.3 Fonctions continues nulle part dérivables

Notons $\mathcal{C}$ l'espace vectoriel des fonctions continues définies sur l'intervalle $I = [0, 1]$ et à valeurs dans $\mathbb{R}$, muni de la norme $\|\cdot\|_\infty$. On rappelle que c'est un espace de Banach, car il est fermé dans l'espace vectoriel des fonctions bornées, qui est complet.

Théorème 6.13. *L'ensemble des fonctions de $\mathcal{C}$ qui sont nulle part dérivables est un résiduel de $\mathcal{C}$, donc en particulier est dense dans $\mathcal{C}$.*

Démonstration. Pour tout $\varepsilon > 0$ et tout $n \geq 1$, on considère l'ensemble $U_{\varepsilon,n}$ des fonctions $f \in \mathcal{C}$ telles que pour tout $x \in I$, il existe un $y \in I$ vérifiant les conditions $0 < |y - x| < \varepsilon$ et $\left| \frac{f(y) - f(x)}{y - x} \right| > n$.

Commençons par montrer que $U_{\varepsilon,n}$ est ouvert. En effet, notons $F_{\varepsilon,n}$ son complémentaire dans $\mathcal{C}$, puis considérons une suite $(f_p)_{p \geq 1}$ dans $F_{\varepsilon,n}$ qui converge vers $f \in \mathcal{C}$. Pour tout $p \geq 1$, comme $f_p \in F_{\varepsilon,n}$, alors il existe $x_p \in I$ tel que $|f_p(y) - f_p(x_p)| \leq n|y - x_p|$ pour tout $y \in I$ vérifiant $|y - x_p| < \varepsilon$. Par compacité de I , on peut extraire une sous-suite $(x_{\varphi(p)})_{p \geq 1}$ qui converge vers un point $x \in I$. Dès lors, pour tout $y \in I$ tel que $0 < |y - x| < \varepsilon$, il existe un entier $p_0 \geq 1$ tel que $0 < |y - x_{\varphi(p)}| < \varepsilon$ pour tout entier $p \geq p_0$. On a donc $|f_{\varphi(p)}(y) - f_{\varphi(p)}(x_{\varphi(p)})| \leq n|y - x_{\varphi(p)}|$. Comme $(f_p)_{p \geq 1}$ converge uniformément vers f et comme $(x_{\varphi(p)})_{p \geq 1}$ converge vers x , on a $f_{\varphi(p)}(x_{\varphi(p)}) \xrightarrow{p \rightarrow \infty} f(x)$, et donc $|f(y) - f(x)| \leq n|y - x|$. On en déduit que f appartient à $F_{\varepsilon,n}$, de sorte que cet ensemble est fermé. Finalement, l'ensemble $U_{\varepsilon,n}$ est bien ouvert.

Montrons ensuite que $U_{\varepsilon,n}$ est dense dans $\mathcal{C}$. Pour $f \in \mathcal{C}$ et $\delta > 0$, il convient de trouver une fonction $g \in U_{\varepsilon,n}$ telle que $\|f - g\|_\infty \leq \delta$. On cherche g sous la forme $x \mapsto f(x) + \delta \sin(Nx)$. D'après le théorème de Heine, la fonction f est uniformément continue sur I , donc il existe $\alpha \in]0, \varepsilon[$ tel que pour tous $x, y \in I$ vérifiant $|x - y| < \alpha$, on ait $|f(x) - f(y)| < \delta/4$. On choisit un entier $N > 8\pi$ tel que $4\pi < N\alpha$ et $\delta N > 8\pi n$, et on pose $g(x) = f(x) + \delta \sin(Nx)$ pour tout $x \in I$. Maintenant, pour tout $x \in I$, on peut trouver $y \in I$ tel que $2\pi \leq |Nx - Ny| \leq 4\pi$ et $|\sin(Nx) - \sin(Ny)| \geq 1$. (En effet, on choisit $z \in [0, N]$ tel que $|\sin(Nx) - \sin(z)| \geq 1$, ainsi que $z \in [Nx + 2\pi, Nx + 4\pi]$ ou $z \in [Nx - 4\pi, Nx - 2\pi]$, puisqu'un de ces deux intervalles est dans $[0, N]$, et on prend finalement $y = z/N$.) On a d'une part

$$\frac{2\pi}{N} \leq |x - y| \leq \frac{4\pi}{N} < \alpha \quad \Rightarrow \quad \left| \frac{f(y) - f(x)}{y - x} \right| < \frac{\delta}{4} \frac{N}{2\pi} = \frac{\delta N}{8\pi}.$$

D'autre part, on a aussi

$$\left| \frac{\delta \sin(Ny) - \delta \sin(Nx)}{y - x} \right| \geq \frac{\delta}{|y - x|} \geq \frac{\delta N}{4\pi}.$$

Enfin, d'après l'inégalité triangulaire, on obtient

$$\left| \frac{g(y) - g(x)}{y - x} \right| \geq \left| \frac{\delta \sin(Ny) - \delta \sin(Nx)}{y - x} \right| - \left| \frac{f(y) - f(x)}{y - x} \right| \geq \frac{\delta N}{8\pi} > n.$$

Il en résulte que $g \in U_{\varepsilon,n}$ et $\|f - g\|_\infty = \delta$, ce qui prouve bien la densité de $U_{\varepsilon,n}$ dans $\mathcal{C}$.

Considérons finalement l'ensemble $R = \bigcap_{n=1}^{\infty} U_{1/n,n}$. D'après le théorème de Baire, il s'agit d'un G_δ dense dans $\mathcal{C}$. De plus, si $f \in R$, alors $f \in U_{1/n,n}$ pour tout entier $n \geq 1$, c'est-à-dire que pour tout $x \in I$, il existe $x_n \in I$ avec $0 < |x - x_n| < \frac{1}{n}$ et $\left| \frac{f(x) - f(x_n)}{x - x_n} \right| > n$. Ainsi, on a $x_n \xrightarrow{n \rightarrow \infty} x$ et $\left| \frac{f(x) - f(x_n)}{x - x_n} \right| \xrightarrow{n \rightarrow \infty} +\infty$, ce qui montre que f n'est pas dérivable au point x . L'ensemble R est donc formé de fonctions nulle part dérivables. $\square$

6.3.4 Caractérisation des polynômes sur $\mathbb{R}$

Théorème 6.14 (Balaguer-Corominas, 1954). Soit $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ une application de classe $\mathcal{C}^\infty$ telle que pour tout $x \in \mathbb{R}$, il existe $n \in \mathbb{N}$ vérifiant $\varphi^{(n)}(x) = 0$. Alors, φ est une fonction polynomiale sur $\mathbb{R}$.

Démonstration. Pour tout $n \in \mathbb{N}$, on note F_n l'ensemble des $x \in \mathbb{R}$ tels que $\varphi^{(n)}(x) = 0$. On note ensuite $\Omega = \bigcup_{n \in \mathbb{N}} \overset{\circ}{F}_n$, ainsi que $F = \mathbb{R} \setminus \Omega$.

Commençons par montrer que φ coïncide avec une fonction polynomiale sur chaque composante connexe de l'ouvert Ω . Soit $]a, b[$ une telle composante connexe, et soit x_0 un point de $]a, b[$. Alors il existe $n \in \mathbb{N}$ et $\alpha > 0$ tels que la fonction $\varphi^{(n)}$ est nulle sur l'intervalle $]x_0 - \alpha, x_0 + \alpha[$. Donc, φ coïncide avec une fonction polynomiale P sur cet intervalle, et l'ensemble Γ des réels $c \in]x_0, b[$ tels que $\varphi(x) = P(x)$ pour tout $x \in [x_0, c]$ est non vide. On note β son supremum. Si $\beta < b$, alors il existe de même un intervalle de la forme $]\beta - \eta, \beta + \eta[$, avec $\beta + \eta < b$, sur lequel la fonction φ coïncide avec une fonction polynomiale Q . On a $P = \varphi = Q$ sur l'intervalle $]x_0, \beta[\cap]\beta - \eta, \beta + \eta[$, si bien que P et Q sont les mêmes fonctions. On en déduit que $\varphi = P$ sur $]\beta, \beta + \eta[$, de sorte que $\beta + \eta \in \Gamma$, ce qui contredit la définition de β . Ainsi, $\beta = b$, et la fonction φ coïncide avec P sur l'intervalle $[x_0, b[$. On peut montrer de même que c'est aussi le cas sur $]a, x_0]$. Finalement, la fonction φ coïncide avec une fonction polynomiale sur $]a, b[$.

Montrons maintenant que le fermé F n'admet pas de point isolé. Supposons par l'absurde l'existence d'un tel point, noté x_0 . Alors, il existe $\varepsilon > 0$ tel que $]x_0 - \varepsilon, x_0 + \varepsilon[\cap F = \{x_0\}$. En particulier, $]x_0 - \varepsilon, x_0[$ est inclus dans Ω , et donc la fonction φ coïncide avec une fonction polynomiale P sur cet intervalle. De même, φ coïncide avec une fonction polynomiale Q sur l'intervalle $]x_0, x_0 + \varepsilon[$. Par continuité des dérivées successives de φ au point x_0 , on a $P^{(n)}(x_0) = \varphi^{(n)}(x_0) = Q^{(n)}(x_0)$ pour tout $n \in \mathbb{N}$. D'après la formule de Taylor appliquée au point x_0 , on déduit que les deux fonctions P et Q sont égales. Ainsi, $\varphi = P$ sur l'intervalle $]x_0 - \varepsilon, x_0 + \varepsilon[$. En choisissant un entier $n \geq 0$ tel que $P^{(n+1)} = 0$, on observe que $\varphi^{(n+1)} = P^{(n+1)} = 0$ sur cet intervalle. Cela prouve que $]x_0 - \varepsilon, x_0 + \varepsilon[$ est inclus dans $\overset{\circ}{F}_{n+1}$, puis que x_0 est inclus dans Ω . On obtient ainsi une contradiction, donc F n'admet pas de point isolé.

Pour finir, montrons que l'ensemble F est forcément vide. Supposons le contraire. Par hypothèse, l'union des ensembles F_n est égale à $\mathbb{R}$ tout entier, donc l'union des ensembles $F_n \cap F$ est égale à F . Pour tout $n \in \mathbb{N}$, l'ensemble $F_n \cap F$ est fermé dans F , qui est complet (car fermé dans $\mathbb{R}$ complet). D'après le théorème de Baire (en passant au complémentaire), il existe alors un entier $n_0 \in \mathbb{N}$ tel que $F \cap F_{n_0}$ est d'intérieur non vide (pour la topologie induite sur F). Ainsi, il existe deux réels a et b tels que $\emptyset \neq]a, b[\cap F \subset F_{n_0}$.

- On considère un point x dans $]a, b[\cap F$. Le point x ne peut être isolé, donc il existe une suite $(x_p)_{p \geq 1}$ de points distincts dans $]a, b[\cap F$ qui converge vers x . Pour tout p , on a en particulier $x_p \in F_{n_0}$, donc $\varphi^{(n_0)}(x_p) = 0$. Quitte à extraire une sous-suite, on peut supposer que la suite $(x_p)_{p \geq 1}$ est strictement monotone, par exemple strictement croissante. Pour tout p , on a $\varphi^{(n_0)}(x_p) = \varphi^{(n_0)}(x_{p+1}) = 0$. D'après le théorème de Rolle, il existe $y_p \in]x_p, x_{p+1}[$ tel que $\varphi^{(n_0+1)}(y_p) = 0$. On obtient alors une suite $(y_p)_{p \geq 1}$ strictement croissante qui converge vers x et qui annule la fonction $\varphi^{(n_0+1)}$. Par continuité, on en déduit que $\varphi^{(n_0+1)}(x) = 0$. En itérant ce processus, on obtient finalement que $\varphi^{(n)}(x) = 0$ pour tout $n \geq n_0$.
- On considère maintenant un point y dans $]a, b[\cap \Omega$. Comme $]a, b[\cap F$ est non vide, la composante connexe Ω_y de Ω qui contient y possède au moins une extrémité dans $]a, b[$, notée x . Ce point x appartient nécessairement à $]a, b[\cap F$, et le raisonnement précédent montre que $\varphi^{(n)}(x) = 0$ pour tout $n \geq n_0$. De plus, la fonction φ coïncide avec une fonction polynomiale P sur Ω_y , et on a donc par continuité des dérivées successives, $P^{(n)}(x) = \varphi^{(n)}(x) = 0$ pour tout $n \geq n_0$. La formule de Taylor pour P donne donc $\varphi^{(n_0)}(y) = P^{(n_0)}(y) = 0$.

Il en ressort que $\varphi^{(n_0)}$ s'annule en tout point de l'intervalle $]a, b[$, que ce point appartienne à F ou à Ω . Cela montre que $]a, b[\subset \overset{\circ}{F}_{n_0}$, ce qui contredit le fait que $]a, b[\cap F$ est non vide. On en conclut que l'ensemble F est vide, ce qui équivaut au fait que Ω est égal à $\mathbb{R}$ tout entier. La fonction φ est donc une fonction polynomiale sur $\mathbb{R}$. $\square$

6.3.5 Séries de Fourier

On note $\mathcal{C}$ l'espace vectoriel des fonctions 2π -périodiques et continues de $\mathbb{R}$ dans $\mathbb{C}$, muni de la norme $\|f\|_\infty = \sup_{t \in]-\pi, \pi]} |f(t)|$. C'est un espace de Banach, car fermé dans l'espace vectoriel complet des fonctions 2π -périodiques et bornées. Les coefficients de Fourier de $f \in \mathcal{C}$ sont donnés par

$$\forall p \in \mathbb{Z} \quad c_p(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ipt} dt.$$

De plus, pour tout entier $n \geq 1$, on considère l'application ℓ_n définie de $\mathcal{C}$ dans $\mathbb{C}$ par $\ell_n(f) = \sum_{p=-n}^n c_p(f)$.

Proposition 6.15. *Pour tout entier $n \geq 1$, l'application ℓ_n est une forme linéaire continue sur $\mathcal{C}$ qui vérifie*

$$\|\ell_n\| = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{\sin((2n+1)t/2)}{\sin(t/2)} \right| dt.$$

Par conséquent, $\|\ell_n\| \xrightarrow{n \rightarrow \infty} +\infty$ et il existe une fonction $f \in \mathcal{C}$ telle que la série de Fourier de f diverge en 0.

Démonstration. Il est clair que chaque application ℓ_n est continue. De plus, pour $f \in \mathcal{C}$, on a

$$\ell_n(f) = \sum_{p=-n}^n c_p(f) = \sum_{p=-n}^n \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ipt} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \left(\sum_{p=-n}^n e^{-ipt} \right) dt.$$

Cependant, pour $t \notin 2\pi\mathbb{Z}$, on rappelle que

$$\sum_{p=-n}^n e^{-ipt} = \sum_{p=-n}^n e^{ipt} = e^{-int} \sum_{p=0}^{2n} e^{ipt} = e^{-int} \frac{e^{i(2n+1)t} - 1}{e^{it} - 1} = \frac{\sin((2n+1)t/2)}{\sin(t/2)} = u_n(t).$$

On prolonge ce quotient par continuité en posant $u_n(2k\pi) = 2n+1$ pour tout $k \in \mathbb{Z}$; il coïncide alors sur $\mathbb{R}$ avec la somme trigonométrique précédente. Ainsi, $\ell_n(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) u_n(t) dt$, puis $|\ell_n(f)| \leq \frac{\|f\|_\infty}{2\pi} \int_{-\pi}^{\pi} |u_n(t)| dt$. De plus, pour tout $\varepsilon > 0$, on considère la fonction $f_\varepsilon(t) = \frac{u_n(t)}{|u_n(t)| + \varepsilon}$. On a clairement $f_\varepsilon \in \mathcal{C}$ et $\|f_\varepsilon\|_\infty \leq 1$. De plus, par convergence dominée, on a $\ell_n(f_\varepsilon) \xrightarrow{\varepsilon \rightarrow 0} \frac{1}{2\pi} \int_{-\pi}^{\pi} |u_n(t)| dt$. On en déduit que $\|\ell_n\| = \frac{1}{2\pi} \int_{-\pi}^{\pi} |u_n(t)| dt$.

Ensuite, comme $|\sin(t/2)| \leq |t/2|$ pour tout réel t , on a

$$0 < |t| \leq \pi \implies |u_n(t)| = \left| \frac{\sin((2n+1)t/2)}{\sin(t/2)} \right| \geq \frac{2}{|t|} \left| \sin\left((2n+1)\frac{t}{2}\right) \right|,$$

de sorte que

$$\begin{aligned} \|\ell_n\| &\geq \frac{1}{\pi} \int_{-\pi}^{\pi} \left| \sin\left((2n+1)\frac{t}{2}\right) \right| \frac{dt}{|t|} = \frac{2}{\pi} \int_0^{\pi} \left| \sin\left((2n+1)\frac{t}{2}\right) \right| \frac{dt}{t} \\ &= \frac{2}{\pi} \int_0^{(2n+1)\pi/2} |\sin x| \frac{dx}{x} \xrightarrow{n \rightarrow \infty} +\infty. \end{aligned}$$

Le théorème de Banach-Steinhaus, appliqué à la famille $H = \{\ell_n, n \geq 1\}$ de l'espace $\mathcal{L}(\mathcal{C}, \mathbb{C})$, implique l'existence d'une fonction $f \in \mathcal{C}$ telle que $\sup_{n \geq 1} |\ell_n(f)| = \infty$. Ainsi, il existe une fonction, en l'occurrence f , dont la série de Fourier diverge en 0. (*Remarque* : On connaît des exemples explicites de telles fonctions, comme l'exemple célèbre dû à Du Bois-Reymond.) $\square$

Théorèmes de Hahn–Banach et applications

Références : [8, 13, 15].

7.1 Forme analytique

On ne considère que des $\mathbb{R}$ -espaces vectoriels. Soit E un tel espace.

Théorème 7.1 (Hahn-Banach, forme analytique). Soit $p : E \rightarrow \mathbb{R}$ une application telle que :

- (i). pour tous $x \in E$ et $\lambda > 0$, on a $p(\lambda x) = \lambda p(x)$;
- (ii). pour tous $x, y \in E$, on a $p(x + y) \leq p(x) + p(y)$.

Soit G un sous-espace vectoriel de E et soit $g : G \rightarrow \mathbb{R}$ une application linéaire telle que $g(x) \leq p(x)$ pour tout $x \in G$. Alors, il existe une forme linéaire f sur E qui prolonge g et qui est dominée par p , c'est-à-dire que :

- (i). pour tout $x \in G$, on a $f(x) = g(x)$;
- (ii). pour tout $x \in E$, on a $f(x) \leq p(x)$.

La preuve fait appel au lemme de Zorn (admis), et nécessite de rappeler quelques notions de théorie des ensembles. Soit P un ensemble muni d'une relation d'ordre (partiel) notée $\leq$. On dit qu'un sous-ensemble $Q \subset P$ est *totale-ment ordonné* si pour tous $a, b \in Q$, on a $a \leq b$ ou $b \leq a$. On dit que $c \in P$ est un *majorant* de Q si $a \leq c$ pour tout $a \in Q$. On dit que $m \in P$ est un *élément maximal* de P si pour tout $x \in P$ tel que $m \leq x$, on a $x = m$. On dit que P est *inductif* si tout sous-ensemble totalement ordonné de P admet un majorant.

Lemme 7.2 (Zorn). Tout sous-ensemble ordonné, inductif et non vide admet un élément maximal.

Prouvons maintenant le théorème. On considère l'ensemble P constitué des formes linéaires h définies sur un sous-espace vectoriel de E noté $D(h)$ qui contient G , qui prolongent g et sont dominées par p sur $D(h)$. On munit P de la relation d'ordre telle que $h_1 \leq h_2$ si et seulement si $D(h_1) \subset D(h_2)$ et h_2 prolonge h_1 . L'ensemble P est non vide car il contient g . De plus, P est inductif. En effet, la chaîne vide est majorée par g . Soit Q un sous-ensemble totalement ordonné non vide de P . On considère le sous-espace vectoriel $H = \bigcup_{q \in Q} D(q)$ et on y

définit une application linéaire $h : H \rightarrow \mathbb{R}$ par $h(x) = q(x)$ si $x \in D(q)$. Il n'y a pas unicité de q , mais la définition a tout de même un sens car Q est totalement ordonné. De plus, on a bien $h \in P$ et h est un majorant de Q . Le lemme de Zorn assure que P admet un élément maximal noté f .

Pour prouver le théorème, il suffit de montrer que l'application f est définie sur E tout entier. Supposons au contraire que $D(f) \neq E$, et notons $x_0 \notin D(f)$. On définit sur l'espace vectoriel $H = D(f) + \mathbb{R}x_0$ l'application linéaire $h(x + t x_0) = f(x) + t \alpha$, pour $x \in D(f)$ et $t \in \mathbb{R}$, avec α fixé ultérieurement de sorte que $h \in P$. On veut alors que $f(x) + t \alpha \leq p(x + t x_0)$ pour tous $x \in D(f)$ et $t \in \mathbb{R}$. Il suffit de vérifier cette relation pour $t = \pm 1$. (En effet, on a $f(x) + t \alpha = t(f(x/t) + \alpha) \leq t p(x/t + x_0) = p(x + t x_0)$ si $t > 0$ et si la relation est vraie pour $t = 1$; et de même, on a $f(x) + t \alpha = -t(f(x/(-t)) - \alpha) \leq -t p(x/(-t) - x_0) = p(x + t x_0)$ si $t < 0$ et si la relation est vraie pour $t = -1$.) Cela revient à choisir α tel que

$$\sup_{y \in D(f)} (f(y) - p(y - x_0)) \leq \alpha \leq \inf_{x \in D(f)} (p(x + x_0) - f(x)).$$

Un tel choix est possible car pour tous $x, y \in D(f)$, on a

$$f(x) + f(y) = f(x + y) \leq p(x + y) = p((x + x_0) + (y - x_0)) \leq p(x + x_0) + p(y - x_0).$$

Finalement, on a construit $h \in P$ telle que $h \neq f$ et $f \leq h$, ce qui contredit la maximalité de f . On en conclut que $D(f) = E$, ce qui prouve le théorème.

On munit désormais E d'une norme $\|\cdot\|$. On note E' le dual topologique de E , c'est-à-dire l'espace vectoriel des formes linéaires continues sur E , muni de la norme $\|f\|_{E'} = \sup_{\substack{x \in E \\ \|x\| \leq 1}} |f(x)| = \sup_{\substack{x \in E \\ \|x\| \leq 1}} f(x)$.

Corollaire 7.3. Soit G un sous-espace vectoriel de E et soit $g : G \rightarrow \mathbb{R}$ une application linéaire continue de norme $\|g\|_{G'} = \sup_{\substack{x \in G \\ \|x\| \leq 1}} g(x)$. Alors, il existe $f \in E'$ qui prolonge g et telle que $\|f\|_{E'} = \|g\|_{G'}$.

Démonstration. Il suffit d'appliquer le théorème de Hahn-Banach avec pour application dominante $p(x) = \|g\|_{G'} \|x\|$. $\square$

Corollaire 7.4. Pour tout $x_0 \in E$, il existe $f_0 \in E'$ tel que $\|f_0\|_{E'} = \|x_0\|$ et $f_0(x_0) = \|x_0\|^2$.

Démonstration. On peut supposer $x_0 \neq 0$. On applique le corollaire précédent avec $G = \mathbb{R}x_0$ et $g \in G'$ définie par $g(tx_0) = t\|x_0\|^2$ pour tout $t \in \mathbb{R}$, ce qui donne bien $\|g\|_{G'} = \left|g\left(\frac{x_0}{\|x_0\|}\right)\right| = \frac{1}{\|x_0\|}\|x_0\|^2 = \|x_0\|$. $\square$

Corollaire 7.5. Pour tout $x \in E$, on a $\|x\| = \sup_{\substack{f \in E' \\ \|f\| \leq 1}} |f(x)| = \max_{\substack{f \in E' \\ \|f\| \leq 1}} |f(x)|$.

Démonstration. On peut supposer $x \neq 0$. Clairement, $\sup_{\substack{f \in E' \\ \|f\| \leq 1}} |f(x)| \leq \|x\|$. Réciproquement, le corollaire précédent assure que l'on peut trouver $f_0 \in E'$ vérifiant $\|f_0\|_{E'} = \|x\|$ et $f_0(x) = \|x\|^2$. On considère $f_1 = \frac{f_0}{\|x\|}$, et on a bien $\|f_1\|_{E'} = 1$ et $f_1(x) = \frac{f_0(x)}{\|x\|} = \|x\|$. $\square$

7.2 Formes géométriques

Commençons par quelques rappels sur les hyperplans. On désigne encore par E un espace vectoriel normé.

Définition 7.1. Un *hyperplan* (affine) est un ensemble de la forme

$$H = \{x \in E \mid f(x) = \alpha\},$$

où f est une forme linéaire non identiquement nulle sur E , et $\alpha \in \mathbb{R}$. On dit que H est l'hyperplan d'équation $f = \alpha$.

Proposition 7.6. Un hyperplan d'équation $f = \alpha$ est fermé si et seulement si la forme linéaire f est continue.

Démonstration. Si f est continue, alors l'hyperplan H d'équation $f = \alpha$ est clairement fermé. Réciproquement, supposons que H est fermé. Son complémentaire H^c est alors ouvert et non vide (puisque f est non identiquement nulle). Considérons $x_0 \in H^c$ et supposons par exemple que $f(x_0) < \alpha$. Notons $r > 0$ tel que la boule ouverte $B(x_0, r)$ soit incluse dans H^c . On a $f(x) < \alpha$ pour tout $x \in B(x_0, r)$. En effet, supposons que $f(x_1) \geq \alpha$ pour un certain $x_1 \in B(x_0, r)$. Le segment formé des points $x_t = (1-t)x_0 + tx_1$, pour $t \in [0, 1]$, est contenu dans la boule $B(x_0, r)$, ce qui donne $f(x_t) \neq \alpha$ pour tout $t \in [0, 1]$. Cependant, on a $f(x_t) = \alpha$ si et seulement si $t = \frac{\alpha - f(x_0)}{f(x_1) - f(x_0)}$, qui appartient à $]0, 1[$, ce qui conduit à une contradiction. On en déduit que $f(x_0 + rz) < \alpha$ pour tout $z \in B(0, 1)$. Cela implique que f est continue et que $\|f\|_{E'} \leq \frac{1}{r}(\alpha - f(x_0))$. $\square$

Définition 7.2. Soit $A, B \subset E$. On dit que l'hyperplan H d'équation $f = \alpha$ sépare A et B au sens large si

$$\begin{cases} \forall x \in A & f(x) \leq \alpha \\ \forall x \in B & f(x) \geq \alpha. \end{cases}$$

On dit que H sépare A et B au sens strict s'il existe $\varepsilon > 0$ tel que

$$\begin{cases} \forall x \in A & f(x) \leq \alpha - \varepsilon \\ \forall x \in B & f(x) \geq \alpha + \varepsilon. \end{cases}$$

Géométriquement, cela traduit le fait que A et B sont situés de part et d'autre de H . On rappelle aussi qu'un ensemble $A \subset E$ est convexe si pour tous $x, y \in A$ et tout $t \in [0, 1]$, on a $tx + (1 - t)y \in A$.

Théorème 7.7 (Hahn-Banach, première forme géométrique). Soit $A, B \subset E$ deux ensembles convexes, non vides et disjoints. On suppose que A est ouvert. Alors, il existe un hyperplan fermé qui sépare A et B au sens large.

La démonstration repose sur deux lemmes.

Lemme 7.8 (jauge d'un convexe). On note $C \subset E$ un convexe ouvert qui contient 0. Pour tout $x \in E$, on pose $p(x) = \inf\{\alpha > 0 \mid x/\alpha \in C\}$; il s'agit de la jauge du convexe C . Alors :

- (i). pour tous $x \in E$ et $\lambda > 0$, on a $p(\lambda x) = \lambda p(x)$;
- (ii). pour tous $x, y \in E$, on a $p(x + y) \leq p(x) + p(y)$;
- (iii). il existe $M > 0$ tel que pour tout $x \in E$, on a $0 \leq p(x) \leq M\|x\|$;
- (iv). $C = \{x \in E \mid p(x) < 1\}$.

Démonstration. Soit $r > 0$ tel que $B(0, r) \subset C$. On a $p(x) \leq \frac{\|x\|}{r}$ pour tout $x \in E$, ce qui donne directement le troisième point. De plus, le premier point découle directement de la définition de p .

Pour prouver le quatrième point, considérons $x \in C$. Comme C est ouvert, on a $(1 + \varepsilon)x \in C$ pour $\varepsilon > 0$ assez petit. Donc, $p(x) \leq \frac{1}{1 + \varepsilon} < 1$. Réciproquement, si $p(x) < 1$, il existe $\alpha \in]0, 1[$ tel que $\frac{x}{\alpha} \in C$, et par convexité, on a $x = \alpha \frac{x}{\alpha} + (1 - \alpha)0 \in C$.

Finalement, pour établir le deuxième point, considérons $x, y \in E$ et $\varepsilon > 0$. On a $p\left(\frac{x}{p(x) + \varepsilon}\right) = \frac{p(x)}{p(x) + \varepsilon} < 1$, donc $\frac{x}{p(x) + \varepsilon} \in C$. C'est également valable pour y , donc par convexité, on a $\frac{tx}{p(x) + \varepsilon} + \frac{(1 - t)y}{p(y) + \varepsilon} \in C$ pour tout $t \in [0, 1]$. Pour $t = \frac{p(x) + \varepsilon}{p(x) + p(y) + 2\varepsilon}$, on obtient $\frac{x + y}{p(x) + p(y) + 2\varepsilon} \in C$. On en déduit que $p\left(\frac{x + y}{p(x) + p(y) + 2\varepsilon}\right) < 1$, puis $p(x + y) < p(x) + p(y) + 2\varepsilon$. On conclut en faisant tendre ε vers 0. $\square$

Lemme 7.9. Soit $C \subset E$ un convexe ouvert non vide et soit $x_0 \in E \setminus C$. Alors, il existe $f \in E'$ vérifiant $f(x) < f(x_0)$ pour tout $x \in C$. En particulier, l'hyperplan d'équation $f = f(x_0)$ sépare $\{x_0\}$ et C au sens large.

Démonstration. Par translation, on peut supposer que C contient 0; on note p la jauge de ce convexe. On considère sur $G = \mathbb{R}x_0$ la forme linéaire g définie par $g(tx_0) = t$ pour tout $t \in \mathbb{R}$. On a $g(x) \leq p(x)$ pour tout $x \in G$. (En effet, $\frac{x}{g(x)} = x_0 \notin C$ si $g(x) > 0$, et l'inégalité est triviale si $g(x) \leq 0$.) D'après la forme analytique du théorème de Hahn-Banach, il existe une forme linéaire sur E qui prolonge g et qui est dominée par p . En particulier, on a $f(x_0) = g(x_0) = 1$ et $f(x) \leq p(x) \leq M\|x\|$, avec les notations du troisième point du lemme précédent, ce qui garantit la continuité de f . De plus, pour $x \in C$, on a $f(x) \leq p(x) < 1$. Finalement, l'hyperplan d'équation $f = f(x_0)$ est fermé et sépare $\{x_0\}$ et C au sens large. $\square$

Démontrons maintenant le théorème. On pose $C = A - B$. Il s'agit d'un convexe ouvert (car union des ouverts $A - y$ pour $y \in B$) qui ne contient pas 0 (car $A \cap B = \emptyset$). D'après le second lemme, il existe une forme linéaire $f \in E'$ qui telle que $f(z) < 0$ pour tout $z \in C$. On a donc $f(x - y) < 0$ pour tous $x \in A$ et $y \in B$, ce qui donne $\sup_{x \in A} f(x) \leq \inf_{y \in B} f(y)$. Il suffit de choisir α entre ces deux valeurs pour conclure que l'hyperplan fermé d'équation $f = \alpha$ sépare A et B au sens large.

Théorème 7.10 (Hahn-Banach, deuxième forme géométrique). *Soit deux ensembles convexes $A, B \subset E$, non vides et disjoints. On suppose que A est fermé et que B est compact. Alors, il existe un hyperplan fermé qui sépare A et B au sens strict.*

Démonstration. Pour tout $\varepsilon > 0$, considérons les ensembles $A_\varepsilon = A + B(0, \varepsilon)$ et $B_\varepsilon = B + B(0, \varepsilon)$. Ce sont des convexes ouverts et non vides. De plus, ils sont disjoints pour $\varepsilon > 0$ assez petit. (En effet, sinon on aurait deux suites $(x_n)_{n \geq 1}$ et $(y_n)_{n \geq 1}$ respectivement dans A et B telles que $\|x_n - y_n\| < \frac{2}{n}$ pour tout $n \geq 1$. Par compacité de B , on peut extraire une sous-suite $(y_{\varphi(n)})_{n \geq 1}$ qui converge vers $y \in B$. On a alors $x_{\varphi(n)} \xrightarrow[n \rightarrow \infty]{} y$, et donc $y \in A$ puisque A est fermé. Cela contredit le fait que A et B sont disjoints.) La première forme géométrique du théorème de Hahn-Banach montre alors l'existence d'un hyperplan fermé d'équation $f = \alpha$ qui sépare les ensembles A_ε et B_ε au sens large. On a donc $f(x + \varepsilon z) \leq \alpha \leq f(y + \varepsilon z)$ pour tous $x \in A$, $y \in B$ et $z \in B(0, 1)$. On en déduit que $f(x) + \varepsilon f(z) \leq \alpha$ pour tout $z \in B(0, 1)$, et donc que $\|f\|_{E'} \leq \frac{1}{\varepsilon}(\alpha - f(x))$. De façon similaire, on a $\alpha \leq f(y) - \varepsilon f(z)$ pour tout $z \in B(0, 1)$. Finalement, on a

$$\begin{cases} \forall x \in A & f(x) \leq \alpha - \varepsilon \|f\|_{E'} \\ \forall y \in B & f(y) \geq \alpha + \varepsilon \|f\|_{E'}, \end{cases}$$

de sorte que l'hyperplan fermé d'équation $f = \alpha$ sépare les ensembles A et B au sens strict. (On a bien $\|f\|_{E'} > 0$.) $\square$

Corollaire 7.11. *Soit $F \subset E$ un sous-espace vectoriel de E tel que $\overline{F} \neq E$. Alors, il existe une forme linéaire $f \in E'$ non identiquement nulle telle que $f(x) = 0$ pour tout $x \in F$.*

Démonstration. On considère $x_0 \in E \setminus \overline{F}$ et on applique la deuxième forme géométrique du théorème de Hahn-Banach avec $A = \overline{F}$ et $B = \{x_0\}$. Il existe alors $f \in E'$ non identiquement nulle telle que l'hyperplan d'équation $f = \alpha$ sépare A et B au sens strict. On a donc $f(x) < \alpha < f(x_0)$ pour tout $x \in \overline{F}$. En particulier, pour $x \in F$ et $\lambda \in \mathbb{R}$, on a $f(\lambda x) < \alpha$, ce qui impose $f(x) = 0$. $\square$

Remarque. Pour prouver qu'un sous-espace vectoriel F de E est dense, il suffit donc de montrer que toute forme linéaire continue nulle sur F est identiquement nulle.

7.3 Applications

7.3.1 Théorème de Krein-Milman

Soit $A \subset E$. On rappelle que l'enveloppe convexe fermée de A , notée $\overline{\text{conv}} A$, est le plus petit ensemble convexe fermé contenant A . De plus, si $K \subset E$ est convexe, on dit qu'un point $x \in K$ est *extrémal* si pour tous $x_0, x_1 \in K$ et $t \in]0, 1[$ vérifiant $x = (1 - t)x_0 + tx_1$, on a $x_0 = x_1 = x$.

Théorème 7.12 (Krein-Milman). *Soit $K \subset E$ un ensemble convexe compact. Alors, K coïncide avec l'enveloppe convexe fermée de ses points extrémaux.*

Démonstration. Si K est vide, le résultat est immédiat. Supposons donc $K \neq \emptyset$. On dit qu'un ensemble $S \subset K$ est *extrémal* si S est compact et non vide, et si pour tous $x_0, x_1 \in K$ et $t \in]0, 1[$ vérifiant $(1 - t)x_0 + tx_1 \in S$, on a $x_0, x_1 \in S$. Ainsi, un point extrémal est un ensemble extrémal réduit à un singleton. On note $\mathcal{P}$ la collection des ensembles extrémaux de K . On a clairement $K \in \mathcal{P}$. De plus, pour toute collection non vide $\mathcal{Q} \subset \mathcal{P}$, l'ensemble

$\bigcap_{T \in \mathcal{Q}} T$ est soit vide, soit extrémal. En effet, supposons que cet ensemble est non vide et considérons $x_0, x_1 \in K$ et $t \in]0, 1[$ tels que $(1-t)x_0 + tx_1 \in \bigcap_{T \in \mathcal{Q}} T$. On a alors $(1-t)x_0 + tx_1 \in T$ pour tout $T \in \mathcal{Q}$, puis $x_0, x_1 \in T$, puisque T est extrémal. Ainsi, $x_0, x_1 \in \bigcap_{T \in \mathcal{Q}} T$, de sorte que ce dernier ensemble est extrémal.

Par ailleurs, si $T \in \mathcal{P}$ et $f \in E'$, alors l'ensemble T_f formé des points x de T en lesquels f est maximale, c'est-à-dire qui vérifient $f(x) = \sup_{y \in T} f(y)$, est un ensemble extrémal. En effet, si $x_0, x_1 \in K$ et $t \in]0, 1[$ vérifient $(1-t)x_0 + tx_1 \in T_f$, alors ils vérifient aussi $(1-t)x_0 + tx_1 \in T$, de sorte que $x_0, x_1 \in T$, puisque T est extrémal. Ainsi, $f(x_0) \leq \mu$ et $f(x_1) \leq \mu$, où μ désigne le supremum de f sur T . Cependant, on a $\mu = f((1-t)x_0 + tx_1) = (1-t)f(x_0) + tf(x_1)$, ce qui impose $f(x_0) = f(x_1) = \mu$, donc $x_0, x_1 \in T_f$.

Considérons maintenant un ensemble extrémal S , ainsi que la collection $\mathcal{P}_S = \{T \in \mathcal{P} \mid T \subset S\}$, munie de l'ordre donné par l'inclusion. On note Ω un sous-ensemble maximal de $\mathcal{P}_S$ qui est totalement ordonné. (Un tel ensemble existe bien car $\mathcal{P}_S$ est non vide, et en vertu du lemme de Zorn appliqué aux parties de $\mathcal{P}_S$.) On considère ensuite $M = \bigcap_{T \in \Omega} T$. Alors, $M \subset K$ et M est non vide. (Sinon, on aurait $K = K \setminus M = \bigcup_{T \in \Omega} K \setminus T \subset \bigcup_{T \in \Omega} E \setminus T$ et, par compacité de K , on pourrait extraire un recouvrement fini $K \subset \bigcup_{i=1}^n E \setminus T_i$ avec $T_1, \dots, T_n \in \Omega$, ce qui donnerait

$K = K \setminus \bigcap_{i=1}^n T_i$, et donc $\bigcap_{i=1}^n T_i = \emptyset$, ce qui est impossible puisque les T_i sont non vides et Ω est totalement ordonné.)

Comme M est non vide, il est extrémal. De plus, comme $M = \bigcap_{T \in \Omega} T$, avec $T \subset S$ pour tout $T \in \Omega$, on a finalement $M \in \mathcal{P}_S$. Par maximalité de Ω , aucun sous-ensemble strict de M ne peut appartenir à $\mathcal{P}_S$. En effet, si U était un tel ensemble, alors $\Omega \cup \{U\}$ serait totalement ordonné puisque $U \subset M \subset S$. En particulier, si $f \in E'$, l'ensemble M_f des points de M en lesquels f est maximale, qui est un sous-ensemble extrémal de M , ne peut être strictement inclus dans M . Ainsi, f est constante sur M . On en déduit que M est réduit à un singleton. (En effet, si M contenait deux points distincts, on pourrait trouver une forme linéaire non constante sur M d'après la deuxième forme géométrique du théorème de Hahn-Banach.) Autrement dit, M est un point extrémal. Finalement, tout ensemble extrémal de K contient un point extrémal.

Pour conclure, on observe d'une part que K est un convexe compact, donc contient l'enveloppe convexe fermée de ses points extrémaux, notée C . Réciproquement, supposons l'existence d'un point $x_0 \in K \setminus C$. Alors, la deuxième forme géométrique du théorème de Hahn-Banach montre l'existence d'une forme linéaire $f \in E'$ telle que l'hyperplan d'équation $f = \alpha$ sépare C et $\{x_0\}$ au sens strict. Alors, on a $f(x) < \alpha < f(x_0)$ pour tout $x \in C$. L'ensemble K_f des points de K en lesquels f est maximale est un ensemble extrémal. Cependant, il ne peut contenir de point extrémal puisque $C \cap K_f = \emptyset$. On obtient une contradiction, ce qui donne enfin $C = K$. $\square$

On peut illustrer ce résultat avec la boule unité fermée de $\mathbb{R}^n$ pour la norme $\|\cdot\|_\infty$, c'est-à-dire l'ensemble B des n -uplets $(x_1, \dots, x_n) \in \mathbb{R}^n$ tels que $|x_i| \leq 1$ pour tout i . Alors, on a $B = \overline{\text{conv}} A$, pour $A = \{-1, 1\}^n$. En effet, comme B est un convexe fermé qui contient A , on a clairement l'inclusion $B \supset \overline{\text{conv}} A$. Réciproquement, considérons un point $x \in B$. On observe que l'ensemble C des combinaisons convexes de points de A est un compact. Si $x \notin C$, la deuxième forme géométrique du théorème de Hahn-Banach montre l'existence d'une forme linéaire f telle qu'un hyperplan d'équation $f = \alpha$ sépare les ensembles C et $\{x\}$ au sens strict. En particulier, on a $f(x) > \alpha > \sup_{y \in C} f(y) \geq \sup_{\varepsilon \in A} f(\varepsilon)$. Ainsi, il existe $(f_1, \dots, f_n) \in \mathbb{R}^n$ tel que, en écrivant $x = (x_1, \dots, x_n)$ et $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$, on ait $\sum_{i=1}^n f_i x_i > \sup_{\varepsilon \in A} \sum_{i=1}^n f_i \varepsilon_i$. Cependant, ce supremum est clairement atteint quand $\varepsilon_i = 1_{\{f_i \geq 0\}} - 1_{\{f_i < 0\}}$, et il vaut alors $\sum_{i=1}^n |f_i|$. On a donc $\sum_{i=1}^n f_i x_i > \sum_{i=1}^n |f_i|$, ce qui est impossible puisque $x \in B$, ce qui impose au contraire $\sum_{i=1}^n f_i x_i \leq \sum_{i=1}^n |f_i|$. On a donc $x \in C$, c'est-à-dire que x est une combinaison convexe d'éléments de A , et appartient donc à $\overline{\text{conv}} A$. Finalement, on a l'inclusion réciproque $B \subset \overline{\text{conv}} A$.

7.3.2 Non-réflexivité de ℓ^1

On considère l'espace $\ell^1 = \ell^1(\mathbb{N}^*)$ des suites réelles $u = (u_n)_{n \geq 1}$ pour lesquelles $\|u\|_1 = \sum_{n=1}^{\infty} |u_n| < \infty$. Il s'agit d'un espace de Banach séparable dont le dual est $\ell^\infty = \ell^\infty(\mathbb{N}^*)$, l'espace des suites réelles $v = (v_n)_{n \geq 1}$ telles que $\|v\|_\infty = \sup_{n \geq 1} |v_n| < \infty$. En effet, si $v \in \ell^\infty$, alors l'application qui à toute suite $u \in \ell^1$ associe la somme $\sum_{n=1}^{\infty} u_n v_n$ définit clairement une forme linéaire sur ℓ^1 . Réciproquement, si φ est une forme linéaire continue sur ℓ^1 , alors on peut la représenter à l'aide de la suite de terme général $v_n = \varphi(e_n)$, où e_n désigne la suite dont tous les termes sont nuls sauf le n -ième qui vaut 1, puisqu'on a alors $|v_n| = |\varphi(e_n)| \leq \|\varphi\|_{(\ell^1)'} \|e_n\|_1 = \|\varphi\|_{(\ell^1)'}$.

Un espace de Banach E est dit *réflexif* si l'application canonique $J : E \rightarrow E'' = (E')'$, définie par $J(x)(\varphi) = \varphi(x)$, est surjective. Ici, ℓ^1 n'est pas réflexif : il existe une forme linéaire continue f sur ℓ^∞ qui n'est pas de la forme $v \mapsto \sum_{n \geq 1} u_n v_n$ avec $u \in \ell^1$. En effet, on définit sur le sous-espace vectoriel formé des suites convergentes de ℓ^∞ une forme linéaire continue g en posant $g(v) = \lim_{n \rightarrow \infty} v_n$. D'après le premier corollaire de la forme analytique du théorème de Hahn-Banach, g se prolonge en une forme linéaire continue sur ℓ^∞ , notée f . Supposons que f provient de ℓ^1 , c'est-à-dire qu'il existe une suite $u \in \ell^1$ telle que $f(v) = \sum_{n=1}^{\infty} u_n v_n$ pour toute suite $v \in \ell^\infty$. On a en particulier $u_n = f(e_n) = g(e_n) = 0$ pour tout $n \geq 1$. Ainsi, la suite u est nulle, de même que les applications f et g , ce qui contredit le fait, par exemple, que g attribue la valeur 1 à la suite constante égale à 1.

7.3.3 Fonctions périodiques

Dans cette application, les espaces sont complexes. La propriété de prolongement utilisée ci-dessous découle du cas réel : on prolonge $\Re \varphi$ en une forme réelle u de même norme, puis on pose $\ell(x) = u(x) - iu(ix)$. Cette forme est complexe linéaire et prolonge φ . En choisissant $|\lambda| = 1$ tel que $\lambda \ell(x) = |\ell(x)|$, on obtient $|\ell(x)| = u(\lambda x) \leq \|u\| \|x\|$, ce qui assure la continuité et la conservation de la norme. Le critère de densité s'étend donc également aux espaces complexes.

On note $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$ le cercle, obtenu en identifiant deux réels dont la différence est un multiple entier de 2π ; les fonctions sur $\mathbb{T}$ s'identifient aux fonctions 2π -périodiques sur $\mathbb{R}$.

Soit $\mathcal{C}$ l'ensemble des fonctions continues et 2π -périodiques $f : \mathbb{R} \rightarrow \mathbb{C}$, muni de la norme $\|f\|_\infty = \sup_{t \in \mathbb{R}} |f(t)|$. Pour $n \in \mathbb{Z}$, on rappelle que le n -ième coefficient de Fourier d'une telle fonction f est donné par

$$c_n(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-int} dt.$$

On note $\text{Sp } f = \{n \in \mathbb{Z} \mid c_n(f) \neq 0\}$ le support en Fourier de f . De plus, pour $\Lambda \subset \mathbb{Z}$, on note $\mathcal{C}_\Lambda = \{f \in \mathcal{C} \mid \text{Sp } f \subset \Lambda\}$. Il s'agit d'un sous-espace vectoriel fermé de $\mathcal{C}$ puisqu'il peut s'écrire $\mathcal{C}_\Lambda = \bigcap_{n \notin \Lambda} \ker c_n$, c'est-à-dire sous la forme

d'une intersection d'hyperplans fermés. En outre, $\mathcal{C}_\Lambda$ est invariant par translation puisque $c_n(f(\cdot + \theta)) = e^{in\theta} c_n(f)$ pour tous $n \in \mathbb{Z}$ et $\theta \in \mathbb{R}$.

On considère une fonction $f \in \mathcal{C}$ et on note $\Lambda = \text{Sp } f$. On veut montrer que les translatées de f forment une partie totale de $\mathcal{C}_\Lambda$, c'est-à-dire que

$$\mathcal{C}_\Lambda = \overline{\text{Vect}\{f(\cdot + \theta), \theta \in \mathbb{R}\}}.$$

D'après la remarque qui suit l'énoncé de la deuxième forme géométrique du théorème de Hahn-Banach, il suffit de montrer que toute forme linéaire continue φ sur $\mathcal{C}_\Lambda$, nulle sur les fonctions de la forme $f(\cdot + \theta)$, pour $\theta \in \mathbb{R}$, est uniformément nulle. D'après le premier corollaire de la forme analytique du théorème de Hahn-Banach, φ se prolonge en une forme linéaire continue sur $\mathcal{C}$, notée ℓ . Le théorème de représentation de Riesz implique l'existence d'une mesure complexe μ sur $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$ telle que $\ell(g) = \int_0^{2\pi} g(-t) \mu(dt)$ pour toute fonction $g \in \mathcal{C}$. Pour tout $\theta \in \mathbb{R}$, on a alors

$$f * \mu(\theta) = \int_0^{2\pi} f(\theta - t) \mu(dt) = \ell(f(\cdot + \theta)) = \varphi(f(\cdot + \theta)) = 0.$$

Si on note $c_n(\mu) = \int_0^{2\pi} e^{-int} \mu(dt)$ les coefficients de Fourier de la mesure μ , on en déduit que $c_n(f)c_n(\mu) = c_n(f * \mu) = 0$ pour tout $n \in \mathbb{Z}$. Par conséquent, si $n \in \Lambda$, on a $c_n(f) \neq 0$, donc $c_n(\mu) = 0$, puis $\ell(e_n) = 0$, et enfin $\varphi(e_n) = 0$, où e_n désigne la fonction $t \mapsto e^{int}$. On peut en conclure que φ est nulle sur $\mathcal{C}_\Lambda$, puisque les fonctions e_n , pour $n \in \Lambda$, forment une famille totale de $\mathcal{C}_\Lambda$, d'après le théorème de Fejér. En effet, si $f \in \mathcal{C}_\Lambda$, ce théorème assure que $\sigma_N(f) = \sum_{n=-N}^N \left(1 - \frac{|n|}{N}\right) c_n(f) e_n \xrightarrow[N^\infty]{\|\cdot\|_\infty} f$, et on observe que $\sigma_N(f) \in \text{Vect}\{e_n, n \in \Lambda\}$ pour tout $N \geq 1$, car $c_n(\sigma_N(f)) = 0$ si $n \notin \Lambda$.

Deuxième partie

Exercices corrigés

Compacité et connexité

Références : [6, 7].

Exercice 8.1 (Distance entre deux parties)

Référence : [6].

1. Soit (E, d) un espace métrique.

(a). Soit K_1 et K_2 deux compacts non vides de E . Montrer l'existence de $x_1 \in K_1$ et $x_2 \in K_2$ tels que

$$d(x_1, x_2) = d(K_1, K_2) = \inf_{\substack{x \in K_1 \\ y \in K_2}} d(x, y).$$

(b). Soit K un compact non vide de E et F un fermé non vide de E . Si $K \cap F = \emptyset$, montrer que $d(K, F) \neq 0$. Ce résultat subsiste-t-il si K est seulement supposé fermé ?

2. Ici $E = \mathbb{R}^n$ muni de la métrique usuelle.

(a). Soit F un fermé non borné de E et $f : F \rightarrow \mathbb{R}$ une application continue telle que

$$\lim_{\substack{\|x\| \rightarrow \infty \\ x \in F}} f(x) = +\infty.$$

Montrer qu'il existe $x \in F$ tel que $f(x) = \inf_{y \in F} f(y)$.

(b). Soient K un compact non vide de E et F un fermé non vide de E . Montrer que

$$\exists x \in K \quad \exists y \in F \quad d(x, y) = d(K, F).$$

Ceci reste-t-il vrai si E est un $\mathbb{R}$ -espace vectoriel normé de dimension infinie ?

Corrigé.

1. (a). Soit (E, d) un espace métrique. Si K_1 et K_2 sont deux compacts non vides de E , alors il existe $x_1 \in K_1$ et $x_2 \in K_2$ tels que $d(x_1, x_2) = d(K_1, K_2) = \inf_{\substack{x \in K_1 \\ y \in K_2}} d(x, y)$. En effet, l'application $(x, y) \mapsto d(x, y)$ est continue sur le compact $K_1 \times K_2$, donc minorée et atteint son minimum.
- (b). Si F est un fermé de E , l'application $x \mapsto d(x, F)$ est continue (et même 1-lipschitzienne). Donc, si K désigne un compact de E , il existe $x_0 \in K$ tel que $d(x_0, F) = \inf_{x \in K} d(x, F) = d(K, F)$. Si $d(K, F) = 0$, alors $d(x_0, F) = 0$, donc $x_0 \in F$ car F est fermé. Ainsi, $K \cap F \neq \emptyset$ implique que $d(K, F) > 0$. Si K est seulement supposé fermé, c'est faux : par exemple, dans $\mathbb{R}^2$, les ensembles K et F de couples (x, y) définis par la condition $y \leq 0$ et $y \geq e^x$, respectivement, sont fermés, disjoints et vérifient $d(K, F) = 0$.
2. (a). On suppose $E = \mathbb{R}^n$. Soit F un fermé non borné de E et $f : F \rightarrow \mathbb{R}$ une application continue telle que $\lim_{\substack{\|x\| \rightarrow \infty \\ x \in F}} f(x) = +\infty$. Soit $a \in F$. L'ensemble Γ des $x \in F$ tels que $f(x) \leq f(a)$ est fermé puisque f est continue et F est fermé. De plus, Γ est borné, puisque f tend vers l'infini à l'infini. Ainsi, Γ est un compact non vide (il contient a), donc f y atteint son minimum en un point noté x , c'est-à-dire que $f(x) = \inf_{y \in \Gamma} f(y)$. Par construction de Γ , on a $\inf_{y \in \Gamma} f(y) = \inf_{y \in F} f(y)$, donc $f(x) = \inf_{y \in F} f(y)$, avec $x \in F$. Autrement dit, la fonction f est minorée sur F et atteint son minimum.
- (b). On suppose encore $E = \mathbb{R}^n$. Soit K un compact non vide de E et F un fermé non vide de E . Alors, il existe $x \in K$ et $y \in F$ tels que $d(x, y) = d(K, F)$. En effet, l'application $x \mapsto d(x, F)$ est continue sur le compact K , donc il existe $x \in K$ tel que $d(x, F) = \inf_{y \in K} d(y, F) = d(K, F)$. De même, l'application $y \mapsto d(x, y)$

est continue sur F . Si F est borné, il est compact, donc le résultat vient de ce qui précède. Sinon, on a $\lim_{\substack{\|y\| \rightarrow +\infty \\ y \in F}} d(x, y) = +\infty$, donc le résultat découle de ce qui est juste au-dessus.

C'est faux en dimension infinie. Par exemple, notons E l'espace vectoriel des suites bornées (u_n) de $\mathbb{R}$, muni de la norme ℓ^∞ . On note $u^{(n)}$ la suite dont tous les termes sont nuls, sauf le n -ième qui vaut $1 + 2^{-n}$. Alors, l'ensemble $F = \{u^{(n)}, n \in \mathbb{N}\}$ est fermé. En effet, si v désigne la limite d'une suite $(u^{(n_j)})_{j \geq 0}$ de suites dans F , alors $\|u^{(n_j)} - u^{(n_k)}\|_\infty \leq 1$ pour j et k assez grands. Cependant, $\|u^{(n_j)} - u^{(n_k)}\|_\infty = 1 + \max\{2^{-n_j}, 2^{-n_k}\} > 1$, sauf si $n_j = n_k$, auquel cas la norme est nulle. On en déduit que n_j est constante à partir d'un certain rang j_0 , ce qui donne $v = u^{(n_{j_0})} \in F$. De plus, si $K = \{0\}$, alors $d(K, F) = 1$, mais $d(u^{(n)}, K) = 1 + 2^{-n} > 1$ pour tout entier n .

Exercice 8.2 (Théorème de Riesz)

Référence : [6].

1. (Lemme de Riesz) Soit E un espace vectoriel normé et soit F un sous-espace vectoriel fermé de E tel que $F \neq E$. Montrer que pour tout $\varepsilon > 0$, il existe $u \in E$ tel que $\|u\| = 1$ et $d(u, F) \geq 1 - \varepsilon$.
2. (Théorème de Riesz) Soit E un espace vectoriel normé. Montrer que si la boule unité fermée de E est compacte, alors E est de dimension finie.

Corrigé.

1. On note E un espace vectoriel normé et F un sous-espace vectoriel fermé de E tel que $F \neq E$. On considère un point quelconque $v \in E \setminus F$. Comme F est fermé, $\{v\}$ est compact et $F \cap \{v\} = \emptyset$, un exercice précédent implique que $d = d(v, F) > 0$. Pour $\varepsilon \geq 1$, tout vecteur unitaire convient. Supposons donc $0 < \varepsilon < 1$. Il existe un vecteur $f_0 \in F$ tel que $d \leq \|v - f_0\| \leq \frac{d}{1 - \varepsilon}$. Alors, le vecteur défini par $u = \frac{v - f_0}{\|v - f_0\|}$ convient. En effet, ce vecteur est de norme 1 et pour tout $f \in F$, on a

$$\|u - f\| = \left\| \frac{v - f_0}{\|v - f_0\|} - f \right\| = \frac{\|v - \overbrace{(f_0 + \|v - f_0\| f)}^{\in F}\|}{\|v - f_0\|} \geq \frac{d}{d/(1 - \varepsilon)} = 1 - \varepsilon.$$

En prenant l'infimum sur tous les vecteurs $f \in F$, on en déduit que $d(u, F) \geq 1 - \varepsilon$.

2. Soit E un espace vectoriel normé. On suppose que la boule unité fermée de E est compacte et, par l'absurde, que E est de dimension infinie. Alors, il existe une suite strictement croissante $(E_n)_{n \geq 0}$ de sous-espaces vectoriels de dimension finie de E . La question précédente permet de construire une suite $(u_n)_{n \geq 1}$ de vecteurs de E , de norme égale à 1 (donc appartenant à la boule unité fermée), vérifiant $u_n \in E_n$ et $d(u_n, E_{n-1}) \geq 1/2$ pour tout $n \geq 1$. En particulier, on a $\|u_n - u_m\| \geq 1/2$ si $m < n$, de sorte que la suite $(u_n)_{n \geq 1}$ ne peut admettre de sous-suite convergente. Cela contredit la propriété de Bolzano-Weierstrass, et donc la compacité de la boule unité de E . Ainsi, l'espace E est de dimension finie.

Exercice 8.3 (Théorèmes de point fixe)

Leçon : 253. Référence : [6].

1. Soit (E, d) un espace métrique compact non vide et $f : E \rightarrow E$ une application vérifiant

$$\forall x, y \in E \quad x \neq y \implies d(f(x), f(y)) < d(x, y).$$

- (a). Montrer que f admet un unique point fixe, noté α .
- (b). Soit $x_0 \in E$. On pose $x_{n+1} = f(x_n)$ pour tout $n \in \mathbb{N}$. Montrer que la suite (x_n) ainsi définie converge vers α .
- (c). Ces résultats restent-ils vrais si on suppose seulement E complet ?

2. Soit K un compact convexe non vide d'un espace vectoriel normé, et $f : K \rightarrow K$ continue telle que

$$\forall x, y \in K \quad \|f(x) - f(y)\| \leq \|x - y\|.$$

Montrer que f admet au moins un point fixe.

Corrigé.

1. (a). L'application $x \mapsto d(x, f(x))$ est continue sur le compact E donc y atteint son minimum en un certain point α , i.e. $d(\alpha, f(\alpha)) = \inf_{x \in E} d(x, f(x))$. On suppose que $\alpha \neq f(\alpha)$ et on pose $\beta = f(\alpha)$. Alors,

$$d(\beta, f(\beta)) = d(f(\alpha), f(\beta)) < d(\alpha, \beta) = d(\alpha, f(\alpha)),$$

ce qui contredit la définition de α . D'où $\alpha = f(\alpha)$. Pour montrer que α est l'unique point fixe, on suppose par l'absurde l'existence de $\beta \neq \alpha$ tel que $f(\beta) = \beta$. On a

$$d(\alpha, \beta) = d(f(\alpha), f(\beta)) < d(\alpha, \beta),$$

ce qui est impossible.

- (b). Si $x_{n_0} = \alpha$ pour un certain n_0 , alors on a $x_n = x_{n_0} = \alpha$ pour tout entier $n \geq n_0$ et le résultat est clair. On peut donc supposer que x_n est toujours distinct de α . On note $u_n = d(\alpha, x_n)$ pour tout entier $n \geq 0$. Alors,

$$u_{n+1} = d(\alpha, x_{n+1}) = d(f(\alpha), f(x_n)) < d(\alpha, x_n) = u_n.$$

La suite (u_n) décroît donc strictement. Comme elle est minorée par 0, elle converge vers une limite notée ℓ . De plus, par compacité de E , on peut extraire de la suite $(x_n)_{n \geq 0}$ une sous-suite convergente $(x_{\varphi(n)})_{n \geq 0}$, de limite notée β . Alors, on a

$$u_{\varphi(n)} = d(\alpha, x_{\varphi(n)}) \xrightarrow{n \rightarrow \infty} d(\alpha, \beta).$$

De plus, par continuité de f , on a

$$u_{\varphi(n)+1} = d(\alpha, x_{\varphi(n)+1}) = d(\alpha, f(x_{\varphi(n)})) \xrightarrow{n \rightarrow \infty} d(\alpha, f(\beta)).$$

On en déduit que les distances $d(\alpha, \beta)$ et $d(\alpha, f(\beta))$ sont toutes deux égales à ℓ . Si on suppose $\alpha \neq \beta$, on a

$$\ell = d(\alpha, f(\beta)) = d(f(\alpha), f(\beta)) < d(\alpha, \beta) = \ell,$$

ce qui est impossible. Ainsi, on a $\alpha = \beta$, ce qui donne finalement $\ell = 0$, et démontre que $x_n \xrightarrow{n \rightarrow \infty} \alpha$.

- (c). Le résultat devient faux si on suppose seulement E complet. Par exemple, la fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ telle que $f(x) = 1$ si $x < 0$, et $f(x) = x + \frac{1}{1+x}$ si $x \geq 0$, est continue, sans point fixe et vérifie l'hypothèse de contractivité.
2. Le résultat est vrai si la fonction f est k -contractante pour $k < 1$. En effet, la suite définie par $y_{n+1} = f(y_n)$ pour tout $n \in \mathbb{N}$, avec $y_0 \in K$ quelconque, vérifie $\|y_{n+1} - y_n\| \leq k^n \|y_1 - y_0\|$ par une récurrence immédiate. Il s'agit donc d'une suite de Cauchy, puisque pour $m < n$,

$$\|y_n - y_m\| \leq \|y_{m+1} - y_m\| + \dots + \|y_n - y_{n-1}\| \leq (k^m + \dots + k^{n-1}) \|y_1 - y_0\| \leq \frac{k^m}{1-k} \|y_1 - y_0\|.$$

Comme K est compact, il est complet. La suite (y_n) converge donc dans K , et par continuité de f , sa limite est un point fixe. Il est même facile de voir que ce point fixe est unique.

En revenant au cas initial, pour $a \in K$ et $n \in \mathbb{N}^*$, on pose $f_n(x) = \frac{a}{n} + \left(1 - \frac{1}{n}\right) f(x)$ pour tout $x \in K$. Par convexité de K , on a bien une application de K dans lui-même. De plus, pour tous $x, y \in K$, on a

$$\|f_n(x) - f_n(y)\| = \left(1 - \frac{1}{n}\right) \|f(x) - f(y)\| \leq \left(1 - \frac{1}{n}\right) \|x - y\|.$$

Ainsi, l'application f_n est $(1 - 1/n)$ -contractante, de sorte qu'elle admet un point fixe $x_n \in K$. Par compacité de K , on peut extraire une sous-suite $x_{\varphi(n)}$ qui converge vers un point $x \in K$. On a alors

$$\|f_n(x_n) - f(x)\| \leq \|f_n(x_n) - f_n(x)\| + \|f_n(x) - f(x)\| \leq \frac{\|a\| + \|f(x)\|}{n} + \|x - x_n\|,$$

ce qui montre que $f_{\varphi(n)}(x_{\varphi(n)})$ converge vers $f(x)$. En passant à la limite dans l'égalité $f_{\varphi(n)}(x_{\varphi(n)}) = x_{\varphi(n)}$, on en déduit que $f(x) = x$. Donc f admet un point fixe dans K .

Exercice 8.4 (Théorème de Stone-Weierstrass et polynômes de Bernstein)

Leçon : 241. Références : [6, 7].

Soit $f : [0, 1] \rightarrow \mathbb{R}$ une fonction continue. On considère la suite de fonctions polynomiales $(P_n)_{n \geq 1}$ définie par

$$P_n(t) = \sum_{k=0}^n \binom{n}{k} f\left(\frac{k}{n}\right) t^k (1-t)^{n-k}.$$

Montrer que (P_n) converge uniformément vers f sur $[0, 1]$.

Corrigé.

On présente ici une preuve de nature probabiliste. Une variable de Bernoulli de paramètre $p \in [0, 1]$ vaut 1 avec probabilité p et 0 avec probabilité $1 - p$; la loi binomiale de paramètres n, p est la loi de la somme de n variables de Bernoulli indépendantes de paramètre p . Pour $x \in [0, 1]$, on note $X_1, \dots, X_n$ des variables aléatoires indépendantes qui suivent toutes la loi de Bernoulli de paramètre x . Alors, $S_n = X_1 + \dots + X_n$ suit la loi binomiale de paramètres n et x , et on remarque que $P_n(x) = \mathbb{E}[f(S_n/n)]$. Pour tout $\delta > 0$, on a donc

$$\begin{aligned} |P_n(x) - f(x)| &= \left| \mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right] - f(x) \right| \leq \mathbb{E}\left[\left|f\left(\frac{S_n}{n}\right) - f(x)\right|\right] \\ &\leq \mathbb{E}\left[\left|f\left(\frac{S_n}{n}\right) - f(x)\right| \mathbf{1}_{\{|S_n/n - x| > \delta\}}\right] + \mathbb{E}\left[\left|f\left(\frac{S_n}{n}\right) - f(x)\right| \mathbf{1}_{\{|S_n/n - x| \leq \delta\}}\right] \\ &\leq 2\|f\|_{\infty} \mathbb{P}(|S_n/n - x| > \delta) + \omega_f(\delta), \end{aligned}$$

où $\|f\|_{\infty} = \sup_{y \in [0, 1]} |f(y)|$ et $\omega_f(\delta) = \sup_{\substack{y, y' \in [0, 1] \\ |y' - y| \leq \delta}} |f(y') - f(y)|$ sont respectivement le supremum et le module d'uniforme

continuité de la fonction continue f (donc bornée et uniformément continue) sur l'intervalle $[0, 1]$. De plus, on peut majorer le premier terme de la borne à l'aide de l'inégalité de Tchebychev par $1/\delta^2$ fois la variance de S_n/n , qui vaut $nx(1-x)/n^2$. En observant que $x(1-x) \leq 1/4$ pour tout $x \in [0, 1]$, on a finalement

$$\|P_n - f\|_{\infty} \leq \frac{\|f\|_{\infty}}{2n\delta^2} + \omega_f(\delta).$$

On en déduit que $\limsup_{n \rightarrow \infty} \|P_n - f\|_{\infty} \leq \omega_f(\delta) \xrightarrow{\delta \rightarrow 0} 0$, et donc que la suite des fonctions polynomiales P_n converge bien uniformément sur $[0, 1]$ vers la fonction f .

(Remarque. On a donc une version explicite du théorème de Weierstrass sur l'intervalle $[0, 1]$, et plus généralement sur tous les segments $[a, b]$ par un changement de variable affine.)

Exercice 8.5 (Courbe sinus du topologue)

Leçon : 204.

On considère l'ensemble $S = \{x + i \sin(1/x), 0 < x \leq 1\}$, et son adhérence $\bar{S}$ dans $\mathbb{C}$.

1. Montrer que $\bar{S}$ est connexe et que $\bar{S} = S \cup [-i, i]$.
2. Montrer qu'un point de $[-i, i]$ ne peut pas être joint à un point de S par un chemin de $\bar{S}$. En déduire que $\bar{S}$ n'est pas connexe par arcs.

Corrigé.

1. L'ensemble S est l'image continue de l'intervalle $]0, 1]$ par l'application continue $x \mapsto x + i \sin(1/x)$, donc c'est un ensemble connexe de $\mathbb{C}$. Il en va de même de son adhérence $\bar{S}$. Pour tout $y \in [-1, 1]$, choisissons $\theta \in \mathbb{R}$ tel que $\sin \theta = y$ et posons $t_n = (\theta + 2\pi n)^{-1}$ pour n assez grand. Alors $t_n \in]0, 1]$, $t_n \rightarrow 0$ et $t_n + i \sin(1/t_n) \rightarrow iy$. Réciproquement, toute limite d'une suite de points de S dont la partie réelle tend vers 0 appartient au segment $[-i, i]$; si cette partie réelle tend vers $x > 0$, la limite appartient à S par continuité de $x \mapsto \sin(1/x)$. Ainsi, $\bar{S} = S \cup [-i, i]$.
2. Soit $\gamma : [0, 1] \rightarrow \bar{S}$ un chemin continu d'origine $\gamma(0) \in [-i, i]$. Sa partie réelle $\varphi = \Re \gamma$ est une fonction continue. Si elle n'est pas identiquement nulle, l'ouvert $\{\varphi \neq 0\}$ admet une composante connexe de la forme $]c, d[$ ou $]c, 1]$ avec $c \geq 0$, puisque $\varphi(0) = 0$. Dans les deux cas, il existe un intervalle $]c, d[$ sur lequel φ est non nulle, et avec $\varphi(c) = 0$. Pour tout x de cet intervalle, on a donc $\gamma(x) \in S$, puis $\gamma(x) = \varphi(x) + i \sin \frac{1}{\varphi(x)}$. Par continuité, cette quantité admet une limite finie quand $x \rightarrow c^+$. Or, pour tout $\eta \in]0, d - c[$, le théorème des valeurs intermédiaires montre que $\varphi([c, c + \eta])$ contient $]0, \varphi(c + \eta)]$. La fonction $\sin(1/\varphi(x))$ prend donc les valeurs 1 et -1 arbitrairement près de c , ce qui exclut cette limite. Ainsi, la fonction φ est identiquement nulle, de sorte que le chemin γ reste dans $[-i, i]$. L'ensemble $\bar{S}$ n'est donc pas connexe par arcs. Cet ensemble admet deux composantes connexes par arcs, à savoir S et $[-i, i]$.

Exercice 8.6 (Valeurs d'adhérence d'une suite)

Leçon : 204. Référence : [6].

1. Soit X un espace métrique compact, et soit (x_n) une suite de points de X telle que $d(x_{n+1}, x_n) \xrightarrow{n \rightarrow \infty} 0$. Montrer que l'ensemble A des valeurs d'adhérence de (x_n) est connexe.
2. On conserve l'hypothèse $d(x_{n+1}, x_n) \rightarrow 0$ et on suppose dans cette question que $X = [0, 1]$, et que la suite (x_n) est définie par un terme initial $x_0 \in X$ et la relation récurrente $x_{n+1} = f(x_n)$, pour une fonction continue $f : X \rightarrow X$. Montrer que A est inclus dans l'ensemble des points fixes de f . En déduire que A est réduit à un élément et que la suite (x_n) converge.
3. On suppose dans cette question que $X = \{z \in \mathbb{C} \mid 1/2 \leq |z| \leq 1\}$, et que la suite (x_n) est définie par un terme initial $x_0 \in X$ et la relation récurrente $x_{n+1} = f(x_n)$, où la fonction $f : X \rightarrow X$ est définie par la formule

$$f(z) = \frac{1}{2 - |z|} \frac{z}{|z|} e^{i(1-|z|)}.$$

- (a). Montrer qu'on peut écrire $x_n = r_n e^{i\theta_n}$, avec les relations de récurrence $r_{n+1} = \frac{1}{2 - r_n}$ et $\theta_{n+1} = \theta_n + 1 - r_n$.
- (b). Montrer que si $|x_0| < 1$, on a $x_{n+1} - x_n \xrightarrow{n \rightarrow \infty} 0$, mais que la suite (x_n) diverge. Comment expliquer cette différence de comportement par rapport à la situation précédente?

Corrigé.

1. L'ensemble $A = \bigcap_{N \geq 0} \overline{\{x_n : n \geq N\}}$ est compact et non vide. Pour tout ouvert U contenant A , tous les x_n appartiennent à U à partir d'un certain rang : sinon, une sous-suite dans le compact $X \setminus U$ aurait une valeur d'adhérence hors de A . Supposons $A = F_1 \cup F_2$, où F_1, F_2 sont deux compacts non vides disjoints. Posons $\varepsilon = d(F_1, F_2) > 0$ et $U_i = \{x : d(x, F_i) < \varepsilon/3\}$. Les deux ouverts U_1, U_2 sont à distance au moins $\varepsilon/3$. À partir d'un certain rang, $x_n \in U_1 \cup U_2$ et $d(x_{n+1}, x_n) < \varepsilon/3$: la suite ne peut donc plus passer d'un de ces ouverts à l'autre. Pourtant, chacun contient une valeur d'adhérence et doit être visité une infinité de fois. Cette contradiction prouve que A est connexe.
2. Soit $a \in A$ une valeur d'adhérence, et $x_{n_k} \xrightarrow{k \rightarrow \infty} a$. On observe que $x_{n_k+1} - x_{n_k} = f(x_{n_k}) - x_{n_k}$ converge à la fois vers zéro et vers $f(a) - a$, par continuité de f . Ainsi, $f(a) = a$, de sorte que A est inclus dans l'ensemble des points fixes de f . Supposons maintenant que le connexe A n'est pas réduit à un point. Alors, A contient un intervalle $[c - h, c + h]$, avec $c \in A$ et $h > 0$. Si un terme de la suite x_n appartient à A , alors celle-ci est

ultimement constante, puisque A est constitué de points fixes de f . On a donc $x_n \notin A$ pour tout n , de sorte que $|x_n - c| > h$. Cela contredit le fait que c est une valeur d'adhérence. Finalement, A est réduit à un point, ce qui montre que la suite (x_n) est convergente.

3. (a). On vérifie rapidement que $f(X) \subset X$, puisque pour tout $z \in X$, le module $|f(z)| = \frac{1}{2 - |z|}$ est compris entre $\frac{1}{2 - 1/2} = \frac{2}{3} \geq \frac{1}{2}$ et $\frac{1}{2 - 1} = 1$. Ensuite, si on écrit $x_n = r_n e^{i\theta_n}$, alors

$$r_{n+1} e^{i\theta_{n+1}} = x_{n+1} = f(x_n) = \frac{1}{2 - |x_n|} \frac{x_n}{|x_n|} e^{i(1 - |x_n|)} = \frac{1}{2 - r_n} e^{i\theta_n} e^{i(1 - r_n)},$$

de sorte que $r_{n+1} = \frac{1}{2 - r_n}$ et $\theta_{n+1} = \theta_n + 1 - r_n$.

- (b). On suppose $r_0 = |x_0| < 1$. Concernant la suite des modules, on observe que $\frac{1}{1 - r_{n+1}} - \frac{1}{1 - r_n} = 1$, si bien que $\frac{1}{1 - r_n} = n + c$ avec $c = \frac{1}{1 - r_0}$. Pour les arguments, on a $\theta_n = \theta_0 + \sum_{k=0}^{n-1} (1 - r_k) = \theta_0 + \sum_{k=0}^{n-1} \frac{1}{k + c}$. On en déduit que

$$\begin{aligned} |x_{n+1} - x_n| &= \left| \left(1 - \frac{1}{n + 1 + c}\right) \exp\left(i\theta_0 + \sum_{k=0}^n \frac{i}{k + c}\right) - \left(1 - \frac{1}{n + c}\right) \exp\left(i\theta_0 + \sum_{k=0}^{n-1} \frac{i}{k + c}\right) \right| \\ &= \left| \left(1 - \frac{1}{n + 1 + c}\right) \exp\left(\frac{i}{n + c}\right) - \left(1 - \frac{1}{n + c}\right) \right| \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

L'ensemble A des valeurs d'adhérence de la suite (x_n) est donc un ensemble connexe inclus dans l'ensemble des points fixes de f , qui est le cercle unité Γ . Cependant, on voit assez facilement que les termes $e^{i\theta_n}$ sont denses dans Γ car $\theta_n \xrightarrow{n \rightarrow \infty} \infty$ et $\theta_{n+1} - \theta_n \xrightarrow{n \rightarrow \infty} 0$. Comme $r_n \xrightarrow{n \rightarrow \infty} 1$, il en résulte que l'ensemble A est le cercle Γ tout entier. En particulier, la suite (x_n) est divergente. La différence avec la situation précédente est qu'il existe dans $\mathbb{C}$ des ensembles connexes non réduits à un point et d'intérieur vide, alors que dans $\mathbb{R}$ non.

Espaces vectoriels normés

Références : [6, 12].

Exercice 9.1 (Formes linéaires continues et norme quotient)

Leçon : 208. Référence : [6].

Soit E un $\mathbb{R}$ -espace vectoriel normé, et soit $\varphi : E \rightarrow \mathbb{R}$ une forme linéaire.

1. Montrer que φ est continue si et seulement si $\ker \varphi$ est un fermé de E .
2. Soit F un sous-espace vectoriel de E . On note $\dot{x} = x + F$ la classe de x dans E/F . Une semi-norme sur un espace vectoriel V est une application $p : V \rightarrow [0, +\infty[$ telle que $p(\lambda u) = |\lambda|p(u)$ et $p(u + v) \leq p(u) + p(v)$, sans imposer que $p(u) = 0$ entraîne $u = 0$. Montrer que l'application N définie sur l'espace quotient E/F par $N(\dot{x}) = \inf_{y \in \dot{x}} \|y\|$ est une semi-norme. Que peut-on dire si F est fermé? Retrouver le résultat de la question précédente.

Corrigé.

1. Si φ est continue, il est clair que son noyau $\ker \varphi$, qui est l'image réciproque du fermé $\{0\}$, est un fermé. Réciproquement, supposons que $\ker \varphi$ est fermé, et supposons par l'absurde que φ n'est pas continue. Ainsi, φ n'est pas bornée sur la sphère unité, et il existe une suite (x_n) d'éléments de E tels que $\|x_n\| = 1$ pour tout n , et $|\varphi(x_n)| \xrightarrow{n \rightarrow \infty} +\infty$. Comme la forme linéaire φ n'est pas nulle (sinon, elle serait trivialement continue), on peut trouver un vecteur $u \in E$ tel que $\varphi(u) = 1$. Pour tout n , en posant $u_n = u - \frac{x_n}{\varphi(x_n)}$, on observe que $\varphi(u_n) = \varphi(u) - \frac{\varphi(x_n)}{\varphi(x_n)} = 0$, c'est-à-dire que $u_n \in \ker \varphi$. Cependant, on a $u_n \xrightarrow{n \rightarrow \infty} u$, et comme $\ker \varphi$ est fermé, on a par passage à la limite $u \in \ker \varphi$, ce qui donne une contradiction. On en conclut que φ est bien continue.
2. Le cas $\lambda = 0$ de l'homogénéité est immédiat puisque $N(\dot{0}) = 0$. Soit maintenant $x \in E$ et $\lambda \in \mathbb{R} \setminus \{0\}$. La classe de x dans l'espace quotient E/F s'écrit $\dot{x} = x + F$. De plus, on a $\lambda \dot{x} = \lambda(x + F) = \lambda x + F = \lambda \dot{x}$ (classe de λx). Ainsi, $N(\lambda \dot{x}) = \inf_{y \in \lambda \dot{x}} \|y\| = \inf_{z \in F} \|\lambda(x + z)\| = |\lambda| \inf_{z \in F} \|x + z\| = |\lambda| \inf_{y \in \dot{x}} \|y\| = |\lambda| N(\dot{x})$. Pour montrer l'inégalité triangulaire, considérons $x, y \in E$, ainsi que $u, v \in F$. En écrivant

$$N(x + y) \leq \|(x + y) + (u + v)\| \leq \|x + u\| + \|y + v\|,$$

et en passant à l'infimum sur u et sur v dans le membre de droite, on obtient bien $N(x + y) \leq N(\dot{x}) + N(\dot{y})$. Ainsi, N est bien une semi-norme. Si le sous-espace vectoriel F est fermé, alors N est une norme. En effet, si $0 = N(\dot{x}) = \inf_{z \in F} \|x - z\|$, alors la distance de x au fermé F est nulle, ce qui montre que $x \in F$, et donc $\dot{x} = 0$.

Pour retrouver le résultat de la question précédente, on suppose que F est le noyau d'une forme linéaire φ , et que F est fermé. On peut bien sûr supposer que φ est non nulle. On peut alors factoriser φ sous la forme $\varphi = \psi \circ s$, où s est la surjection canonique de E dans E/F , et ψ est une forme linéaire sur E/F . La forme linéaire ψ est continue, puisque E/F est de dimension finie. De même, la surjection s est clairement continue, puisque $N(\dot{x} - \dot{y}) \leq \|x - y\|$ pour tous $x, y \in E$. On en déduit que φ est continue.

Exercice 9.2 (Sur l'exponentielle complexe)

Référence : [12].

1. Soit (x_n) une suite de réels. Montrer que les deux propositions suivantes sont équivalentes :
 - (a). la suite (e^{itx_n}) converge pour tout $t \in \mathbb{R}$;
 - (b). la suite (x_n) converge.
2. La loi gaussienne $\mathcal{N}(\mu, \sigma^2)$, pour $\sigma > 0$, a pour densité $x \mapsto (\sigma\sqrt{2\pi})^{-1} \exp(-(x - \mu)^2/(2\sigma^2))$; pour $\sigma = 0$, elle désigne la masse de Dirac en μ . Soit (X_n) une suite de variables aléatoires réelles gaussiennes et soit X une variable aléatoire réelle. On suppose que (X_n) converge en loi vers X , c'est-à-dire que $\mathbb{E}[h(X_n)] \rightarrow \mathbb{E}[h(X)]$ pour toute fonction continue bornée $h : \mathbb{R} \rightarrow \mathbb{R}$. Dédurre de la question précédente que la variable X est gaussienne, éventuellement dégénérée (c'est-à-dire constante presque sûrement).

Corrigé.

La fonction caractéristique d'une variable réelle Y est $t \mapsto \mathbb{E}[e^{itY}]$.

1. Si la suite (x_n) converge, alors il est clair que la suite (e^{itx_n}) converge quel que soit $t \in \mathbb{R}$. Pour établir la réciproque, supposons que la suite (x_n) admette deux valeurs d'adhérence ℓ et ℓ' . Alors, pour tout $t \in \mathbb{R}$, la suite convergente (e^{itx_n}) converge le long d'une sous-suite vers $e^{it\ell}$ et le long d'une autre sous-suite vers $e^{it\ell'}$. On a donc $e^{it\ell} = e^{it\ell'}$ pour tout réel t . En dérivant en 0, on en déduit que $\ell = \ell'$. Donc, la suite (x_n) ne peut admettre qu'une seule valeur d'adhérence.

Montrons maintenant que cette suite est bornée. Si ce n'était pas le cas, elle admettrait une sous-suite qui tendrait vers $+\infty$, ou vers $-\infty$. Supposons donc, par exemple, qu'il existe une extraction de la suite pour laquelle $x_{\varphi(n)} \xrightarrow{n \rightarrow \infty} +\infty$. On remarque que la fonction définie par $f(t) = \lim_{n \rightarrow \infty} e^{itx_{\varphi(n)}}$ est intégrable sur $[0, 1]$, a pour module 1, et vérifie par convergence dominée

$$1 = \int_0^1 |f(t)|^2 dt = \lim_{n \rightarrow \infty} \int_0^1 \overline{f(t)} e^{itx_{\varphi(n)}} dt.$$

Cependant, le lemme de Riemann-Lebesgue implique que cette dernière limite est nulle, ce qui donne une contradiction. Ainsi, la suite (x_n) est bornée, et comme on sait qu'elle admet au plus une valeur d'adhérence, elle est convergente.

2. On note respectivement μ_n et σ_n^2 l'espérance et la variance de chaque variable aléatoire X_n . L'hypothèse implique que les fonctions caractéristiques $\mathbb{E}[e^{itX_n}] = \exp\left(i\mu_n t - \frac{\sigma_n^2 t^2}{2}\right)$ convergent simplement pour tout $t \in \mathbb{R}$ vers la fonction caractéristique $\mathbb{E}[e^{itX}]$. La fonction caractéristique de X est continue en 0 et vaut 1 en ce point. On peut donc choisir $t_0 \neq 0$ tel que $|\mathbb{E}[e^{it_0 X}]| > 0$. En prenant le module à $t = t_0$, on obtient $\exp(-\sigma_n^2 t_0^2/2) \rightarrow |\mathbb{E}[e^{it_0 X}]| > 0$. En prenant le logarithme, on en déduit que la suite des variances (σ_n^2) converge vers une limite finie notée σ^2 . En multipliant par $\exp\left(\frac{\sigma_n^2 t^2}{2}\right)$, on en déduit que la suite $(e^{i\mu_n t})$ converge pour tout $t \in \mathbb{R}$. D'après la question précédente, la suite (μ_n) converge donc, vers une limite notée μ . Il en ressort par passage à la limite dans les fonctions caractéristiques que

$$\mathbb{E}[e^{itX}] = \lim_{n \rightarrow \infty} \mathbb{E}[e^{itX_n}] = \lim_{n \rightarrow \infty} \exp\left(i\mu_n t - \frac{\sigma_n^2 t^2}{2}\right) = \exp\left(i\mu t - \frac{\sigma^2 t^2}{2}\right).$$

Cela prouve que X suit la loi $\mathcal{N}(\mu, \sigma^2)$ si $\sigma^2 > 0$. Si $\sigma^2 = 0$, sa fonction caractéristique est celle de la constante μ , donc $X = \mu$ presque sûrement. Ce cas peut effectivement se produire, par exemple pour $X_n = Z/n$, avec Z de loi $\mathcal{N}(0, 1)$.

Calcul différentiel et extrema

Références : [5, 6, 10, 12].

Exercice 10.1 (Dérivées partielles sans continuité)

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction définie par

$$f(x, y) = \begin{cases} \left(\frac{x^2 y}{x^4 + y^2} \right)^2 & \text{si } y \neq 0 \\ 0 & \text{si } y = 0. \end{cases}$$

Montrer que f admet des dérivées partielles en $(0, 0)$ mais qu'elle n'est pas continue en ce point.

Corrigé.

Pour tout $x \neq 0$,

$$\frac{f(x, 0) - f(0, 0)}{x - 0} = 0 \quad \implies \quad \frac{\partial f}{\partial x}(0, 0) = 0$$

et pour tout $y \neq 0$,

$$\frac{f(0, y) - f(0, 0)}{y - 0} = 0 \quad \implies \quad \frac{\partial f}{\partial y}(0, 0) = 0.$$

Par ailleurs, f n'est pas continue en $(0, 0)$, car par exemple $f(x, x^2) = 1/4$ pour tout $x \neq 0$.

Exercice 10.2 (Dérivées directionnelles et gradient)

Référence : [6].

Soient U un ouvert de $\mathbb{R}^n$, $a \in U$ et $v \in \mathbb{R}^n$. On dit qu'une fonction $f : U \rightarrow \mathbb{R}$ est *dérivable en a suivant v* si

$$\lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t} \text{ existe.}$$

Dans ce cas, ce nombre est appelé la *dérivée de f en a suivant v* et est noté $\partial_v f(a)$. On suppose dans la suite que f est dérivable en a suivant v .

1. À quoi correspondent les dérivées partielles de f ?
2. Supposons f différentiable en a . Montrer que, pour tout vecteur $v = (v_1, \dots, v_n)$,

$$\partial_v f(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) v_i = \nabla f(a) \cdot v,$$

où $\nabla f(a)$ désigne le gradient de f au point a .

3. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction définie par $f(x, y) = \frac{y^2}{x}$ si $x \neq 0$ et $f(0, y) = y$, pour tout $y \in \mathbb{R}$. Montrer que f est dérivable au point $(0, 0)$ suivant n'importe quel vecteur $v \in \mathbb{R}^2$. Calculer $\partial_v f(0, 0)$. L'application f est-elle différentiable ?
4. Soit $\mathbb{S}^{n-1} = \{v \in \mathbb{R}^n \mid \|v\| = 1\}$ la sphère unité de $\mathbb{R}^n$ pour la norme euclidienne. On suppose que f est différentiable en a . Notons

$$M = \sup_{v \in \mathbb{S}^{n-1}} \partial_v f(a).$$

Donner la valeur de M et le vecteur v pour lequel elle est atteinte. Interpréter géométriquement cela.

5. Vous êtes sur une montagne dont la surface est donnée par l'équation $z = \max\{-x^2 - y^2 + 1800, 0\}$, au point a de coordonnées $(20, 20, 1000)$. Dans quelle direction devez-vous marcher pour atteindre le sommet au plus vite ?

Corrigé.

1. On a $\frac{\partial f}{\partial x_i} = \partial_{e_i} f$.
2. Si $v = 0$, le résultat est immédiat. Sinon, dans la définition de la différentiabilité, on prend $h = tv$, il vient

$$\lim_{t \rightarrow 0} \frac{|f(a + tv) - f(a) - t df(a)(v)|}{|t||v|} = 0,$$

ce qui montre que

$$\lim_{t \rightarrow 0} \left| \frac{f(a + tv) - f(a)}{t} - df(a)(v) \right| = 0.$$

On en déduit que $\partial_v f(a) = df(a)(v)$.

3. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction définie pour tout $(x, y) \in \mathbb{R}^2$ par

$$f(x, y) = \begin{cases} \frac{y^2}{x} & \text{si } x \neq 0 \\ y & \text{si } x = 0. \end{cases}$$

Pour tout $v = (v_1, v_2) \in \mathbb{R}^2$, on a alors

$$f(tv) = \begin{cases} t \frac{v_2^2}{v_1} & \text{si } tv_1 \neq 0 \\ tv_2 & \text{si } tv_1 = 0, \end{cases}$$

et donc, si $t \neq 0$,

$$\frac{f(tv) - f(0, 0)}{t} = \frac{f(tv)}{t} = \begin{cases} \frac{v_2^2}{v_1} & \text{si } tv_1 \neq 0 \\ v_2 & \text{si } tv_1 = 0. \end{cases}$$

On en déduit que

$$\partial_v f(0, 0) = \begin{cases} \frac{v_2^2}{v_1} & \text{si } v_1 \neq 0 \\ v_2 & \text{si } v_1 = 0. \end{cases}$$

Comme l'application $v \mapsto \partial_v f(0, 0)$ n'est pas linéaire, la fonction f n'est pas différentiable en $(0, 0)$.

4. On a

$$M = \sup_{\|v\|=1} \partial_v f(a) = \sup_{\|v\|=1} \nabla f(a) \cdot v = \|\nabla f(a)\|$$

et, si $\nabla f(a) \neq 0$, le supremum est atteint en $v = \frac{\nabla f(a)}{\|\nabla f(a)\|}$. Si $\nabla f(a) = 0$, toute direction unitaire réalise le supremum $M = 0$. On interprète donc $\nabla f(a)$ comme la direction de plus grande pente de f en a .

5. On pose $f(x, y) = \max\{-x^2 - y^2 + 1800, 0\}$ pour tout $(x, y) \in \mathbb{R}^2$. Le sommet de la montagne est représenté par le maximum de la fonction f qui est atteint en $(x, y) = (0, 0)$ et la valeur du maximum est 1800. Au point a de coordonnées $(20, 20)$, la fonction f est différentiable car $f(20, 20) = 1000$ et donc $f(x, y) = -x^2 - y^2 + 1800$ pour (x, y) dans un voisinage de a . Afin d'atteindre le sommet le plus rapidement possible, il convient de prendre la direction de la plus grande pente, i.e. la direction $\nabla f(20, 20) = (-40, -40)$.

Exercice 10.3 (Laplacien des fonctions radiales)

Soit $g :]0, +\infty[\rightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^2$. On considère la fonction f définie sur l'ouvert $U = \mathbb{R}^2 \setminus \{(0, 0)\}$ par la formule $f(x, y) = g(\sqrt{x^2 + y^2})$ pour tout $(x, y) \in U$.

1. Expliquez pourquoi la fonction f est de classe $\mathcal{C}^2$ sur U . Calculer les dérivées partielles premières $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$, et les dérivées partielles secondes $\frac{\partial^2 f}{\partial x^2}$ et $\frac{\partial^2 f}{\partial y^2}$.
2. On rappelle que le laplacien de f est défini par $\Delta f(x, y) = \frac{\partial^2 f}{\partial x^2}(x, y) + \frac{\partial^2 f}{\partial y^2}(x, y)$ pour tout $(x, y) \in U$. Montrer que pour tout $(x, y) \in U$,

$$\Delta f(x, y) = g''(\sqrt{x^2 + y^2}) + \frac{g'(\sqrt{x^2 + y^2})}{\sqrt{x^2 + y^2}}.$$

3. En déduire que $\Delta f(x, y) = 0$ pour tout $(x, y) \in U$ si et seulement si $\frac{1}{r} \frac{d}{dr}(rg'(r)) = 0$ pour tout $r > 0$.
4. Montrer alors que $\Delta f(x, y) = 0$ pour tout $(x, y) \in U$ si et seulement si on peut trouver deux réels a et b tels que pour tout $(x, y) \in U$,

$$f(x, y) = a \ln(\sqrt{x^2 + y^2}) + b.$$

Corrigé.

1. La fonction f est de classe $\mathcal{C}^2$ sur U comme composée de fonctions de classe $\mathcal{C}^2$. Pour tout $(x, y) \neq (0, 0)$, on a

$$\frac{\partial f}{\partial x}(x, y) = g'(\sqrt{x^2 + y^2}) \frac{x}{\sqrt{x^2 + y^2}} \quad \text{et} \quad \frac{\partial f}{\partial y}(x, y) = g'(\sqrt{x^2 + y^2}) \frac{y}{\sqrt{x^2 + y^2}},$$

ainsi que

$$\frac{\partial^2 f}{\partial x^2}(x, y) = g''(\sqrt{x^2 + y^2}) \frac{x^2}{x^2 + y^2} + g'(\sqrt{x^2 + y^2}) \frac{y^2}{(x^2 + y^2)^{3/2}}$$

et

$$\frac{\partial^2 f}{\partial y^2}(x, y) = g''(\sqrt{x^2 + y^2}) \frac{y^2}{x^2 + y^2} + g'(\sqrt{x^2 + y^2}) \frac{x^2}{(x^2 + y^2)^{3/2}}.$$

2. En sommant, on obtient effectivement

$$\Delta f(x, y) = g''(\sqrt{x^2 + y^2}) + \frac{g'(\sqrt{x^2 + y^2})}{\sqrt{x^2 + y^2}}.$$

3. En posant $r = \sqrt{x^2 + y^2}$, on obtient que $\Delta f(x, y) = 0$ pour tout $(x, y) \in U$ si et seulement si

$$\forall r > 0 \quad \frac{1}{r} \frac{d}{dr}(rg'(r)) = g''(r) + \frac{1}{r}g'(r) = 0.$$

4. On intègre l'équation différentielle précédente : il existe un réel a tel que $rg'(r) = a$ pour tout $r > 0$, soit $g'(r) = \frac{a}{r}$, puis enfin il existe un autre réel b tel que $g(r) = a \ln r + b$ pour tout $r > 0$. Finalement,

$$\forall (x, y) \in U \quad f(x, y) = a \ln(\sqrt{x^2 + y^2}) + b.$$

Exercice 10.4 (Dérivées croisées et théorème de Schwarz)

Référence : [6].

On pose $\Omega = \mathbb{R}^2 \setminus \{(0, 0)\}$. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction définie par

$$f(x, y) = \begin{cases} xy \frac{x^2 - y^2}{x^2 + y^2} & \text{si } (x, y) \in \Omega \\ 0 & \text{si } (x, y) = (0, 0). \end{cases}$$

1. Montrer que la fonction f est différentiable sur Ω et calculer sa différentielle.
2. Montrer que f est différentiable en $(0, 0)$ et que sa différentielle en ce point est nulle.
3. Montrer que f admet en tout point des dérivées partielles secondes $\frac{\partial^2 f}{\partial x \partial y}$ et $\frac{\partial^2 f}{\partial y \partial x}$, et calculer leurs valeurs en $(0, 0)$. Qu'en déduit-on pour la continuité de ces dérivées partielles secondes en $(0, 0)$?

Corrigé.

1. La fonction f est bien définie, continue et admet des dérivées partielles sur $\mathbb{R}^2 \setminus \{(0, 0)\}$. De plus, pour tout $(x, y) \neq (0, 0)$, on a

$$\frac{\partial f}{\partial x}(x, y) = y \frac{x^4 + 4x^2y^2 - y^4}{(x^2 + y^2)^2} \quad \text{et} \quad \frac{\partial f}{\partial y}(x, y) = x \frac{x^4 - 4x^2y^2 - y^4}{(x^2 + y^2)^2}.$$

Les fonctions $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ étant continues sur Ω , on en déduit que f est différentiable sur Ω .

2. Le calcul des dérivées partielles en $(0, 0)$ donne

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{x \rightarrow 0} \frac{f(x, 0) - f(0, 0)}{x - 0} = 0 \quad \text{et} \quad \frac{\partial f}{\partial y}(0, 0) = \lim_{y \rightarrow 0} \frac{f(0, y) - f(0, 0)}{y - 0} = 0.$$

Les fonctions $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont continues $\mathbb{R}^2 \setminus \{(0, 0)\}$. Pour étudier la continuité des dérivées partielles en $(0, 0)$ on pose $x = r \cos \theta$ et $y = r \sin \theta$ où $r = \sqrt{x^2 + y^2} > 0$ et $\theta \in [0, 2\pi[$. Notons que $(x, y) \rightarrow 0$ si et seulement si $r \rightarrow 0$. Il vient

$$\left| \frac{\partial f}{\partial x}(x, y) \right| \leq r |\sin \theta| |\cos^4 \theta + 4 \cos^2 \theta \sin^2 \theta - \sin^4 \theta| \leq 6r \xrightarrow{(x,y) \rightarrow (0,0)} 0 = \frac{\partial f}{\partial x}(0, 0),$$

ce qui prouve que $\frac{\partial f}{\partial x}$ est continue en $(0, 0)$. On montre de même que $\frac{\partial f}{\partial y}$ est continue en $(0, 0)$. Finalement, les dérivées partielles de f sont continues sur $\mathbb{R}^2$ ce qui assure que f est différentiable sur $\mathbb{R}^2$ et que $df(0, 0)$ est l'application linéaire nulle.

3. On calcule

$$\frac{\frac{\partial f}{\partial x}(0, y) - \frac{\partial f}{\partial x}(0, 0)}{y - 0} = -1 \quad \implies \quad \frac{\partial^2 f}{\partial y \partial x}(0, 0) = -1$$

et

$$\frac{\frac{\partial f}{\partial y}(x, 0) - \frac{\partial f}{\partial y}(0, 0)}{x - 0} = 1 \quad \implies \quad \frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1.$$

Les deux dérivées croisées coïncident sur Ω , mais leur valeur commune est 1 sur l'axe $y = 0$ et -1 sur l'axe $x = 0$, hors de l'origine. Aucune des deux n'a donc de limite en $(0, 0)$. Les hypothèses de continuité du théorème de Schwarz ne sont pas satisfaites.

Exercice 10.5 (Différentielle du déterminant)

Référence : [6].

Soit $f : \mathcal{M}_n(\mathbb{R}) \rightarrow \mathbb{R}$ l'application définie par $f(A) = \det(A)$. Montrer que f est de classe $\mathcal{C}^\infty$ et calculer sa différentielle.

Corrigé.

On pose $f(A) = \det(A)$ pour tout $A \in \mathcal{M}_n(\mathbb{R})$. La fonction f est de classe $\mathcal{C}^\infty$ sur $\mathcal{M}_n(\mathbb{R})$ car c'est une fonction polynomiale des coefficients de A . Calculons les dérivées partielles de f . Soit $H^{(i,j)}$ la matrice dont tous les coefficients sont nuls, excepté le coefficient d'indice (i, j) qui vaut 1. Alors, d'après la multilinéarité du déterminant, on a

$$f(A + tH^{(i,j)}) = \det(A + tH^{(i,j)}) = \det(A) + t \operatorname{Com}(A)_{i,j},$$

ce qui montre que

$$\frac{\partial f}{\partial A_{i,j}}(A) = \lim_{t \rightarrow 0} \frac{f(A + tH^{(i,j)}) - f(A)}{t} = \operatorname{Com}(A)_{i,j}.$$

On en déduit que pour tout $H \in \mathcal{M}_n(\mathbb{R})$,

$$df(A)(H) = \sum_{1 \leq i, j \leq n} \operatorname{Com}(A)_{i,j} H_{i,j} = \operatorname{Tr}({}^t \operatorname{Com}(A) H).$$

Quand $A \in \operatorname{GL}_n(\mathbb{R})$, on a $A^{-1} = \frac{1}{\det(A)} {}^t \operatorname{Com}(A)$, ce qui montre alors que $df(A)(H) = \det(A) \operatorname{Tr}(A^{-1} H)$.

Exercice 10.6 (Différentielle de l'inverse)

Référence : [6].

Soit $\mathcal{L}(\mathbb{R}^n)$ l'ensemble des applications linéaires continues de $\mathbb{R}^n$ dans lui-même, muni de la norme d'opérateur, et soit $\operatorname{GL}_n(\mathbb{R})$ le sous-ensemble de $\mathcal{L}(\mathbb{R}^n)$ constitué des éléments inversibles.

1. Montrer que $\operatorname{GL}_n(\mathbb{R})$ est un ouvert de $\mathcal{L}(\mathbb{R}^n)$.
2. Montrer que l'application $\operatorname{inv} : u \in \operatorname{GL}_n(\mathbb{R}) \mapsto u^{-1}$ est différentiable et calculer sa différentielle.

Corrigé.

1. Soit $h \in \mathcal{L}(\mathbb{R}^n)$ tel que

$$\|h\| = \sup_{x \in \mathbb{R}^n \setminus \{0\}} \frac{\|h(x)\|}{\|x\|} < 1.$$

Alors, $\|h^k\| \leq \|h\|^k$ pour tout $k \in \mathbb{N}$, et donc $\sum_{k \geq 0} h^k$ converge absolument dans $\mathcal{L}(\mathbb{R}^n)$. De plus,

$$(I - h) \sum_{k=0}^m h^k = \sum_{k=0}^m h^k - \sum_{k=1}^{m+1} h^k = I - h^{m+1}.$$

Sachant que $\|h\| < 1$, il vient $\|h^{m+1}\| \leq \|h\|^{m+1} \rightarrow 0$, ce qui montre que

$$(I - h) \sum_{k=0}^{\infty} h^k = I.$$

De même, on a

$$\left(\sum_{k \geq 0} h^k \right) (I - h) = I.$$

On en déduit que $I - h \in \operatorname{GL}_n(\mathbb{R})$ et $(I - h)^{-1} = \sum_{k=0}^{\infty} h^k$.

Soit maintenant $u \in \operatorname{GL}_n(\mathbb{R})$ et $h \in \mathcal{L}(\mathbb{R}^n)$ tels que $\|h\| < \frac{1}{\|u^{-1}\|} = R$. Comme $u + h = u(I + u^{-1}h)$ et

$\|u^{-1}h\| \leq \|u^{-1}\| \|h\| < 1$, on en déduit que $I + u^{-1}h \in \operatorname{GL}_n(\mathbb{R})$ puis que $u + h \in \operatorname{GL}_n(\mathbb{R})$. Par conséquent, la boule ouverte $B(u, R)$ est incluse dans $\operatorname{GL}_n(\mathbb{R})$. Donc, $\operatorname{GL}_n(\mathbb{R})$ est ouvert dans $\mathcal{L}(\mathbb{R}^n)$.

2. On a

$$(u + h)^{-1} = (I + u^{-1}h)^{-1}u^{-1} = \sum_{k=0}^{\infty} (-1)^k (u^{-1}h)^k u^{-1} = u^{-1} - u^{-1}hu^{-1} + \sum_{k=2}^{\infty} (-1)^k (u^{-1}h)^k u^{-1}.$$

Comme $h \in \mathcal{L}(\mathbb{R}^n) \mapsto -u^{-1}hu^{-1} \in \mathcal{L}(\mathbb{R}^n)$ est linéaire et comme

$$\left\| \sum_{k=2}^{\infty} (-1)^k (u^{-1}h)^k u^{-1} \right\| \leq \sum_{k=2}^{\infty} \|u^{-1}\|^{k+1} \|h\|^k = \|u^{-1}\|^3 \|h\|^2 \sum_{l=0}^{\infty} \|u^{-1}\|^l \|h\|^l = \frac{\|u^{-1}\|^3 \|h\|^2}{1 - \|u^{-1}\| \|h\|} = o(\|h\|),$$

on en déduit que inv est différentiable et que $d \text{inv}(u)(h) = -u^{-1}hu^{-1}$ pour tout $h \in \mathcal{L}(\mathbb{R}^n)$.

Exercice 10.7 (Coordonnées polaires et cylindriques)

Leçon : 214. Référence : [5].

Montrer que les fonctions suivantes sont des $\mathcal{C}^1$ -difféomorphismes globaux

$$\begin{aligned} \Phi :]0, +\infty[\times]0, 2\pi[&\rightarrow \mathbb{R}^2 \setminus ([0, +\infty[\times \{0\}) \\ (r, \theta) &\mapsto (r \cos \theta, r \sin \theta) \end{aligned}$$

et

$$\begin{aligned} \Psi :]0, +\infty[\times]0, 2\pi[\times \mathbb{R} &\rightarrow \mathbb{R}^3 \setminus ([0, +\infty[\times \{0\} \times \mathbb{R}) \\ (r, \theta, z) &\mapsto (r \cos \theta, r \sin \theta, z). \end{aligned}$$

Corrigé.

Soit $\Phi :]0, +\infty[\times]0, 2\pi[\rightarrow \mathbb{R}^2 \setminus ([0, +\infty[\times \{0\})$ la fonction définie par

$$\forall (r, \theta) \in]0, +\infty[\times]0, 2\pi[\quad \Phi(r, \theta) = (r \cos \theta, r \sin \theta).$$

La fonction Φ est de classe $\mathcal{C}^\infty$ et pour tout $(r, \theta) \in]0, +\infty[\times]0, 2\pi[$, le jacobien s'écrit

$$J_\Phi(r, \theta) = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}.$$

Par conséquent,

$$\det J_\Phi(r, \theta) = r \cos^2 \theta + r \sin^2 \theta = r > 0.$$

De plus, si $\Phi(r, \theta) = \Phi(r', \theta')$, on obtient $r^2 = r'^2$, soit $r = r'$ puisque $r, r' > 0$. Aussi, $(\cos \theta, \sin \theta) = (\cos \theta', \sin \theta')$, ce qui implique que $\theta = \theta' \pmod{2\pi}$, et donc $\theta = \theta'$ puisque $\theta, \theta' \in]0, 2\pi[$. On en déduit que Φ est injective sur $]0, +\infty[\times]0, 2\pi[$. Le théorème d'inversion globale montre alors que Φ réalise un $\mathcal{C}^\infty$ -difféomorphisme de $]0, +\infty[\times]0, 2\pi[$ sur son image, qu'il s'agit d'identifier.

Clairement, on a

$$\Phi([0, +\infty[\times]0, 2\pi[) \subset \mathbb{R}^2 \setminus ([0, +\infty[\times \{0\}).$$

Soit (x, y) dans l'ensemble de droite. On pose $r = \sqrt{x^2 + y^2} > 0$ de sorte que $\left(\frac{x}{r}, \frac{y}{r}\right) \in \mathbb{S}^1 \setminus \{(1, 0)\}$. Il existe donc un unique $\theta \in]0, 2\pi[$ tel que $\left(\frac{x}{r}, \frac{y}{r}\right) = (\cos \theta, \sin \theta)$, c'est-à-dire

$$(x, y) = (r \cos \theta, r \sin \theta) = \Phi(r, \theta) \in \Phi([0, +\infty[\times]0, 2\pi[).$$

Il s'agit du changement de variables en coordonnées polaires.

Soit $\Psi :]0, +\infty[\times]0, 2\pi[\times \mathbb{R} \rightarrow \mathbb{R}^3 \setminus ([0, +\infty[\times \{0\} \times \mathbb{R})$ la fonction définie par

$$\forall (r, \theta, z) \in]0, +\infty[\times]0, 2\pi[\times \mathbb{R} \quad \Psi(r, \theta, z) = (r \cos \theta, r \sin \theta, z).$$

La fonction Ψ est de classe $\mathcal{C}^\infty$ et pour tout $(r, \theta, z) \in]0, +\infty[\times]0, 2\pi[\times \mathbb{R}$, on a comme jacobien

$$J_\Psi(r, \theta, z) = \begin{pmatrix} \cos \theta & -r \sin \theta & 0 \\ \sin \theta & r \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Par conséquent,

$$\det J_\Psi(r, \theta, z) = r \cos^2 \theta + r \sin^2 \theta = r > 0.$$

L'injectivité sur $]0, +\infty[\times]0, 2\pi[\times \mathbb{R}$ est identique au cas précédent. Le théorème d'inversion globale montre alors que Ψ réalise un $\mathcal{C}^\infty$ -difféomorphisme de $]0, +\infty[\times]0, 2\pi[\times \mathbb{R}$ sur son image et on montre de même que

$$\Psi([0, +\infty[\times]0, 2\pi[\times \mathbb{R}) = \mathbb{R}^3 \setminus ([0, +\infty[\times \{0\} \times \mathbb{R}).$$

Il s'agit du changement de variables en coordonnées cylindriques.

Exercice 10.8 (Inversion locale sans inversion globale)

Leçon : 214. Référence : [6].

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ le champ de vecteurs défini par

$$f(x, y) = (x^2 - y^2, 2xy).$$

Montrer que f est un $\mathcal{C}^1$ -difféomorphisme local en tout point de $\mathbb{R}^2 \setminus \{(0, 0)\}$, mais que ça n'est pas un $\mathcal{C}^1$ -difféomorphisme global sur $\mathbb{R}^2 \setminus \{(0, 0)\}$.

Corrigé.

Chacune des composantes de f est une fonction polynomiale. La fonction f est donc de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^2$. De plus, pour tout $(x, y) \in \mathbb{R}^2$, le jacobien s'écrit

$$J_f(x, y) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}.$$

Par conséquent,

$$\det J_f(x, y) = 4(x^2 + y^2) = 0 \iff (x, y) = (0, 0).$$

Le théorème d'inversion locale montre que la fonction f est un $\mathcal{C}^\infty$ -difféomorphisme local au voisinage de tout point $(x, y) \neq (0, 0)$.

De plus, comme $f(1, 1) = f(-1, -1)$, on observe que f n'est pas injective sur $\mathbb{R}^2 \setminus \{(0, 0)\}$. Par conséquent, f ne peut pas être un $\mathcal{C}^1$ -difféomorphisme global sur $\mathbb{R}^2 \setminus \{(0, 0)\}$.

Exercice 10.9 (Folium de Descartes)

Leçons : 214, 267. Référence : [5].

On considère l'ensemble

$$\mathcal{C} = \{(x, y) \in \mathbb{R}^2 \mid x^3 + y^3 - 3xy = 0\}.$$

En quels points cette équation définit-elle localement y comme fonction de classe $\mathcal{C}^1$ de x ? Si oui, calculer la dérivée de la fonction implicite et écrire l'équation de la tangente à $\mathcal{C}$.

Corrigé.

On pose $f(x, y) = x^3 + y^3 - 3xy$ pour tout $(x, y) \in \mathbb{R}^2$. Il s'agit d'une fonction de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^2$, puisque c'est une fonction polynomiale. De plus, pour tout $(x, y) \in \mathbb{R}^2$, on a

$$\frac{\partial f}{\partial y}(x, y) = 3(y^2 - x),$$

de sorte que $\frac{\partial f}{\partial y}(x, y) \neq 0$ si et seulement si $y^2 \neq x$. Cependant, un point (y^2, y) appartient à $\mathcal{C}$ si et seulement si $f(y^2, y) = 0$, c'est-à-dire $y^6 - 2y^3 = 0$, ce qui est encore équivalent à $y = 0$ ou $y = 2^{1/3}$. Le théorème des fonctions implicites s'applique en tout point $(a, b) \in \mathcal{C} \setminus \{(0, 0), (2^{2/3}, 2^{1/3})\}$.

Il existe donc un intervalle ouvert I contenant a , un intervalle ouvert J contenant b et une fonction $\varphi : I \rightarrow J$ de classe $\mathcal{C}^\infty$ telle que

$$(x, y) \in (I \times J) \cap \mathcal{C} \iff x \in I \text{ et } y = \varphi(x).$$

On a alors $f(x, \varphi(x)) = 0$ pour tout $x \in I$, puis en dérivant,

$$\frac{\partial f}{\partial x}(x, \varphi(x)) + \frac{\partial f}{\partial y}(x, \varphi(x))\varphi'(x) = 0, \implies 3(x^2 - \varphi(x)) + 3(\varphi(x)^2 - x)\varphi'(x) = 0.$$

Pour $x = a$, on a

$$3(a^2 - b) + 3(b^2 - a)\varphi'(a) = 0 \implies \varphi'(a) = \frac{b - a^2}{b^2 - a}.$$

L'équation de la tangente à $\mathcal{C}$ en (a, b) est donc

$$y = \varphi(a) + \varphi'(a)(x - a) = b + \frac{b - a^2}{b^2 - a}(x - a).$$

Aux deux points exclus, la courbe n'est pas localement le graphe d'une fonction de classe $\mathcal{C}^1$ de x . Au point $(a, b) = (2^{2/3}, 2^{1/3})$, on a $f_y(a, b) = 0$ et $f_x(a, b) = 3b \neq 0$: l'identité $f(x, \varphi(x)) = 0$ donnerait donc $f_x(a, b) + f_y(a, b)\varphi'(a) = 0$, une contradiction. Le théorème des fonctions implicites permet en revanche d'écrire x en fonction de y ; la tangente est la droite verticale $x = 2^{2/3}$.

À l'origine, considérons l'arc

$$\gamma(t) = \left(\frac{3t^2}{1+t^3}, \frac{3t}{1+t^3} \right), \quad |t| < \frac{1}{2}.$$

On vérifie que $\gamma(t) \in \mathcal{C}$ et que $\gamma'(0) = (0, 3)$. Si la courbe était localement un graphe $y = \varphi(x)$ de classe $\mathcal{C}^1$, la dérivation de $\gamma_2(t) = \varphi(\gamma_1(t))$ en 0 donnerait $3 = 0$. En échangeant les deux coordonnées, on obtient aussi un arc de tangente horizontale : l'origine est un point double dont les deux tangentes sont les axes de coordonnées.

Exercice 10.10 (Un point critique sans extremum)

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction définie par $f(x, y) = x^2 - y^2$ pour tout $(x, y) \in \mathbb{R}^2$. Montrer que $(0, 0)$ est un point critique de f , mais que ce n'est pas un extremum local.

Corrigé.

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction définie par $f(x, y) = x^2 - y^2$ pour tout $(x, y) \in \mathbb{R}^2$. Il s'agit d'une fonction de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^2$ puisque c'est une fonction polynomiale. Cherchons d'abord les points critiques de f . Pour ce faire, on résout l'équation $\nabla f(x, y) = 0$, soit

$$\frac{\partial f}{\partial x}(x, y) = 2x = 0 \quad \text{et} \quad \frac{\partial f}{\partial y}(x, y) = -2y = 0,$$

ce qui donne $(x, y) = (0, 0)$. La fonction f admet donc un unique point critique qui est $(0, 0)$.

On constate que f n'admet ni un minimum local ni un maximum local sur $\mathbb{R}^2$ en $(0, 0)$. En effet, si U est un ouvert de $\mathbb{R}^2$ contenant $(0, 0)$, alors la boule fermée $\bar{B}((0, 0), \varepsilon)$ est incluse dans U pour $\varepsilon > 0$ assez petit. Dès lors,

$$(\varepsilon, 0) \in U \implies f(\varepsilon, 0) = \varepsilon^2 > 0 = f(0, 0),$$

et

$$(0, \varepsilon) \in U \implies f(0, \varepsilon) = -\varepsilon^2 < 0 = f(0, 0).$$

Cela montre respectivement que $(0, 0)$ n'est pas un point de maximum local, ni un point de minimum local.

Exercice 10.11 (Extrema sur un disque)

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction définie par $f(x, y) = x^3 - 3x(1 + y^2)$.

1. Étudier les extrema locaux de f .
2. Soit $D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$. Montrer que f admet un maximum M et un minimum m sur D .
3. Soit $(x, y) \in D$. Montrer que si $f(x, y) = M$ ou $f(x, y) = m$, alors $x^2 + y^2 = 1$.
4. Étudier la fonction $t \mapsto f(\cos t, \sin t)$. En déduire les valeurs de M et m .

Corrigé.

1. La fonction f est polynomiale, et donc de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^2$. Pour tout $(x, y) \in \mathbb{R}^2$,

$$\frac{\partial f}{\partial x}(x, y) = 3x^2 - 3(1 + y^2) \quad \text{et} \quad \frac{\partial f}{\partial y}(x, y) = -6xy,$$

de sorte que $\nabla f(x, y) = 0$ si et seulement si

$$\begin{cases} x^2 - y^2 = 1 \\ x = 0 \text{ ou } y = 0 \end{cases} \iff (x, y) = (1, 0) \quad \text{ou} \quad (x, y) = (-1, 0).$$

La fonction f admet donc deux points critiques. Pour déterminer leur nature, calculons la matrice hessienne. Pour tout $(x, y) \in \mathbb{R}^2$,

$$D^2 f(x, y) = \begin{pmatrix} 6x & -6y \\ -6y & -6x \end{pmatrix}.$$

En particulier,

$$D^2 f(\pm 1, 0) = \begin{pmatrix} \pm 6 & 0 \\ 0 & \mp 6 \end{pmatrix}.$$

Dans les deux cas, comme les valeurs propres de $D^2 f(\pm 1, 0)$ ne sont pas de même signe, on en déduit que $(\pm 1, 0)$ sont des points selles.

2. L'ensemble D est compact car borné et fermé, et la fonction f est continue. Par conséquent, f atteint ses bornes sur D .
3. Soit $(x, y) \in D$ tel que $f(x, y) = m$ ou $f(x, y) = M$. Si (x, y) est dans l'intérieur $\overset{\circ}{D}$, alors $\nabla f(x, y) = \vec{0}$. Cela implique d'après la première question que $(x, y) = (\pm 1, 0) \notin \overset{\circ}{D}$, ce qui est impossible. Par conséquent, f atteint ses bornes sur D sur le bord.
4. Pour étudier les variations de f sur le bord ∂D , on paramétrise ∂D par l'application $t \in [0, 2\pi[\mapsto (\cos t, \sin t)$. On pose pour tout $t \in \mathbb{R}$,

$$g(t) = f(\cos t, \sin t) = \cos^3 t - 3(\cos t)(1 + \sin^2 t).$$

Clairement, la fonction g est paire et 2π -périodique. Il suffit donc de l'étudier sur l'intervalle $]0, \pi[$. Pour tout $t \in]0, \pi[$, on a

$$g'(t) = -9 \cos^2 t \sin t + 3 \sin t(1 + \sin^2 t) = 6 \sin t(1 - 2 \cos^2 t) = -6 \sin t \cos(2t).$$

On peut alors dresser le tableau de variations de g :

t	0	$\pi/4$	$3\pi/4$	π			
$\sin t$	0	+	+	0			
$\cos 2t$		+	0	-	0	+	
$g'(t)$	0	-	0	+	0	-	0
g	-2	$\longrightarrow$	$-2\sqrt{2}$	$\longrightarrow$	$2\sqrt{2}$	$\longrightarrow$	2

On déduit du tableau et de la parité de g que $m = -2\sqrt{2}$ est atteint pour $t = \pm \frac{\pi}{4}$, et donc $(x, y) = \left(\frac{\sqrt{2}}{2}, \pm \frac{\sqrt{2}}{2}\right)$, et que $M = 2\sqrt{2}$ est atteint pour $t = \pm \frac{3\pi}{4}$, et donc $(x, y) = \left(-\frac{\sqrt{2}}{2}, \pm \frac{\sqrt{2}}{2}\right)$.

Exercice 10.12 (Principe du maximum)

Référence : [6].

Soit $\Omega \subset \mathbb{R}^n$ un ouvert borné non vide et $u : \overline{\Omega} \rightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^2$ sur l'ouvert Ω et continue sur le fermé $\overline{\Omega}$ telle que

$$\forall x \in \Omega \quad \Delta u(x) = 0.$$

Pour tout $\varepsilon > 0$, posons $u_\varepsilon(x) = u(x) + \varepsilon e^{x_1}$.

1. Pourquoi les fonctions u et u_ε atteignent-elles leur maximum sur $\overline{\Omega}$?
2. Montrer que $\Delta u_\varepsilon(x) > 0$ pour tout $x \in \Omega$.
3. Soit $x_\varepsilon \in \overline{\Omega}$ un point de maximum de u_ε sur $\overline{\Omega}$. Montrer que $x_\varepsilon \in \partial\Omega$, où $\partial\Omega = \overline{\Omega} \setminus \Omega$ est le *bord* de Ω .
4. En déduire que le maximum de u sur $\overline{\Omega}$ est atteint sur le bord $\partial\Omega$.
5. Montrer que si $u_1, u_2 \in \mathcal{C}(\overline{\Omega}) \cap \mathcal{C}^2(\Omega)$ sont deux fonctions harmoniques (*i.e.* de laplacien nul) sur Ω telles que $u_1 = u_2$ sur $\partial\Omega$, alors $u_1 = u_2$ sur $\overline{\Omega}$.

Corrigé.

1. Les fonctions u et u_ε sont continues sur le compact $\overline{\Omega}$ (fermé borné dans $\mathbb{R}^n$). Elles atteignent bien leur maximum sur $\overline{\Omega}$.
2. Pour tout $x \in \Omega$, on a $\Delta u_\varepsilon(x) = \Delta u(x) + \varepsilon e^{x_1} = \varepsilon e^{x_1} > 0$.
3. Si $x_\varepsilon \in \Omega$, alors toutes les valeurs propres de la matrice hessienne $D^2 u_\varepsilon(x_\varepsilon)$ sont négatives ou nulles. En particulier, on a que $\text{Tr}(D^2 u_\varepsilon(x_\varepsilon)) \leq 0$. Cela contredit le fait que $\text{Tr}(D^2 u_\varepsilon(x_\varepsilon)) = \Delta u_\varepsilon(x_\varepsilon) > 0$. On en déduit que $x_\varepsilon \in \partial\Omega$.
4. Comme $\partial\Omega$ est compact, quitte à extraire une sous-suite, on a $x_\varepsilon \rightarrow \bar{x} \in \partial\Omega$. Montrons que $\bar{x}$ est un point de maximum de u sur $\overline{\Omega}$. En effet, pour tout $y \in \overline{\Omega}$, on a

$$u(x_\varepsilon) + \varepsilon e^{(x_\varepsilon)_1} = u_\varepsilon(x_\varepsilon) \geq u_\varepsilon(y) \geq u(y).$$

Par passage à la limite quand $\varepsilon \rightarrow 0$, on obtient $u(\bar{x}) \geq u(y)$ ce qui montre que $\bar{x} \in \partial\Omega$ est un point de maximum de u sur $\overline{\Omega}$.

5. On a $u_1 - u_2 \in \mathcal{C}(\overline{\Omega}) \cap \mathcal{C}^2(\Omega)$ et $\Delta(u_1 - u_2) = 0$ sur Ω . La question précédente montre qu'il existe $\bar{x} \in \partial\Omega$ tel que

$$\forall x \in \overline{\Omega} \quad u_1(x) - u_2(x) \leq u_1(\bar{x}) - u_2(\bar{x}) = 0.$$

On en déduit que $u_1 \leq u_2$ sur $\overline{\Omega}$. L'autre inégalité s'obtient en inversant les rôles de u_1 et u_2 .

Exercice 10.13 (Extrema sur le cercle et multiplicateurs de Lagrange)

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^1$. On s'intéresse à la recherche d'extrema locaux de f sur le cercle unité $S = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$.

1. On considère une paramétrisation de S donnée par $\gamma : t \in \mathbb{R} \mapsto (\cos t, \sin t)$. Rappeler l'expression du vecteur tangent à S au point $(x, y) = (\cos t, \sin t)$.
2. Pour tout $t \in \mathbb{R}$, on pose $g(t) = f(\cos t, \sin t)$. Montrer que g est de classe $\mathcal{C}^1$ sur $\mathbb{R}$, et que si $(x_0, y_0) = (\cos t_0, \sin t_0)$ est un extremum local de f sur S , alors $g'(t_0) = 0$.

3. En déduire que

$$-\frac{\partial f}{\partial x}(\cos t_0, \sin t_0) \sin t_0 + \frac{\partial f}{\partial y}(\cos t_0, \sin t_0) \cos t_0 = 0.$$

4. Établir alors que le gradient $\nabla f(x_0, y_0)$ est orthogonal au vecteur tangent à S en (x_0, y_0) . En déduire l'existence d'un réel λ tel que

$$\nabla f(x_0, y_0) = \lambda \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

Corrigé.

1. Le vecteur tangent est donné par $\gamma'(t) = (-\sin t, \cos t)$ pour tout $t \in \mathbb{R}$.
2. La fonction g est de classe $\mathcal{C}^1$ comme composée de fonctions de classe $\mathcal{C}^1$. Si $(x_0, y_0) = (\cos t_0, \sin t_0)$ est un extremum local de f sur S , alors t_0 est un extremum local de g sur $\mathbb{R}$ et on a alors la condition nécessaire $g'(t_0) = 0$.
3. Pour tout $t \in \mathbb{R}$, on a $g'(t) = \nabla f(\gamma(t)) \cdot \gamma'(t)$ et on remplace dans l'expression précédente.
4. Les vecteurs $\nabla f(\cos t_0, \sin t_0)$ et $(-\sin t_0, \cos t_0)$ sont orthogonaux. Par conséquent $\nabla f(x_0, y_0)$ est orthogonal au vecteur tangent à S en (x_0, y_0) . Autrement dit, $\nabla f(x_0, y_0)$ est colinéaire au vecteur normal $(\cos t_0, \sin t_0) = (x_0, y_0)$.

Exercice 10.14 (Inversion globale)

Leçon : 214. Référence : [5].

On considère un réel $k > 0$ et une application $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ de classe $\mathcal{C}^1$ et k -dilatante, c'est-à-dire telle que

$$\forall x, y \in \mathbb{R}^n \quad \|f(x) - f(y)\| \geq k \|x - y\|.$$

1. Montrer que l'application f est injective, et que l'image $f(\mathbb{R}^n)$ est fermée dans $\mathbb{R}^n$.
2. Montrer que la différentielle $df(x)$ est inversible pour tout x .
3. Montrer que l'image $f(\mathbb{R}^n)$ est ouverte dans $\mathbb{R}^n$, puis que l'application f est un $\mathcal{C}^1$ -difféomorphisme global de $\mathbb{R}^n$ dans lui-même.
4. Établir un lien avec le théorème de Hadamard-Lévy.

Corrigé.

1. Si $f(x) = f(y)$, on a $0 = \|f(x) - f(y)\| \geq k \|x - y\| \geq 0$, donc $x = y$. L'injectivité de f s'en déduit. Pour montrer que l'image de f est fermée, considérons une suite de points $f(x_n)$ qui converge vers une limite ℓ . C'est en particulier une suite de Cauchy, et l'inégalité vérifiée par f implique que la suite des points x_n est aussi de Cauchy. Par complétude de $\mathbb{R}^n$, on en déduit que x_n tend vers une limite x . Par continuité de f , on a ensuite $\ell = f(x)$, ce qui montre que ℓ est dans l'image de f .
2. On sait que f est différentiable, et par définition, sa différentielle en un point x quelconque vérifie pour tout h ,

$$\|df(x)(h)\| = \lim_{t \rightarrow 0} \frac{1}{|t|} \|f(x + th) - f(x)\| \geq k \|h\|.$$

La différentielle de f est alors injective, donc inversible, en tout point. L'inverse a pour norme au plus $1/k$.

3. Pour tout point $y_0 = f(x_0)$ appartenant à l'image, le théorème d'inversion locale montre l'existence d'un voisinage ouvert V de x_0 et d'un voisinage ouvert W de y_0 tels que f soit un difféomorphisme de V sur W . En particulier, $f(\mathbb{R}^n)$ contient W , et donc un voisinage ouvert de y_0 . Il en résulte que l'image de f est ouverte. On a donc montré que l'image de f est à la fois fermée et ouverte dans l'ensemble connexe $\mathbb{R}^n$. Elle est donc égale à $\mathbb{R}^n$ tout entier. Ainsi, l'application f est surjective. On en conclut que f est une application bijective de $\mathbb{R}^n$ dans lui-même, de classe $\mathcal{C}^1$, et de réciproque de classe $\mathcal{C}^1$ également, par inversion locale. Finalement, f est un $\mathcal{C}^1$ -difféomorphisme global de $\mathbb{R}^n$ dans lui-même.
4. On a vu que l'inégalité portant sur f implique que la différentielle df est inversible en tout point. De plus, l'application f est propre car pour tout x , l'inégalité triangulaire donne

$$\|f(x)\| \geq \|f(x) - f(0)\| - \|f(0)\| \geq k \|x\| - \|f(0)\|,$$

ce qui montre que l'image réciproque d'un ensemble borné est bornée. Les hypothèses sur f impliquent donc celles du théorème de Hadamard-Lévy, et ce résultat permet de retrouver que f est un $\mathcal{C}^1$ -difféomorphisme global de $\mathbb{R}^n$ dans lui-même.

Exercice 10.15 (Théorème du rang constant)

Leçon : 214. Référence : [5].

Soit U un ouvert de $\mathbb{R}^n$ et $f : U \rightarrow \mathbb{R}^p$ une application de classe $\mathcal{C}^1$. On suppose que le rang de la différentielle df est constant égal à r sur U . Le but de l'exercice est de montrer l'existence d'un $\mathcal{C}^1$ -difféomorphisme local $x \mapsto X$ au voisinage de $a \in U$ quelconque, et d'un $\mathcal{C}^1$ -difféomorphisme local $y \mapsto Y$ au voisinage de $f(a)$, qui transforment f en l'application linéaire $A : \mathbb{R}^n \rightarrow \mathbb{R}^p$ de rang r donnée par

$$A : X = (X_1, \dots, X_r, X_{r+1}, \dots, X_n) \mapsto Y = (X_1, \dots, X_r, 0, \dots, 0).$$

1. Montrer qu'on peut supposer $a = 0$, $f(a) = 0$ et $df(a) = A$.
2. Vérifier qu'on obtient un $\mathcal{C}^1$ -difféomorphisme local $x \mapsto X$ sur $\mathbb{R}^n$ au voisinage de 0 en posant

$$X_1 = f_1(x_1, \dots, x_n), \quad \dots \quad X_r = f_r(x_1, \dots, x_n), \quad X_{r+1} = x_{r+1}, \quad \dots \quad X_n = x_n.$$

3. L'application $f : x \mapsto y$ devient $\varphi : X \mapsto y$ quand on remplace les coordonnées x par X au départ. Montrer que l'application φ ne dépend que des coordonnées $X_1, \dots, X_r$, et expliciter ses r premières composantes.
4. Montrer qu'on obtient un $\mathcal{C}^1$ -difféomorphisme local $Y \mapsto y$ sur $\mathbb{R}^p$ au voisinage de 0 en posant

$$y_1 = Y_1, \quad \dots \quad y_r = Y_r, \quad y_{r+1} = \varphi_{r+1}(Y_1, \dots, Y_r) + Y_{r+1}, \quad \dots \quad y_p = \varphi_p(Y_1, \dots, Y_r) + Y_p,$$

et que les $\mathcal{C}^1$ -difféomorphismes locaux $x \mapsto X$ et $y \mapsto Y$ ainsi construits conviennent.

Corrigé.

1. La fonction définie par $g(x) = f(a + x) - f(a)$ vérifie $g(0) = 0$, est de classe $\mathcal{C}^1$ sur un voisinage de 0, et sa différentielle dans un tel voisinage vérifie $dg(x) = df(a + x)$, donc est de rang constant égal à r . En particulier, la différentielle $dg(0)$ admet dans certaines bases $(e_1, \dots, e_n)$ et $(f_1, \dots, f_p)$, respectivement de $\mathbb{R}^n$ et de $\mathbb{R}^p$, pour matrice $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$. Il existe donc des applications linéaires u et v de changements de bases, respectivement de $\mathbb{R}^n$ et de $\mathbb{R}^p$, telles que $v^{-1}dg(0)u$ admette pour matrice dans les bases canoniques la matrice précédente. On en déduit que l'application $h = v^{-1} \circ g \circ u$ vérifie $h(0) = 0$, est de classe $\mathcal{C}^1$ sur un voisinage de 0, avec une différentielle $dh(x) = v^{-1}dg(u(x))u$ de rang constant égal à r , et telle que la différentielle $dh(0) = v^{-1}dg(0)u$ soit l'application A dont la matrice dans les bases canoniques est la matrice précédente. En notant f au lieu de h , on peut donc supposer que $a = 0$, $f(a) = 0$ et $df(a) = A$.
2. La matrice jacobienne de l'application $x \mapsto X$ en 0 est $\begin{pmatrix} I_r & 0 \\ 0 & I_{n-r} \end{pmatrix}$, donc est inversible. Cette application est de plus de classe $\mathcal{C}^1$ au voisinage de l'origine. Le théorème d'inversion locale implique qu'il s'agit d'un $\mathcal{C}^1$ -difféomorphisme local au voisinage de 0 dans $\mathbb{R}^n$.

3. Dans les coordonnées X , l'application f devient φ et s'écrit

$$y_1 = X_1, \quad \dots \quad y_r = X_r, \quad y_{r+1} = \varphi_{r+1}(X_1, \dots, X_n), \quad \dots \quad y_p = \varphi_p(X_1, \dots, X_n),$$

avec $\varphi_{r+1}, \dots, \varphi_p$ des applications de classe $\mathcal{C}^1$ au voisinage de 0, obtenues en composant f et $X \mapsto x$. La matrice jacobienne de φ est donc $\begin{pmatrix} I_r & 0 \\ * & (\partial\varphi_i/\partial X_j)(X) \end{pmatrix}$. Son rang au voisinage de l'origine est le même que celui de la différentielle df , c'est-à-dire constant égal à r . Ainsi, les dérivées partielles $\frac{\partial\varphi_i}{\partial X_j}$, avec $i, j > r$ sont nulles sur un voisinage de l'origine. Quitte à restreindre ce voisinage à un pavé ouvert centré en 0, on en déduit que les fonctions φ_i ne dépendent que des coordonnées $X_1, \dots, X_r$.

4. Les relations de l'énoncé impliquent directement

$$Y_1 = y_1, \quad \dots \quad Y_r = y_r, \quad Y_{r+1} = y_{r+1} - \varphi_{r+1}(y_1, \dots, y_r), \quad \dots \quad Y_p = y_p - \varphi_p(y_1, \dots, y_r).$$

On a donc $y = f(x)$ si et seulement si $Y_1 = X_1, \dots, Y_r = X_r$ et $Y_{r+1} = \dots = Y_p = 0$, c'est-à-dire $Y = A(X)$.

Exercice 10.16 (Asymptotique d'une équation du troisième degré)

Leçon : 214. Référence : [5].

On considère $f(x, \varepsilon) = (x - a)(b - x) + \varepsilon x^3$, avec $a < b$.

- Montrer que pour $\varepsilon > 0$ assez petit, l'équation $f(x, \varepsilon) = 0$ admet trois racines distinctes $x_1(\varepsilon) < x_2(\varepsilon) < x_3(\varepsilon)$. Donner un développement asymptotique de ces racines quand $\varepsilon \rightarrow 0$ avec pour reste $O(\varepsilon^2)$.
- En déduire un développement asymptotique de l'intégrale elliptique

$$I(\varepsilon) = \int_{x_1(\varepsilon)}^{x_2(\varepsilon)} \frac{dx}{\sqrt{f(x, \varepsilon)}}$$

quand $\varepsilon \rightarrow 0$ avec pour reste $O(\varepsilon^2)$.

Corrigé.

- On a $f(a, 0) = 0$ et $\frac{\partial f}{\partial x}(x, \varepsilon) = a + b - 2x + 3\varepsilon x^2$, donc $\frac{\partial f}{\partial x}(a, 0) = b - a > 0$. Le théorème des fonctions implicites montre qu'il existe un voisinage V_1 de 0, un voisinage W_1 de a , et une application $x_1 : V_1 \rightarrow W_1$ vérifiant

$$\begin{cases} f(x, \varepsilon) = 0 \\ (x, \varepsilon) \in W_1 \times V_1 \end{cases} \iff \begin{cases} x = x_1(\varepsilon) \\ \varepsilon \in V_1. \end{cases}$$

De plus, l'application x_1 est de classe $\mathcal{C}^2$. On a $x_1(0) = a$ et, en dérivant $f(x_1(\varepsilon), \varepsilon) = 0$, on obtient la relation $x_1'(\varepsilon) \frac{\partial f}{\partial x}(x_1(\varepsilon), \varepsilon) + \frac{\partial f}{\partial \varepsilon}(x_1(\varepsilon), \varepsilon) = 0$, puis $(b - a)x_1'(0) + a^3 = 0$. Il en ressort que, quand $\varepsilon \rightarrow 0$,

$$x_1(\varepsilon) = a - \frac{a^3}{b - a} \varepsilon + O(\varepsilon^2).$$

De même, on obtient un voisinage V_2 de 0, un voisinage W_2 de b , et une application $x_2 : V_2 \rightarrow W_2$ vérifiant

$$\begin{cases} f(x, \varepsilon) = 0 \\ (x, \varepsilon) \in W_2 \times V_2 \end{cases} \iff \begin{cases} x = x_2(\varepsilon) \\ \varepsilon \in V_2, \end{cases}$$

puis quand $\varepsilon \rightarrow 0$,

$$x_2(\varepsilon) = b + \frac{b^3}{b - a} \varepsilon + O(\varepsilon^2).$$

Enfin, en développant, on a $f(x, \varepsilon) = \varepsilon x^3 - x^2 + (a + b)x - ab$. Les relations coefficients racines impliquent que $x_1(\varepsilon) + x_2(\varepsilon) + x_3(\varepsilon) = 1/\varepsilon$, et finalement quand $\varepsilon \rightarrow 0$,

$$x_3(\varepsilon) = \frac{1}{\varepsilon} - (a + b) - (a^2 + ab + b^2)\varepsilon + O(\varepsilon^2).$$

2. Pour simplifier, on omet d'écrire la dépendance en ε dans les racines x_1 , x_2 et x_3 . En factorisant et en exprimant x_3 comme $1/\varepsilon - x_1 - x_2$, on a

$$f(x, \varepsilon) = \varepsilon(x - x_1)(x_2 - x)(x_3 - x) = (x - x_1)(x_2 - x)(1 - \varepsilon(x + x_1 + x_2)),$$

ce qui permet d'exprimer l'intégrale elliptique sous la forme

$$I(\varepsilon) = \int_{x_1}^{x_2} \frac{(1 - \varepsilon(x + x_1 + x_2))^{-1/2}}{\sqrt{(x - x_1)(x_2 - x)}} dx.$$

On remarque que cette intégrale converge pour ε assez petit car $x_1 < x_2 < x_3$. On effectue le changement de variable $x = u + v \sin t$ avec $u = (x_1 + x_2)/2$ et $v = (x_2 - x_1)/2$, ce qui donne

$$I(\varepsilon) = \int_{-\pi/2}^{\pi/2} (1 - \varepsilon(3u + v \sin t))^{-1/2} dt.$$

On a $3u + v \sin t = \frac{3}{2}(a + b) + \frac{b-a}{2} \sin t + r(t, \varepsilon)$, avec $|r(t, \varepsilon)| \leq C\varepsilon$ uniformément en t , pour ε petit. De plus, $(1-y)^{-1/2} = 1 + \frac{y}{2} + \frac{y^2}{2} + \alpha \frac{y^3}{2}$ avec $|\alpha| \leq 3\sqrt{2}$ pour $|y| \leq 1/2$. Ainsi, $(1 - \varepsilon(3u + v \sin t))^{-1/2} = 1 + \frac{\varepsilon}{2}(3u + v \sin t) + s(t, \varepsilon)$, avec $|s(t, \varepsilon)| \leq C'\varepsilon^2$ uniformément en t , pour ε petit. Par uniformité, on peut intégrer les développements limités, ce qui donne quand $\varepsilon \rightarrow 0$,

$$I(\varepsilon) = \int_{-\pi/2}^{\pi/2} \left(1 + \frac{\varepsilon}{2} \left(\frac{3}{2}(a + b) + \frac{b-a}{2} \sin t \right) \right) dt + O(\varepsilon^2) = \pi + \frac{3\pi}{4}(a + b)\varepsilon + O(\varepsilon^2).$$

Exercice 10.17 (Ellipsoïde de John-Loewner)

Leçon : 219. Références : [10, 12].

On munit $\mathbb{R}^d$ de sa structure euclidienne canonique. Pour toute forme quadratique positive q sur $\mathbb{R}^d$, on note

$$E_q = \{x \in \mathbb{R}^d \mid q(x) \leq 1\}.$$

On appelle *ellipsoïde* tout ensemble de la forme E_q , où q est une forme quadratique définie positive.

1. Montrer que tous les ensembles E_q sont convexes.
2. Montrer que pour toutes formes quadratiques q et q' , avec q définie positive et q' positive, si $E_q \subset E_{q'}$, alors $q' \leq q$. En déduire que chaque ellipsoïde est issu d'une unique forme quadratique définie positive.
3. Soit q une forme quadratique définie positive, et soit S sa matrice dans la base canonique. On note V_q le volume de l'ellipsoïde E_q , et κ le volume de la boule unité (pour la mesure de Lebesgue). Montrer que

$$V_q = \frac{\kappa}{\sqrt{\det S}}.$$

Interpréter cette formule dans le cas de la dimension $d = 2$.

4. Soit K un compact de $\mathbb{R}^d$ d'intérieur non vide. Montrer que l'ensemble C_K des matrices symétriques positives S telles que ${}^tXSX \leq 1$ pour tout vecteur-colonne X appartenant à K est un compact non vide de l'espace vectoriel des matrices symétriques muni de la norme subordonnée à la norme euclidienne.
5. Montrer que le déterminant atteint son maximum sur une matrice de C_K , qui est symétrique définie positive.
6. Montrer que l'ensemble C_K est convexe, puis que pour tout $0 < \lambda < 1$ et toutes matrices symétriques définies positives S_1 et S_2 distinctes,

$$\det(\lambda S_1 + (1 - \lambda)S_2) > \det(S_1)^\lambda \det(S_2)^{1-\lambda},$$

ce qui indique que le déterminant est strictement log-concave.

7. En déduire l'unicité de la matrice symétrique définie positive sur laquelle le déterminant atteint son maximum.

8. Dédurre de ce qui précède le *théorème de John-Loewner* : pour tout compact K de $\mathbb{R}^d$ d'intérieur non vide, il existe un unique ellipsoïde centré en l'origine et de volume minimal qui contienne K . On appelle cet ellipsoïde l'*ellipsoïde de John-Loewner de K* .
9. Plus généralement, pour tout point c de $\mathbb{R}^d$ et toute forme quadratique positive q sur $\mathbb{R}^d$, on note

$$E_{q,c} = \{x \in \mathbb{R}^d \mid q(x - c) \leq 1\},$$

et on appelle *ellipsoïde centré en c* tout ensemble de la forme $E_{q,c}$, où q est une forme quadratique définie positive. Montrer que $E_{q,c}$ se déduit de E_q par translation. En déduire que pour tout compact K de $\mathbb{R}^d$ d'intérieur non vide, il existe un unique ellipsoïde centré en c et de volume minimal qui contienne K . On appelle cet ellipsoïde l'*ellipsoïde de John-Loewner de K centré en c* .

En complément, on peut déduire de cet exercice que pour tout compact non vide K de $\mathbb{R}^d$ d'intérieur non vide, il existe un unique ellipsoïde de volume minimal qui contienne K . Cela signifie que l'on peut aussi optimiser selon le centre c de l'ellipsoïde. Enfin, une question classique à laquelle il faut avoir réfléchi est la détermination de cet ellipsoïde dans le cas d'une figure géométrique simple du plan, comme par exemple un triangle. Une autre application est l'exercice ci-dessous.

Corrigé.

1. Soit q une forme quadratique positive. L'inégalité de Minkowski implique que pour $\lambda \in [0, 1]$ et $x, y \in \mathbb{R}^d$,

$$\sqrt{q(\lambda x + (1 - \lambda)y)} \leq \sqrt{q(\lambda x)} + \sqrt{q((1 - \lambda)y)} = \lambda \sqrt{q(x)} + (1 - \lambda) \sqrt{q(y)}.$$

Cela conduit directement à la convexité de l'ensemble E_q .

2. Pour tout $x \neq 0$, le point $q(x)^{-1/2}x$ appartient à E_q , donc à $E_{q'}$ quand $E_q \subset E_{q'}$. D'où $q'(q(x)^{-1/2}x) \leq 1$, ce qui donne $q'(x) \leq q(x)$. Par conséquent, si $E_q \subset E_{q'}$, on a $q' \leq q$. Chaque ellipsoïde est donc issu d'une unique forme quadratique.
3. D'après le théorème spectral, il existe une base $(e_1, \dots, e_d)$ orthonormée pour le produit scalaire canonique, et orthogonale pour la forme polaire associée à q . La matrice de q dans cette base est diagonale, avec des coefficients diagonaux notés $\mu_i > 0$. On note P la matrice dont les colonnes sont les vecteurs $e_1, \dots, e_d$ écrits dans la base canonique. La matrice P est orthogonale de déterminant ± 1 . De plus, l'application définie par $\Phi(x) = P \text{diag}(\mu_1^{-1/2}, \dots, \mu_d^{-1/2})x$, pour tout vecteur-colonne x de $\mathbb{R}^d$, est un $\mathcal{C}^1$ -difféomorphisme de $\mathbb{R}^d$ dont la valeur absolue du jacobien est $(\mu_1 \cdots \mu_d)^{-1/2}$ qui vérifie

$$\|x\| \leq 1 \iff \sum_{i=1}^d \mu_i (\mu_i^{-1/2} x_i)^2 \leq 1 \iff q(\Phi(x)) \leq 1.$$

En effet, $\Phi(x)$ donne les coordonnées dans la base canonique du vecteur de coordonnées $(\mu_1^{-1/2}x_1, \dots, \mu_d^{-1/2}x_d)$ dans la base $(e_1, \dots, e_d)$. Ainsi, Φ met en bijection la boule unité fermée B et l'ellipsoïde E_q , de sorte que

$$V_q = \int_{E_q} dy = \int_B (\mu_1 \cdots \mu_d)^{-1/2} dx = \kappa (\mu_1 \cdots \mu_d)^{-1/2} = \frac{\kappa}{\sqrt{\det S}}.$$

La dernière égalité provient du fait que ${}^tPSP = \text{diag}(\mu_1, \dots, \mu_d)$, et donc $\det S = \mu_1 \cdots \mu_d$.

En dimension $d = 2$, l'ellipsoïde est un disque elliptique dont les axes ont pour longueur $2\mu_1^{-1/2}$ et $2\mu_2^{-1/2}$, et on retrouve bien que son aire est $\pi \mu_1^{-1/2} \mu_2^{-1/2} = \pi ab$, si a et b désignent les longueurs des demi-axes.

4. Montrons que C_K est fermé. Considérons une suite de matrices S_n dans C_K qui converge vers une matrice symétrique S . D'après l'inégalité de Cauchy-Schwarz, on a pour tout X dans $\mathbb{R}^d$,

$$\|{}^tXS_nX - {}^tXSX\| = \|{}^tX(S_n - S)X\| \leq \|(S_n - S)X\| \|X\| \leq \|S_n - S\| \|X\|^2 \xrightarrow{n \rightarrow \infty} 0.$$

Ainsi, les conditions ${}^tXS_nX \geq 0$ et ${}^tXS_nX \leq 1$ passent à la limite. On en déduit que S appartient à C_K .

Montrons que C_K est borné. Soit $S \in C_K$, de forme quadratique associée notée q . On a alors $K \subset E_q$ et, comme il est d'intérieur non vide, K contient une boule fermée $\bar{B}(x, r)$. L'ensemble E_q contient donc la boule $\bar{B}(x, r)$,

et par symétrie la boule $\bar{B}(-x, r)$, puis par convexité la boule $\bar{B}(0, r)$. Cependant, cette dernière boule est l'ellipsoïde associé à la forme quadratique définie positive $r^{-2}\|\cdot\|^2$. On a donc $q \leq r^{-2}\|\cdot\|^2$. D'après le théorème spectral, on a alors

$$\|S\| = \sup_{\|X\|=1} \|XSX\| = \sup_{\|X\|=1} q(X) \leq \sup_{\|X\|=1} r^{-2}\|X\|^2 = r^{-2}.$$

Ainsi, le fermé C_K est borné, donc est compact.

Enfin, comme l'ensemble K est compact, il est borné par un réel $m > 0$, et la matrice $m^{-2}I_d$ appartient à C_K . L'ensemble C_K est donc non vide.

5. L'application déterminant est continue sur le compact non vide C_K , donc est majorée et atteint son maximum sur C_K , en une matrice S . Comme $m^{-2}I_d$ appartient aussi à C_K , on a $\det S \geq \det(m^{-2}I_d) = m^{-2d} > 0$. Ainsi, la matrice symétrique S est définie positive.
6. Considérons $\lambda \in [0, 1]$ et deux matrices S_1 et S_2 dans C_K . La matrice $\lambda S_1 + (1 - \lambda)S_2$ est symétrique. De plus,

$${}^tX(\lambda S_1 + (1 - \lambda)S_2)X = \lambda {}^tXS_1X + (1 - \lambda){}^tXS_2X.$$

Ainsi, ${}^tX(\lambda S_1 + (1 - \lambda)S_2)X$ est positif quand ${}^tXS_1X$ et ${}^tXS_2X$ le sont. De même, ${}^tX(\lambda S_1 + (1 - \lambda)S_2)X$ est inférieur ou égal à 1 quand ${}^tXS_1X$ et ${}^tXS_2X$ le sont. La première propriété est vraie pour tout X de $\mathbb{R}^d$, et la seconde pour tout X de K . On en conclut que la matrice $\lambda S_1 + (1 - \lambda)S_2$ appartient à C_K . Ainsi, C_K est convexe.

Pour la stricte log-concavité, supposons $0 < \lambda < 1$ et $S_1 \neq S_2$. Comme les deux matrices S_1 et S_2 sont symétriques définies positives, le théorème spectral montre qu'il existe une base orthonormée pour le produit scalaire associé à S_1 qui est orthogonale pour le produit scalaire associé à S_2 . Si Q est la matrice dont les colonnes sont les vecteurs de cette base dans la base canonique, on pose $P = Q^{-1}$. Ainsi, on obtient $\mu_1, \dots, \mu_d > 0$ tels que

$$S_1 = {}^tPP \quad \text{et} \quad S_2 = {}^tP\text{diag}(\mu_1, \dots, \mu_d)P,$$

de sorte que $\det S_1 = (\det P)^2$, ainsi que $\det S_2 = (\det P)^2 \mu_1 \cdots \mu_d$, puis

$$\det(\lambda S_1 + (1 - \lambda)S_2) = \det({}^tP\text{diag}(\lambda + (1 - \lambda)\mu_1, \dots, \lambda + (1 - \lambda)\mu_d)P) = (\det P)^2 \prod_{i=1}^d (\lambda + (1 - \lambda)\mu_i).$$

La stricte concavité du logarithme assure que

$$\log(\lambda + (1 - \lambda)\mu_i) \geq \lambda \log 1 + (1 - \lambda) \log \mu_i \quad \text{et donc} \quad \prod_{i=1}^d (\lambda + (1 - \lambda)\mu_i) > \prod_{i=1}^d \mu_i^{1-\lambda}.$$

La dernière inégalité est stricte car au moins un des μ_i est différent de 1, sinon S_1 et S_2 seraient les mêmes. On en conclut que

$$\det(\lambda S_1 + (1 - \lambda)S_2) > (\det P)^2 \prod_{i=1}^d \mu_i^{1-\lambda} = ((\det P)^2)^\lambda \left((\det P)^2 \prod_{i=1}^d \mu_i \right)^{1-\lambda} = (\det S_1)^\lambda (\det S_2)^{1-\lambda}.$$

7. Si S_1 et S_2 désignent deux matrices distinctes de C_K qui maximisent le déterminant, on observe avec le résultat ci-dessus que la matrice $(S_1 + S_2)/2$ appartient aussi à C_K et vérifie

$$\det\left(\frac{S_1 + S_2}{2}\right) > (\det S_1)^{1/2}(\det S_2)^{1/2} = \max_{C_K} \det,$$

ce qui donne une contradiction. Finalement, il existe une unique matrice de C_K qui maximise le déterminant.

8. Les ellipsoïdes centrés à l'origine sont de la forme E_q , avec q une forme quadratique définie positive, et ceux qui contiennent K ont pour matrice associée à q une matrice S qui appartient à C_K . De plus, minimiser le volume de E_q revient à maximiser le déterminant de S . L'unique matrice de C_K qui maximise le déterminant correspond donc à une unique forme quadratique q définie positive qui minimise le volume de E_q et telle que E_q contienne K , et enfin correspond à un unique ellipsoïde centré en l'origine et de volume minimal et qui contienne K .

9. On a clairement $x \in E_{q,c}$ si et seulement si $x - c \in E_q$, ce qui signifie que $E_{q,c} = c + E_q$. Si K désigne un compact d'intérieur non vide, alors le translaté $-c + K$ est également un compact d'intérieur non vide. On note E_q l'ellipsoïde de John-Loewner de $-c + K$. Alors, $E_{q,c} = c + E_q$ est un ellipsoïde centré en c qui contient K , et de volume minimal, car le volume est préservé par translation, ce qui garantit l'unicité.

Exercice 10.18 (Sous-groupes compacts maximaux de $GL_d(\mathbb{R})$)

Leçon : 219. Référence : [10].

1. Soit G un sous-groupe compact de $GL_d(\mathbb{R})$. Montrer qu'il existe un produit scalaire sur $\mathbb{R}^d$, associé à une forme quadratique q , tel que $G \subset O(q)$, où $O(q)$ désigne le groupe orthogonal de q , c'est-à-dire l'ensemble des matrices $M \in GL_d(\mathbb{R})$ telles que $q(Mx) = q(x)$ pour tout $x \in \mathbb{R}^d$.
2. En déduire que les sous-groupes compacts de $GL_d(\mathbb{R})$ qui sont maximaux pour l'inclusion sont exactement les groupes orthogonaux $O(q)$, avec q une forme quadratique définie positive.
3. *Application.* Soit $\|\cdot\|$ une norme sur $\mathbb{R}^d$ pour laquelle le groupe des isométries agit transitivement sur la sphère unité. Montrer qu'il s'agit nécessairement d'une norme euclidienne.

Corrigé.

1. On note B la boule unité fermée de $\mathbb{R}^d$ pour la norme euclidienne canonique. L'ensemble $K = \{Mx, M \in G, x \in B\}$ est un compact de $\mathbb{R}^d$, en tant qu'image du compact $G \times B$ par l'application continue $(M, x) \mapsto Mx$. Il est d'intérieur non vide car il contient B , et il est invariant par tout élément de G . On peut donc considérer l'ellipsoïde de John-Loewner de K , noté E_q pour une unique forme quadratique définie positive q . Pour toute matrice $M \in G$, la forme quadratique $q'(x) = q(Mx)$ est définie positive, et vérifie $K \subset E_{q'}$ car $MK = K$. De plus, le déterminant est borné sur le groupe compact G , donc en particulier sur $\{M^p, p \in \mathbb{Z}\}$, ce qui impose $|\det M| = 1$. On en déduit que les matrices symétriques définies positives associées à q et q' ont le même déterminant, c'est-à-dire que E_q et $E_{q'}$ ont le même volume. Par unicité, on a donc $E_q = E_{q'}$, puis $q = q'$. On en conclut que $G \subset O(q)$.
2. On note G un sous-groupe compact de $GL_d(\mathbb{R})$ maximal pour l'inclusion. La question précédente montre que $G \subset O(q)$ pour une forme quadratique définie positive q . Comme $O(q)$ est compact (car fermé et borné), on en déduit par maximalité que $G = O(q)$.
Réciproquement, supposons $O(q) \subset G'$, où G' est un sous-groupe compact. La première question fournit une forme quadratique définie positive q' telle que

$$O(q) \subset G' \subset O(q').$$

Le groupe $O(q)$ agit transitivement sur la sphère $\{x : q(x) = 1\}$: cela découle du même résultat pour la sphère euclidienne après un changement de base. Comme q' est invariante par ce groupe, elle est constante sur cette sphère, de valeur $c > 0$. Par homogénéité, $q' = cq$ sur $\mathbb{R}^d$, donc $O(q') = O(q)$. Les inclusions précédentes donnent $G' = O(q)$, ce qui prouve la maximalité.

3. Le groupe des isométries G pour la norme $\|\cdot\|$ est par définition constitué des matrices M de $GL_d(\mathbb{R})$ telles que $\|Mx\| = \|x\|$ pour tout x de $\mathbb{R}^d$. Il s'agit d'un sous-groupe compact (car fermé et borné) de $GL_d(\mathbb{R})$, de sorte que $G \subset O(q)$ pour une forme quadratique définie positive q . De plus, comme G agit transitivement sur la sphère unité, on a pour tous x_0 et x de cette sphère, une matrice M dans G telle que $x = Mx_0$. Ainsi, $q(x) = q(Mx_0) = q(x_0)$, puisque M est dans $O(q)$. Il en ressort que q est constante sur la sphère unité, donc pour tout $x \neq 0$, on a $q(x/\|x\|) = q(x_0)$, et donc $q(x) = q(x_0)\|x\|^2$, soit encore $\|x\| = (q(x)/q(x_0))^{1/2}$, ce qui montre que $\|\cdot\|$ découle du produit scalaire associé à $q/q(x_0)$. C'est donc une norme euclidienne.

Équations différentielles

Références : [6, 9].

Exercice 11.1 (Solutions d'une équation différentielle d'ordre n à coefficients constants)

Référence : [6].

Pour tout polynôme complexe $P = \sum_{k=0}^n a_k X^k$ et pour toute fonction $f : \mathbb{R} \rightarrow \mathbb{C}$ de classe $\mathcal{C}^n$, on note

$$P(D)f = \sum_{k=0}^n a_k f^{(k)} = a_n f^{(n)} + \dots + a_1 f' + a_0 f.$$

(Il s'agit d'un polynôme de l'opérateur D de dérivation.)

1. Pour tout polynôme $P \in \mathbb{C}[X]$, notons $\mathcal{S}_P$ l'espace des solutions de l'équation différentielle $P(D)y = 0$. Si $P_1, \dots, P_j \in \mathbb{C}[X]$ sont des polynômes non nuls premiers entre eux deux à deux, montrer que $\mathcal{S}_{P_1 \dots P_j} = \mathcal{S}_{P_1} \oplus \dots \oplus \mathcal{S}_{P_j}$.
2. Si $P(X) = (X - \alpha)^n$, déterminer la forme des solutions de l'équation différentielle $P(D)y = 0$.
3. En déduire, pour tout polynôme $P \in \mathbb{C}[X]$ de degré ≥ 1 , la forme des solutions de l'équation différentielle $P(D)y = 0$.

Corrigé.

1. Les solutions des équations différentielles considérées sont des fonctions de classe $\mathcal{C}^\infty$, donc on peut voir l'opérateur de dérivation D comme l'endomorphisme de $\mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$ qui à toute fonction f associe sa dérivée f' . Pour tout polynôme $P \in \mathbb{C}[X]$, l'opérateur $P(D)$ est donc un polynôme d'endomorphisme et l'ensemble des solutions de l'équation différentielle $P(D)f = 0$ est égal au noyau de $P(D)$, c'est-à-dire que $\mathcal{S}_P = \ker P(D)$. Par conséquent, le lemme des noyaux assure que si $P_1, \dots, P_j \in \mathbb{C}[X]$ sont des polynômes non nuls premiers entre eux deux à deux, on a $\ker P_1 \dots P_j(D) = \ker P_1(D) \oplus \dots \oplus \ker P_j(D)$, soit encore $\mathcal{S}_{P_1 \dots P_j} = \mathcal{S}_{P_1} \oplus \dots \oplus \mathcal{S}_{P_j}$.
2. Soit $\alpha \in \mathbb{C}$. Notons $Q_{n,\alpha} = (X - \alpha)^n$ et montrons par récurrence sur $n \in \mathbb{N}^*$ que les solutions de l'équation différentielle $Q_{n,\alpha}(D)f = 0$ sont les fonctions $f : t \mapsto e^{\alpha t} \varphi(t)$, où φ est une fonction polynomiale complexe de degré $< n$. Pour $n = 1$, l'équation $Q_{1,\alpha}(D)f = 0$ s'écrit $f' - \alpha f = 0$, et ses solutions sont les fonctions définies sur $\mathbb{R}$ de la forme $t \mapsto \lambda e^{\alpha t}$, avec $\lambda \in \mathbb{C}$. L'assertion est donc vraie au rang $n = 1$. Soit ensuite un entier $n \in \mathbb{N}^*$ pour lequel l'assertion est vraie. On a $Q_{n+1,\alpha}(D)f = 0$ si et seulement si $Q_{n,\alpha}(D)Q_{1,\alpha}(D)f = 0$. D'après l'hypothèse de récurrence, cela équivaut à l'existence d'une fonction polynomiale complexe φ de degré $< n$ telle que $f' - \alpha f = e^{\alpha t} \varphi(t)$, c'est-à-dire $(e^{-\alpha t} f)' = \varphi(t)$. Ainsi, $f(t) = e^{\alpha t} \Phi(t)$, où Φ est une primitive de φ , donc une fonction polynomiale complexe de degré $< n + 1$. L'assertion est vraie au rang $n + 1$ et on conclut grâce au principe de récurrence.
3. Désignons par $P = \lambda Q_{n_1,\alpha_1} \dots Q_{n_j,\alpha_j}$ la décomposition du polynôme $P \in \mathbb{C}[X]$ en facteurs irréductibles. D'après la première question, on a $\mathcal{S}_P = \mathcal{S}_{Q_{n_1,\alpha_1}} \oplus \dots \oplus \mathcal{S}_{Q_{n_j,\alpha_j}}$. Et d'après la deuxième question, les espaces intervenant dans cette décomposition sont formés respectivement des fonctions de la forme $t \mapsto e^{\alpha_1 t} \varphi_1(t), \dots, t \mapsto e^{\alpha_j t} \varphi_j(t)$, avec des fonctions polynomiales complexes de degrés strictement inférieurs à $n_1, \dots, n_j$, respectivement. Les solutions de l'équation différentielle $P(D)f = 0$ sont donc de la forme

$$f : t \mapsto e^{\alpha_1 t} \varphi_1(t) + \dots + e^{\alpha_j t} \varphi_j(t),$$

où les fonctions $\varphi_1, \dots, \varphi_j$ sont des fonctions polynomiales complexes de degrés strictement inférieurs à $n_1, \dots, n_j$, respectivement.

Exercice 11.2 (Théorème de Floquet)

Référence : [6].

On considère sur $\mathbb{C}^n$ le système différentiel (E) : $Y' = A(t)Y$, où A est une fonction continue et T -périodique de la variable réelle, à valeurs dans $\mathcal{M}_n(\mathbb{C})$.

1. Montrer que (E) admet une solution V non nulle telle que

$$\exists \lambda \in \mathbb{C} \quad \forall t \in \mathbb{R} \quad V(t+T) = \lambda V(t).$$

2. On considère n solutions $V_1, \dots, V_n$ de (E) linéairement indépendantes. En notant $M(t)$ la matrice carrée dont les vecteurs colonnes sont $V_1(t), \dots, V_n(t)$, montrer que

$$\exists C \in \text{GL}_n(\mathbb{C}) \quad \forall t \in \mathbb{R} \quad M(t+T) = M(t)C.$$

Corrigé.

1. L'ensemble $\mathcal{S}$ des solutions de (E) est un $\mathbb{C}$ -espace vectoriel de dimension n . Comme la fonction A est T -périodique, partant de n'importe quelle solution V , on observe que la fonction $V_T : t \mapsto V(t+T)$ est aussi solution. L'application $\Phi : V \mapsto V_T$ définit donc un endomorphisme de $\mathcal{S}$. Comme $\mathcal{S}$ est de dimension finie, l'endomorphisme Φ admet une valeur propre λ , et un vecteur propre associé V . Ainsi, V est une solution non nulle de (E) qui vérifie $V_T = \lambda V$, c'est-à-dire $V(t+T) = \lambda V(t)$ pour tout $t \in \mathbb{R}$.
2. Les vecteurs $V_1, \dots, V_n$ forment une base de $\mathcal{S}$. L'application Φ est inversible, son inverse étant la translation de $-T$. Notons $C = (c_{k,i})$ sa matrice dans cette base. Pour tout i et tout $t \in \mathbb{R}$,

$$V_i(t+T) = \sum_{k=1}^n c_{k,i} V_k(t).$$

Cette identité exprime l'égalité des i -èmes colonnes de $M(t+T)$ et $M(t)C$. Ainsi, $M(t+T) = M(t)C$, avec $C \in \text{GL}_n(\mathbb{C})$.

Exercice 11.3 (Calcul du Wronskien)

Référence : [6].

1. Soient I un intervalle de $\mathbb{R}$ et $p, q : I \rightarrow \mathbb{C}$ deux fonctions continues. On considère deux solutions u et v de l'équation différentielle (L) : $y'' + p(t)y' + q(t)y = 0$. Calculer le wronskien de u et v en fonction de sa valeur en un point $a \in I$.
2. On souhaite étendre ce résultat aux systèmes linéaires. Soient I un intervalle de $\mathbb{R}$ et $A : I \rightarrow \mathcal{M}_n(\mathbb{C})$ une fonction continue. On considère n solutions $f_1, \dots, f_n$ du système différentiel (S) : $Y' = A(t)Y$. Calculer le wronskien de $f_1, \dots, f_n$ en fonction de sa valeur en un point $a \in I$.

Corrigé.

1. Le wronskien des deux solutions u et v s'écrit $w = u v' - u' v$. En particulier, on a $w' = u v'' - u'' v$. De plus, on sait que

$$\begin{cases} u(v'' + p(t)v' + q(t)v) = 0 \\ v(u'' + p(t)u' + q(t)u) = 0, \end{cases}$$

En soustrayant ces deux équations, on obtient $w' + p(t)w = 0$. On en déduit pour tout $t \in I$,

$$w(t) = w(a) \exp\left(-\int_a^t p(u) du\right).$$

2. Pour un système linéaire de taille n , le wronskien de n solutions s'écrit $w = \det(f_1, \dots, f_n)$. En dérivant, on obtient

$$w' = \sum_{k=1}^n \det(f_1, \dots, f_{k-1}, f'_k, f_{k+1}, \dots, f_n) = \sum_{k=1}^n \det(f_1, \dots, f_{k-1}, A(t) f_k, f_{k+1}, \dots, f_n).$$

à $t \in I$ fixé, le membre de droite est une forme n -linéaire alternée des vecteurs $f_1, \dots, f_n$, donc est proportionnel au déterminant. Cependant, si on remplace ces vecteurs par la base canonique de $\mathbb{C}^n$, on montre que le coefficient de proportionnalité est égal à la trace de $A(t)$. Il en résulte que $w' = \text{Tr } A(t) \det(f_1, \dots, f_n) = \text{Tr } A(t) w$, soit finalement pour tout $t \in I$,

$$w(t) = w(a) \exp\left(\int_a^t \text{Tr } A(u) du\right).$$

Exercice 11.4 (Lemme de Gronwall et applications)

Référence : [6].

1. Soient φ, ψ et y trois fonctions continues sur un segment $[a, b]$, à valeurs positives et telles que

$$\forall t \in [a, b] \quad y(t) \leq \varphi(t) + \int_a^t \psi(s) y(s) ds.$$

(a). Pour $F(t) = \int_a^t \psi(s) y(s) ds$, montrer que $F'(t) - \psi(t) F(t) \leq \varphi(t) \psi(t)$ pour tout $t \in [a, b]$.

(b). En déduire une majoration de $G(t) = F(t) \exp\left(-\int_a^t \psi(s) ds\right)$.

(c). Conclure que pour tout $t \in [a, b]$,

$$y(t) \leq \varphi(t) + \int_a^t \varphi(s) \psi(s) \exp\left(\int_s^t \psi(u) du\right) ds.$$

2. Soient ψ et y deux fonctions continues sur un segment $[a, b]$, à valeurs positives et telles que

$$\exists c \geq 0 \quad \forall t \in [a, b] \quad y(t) \leq c + \int_a^t \psi(s) y(s) ds.$$

Montrer que pour tout $t \in [a, b]$,

$$y(t) \leq c \exp\left(\int_a^t \psi(s) ds\right).$$

3. Soit $y : [a, b] \rightarrow \mathbb{R}^n$ une fonction de classe $\mathcal{C}^1$ telle que

$$\exists \alpha > 0 \quad \exists \beta \geq 0 \quad \forall t \in [a, b] \quad \|y'(t)\| \leq \beta + \alpha \|y(t)\|.$$

Montrer que pour tout $t \in [a, b]$,

$$\|y(t)\| \leq \|y(a)\| e^{\alpha(t-a)} + \frac{\beta}{\alpha} (e^{\alpha(t-a)} - 1).$$

Corrigé.

1. (a). En multipliant l'inégalité de l'énoncé par $\psi(t) \geq 0$ et en passant l'intégrale dans le membre de gauche, on obtient directement $F'(t) - \psi(t) F(t) \leq \varphi(t) \psi(t)$ pour tout $t \in [a, b]$.

(b). Ensuite, en multipliant par $\exp\left(-\int_a^t \psi(s) ds\right)$, on obtient pour tout $t \in [a, b]$,

$$G'(t) \leq \varphi(t) \psi(t) \exp\left(-\int_a^t \psi(s) ds\right).$$

(c). Comme $G(a) = F(a) = 0$, on en déduit par intégration que

$$G(t) = \int_a^t G'(s) ds \leq \int_a^t \varphi(s) \psi(s) \exp\left(-\int_a^s \psi(u) du\right) ds.$$

Pour conclure, on combine cette majoration avec le fait que pour tout $t \in [a, b]$,

$$y(t) \leq \varphi(t) + F(t) = \varphi(t) + G(t) \exp\left(\int_a^t \psi(u) du\right).$$

2. On applique la question précédente avec $\varphi(t) = c$ pour tout $t \in [a, b]$. Ainsi,

$$y(t) \leq c + c \int_a^t \psi(s) \exp\left(\int_s^t \psi(u) du\right) ds = c \exp\left(\int_a^t \psi(s) ds\right).$$

3. Pour tout $t \in [a, b]$, on a

$$\|y(t)\| \leq \|y(a)\| + \int_a^t \|y'(s)\| ds \leq \|y(a)\| + \beta(t-a) + \alpha \int_a^t \|y(s)\| ds.$$

On applique ensuite le résultat de la première question avec $\varphi(t) = \|y(a)\| + \beta(t-a)$ et $\psi(t) = \alpha$ pour tout $t \in [a, b]$. Ainsi,

$$\|y(t)\| \leq (\|y(a)\| + \beta(t-a)) + \int_a^t (\|y(a)\| + \beta(s-a)) \alpha e^{\alpha(t-s)} ds.$$

à l'aide d'une intégration par parties, on voit que l'intégrale devient

$$\begin{aligned} & \left[(\|y(a)\| + \beta(s-a)) (-e^{\alpha(t-s)}) \right]_a^t + \beta \int_a^t e^{\alpha(t-s)} ds \\ &= -(\|y(a)\| + \beta(t-a)) + \|y(a)\| e^{\alpha(t-a)} + \beta \left[-\frac{e^{\alpha(t-s)}}{\alpha} \right]_{s=a}^{s=t} \\ &= -(\|y(a)\| + \beta(t-a)) + \|y(a)\| e^{\alpha(t-a)} + \frac{\beta}{\alpha} (e^{\alpha(t-a)} - 1). \end{aligned}$$

Le résultat s'ensuit directement.

Exercice 11.5 (Équations différentielles d'ordre 2)

Référence : [6].

- On considère l'équation différentielle linéaire homogène (E) : $y'' + p(t)y' + q(t)y = 0$, où p est une fonction de classe $\mathcal{C}^1$ définie sur un intervalle I de $\mathbb{R}$, et q est une fonction continue sur I . En procédant au changement de fonction $x = y \exp\left(\frac{1}{2} \int_{t_0}^t p(s) ds\right)$, avec $t_0 \in I$ quelconque, montrer qu'on peut se ramener au cas où $p = 0$.
- On considère l'équation différentielle (L) : $y'' + q(t)y = 0$, où $q : \mathbb{R}^+ \rightarrow \mathbb{R}$ est une fonction continue telle que $\int_0^{+\infty} |q(t)| dt$ converge.
 - Si y désigne une solution bornée de (L), montrer que y' tend vers 0 en $+\infty$.
 - En déduire que (L) admet des solutions non bornées.
- On considère la même équation (L), mais on suppose maintenant que la fonction q est $\mathcal{C}^1$, strictement positive et croissante sur $\mathbb{R}^+$. Montrer que les solutions réelles sont bornées sur $\mathbb{R}^+$.
- On suppose dans cette question que la fonction q est définie sur $\mathbb{R}$ tout entier, continue et strictement négative.
 - Montrer que la seule solution réelle de (L) bornée sur $\mathbb{R}$ est la fonction nulle.
 - Montrer qu'une solution non nulle s'annule au plus une fois sur $\mathbb{R}$.
- On suppose que la fonction q est définie et continue sur un intervalle $[a, b]$, et que l'équation (L) admet une solution y nulle en a et en b , et strictement positive sur $]a, b[$. Montrer que

$$\int_a^b |q(t)| dt > \frac{4}{b-a}.$$

Corrigé.

- Le changement de fonction proposé correspond à $y = xz$, avec $z = \exp\left(-\frac{1}{2} \int_{t_0}^t p(s) ds\right)$. On a clairement $2z' + p(t)z = 0$. De plus, $y' = x'z + xz'$ et $y'' = x''z + 2x'z' + xz''$. En remplaçant dans l'équation différentielle (E), on obtient

$$0 = y'' + py' + qy = xz'' + \underbrace{(2z' + pz)}_{=0} x' + (z'' + pz' + qz)x,$$

ce qui équivaut, en divisant par la fonction z qui est non nulle, à l'équation

$$x'' + \left(q(t) - \frac{p(t)^2}{4} - \frac{p'(t)}{2}\right)x = 0.$$

La fonction x est donc solution d'une équation linéaire homogène d'ordre 2 dont le terme en x' est nul. On peut donc toujours se ramener au cas où $p = 0$.

- Comme y est bornée et q est intégrable, alors $y'' = -qy$ est intégrable aussi. La fonction y' admet donc une limite finie, notée α , en $+\infty$. Si α était non nul, on aurait $y(t) \underset{t \rightarrow +\infty}{\sim} \alpha t$, ce qui est impossible puisque y est bornée. On a donc $\alpha = 0$, c'est-à-dire que $y'(t) \xrightarrow{t \rightarrow +\infty} 0$.
 - Supposons que toutes les solutions de (L) soient bornées. En particulier, si (u, v) désigne une base de solutions, alors les deux dérivées u' et v' tendent vers 0 en $+\infty$, donc le wronskien $w = u'v - uv'$ fait de même. Cependant, l'exercice 11.3 montre que le wronskien est ici constant. Il est donc partout nul, ce qui contredit le fait que (u, v) est une base.
- Soit y une solution. En multipliant l'équation (L) par $2y'$, on obtient $2y''y' + 2q(t)y'y = 0$, donc par intégration, pour tout $t \geq 0$,

$$y'(t)^2 - y'(0)^2 + 2 \int_0^t q(s)y'(s)y(s) ds = 0.$$

à l'aide d'une intégration par parties, le dernier terme devient

$$[q(s)y(s)^2]_0^t - \int_0^t q'(s)y(s)^2 ds = q(t)y(t)^2 - q(0)y(0)^2 - \int_0^t q'(s)y(s)^2 ds,$$

si bien que pour tout $t \geq 0$,

$$y'(t)^2 + q(t)y(t)^2 = \underbrace{y'(0)^2 + q(0)y(0)^2}_K + \int_0^t q'(s)y(s)^2 ds.$$

En particulier, pour tout $t \geq 0$,

$$q(t)y(t)^2 \leq K + \int_0^t \frac{q'(s)}{q(s)} q(s)y(s)^2 ds.$$

L'exercice 11.4, appliqué à la fonction qy^2 , montre finalement que pour tout $t \geq 0$,

$$q(t)y(t)^2 \leq K \exp\left(\int_0^t \frac{q'(s)}{q(s)} ds\right) = K \frac{q(t)}{q(0)}.$$

On en conclut bien que la fonction y est bornée sur $\mathbb{R}^+$.

4. (a). Soit f une solution de (L) et soit $z = f^2$. Comme q est strictement négative, on a

$$z'' = (2f'f)' = 2(f''f + (f')^2) = 2((f')^2 - q(t)f^2) \geq 0.$$

La fonction z est donc convexe. Si elle est constante, alors f est aussi constante, et donc elle est nulle : $f = -f''/q(t) = 0$. Sinon, on a $z(t) \geq z(t_0) + z'(t_0)(t - t_0)$ pour tout $t \in \mathbb{R}$, avec t_0 un réel tel que $z'(t_0) \neq 0$. Selon le signe de $z'(t_0)$, on observe que la limite en $-\infty$ ou en $+\infty$ de la fonction z vaut $+\infty$. Ainsi, z et donc f ne sont pas bornées. La seule solution bornée est donc la fonction nulle.

- (b). Si une solution f s'annule en deux points t_1 et t_2 , alors $z = f^2$ aussi. Comme z est convexe et positive, elle est forcément nulle sur tout l'intervalle fermé dont les bornes sont t_1 et t_2 . C'est donc aussi le cas de f . Le théorème de Cauchy-Lipschitz assure alors que f est nulle sur $\mathbb{R}$ tout entier. Donc toute solution non nulle s'annule au plus une fois.
5. La fonction y est continue sur l'intervalle $[a, b]$, donc admet son maximum en un point c de cet intervalle. En fait, c est intérieur à l'intervalle puisque y est nul aux bornes et strictement positive à l'intérieur. Si $a < \alpha < \beta < b$, on a donc

$$\int_a^b |q(t)| dt = \int_a^b \left| \frac{y''(t)}{y(t)} \right| dt > \frac{1}{y(c)} \int_a^b |y''(t)| dt \geq \left| \frac{y'(\beta) - y'(\alpha)}{y(c)} \right|.$$

Pour justifier l'inégalité stricte, remarquons que y'' n'est pas identiquement nulle, sinon y serait affine et nulle aux deux bornes. Il existe alors un sous-intervalle sur lequel $|y''| > 0$ et $y < y(c)$: si un point où $y'' \neq 0$ est un maximum, des points voisins ont une valeur strictement inférieure au maximum et une dérivée seconde encore non nulle. D'après le théorème des accroissements finis, on peut choisir α et β tels que $a < \alpha < c < \beta < b$ et

$$y'(\alpha) = \frac{y(c) - y(a)}{c - a} = \frac{y(c)}{c - a} \quad \text{et} \quad y'(\beta) = \frac{y(b) - y(c)}{b - c} = -\frac{y(c)}{b - c}.$$

On en déduit alors que

$$\int_a^b |q(t)| dt > \frac{1}{c - a} + \frac{1}{b - c} \geq \frac{4}{b - a},$$

où la dernière borne correspond à la plus petite valeur possible selon la position de c , qui s'obtient quand c est au milieu de l'intervalle.

Exercice 11.6 (Principe de majoration a priori)

Référence : [6].

1. Soit Ω un ouvert de $\mathbb{R} \times \mathbb{R}^n$ et soit $F : \Omega \rightarrow \mathbb{R}^n$ une application continue, localement lipschitzienne en la seconde variable. Soit $(t_0, x_0) \in \Omega$. Montrer qu'il existe un voisinage V de x_0 dans $\mathbb{R}^n$ et un intervalle ouvert I centré en t_0 tels que $I \times V \subset \Omega$ et pour tout $(t_1, x) \in I \times V$, il existe une solution φ de $y' = F(t, y)$ définie sur I et vérifiant $\varphi(t_1) = x$.
2. Soit $]a, b[$ un intervalle ouvert, O un ouvert de $\mathbb{R}^n$ et $F :]a, b[\times O \rightarrow \mathbb{R}^n$ une fonction continue et localement lipschitzienne en la seconde variable. Soit $\varphi :]\alpha, \beta[\rightarrow \mathbb{R}^n$ une solution maximale de $y' = F(t, y)$. Montrer que si $\beta < b$ (resp. $\alpha > a$), alors pour tout compact $K \subset O$, il existe $\gamma \in]\alpha, \beta[$ tel que $\varphi(t) \notin K$ pour tout $t \in]\gamma, \beta[$ (resp. $t \in]\alpha, \gamma[$).
3. Quel énoncé plus simple peut-on formuler quand $O = \mathbb{R}^n$? Et dans le cas d'une équation de la forme $y' = F(y)$, avec $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ localement lipschitzienne?
4. On considère une équation différentielle (L) : $X' = A(t)X + B(t)$, avec $A :]a, b[\rightarrow \mathcal{M}_n(\mathbb{R})$ et $B :]a, b[\rightarrow \mathbb{R}^n$ des applications continues. Montrer que les solutions maximales de (L) sont définies sur $]a, b[$ tout entier.

Corrigé.

1. Pour $t \in \mathbb{R}$ et $\alpha > 0$, on note $I_\alpha(t) =]t - \alpha, t + \alpha[$. De plus, pour $x \in \mathbb{R}^n$ et $r > 0$, on désigne par $B_r(x) = B(x, r)$ la boule ouverte de centre x et de rayon r . On rappelle que si $C = I_\alpha(t) \times B_r(x)$ est un cylindre de sécurité pour F en (t, x) , c'est-à-dire que F est lipschitzienne en la seconde variable sur C et il existe un majorant $M > 0$ de F sur C tel que $\alpha M < r$, alors il existe une solution φ de $y' = F(t, y)$ définie sur $I_\alpha(t)$ et vérifiant $\varphi(t) = x$, voir la preuve du théorème de Cauchy-Lipschitz. Considérons alors un voisinage compact $K \subset \Omega$ de (t_0, x_0) tel que F soit lipschitzienne en la seconde variable sur K . Notons M un majorant de la fonction continue F sur le compact K et choisissons $\alpha, r > 0$ tels que $C = I_{\alpha/3}(t_0) \times B_r(x_0) \subset K$ et $\alpha M < r$, si bien que C est un cylindre de sécurité pour F en (t_0, x_0) . Pour conclure, posons $I = I_{\alpha/3}(t_0)$ et $V = B_r(x_0)$. Pour $(t_1, x) \in I \times V$, on observe que $I_{2\alpha/3}(t_1) \subset I_\alpha(t_0)$, que $B_{2r/3}(x) \subset B_r(x_0)$ et que $(2\alpha/3)M < 2r/3$, ce qui implique que $I_{2\alpha/3}(t_1) \times B_{2r/3}(x)$ est un cylindre de sécurité pour F en (t_1, x) . Ainsi, il existe une solution φ de $y' = F(t, y)$ vérifiant $\varphi(t_1) = x$ et définie sur $I_{2\alpha/3}(t_1)$, qui contient l'intervalle I .
2. Considérons $\beta < b$ et supposons par l'absurde qu'il existe un compact $K \subset O$ et une suite croissante (γ_n) qui converge vers β telle que $\varphi(\gamma_n) \in K$ pour tout n . Par compacité, on peut en extraire une sous-suite vérifiant $\varphi(\gamma_{n_k}) \xrightarrow[k \rightarrow \infty]{} x_0$. Comme $(\beta, x_0) \in]a, b[\times O$, la question précédente montre l'existence d'un réel $\varepsilon > 0$ et d'un ouvert $V \subset O$ contenant x_0 tels que pour tous $t_1 \in]\beta - \varepsilon, \beta + \varepsilon[$ et $x \in V$, il existe une solution ψ de $y' = F(t, y)$ définie sur $]\beta - \varepsilon, \beta + \varepsilon[$ et vérifiant $\psi(t_1) = x$. Pour k assez grand, on a $\gamma_{n_k} \in]\beta - \varepsilon, \beta + \varepsilon[$ et $\varphi(\gamma_{n_k}) \in V$, donc on peut considérer une telle solution ψ vérifiant $\psi(\gamma_{n_k}) = \varphi(\gamma_{n_k})$. D'après l'unicité dans le théorème de Cauchy-Lipschitz, les fonctions ψ et φ coïncident sur l'intervalle $]\beta - \varepsilon, \beta + \varepsilon[$. La fonction définie sur l'intervalle $]\alpha, \beta + \varepsilon[$ comme étant égale à φ sur $]\alpha, \beta[$ et égale à ψ sur $[\beta, \beta + \varepsilon[$ est donc une solution de $y' = F(t, y)$ définie sur un intervalle strictement plus grand que $]\alpha, \beta[$. Cela contredit le fait que φ est une solution maximale, et prouve donc le résultat en β . Le résultat en α se démontre de la même manière.
3. Si $O = \mathbb{R}^n$, le résultat montre que la solution maximale φ définie sur $]\alpha, \beta[$ vérifie $\|\varphi(t)\| \xrightarrow[t \rightarrow \beta]{} +\infty$ si $\beta < b$, et que $\|\varphi(t)\| \xrightarrow[t \rightarrow \alpha]{} +\infty$ si $\alpha > a$. De plus, si F ne dépend pas de t , alors on peut appliquer ce qui précède à la fonction définie par $F(t, y) = F(y)$, pour tout $t \in \mathbb{R}$. On a donc $a = -\infty$ et $b = +\infty$. Ainsi, la solution maximale φ définie sur $]\alpha, \beta[$ vérifie $\|\varphi(t)\| \xrightarrow[t \rightarrow \beta]{} +\infty$ si $\beta < +\infty$, et que $\|\varphi(t)\| \xrightarrow[t \rightarrow \alpha]{} +\infty$ si $\alpha > -\infty$.
4. On applique ce qui précède avec la fonction définie par $F(t, X) = A(t)X + B(t)$ pour $t \in]a, b[$ et $X \in \mathbb{R}^n$. Cette fonction est clairement continue et localement lipschitzienne en la seconde variable. Notons $X :]\alpha, \beta[\rightarrow \mathbb{R}^n$ une solution maximale de l'équation (L). Supposons $\beta < b$, et pour $t_0 \in]\alpha, \beta[$ quelconque, notons $M = 1 + \sup_{t_0 \leq t \leq \beta} \|A(t)\|$ et $N = \sup_{t_0 \leq t \leq \beta} \|B(t)\|$. On a donc $\|X'(t)\| \leq M\|X(t)\| + N$ pour tout $t \in [t_0, \beta[$, ce qui implique que $\|X(t)\| \leq \|X(t_0)\| e^{M(t-t_0)} + (e^{M(t-t_0)} - 1)N/M$ d'après l'exercice 11.4. En particulier, la fonction X est

bornée au voisinage à gauche de β . Cependant, la question précédente montre que $\|X(t)\| \xrightarrow[t \rightarrow \beta]{} +\infty$. Il y a donc une contradiction, ce qui impose d'avoir $\beta = b$. De même, on peut montrer que $\alpha = a$. Il s'ensuit que la solution maximale X est définie sur $]a, b[$ tout entier.

Exercice 11.7 (Théorème de Cauchy-Peano par la méthode d'Euler explicite)

Leçon : 226. Référence : [9].

Soit Ω un ouvert de $\mathbb{R} \times \mathbb{R}^n$ et soit $F : \Omega \rightarrow \mathbb{R}^n$ une application continue. L'objectif de l'exercice est de démontrer que pour tout $(t_0, y_0) \in \Omega$, il existe une solution $\varphi : I \rightarrow \mathbb{R}^n$, avec I un intervalle dont l'intérieur contient t_0 , au problème de Cauchy
$$\begin{cases} y' = F(t, y) \\ y(t_0) = y_0. \end{cases}$$

- Rappeler pourquoi une fonction $\varphi : I \rightarrow \mathbb{R}^n$ est solution de ce problème de Cauchy si et seulement si φ est continue et vérifie les conditions $(t, \varphi(t)) \in \Omega$ et $\varphi(t) = y_0 + \int_{t_0}^t F(s, \varphi(s)) ds$ pour tout $t \in I$.
- Montrer que l'ensemble Ω contient un cylindre de la forme $C_0 = [t_0 - T_0, t_0 + T_0] \times \bar{B}(y_0, r_0)$. Expliquer pourquoi $M = 1 + \sup_{(t,y) \in C_0} \|F(t, y)\|$ est fini, puis montrer que $C = [t_0 - T, t_0 + T] \times \bar{B}(y_0, r_0)$ est un *cylindre de sécurité* dès que $0 < T \leq \min\{T_0, r_0/M\}$, c'est-à-dire que toute solution $\varphi : I \rightarrow \mathbb{R}^n$ du problème de Cauchy définie sur un intervalle $I \subset [t_0 - T, t_0 + T]$ garde ses valeurs dans $\bar{B}(y_0, r_0)$.
- Soit $t_0 < t_1 < \dots < t_N = t_0 + T$ une subdivision de pas constant $h = T/N$. Partant de y_0 , on définit selon la méthode d'Euler explicite un N -uplet $(y_1, \dots, y_N)$ tel que $y_{n+1} = y_n + h F(t_n, y_n)$ pour tout $n \in \{0, \dots, N-1\}$, puis on note φ_N la fonction affine par morceaux sur $[t_0, t_0 + T]$ vérifiant $\varphi_N(t_n) = y_n$ pour tout $n \in \{0, \dots, N\}$. (On confond donc la solution sur chaque intervalle $[t_n, t_{n+1}]$ avec sa tangente au point (t_n, y_n) .) Expliquer comment construire selon la même méthode φ_N sur l'intervalle $[t_0 - T, t_0]$, avec une subdivision $t_0 - T = t_{-N} < t_{-N+1} < \dots < t_0$ de même pas constant h .
- Montrer que pour $0 < T \leq \min\{T_0, r_0/M\}$, les fonctions φ_N ainsi définies sur l'intervalle $[t_0 - T, t_0 + T]$ prennent leurs valeurs dans $\bar{B}(y_0, r_0)$.
- Toujours pour $0 < T \leq \min\{T_0, r_0/M\}$ et $C = [t_0 - T, t_0 + T] \times \bar{B}(y_0, r_0)$, on considère le module de continuité de F sur C défini par $\omega(\delta) = \sup_{\substack{(t,y),(s,x) \in C \\ |t-s| + \|y-x\| \leq \delta}} \|F(t, y) - F(s, x)\|$. Montrer que, pour $\varepsilon = \omega((M+1)h)$, la fonction affine par morceaux φ_N est une *solution ε -approchée* du problème de Cauchy, c'est-à-dire que :
 - $(t, \varphi_N(t)) \in \Omega$ pour tout $t \in [t_0 - T, t_0 + T]$;
 - $\|\varphi'_N(t) - F(t, \varphi_N(t))\| \leq \varepsilon$ pour tout $t \in]t_n, t_{n+1}[$ et tout $n \in \{-N, \dots, N-1\}$.
- On suppose que la suite de fonctions (φ_N) converge uniformément sur l'intervalle $[t_0 - T, t_0 + T]$ vers une fonction φ . Montrer que φ est solution du problème de Cauchy.
- Montrer que toutes les fonctions φ_N sont M -lipschitziennes. Conclure à l'aide du théorème d'Ascoli.

Corrigé.

- Si une fonction $\varphi : I \rightarrow \mathbb{R}^n$ est solution du problème de Cauchy, alors elle est par définition dérivable, donc continue sur I . De plus, on a $(t, \varphi(t)) \in \Omega$ pour tout $t \in I$. Enfin, comme φ est de classe $\mathcal{C}^1$, il est clair que pour tout $t \in I$,

$$\varphi(t) = \varphi(t_0) + \int_{t_0}^t \varphi'(s) ds = y_0 + \int_{t_0}^t F(s, \varphi(s)) ds.$$

Réciproquement, si la fonction $\varphi : I \rightarrow \mathbb{R}^n$ est continue et vérifie les conditions $(t, \varphi(t)) \in \Omega$ et $\varphi(t) = y_0 + \int_{t_0}^t F(s, \varphi(s)) ds$ pour tout $t \in I$, alors φ est une primitive de la fonction continue $s \mapsto F(s, \varphi(s))$, donc elle est dérivable sur I et vérifie $\varphi'(t) = F(t, \varphi(t))$ pour tout $t \in I$. En outre, $\varphi(t_0) = y_0$. Ainsi, φ est solution du problème de Cauchy.

2. Comme l'ensemble Ω est ouvert qui contient le point (t_0, y_0) , alors $C_0 = [t_0 - T_0, t_0 + T_0] \times \bar{B}(y_0, r_0)$ est contenu dans Ω quitte à choisir $T_0, r_0 > 0$ suffisamment petits. Le cylindre C_0 est compact et l'application F est continue, donc $M = 1 + \sup_{(t,y) \in C_0} \|F(t, y)\|$ est fini. Enfin, pour $0 < T \leq \min\{T_0, r_0/M\}$ quelconque, considérons le cylindre $C = [t_0 - T, t_0 + T] \times \bar{B}(y_0, r_0)$, et une solution $\varphi : I \rightarrow \mathbb{R}^n$ du problème de Cauchy définie sur un intervalle $I \subset [t_0 - T, t_0 + T]$. Supposons que la fonction φ ne garde pas ses valeurs dans la boule $\bar{B}(y_0, r_0)$. Ainsi,

$$S = \{t \in I \mid \|\varphi(t) - y_0\| > r_0\} \neq \emptyset.$$

Par continuité de φ , on a $\tau = \inf_{t \in S} |t - t_0| > 0$, et de plus on a $t_1 \in I$ et $\|\varphi(t_1) - y_0\| = r_0$ pour $t_1 \in \{t_0 - \tau, t_0 + \tau\}$. La question précédente implique que

$$r_0 = \|\varphi(t_1) - y_0\| = \left\| \int_{t_0}^{t_1} F(s, \varphi(s)) \, ds \right\| \leq \left| \int_{t_0}^{t_1} \|F(s, \varphi(s))\| \, ds \right|.$$

Pour s compris entre t_0 et t_1 , on a $s \in I$ et $\varphi(s) \in \bar{B}(y_0, r_0)$, ce qui montre que le point $(s, \varphi(s))$ appartient au cylindre C_0 , et donc que $\|F(s, \varphi(s))\| \leq M$. D'où $r_0 \leq M |t_1 - t_0| = M\tau$. Cependant, on a $\tau < T$, donc $M\tau < MT \leq r_0$, ce qui donne une contradiction. On en conclut que fonction φ garde ses valeurs dans la boule $\bar{B}(y_0, r_0)$.

3. On note $t_n = t_0 + nh$ pour $n \in \{-N, \dots, -1\}$, de sorte que $t_0 - T = t_{-N} < t_{-N+1} < \dots < t_0$ est bien une subdivision de pas constant $h = T/N$. Partant de y_0 , on construit ensuite un N -uplet $(y_{-1}, \dots, y_{-N})$ tel que $y_{n-1} = y_n - h F(t_n, y_n)$ pour tout $n \in \{0, \dots, -N+1\}$. On étend finalement la fonction φ_N à l'intervalle $[t_0 - T, t_0]$ comme étant affine par morceaux et vérifiant $\varphi_N(t_n) = y_n$ pour tout $n \in \{-N, \dots, 0\}$.
4. On considère $0 < T \leq \min\{T_0, r_0/M\}$ et la fonction φ_N construite précédemment. Montrons par récurrence (finie) que pour tout entier $n \in \{0, \dots, N\}$, on a $\|\varphi_N(t) - y_0\| \leq M |t - t_0|$ pour tout $t \in [t_{-n}, t_n]$. C'est trivial pour $n = 0$. Supposons que cela soit vrai pour $n \in \{0, \dots, N-1\}$ fixé. En particulier, on a $\|y_n - y_0\| = \|\varphi_N(t_n) - y_0\| \leq M(t_n - t_0) \leq MT \leq r_0$, de sorte que $(t_n, y_n) \in C$, et donc $\|F(t_n, y_n)\| \leq M$. Pour $t \in [t_n, t_{n+1}]$, on a explicitement $\varphi_N(t) = F(t_n, y_n)(t - t_n) + y_n$, donc l'inégalité triangulaire donne

$$\|\varphi_N(t) - y_0\| \leq \|F(t_n, y_n)\| (t - t_n) + \|y_n - y_0\| \leq M(t - t_n) + M(t_n - t_0) = M(t - t_0).$$

Le raisonnement précédent reste vrai si on remplace l'entier n par $-n$, et l'intervalle $[t_n, t_{n+1}]$ par $[t_{-n-1}, t_{-n}]$. On en déduit le résultat au rang $n+1$. En particulier, au rang final N , on a

$$\forall t \in [t_0 - T, t_0 + T] \quad \|\varphi_N(t) - y_0\| \leq M |t - t_0| \leq MT \leq r_0.$$

Ainsi, la fonction φ_N prend ses valeurs dans $\bar{B}(y_0, r_0)$.

5. Pour tout $t \in [t_0 - T, t_0 + T]$, on a $\varphi_N(t) \in \bar{B}(y_0, r_0)$, donc $(t, \varphi_N(t))$ appartient à C , donc à Ω . De plus, pour tout $n \in \{0, \dots, N-1\}$ et tout $t \in]t_n, t_{n+1}[$, on a $\varphi'_N(t) = F(t_n, y_n)$, donc

$$\|\varphi'_N(t) - F(t, \varphi_N(t))\| = \|F(t_n, y_n) - F(t, \varphi_N(t))\|.$$

Sachant que les points (t_n, y_n) et $(t, \varphi_N(t))$ appartiennent tous deux à C et vérifient

$$|t_n - t| + \|y_n - \varphi_N(t)\| = |t - t_n| + \|F(t_n, y_n)(t - t_n)\| \leq (M+1)h,$$

on en déduit par définition du module de continuité que $\|\varphi'_N(t) - F(t, \varphi_N(t))\|$ est majoré par $\varepsilon = \omega((M+1)h)$. On peut reproduire le même raisonnement en remplaçant l'entier n par $-n$, et l'intervalle $]t_n, t_{n+1}[$ par $]t_{-n-1}, t_{-n}[$. Finalement, φ_N est bien une solution ε -approchée.

6. Tout d'abord, en tant que limite uniforme d'une suite de fonctions continues, la fonction φ est continue sur $[t_0 - T, t_0 + T]$. De plus, pour tout réel t de cet intervalle et pour tout entier N , le point $(t, \varphi_N(t))$ appartient au fermé C . Le point limite $(t, \varphi(t))$ appartient donc aussi à C , et donc à Ω . Enfin, par intégration, le fait que φ_N est bien une solution ε -approchée donne

$$\left\| \varphi_N(t) - y_0 - \int_{t_0}^t F(s, \varphi_N(s)) \, ds \right\| \leq \varepsilon |t - t_0|.$$

Ainsi, en notant δ_N le supremum de $\|\varphi_N - \varphi\|$ sur l'intervalle $[t_0 - T, t_0 + T]$, on obtient

$$\left\| \varphi(t) - y_0 - \int_{t_0}^t F(s, \varphi(s)) ds \right\| \leq \varepsilon |t - t_0| + \delta_N + \left| \int_{t_0}^t \|F(s, \varphi_N(s)) - F(s, \varphi(s))\| ds \right|.$$

Par définition du module de continuité, le dernier terme est majoré par $\omega(\delta_N) |t - t_0|$. On en conclut que pour tout $t \in [t_0 - T, t_0 + T]$,

$$\left\| \varphi(t) - y_0 - \int_{t_0}^t F(s, \varphi(s)) ds \right\| \leq \delta_N + \left(\omega(\delta_N) + \omega\left((M+1)\frac{T}{N}\right) \right) T.$$

Comme le module de continuité ω tend vers 0 en 0, on en déduit que le membre de droite tend vers 0 quand N tend vers l'infini. Finalement, φ satisfait l'équation intégrale qui montre qu'elle est solution au problème de Cauchy.

7. On a vu que pour tout $n \in \{0, \dots, N-1\}$ et tout $t \in]t_n, t_{n+1}[$, on a $\varphi'_N(t) = F(t_n, y_n)$, avec $(t_n, y_n) \in C$, de sorte que $\|\varphi'_N(t)\| \leq M$. Cette conclusion reste valable avec $-n$ à la place de n , et $]t_{-n-1}, t_{-n}[$ à la place de $]t_n, t_{n+1}[$. Par intégration, on en déduit que la fonction φ_N est M -lipschitzienne. Ainsi, (φ_N) est une suite de fonctions de $[t_0 - T, t_0 + T]$ dans $\overline{B}(y_0, r_0)$, qui sont deux espaces métriques compacts. Le théorème d'Ascoli implique alors l'existence d'une sous-suite de (φ_N) qui converge uniformément. Le raisonnement de la question précédente s'applique à cette sous-suite, ce qui montre l'existence d'une solution au problème de Cauchy.

Exercice 11.8 (Équations différentielles de type Fuchs)

Référence : [6].

Partie I

Pour $a, b \in \mathbb{R}$, on considère l'équation différentielle linéaire

$$x^2 y'' + ax y' + by = 0. \quad (E)$$

- (a). Rappeler pourquoi l'ensemble des solutions réelles de (E) sur $]0, +\infty[$ est un sous-espace de dimension 2 de l'espace vectoriel des fonctions réelles de classe C^2 sur $]0, +\infty[$.
- (b). On rappelle qu'une application Φ de $]0, +\infty[$ dans $\mathbb{C}$ est une solution complexe de (E) si et seulement si les fonctions $\Re \Phi$ et $\Im \Phi$ sont des solutions de (E). Montrer que l'ensemble des solutions complexes de (E) sur $]0, +\infty[$ est un espace vectoriel complexe de dimension 2.
- à quelle condition l'équation (E) admet-elle des solutions polynomiales non nulles ? Préciser alors ces solutions.
- En effectuant le changement de variable $x = e^t$, résoudre l'équation différentielle (E) ; préciser la forme générale des solutions réelles sur $]0, +\infty[$.

Partie II

On considère deux séries entières $a(x) = \sum_{n=0}^{\infty} a_n x^n$ et $b(x) = \sum_{n=0}^{\infty} b_n x^n$ à coefficients réels, toutes deux de rayon de convergence supérieur ou égal à $R > 0$. On s'intéresse sur $]0, R[$ à l'équation différentielle linéaire

$$x^2 y'' + a(x)x y' + b(x)y = 0. \quad (F)$$

Pour $x > 0$ et $k \in \mathbb{C}$, on note comme d'habitude $x^k = \exp(k \log x)$.

- Soit $k \in \mathbb{C}$ et $c(x) = \sum_{n=0}^{\infty} c_n x^n$ une série entière à coefficients complexes de rayon de convergence non nul. On suppose que $c_0 = 1$.
 - On suppose que la fonction $y = c(x)x^k$ est une solution complexe de (F) sur un intervalle de la forme $]0, R'[$. Montrer que le nombre complexe k est racine d'une équation du second degré que l'on précisera.

- (b). Montrer que, dans ce cas, les nombres c_n vérifient une relation de récurrence que l'on précisera et que l'on notera (P).

2. Soit k une racine de l'équation

$$k^2 + (a_0 - 1)k + b_0 = 0. \quad (C)$$

- (a). On suppose que $\Delta = (a_0 - 1)^2 - 4b_0$ n'est pas le carré d'un nombre entier non nul. Montrer qu'il existe une unique suite de nombres complexes $(c_n)_{n \geq 0}$ vérifiant $c_0 = 1$ et la relation de récurrence (P) établie ci-dessus.
- (b). Si $\Delta = p^2$, avec p entier strictement positif, donner une condition nécessaire et suffisante pour qu'il existe une suite $(c_n)_{n \geq 0}$ vérifiant $c_0 = 1$ et la relation (P) associée à k , où k est la plus petite racine réelle de l'équation (C). Montrer qu'en outre, il existe toujours une suite $(c_n)_{n \geq 0}$ vérifiant la relation (P) associée à une autre valeur k' convenablement choisie.
3. (a). Soit $(c_n)_{n \geq 0}$ une des suites construites précédemment. On note $d_n = |c_n|$ pour tout entier $n \geq 0$. Montrer que pour tout entier $n \geq 1$,

$$|n + 2k + a_0 - 1|nd_n \leq \sum_{m=0}^{n-1} (|b_{n-m}| + |k + m| \cdot |a_{n-m}|)d_m.$$

- (b). Montrer que pour tout $\delta \in]0, R[$, il existe un entier n_0 tel que pour tout entier $n \geq n_0$,

$$d_n \delta^n \leq \sup_{0 \leq m \leq n-1} d_m \delta^m.$$

4. (a). Montrer qu'il existe un nombre complexe k et une série entière complexe $c(x) = \sum_{n=0}^{\infty} c_n x^n$ de rayon de convergence $\geq R$, tels que la fonction $y = c(x)x^k$ soit solution de (F) sur $]0, R[$.
- (b). Décrire alors la forme de toutes les solutions de (F) au voisinage de 0 selon les cas rencontrés.

Partie III

Pour p nombres réels distincts $a_1, \dots, a_p$, notons $P(x) = (x - a_1) \dots (x - a_p)$, et considérons deux autres polynômes à coefficients réels Q et R . On s'intéresse à l'équation différentielle linéaire

$$(P(x))^2 y'' + P(x)Q(x)y' + R(x)y = 0. \quad (\Phi)$$

1. Soit $j \in \{1, \dots, p\}$ et soit $\varepsilon > 0$ assez petit pour que $]a_j, a_j + \varepsilon[$ ne contienne aucune autre racine de P .
- (a). Montrer que y est une solution de (Φ) sur $]a_j, a_j + \varepsilon[$ si et seulement si la fonction définie par $z(t) = y(t + a_j)$ est solution sur l'intervalle $]0, \varepsilon[$ de l'équation différentielle

$$t^2 z'' + A_j(t)t z' + B_j(t)z = 0, \quad (F_j)$$

où A_j et B_j sont des fractions rationnelles que l'on précisera.

- (b). Montrer qu'il existe $R_j > 0$ tel que les fonctions A_j et B_j soient développables en série entière sur $] -R_j, R_j[$.
- (c). On appelle valeurs caractéristiques en a_j de l'équation (Φ) les deux racines de l'équation du second degré

$$k^2 + (a_0 - 1)k + b_0 = 0,$$

où $a_0 = A_j(0)$ et $b_0 = B_j(0)$. Exprimer la somme des deux valeurs caractéristiques en a_j à l'aide des polynômes P et Q .

2. (a). Montrer que y est solution de (Φ) sur $]S, +\infty[$, où $S = \max\{0, a_1, \dots, a_p\}$ si et seulement si la fonction z définie par $z(t) = y(1/t)$ est solution sur l'intervalle $]0, 1/S[$ d'une équation différentielle de la forme

$$t^2 z'' + \tilde{A}(t)t z' + \tilde{B}(t)z = 0, \quad (\tilde{\Phi})$$

où $\tilde{A}$ et $\tilde{B}$ sont des fractions rationnelles. Ici, on convient que $1/S = +\infty$ si $S = 0$.

- (b). Trouver une condition nécessaire et suffisante sur les degrés de P , Q et R pour que les fonctions $\tilde{A}$ et $\tilde{B}$ soient développables en série entière au voisinage de 0 (avec $\deg 0 = -\infty$).
- (c). On suppose que cette condition est remplie, et on appelle valeur caractéristique à l'infini de l'équation (Φ) les valeurs caractéristiques en 0 de l'équation $(\tilde{\Phi})$. Exprimer leur somme.
3. Quelle est la somme de toutes les valeurs caractéristiques de l'équation différentielle (Φ) en tous les points a_j et à l'infini ?
4. On considère le polynôme $P(x) = x(x-1)$. Déterminer des polynômes Q et R tels que les valeurs caractéristiques de l'équation différentielle (Φ) correspondante soient respectivement :
- 0 et 0 au point 0 ;
 - 0 et 0 au point 1 ;
 - $1/2$ et $1/2$ à l'infini.
5. On se place dans le cas particulier obtenu à la question précédente. Trouver une solution de (Φ) développable en série entière au voisinage de 1. Existe-t-il une solution (non nulle) définie sur $]0, +\infty[$?

Corrigé.

Partie I

1. (a). Sur l'intervalle $]0, +\infty[$, l'équation (E) est équivalente à

$$y'' = -\frac{a}{x}y' - \frac{b}{x^2}y,$$

c'est-à-dire $X' = AX$, avec $X = \begin{pmatrix} y \\ y' \end{pmatrix}$ et $A = \begin{pmatrix} 0 & 1 \\ -b/x^2 & -a/x \end{pmatrix}$. On peut donc appliquer le théorème de Cauchy-Lipschitz linéaire et en déduire que l'ensemble des solutions réelles de (E) sur $]0, +\infty[$ est un sous-espace de dimension 2 de l'espace vectoriel des fonctions réelles de classe C^2 sur $]0, +\infty[$.

- (b). Comme les coefficients a et b de l'équation différentielles sont réels, une application Φ de $]0, +\infty[$ dans $\mathbb{C}$ en est une solution complexe si et seulement si les fonctions $\Re\Phi$ et $\Im\Phi$ sont solutions. Si φ_1 et φ_2 désignent deux solutions réelles linéairement indépendantes, alors une fonction Φ est solution complexe si et seulement si on peut trouver des coefficients $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ tels que

$$\begin{cases} \Re\Phi = \alpha\varphi_1 + \beta\varphi_2 \\ \Im\Phi = \gamma\varphi_1 + \delta\varphi_2, \end{cases}$$

soit encore

$$\Phi = (\alpha + i\gamma)\varphi_1 + (\beta + i\delta)\varphi_2.$$

Les solutions φ_1 et φ_2 sont également linéairement indépendantes sur $\mathbb{C}$, et le raisonnement ci-dessus montre qu'elles engendrent toutes les solutions complexes. L'ensemble des solutions complexes est donc un espace vectoriel complexe de dimension 2.

2. La fonction monomiale $y = c_n x^n$ est solution si et seulement si $(n(n-1) + an + b)c_n x^n = 0$ pour tout réel $x > 0$. Par linéarité, on en déduit que si l'équation $n^2 + (a-1)n + b = 0$ admet :
- deux solutions distinctes $n_1, n_2 \in \mathbb{N}$, alors l'équation différentielle n'a que des solutions polynomiales, engendrées par les fonctions monomiales x^{n_1} et x^{n_2} ;

- une seule solution $n \in \mathbb{N}$, alors l'ensemble des solutions polynomiales de l'équation différentielle est la droite engendrée par la fonction monomiale x^n ;
 - zéro solutions dans $\mathbb{N}$, alors l'équation différentielle n'admet que la solution polynomiale nulle.
3. On effectue le changement de variable indiqué $x = e^t$. L'équation différentielle (E) est donc équivalente à

$$e^{2t}y''(e^t) + ae^ty'(e^t) + by(e^t) = 0,$$

soit encore, en posant $z(t) = y(e^t)$, équivalente à

$$z'' + (a-1)z' + bz = 0. \quad (\tilde{E})$$

On obtient ainsi une équation linéaire à coefficients constants, dont l'équation caractéristique est $r^2 + (a-1)r + b = 0$, et dont les solutions diffèrent selon le signe du discriminant $(a-1)^2 - 4b$:

- si ce discriminant est strictement positif, l'équation caractéristique admet deux racines distinctes $r_1, r_2 \in \mathbb{R}$ et les solutions réelles de l'équation différentielle ($\tilde{E}$) sont engendrées par les deux fonctions $t \mapsto e^{r_1 t}$ et $t \mapsto e^{r_2 t}$, ce qui fait que les solutions de l'équation différentielle (E) s'écrivent sous la forme $\varphi(x) = \alpha x^{r_1} + \beta x^{r_2}$, pour $x > 0$, avec $\alpha, \beta \in \mathbb{R}$;
- si le discriminant est nul, l'équation caractéristique admet une unique racine $r \in \mathbb{R}$ et les solutions réelles de l'équation différentielle ($\tilde{E}$) sont engendrées par les deux fonctions $t \mapsto e^{rt}$ et $t \mapsto t e^{rt}$, ce qui fait que les solutions de l'équation différentielle (E) s'écrivent sous la forme $\varphi(x) = (\alpha + \beta \log x) x^r$, pour $x > 0$, avec $\alpha, \beta \in \mathbb{R}$;
- si le discriminant est strictement négatif, l'équation caractéristique admet deux racines complexes conjuguées $(r \pm is) \in \mathbb{C}$ et les solutions réelles de l'équation différentielle ($\tilde{E}$) sont engendrées par les deux fonctions $t \mapsto e^{rt} \cos(st)$ et $t \mapsto e^{rt} \sin(st)$, ce qui fait que les solutions de l'équation différentielle (E) s'écrivent sous la forme $\varphi(x) = (\alpha \cos(s \log x) + \beta \sin(s \log x)) x^r$, pour $x > 0$, avec $\alpha, \beta \in \mathbb{R}$.

Partie II

1. (a). Soit $R' > 0$ un réel strictement inférieur à R et au rayon de convergence de la série entière $c(x)$. La fonction $y = c(x)x^k$ est deux fois dérivable sur l'intervalle $]0, R'[,$ et on a

$$xy' = x^k \sum_{n=0}^{\infty} (n+k)c_n x^n \quad \text{et} \quad x^2 y'' = x^k \sum_{n=0}^{\infty} (n+k)(n+k-1)c_n x^n.$$

En injectant ces expressions dans l'équation différentielle (F), divisée par x^k , on est amené à effectuer des produits de Cauchy de séries entières, et on observe finalement que l'équation est équivalente au fait que pour tout entier $n \geq 0$,

$$(n+k)(n+k-1)c_n + \sum_{m=0}^n a_{n-m}(m+k)c_m + \sum_{m=0}^n b_{n-m}c_m = 0,$$

soit encore

$$((n+k)(n+k-1+a_0) + b_0)c_n + \sum_{m=0}^{n-1} ((m+k)a_{n-m} + b_{n-m})c_m = 0.$$

En particulier, pour $n = 0$, et sachant que $c_0 = 1$, on voit que k est solution de l'équation

$$T(k) = k(k-1+a_0) + b_0 = 0.$$

- (b). Si k désigne une solution de cette équation, on peut remplacer b_0 par $-k(k-1+a_0)$ dans la relation qui est au-dessus, et on obtient pour tout entier $n \geq 1$,

$$n(n+2k-1+a_0)c_n + \sum_{m=0}^{n-1} ((m+k)a_{n-m} + b_{n-m})c_m = 0. \quad (P)$$

2. (a). Le nombre complexe k correspond à une racine du polynôme T défini ci-dessus. Le discriminant de ce polynôme est alors le nombre $\Delta = (a_0 - 1)^2 - 4b_0$ défini dans l'énoncé. De plus, le carré de la différence des deux racines de T vaut Δ . Ainsi, comme k est racine, alors $n + k$ ne peut jamais être racine quand n est entier strictement positif et quand Δ n'est pas le carré d'un entier non nul. Ainsi, on a $T(n + k) \neq 0$ pour tout entier $n \geq 1$. Comme $T(n + k)$ est le coefficient devant c_n dans la relation (P), alors on peut toujours exprimer le coefficient c_n en fonction des coefficients précédents $c_0, \dots, c_{n-1}$, et il existe donc une unique suite de nombres complexes $(c_n)_{n \geq 0}$ vérifiant $c_0 = 1$ et la relation de récurrence (P).
- (b). Dans le cas où $\Delta = p^2$, pour un entier $p \geq 1$, alors les racines du polynôme T , sont k et $p + k$, avec $k = \frac{1 - a_0 - p}{2}$. Pour tout $n \in \{1, \dots, p - 1\}$, le coefficient devant c_n dans la relation (P) associée à k , c'est-à-dire $T(n + k)$, est non nul. Par conséquent, partant de $c_0 = 1$, il y a une unique façon de déterminer les coefficients $c_1, \dots, c_{p-1}$ de façon à ce que la relation (P) associée à k soit vérifiée pour $n \in \{1, \dots, p - 1\}$. Ensuite, pour $n = p$, la relation (P) associée à k s'écrit simplement

$$\sum_{m=0}^{p-1} ((m+k)a_{p-m} + b_{p-m})c_m = 0. \quad (L)$$

Cela donne une condition nécessaire et suffisante pour l'existence d'une suite $(c_n)_{n \geq 0}$ vérifiant pour tout n la relation (P) associée à k . En effet, si cette condition est satisfaite, c_p est libre, puis les coefficients $c_{p+1}, c_{p+2}, \dots$ sont déterminés successivement par (P), puisque $T(n + k) = n(n - p) \neq 0$ pour $n > p$. Ils ne sont en général pas nuls. Enfin, la relation (P) associée à $k' = p + k$ et partant de $c'_0 = 1$ conduit toujours à une unique suite $(c'_n)_{n \geq 0}$, puisque les coefficients $T(n + k')$ devant c'_n ne sont jamais nuls.

3. (a). En prenant le module dans la relation (P), en remplaçant par d_n le module $|c_n|$, et en utilisant l'inégalité triangulaire, on obtient comme souhaité pour tout entier $n \geq 1$,

$$|n + 2k + a_0 - 1|nd_n \leq \sum_{m=0}^{n-1} (|b_{n-m}| + |k + m| \cdot |a_{n-m}|)d_m.$$

- (b). Pour $\delta \in]0, R[$, les séries $A = \sum_{n=0}^{\infty} |a_n| \delta^n$ et $B = \sum_{n=0}^{\infty} |b_n| \delta^n$ sont convergentes. En multipliant par δ^n l'inégalité obtenue à la question précédente, on a

$$\begin{aligned} |n + 2k + a_0 - 1|nd_n \delta^n &\leq \sum_{m=0}^{n-1} (|b_{n-m}| \delta^{n-m} + (|k| + |m|) \cdot |a_{n-m}| \delta^{n-m})d_m \delta^m \\ &\leq (B + (|k| + n)A) \sup_{0 \leq m \leq n-1} d_m \delta^m. \end{aligned}$$

Pour n assez grand, on peut diviser par le coefficient devant $d_n \delta^n$ dans le membre de gauche, on a donc

$$d_n \delta^n \leq \frac{B + (|k| + n)A}{|n + 2k + a_0 - 1|n} \sup_{0 \leq m \leq n-1} d_m \delta^m.$$

Le coefficient devant le supremum dans le membre de droite tend vers 0 quand n tend vers l'infini, donc il est inférieur à 1 pour n supérieur à un certain entier n_0 . Ainsi, pour $n \geq n_0$,

$$d_n \delta^n \leq \sup_{0 \leq m \leq n-1} d_m \delta^m.$$

4. (a). Montrons par une récurrence forte sur $n \geq n_0$ que $d_n \delta^n \leq \sup_{0 \leq m \leq n_0-1} d_m \delta^m$. La question précédente indique que cette majoration est vraie pour $n = n_0$. Considérons un entier $n \geq n_0$ pour lequel elle est vraie du rang n_0 au rang n . Au rang $n + 1$, on a alors

$$d_{n+1} \delta^{n+1} \leq \sup_{0 \leq m \leq n} d_m \delta^m = \max \left\{ \sup_{0 \leq m \leq n_0-1} d_m \delta^m, \sup_{n_0 \leq m \leq n} d_m \delta^m \right\} \leq \sup_{0 \leq m \leq n_0-1} d_m \delta^m,$$

toujours d'après la question précédente. On conclut donc grâce au principe de récurrence. En particulier, on observe que la suite de terme général $d_n \delta^n$ est bornée. Cela implique que la série entière complexe $c(x) = \sum_{n=0}^{\infty} c_n x^n$ converge sur l'intervalle $] -\delta, \delta[$ au moins. Cela étant vrai pour tout $\delta \in]0, R[$, on en déduit que cette série entière a un rayon de convergence supérieur ou égal à R . En résumé, il existe une racine k du polynôme T et une suite $(c_n)_{n \geq 0}$ associée via la relation (P) telles que la série entière $c(x) = \sum_{n=0}^{\infty} c_n x^n$ ait un rayon de convergence $\geq R$ et la fonction $y = c(x)x^k$ soit solution de l'équation différentielle (F) sur l'intervalle ouvert $]0, R[$.

(b). Distinguons les racines de l'équation caractéristique.

- Si $\Delta \neq 0$ et si Δ n'est pas le carré d'un entier strictement positif, les racines k_1, k_2 sont distinctes et leur différence n'est pas entière. Les solutions $y_j(x) = x^{k_j} c_j(x)$ construites précédemment, avec $c_j(0) = 1$, sont indépendantes : leur wronskien vérifie

$$y_1 y_2' - y_1' y_2 = x^{k_1+k_2-1} (k_2 - k_1 + O(x)).$$

Elles forment donc une base complexe de solutions sur $]0, R[$.

- Si $\Delta = p^2$, avec $p \geq 1$ entier, notons $k_1 < k_2$ les racines, avec $k_2 - k_1 = p$. La solution $y_2 = x^{k_2} c_2(x)$ existe toujours. Puisque $c_2(0) = 1$, choisissons $0 < \rho < R$ assez petit pour que c_2 ne s'annule pas sur le disque $|z| < \rho$. Sur $]0, \rho[$, cherchons une seconde solution sous la forme $y_1 = \lambda y_2$. En substituant, on obtient

$$x y_2 \lambda'' + (2x y_2' + a(x) y_2) \lambda' = 0.$$

Posons $A(x) = (a(x) - a_0)/x$, prolongée analytiquement en 0. Une solution pour λ' est

$$\lambda'(x) = x^{-p-1} H(x), \quad H(x) = \frac{\exp\left(-\int_0^x A(t) dt\right)}{c_2(x)^2} = \sum_{n=0}^{\infty} h_n x^n,$$

où H est analytique près de 0 et $h_0 = 1$. L'intégration terme à terme sur $]0, \rho[$, en diminuant ρ si nécessaire, donne

$$\lambda(x) = h_p \log x + \sum_{\substack{n \geq 0 \\ n \neq p}} \frac{h_n}{n-p} x^{n-p}.$$

Ainsi, $y_1 = x^{k_1} d(x) + h_p y_2 \log x$, où d est analytique et $d(0) = -1/p$. Comme λ' ne s'annule pas, le wronskien $y_2^2 \lambda'$ est non nul : y_1, y_2 forment une base locale de solutions. Si la condition (L) est satisfaite, on peut aussi construire directement une solution $x^{k_1} c_1(x)$ sans logarithme, avec $c_1(0) = 1$; sinon aucune solution de cette forme n'existe.

- Si $\Delta = 0$, la racine $k = (1 - a_0)/2$ est double. La solution normalisée $y_0 = x^k c(x)$ existe et est unique parmi celles avec $c(0) = 1$. La même réduction donne $\lambda'(x) = x^{-1} H(x)$, avec H analytique et $H(0) = 1$. On obtient donc une seconde solution

$$y_1(x) = y_0(x)(\log x + g(x)),$$

où g est analytique près de 0. Le wronskien est à nouveau non nul ; le terme logarithmique ne peut pas être omis.

Dans chaque cas, les combinaisons linéaires des deux solutions décrites donnent toutes les solutions complexes au voisinage de 0. Pour les solutions réelles, on prend des coefficients réels lorsque les racines sont réelles ; lorsque les racines sont conjuguées non réelles, les parties réelle et imaginaire d'une solution constituent une base réelle.

1. (a). Quitte à réindexer, on peut supposer $a_1 < \dots < a_p$. On considère $j \in \{1, \dots, p\}$ et $\varepsilon > 0$, avec $\varepsilon < a_{j+1} - a_j$ lorsque $j < p$; lorsque $j = p$, tout $\varepsilon > 0$ convient. Alors, le polynôme P ne s'annule pas sur l'intervalle $]a_j, a_j + \varepsilon[$, et l'ensemble des solutions de l'équation (Φ) sur cet intervalle est un espace vectoriel de dimension 2. Si y désigne une fonction deux fois dérivable définie sur $]a_j, a_j + \varepsilon[$, et si on pose $z(t) = y(t + a_j)$ pour $t \in]0, \varepsilon[$, alors y est solution de (Φ) si et seulement si

$$(P(t + a_j))^2 y''(t + a_j) + P(t + a_j)Q(t + a_j) y'(t + a_j) + R(t + a_j) y(t + a_j) = 0,$$

c'est-à-dire si

$$t^2 z''(t) \prod_{k \neq j} (t + a_j - a_k)^2 + t z'(t) Q(t + a_j) \prod_{k \neq j} (t + a_j - a_k) + R(t + a_j) z(t) = 0,$$

soit encore

$$t^2 z'' + \frac{Q(t + a_j)}{\prod_{k \neq j} (t + a_j - a_k)} t z' + \frac{R(t + a_j)}{\prod_{k \neq j} (t + a_j - a_k)^2} z = 0.$$

En notant les fractions rationnelles $A_j(t) = \frac{Q(t + a_j)}{\prod_{k \neq j} (t + a_j - a_k)}$ et $B_j(t) = \frac{R(t + a_j)}{\prod_{k \neq j} (t + a_j - a_k)^2}$, on conclut qu'une

fonction y deux fois dérivable définie sur $]a_j, a_j + \varepsilon[$ est solution de l'équation différentielle (Φ) si et seulement si la fonction définie par $z(t) = y(t + a_j)$, pour $t \in]0, \varepsilon[$, est solution de l'équation différentielle

$$t^2 z'' + A_j(t) t z' + B_j(t) z = 0. \quad (F_j)$$

- (b). Les deux fonctions A_j et B_j sont des fractions rationnelles qui n'admettent pas 0 comme pôle. Elles sont donc développables en séries entières sur un intervalle de la forme $] -R_j, R_j[$, pour un certain réel $R_j > 0$. En effet, on commence par décomposer la fraction rationnelle en éléments simples, et comme 0 n'est pas pôle, on obtient des éléments de la forme

$$\frac{1}{(t - a)^\nu} = \frac{(-1)^\nu}{a^\nu} \frac{1}{(1 - t/a)^\nu} = \frac{(-1)^\nu}{a^\nu} \sum_{n=0}^{\infty} \binom{n + \nu - 1}{\nu - 1} \left(\frac{t}{a}\right)^n,$$

où le développement est valable pour $|t| < |a|$. En combinant tous les éléments simples issus de chacune des fractions rationnelles A_j et B_j , on obtient le résultat.

- (c). La somme des deux valeurs caractéristiques en a_j de l'équation (Φ) est égale à la somme des racines du polynôme $k^2 + (a_0 - 1)k + b_0$, où $a_0 = A_j(0)$ et $b_0 = B_j(0)$, c'est-à-dire

$$1 - a_0 = 1 - A_j(0) = 1 - \frac{Q(a_j)}{\prod_{k \neq j} (a_j - a_k)} = 1 - \frac{Q(a_j)}{P'(a_j)}.$$

2. (a). Comme le polynôme P ne s'annule pas sur l'intervalle $]S, +\infty[$, alors l'ensemble des solutions de l'équation différentielle (Φ) sur cet intervalle est un espace vectoriel de dimension 2. De plus, si y désigne une fonction deux fois dérivable définie sur $]S, +\infty[$, et si on pose $z(t) = y(1/t)$ pour $t \in]0, 1/S[$, avec la convention que $1/S = +\infty$ si $S = 0$, alors y est solution de (Φ) si et seulement si

$$t^2 z'' + \left(2 - \frac{Q(1/t)}{t P(1/t)}\right) t z' + \frac{R(1/t)}{t^2 P(1/t)^2} z = 0.$$

En notant les fractions rationnelles $\tilde{A}(t) = 2 - \frac{Q(1/t)}{t P(1/t)}$ et $\tilde{B}(t) = \frac{R(1/t)}{t^2 P(1/t)^2}$, on conclut qu'une fonction y deux fois dérivable définie sur $]S, +\infty[$ est solution de l'équation (Φ) si et seulement si la fonction définie par $z(t) = y(1/t)$, pour $t \in]0, 1/S[$, est solution de l'équation différentielle

$$t^2 z'' + \tilde{A}(t) t z' + \tilde{B}(t) z = 0. \quad (\tilde{\Phi})$$

- (b). Les fractions rationnelles $\tilde{A}$ et $\tilde{B}$ sont développables en séries entières au voisinage de 0 si et seulement si elles n'admettent pas 0 comme pôle. Si Q ou R est nul, le terme correspondant dans ces fractions est nul et ne crée aucun pôle. Pour les polynômes non nuls, on peut écrire $P(1/t) = \frac{P_1(t)}{t^p}$, où p est le degré de P et P_1 est un polynôme tel que $P_1(0) \neq 0$, et de façon similaire $Q(1/t) = \frac{Q_1(t)}{t^q}$ et $R(1/t) = \frac{R_1(t)}{t^r}$. Il en ressort que

$$\tilde{A}(t) = 2 - \frac{Q_1(t)/t^q}{t P_1(t)/t^p} = 2 - \frac{Q_1(t)}{P_1(t)} t^{p-q-1} \quad \text{et} \quad \tilde{B}(t) = \frac{R(1/t)}{t^2 P(1/t)^2} = \frac{R_1(t)}{P_1(t)^2} t^{2p-r-2}.$$

Ainsi, les fonctions $\tilde{A}$ et $\tilde{B}$ sont développables en séries entières au voisinage de 0 si et seulement si les puissances sur t , c'est-à-dire $p - q - 1$ et $2p - r - 2$, sont positives ou nulles, soit

$$\deg(P) - 1 \geq \max \left\{ \deg(Q), \frac{\deg(R)}{2} \right\}.$$

- (c). On suppose que cette dernière condition est satisfaite. Les valeurs caractéristiques en l'infini de l'équation différentielle (Φ) sont les valeurs caractéristiques en 0 de l'équation différentielle $(\tilde{\Phi})$, c'est-à-dire les deux racines de l'équation du second degré

$$k^2 + (a_0 - 1)k + b_0 = 0,$$

où $a_0 = \tilde{A}(0)$ et $b_0 = \tilde{B}(0)$. Leur somme est alors égale à $1 - a_0$. Si $Q = 0$, cette somme vaut -1 . Sinon,

$$a_0 = \tilde{A}(0) = 2 - \frac{Q_1(0)}{P_1(0)} \lim_{t \rightarrow 0} t^{p-q-1}.$$

Comme le coefficient dominant de P est égal à 1, on a $P_1(0) = 1$. On en conclut que la somme des valeurs caractéristiques en l'infini de l'équation différentielle (Φ) vaut

$$1 - a_0 = \begin{cases} Q_1(0) - 1 & \text{si } p = q + 1 \\ -1 & \text{si } p > q + 1. \end{cases}$$

3. La somme de toutes les valeurs caractéristiques de l'équation différentielle (Φ) en tous les points a_j est égale à

$$\sum_{j=1}^p \left(1 - \frac{Q(a_j)}{P'(a_j)} \right) = p - \sum_{j=1}^p \frac{Q(a_j)}{P'(a_j)}.$$

Si $Q = 0$, la somme cherchée vaut directement $p - 1$. Sinon, comme la condition $p \geq q + 1$ est satisfaite, et comme P admet les a_j pour racines distinctes, on peut décomposer la fraction rationnelle Q/P comme suit :

$$\frac{Q(x)}{P(x)} = \sum_{j=1}^p \frac{Q(a_j)}{P'(a_j)(x - a_j)}.$$

En multipliant par x les deux membres de cette identité et en faisant ensuite tendre x vers l'infini, le membre de droite tend vers $\sum_{j=1}^p \frac{Q(a_j)}{P'(a_j)}$ et le membre de gauche vers 0 si $p > q + 1$, et vers $Q_1(0)$ si $p = q + 1$. Ainsi,

$$\sum_{j=1}^p \frac{Q(a_j)}{P'(a_j)} = \begin{cases} Q_1(0) & \text{si } p = q + 1 \\ 0 & \text{si } p > q + 1. \end{cases}$$

De plus, la somme de toutes les valeurs caractéristiques de l'équation différentielle (Φ) en l'infini est égale à $Q_1(0) - 1$ dans le premier cas, et à -1 dans le second. Finalement, dans tous les cas, la somme de toutes les valeurs caractéristiques de l'équation différentielle (Φ) en tous les points a_j et à l'infini vaut $p - 1$.

4. On suppose que $P(x) = x(x-1)$. Alors, les valeurs caractéristiques en $a_1 = 0$ de l'équation différentielle (Φ) sont les racines de l'équation du second degré

$$k^2 + (a_0 - 1)k + b_0 = 0,$$

avec $a_0 = A_1(0) = -Q(0)$ et $b_0 = B_1(0) = R(0)$. Pour que ces deux valeurs caractéristiques soient égales à 0, il faut et il suffit que 0 soit racine double de l'équation du second degré

$$k^2 - (Q(0) + 1)k + R(0) = 0,$$

ce qui équivaut aux conditions $Q(0) = -1$ et $R(0) = 0$. De même, les valeurs caractéristiques en $a_2 = 1$ de l'équation différentielle (Φ) sont les racines de l'équation du second degré

$$k^2 + (a_0 - 1)k + b_0 = 0,$$

avec $a_0 = A_2(0) = Q(1)$ et $b_0 = B_2(0) = R(1)$. Pour que ces deux valeurs caractéristiques soient égales à 0, il faut et il suffit que 0 soit racine double de l'équation du second degré

$$k^2 + (Q(1) - 1)k + R(1) = 0,$$

ce qui équivaut aux conditions $Q(1) = 1$ et $R(1) = 0$. Enfin, d'après les conditions sur les degrés des polynômes, on voit que Q est nécessairement de degré 1, et que R est de degré au plus 2. D'après ce qui précède, on voit d'abord que $Q(x) = 2x - 1$. Les valeurs caractéristiques en l'infini de l'équation différentielle (Φ) sont alors les racines de l'équation du second degré

$$k^2 + (a_0 - 1)k + b_0 = 0,$$

avec $a_0 = \tilde{A}(0) = 2 - \lim_{t \rightarrow 0} \frac{Q(1/t)}{t P(1/t)} = 0$ et $b_0 = \tilde{B}(0) = \lim_{t \rightarrow 0} \frac{R(1/t)}{t^2 P(1/t)^2} = \lim_{t \rightarrow 0} t^2 R(1/t)$. Pour que ces deux valeurs caractéristiques soient égales à $1/2$, il faut et il suffit que $1/2$ soit racine double de l'équation du second degré

$$k^2 - k + \lim_{t \rightarrow 0} t^2 R(1/t) = 0,$$

ce qui équivaut à la condition $\lim_{t \rightarrow 0} t^2 R(1/t) = 1/4$. Cette condition correspond au fait que R est de degré 2, avec pour coefficient dominant $1/4$. Sachant que $R(0) = R(1) = 0$, on en déduit que $R(x) = x(x-1)/4$.

5. Dans le cas de la question précédente, l'équation (Φ) s'écrit

$$x^2(x-1)^2 y'' + x(x-1)(2x-1) y' + \frac{x(x-1)}{4} y = 0,$$

dont l'ensemble des solutions inclut en particulier les solutions de l'équation

$$(x(x-1)y')' + \frac{y}{4} = 0.$$

Pour en trouver une solution développable en série entière au voisinage de $a_2 = 1$, on s'appuie sur ce qui précède et on se ramène au fait que $z(t) = y(t+1)$ est solution de

$$(t+1)t z'' + (2t+1) z' + \frac{z}{4} = 0.$$

En cherchant z sous la forme d'une série entière $z(t) = \sum_{n=0}^{\infty} c_n t^n$, on voit que cela équivaut à la condition $(n+1)^2 c_{n+1} = -\left(n + \frac{1}{2}\right)^2 c_n$ pour tout entier $n \geq 0$. En prenant par exemple $c_0 = 1$, on trouve $c_n = (-1)^n \left(\frac{(2n)!}{(2^n n!)^2}\right)^2$.

La série entière correspondante à un rayon de convergence égal à 1, et donc on obtient une solution sur l'intervalle $]0, 2[$ en considérant

$$y(x) = \sum_{n=0}^{\infty} (-1)^n \left(\frac{(2n)!}{(2^n n!)^2}\right)^2 (x-1)^n.$$

Enfin, le théorème de Cauchy-Lipschitz s'applique sur l'intervalle $]1, +\infty[$, puisque P ne s'y annule pas, et donc on peut prolonger la solution précédente, définie sur $]0, 2[$, en une solution non nulle définie sur $]0, +\infty[$ tout entier.

Espaces de Hilbert

Références : [7, 8, 12, 14].

Exercice 12.1 (Théorème de Lax-Milgram)

Leçon : 208. Référence : [8].

On dit qu'une forme bilinéaire a de $H \times H$ dans $\mathbb{R}$ est *continue* s'il existe $C > 0$ tel que $|a(u, v)| \leq C \|u\| \|v\|$ pour tous $u, v \in H$; et *coercive* s'il existe $\alpha > 0$ tel que $a(v, v) \geq \alpha \|v\|^2$ pour tout $v \in H$. On considère un espace de Hilbert réel H et une forme bilinéaire a continue et coercive sur H .

1. Montrer qu'il existe un unique opérateur linéaire continu $T \in \mathcal{L}(H)$ tel que $a(u, v) = \langle Tu, v \rangle$ pour tous $u, v \in H$.
2. Montrer que l'opérateur T est injectif.
3. Montrer que l'image de T est fermée dans H .
4. Montrer que l'image de T est dense dans H .
5. En déduire que l'opérateur T est bijectif, puis que pour toute forme linéaire $\varphi \in H'$, il existe un unique vecteur $u \in H$ tel que $a(u, v) = \varphi(v)$ pour tout vecteur $v \in H$.
6. On suppose désormais que la forme bilinéaire a est symétrique. Montrer que le vecteur u vérifie

$$\frac{1}{2}a(u, u) - \varphi(u) = \min_{v \in H} \left(\frac{1}{2}a(v, v) - \varphi(v) \right).$$

7. Montrer réciproquement que si un vecteur $u \in H$ est solution du problème de minimisation ci-dessus, alors il vérifie $a(u, v) = \varphi(v)$ pour tout $v \in H$.

Corrigé.

1. Pour tout vecteur $u \in H$, l'application $v \mapsto a(u, v)$ est une forme linéaire sur H . Le théorème de Riesz implique l'existence d'un unique vecteur $Tu \in H$ tel que $a(u, v) = \langle Tu, v \rangle$ pour tout $v \in H$. Il est clair qu'on définit ainsi un unique opérateur linéaire continu $T \in \mathcal{L}(H)$.
2. La coercivité de a et l'inégalité de Cauchy-Schwarz impliquent la minoration $\|Tv\| \geq \alpha \|v\|$ pour tout vecteur $v \in H$, puisqu'en effet

$$\alpha \|v\|^2 \leq a(v, v) = \langle Tv, v \rangle \leq \|Tv\| \|v\|.$$

Cela impose en particulier l'injectivité de T .

3. Supposons que $w = \lim_{n \rightarrow \infty} w_n$ est limite d'une suite d'éléments $w_n \in \text{im } T$. On peut écrire $w_n = Tv_n$ pour tout n , et il résulte comme ci-dessus de la coercivité et de l'inégalité de Cauchy-Schwarz que

$$\alpha \|v_n - v_m\| \leq \|T(v_n - v_m)\| = \|w_n - w_m\| \xrightarrow{m, n \rightarrow \infty} 0.$$

La suite (v_n) est donc une suite de Cauchy dans H , qui est complet. Elle admet donc une limite, notée v . On a clairement $w = Tv \in \text{im } T$. Finalement, l'image de l'opérateur T est fermée dans H .

4. Si un vecteur v appartient à $(\text{im } T)^\perp$, alors on a $0 = \langle Tu, v \rangle = a(u, v)$ pour tout $u \in H$. En particulier, toujours par coercivité, on a

$$\alpha \|v\|^2 \leq a(v, v) = \langle Tv, v \rangle = 0.$$

Ainsi $v = 0$, ce qui prouve que l'orthogonal de $\text{im } T$ se réduit à $\{0\}$. Le critère de densité classique dans les espaces de Hilbert permet de conclure que l'image de T est dense dans H .

5. On déduit des arguments précédents que l'opérateur T est une bijection de H dans lui-même. Par conséquent, pour $\varphi \in H'$, le théorème de Riesz fournit un unique représentant $w \in H$. Si on pose $u = T^{-1}w$, alors pour tout vecteur $v \in H$, on a

$$\varphi(v) = \langle w, v \rangle = \langle Tu, v \rangle = a(u, v).$$

Cela prouve l'existence de u . Son unicité est claire puisque si u_1 et u_2 conviennent, on a $a(u_1, v) = \varphi(v) = a(u_2, v)$ pour tout $v \in H$, et donc $a(u_1 - u_2, v) = 0$ par bilinéarité. À nouveau par coercivité, on a

$$\alpha \|u_1 - u_2\|^2 \leq a(u_1 - u_2, u_1 - u_2) = 0,$$

ce qui donne $u_1 = u_2$.

6. Supposons que la forme bilinéaire a est symétrique, et écrivons alors que pour $u, h \in H$,

$$\frac{1}{2}a(u+h, u+h) - \varphi(u+h) = \frac{1}{2}a(u, u) - \varphi(u) + \frac{1}{2}a(h, h) - \varphi(h) + a(u, h).$$

Si le vecteur u représente la forme linéaire φ pour la forme bilinéaire a , on a $a(u, h) = \varphi(h)$. De plus, par coercivité, on a clairement $a(h, h) \geq 0$. Ainsi, pour tout $h \in H$,

$$\frac{1}{2}a(u+h, u+h) - \varphi(u+h) \geq \frac{1}{2}a(u, u) - \varphi(u).$$

Comme le membre de gauche coïncide avec celui de droite pour $h = 0$, on en déduit que u est solution du problème de minimisation.

7. Réciproquement, si u est solution du problème de minimisation, on a pour tout $h \in H$,

$$\frac{1}{2}a(h, h) - \varphi(h) + a(u, h) \geq 0.$$

En remplaçant h par son opposé, on obtient aussi

$$\frac{1}{2}a(-h, -h) - \varphi(-h) + a(u, -h) \geq 0.$$

En utilisant la linéarité et en regroupant ces deux inégalités, on en déduit que pour tout $h \in H$,

$$|a(u, h) - \varphi(h)| \leq \frac{1}{2}a(h, h).$$

Enfin, en remplaçant h par εh , avec $\varepsilon > 0$, cela donne

$$|a(u, h) - \varphi(h)| = \frac{1}{\varepsilon} |a(u, \varepsilon h) - \varphi(\varepsilon h)| \leq \frac{1}{2\varepsilon} a(\varepsilon h, \varepsilon h) = \frac{\varepsilon}{2} a(h, h).$$

En faisant $\varepsilon \rightarrow 0$, on conclut que $a(u, h) = \varphi(h)$ pour tout $h \in H$.

Exercice 12.2 (Théorème ergodique de von Neumann)

Leçon : 208. Référence : [7].

On considère un espace de Hilbert H et un opérateur linéaire continu $U \in \mathcal{L}(H)$ sur l'espace H dont la norme d'opérateur vérifie $\|U\| \leq 1$. On désigne par P le projecteur orthogonal sur $\text{Ker}(\text{Id} - U)$. Le but de l'exercice est de montrer qu'au sens de la topologie de H , on a pour tout $x \in H$,

$$\frac{1}{n} \sum_{k=0}^{n-1} U^k x \xrightarrow{n \rightarrow \infty} Px. \quad (12.1)$$

1. On note F l'ensemble des $x \in H$ qui vérifient (12.1). Montrer que F est un sous-espace vectoriel fermé de H .
2. Montrer qu'un vecteur $x \in H$ vérifie $Ux = x$ si et seulement si $\langle Ux, x \rangle = \|x\|^2$.
3. En déduire que les opérateurs $\text{Id} - U$ et $\text{Id} - U^*$ possèdent le même noyau.
4. Montrer que F contient $\text{Ker}(\text{Id} - U)$ et $\overline{\mathfrak{F}(\text{Id} - U)}$.
5. Montrer que $(\mathfrak{F}(\text{Id} - U))^\perp$ coïncide avec $\text{Ker}(\text{Id} - U^*)$.
6. En déduire que $F = H$.

7. *Application aux systèmes dynamiques.* On considère un espace mesuré $(X, \mathcal{B}, \mu)$, avec μ une mesure de probabilité, et une application $T : X \rightarrow X$ mesurable telle que μ est T -invariante (i.e. $\mu \circ T^{-1} = \mu$) et T -ergodique (i.e. les seules fonctions $f \in L^2$ avec $f \circ T = f$ sont les constantes).
- (a). Transcrire le résultat précédent quand $H = L^2$ et U est défini par $Uf = f \circ T$ pour $f \in L^2$.
- (b). Traiter le cas particulier où $X = \mathbb{R}/2\pi\mathbb{Z}$, $\mu(d\theta) = d\theta/(2\pi)$ et T est la rotation d'angle α tel que $\alpha/(2\pi) \notin \mathbb{Q}$.

Corrigé.

1. Il est clair que F est un sous-espace vectoriel de H . Pour montrer qu'il est fermé, notons (x_q) une suite d'éléments de F qui converge vers un vecteur x . Comme les opérateurs U et P sont de norme au plus 1, on a

$$\begin{aligned} \left\| \frac{1}{n} \sum_{k=0}^{n-1} U^k x - Px \right\| &\leq \left\| \frac{1}{n} \sum_{k=0}^{n-1} U^k x - \frac{1}{n} \sum_{k=0}^{n-1} U^k x_q \right\| + \left\| \frac{1}{n} \sum_{k=0}^{n-1} U^k x_q - Px_q \right\| + \|Px_q - Px\| \\ &\leq 2 \|x_q - x\| + \left\| \frac{1}{n} \sum_{k=0}^{n-1} U^k x_q - Px_q \right\|. \end{aligned}$$

Comme chaque x_q appartient à F , on a en faisant tendre n vers l'infini, puis q vers l'infini,

$$\limsup_{n \rightarrow \infty} \left\| \frac{1}{n} \sum_{k=0}^{n-1} U^k x - Px \right\| \leq 2 \|x_q - x\| \xrightarrow{q \rightarrow \infty} 0.$$

Cela prouve bien que x appartient à F , et donc que le sous-espace vectoriel F est fermé.

2. Il est clair que si un vecteur $x \in H$ vérifie $Ux = x$, alors $\langle Ux, x \rangle = \langle x, x \rangle = \|x\|^2$. Pour la réciproque, on observe que si $\langle Ux, x \rangle = \|x\|^2$, alors on a

$$|\langle Ux, x \rangle| = \|x\|^2 \geq \|Ux\| \|x\|,$$

puisque l'opérateur U est de norme au plus 1. On est donc dans le cas d'égalité de l'inégalité de Cauchy-Schwarz. Par conséquent, les vecteurs Ux et x sont colinéaires. On peut clairement supposer que x est non nul. Ainsi, il existe un scalaire non nul λ tel que $Ux = \lambda x$. L'hypothèse de départ sur x conduit à $\lambda \|x\|^2 = \langle \lambda x, x \rangle = \|x\|^2$, ce qui donne $\lambda = 1$, et finalement $Ux = x$.

3. Le résultat de la question précédente s'applique également à l'opérateur adjoint de U puisque $\|U^*\| = \|U\| \leq 1$. Par conséquent, pour tout vecteur $x \in H$, on a

$$Ux = x \iff \langle Ux, x \rangle = \|x\|^2 \iff \langle x, U^*x \rangle = \|x\|^2 \iff U^*x = x.$$

On en conclut que les opérateurs $\text{Id} - U$ et $\text{Id} - U^*$ possèdent bien le même noyau.

4. D'une part, il est clair que si un vecteur $x \in H$ vérifie $Ux = x$, alors x appartient à F . Le sous-espace vectoriel F contient donc $\ker(\text{Id} - U)$. D'autre part, il contient également l'image $\text{im}(\text{Id} - U)$, et donc son adhérence, puisque F est fermé. En effet, si un vecteur s'écrit sous la forme $y = x - Ux$, alors

$$\frac{1}{n} \sum_{k=0}^{n-1} U^k y = \frac{1}{n} \sum_{k=0}^{n-1} U^k x - \frac{1}{n} \sum_{k=0}^{n-1} U^{k+1} x = \frac{1}{n} \sum_{k=0}^{n-1} U^k x - \frac{1}{n} \sum_{k=1}^n U^k x = \frac{1}{n} (x - U^n x) \xrightarrow{n \rightarrow \infty} 0,$$

où la limite est nulle du fait que $\|U\| \leq 1$. Cependant, on a aussi $Py = 0$ quand $y \in \text{im}(\text{Id} - U)$. En effet, on a grâce au résultat de la question précédente,

$$\text{im}(\text{Id} - U) \subset (\ker((\text{Id} - U)^*))^\perp = (\ker(\text{Id} - U^*))^\perp = (\ker(\text{Id} - U))^\perp = \ker P.$$

Tout vecteur y de $\text{im}(\text{Id} - U)$ appartient donc bien à F .

5. Tout vecteur $x \in H$ appartient à $(\text{im}(\text{Id} - U))^\perp$ si et seulement si $\langle x, z - Uz \rangle = 0$ pour tout $z \in H$, c'est-à-dire $\langle x - U^*x, z \rangle = 0$ pour tout $z \in H$, ce qui équivaut à $x - U^*x = 0$, soit encore $x \in \ker(\text{Id} - U^*)$. On en déduit que $(\text{im}(\text{Id} - U))^\perp = \ker(\text{Id} - U^*)$.

6. D'après ce qui précède, on a $(\text{im}(\text{Id} - U))^\perp = \ker(\text{Id} - U^*) = \ker(\text{Id} - U)$. De plus, le sous-espace vectoriel F contient à la fois $\ker(\text{Id} - U)$ et $\overline{\text{im}(\text{Id} - U)}$. Finalement,

$$F \supset \ker(\text{Id} - U) + \overline{\text{im}(\text{Id} - U)} = (\text{im}(\text{Id} - U))^\perp + \overline{\text{im}(\text{Id} - U)} = H,$$

puisque la dernière identité est vraie pour tout sous-espace vectoriel de H . Ainsi, $F = H$.

7. (a). On vérifie bien que l'opérateur U est de norme (au plus) 1 puisque la mesure μ est T -invariante. En effet, pour toute fonction $f \in L^2$, on a

$$\|Uf\|^2 = \int_X |f \circ T|^2 d\mu = \int_X |f|^2 d\mu = \|f\|^2.$$

De plus, une fonction $f \in L^2$ vérifie $Uf = f$ si et seulement si $f \circ T = f$, ce qui par ergodicité signifie que f est constante. Ainsi, le noyau de l'opérateur $\text{Id} - U$ est l'ensemble des fonctions constantes, et le projecteur orthogonal P sur ce noyau est donc donné par

$$Pf = \int_X f d\mu, \quad f \in L^2.$$

Le théorème ergodique de Von Neumann s'applique bien à la situation considérée, et implique que pour toute fonction $f \in L^2$, on a

$$\frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k \xrightarrow[n \rightarrow \infty]{L^2} \int_X f d\mu.$$

- (b). Il est clair que, dans $X = \mathbb{R}/2\pi\mathbb{Z}$, la mesure de Lebesgue normalisée $\mu(d\theta) = \frac{d\theta}{2\pi}$ est invariante par la rotation d'angle α tel que $\frac{\alpha}{2\pi} \notin \mathbb{Q}$, c'est-à-dire l'application $T : \theta \mapsto \theta + \alpha$. Il y a de plus ergodicité. En effet, pour toute fonction $f \in L^2$ et tout entier $p \in \mathbb{Z}$, le coefficient de Fourier d'indice p de f vérifie

$$c_p(f \circ T) = \int_X f(\theta + \alpha) e^{-ip\theta} \frac{d\theta}{2\pi} = \int_X f(\varphi) e^{-ip(\varphi - \alpha)} \frac{d\varphi}{2\pi} = e^{ip\alpha} c_p(f).$$

Dès lors, si la fonction f vérifie $f \circ T = f$, on a nécessairement $(1 - e^{ip\alpha})c_p(f) = 0$ pour tout $p \in \mathbb{Z}$. Puisque $\frac{\alpha}{2\pi}$ est irrationnel, il en ressort que $c_p(f) = 0$ pour tout entier relatif $p \neq 0$, ce qui implique que f est constante presque partout. Le théorème ergodique de Von Neumann permet de conclure que pour toute fonction $f \in L^2$, on a

$$\frac{1}{n} \sum_{k=0}^{n-1} f(\cdot + k\alpha) \xrightarrow[n \rightarrow \infty]{L^2} \int_X f(\theta) \mu(d\theta).$$

Exercice 12.3 (Fonctions d'Hermite)

Leçons : 201, 265. Référence : [7].

Pour tout entier $n \in \mathbb{N}$, on considère la fonction d'Hermite $h_n(t) = e^{t^2/2} \frac{d^n}{dt^n}(e^{-t^2})$.

1. Montrer que les fonctions h_n , pour $n \in \mathbb{N}$, sont orthogonales dans $L^2(\mathbb{R})$.
2. Montrer que la famille $\{h_n, n \in \mathbb{N}\}$ est totale dans $L^2(\mathbb{R})$.
3. Montrer que les fonctions h_n , pour $n \in \mathbb{N}$, sont vecteurs propres de la transformée de Fourier.

Corrigé.

1. Notons H_n le polynôme d'Hermite défini par $H_n(t) = e^{t^2} \frac{d^n}{dt^n}(e^{-t^2})$, de sorte que la fonction d'Hermite donnée dans l'énoncé s'écrit $h_n(t) = H_n(t) e^{-t^2/2}$. Pour $n > m$, on considère le produit scalaire

$$\langle h_n, h_m \rangle = \int_{-\infty}^{+\infty} e^{t^2/2} \frac{d^n}{dt^n}(e^{-t^2}) H_m(t) e^{-t^2/2} dt = \int_{-\infty}^{+\infty} \frac{d^n}{dt^n}(e^{-t^2}) H_m(t) dt.$$

En procédant à $m + 1$ intégrations par parties, on observe que cette dernière intégrale prend les formes successives

$$-\int_{-\infty}^{+\infty} \frac{d^{n-1}}{dt^{n-1}}(e^{-t^2}) H'_m(t) dt = \dots = (-1)^{m+1} \int_{-\infty}^{+\infty} \frac{d^{n-m-1}}{dt^{n-m-1}}(e^{-t^2}) H_m^{(m+1)}(t) dt.$$

Comme H_m est un polynôme de degré m , la dérivée $H_m^{(m+1)}$ est nulle, et donc l'intégrale aussi. D'où $\langle h_n, h_m \rangle = 0$. La famille de fonctions $\{h_n, n \geq 0\}$ est donc orthogonale dans $L^2(\mathbb{R})$.

2. Pour montrer que cette famille est totale, considérons une fonction $f \in L^2(\mathbb{R})$ et supposons qu'elle est orthogonale à toutes les fonctions h_n . On rappelle que $h_n(t) = H_n(t) e^{-t^2/2}$, avec H_n un polynôme de degré n . Cela implique que la fonction f est orthogonale à toutes les fonctions de la forme $t^n e^{-t^2/2}$, pour $n \in \mathbb{N}$. On considère la fonction g définie sur le plan complexe par

$$g(z) = \int_{-\infty}^{+\infty} f(t) e^{-t^2/2} e^{tz} dt.$$

On montre à l'aide du théorème d'holomorphie des intégrales à paramètre, et de l'inégalité de Cauchy-Schwarz, que la fonction g est holomorphe sur le plan $\mathbb{C}$ tout entier. De plus, ses dérivées successives vérifient $g^{(n)}(0) = 0$, pour tout $n \in \mathbb{N}$. On en déduit que la fonction g est nulle. En considérant cette fonction sur l'axe imaginaire pur, on conclut que la transformée de Fourier de $t \mapsto f(t) e^{-t^2/2}$ est nulle. Ainsi, la fonction f est nulle presque partout. Finalement, la famille de fonctions $\{h_n, n \geq 0\}$ est totale.

3. On commence par observer par le calcul que $h_0(t) = e^{-t^2/2}$ est vecteur propre de la transformée de Fourier $\mathcal{F}$. Plus précisément, on a $\mathcal{F}(h_0) = \sqrt{2\pi} h_0$ car pour tout nombre réel ξ ,

$$\int_{-\infty}^{+\infty} e^{-t^2/2} e^{-it\xi} dt = \sqrt{2\pi} e^{-\xi^2/2}.$$

On considère plus généralement l'espace vectoriel E_n formé des fonctions de la forme $P_n h_0$, où P_n est une fonction polynomiale de degré au plus n à coefficients dans $\mathbb{C}$. On vient de voir que E_0 est stable par $\mathcal{F}$. Ce résultat reste vrai pour $n \geq 0$ quelconque. En effet, comme la transformée de Fourier échange multiplication par un polynôme et dérivations successives, si P_n est une fonction polynomiale de degré au plus n à coefficients dans $\mathbb{C}$, alors la transformée de Fourier de la fonction $P_n h_0$ est une combinaison linéaire des dérivées d'ordre au plus n de la fonction h_0 . Cependant, $h_0^{(k)}(t) = 2^{-k/2} H_k\left(\frac{t}{\sqrt{2}}\right) h_0(t)$ pour tous $t \in \mathbb{R}$ et $k \in \mathbb{N}$. Ainsi, $\mathcal{F}(P_n h_0)$ est le produit de la fonction h_0 par une fonction polynomiale de degré au plus n . L'espace vectoriel E_n est donc stable par $\mathcal{F}$. Comme $\mathcal{F}/\sqrt{2\pi}$ est un opérateur unitaire et E_n est de dimension finie, elle laisse donc aussi stable l'espace $E_n^\perp$. Pour $n \geq 1$, il en ressort que $E_n \cap E_{n-1}^\perp = \text{Vect}\{h_n\}$ est stable par $\mathcal{F}$, donc que la fonction h_n est vecteur propre de la transformée de Fourier.

Exercice 12.4 (Théorie de Paley-Wiener et théorème d'échantillonnage de Shannon)

Leçons : 201, 208, 250. Références : [12, 14].

On désigne par sinc la fonction sinus cardinal (définie par $\text{sinc } t = \sin t/t$ pour $t \neq 0$ et par $\text{sinc } 0 = 1$) et par τ_a l'opérateur de translation par a de la variable (défini par $\tau_a f(t) = f(t - a)$). On considère aussi l'espace de Paley-Wiener

$$PW(\mathbb{R}) = \{f \in L^2(\mathbb{R}) \mid \text{supp } \hat{f} \subset [-1, 1]\}.$$

Ici, $\text{supp } \hat{f}$ est le support essentiel : le complémentaire du plus grand ouvert sur lequel $\hat{f}$ est nulle presque partout. Ainsi, la condition affichée signifie $\hat{f} = 0$ presque partout hors de $[-1, 1]$. On identifie chaque élément de $PW(\mathbb{R})$ à son représentant continu $f(x) = (2\pi)^{-1} \int_{-1}^1 \hat{f}(t) e^{itx} dt$, ce qui donne un sens aux valeurs $f(n\pi)$.

1. Montrer que les fonctions $t \mapsto \mathbf{1}_{[-1,1]}(t) e^{in\pi t} / \sqrt{2}$, pour $n \in \mathbb{Z}$, forment une base hilbertienne de $L^2([-1, 1])$.
2. En déduire que les fonctions $\tau_{n\pi} \text{sinc} / \sqrt{\pi}$, pour $n \in \mathbb{Z}$, forment une base hilbertienne de $PW(\mathbb{R})$.
3. Pour f dans $PW(\mathbb{R})$, calculer $\langle f, \tau_{n\pi} \text{sinc} \rangle$ pour tout entier $n \in \mathbb{Z}$.

4. En déduire le théorème d'échantillonnage de Shannon : pour toute fonction f de l'espace $PW(\mathbb{R})$, on a

$$f = \sum_{n \in \mathbb{Z}} f(n\pi) \tau_{n\pi} \text{sinc}.$$

Préciser en quel sens la convergence a lieu.

Corrigé.

1. Pour tout entier $n \in \mathbb{Z}$, on pose $e_n(t) = \mathbf{1}_{[-1,1]}(t) \frac{e^{in\pi t}}{\sqrt{2}}$. On observe que $\langle e_n, e_m \rangle = \frac{1}{2} \int_{-1}^1 e^{i(n-m)\pi t} dt$; ce produit scalaire est nul si $m \neq n$, et vaut 1 sinon. Les fonctions e_n , pour $n \in \mathbb{Z}$, forment donc une famille orthonormale de $L^2([-1, 1])$. De plus, l'espace $L^2([-1, 1])$ s'identifie par restriction à l'espace des fonctions mesurables 2-périodiques dont le module au carré est intégrable sur une période, identifiées lorsqu'elles sont égales presque partout. Dans cet espace, les fonctions $t \mapsto e^{in\pi t}$, pour $n \in \mathbb{Z}$, forment une famille totale d'après le théorème de Fejér. On en déduit que la famille des fonctions e_n , pour $n \in \mathbb{Z}$, est totale. Il s'agit donc d'une base hilbertienne de l'espace $L^2([-1, 1])$.
2. On rappelle que si la transformée de Fourier sur $L^2(\mathbb{R})$ est définie en partant de la formule

$$\widehat{f}(t) = \int_{-\infty}^{+\infty} f(\xi) e^{-i\xi t} d\xi,$$

valable quand $f \in L^1(\mathbb{R})$, alors l'application $\mathcal{F} : f \mapsto \frac{\widehat{f}}{\sqrt{2\pi}}$ est une isométrie linéaire dont l'inverse $\mathcal{F}^{-1}$ vérifie

$$\mathcal{F}^{-1}(g)(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(t) e^{it\xi} dt$$

quand $g \in L^1(\mathbb{R})$. De plus, $\mathcal{F}$ établit une correspondance isométrique entre $PW(\mathbb{R})$ et $L^2([-1, 1])$. Comme les fonctions e_n , pour $n \in \mathbb{Z}$, forment une base hilbertienne de $L^2([-1, 1])$, leurs images inverses $f_n = \mathcal{F}^{-1}(e_n)$ forment une base hilbertienne de $PW(\mathbb{R})$. On sait aussi que l'opérateur de translation $\tau_{n\pi}$ sur la variable ξ correspond dans le domaine de Fourier à une multiplication par la fonction $t \mapsto e^{-in\pi t}$. Ainsi, pour tout $n \in \mathbb{Z}$,

$$f_n = \mathcal{F}^{-1}(e_n) = \frac{1}{\sqrt{2}} \tau_{n\pi} \mathcal{F}^{-1}(\mathbf{1}_{[-1,1]}).$$

Cependant, comme l'indicatrice de $[-1, 1]$ est dans $L^1(\mathbb{R})$, on calcule aisément pour $\xi \neq 0$

$$\mathcal{F}^{-1}(\mathbf{1}_{[-1,1]})(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathbf{1}_{[-1,1]}(t) e^{it\xi} dt = \frac{1}{\sqrt{2\pi}} \int_{-1}^1 e^{it\xi} dt = \frac{1}{\sqrt{2\pi}} \left[\frac{e^{it\xi}}{i\xi} \right]_{-1}^1 = \sqrt{\frac{2}{\pi}} \text{sinc } \xi.$$

L'égalité se prolonge en $\xi = 0$ par continuité. On en déduit que $f_n = \frac{\tau_{n\pi} \text{sinc}}{\sqrt{\pi}}$ pour tout $n \in \mathbb{Z}$, et ces fonctions forment une base hilbertienne de $PW(\mathbb{R})$.

3. Pour toute fonction f dans $PW(\mathbb{R})$, on commence par calculer, en utilisant l'isométrie de Fourier et ce qui précède, le produit scalaire

$$\begin{aligned} \langle f, \text{sinc} \rangle &= \left\langle \frac{1}{\sqrt{2\pi}} \mathcal{F}^{-1}(\widehat{f}), \sqrt{\frac{\pi}{2}} \mathcal{F}^{-1}(\mathbf{1}_{[-1,1]}) \right\rangle \\ &= \frac{1}{2} \langle \widehat{f}, \mathbf{1}_{[-1,1]} \rangle = \frac{1}{2} \int_{-1}^1 \widehat{f}(t) dt = \frac{1}{2} \int_{-\infty}^{+\infty} \widehat{f}(t) dt \\ &= \frac{\sqrt{2\pi}}{2} \mathcal{F}^{-1}(\widehat{f})(0) = \frac{\sqrt{2\pi}}{2} \sqrt{2\pi} \mathcal{F}^{-1}(\mathcal{F}(f))(0) = \pi f(0). \end{aligned}$$

Ensuite, pour tout entier $n \in \mathbb{Z}$, on observe que la fonction $\tau_{-n\pi} f$ appartient aussi à l'espace $PW(\mathbb{R})$, et donc le calcul précédent fournit

$$\langle f, \tau_{n\pi} \text{sinc} \rangle = \langle \tau_{-n\pi} f, \text{sinc} \rangle = \pi \tau_{-n\pi} f(0) = \pi f(n\pi).$$

4. Soit f une fonction de $PW(\mathbb{R})$. En décomposant la fonction f sur la base hilbertienne formée par les fonctions f_n , pour $n \in \mathbb{Z}$, on peut écrire

$$f = \sum_{n=-\infty}^{+\infty} \langle f, f_n \rangle f_n = \sum_{n=-\infty}^{+\infty} \frac{1}{\pi} \langle f, \tau_{n\pi} \text{sinc} \rangle \tau_{n\pi} \text{sinc} = \sum_{n=-\infty}^{+\infty} f(n\pi) \tau_{n\pi} \text{sinc}.$$

Cette convergence a lieu dans $PW(\mathbb{R})$, donc au sens de la norme L^2 , et également uniformément sur $\mathbb{R}$. En effet, pour tout $g \in PW(\mathbb{R})$, la formule d'inversion et Cauchy-Schwarz donnent $\|g\|_\infty \leq \|\hat{g}\|_2 / (\sqrt{2\pi}) = \|g\|_2 / \sqrt{\pi}$. On applique cette borne aux restes des sommes partielles. Enfin, plus explicitement, la formule s'écrit

$$f(\xi) = \sum_{n=-\infty}^{+\infty} f(n\pi) \text{sinc}(\xi - n\pi).$$

On peut aussi remarquer que les fonctions f de l'espace $PW(\mathbb{R})$ sont très régulières puisqu'elles s'étendent en des fonctions holomorphes sur tout le plan complexe grâce à la formule

$$f(z) = \frac{1}{2\pi} \int_{-1}^1 \hat{f}(t) e^{itz} dt.$$

Exercice 12.5 (Théorème de Müntz-Szasz)

Leçon : 245. Référence : [14].

Rappel : Une partie A d'un espace de Hilbert H est appelée famille totale si $\text{Vect } A$ est dense dans H .

1. Montrer que la partie $\{t \mapsto t^n, n \in \mathbb{N}\}$ est une famille totale dans $L^2([0, 1])$.
2. On note $(\lambda_n)_{n \in \mathbb{N}^*}$ une suite de nombres réels deux à deux distincts, tous strictement supérieurs à $-1/2$, qui converge vers un réel $\lambda > -1/2$. Montrer que la partie $\{t \mapsto t^{\lambda_n}, n \in \mathbb{N}^*\}$ est une famille totale dans $L^2([0, 1])$.
3. Soit $(\lambda_n)_{n \in \mathbb{N}^*}$ une suite strictement croissante de nombres réels strictement positifs qui diverge vers l'infini. Le but de la fin de l'exercice est de montrer que la partie $\{t \mapsto t^{\lambda_n}, n \in \mathbb{N}^*\}$ est une famille totale dans $L^2([0, 1])$ si et seulement si la série $\sum_n (1/\lambda_n)$ diverge.

- (a). Une fonction f de $L^2([0, 1])$ étant fixée, on introduit la fonction de la variable complexe

$$h_f(z) = \int_0^1 t^z f(t) dt.$$

Montrer que la fonction h_f est holomorphe sur le demi-plan défini par $\Re z > -1/2$ et bornée sur le demi-plan défini par $\Re z \geq 0$.

- (b). On note $\varphi(z) = (1+z)/(1-z)$ l'homographie qui envoie le disque unité ouvert D vers le demi-plan défini par $\Re z > 0$. Montrer que la fonction $K = h_f \circ \varphi$ est holomorphe dans D et continue sur $\overline{D} \setminus \{1\}$.
- (c). Montrer que la fonction f est nulle si l'ensemble des zéros de K admet un point d'accumulation dans D .
- (d). Montrer que pour tout $\alpha \in D$, la fonction $\psi_\alpha(z) = \frac{z-\alpha}{1-\bar{\alpha}z}$ est une bijection de D dans D , et de ∂D dans ∂D , d'inverse $\psi_{-\alpha}$.

On suppose que la série $\sum_n (1/\lambda_n)$ diverge et que la fonction f vérifie $h_f(\lambda_n) = 0$ pour tout entier $n \geq 1$. On suppose que l'ensemble des zéros de K n'admet pas de point d'accumulation dans D . Pour les questions qui suivent, on omet les termes initiaux tels que $\lambda_n \leq 1$, puis on réindexe la suite : cela ne change pas la divergence de $\sum_n 1/\lambda_n$ et assure que $\alpha_n \neq 0$.

- (e). On note $m \in \mathbb{N}$ la multiplicité de 0 comme zéro éventuel de K et on définit $\alpha_n = \varphi^{-1}(\lambda_n)$ pour tout $n \geq 1$.

Montrer que la fonction $K_N(z) = \frac{K(z)}{z^m} \prod_{n=1}^N \frac{1}{\psi_{\alpha_n}(z)}$ est holomorphe sur D et continue sur $\overline{D} \setminus \{1\}$, quel que soit l'entier $N \geq 1$.

- (f). Pour $0 < \varepsilon < 1$, assez petit pour que le disque $|z - 1| \leq \varepsilon$ ne contienne aucun des $\alpha_1, \dots, \alpha_N$, on note Ω_ε l'ouvert constitué des points $z \in D$ tels que $|z - 1| > \varepsilon$. En appliquant le principe du maximum à la fonction K_N sur l'ouvert Ω_ε , montrer que

$$0 < \left(\lim_{z \rightarrow 0} \left| \frac{K(z)}{z^m} \right| \right) \prod_{n=1}^N \frac{1}{|\alpha_n|} \leq \sup_{z \in \partial \Omega_\varepsilon} |K_N(z)|.$$

- (g). En notant $\|f\|$ la norme de f dans $L^2([0, 1])$, établir la borne

$$\sup_{z \in \partial \Omega_\varepsilon} |K_N(z)| \leq \frac{\|f\|}{(1 - \varepsilon)^m} \sup_{|z-1| \leq \varepsilon} \prod_{n=1}^N \frac{1}{|\psi_{\alpha_n}(z)|}.$$

- (h). En faisant tendre ε vers zéro, en déduire que

$$0 < \lim_{z \rightarrow 0} \left| \frac{K(z)}{z^m} \right| \leq \|f\| \prod_{n=1}^N |\alpha_n|.$$

- (i). En faisant tendre N vers l'infini, en déduire que l'ensemble des zéros de K admet nécessairement un point d'accumulation dans D , puis conclure.

Réciproquement, on suppose que la série $\sum_n (1/\lambda_n)$ converge. On va construire une fonction $f \in L^2([0, 1])$ non nulle telle que $h_f(\lambda_n) = 0$ pour tout entier $n \geq 1$.

- (j). Montrer que la fonction $h(z) = \frac{z}{(2+z)^3} \prod_{n=1}^{\infty} \frac{\lambda_n - z}{2 + \lambda_n + z}$ est méromorphe sur $\mathbb{C}$, avec pour pôles les points -2 et $-2 - \lambda_n$, pour $n \in \mathbb{N}^*$, et pour zéros les points 0 et λ_n , pour $n \in \mathbb{N}^*$. Montrer que $|h(z)| \leq 1$ si $\Re z \geq -1$ et que la fonction $s \mapsto h(-1 + is)$ est intégrable sur $\mathbb{R}$.

- (k). En utilisant la formule de Cauchy, montrer que pour tout $z \in \mathbb{C}$ tel que $\Re z > -1$, on a

$$h(z) = -\frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{h(-1 + is)}{-1 + is - z} ds.$$

- (l). Montrer que $h(z) = \int_0^1 t^z \left(\frac{1}{2\pi} \int_{-\infty}^{+\infty} h(-1 + is) e^{-is \log t} ds \right) dt$.

- (m). Montrer que la fonction définie par $f(t) = \frac{1}{2\pi} \hat{g}(\log t)$, avec $g(s) = h(-1 + is)$ et $\hat{g}(\xi) = \int_{\mathbb{R}} g(s) e^{-is\xi} ds$, vérifie $h_f = h$ et donc convient. Conclure.

Corrigé.

- La partie $\{t \mapsto t^n, n \in \mathbb{N}\}$ est une famille totale dans $L^2([0, 1])$ du fait de la densité de $\mathcal{C}^0([0, 1])$ dans cet espace, et du théorème de Stone-Weierstrass qui assure la densité des fonctions polynomiales dans $\mathcal{C}^0([0, 1])$.
- Notons F l'espace vectoriel engendré par les fonctions $t \mapsto t^{\lambda_n}$, pour $n \in \mathbb{N}^*$. Il suffit de montrer que $F^\perp = \{0\}$. Dans ce but, considérons une fonction $f \in F^\perp$, puis la fonction définie par

$$h_f(z) = \int_0^1 t^z f(t) dt.$$

Il s'agit d'une fonction holomorphe dans le domaine ouvert $\{z \in \mathbb{C} \mid \Re z > -1/2\}$. Cela provient en particulier de l'inégalité de Cauchy-Schwarz qui assure que

$$\int_0^1 |t^z f(t)| dt \leq \|f\| \left(\int_0^1 t^{2\Re z} dt \right)^{1/2}.$$

De plus, cette fonction est nulle sur tous les points λ_n , qui présentent un point d'accumulation dans le domaine, puisque $\lambda_n \xrightarrow{n \rightarrow \infty} \lambda > -1/2$. On en déduit que la fonction h_f est nulle sur tout son domaine. Cela implique en particulier que $h_f(n) = 0$ pour tout entier $n \in \mathbb{N}$. Autrement dit, pour tout entier $n \in \mathbb{N}$, on a

$$\int_0^1 t^n f(t) dt = 0.$$

Comme la famille des fonctions $t \mapsto t^n$, pour $n \in \mathbb{N}$, est totale d'après la question précédente, on en déduit que la fonction f est nulle. Finalement, l'espace vectoriel F est bien dense dans $L^2([0, 1])$.

3. (a). Comme vu précédemment, la fonction h_f est holomorphe sur le demi-plan ouvert défini par la condition $\Re z > -1/2$. De plus, elle est bornée sur le demi-plan défini par la condition $\Re z \geq 0$, puisque dans ce cas, l'inégalité triangulaire et l'inégalité de Cauchy-Schwarz impliquent que

$$|h_f(z)| \leq \int_0^1 |t^z f(t)| dt \leq \|f\| \left(\int_0^1 \underbrace{t^{2\Re z}}_{\leq 1} dt \right)^{1/2} \leq \|f\|.$$

- (b). On précise que l'homographie φ qui envoie le disque unité ouvert D vers le demi-plan défini par $\Re z > 0$ est définie par $\varphi(z) = \frac{1+z}{1-z}$ pour $z \neq 1$. Son inverse est donnée par $\varphi^{-1}(z) = \frac{z-1}{z+1}$ pour $z \neq -1$. La fonction φ est clairement holomorphe sur D , et envoie ce domaine dans le demi-plan défini par $\Re z > 0$, sur lequel la fonction h_f est holomorphe. La composée $K = h_f \circ \varphi$ est donc holomorphe sur D . De même, la fonction φ est clairement continue sur le domaine $\overline{D} \setminus \{1\}$, et envoie ce domaine dans le demi-plan défini par $\Re z \geq 0$, sur lequel la fonction h_f est continue. La composée K est donc continue sur $\overline{D} \setminus \{1\}$.
- (c). Il est clair que si la fonction f est nulle, alors la fonction h_f est nulle aussi, de même que la fonction K . L'ensemble des zéros de K admet donc évidemment un point d'accumulation dans D . Pour la réciproque, supposons que K vérifie cette propriété. Comme K est holomorphe sur le domaine D , cela signifie que K est nulle sur D . En composant par φ^{-1} , on en déduit que la fonction h_f est nulle sur le demi-plan défini par $\Re z > 0$. En particulier, on a $h_f(n) = 0$ pour tout entier $n \in \mathbb{N}$, c'est-à-dire qu'on retrouve la propriété

$$\int_0^1 t^n f(t) dt = 0$$

pour tout $n \in \mathbb{N}$. Comme vu précédemment, la famille des fonctions $t \mapsto t^n$, pour $n \in \mathbb{N}$ est totale, ce qui permet de conclure que la fonction f est nulle.

- (d). Pour $\alpha = 0$, la fonction ψ_0 est l'identité et le résultat est immédiat. Pour tout nombre complexe $\alpha \in D \setminus \{0\}$, on observe d'abord que la fonction ψ_α est holomorphe sur tout le plan complexe, excepté au pôle $1/\bar{\alpha}$. De plus, pour tout nombre complexe z distinct de ce point, on calcule

$$\psi_{-\alpha}(\psi_\alpha(z)) = \frac{\frac{z-\alpha}{1-\bar{\alpha}z} + \alpha}{1 + \bar{\alpha} \frac{z-\alpha}{1-\bar{\alpha}z}} = \frac{(z-\alpha) + (1-\bar{\alpha}z)\alpha}{(1-\bar{\alpha}z) + (z-\alpha)\bar{\alpha}} = \frac{z-\alpha + \alpha - |\alpha|^2 z}{1-\bar{\alpha}z + \bar{\alpha}z - |\alpha|^2} = z.$$

En remplaçant α par son opposé, et en supposant z distinct du pôle $-1/\bar{\alpha}$, on obtient aussi $\psi_\alpha(\psi_{-\alpha}(z)) = z$. On en déduit que les fonctions ψ_α et $\psi_{-\alpha}$ sont inversibles, et inverses l'une de l'autre. Ensuite, pour tout nombre réel θ , on a

$$|\psi_\alpha(e^{i\theta})| = \left| \frac{e^{i\theta} - \alpha}{1 - \bar{\alpha}e^{i\theta}} \right| = \frac{|e^{i\theta} - \alpha|}{|e^{-i\theta} - \bar{\alpha}|} = 1,$$

de sorte que ψ_α est une bijection du cercle unité ∂D dans lui-même. Comme le pôle $1/\bar{\alpha}$ de la fonction ψ_α n'appartient pas au disque unité ouvert D , on en déduit que ψ_α est holomorphe sur ce domaine. D'après le principe du maximum, le module ψ_α atteint son maximum sur ∂D , et ce maximum vaut donc 1. On a ainsi $|\psi_\alpha(z)| \leq 1$ pour tout $z \in D$, et l'inégalité est en fait stricte à nouveau en vertu du principe du maximum. Par conséquent, l'application ψ_α envoie le disque unité ouvert D dans lui-même. En remplaçant α par son opposé, on voit que $\psi_{-\alpha}$ vérifie la même propriété. On en conclut que ψ_α est une bijection de D dans lui-même.

- (e). On a supposé que la fonction h_f est nulle en tout point de la forme λ_n , pour $n \in \mathbb{N}^*$. Ainsi, la fonction K , holomorphe sur D , est nulle en tout point de la forme

$$\alpha_n = \varphi^{-1}(\lambda_n) = \frac{\lambda_n - 1}{\lambda_n + 1} \in D.$$

De plus, on a supposé que l'ensemble des zéros de K n'admet pas de point d'accumulation. Pour tout entier $N \geq 1$, on peut alors présenter K sous la forme

$$K(z) = z^m (z - \alpha_1) \cdots (z - \alpha_N) \tilde{K}(z),$$

où m désigne la multiplicité de 0 comme zéro éventuel de K , et $\tilde{K}$ est une fonction holomorphe sur D . Ainsi, la fonction définie par

$$K_N(z) = \frac{K(z)}{z^m} \prod_{n=1}^N \frac{1}{\psi_{\alpha_n}(z)} = (1 - \bar{\alpha}_1 z) \cdots (1 - \bar{\alpha}_N z) \tilde{K}(z)$$

est holomorphe sur le disque unité ouvert D . De plus, il est clair que cette fonction K_N est continue sur le domaine $\bar{D} \setminus \{1\}$ puisque c'est le cas de la fonction K et puisque toutes les fonctions ψ_{α} sont continues sur le disque unité fermé $\bar{D}$.

- (f). La fonction K_N est holomorphe sur l'ouvert Ω_ε , et continue sur son adhérence $\bar{\Omega}_\varepsilon$. D'après le principe du maximum, le module de la fonction K_N atteint son maximum au bord, et comme l'ouvert Ω_ε contient 0, on a $|K_N(0)| \leq \sup_{z \in \partial\Omega_\varepsilon} |K_N(z)|$. Comme on a pris soin de diviser par z^m pour définir $K_N(z)$, on a nécessairement $K_N(0) \neq 0$. De plus, on a $\psi_{\alpha_n}(0) = -\alpha_n$ pour tout n . En évaluant $|K_N(0)|$, on en déduit que

$$0 < \left(\lim_{z \rightarrow 0} \left| \frac{K(z)}{z^m} \right| \right) \prod_{n=1}^N \frac{1}{|\alpha_n|} \leq \sup_{z \in \partial\Omega_\varepsilon} |K_N(z)|.$$

- (g). Pour tout point $z \in \partial\Omega_\varepsilon$, on a $|z| \geq 1 - \varepsilon$, ainsi que $\Re \varphi(z) \geq 0$, ce qui indique que $\varphi(z)$ appartient à un demi-plan sur lequel la fonction h_f est bornée par $\|f\|$. Par conséquent,

$$|K_N(z)| = \frac{|h_f(\varphi(z))|}{|z|^m} \prod_{n=1}^N \frac{1}{|\psi_{\alpha_n}(z)|} \leq \frac{\|f\|}{(1 - \varepsilon)^m} \prod_{n=1}^N \frac{1}{|\psi_{\alpha_n}(z)|}.$$

On a maintenant deux cas possibles : si $z \in \partial D$, alors $|\psi_{\alpha_n}(z)| = 1$; sinon, on a $z \in D$ et $|z - 1| = \varepsilon$, et donc $|\psi_{\alpha_n}(z)| \leq 1$. Cela étant valable pour tout entier n , on en déduit que le supremum de $\prod_{n=1}^N \frac{1}{|\psi_{\alpha_n}(z)|}$ sur le bord $\partial\Omega_\varepsilon$ est en fait atteint sur l'ensemble des $z \in D$ vérifiant $|z - 1| = \varepsilon$. *A fortiori*, on a la borne

$$\sup_{z \in \partial\Omega_\varepsilon} |K_N(z)| \leq \frac{\|f\|}{(1 - \varepsilon)^m} \sup_{|z-1| \leq \varepsilon} \prod_{n=1}^N \frac{1}{|\psi_{\alpha_n}(z)|}.$$

- (h). Chaque fonction ψ_{α_n} étant continue et non nulle au point 1, le membre de droite tend vers $\|f\| \prod_{n=1}^N \frac{1}{|\psi_{\alpha_n}(1)|}$ quand $\varepsilon \rightarrow 0$. Cependant, on sait que $|\psi_{\alpha_n}(1)| = 1$ pour tout n , donc cette limite se résume à $\|f\|$. Finalement,

$$0 < \lim_{z \rightarrow 0} \left| \frac{K(z)}{z^m} \right| \leq \|f\| \prod_{n=1}^N |\alpha_n|.$$

- (i). On observe que $\log |\alpha_n| = \log \left| 1 - \frac{2}{\lambda_n + 1} \right| \sim -\frac{2}{\lambda_n}$. La série $\sum \log |\alpha_n|$ est donc divergente, tout comme la série $\sum 1/\lambda_n$. On en déduit que $\|f\| \prod_{n=1}^N |\alpha_n| \xrightarrow{N \rightarrow \infty} 0$. Cela contredit le fait que cette quantité est uniformément minorée par une borne strictement positive. Il en ressort que l'ensemble des zéros de K admet un

point d'accumulation. Comme vu précédemment, cela implique que la fonction f est nulle. On a donc démontré que si la série $\sum 1/\lambda_n$ diverge, alors l'unique fonction $f \in L^2([0, 1])$ telle que pour tout $n \in \mathbb{N}^*$,

$$\int_0^1 t^{\lambda_n} f(t) dt = 0$$

est la fonction nulle. Ainsi, la partie $\{t \mapsto t^{\lambda_n}, n \in \mathbb{N}^*\}$ est une famille totale dans $L^2([0, 1])$.

- (j). En observant que $\frac{\lambda_n - z}{2 + \lambda_n + z} = 1 - \frac{1}{\lambda_n} \frac{2z + 2}{1 + \frac{2+z}{\lambda_n}}$ pour tout entier n , et en invoquant la convergence de la série $\sum 1/\lambda_n$, on montre que le produit infini $\prod_{n=1}^{\infty} \frac{\lambda_n - z}{2 + \lambda_n + z}$ converge uniformément sur tout compact de $\mathbb{C}$ qui ne contient aucun point de la forme $-2 - \lambda_n$. La fonction $h(z) = \frac{z}{(2+z)^3} \prod_{n=1}^{\infty} \frac{\lambda_n - z}{2 + \lambda_n + z}$ est donc méromorphe sur $\mathbb{C}$, avec pour pôles le point -2 et les points de la forme $-2 - \lambda_n$, pour $n \in \mathbb{N}^*$, et pour zéros le point 0 et les points de la forme λ_n , pour $n \in \mathbb{N}^*$. De plus, on vérifie que chaque facteur dans le produit infini est de module au plus 1 sous la condition $\Re z \geq -1$. C'est aussi le cas du facteur initial. On a donc $|h(z)| \leq 1$. On a même, si $z = -1 + is$, avec $s \in \mathbb{R}$,

$$|h(-1 + is)| \leq \left| \frac{-1 + is}{(2 - 1 + is)^3} \right| = \frac{1}{1 + s^2},$$

de sorte que la fonction $s \mapsto h(-1 + is)$ est intégrable sur $\mathbb{R}$.

- (k). On fixe un nombre complexe $z \in \mathbb{C}$ tel que $\Re z > -1$. On note $R > 1 + |z|$ et on considère le chemin fermé γ_R constitué du demi-cercle allant du point $-1 - iR$ au point $-1 + iR$ en passant par le point $-1 + R$, puis du segment allant du point $-1 + iR$ au point $-1 - iR$. Puisque la fonction h ne possède aucun pôle à l'intérieur du domaine délimité par ce chemin, la formule de Cauchy donne

$$h(z) = \int_{\gamma_R} \frac{h(\xi)}{\xi - z} \frac{d\xi}{2i\pi}.$$

Sur le demi-cercle, l'intégrale vérifie

$$\left| \int_{-\pi/2}^{\pi/2} \frac{h(-1 + Re^{i\theta})}{-1 + Re^{i\theta} - z} iRe^{i\theta} \frac{d\theta}{2i\pi} \right| \leq \frac{R}{2\pi} \int_{-\pi/2}^{\pi/2} \frac{\left| \frac{-1 + Re^{i\theta}}{(1 + Re^{i\theta})^3} \right|}{R - |1 + z|} d\theta \xrightarrow{R \rightarrow \infty} 0.$$

Il reste donc la limite de l'intégrale sur le segment, c'est-à-dire

$$h(z) = -\frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{h(-1 + is)}{-1 + is - z} ds.$$

- (l). Pour $\Re z > -1$ et $s \in \mathbb{R}$, on a clairement $\frac{1}{1 + z - is} = \int_0^1 t^{z-is} dt$, de sorte qu'après usage du théorème de Fubini, on obtient

$$h(z) = \int_0^1 t^z \left(\frac{1}{2\pi} \int_{-\infty}^{+\infty} h(-1 + is) e^{-is \log t} ds \right) dt.$$

- (m). Si on pose $g(s) = h(-1 + is)$, on voit que $h(z) = \frac{1}{2\pi} \int_0^1 t^z \hat{g}(\log t) dt$. Autrement dit, la fonction f définie par $f(t) = \frac{1}{2\pi} \hat{g}(\log t)$ appartient à $L^2([0, 1])$, puisque $\hat{g}$ est continue et bornée sur $\mathbb{R}$, en tant que transformée de Fourier de la fonction $g \in L^1(\mathbb{R})$, et de plus cette fonction f vérifie bien

$$h(z) = \int_0^1 t^z f(t) dt = h_f(z)$$

sous la condition $\Re z > -1$. Finalement, comme la fonction h s'annule en tout point de la forme λ_n , on en déduit que la fonction f vérifie pour tout $n \in \mathbb{N}^*$,

$$\int_0^1 t^{\lambda_n} f(t) dt = 0.$$

On a donc trouvé une fonction f non nulle qui est orthogonale dans $L^2([0, 1])$ à toutes les fonctions $t \mapsto t^{\lambda_n}$, pour $n \in \mathbb{N}^*$. Ces fonctions ne peuvent donc pas former une famille totale.

Théorème de Baire

Référence : [6].

Exercice 13.1 (Théorème de Balaguer-Corominas)

Leçon : 208. Référence : [6].

Soit $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ une application de classe $\mathcal{C}^\infty$ vérifiant

$$\forall x \in \mathbb{R} \quad \exists n \in \mathbb{N} \quad \varphi^{(n)}(x) = 0.$$

On veut montrer que φ est une fonction polynomiale sur $\mathbb{R}$. Pour tout $n \in \mathbb{N}$, on note F_n l'ensemble des $x \in \mathbb{R}$ tels que $\varphi^{(n)}(x) = 0$. On note ensuite $\Omega = \bigcup_{n \in \mathbb{N}} \overset{\circ}{F}_n$, l'union des intérieurs des ensembles F_n , ainsi que $F = \mathbb{R} \setminus \Omega$.

1. Soit $(x_p)_{p \in \mathbb{N}}$ une suite de points distincts de $\mathbb{R}$ tendant vers $x \in \mathbb{R}$ telle qu'il existe $n_0 \in \mathbb{N}$ vérifiant $\varphi^{(n_0)}(x_p) = 0$ pour tout p . Montrer que $\varphi^{(n)}(x) = 0$ pour tout $n \geq n_0$.
2. Montrer que sur toute composante connexe de Ω , la fonction φ coïncide avec une fonction polynomiale.
3. Montrer que F ne possède aucun point isolé.
4. En raisonnant par l'absurde, montrer que l'ensemble F est vide, puis conclure.

Corrigé.

1. Quitte à extraire une sous-suite, on peut supposer que la suite $(x_p)_{p \in \mathbb{N}}$ est strictement monotone, par exemple strictement croissante. Pour tout entier p , on a $\varphi^{(n_0)}(x_p) = \varphi^{(n_0)}(x_{p+1}) = 0$. D'après le théorème de Rolle, il existe $y_p \in]x_p, x_{p+1}[$ tel que $\varphi^{(n_0+1)}(y_p) = 0$. On obtient alors une suite $(y_p)_{p \in \mathbb{N}}$ strictement croissante qui converge vers x et qui annule la fonction $\varphi^{(n_0+1)}$. Par continuité, on en déduit que $\varphi^{(n_0+1)}(x) = 0$. En itérant ce processus, on obtient finalement que $\varphi^{(n)}(x) = 0$ pour tout $n \geq n_0$.
2. Soit $]a, b[$ une composante connexe de Ω , et soit x_0 un point de $]a, b[$. Alors il existe $n \in \mathbb{N}$ et $\alpha > 0$ tels que la fonction $\varphi^{(n)}$ est nulle sur l'intervalle $]x_0 - \alpha, x_0 + \alpha[$. Donc, φ coïncide avec une fonction polynomiale P sur cet intervalle, et l'ensemble Γ des réels $c \in]x_0, b[$ tels que $\varphi(x) = P(x)$ pour tout $x \in [x_0, c]$ est non vide. On note β son supremum. Si $\beta < b$, alors il existe de même un intervalle de la forme $]\beta - \eta, \beta + \eta[$, avec $\beta + \eta < b$, sur lequel la fonction φ coïncide avec une fonction polynomiale Q . On a $P = \varphi = Q$ sur l'intervalle $]x_0, \beta[\cap]\beta - \eta, \beta + \eta[$, si bien que P et Q sont les mêmes fonctions. On en déduit que $\varphi = P$ sur $]\beta, \beta + \eta[$, de sorte que $\beta + \eta \in \Gamma$, ce qui contredit la définition de β . Ainsi, $\beta = b$, et la fonction φ coïncide avec P sur l'intervalle $[x_0, b[$. On peut montrer de même que c'est aussi le cas sur l'intervalle $]a, x_0]$. Finalement, φ coïncide avec une fonction polynomiale sur chaque composante connexe de l'ouvert Ω .
3. Supposons que F admette un point isolé, noté x_0 . Alors, il existe $\varepsilon > 0$ tel que $]x_0 - \varepsilon, x_0 + \varepsilon[\cap F = \{x_0\}$. En particulier, $]x_0 - \varepsilon, x_0[$ est inclus dans Ω , et donc la fonction φ coïncide avec une fonction polynomiale P sur cet intervalle. De même, φ coïncide avec une fonction polynomiale Q sur l'intervalle $]x_0, x_0 + \varepsilon[$. Par continuité des dérivées successives de φ au point x_0 , on a $P^{(n)}(x_0) = \varphi^{(n)}(x_0) = Q^{(n)}(x_0)$ pour tout $n \in \mathbb{N}$. D'après la formule de Taylor appliquée au point x_0 , on déduit que les deux fonctions P et Q sont égales. Ainsi, $\varphi = P$ sur l'intervalle $]x_0 - \varepsilon, x_0 + \varepsilon[$. En choisissant un entier $n \geq 0$ tel que $P^{(n+1)} = 0$, on observe que $\varphi^{(n+1)} = P^{(n+1)} = 0$ sur cet intervalle. Cela prouve que $]x_0 - \varepsilon, x_0 + \varepsilon[$ est inclus dans $\overset{\circ}{F}_{n+1}$, puis que x_0 est inclus dans Ω . On obtient une contradiction, donc F ne peut admettre de point isolé.
4. Supposons que F est non vide. Par hypothèse, l'union des ensembles F_n est égale à $\mathbb{R}$ tout entier, donc l'union des ensembles $F_n \cap F$ est égale à F . Pour tout $n \in \mathbb{N}$, l'ensemble $F_n \cap F$ est fermé dans F , qui est complet (car fermé dans $\mathbb{R}$ complet). D'après le lemme de Baire (en passant au complémentaire), il existe alors un entier $n_0 \in \mathbb{N}$ tel que $F \cap \overset{\circ}{F}_{n_0}$ est d'intérieur non vide (pour la topologie induite sur F). Ainsi, il existe deux réels a et b tels que $\emptyset \neq]a, b[\cap F \subset \overset{\circ}{F}_{n_0}$.

- On considère un point x dans $]a, b[\cap F$. Le point x ne peut être isolé, donc il existe une suite $(x_p)_{p \in \mathbb{N}}$ de points distincts dans $]a, b[\cap F$ qui converge vers x . Pour tout p , on a en particulier $x_p \in F_{n_0}$, donc $\varphi^{(n_0)}(x_p) = 0$. Le résultat de la première question implique que $\varphi^{(n)}(x) = 0$ pour tout $n \geq n_0$.
- On considère maintenant un point y dans $]a, b[\cap \Omega$. Comme $]a, b[\cap F$ est non vide, la composante connexe Ω_y de Ω qui contient y possède au moins une extrémité dans $]a, b[$, notée x . Ce point x appartient nécessairement à $]a, b[\cap F$, et le point précédent montre que $\varphi^{(n)}(x) = 0$ pour tout $n \geq n_0$. De plus, la fonction φ coïncide avec une fonction polynomiale P sur Ω_y , et on a donc par continuité des dérivées successives, $P^{(n)}(x) = \varphi^{(n)}(x) = 0$ pour tout $n \geq n_0$. La formule de Taylor pour P donne donc $\varphi^{(n_0)}(y) = P^{(n_0)}(y) = 0$.

Il en ressort que $\varphi^{(n_0)}$ s'annule en tout point de l'intervalle $]a, b[$, que ce point appartienne à F ou à Ω . Cela montre que $]a, b[$ est inclus dans l'intérieur de F_{n_0} , ce qui contredit le fait que $]a, b[\cap F$ est non vide. On en conclut que l'ensemble F est vide, ce qui équivaut au fait que Ω est égal à $\mathbb{R}$ tout entier. La fonction φ est donc une fonction polynomiale sur $\mathbb{R}$.

Suites numériques et récurrences

Références : [6, 7, 13].

Exercice 14.1 (Lemme de Cesàro et lemme de Kronecker)

Leçon : 223. Références : [6, 13].

1. *Lemme de Cesàro.* Soit $(a_n)_{n \geq 1}$ une suite complexe convergente, de limite notée ℓ . Montrer que la suite $(b_n)_{n \geq 1}$ définie par la formule $b_n = \frac{a_1 + \dots + a_n}{n}$ converge aussi vers ℓ .
2. *Généralisation.* On considère une suite $(\varepsilon_n)_{n \geq 1}$ de réels positifs, avec $\varepsilon_1 > 0$, tels que la série $\sum \varepsilon_n$ diverge. Montrer que

$$\frac{\varepsilon_1 a_1 + \dots + \varepsilon_n a_n}{\varepsilon_1 + \dots + \varepsilon_n} \xrightarrow{n \rightarrow \infty} \ell.$$

3. *Lemme de Kronecker.* On considère une série convergente $\sum u_n$ et une suite croissante $(b_n)_{n \geq 1}$ de réels strictement positifs qui diverge vers l'infini. Montrer que

$$\frac{1}{b_n} \sum_{k=1}^n b_k u_k \xrightarrow{n \rightarrow \infty} 0.$$

4. *Application.* Montrer que pour toute suite $(x_n)_{n \geq 1}$, si la série de terme général x_n/n converge, alors

$$\frac{1}{n} \sum_{k=1}^n x_k \xrightarrow{n \rightarrow \infty} 0.$$

Remarque. Ce résultat est utilisé dans une preuve classique de la loi forte des grands nombres de Kolmogorov.

Corrigé.

1. Pour $\varepsilon > 0$, il existe un entier $N \in \mathbb{N}$ tel que $|a_n - \ell| \leq \varepsilon$ pour tout entier $n \geq N$. On écrit alors

$$b_n - \ell = \frac{a_1 + \dots + a_{N-1}}{n} + \frac{a_N + \dots + a_n}{n} - \ell \frac{N-1}{n} - \ell \frac{n-N+1}{n},$$

puis grâce à l'inégalité triangulaire,

$$\begin{aligned} |b_n - \ell| &\leq \frac{|a_1| + \dots + |a_{N-1}|}{n} + |\ell| \frac{N-1}{n} + \frac{1}{n} (|a_N - \ell| + \dots + |a_n - \ell|) \\ &\leq \frac{|a_1| + \dots + |a_{N-1}|}{n} + |\ell| \frac{N-1}{n} + \frac{1}{n} (n - N + 1)\varepsilon. \end{aligned}$$

En faisant tendre n vers l'infini, on obtient $\limsup_{n \rightarrow \infty} |b_n - \ell| \leq \varepsilon$. En faisant tendre ε vers 0, on conclut que la suite (b_n) converge vers ℓ .

2. On note $\sigma_n = \varepsilon_1 + \dots + \varepsilon_n$, et on écrit selon le même principe que précédemment

$$\frac{\varepsilon_1 a_1 + \dots + \varepsilon_n a_n}{\varepsilon_1 + \dots + \varepsilon_n} - \ell = \frac{\varepsilon_1 a_1 + \dots + \varepsilon_{N-1} a_{N-1}}{\sigma_n} + \frac{\varepsilon_N a_N + \dots + \varepsilon_n a_n}{\sigma_n} - \ell \frac{\sigma_{N-1}}{\sigma_n} - \ell \frac{\sigma_n - \sigma_{N-1}}{\sigma_n},$$

puis à nouveau grâce à l'inégalité triangulaire,

$$\begin{aligned} \left| \frac{\varepsilon_1 a_1 + \dots + \varepsilon_n a_n}{\varepsilon_1 + \dots + \varepsilon_n} - \ell \right| &\leq \frac{\varepsilon_1 |a_1| + \dots + \varepsilon_{N-1} |a_{N-1}|}{\sigma_n} + |\ell| \frac{\sigma_{N-1}}{\sigma_n} + \frac{1}{\sigma_n} (\varepsilon_N |a_N - \ell| + \dots + \varepsilon_n |a_n - \ell|) \\ &\leq \frac{\varepsilon_1 |a_1| + \dots + \varepsilon_{N-1} |a_{N-1}|}{\sigma_n} + |\ell| \frac{\sigma_{N-1}}{\sigma_n} + \frac{1}{\sigma_n} (\sigma_n - \sigma_{N-1})\varepsilon. \end{aligned}$$

Comme $\sigma_n \xrightarrow{n \rightarrow \infty} +\infty$, on en déduit comme précédemment que $\limsup_{n \rightarrow \infty} \left| \frac{\varepsilon_1 a_1 + \dots + \varepsilon_n a_n}{\varepsilon_1 + \dots + \varepsilon_n} - \ell \right| \leq \varepsilon$. Enfin, en faisant tendre ε vers 0, on obtient $\frac{\varepsilon_1 a_1 + \dots + \varepsilon_n a_n}{\varepsilon_1 + \dots + \varepsilon_n} \xrightarrow{n \rightarrow \infty} \ell$.

3. On note $r_k = \sum_{p=k+1}^{\infty} u_p$ le reste de la série convergente $\sum u_n$, ainsi que $b_0 = 0$. Alors, une transformation d'Abel donne que pour tout entier $n \geq 1$,

$$\sum_{k=1}^n b_k u_k = \sum_{k=1}^n b_k (r_{k-1} - r_k) = -b_n r_n + \sum_{k=1}^n (b_k - b_{k-1}) r_{k-1},$$

de sorte que

$$\frac{1}{b_n} \sum_{k=1}^n b_k u_k = -r_n + \frac{1}{b_n} \sum_{k=1}^n (b_k - b_{k-1}) r_{k-1}.$$

Pour conclure, on observe d'abord que $r_n \xrightarrow{n \rightarrow \infty} 0$, et ensuite que la question précédente appliquée avec $a_k = r_{k-1}$

et $\varepsilon_k = b_k - b_{k-1}$ donne aussi $\frac{1}{b_n} \sum_{k=1}^n (b_k - b_{k-1}) r_{k-1} \xrightarrow{n \rightarrow \infty} 0$.

4. Il suffit d'appliquer le résultat de la question précédente avec $b_n = n$ et $u_k = x_k/k$.

Exercice 14.2 (Approximation rationnelle)

Leçon : 223. Référence : [6].

Soit $\alpha > 0$ un nombre irrationnel, et soit $(p_n/q_n)_{n \geq 0}$ une suite de nombres rationnels (avec $p_n \in \mathbb{Z}$ et $q_n \in \mathbb{N}^*$) qui converge vers α . Montrer que

$$\lim_{n \rightarrow \infty} p_n = \lim_{n \rightarrow \infty} q_n = +\infty.$$

Corrigé.

On pose $r_n = p_n/q_n$. Un entier $N \in \mathbb{N}^*$ étant fixé, on note Γ l'ensemble des nombres rationnels de la forme a/b situés dans l'intervalle $[\alpha - 1, \alpha + 1]$ et vérifiant $a \in \mathbb{Z}$ et $b \in \mathbb{N}$ avec $1 \leq b \leq N$. L'ensemble Γ est fini puisque

$$\Gamma = \bigcup_{1 \leq q \leq N} \Gamma_q \quad \text{avec} \quad \Gamma_q = \left\{ r \in [\alpha - 1, \alpha + 1] \mid \exists p \in \mathbb{Z} : r = \frac{p}{q} \right\},$$

où chaque ensemble Γ_q est fini. En effet, à q fixé, les entiers p tels que $r = p/q \in \Gamma_q$ vérifient $(\alpha - 1)q \leq p \leq (\alpha + 1)q$, et donc sont en nombre fini.

De plus, comme α est irrationnel, il n'appartient pas à l'ensemble Γ . Par finitude de Γ , on a donc $\rho = \min_{x \in \Gamma} |x - \alpha| > 0$.

Comme la suite (r_n) converge vers α , on a en outre $|r_n - \alpha| < \min(\rho, 1)$ pour tout entier $n \geq n_0$, et pour un certain entier n_0 . Ainsi, $r_n \notin \Gamma$ pour tout $n \geq n_0$, de sorte que r_n a un dénominateur $q_n > N$. Il en ressort que $q_n \xrightarrow{n \rightarrow \infty} +\infty$.

Enfin, on a également $p_n \xrightarrow{n \rightarrow \infty} +\infty$, puisque $p_n \sim \alpha q_n$.

Exercice 14.3 (Suite arithmético-géométrique)

Leçons : 223, 226. Référence : [6].

Étudier les deux suites réelles (u_n) et (v_n) vérifiant $0 < v_0 < u_0$, et pour tout $n \in \mathbb{N}$,

$$u_{n+1} = \frac{u_n + v_n}{2} \quad \text{et} \quad v_{n+1} = \sqrt{u_n v_n}.$$

Corrigé.

Une récurrence immédiate montre que toutes les valeurs prises par les suites sont strictement positives. De plus, on a $v_n \leq u_n$ pour tout entier $n \in \mathbb{N}$. En effet, le résultat est vrai pour $n = 0$. Ensuite, pour $n \in \mathbb{N}^*$, on a

$$u_n - v_n = \frac{u_{n-1} + v_{n-1}}{2} - \sqrt{u_{n-1} v_{n-1}} = \frac{1}{2} (\sqrt{u_{n-1}} - \sqrt{v_{n-1}})^2 \geq 0.$$

On remarque aussi que la suite (u_n) décroît, et que la suite (v_n) croît : pour tout $n \in \mathbb{N}$,

$$u_{n+1} - u_n = \frac{1}{2}(v_n - u_n) \leq 0 \quad \text{et} \quad \frac{v_{n+1}}{v_n} = \sqrt{\frac{u_n}{v_n}} \geq 1.$$

On en déduit que pour tout entier $n \in \mathbb{N}$,

$$v_0 \leq v_1 \leq \dots \leq v_n \leq u_n \leq \dots \leq u_1 \leq u_0.$$

La suite (u_n) est donc décroissante et minorée, et la suite (v_n) est croissante et majorée. Ces deux suites sont donc convergentes. Notons U et V les limites respectives. Par passage à la limite dans les relations de récurrence, on a

$$U = \frac{U+V}{2} \quad \text{et} \quad V = \sqrt{UV},$$

ce qui conduit nécessairement à $U = V$. On a montré au passage que les suites (u_n) et (v_n) sont adjacentes.

Remarque. Si on part des valeurs $u_0 = a$ et $v_0 = b$, la limite commune des deux suites, notée $M(a, b)$, est appelée *moyenne arithmético-géométrique* de a et b . Gauss a établi une relation entre la moyenne arithmético-géométrique et une intégrale elliptique de première espèce. Plus précisément,

$$\int_0^{\pi/2} \frac{d\theta}{\sqrt{a^2 \cos^2 \theta + b^2 \sin^2 \theta}} = \frac{\pi}{2 M(a, b)}.$$

Exercice 14.4 (Récurrence linéaire)

Leçons : 223, 226.

Soit $(u_n)_{n \geq 0}$ une suite complexe vérifiant une récurrence linéaire d'ordre $h \in \mathbb{N}^*$, c'est-à-dire qu'il existe $a_1, \dots, a_h \in \mathbb{C}$, avec $a_h \neq 0$, tels que pour tout $n \geq h$,

$$u_n = a_1 u_{n-1} + \dots + a_h u_{n-h}.$$

On considère l'équation caractéristique associée $x^h - a_1 x^{h-1} - \dots - a_h = 0$, dont les racines sont notées $r_1, \dots, r_q$, avec pour ordres de multiplicité $\alpha_1, \dots, \alpha_q$. En considérant la série entière $\sum u_n z^n$, démontrer que l'ensemble des suites vérifiant la relation de récurrence est constitué des suites de la forme

$$u_n = P_1(n)r_1^n + \dots + P_q(n)r_q^n,$$

où chaque P_i est un polynôme de degré $< \alpha_i$.

Corrigé.

Si la série entière $f(z) = \sum_{n=0}^{\infty} u_n z^n$ possède un rayon de convergence non nul, alors en multipliant par z^n la relation de récurrence vérifiée par la suite (u_n) et en sommant sur n , on obtient

$$f(z) - P_h(z) = a_1 z(f(z) - P_{h-1}(z)) + \dots + a_{h-1} z^{h-1}(f(z) - P_1(z)) + a_h z^h f(z),$$

où l'on a posé $P_k(z) = u_0 + u_1 z + \dots + u_{k-1} z^{k-1}$ pour tout entier $k \in \{1, \dots, h\}$. L'équation précédente est équivalente à $f(z)Q(z) = P(z)$, avec

$$\begin{cases} Q(z) = 1 - a_1 z - \dots - a_h z^h \\ P(z) = P_h(z) - a_1 z P_{h-1}(z) - \dots - a_{h-1} z^{h-1} P_1(z). \end{cases}$$

Il en ressort que $f(z) = \frac{P(z)}{Q(z)}$ est une fraction rationnelle, qui est développable en série entière en 0 car ce point

n'est pas racine de Q . Il y a donc une série entière $g(z) = \sum_{n=0}^{\infty} v_n z^n$ de rayon de convergence strictement positif dont la somme vérifie $g(z)Q(z) = P(z)$. D'après ce qui précède, cette équation implique que la suite de coefficients (v_n) vérifie aussi la relation de récurrence linéaire. De plus, l'équation est équivalente à

$$g(z) - P_h(z) = a_1 z(g(z) - P_{h-1}(z)) + \dots + a_{h-1} z^{h-1}(g(z) - P_1(z)) + a_h z^h g(z).$$

Le membre de droite est nul quand $z = 0$, ce qui donne $g(0) = P_h(0)$, puis $v_0 = u_0$. On peut donc diviser par z pour $z \neq 0$, puis prolonger en 0 les deux membres de l'équation, puis à nouveau fixer $z = 0$ pour obtenir $v_1 = u_1$. En itérant ce raisonnement, on en déduit que les termes initiaux $v_0, v_1, \dots, v_{h-1}$ coïncident avec $u_0, u_1, \dots, u_{h-1}$. On en déduit par récurrence que les deux suites (u_n) et (v_n) sont identiques. Ainsi, la série entière définissant f a bien un rayon de convergence non nul, et vérifie bien $f = P/Q$ sur le disque de convergence.

Pour calculer explicitement les termes de la suite (u_n) , on part des racines $r_1, \dots, r_q$ de l'équation caractéristique, et de leurs ordres de multiplicité respectifs $\alpha_1, \dots, \alpha_q$. Le polynôme correspondant s'écrit donc

$$R(X) = X^h - a_1 X^{h-1} - \dots - a_h = \prod_{i=1}^q (X - r_i)^{\alpha_i}.$$

Ainsi, le polynôme Q défini précédemment vérifie

$$Q(X) = X^h R\left(\frac{1}{X}\right) = \prod_{i=1}^q (1 - r_i X)^{\alpha_i}.$$

Comme le polynôme P vérifie $\deg P < h = \deg Q$, la décomposition en éléments simples de la fraction rationnelle P/Q s'écrit sous la forme

$$\frac{P(X)}{Q(X)} = \sum_{i=1}^q \sum_{j=1}^{\alpha_i} \frac{c_{i,j}}{(X - 1/r_i)^j},$$

avec des coefficients complexes $c_{i,j}$. En développant en série entière chaque fraction de l'expression précédente, on en déduit que le développement en série entière de la fonction f s'écrit

$$f(z) = \sum_{i=1}^q \sum_{j=1}^{\alpha_i} \frac{(-1)^j c_{i,j} r_i^j}{(j-1)!} \sum_{n=0}^{\infty} \frac{(n+j-1)!}{n!} r_i^n z^n.$$

En identifiant le coefficient devant z^n dans ce développement, on en déduit que

$$u_n = \sum_{i=1}^q \sum_{j=1}^{\alpha_i} \frac{(-1)^j c_{i,j} r_i^j}{(j-1)!} (n+j-1) \dots (n+1) r_i^n.$$

Posons $T_i(n) = \sum_{j=1}^{\alpha_i} \frac{(-1)^j c_{i,j} r_i^j}{(j-1)!} (n+j-1) \dots (n+1)$ pour tout entier $i \in \{1, \dots, q\}$. Il s'agit d'un polynôme en n de degré strictement inférieur à α_i . De plus, chaque terme u_n s'écrit alors sous la forme

$$u_n = T_1(n) r_1^n + \dots + T_q(n) r_q^n. \quad (14.1)$$

Notons Γ l'espace vectoriel des suites (u_n) dont le terme général s'écrit sous la forme (14.1), où chaque T_i est un polynôme de degré strictement inférieur à α_i . On vient de montrer que l'espace vectoriel, noté U , des suites vérifiant la récurrence linéaire vérifie $U \subset \Gamma$. Cependant, on observe que U est en correspondance bijective avec $\mathbb{C}^h$ via l'application qui à une suite (u_n) associe ses termes initiaux $(u_0, \dots, u_{h-1})$, puisque la donnée des h termes initiaux détermine entièrement la suite récurrente. Ainsi, l'espace vectoriel U est de dimension h . Or Γ est engendré par les h suites $(n^j r_i^n)_n$, pour $1 \leq i \leq q$ et $0 \leq j < \alpha_i$, donc sa dimension est au plus h . L'inclusion $U \subset \Gamma$ impose alors $\dim \Gamma = h$ et $U = \Gamma$. Cela prouve que l'ensemble des suites s'écrivant sous la forme (14.1) est exactement l'ensemble des suites qui vérifient la relation de récurrence linéaire.

Exercice 14.5 (Suite récurrente)

Leçons : 223, 226. Référence : [6].

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ une application continue. Un point $x_0 \in \mathbb{R}^n$ étant donné, on définit la suite (x_n) par $x_{n+1} = f(x_n)$. On suppose que (x_n) admet une et une seule valeur d'adhérence. Montrer que la suite (x_n) converge.

Corrigé.

Il suffit de montrer que la suite (x_n) est bornée. En effet, dans ce cas, elle prendra ses valeurs dans un compact, et on sait qu'une suite à valeurs dans un compact qui possède une et une seule valeur d'adhérence converge.

On suppose par l'absurde que la suite (x_n) n'est pas bornée. Notons a l'unique valeur d'adhérence de la suite, et K le compact $\overline{B}(a, 1)$. Soit $N \in \mathbb{N}$. Comme a est valeur d'adhérence, il existe un entier $n_0 \geq N$ tel que $x_{n_0} \in K$. De plus, comme la suite n'est pas bornée, il existe $n > n_0$ tel que $x_n \notin K$. Ainsi, on peut trouver un indice $m \geq n_0$ tel que $x_m \in K$ et $x_{m+1} \notin K$. On a en outre $x_{m+1} = f(x_m) \in f(K)$. On vient donc de démontrer que

$$\forall N \in \mathbb{N} \quad \exists m \geq N \quad x_{m+1} \in f(K) \setminus K.$$

On en déduit l'existence d'une sous-suite $(x_{\varphi(n)})$ à valeurs dans $f(K) \setminus K$, donc en particulier à valeurs dans $f(K)$. Comme $f(K)$ est compact (car image d'un compact par une fonction continue), on extrait de cette sous-suite une autre sous-suite $(x_{\varphi \circ \psi(n)})$ qui converge, vers une limite notée b . Pour tout entier n , on sait que $x_{\varphi(n)} \notin K$, c'est-à-dire que $\|x_{\varphi(n)} - a\| > 1$. Par passage à la limite dans la nouvelle sous-suite, on en déduit que $\|b - a\| \geq 1$. Cependant, b est aussi une valeur d'adhérence de la suite (x_n) . Cela contredit l'unicité de a . On en déduit que la suite (x_n) est bornée, et donc converge vers a .

Exercice 14.6 (Critère de Weyl)

Leçon : 223. Référence : [7].

Une suite $(u_n)_{n \geq 1}$ de $[0, 1]$ est dite *équirépartie* si on a $\frac{S_n(a, b)}{n} \xrightarrow{n \rightarrow \infty} b - a$ pour tous $a, b \in \mathbb{R}$ tels que $0 \leq a < b \leq 1$, où $S_n(a, b) = \#\{k \in \{1, \dots, n\} \mid u_k \in [a, b]\}$. Montrer l'équivalence des faits suivants :

- la suite (u_n) est équirépartie ;
- pour toute fonction continue $f : [0, 1] \rightarrow \mathbb{R}$, on a $\frac{1}{n} \sum_{k=1}^n f(u_k) \xrightarrow{n \rightarrow \infty} \int_0^1 f(t) dt$;
- pour tout entier $p \geq 1$, on a $\frac{1}{n} \sum_{k=1}^n e^{2i\pi p u_k} \xrightarrow{n \rightarrow \infty} 0$.

Corrigé.

(1) $\Rightarrow$ (2) : La propriété (1) donne la convergence de ces moyennes pour les indicatrices de segments non réduits à un point. Pour un singleton $\{a\}$, on le majore par l'indicatrice d'un segment de longueur arbitrairement petite contenant a et inclus dans $[0, 1]$; sa fréquence tend donc vers zéro. La convergence vaut ainsi pour tous les segments, y compris les singletons. Par linéarité, la propriété (2) est donc valable pour les fonctions en escalier. On passe par densité aux fonctions continues. En effet, soit un nombre réel $\varepsilon > 0$ et une fonction continue $f : [0, 1] \rightarrow \mathbb{R}$. Il existe une fonction en escalier $g : [0, 1] \rightarrow \mathbb{R}$ telle que $\|f - g\|_\infty \leq \varepsilon$. D'après l'inégalité triangulaire, on a pour tout entier $n \geq 1$,

$$\left| \frac{1}{n} \sum_{k=1}^n f(u_k) - \int_0^1 f(t) dt \right| \leq \left| \frac{1}{n} \sum_{k=1}^n (f(u_k) - g(u_k)) \right| + \left| \frac{1}{n} \sum_{k=1}^n g(u_k) - \int_0^1 g(t) dt \right| + \left| \int_0^1 g(t) dt - \int_0^1 f(t) dt \right|.$$

Le premier et le troisième termes dans la borne sont majorés par ε , tandis que le second terme converge vers 0 quand n tend vers l'infini. On a donc

$$\limsup_{n \rightarrow \infty} \left| \frac{1}{n} \sum_{k=1}^n f(u_k) - \int_0^1 f(t) dt \right| \leq 2\varepsilon.$$

On conclut que (2) est vérifiée pour toute fonction continue $f : [0, 1] \rightarrow \mathbb{R}$ en faisant finalement tendre ε vers 0.

(2) $\Rightarrow$ (1) : Considérons d'abord un segment $I = [a, b]$ tel que $0 < a < b < 1$. Pour tout entier p assez grand, notons φ_p la fonction affine par morceaux nulle sur l'ensemble $[0, a] \cup [b, 1]$ et égale à 1 sur l'intervalle $[a + 1/p, b - 1/p]$, ainsi que ψ_p la fonction affine par morceaux nulle sur l'ensemble $[0, a - 1/p] \cup [b + 1/p, 1]$ et égale à 1 sur l'intervalle $[a, b]$. On observe que $\varphi_p \leq \mathbf{1}_{[a, b]} \leq \psi_p$. Ainsi, les moyennes de ces fonctions sur les premiers termes de (u_n) vérifient

$$\frac{1}{n} \sum_{k=1}^n \varphi_p(u_k) \leq \frac{S_n(a, b)}{n} \leq \frac{1}{n} \sum_{k=1}^n \psi_p(u_k).$$

Comme les fonctions φ_p et ψ_p sont continues, on sait que

$$\frac{1}{n} \sum_{k=1}^n \varphi_p(u_k) \xrightarrow{n \rightarrow \infty} \int_0^1 \varphi_p(t) dt = b - a - \frac{1}{p} \quad \text{et} \quad \frac{1}{n} \sum_{k=1}^n \psi_p(u_k) \xrightarrow{n \rightarrow \infty} \int_0^1 \psi_p(t) dt = b - a + \frac{1}{p}.$$

Il en ressort que

$$b - a - \frac{1}{p} \leq \liminf_{n \rightarrow \infty} \frac{S_n(a, b)}{n} \leq \limsup_{n \rightarrow \infty} \frac{S_n(a, b)}{n} \leq b - a + \frac{1}{p}.$$

En faisant tendre p vers l'infini, on en conclut que $\frac{S_n(a, b)}{n} \xrightarrow{n \rightarrow \infty} b - a$. Si $a = 0$ ou $b = 1$, on adapte les deux fonctions d'encadrement : on conserve les rampes affines aux bornes intérieures et on prolonge par la valeur 1 jusqu'à toute borne de $[0, 1]$ qui appartient au segment considéré. On obtient ainsi des fonctions continues $0 \leq \varphi_p \leq \mathbf{1}_{[a, b]} \leq \psi_p \leq 1$ dont les intégrales tendent vers $b - a$. Le même encadrement des moyennes permet de conclure. Pour $[a, b] = [0, 1]$, on prend simplement $\varphi_p = \psi_p = 1$. On en déduit que la suite est équirépartie.

(2) $\Rightarrow$ (3) : C'est immédiat en considérant la partie réelle et la partie imaginaire de l'exponentielle complexe $e^{2i\pi p u_k}$.

(3) $\Rightarrow$ (2) : Par linéarité, on observe que la propriété est vraie pour tous les polynômes trigonométriques de la forme

$$x \mapsto a_0 + \sum_{k=1}^n a_k \cos(2k\pi x) + \sum_{k=1}^n b_k \sin(2k\pi x).$$

De plus, le théorème de Stone-Weierstrass trigonométrique assure que toute fonction continue $f : [0, 1] \rightarrow \mathbb{R}$ telle que $f(0) = f(1)$ est limite uniforme d'une suite de tels polynômes trigonométriques. En appliquant les raisonnements de densité utilisés ci-dessus, on peut donc montrer que la propriété est vraie pour de telles fonctions f . Il reste à étendre la propriété aux fonctions continues f qui ne vérifient pas la condition $f(0) = f(1)$. Supposons par exemple $f(0) > f(1)$. Pour $\delta \in]0, 1[$ fixé, on note g la fonction qui coïncide avec f sur l'intervalle $[\delta, 1]$, et qui est donnée par la formule

$$g(x) = \min \left\{ f(x), \frac{f(\delta) - f(1)}{\delta} x + f(1) \right\}$$

pour x dans l'intervalle $[0, \delta]$. Il s'agit d'une fonction continue, en tant que minimum de deux fonctions continues. De plus, elle vérifie $g(0) = \min\{f(0), f(1)\} = f(1) = g(1)$, ainsi que $g \leq f$ sur $[0, 1]$ et

$$0 \leq \int_0^1 (f(x) - g(x)) dx = \int_0^\delta (f(x) - g(x)) dx \leq \int_0^\delta \left| f(x) - \frac{f(\delta) - f(1)}{\delta} x - f(1) \right| dx \leq 4\|f\|_\infty \delta.$$

De même, on note h la fonction qui coïncide avec f sur l'intervalle $[0, 1 - \delta]$, et qui est donnée par la formule

$$h(x) = \max \left\{ f(x), \frac{f(0) - f(1 - \delta)}{\delta} (x - (1 - \delta)) + f(1 - \delta) \right\}$$

pour x dans l'intervalle $[1 - \delta, 1]$. Il s'agit d'une fonction continue, en tant que maximum de deux fonctions continues. De plus, elle vérifie $h(1) = \max\{f(1), f(0)\} = f(0) = h(0)$, ainsi que $h \geq f$ sur $[0, 1]$ et

$$\begin{aligned} 0 &\leq \int_0^1 (h(x) - f(x)) dx = \int_{1-\delta}^1 (h(x) - f(x)) dx \\ &\leq \int_{1-\delta}^1 \left| \frac{f(0) - f(1 - \delta)}{\delta} (x - (1 - \delta)) + f(1 - \delta) - f(x) \right| dx \leq 4\|f\|_\infty \delta. \end{aligned}$$

On a alors l'encadrement

$$\frac{1}{n} \sum_{k=1}^n g(u_k) \leq \frac{1}{n} \sum_{k=1}^n f(u_k) \leq \frac{1}{n} \sum_{k=1}^n h(u_k),$$

et comme la propriété de convergence est vraie pour les fonctions g et h , on en déduit que

$$\int_0^1 g(t) dt = \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n g(u_k) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n f(u_k) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n f(u_k) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n h(u_k) = \int_0^1 h(t) dt,$$

puis

$$-4\|f\|_\infty \delta + \int_0^1 f(t) dt \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n f(u_k) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n f(u_k) \leq \int_0^1 f(t) dt + 4\|f\|_\infty \delta.$$

En faisant tendre δ vers zéro, on en conclut que la propriété de convergence est vraie pour la fonction f .

Exercice 14.7 (Sous-groupes additifs de $\mathbb{R}$)

Leçon : 223. Référence : [6].

Soit G un sous-groupe de $(\mathbb{R}, +)$ non réduit à $\{0\}$. On note $a = \inf(G \cap]0, +\infty[)$.

1. Montrer que a définit bien un nombre réel positif.
2. On suppose $a > 0$. Montrer que $a \in G$, puis que $G = a\mathbb{Z}$.
3. On suppose $a = 0$. Montrer que G est dense dans $\mathbb{R}$.
4. *Applications.*
 - (a). Montrer que pour tous réels non nuls a et b , le sous-groupe $a\mathbb{Z} + b\mathbb{Z}$ est dense dans $\mathbb{R}$ si et seulement si a/b est irrationnel.
 - (b). Soit θ un réel non nul tel que π/θ est irrationnel. En déduire que le sous-groupe $\theta\mathbb{Z} + 2\pi\mathbb{Z}$ est dense dans $\mathbb{R}$.
 - (c). En déduire que la suite $(\cos n)_{n \geq 0}$ est dense dans l'intervalle $[-1, 1]$. Montrer qu'il en va de même pour la suite $(\sin n)_{n \in \mathbb{Z}}$.
 - (d). Montrer que si θ est un réel non nul tel que π/θ est irrationnel, alors l'ensemble $\theta\mathbb{N} + 2\pi\mathbb{Z}$ est dense dans $\mathbb{R}$. En déduire que la suite $(\sin n)_{n \geq 0}$ est dense dans l'intervalle $[-1, 1]$.

Corrigé.

1. Comme G n'est pas réduit à $\{0\}$, il contient un élément non nul x . De plus, comme G est un sous-groupe, on sait que x et $-x$ appartiennent tous deux à G , de sorte que G contient un élément strictement positif. Ainsi, l'ensemble $G \cap]0, +\infty[$ est non vide et minoré par 0, et son infimum a est donc bien un réel positif.
2. On suppose $a > 0$. Alors, a appartient nécessairement à G . Sinon, par définition de l'infimum, il existe un réel $x \in G$ tel que $a < x < 2a$, puis un réel $y \in G$ tel que $a < y < x$. On a donc $x - y \in G$ et $0 < x - y < a$, ce qui contredit la définition de a . Dès lors, comme G est un groupe qui contient a , il contient nécessairement $a\mathbb{Z}$. Réciproquement, si $x \in G$, alors on peut écrire $x = na + \delta$, avec $0 \leq \delta < a$. On observe que $\delta = x - na \in G$, ce qui par définition de a implique $\delta = 0$. Finalement, $G = a\mathbb{Z}$.
3. On suppose $a = 0$. On considère un intervalle ouvert non vide $I =]u, v[$ de longueur $\varepsilon = v - u > 0$. Comme a est nul, il existe un réel $x \in G$ tel que $0 < x < \varepsilon$. De plus, il existe un entier n tel que $(n-1)x \leq u < nx$. On a $nx \in G$, ainsi que $u < nx = (n-1)x + x < u + \varepsilon = v$, ce qui prouve que $nx \in I$. On en conclut que G est dense.
4. (a). On note a et b deux réels non nuls. Il est clair que $a\mathbb{Z} + b\mathbb{Z}$ est un sous-groupe additif de $\mathbb{R}$ non réduit à $\{0\}$. Supposons qu'il est de la forme $c\mathbb{Z}$, pour $c > 0$. Comme a et b appartiennent au sous-groupe, ils s'écrivent respectivement $a = c\alpha$ et $b = c\beta$, avec $\alpha, \beta \in \mathbb{Z} \setminus \{0\}$. On a donc $a/b = \alpha/\beta \in \mathbb{Q}$. On en déduit par contraposée que le sous-groupe $a\mathbb{Z} + b\mathbb{Z}$ est dense quand a/b est irrationnel. Réciproquement, supposons que a/b est rationnel et considérons une écriture irréductible p/q . On note $r = a/p = b/q$. On observe que $a\mathbb{Z} + b\mathbb{Z} = rp\mathbb{Z} + rq\mathbb{Z} \subset r\mathbb{Z}$. De plus, comme p et q sont premiers entre eux, il existe $s, t \in \mathbb{Z}$ tels que $ps + qt = 1$, ce qui donne $r = prs + qrt = as + bt$, et finalement $r\mathbb{Z} \subset a\mathbb{Z} + b\mathbb{Z}$. On en conclut que le sous-groupe $a\mathbb{Z} + b\mathbb{Z}$ n'est pas dense quand a/b est rationnel.
 - (b). Si θ désigne un réel non nul tel que π/θ est irrationnel, alors $(2\pi)/\theta$ est irrationnel, et le résultat de la question précédente implique que le sous-groupe $\theta\mathbb{Z} + 2\pi\mathbb{Z}$ est dense dans $\mathbb{R}$.
 - (c). Comme π est irrationnel, le résultat de la question précédente montre que le sous-groupe $\mathbb{Z} + 2\pi\mathbb{Z}$ est dense dans $\mathbb{R}$. La propriété de densité est conservée en prenant l'image par une application continue. En particulier, les nombres de la forme $\cos(n + 2m\pi)$ et $\sin(n + 2m\pi)$, pour $m, n \in \mathbb{Z}$, sont denses dans l'intervalle $[-1, 1]$. En simplifiant et en utilisant la parité du cosinus, on en déduit que la suite $(\cos n)_{n \geq 0}$ est dense dans $[-1, 1]$. La même propriété est vraie pour le sinus, mais il convient de conserver les termes d'indices négatifs car le sinus est impair : la suite $(\sin n)_{n \in \mathbb{Z}}$ est dense dans $[-1, 1]$.
 - (d). Soit $I =]u, v[$ un intervalle ouvert de longueur $\varepsilon = v - u > 0$. Comme l'ensemble $\theta\mathbb{N} + 2\pi\mathbb{Z}$ contient 0, il rencontre clairement I si cet intervalle contient 0 également. On peut donc supposer que $u \geq 0$ ou $v \leq 0$.

- Supposons $u \geq 0$. Il suffit de montrer que l'ensemble $\theta\mathbb{N} + 2\pi\mathbb{Z}$ rencontre l'intervalle $]0, \varepsilon[$. En effet, si x désigne un point commun à ces deux ensembles, alors il existe un entier $n \geq 0$ tel que $(n-1)x \leq u < nx$, et on a donc $nx \in (\theta\mathbb{N} + 2\pi\mathbb{Z})$ et $u < nx = (n-1)x + x < u + \varepsilon = v$, de sorte que le point nx est commun à l'ensemble $\theta\mathbb{N} + 2\pi\mathbb{Z}$ et l'intervalle I .

Pour montrer que l'ensemble $\theta\mathbb{N} + 2\pi\mathbb{Z}$ rencontre l'intervalle $]0, \varepsilon[$, on utilise la densité du sous-groupe $\theta\mathbb{Z} + 2\pi\mathbb{Z}$. Il existe donc une infinité de couples d'entiers $n, m \in \mathbb{Z}$ tels que $0 < n\theta + 2m\pi < \varepsilon$. Si on peut choisir n positif, on a le résultat voulu. Par conséquent, on peut supposer que pour cette infinité de couples, on a toujours $n < 0$. On considère un de ces couples, et on note $x = n\theta + 2m\pi$ l'élément correspondant. À nouveau par densité, il existe une infinité d'éléments $0 < p\theta + 2q\pi < x$ avec $p, q \in \mathbb{Z}$, et on a nécessairement $p < 0$. Pour chaque p fixé, il n'y a qu'un nombre fini d'entiers q pour lesquels cet encadrement est possible. On peut donc trouver $p \leq n$ et $q \in \mathbb{Z}$ tel que $x' = p\theta + 2q\pi$ vérifie $0 < x' < x$. On a $0 < x - x' < \varepsilon$ et $x - x' = (n-p)\theta + 2(m-q)\pi$ appartient à l'ensemble $\theta\mathbb{N} + 2\pi\mathbb{Z}$.

- Si $v \leq 0$, on peut adapter le raisonnement à l'intervalle $]-\varepsilon, 0[$, et on obtient la même conclusion.

On vient de montrer que l'ensemble $\theta\mathbb{N} + 2\pi\mathbb{Z}$ est dense dans $\mathbb{R}$. En prenant l'image par la fonction continue sinus, on en déduit que les nombres de la forme $\sin(n\theta + 2m\pi)$, avec $n \in \mathbb{N}$ et $m \in \mathbb{Z}$, sont denses dans l'intervalle $[-1, 1]$. En prenant $\theta = 1$ et en utilisant la 2π -périodicité du sinus, on en déduit la densité de la suite $(\sin n)_{n \geq 0}$.

Espaces de fonctions, suites et séries de fonctions

Références : [6, 7, 17].

Exercice 15.1 (Théorèmes de Dini)

Leçon : 241. Références : [6, 17].

1. Soit (f_n) une suite monotone de fonctions réelles continues et définies sur un espace compact X . On suppose que (f_n) converge simplement vers une fonction f continue sur X . Montrer que la convergence est uniforme.
2. *Application.* Montrer que la fonction $t \mapsto \sqrt{t}$ est limite uniforme sur $[0, 1]$ de la suite (P_n) de fonctions polynomiales définie par $P_0 = 0$ et $P_{n+1}(t) = P_n(t) + (t - P_n(t)^2)/2$.
3. Soit (f_n) une suite de fonctions croissantes définies sur $I = [a, b] \subset \mathbb{R}$. On suppose que (f_n) converge simplement vers une fonction f continue sur I . Montrer que la convergence est uniforme.
(Remarque : Ce résultat intervient dans la preuve du théorème de Glivenko-Cantelli lorsque la fonction de répartition limite est continue ; les fonctions de répartition empiriques ne sont pas nécessairement continues.)

Corrigé.

1. Quitte à remplacer chaque fonction f_n par $\pm(f_n - f)$, on peut supposer que la suite (f_n) est décroissante et converge simplement vers 0. En particulier, on a $f_n(x) \geq 0$ pour tous n et x . Pour tout $\varepsilon > 0$, on considère $V_n(\varepsilon) = \{x \in X \mid f_n(x) < \varepsilon\}$. D'après les hypothèses sur les fonctions f_n , les ensembles $V_n(\varepsilon)$ forment une suite croissante d'ouverts qui recouvrent X . Par compacité de X et croissance de ces ouverts, il existe donc un entier N_ε tel que $X = V_{N_\varepsilon}(\varepsilon)$. À nouveau par croissance des ouverts, on a en fait $X = V_n(\varepsilon)$ pour tout $n \geq N_\varepsilon$, ce qui signifie que $0 \leq f_n(x) < \varepsilon$ pour tous $x \in X$ et $n \geq N_\varepsilon$. D'où la convergence uniforme.
2. On considère, pour $t \in [0, 1]$, la fonction f_t définie sur $\mathbb{R}$ par $f_t(x) = x + \frac{1}{2}(t - x^2)$, qui atteint son maximum en $x = 1$, de sorte que pour tout $x \in \mathbb{R}$,

$$f_t(x) \leq f_t(1) = 1 + \frac{1}{2}(t - 1) = \frac{1+t}{2} \leq 1.$$

Ainsi, pour tout $t \in [0, 1]$, on a $P_n(t) = f_t(P_{n-1}(t)) \leq 1$. De plus, on a $P_{n+1}(t) - P_n(t) = f_t(P_n(t)) - f_t(P_{n-1}(t))$, et f_t croît sur $]-\infty, 1]$, intervalle dans lequel se situe $P_n(t)$. Donc, $P_{n+1}(t) - P_n(t)$ possède le même signe que $P_n(t) - P_{n-1}(t)$, c'est-à-dire le même signe que $P_1(t) - P_0(t) = \frac{t}{2} \geq 0$. On en déduit que la suite $(P_n(t))$ est croissante et majorée par 1, donc converge simplement vers $\varphi(t)$ tel que $\varphi(t) = f_t(\varphi(t))$, i.e. $\varphi(t) = \sqrt{t}$. D'après le théorème de Dini, la convergence est uniforme.

3. La fonction f est continue sur $[a, b]$, mais aussi croissante comme limite simple de fonctions croissantes. D'après le théorème de Heine, la fonction f est uniformément continue sur $[a, b]$. Pour $\varepsilon > 0$, on note $\eta > 0$ tel que $|f(x) - f(x')| < \varepsilon$ dès que $|x - x'| < \eta$, puis on considère une subdivision $a = x_0 < x_1 < \dots < x_{k-1} < x_k = b$ telle que $x_{i+1} - x_i < \eta$ pour tout i . En utilisant la convergence simple en tout point x_i , on trouve un entier N tel que $|f_n(x_i) - f(x_i)| < \varepsilon$ pour tous $i \in \{0, \dots, k\}$ et $n \geq N$. Pour $x \in [a, b]$, on a $x_i \leq x \leq x_{i+1}$ pour un certain i , et par croissance des f_n et f , on a donc

$$f_n(x_i) \leq f_n(x) \leq f_n(x_{i+1}) \quad \text{et} \quad f(x_i) \leq f(x) \leq f(x_{i+1}),$$

ce qui donne

$$f(x_i) - f_n(x_{i+1}) \leq f(x) - f_n(x) \leq f(x_{i+1}) - f_n(x_i).$$

Cependant, pour $n \geq N$, on a

$$|f(x_{i+1}) - f_n(x_i)| \leq |f(x_{i+1}) - f(x_i)| + |f(x_i) - f_n(x_i)| < 2\varepsilon,$$

et de même $|f(x_i) - f_n(x_{i+1})| < 2\varepsilon$. On en déduit que $|f(x) - f_n(x)| < 2\varepsilon$, pour tous $x \in [a, b]$ et $n \geq N$, ce qui prouve la convergence uniforme.

Remarque. Lorsque la fonction de répartition F est continue, la loi forte des grands nombres sur les rationnels, la monotonie, puis le résultat précédent sur les segments et un contrôle des queues permettent d'obtenir le théorème de Glivenko-Cantelli ci-dessous. Le cas où F présente des sauts demande un argument supplémentaire ; il ne découle pas directement de cette version du théorème de Dini. L'énoncé général est le suivant : Soit $X_1, X_2, \dots$ une suite de variables aléatoires réelles indépendantes et identiquement distribuées définies sur un même espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. On note F la fonction de répartition des variables X_i . Pour tous $n \in \mathbb{N}^*$

et $t \in \mathbb{R}$, on considère la fonction de répartition empirique $F_n(t) = \frac{1}{n} \sum_{k=1}^n \mathbf{1}_{\{X_k \leq t\}}$. Alors, presque sûrement,

$$\sup_{t \in \mathbb{R}} |F_n(t) - F(t)| \xrightarrow[n \rightarrow \infty]{} 0.$$

Exercice 15.2 (Théorème de Fejér)

Leçon : 241. Référence : [7].

Soit $f : \mathbb{R} \rightarrow \mathbb{C}$ une fonction continue et 2π -périodique. Pour tout $k \in \mathbb{Z}$, on pose $e_k(t) = e^{ikt}$ et on considère le coefficient de Fourier $c_k(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e_k(-t) dt$. Pour tout $n \in \mathbb{N}$, on définit

$$S_n = \sum_{k=-n}^n c_k(f) e_k \quad \text{et} \quad C_n = \frac{1}{n+1} (S_0 + S_1 + \dots + S_n),$$

ainsi que

$$\tilde{S}_n = \sum_{k=-n}^n e_k \quad \text{et} \quad \tilde{C}_n = \frac{1}{n+1} (\tilde{S}_0 + \tilde{S}_1 + \dots + \tilde{S}_n).$$

1. Vérifier que $\frac{1}{2\pi} \int_{-\pi}^{\pi} \tilde{C}_n(t) dt = 1$ pour tout $n \in \mathbb{N}$.
2. Montrer que pour tout $\alpha \in]0, \pi[$, la suite de fonctions $(\tilde{C}_n)$ converge uniformément vers 0 sur $[-\pi, \pi] \setminus [-\alpha, \alpha]$.
3. (Théorème de Fejér) En déduire que la suite de fonctions (C_n) converge uniformément vers f sur $\mathbb{R}$.

Corrigé.

1. On a $\frac{1}{2\pi} \int_{-\pi}^{\pi} e_k(t) dt = 0$ pour tout $k \in \mathbb{Z}^*$, $\frac{1}{2\pi} \int_{-\pi}^{\pi} e_0(t) dt = 1$. On en conclut que pour tout $n \in \mathbb{N}$,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \tilde{S}_n(t) dt = 1 \quad \text{et} \quad \frac{1}{2\pi} \int_{-\pi}^{\pi} \tilde{C}_n(t) dt = \frac{1}{n+1} \sum_{k=0}^n \frac{1}{2\pi} \int_{-\pi}^{\pi} \tilde{S}_k(t) dt = 1.$$

2. On reconnaît en $\tilde{S}_n$ le noyau de Dirichlet dont le calcul est classique : pour tout $n \in \mathbb{N}$ et tout $t \in \mathbb{R} \setminus 2\pi\mathbb{Z}$,

$$\tilde{S}_n(t) = e^{-int} \frac{e^{i(2n+1)t} - 1}{e^{it} - 1} = \frac{\sin((n+1/2)t)}{\sin(t/2)}.$$

De plus, on a aussi

$$\sum_{k=0}^n e^{i(k+1/2)t} = e^{it/2} \frac{e^{i(n+1)t} - 1}{e^{it} - 1} = e^{i(n+1)t/2} \frac{\sin((n+1)t/2)}{\sin(t/2)},$$

ce qui donne en prenant la partie imaginaire :

$$\sum_{k=0}^n \sin((k+1/2)t) = \frac{\sin^2((n+1)t/2)}{\sin(t/2)}.$$

Finalement, pour tout $n \in \mathbb{N}$ et tout $t \in \mathbb{R} \setminus 2\pi\mathbb{Z}$,

$$\tilde{C}_n(t) = \frac{1}{n+1} \sum_{k=0}^n \tilde{S}_k(t) = \frac{1}{n+1} \sum_{k=0}^n \frac{\sin((k+1/2)t)}{\sin(t/2)} = \frac{1}{n+1} \left(\frac{\sin((n+1)t/2)}{\sin(t/2)} \right)^2.$$

(On note que $\tilde{C}_n$, le noyau de Fejér, est positif.) Pour conclure, on observe alors que si $0 < \alpha < \pi$, alors

$$0 \leq \sup_{\substack{t \in [-\pi, \pi] \\ |t| > \alpha}} \tilde{C}_n(t) \leq \frac{1}{(n+1) \sin^2(\alpha/2)} \xrightarrow{n \rightarrow \infty} 0.$$

3. On remarque que pour tout $n \in \mathbb{N}$ et tout $x \in \mathbb{R}$,

$$S_n(x) = \frac{1}{2\pi} \sum_{k=-n}^n \left(\int_{-\pi}^{\pi} f(t) e^{-ikt} dt \right) e^{ikx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \tilde{S}_n(x-t) dt,$$

et donc par linéarité,

$$C_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \tilde{C}_n(x-t) dt,$$

de sorte que C_n s'écrit comme un produit de convolution. Par périodicité, et grâce à un simple changement de variable, on a aussi

$$C_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) \tilde{C}_n(t) dt.$$

Comme la fonction f est continue et 2π -périodique, elle est uniformément continue sur $\mathbb{R}$ tout entier (d'après le théorème de Heine). Ainsi, pour tout $\varepsilon > 0$, il existe $\alpha \in]0, \pi[$ tel que pour tous $x, y \in \mathbb{R}$ tels que $|x - y| < \alpha$, on a $|f(x) - f(y)| < \varepsilon$. En notant M un majorant de $|f|$ sur $\mathbb{R}$, et en partant de la formule de convolution, on obtient pour tout $n \in \mathbb{N}$ et tout $x \in \mathbb{R}$,

$$\begin{aligned} |f(x) - C_n(x)| &= \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} (f(x-t) - f(x)) \tilde{C}_n(t) dt \right| \\ &\leq \frac{1}{2\pi} \int_{\alpha \leq |t| \leq \pi} 2M \tilde{C}_n(t) dt + \frac{1}{2\pi} \int_{|t| < \alpha} \varepsilon \tilde{C}_n(t) dt \\ &\leq \frac{2M}{(n+1) \sin^2(\alpha/2)} + \frac{\varepsilon}{2\pi} \int_{-\pi}^{\pi} \tilde{C}_n(t) dt \leq \frac{2M}{(n+1) \sin^2(\alpha/2)} + \varepsilon. \end{aligned}$$

On en déduit que

$$\lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |f(x) - C_n(x)| \leq \varepsilon \xrightarrow{\varepsilon \rightarrow 0} 0,$$

si bien que la suite de fonctions (C_n) converge uniformément vers f sur $\mathbb{R}$.

Exercice 15.3 (Limite uniforme de polynômes sur la droite réelle)

Leçon : 241.

Soit f une fonction de $\mathbb{R}$ dans lui-même. On suppose que f est limite uniforme sur $\mathbb{R}$ tout entier d'une suite de fonctions polynomiales (P_n) . Montrer que f est une fonction polynomiale.

Corrigé.

Le critère de Cauchy uniforme implique qu'il existe un entier N tel que pour tout entier $n \geq N$ et tout $x \in \mathbb{R}$, on a $|P_n(x) - P_N(x)| \leq 1$. Donc, pour tout entier $n \geq N$, la fonction $P_n - P_N$ est une fonction polynomiale bornée sur $\mathbb{R}$, donc elle est égale à une constante notée α_n . On a $\alpha_n = P_n(0) - P_N(0)$, qui converge vers une limite notée α quand $n \rightarrow \infty$, puisque $P_n(0)$ converge. Cependant, on a $\alpha_n = P_n(x) - P_N(x)$ pour tout x , donc en passant à la limite, on trouve $\alpha = f(x) - P_N(x)$, ce qui montre que f est une fonction polynomiale.

Exercice 15.4 (Convergence uniforme, équicontinuité et convexité)

Leçon : 241. Référence : [6].

1. Soit (f_n) une suite de fonctions d'un segment $I = [a, b]$ de $\mathbb{R}$ à valeurs dans un $\mathbb{R}$ -espace vectoriel normé E . On suppose qu'il existe $K > 0$ tel que toutes les fonctions f_n soient K -lipschitziennes. Si (f_n) converge simplement vers une fonction $f : I \rightarrow E$, montrer que la convergence est uniforme.
2. Soit (f_n) une suite de fonctions convexes d'un intervalle $]a, b[$ de $\mathbb{R}$ dans $\mathbb{R}$ qui converge simplement vers une fonction f sur $]a, b[$. Montrer que la suite (f_n) converge uniformément vers f sur tout segment $[\alpha, \beta] \subset]a, b[$. La convergence est-elle uniforme sur $]a, b[$ tout entier ?

Corrigé.

1. La méthode est semblable à celle utilisée dans la preuve du deuxième théorème de Dini. On observe qu'une limite simple de fonctions K -lipschitziennes est une fonction K -lipschitzienne. C'est donc le cas de f . On considère ensuite une subdivision de $a = x_0 < x_1 < \dots < x_{k-1} < x_k = b$ telle que $x_{i+1} - x_i < \varepsilon$, avec $\varepsilon > 0$ fixé au préalable. En utilisant la convergence simple en tout point x_i , on trouve un entier N tel que $\|f_n(x_i) - f(x_i)\| < \varepsilon$ pour tous $i \in \{0, \dots, k\}$ et $n \geq N$. Enfin, pour tout $x \in [a, b]$, on a $x_i \leq x \leq x_{i+1}$ pour un certain i , et donc par inégalité triangulaire,

$$\begin{aligned} \|f_n(x) - f(x)\| &\leq \|f_n(x) - f_n(x_i)\| + \|f_n(x_i) - f(x_i)\| + \|f(x_i) - f(x)\| \\ &\leq K(x - x_i) + \varepsilon + K(x - x_i) \leq (2K + 1)\varepsilon. \end{aligned}$$

Cela étant vrai pour tous $x \in [a, b]$ et $n \geq N$, on en déduit la convergence uniforme.

2. On considère un segment $[\alpha, \beta] \subset]a, b[$, ainsi que deux réels α' et β' tels que $a < \alpha' < \alpha < \beta < \beta' < b$. Les deux suites dont le terme général est respectivement donné par $\frac{f_n(\alpha') - f_n(\alpha)}{\alpha' - \alpha}$ et $\frac{f_n(\beta) - f_n(\beta')}{\beta - \beta'}$ sont convergentes, et donc bornées. On note K une majorant des valeurs absolues de tous les termes de ces deux suites. La convexité donne une inégalité sur les pentes qui implique que pour tous $x, y \in [\alpha, \beta]$ distincts et tout n ,

$$-K \leq \frac{f_n(\alpha') - f_n(\alpha)}{\alpha' - \alpha} \leq \frac{f_n(x) - f_n(y)}{x - y} \leq \frac{f_n(\beta) - f_n(\beta')}{\beta - \beta'} \leq K.$$

Ainsi, toutes les fonctions f_n sont K -lipschitziennes sur $[\alpha, \beta]$. La question précédente montre que leur convergence vers f est uniforme sur $[\alpha, \beta]$.

Enfin, on observe que la suite de fonctions convexes $f_n(x) = x^n$ converge simplement vers 0 sur l'intervalle $]0, 1[$, mais ne converge pas uniformément (car $f_n(1 - 1/n)$ tend vers e^{-1}). Il peut donc ne pas y avoir convergence uniforme sur tout l'intervalle ouvert.

Exercice 15.5 (Série entière bornée à coefficients entiers)

Leçon : 241.

Soit f la somme d'une série entière $\sum a_n z^n$ de rayon de convergence au moins 1 telle que a_n est entier pour tout n . On suppose que f est bornée sur le disque unité ouvert. Montrer que f est une fonction polynomiale.

Corrigé.

D'après la formule de Parseval, on a $\sum_{n=0}^{\infty} |a_n|^2 r^{2n} = \frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\theta})|^2 d\theta$ pour tout $r \in]0, 1[$. En notant M un majorant de $|f|$ sur le disque unité ouvert, on a donc $\sum_{n=0}^{\infty} |a_n|^2 r^{2n} \leq M^2$. Les sommes partielles sont donc toutes majorées par M^2 , et admettent une limite quand $r \rightarrow 1^-$, puisque ce sont des sommes finies. En reprenant la somme totale, on a donc $\sum_{n=0}^{\infty} |a_n|^2 \leq M^2$. Cela implique que la suite (a_n) converge vers 0. Comme elle est constituée d'entiers, elle est donc nulle à partir d'un certain rang, ce qui implique que f est une fonction polynomiale.

Exercice 15.6 (Série exponentielle et transformation de Laplace)

Leçon : 241.

Soit Φ la somme d'une série entière $\sum a_n z^n$ de rayon de convergence $R > 0$.

1. Montrer que la série entière $\sum a_n z^n / n!$ a un rayon de convergence infini. On note φ sa somme.
2. Montrer que pour tout nombre complexe z tel que $|z| < R$, on a $\Phi(z) = \int_0^{+\infty} \varphi(zx) e^{-x} dx$.

Corrigé.

1. En prenant $r_0 \in]0, R[$, on voit que la suite $(a_n r_0^n)$ est bornée. De plus, pour tout $r > 0$, la suite $((r/r_0)^n / n!)$ est bornée. Par produit, la suite $(a_n r^n / n!)$ est donc bornée. On en déduit que la série entière $\sum a_n r^n / n!$ a un rayon de convergence infini.
2. Soit z un nombre complexe tel que $|z| < R$. L'intégrale s'écrit au moins formellement

$$\int_0^{+\infty} \varphi(zx) e^{-x} dx = \int_0^{+\infty} \sum_{n=0}^{\infty} a_n \frac{(zx)^n}{n!} e^{-x} dx.$$

Afin d'utiliser le théorème de Fubini, on écrit d'abord que

$$\sum_{n=0}^{\infty} \int_0^{+\infty} \left| a_n \frac{(zx)^n}{n!} e^{-x} \right| dx = \sum_{n=0}^{\infty} |a_n| \frac{|z|^n}{n!} \underbrace{\int_0^{+\infty} x^n e^{-x} dx}_{=n!} = \sum_{n=0}^{\infty} |a_n| |z|^n < \infty.$$

Pour le calcul de l'intégrale, on peut procéder par récurrence et intégrations par parties, ou se souvenir du lien avec le fonction Γ , puisque l'intégrale est $\Gamma(n+1) = n!$. Le théorème de Fubini, permet donc d'intervertir les symboles $\sum$ et $\int$, de sorte que

$$\int_0^{+\infty} \varphi(zx) e^{-x} dx = \sum_{n=0}^{\infty} \int_0^{+\infty} a_n \frac{(zx)^n}{n!} e^{-x} dx = \sum_{n=0}^{\infty} a_n \frac{z^n}{n!} \underbrace{\int_0^{+\infty} x^n e^{-x} dx}_{=n!} = \sum_{n=0}^{\infty} a_n z^n = \Phi(z).$$

Exercice 15.7 (Inégalité de Wirtinger)

Leçons : 241, 267. Référence : [7].

Soit f une fonction à valeurs complexes, 2π -périodique, de classe $\mathcal{C}^1$ et d'intégrale nulle sur une période. Montrer que

$$\int_0^{2\pi} |f(t)|^2 dt \leq \int_0^{2\pi} |f'(t)|^2 dt$$

et étudier le cas d'égalité.

Corrigé.

Pour tout entier $n \in \mathbb{Z}$, une simple intégration par parties montre clairement que $c_n(f') = in c_n(f)$. En appliquant l'identité de Bessel-Parseval, et en observant que le coefficient de Fourier $c_0(f)$ est nul puisque la fonction f est d'intégrale nulle, on obtient

$$\int_0^{2\pi} |f(t)|^2 dt = 2\pi \sum_{n \in \mathbb{Z} \setminus \{0\}} |c_n(f)|^2 \leq 2\pi \sum_{n \in \mathbb{Z}} |in c_n(f)|^2 = 2\pi \sum_{n \in \mathbb{Z} \setminus \{0\}} |c_n(f')|^2 = \int_0^{2\pi} |f'(t)|^2 dt,$$

ce qui donne l'inégalité annoncée. De plus, on observe que la différence des deux intégrales s'écrit simplement

$$2\pi \sum_{n \in \mathbb{Z} \setminus \{0\}} (n^2 - 1) |c_n(f)|^2.$$

Le cas d'égalité correspond donc au fait que $c_n(f) = 0$ dès que $|n| \geq 2$. Les seules fonctions possibles sont donc $t \mapsto a e^{it} + b e^{-it}$, avec $a, b \in \mathbb{C}$.

Intégration et intégrales à paramètre

Références : [6, 7, 14].

La fonction Gamma d'Euler est définie pour $\Re z > 0$ par $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$, avec $t^{z-1} = \exp((z-1)\log t)$. Elle vérifie $\Gamma(z+1) = z\Gamma(z)$ et $\Gamma(n+1) = n!$ pour $n \in \mathbb{N}$.

Exercice 16.1 (Calcul de l'intégrale de Gauss)

Leçons : 236, 239, 265. Référence : [6].

Le but de cet exercice est de calculer la valeur de l'intégrale de Gauss $I = \int_0^{+\infty} e^{-x^2} dx$.

1. *Méthode utilisant une relation fonctionnelle* : On considère les fonctions f et g définies sur $[0, +\infty[$ par

$$f(t) = \int_0^t e^{-x^2} dx \quad \text{et} \quad g(t) = \int_0^1 \frac{e^{-(x^2+1)t^2}}{x^2+1} dx.$$

Donner une relation fonctionnelle entre f et g , et en déduire la valeur de I .

2. *Méthode utilisant une intégrale double* : Exprimer I^2 comme l'intégrale d'une fonction de deux variables, puis passer en coordonnées polaires dans l'intégrale. En déduire la valeur de I .

3. Déduire de ce qui précède la valeur de $\Gamma\left(\frac{1}{2}\right) = \int_0^\infty \frac{e^{-x}}{\sqrt{x}} dx$.

Corrigé.

1. À l'aide du théorème de dérivation des intégrales à paramètre, on peut montrer que la fonction g est dérivable sur $[0, +\infty[$, et que pour tout réel t de cet intervalle,

$$g'(t) = \int_0^1 -2te^{-(x^2+1)t^2} dx = -2e^{-t^2} \int_0^1 e^{-x^2 t^2} dx = -2e^{-t^2} \int_0^t e^{-u^2} du = -2f'(t)f(t).$$

En intégrant, on en déduit que pour tout réel $t \geq 0$,

$$g(t) - g(0) = -f(t)^2 + f(0)^2, \quad \text{donc} \quad g(t) = -f(t)^2 + \int_0^1 \frac{dx}{x^2+1} = \frac{\pi}{4} - f(t)^2.$$

On observe de plus que $g(t) \xrightarrow{t \rightarrow +\infty} 0$ car pour tout réel $t \geq 0$,

$$0 \leq g(t) \leq e^{-t^2} \int_0^1 \frac{dx}{x^2+1} = \frac{\pi}{4} e^{-t^2}.$$

Comme la fonction f est positive, on en conclut que $f(t) \xrightarrow{t \rightarrow +\infty} \sqrt{\frac{\pi}{4}}$, ce qui donne $I = \frac{\sqrt{\pi}}{2}$.

2. En procédant au changement de variables usuel pour passer en coordonnées polaires, on peut écrire que

$$I^2 = \int_{[0, \infty[^2} e^{-(x^2+y^2)} dx dy = \int_0^{\pi/2} \int_0^\infty e^{-r^2} r dr d\theta = \int_0^{\pi/2} \left[-\frac{e^{-r^2}}{2} \right]_0^\infty d\theta = \frac{\pi}{4},$$

de sorte qu'on retrouve $I = \frac{\sqrt{\pi}}{2}$.

3. Pour faire le lien avec $\Gamma(1/2)$, on effectue le changement de variable $x = t^2$ dans l'intégrale

$$\Gamma\left(\frac{1}{2}\right) = \int_0^\infty \frac{e^{-x}}{\sqrt{x}} dx = \int_0^\infty \frac{e^{-t^2}}{t} 2t dt = 2 \int_0^\infty e^{-t^2} dt = \int_{-\infty}^{+\infty} e^{-t^2} dt = 2I = \sqrt{\pi}.$$

Exercice 16.2 (Transformée de Fourier de la gaussienne)

Leçons : 236, 239, 250. Référence : [14].

La transformée de Fourier d'une fonction $f \in L^1(\mathbb{R})$ est définie ici par la formule $\widehat{f}(\xi) = \int_{-\infty}^{+\infty} f(x) e^{-i\xi x} dx$. On propose trois méthodes pour montrer que, pour $\alpha > 0$, la transformée de Fourier de la fonction gaussienne $\varphi_\alpha : x \mapsto e^{-\alpha x^2}$ est la fonction $\widehat{\varphi}_\alpha : \xi \mapsto \sqrt{\frac{\pi}{\alpha}} e^{-\xi^2/(4\alpha)}$. On propose ensuite d'étendre ce résultat au cas d'un paramètre complexe, ainsi qu'au cas vectoriel.

1. Utiliser le résultat de l'exercice précédent pour indiquer la valeur de l'intégrale $\int_{-\infty}^{+\infty} e^{-x^2} dx$.
2. *Méthode utilisant une équation différentielle* : Montrer que la transformée de Fourier $\widehat{\varphi}_\alpha$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$, et solution d'une équation différentielle simple. En résolvant cette équation, déterminer l'expression de $\widehat{\varphi}_\alpha$.
3. *Méthode utilisant un prolongement analytique* : Montrer que la fonction $\Phi_\alpha : z \mapsto \int_{-\infty}^{+\infty} e^{-\alpha x^2} e^{-xz} dx$ est holomorphe sur $\mathbb{C}$. En donner une expression simple quand z est réel. Déterminer alors l'expression de $\widehat{\varphi}_\alpha$ par prolongement analytique.
4. *Méthode utilisant une intégrale de chemin* : Retrouver l'expression de $\widehat{\varphi}_\alpha$ en intégrant la fonction $z \mapsto e^{-\alpha z^2}$ sur tout rectangle du plan complexe défini par les points $-R$ et $R + i\frac{\xi}{2\alpha}$, orienté arbitrairement, avec $R \rightarrow +\infty$.
5. *Extension au cas d'un paramètre complexe* : On choisit la détermination principale $\text{Log } z = \log |z| + i \text{Arg } z$ sur $\mathbb{C} \setminus]-\infty, 0]$, avec $-\pi < \text{Arg } z < \pi$, et $\sqrt{z} = \exp(\text{Log } z/2)$. Avec cette convention, montrer que la formule $\widehat{\varphi}_\alpha(\xi) = \sqrt{\frac{\pi}{\alpha}} e^{-\xi^2/(4\alpha)}$ reste vraie pour $\alpha \in \mathbb{C}$ tel que $\Re \alpha > 0$.
6. *Extension au cas vectoriel* : La transformée de Fourier d'une fonction f sur $\mathbb{R}^n$ est définie ici par la formule $\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-i\xi \cdot x} dx$ pour $\xi \in \mathbb{R}^n$. Soit A une matrice symétrique définie positive. Déterminer la transformée de Fourier de la fonction définie sur $\mathbb{R}^n$ par $\Phi_A : x \mapsto e^{-Ax \cdot x}$.

Corrigé.

1. La fonction $x \mapsto e^{-x^2}$ étant paire, son intégrale sur $\mathbb{R}$ est le double de son intégrale sur $[0, \infty[$. On a donc $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$.
2. La transformée de Fourier considérée s'écrit $\widehat{\varphi}_\alpha(\xi) = \int_{-\infty}^{+\infty} e^{-\alpha x^2} e^{-i\xi x} dx$. On peut lui appliquer le théorème de dérivation sous le signe intégrale, ce qui montre que la fonction $\widehat{\varphi}_\alpha$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et vérifie pour tout ξ ,

$$\widehat{\varphi}_\alpha'(\xi) = \int_{-\infty}^{+\infty} -ix e^{-\alpha x^2} e^{-i\xi x} dx = \left[\frac{i}{2\alpha} e^{-\alpha x^2} e^{-i\xi x} \right]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} \frac{\xi}{2\alpha} e^{-\alpha x^2} e^{-i\xi x} dx = -\frac{\xi}{2\alpha} \widehat{\varphi}_\alpha(\xi).$$

On en déduit l'existence d'une constante C telle que $\widehat{\varphi}_\alpha(\xi) = C e^{-\xi^2/(4\alpha)}$ pour tout $\xi \in \mathbb{R}$. Afin de déterminer la valeur de C , il suffit de fixer $\xi = 0$, ce qui donne $C = \widehat{\varphi}_\alpha(0) = \int_{-\infty}^{+\infty} e^{-\alpha x^2} dx = \int_{-\infty}^{+\infty} e^{-y^2} \frac{dy}{\sqrt{\alpha}} = \sqrt{\frac{\pi}{\alpha}}$, et finalement l'expression $\widehat{\varphi}_\alpha(\xi) = \sqrt{\frac{\pi}{\alpha}} e^{-\xi^2/(4\alpha)}$ pour tout $\xi \in \mathbb{R}$.

3. On peut montrer que la fonction Φ_α est holomorphe sur $\mathbb{C}$ en utilisant le théorème d'holomorphicité des intégrales à paramètre. De plus, si z désigne un nombre réel, on peut écrire que

$$\Phi_\alpha(z) = \int_{-\infty}^{+\infty} e^{-\alpha x^2} e^{-xz} dx = \int_{-\infty}^{+\infty} \exp\left(\frac{z^2}{4\alpha} - \alpha\left(x + \frac{z}{2\alpha}\right)^2\right) dx = e^{z^2/(4\alpha)} \int_{-\infty}^{+\infty} e^{-\alpha y^2} dy = \sqrt{\frac{\pi}{\alpha}} e^{z^2/(4\alpha)}.$$

Le membre de gauche et le membre de droite définissent deux fonctions holomorphes sur $\mathbb{C}$ qui coïncident sur $\mathbb{R}$. Par unicité du prolongement analytique, ces deux fonctions coïncident sur $\mathbb{C}$ tout entier. En prenant alors $z = i\xi$, pour $\xi \in \mathbb{R}$, on retrouve l'expression $\widehat{\varphi}_\alpha(\xi) = \sqrt{\frac{\pi}{\alpha}} e^{-\xi^2/(4\alpha)}$.

4. La fonction $z \mapsto e^{-\alpha z^2}$ est holomorphe sur $\mathbb{C}$, donc pour tout $R > 0$, son intégrale sur le rectangle défini par les points $-R$ et $R + i\frac{\xi}{2\alpha}$, parcouru successivement par les sommets $-R$, R , $R + i\xi/(2\alpha)$ et $-R + i\xi/(2\alpha)$ (le cas $\xi = 0$ étant déjà connu), est nulle. Cette intégrale se décompose comme

$$\int_{-R}^R e^{-\alpha x^2} dx + \int_0^{\xi/(2\alpha)} e^{-\alpha(R+it)^2} i dt - \int_{-R}^R e^{-\alpha\left(x+i\frac{\xi}{2\alpha}\right)^2} dx - \int_0^{\xi/(2\alpha)} e^{-\alpha(-R+it)^2} i dt = 0.$$

Quand $R \rightarrow +\infty$, le premier terme tend vers l'intégrale sur $\mathbb{R}$ tout entier, soit $\sqrt{\pi/\alpha}$, comme calculé dans les questions précédentes. La troisième intégrale tend vers

$$\int_{-\infty}^{+\infty} e^{-\alpha\left(x+i\frac{\xi}{2\alpha}\right)^2} dx = e^{\xi^2/(4\alpha)} \int_{-\infty}^{+\infty} e^{-\alpha x^2} e^{-ix\xi} dx = e^{\xi^2/(4\alpha)} \widehat{\varphi}_\alpha(\xi).$$

Enfin, les deux intégrales restantes vérifient

$$\left| \int_0^{\xi/(2\alpha)} e^{-\alpha(R+it)^2} i dt \right| \leq e^{-\alpha R^2} \int_0^{|\xi|/(2\alpha)} e^{\alpha t^2} dt \xrightarrow{R \rightarrow \pm\infty} 0.$$

On en conclut que $\sqrt{\frac{\pi}{\alpha}} - e^{\xi^2/(4\alpha)} \widehat{\varphi}_\alpha(\xi) = 0$ pour tout $\xi \in \mathbb{R}$, ce qui donne à nouveau $\widehat{\varphi}_\alpha(\xi) = \sqrt{\frac{\pi}{\alpha}} e^{-\xi^2/(4\alpha)}$.

5. En utilisant la détermination principale du logarithme, on peut poser $\sqrt{z} = e^{(\log z)/2}$ pour $\Re z > 0$. Ainsi, pour $\xi \in \mathbb{R}$ fixé, la fonction $\alpha \mapsto \sqrt{\frac{\pi}{\alpha}} e^{-\xi^2/(4\alpha)}$ s'étend en une fonction holomorphe sur le domaine $\Re \alpha > 0$. De plus, en utilisant le théorème d'holomorphie des intégrales à paramètre, on peut montrer que la fonction $\alpha \mapsto \int_{-\infty}^{+\infty} e^{-\alpha x^2} e^{-i\xi x} dx$ est holomorphe sur le même domaine. On conclut par unicité du prolongement analytique, puisque la formule est vraie quand α est réel strictement positif.
6. La matrice A est diagonalisable dans une base orthonormée. On note $\alpha_1, \dots, \alpha_n > 0$ ses valeurs propres et $x_1, \dots, x_n$ les coordonnées d'un point générique $x \in \mathbb{R}^n$ dans une base orthonormée dans laquelle A est diagonalisable. Le changement de base est donné par une matrice orthogonale, donc induit un changement de variable dont la valeur absolue du jacobien vaut 1. On a donc

$$\begin{aligned} \widehat{\Phi}_A(\xi) &= \int_{\mathbb{R}^n} e^{-Ax \cdot x} e^{-i\xi \cdot x} dx = \int_{\mathbb{R}^n} e^{-(\alpha_1 x_1^2 + \dots + \alpha_n x_n^2)} e^{-i(\xi_1 x_1 + \dots + \xi_n x_n)} dx \\ &= \prod_{j=1}^n \int_{-\infty}^{+\infty} e^{-\alpha_j x_j^2} e^{-i\xi_j x_j} dx_j = \prod_{j=1}^n \widehat{\varphi}_{\alpha_j}(\xi_j) = \prod_{j=1}^n \sqrt{\frac{\pi}{\alpha_j}} e^{-\xi_j^2/(4\alpha_j)} = \frac{\pi^{n/2}}{\sqrt{\det A}} e^{-A^{-1}\xi \cdot \xi/4}. \end{aligned}$$

Exercice 16.3 (Calcul du volume de la boule unité)

Leçon : 236.

1. Montrer que le volume de la boule euclidienne unité fermée de $\mathbb{R}^n$ s'écrit $v_n = \int_{\mathbb{R}^n} \mathbf{1}_{\{x_1^2 + \dots + x_n^2 \leq 1\}} dx_1 \dots dx_n$.
2. Montrer que pour tout entier $n \geq 1$, on a $v_{n+1} = I_n v_n$, avec $I_n = \int_{-1}^1 (1 - x^2)^{n/2} dx$.
3. Montrer que pour tout entier $n \geq 2$, on a $I_n = \frac{n}{n+1} I_{n-2}$.
4. En déduire l'expression de v_n .

Corrigé.

1. La boule euclidienne unité fermée de $\mathbb{R}^n$ est l'ensemble des points $(x_1, \dots, x_n)$ tels que $x_1^2 + \dots + x_n^2 \leq 1$. Sachant que le volume est donné par la mesure de Lebesgue, on en déduit que le volume de la boule euclidienne unité fermée s'exprime comme $v_n = \int_{\mathbb{R}^n} \mathbf{1}_{\{x_1^2 + \dots + x_n^2 \leq 1\}} dx_1 \dots dx_n$.
2. En utilisant le théorème de Fubini, on peut écrire que

$$v_{n+1} = \int_{\mathbb{R}^{n+1}} \mathbf{1}_{\{x_1^2 + \dots + x_{n+1}^2 \leq 1\}} dx_1 \dots dx_{n+1} = \int_{-1}^1 \int_{\mathbb{R}^n} \mathbf{1}_{\{x_1^2 + \dots + x_n^2 \leq 1-x^2\}} dx_1 \dots dx_n dx.$$

On reconnaît à l'intérieur le volume de la boule euclidienne de $\mathbb{R}^n$, centrée à l'origine et de rayon $(1 - x^2)^{1/2}$. Ce volume vaut $(1 - x^2)^{n/2}$ fois celui de la boule unité fermée. On a donc

$$v_{n+1} = \int_{-1}^1 (1 - x^2)^{n/2} v_n dx = I_n v_n \quad \text{avec} \quad I_n = \int_{-1}^1 (1 - x^2)^{n/2} dx.$$

3. Pour $n \geq 2$, on observe que

$$I_n = \int_{-1}^1 (1 - x^2)^{n/2} dx = \int_{-1}^1 (1 - x^2)(1 - x^2)^{n/2-1} dx = I_{n-2} - \int_{-1}^1 x^2(1 - x^2)^{n/2-1} dx.$$

De plus, à l'aide d'une intégration par parties, on a

$$\int_{-1}^1 x^2(1 - x^2)^{n/2-1} dx = -\frac{1}{n} \left[x(1 - x^2)^{n/2} \right]_{-1}^1 + \frac{1}{n} \int_{-1}^1 (1 - x^2)^{n/2} dx = \frac{I_n}{n}.$$

On en déduit que $I_n = I_{n-2} - \frac{I_n}{n}$, d'où l'on tire finalement $I_n = \frac{n}{n+1} I_{n-2}$.

4. Montrons par récurrence que pour tout entier $n \geq 2$, on a $I_{n-1} I_{n-2} = \frac{2\pi}{n}$. En effet, au rang initial, on observe que $I_1 I_0 = \frac{\pi}{2} 2 = \pi = \frac{2\pi}{2}$. Ensuite, si la relation est vraie au rang n , alors la relation obtenue ci-dessus implique

$$I_n I_{n-1} = \frac{n}{n+1} I_{n-2} I_{n-1} = \frac{n}{n+1} \frac{2\pi}{n} = \frac{2\pi}{n+1}.$$

Par conséquent, on a $v_n = I_{n-1} I_{n-2} v_{n-2} = \frac{2\pi}{n} v_{n-2}$ pour tout entier $n \geq 3$. Selon la parité de n , on a donc

$$v_{2k} = \frac{2\pi}{2k} \frac{2\pi}{2(k-1)} \dots \frac{2\pi}{4} v_2 = \frac{\pi^k}{k!} \quad \text{et} \quad v_{2k+1} = \frac{2\pi}{2k+1} \frac{2\pi}{2k-1} \dots \frac{2\pi}{3} v_1 = \frac{\pi^k}{(k+1/2)(k-1/2) \dots (3/2)(1/2)},$$

car on sait que $v_2 = \pi$ et $v_1 = 2$. On peut condenser ces résultats dans la formule $v_n = \frac{\pi^{n/2}}{\Gamma(n/2 + 1)}$.

Exercice 16.4 (Calcul de l'aire de la sphère unité)

Leçon : 236.

1. Donner la valeur de l'intégrale

$$I_n = \int_{\mathbb{R}^n} e^{-(x_1^2 + \dots + x_n^2)} dx_1 \dots dx_n.$$

2. À l'aide d'un passage en coordonnées sphériques, déterminer la valeur de $\sigma(\mathbb{S}^{n-1})$, où $\mathbb{S}^{n-1}$ désigne la sphère euclidienne unité et σ est la mesure de surface correspondante.
3. En déduire le volume de la boule euclidienne unité fermée de $\mathbb{R}^n$.

Corrigé.

1. À l'aide du théorème de Fubini et de la valeur de l'intégrale de Gauss déterminée dans un exercice précédent, on écrit

$$I_n = \int_{\mathbb{R}^n} e^{-(x_1^2 + \dots + x_n^2)} dx_1 \dots dx_n = \left(\int_{-\infty}^{\infty} e^{-x^2} dx \right)^n = \pi^{n/2}.$$

2. En passant en coordonnées sphériques dans l'intégrale, on a

$$I_n = \int_{\mathbb{S}^{n-1}} \int_0^{\infty} e^{-r^2} r^{n-1} dr \sigma(ds) = \sigma(\mathbb{S}^{n-1}) \int_0^{\infty} e^{-r^2} r^{n-1} dr.$$

Il reste à calculer cette dernière intégrale. Le changement de variable $t = r^2$ donne

$$\int_0^{\infty} e^{-r^2} r^{n-1} dr = \int_0^{\infty} t^{(n-1)/2} e^{-t} \frac{dt}{2\sqrt{t}} = \frac{1}{2} \int_0^{\infty} t^{n/2-1} e^{-t} dt = \frac{1}{2} \Gamma\left(\frac{n}{2}\right).$$

On en conclut que $\sigma(\mathbb{S}^{n-1}) = \frac{2\pi^{n/2}}{\Gamma(n/2)}$.

3. En notant
- v_n
- le volume de la boule euclidienne unité fermée, on a en coordonnées sphériques

$$v_n = \int_{\mathbb{S}^{n-1}} \int_0^1 r^{n-1} dr \sigma(ds) = \sigma(\mathbb{S}^{n-1}) \left[\frac{r^n}{n} \right]_0^1 = \frac{2\pi^{n/2}}{n\Gamma(n/2)} = \frac{\pi^{n/2}}{\Gamma(n/2 + 1)}.$$

Exercice 16.5 (Intégrale de Frullani)

Leçon : 239. Référence : [6].

Soit $f : [0, +\infty[\rightarrow \mathbb{R}$ une fonction continue telle que l'intégrale $\int_1^{+\infty} \frac{f(t)}{t} dt$ converge.

1. Soient deux nombres réels $a, b > 0$. Montrer que l'intégrale $\int_0^{+\infty} \frac{f(at) - f(bt)}{t} dt$ converge et calculer sa valeur.
2. *Application.* Pour $a, b > 0$, calculer la valeur de l'intégrale $\int_0^{+\infty} \frac{e^{-at} - e^{-bt}}{t} dt$.

Corrigé.

1. On pose $g(t) = \frac{f(at) - f(bt)}{t}$ pour tout réel $t > 0$. Un simple changement de variable montre que pour tout $\alpha > 0$, l'intégrale $\int_1^{+\infty} \frac{f(\alpha t)}{t} dt$ est de même nature que l'intégrale $\int_{\alpha}^{+\infty} \frac{f(t)}{t} dt$, c'est-à-dire convergente. On en déduit que l'intégrale $\int_1^{+\infty} g(t) dt$ converge. Il reste à prouver que l'intégrale converge jusqu'à la borne 0. Pour tout réel $x > 0$, en effectuant de simples changements de variables, on a

$$\int_x^{+\infty} g(t) dt = \int_{ax}^{+\infty} \frac{f(u)}{u} du - \int_{bx}^{+\infty} \frac{f(u)}{u} du = \int_{ax}^{bx} \frac{f(u)}{u} du.$$

Comme la fonction f est continue, la formule de la moyenne sur le segment de bornes ax et bx (avec intégrales orientées si $a > b$) montre l'existence d'un réel c_x de cet intervalle pour lequel

$$\int_{ax}^{bx} \frac{f(u)}{u} du = f(c_x) \int_{ax}^{bx} \frac{du}{u} = f(c_x) \log \frac{b}{a} \xrightarrow{x \rightarrow 0} f(0) \log \frac{b}{a}.$$

On en déduit que l'intégrale converge et vaut $\int_0^{+\infty} \frac{f(at) - f(bt)}{t} dt = f(0) \log \frac{b}{a}$.

2. En appliquant ce qui précède à la fonction $f(t) = e^{-t}$, on obtient directement $\int_0^{+\infty} \frac{e^{-at} - e^{-bt}}{t} dt = \log \frac{b}{a}$.

Exercice 16.6 (Intégrales à paramètre et fonction Gamma)

Leçon : 239. Référence : [6].

1. Pour tout réel $x \geq 0$, calculer la valeur de l'intégrale $I(x) = \int_0^{+\infty} \frac{\sin(xt)}{t} e^{-t} dt$.
2. Pour tous réels $\alpha > 0$ et $x \geq 0$, exprimer la valeur de l'intégrale $I_\alpha(x) = \int_0^{+\infty} e^{-t} e^{ixt} t^{\alpha-1} dt$ à l'aide de la fonction Γ . En déduire la valeur de $J(\alpha) = \int_0^{+\infty} t^{\alpha-1} e^{it} dt$ quand $0 < \alpha < 1$.

Corrigé.

1. En utilisant le théorème de dérivation des intégrales à paramètres, on peut montrer que la fonction I est dérivable sur $[0, +\infty[$, et que pour tout réel x de cet intervalle,

$$I'(x) = \int_0^{+\infty} \cos(xt) e^{-t} dt = \Re \left(\int_0^{+\infty} e^{(ix-1)t} dt \right) = \Re \left(\left[\frac{e^{(ix-1)t}}{ix-1} \right]_0^{+\infty} \right) = \Re \left(\frac{1}{1-ix} \right) = \frac{1}{1+x^2}.$$

On en déduit par intégration que pour tout réel $x \geq 0$,

$$I(x) = I(0) + \int_0^x \frac{dt}{1+t^2} = \arctan x.$$

2. À nouveau en utilisant le théorème de dérivation des intégrales à paramètres, on peut montrer que la fonction I_α est dérivable sur $[0, +\infty[$, et que pour tout réel x de cet intervalle,

$$I'_\alpha(x) = \int_0^{+\infty} it e^{-t} e^{ixt} t^{\alpha-1} dt.$$

En outre, une intégration par parties sur l'intégrale définissant $I_\alpha(x)$ donne pour tout réel $x \geq 0$,

$$I_\alpha(x) = \left[e^{(ix-1)t} \frac{t^\alpha}{\alpha} \right]_0^{+\infty} - \frac{ix-1}{\alpha} \int_0^{+\infty} e^{-t} e^{ixt} t^\alpha dt = -\frac{i+x}{\alpha} I'_\alpha(x).$$

On peut intégrer l'équation différentielle ainsi obtenue en calculant

$$\int \frac{dx}{i+x} = \int \frac{x-i}{x^2+1} dx = \frac{1}{2} \log(x^2+1) - i \arctan x + C.$$

On en déduit qu'il existe une constante $\lambda \in \mathbb{C}$ telle que pour tout réel $x \geq 0$,

$$I_\alpha(x) = \lambda \exp \left(-\frac{\alpha}{2} \log(x^2+1) + i\alpha \arctan x \right) = \frac{\lambda e^{i\alpha \arctan x}}{(x^2+1)^{\alpha/2}}.$$

Pour déterminer la valeur de la constante, on évalue I_α en 0, ce qui donne $\lambda = I_\alpha(0) = \Gamma(\alpha)$. Finalement, pour tout réel $x \geq 0$, on a

$$I_\alpha(x) = \frac{\Gamma(\alpha) e^{i\alpha \arctan x}}{(x^2+1)^{\alpha/2}}.$$

Pour faire ensuite le lien avec la fonction $J(\alpha)$, on procède au changement de variable $u = xt$ dans l'intégrale définissant $I_\alpha(x)$. Ainsi, pour $x > 0$, on a

$$I_\alpha(x) = \frac{1}{x^\alpha} \int_0^{+\infty} e^{-u/x} e^{iu} u^{\alpha-1} du.$$

L'idée est maintenant de faire tendre x vers l'infini dans l'intégrale, puisqu'on voit formellement que la limite est égale à $J(\alpha)$. On ne peut pas utiliser directement le théorème de convergence dominée sur l'intervalle $]0, +\infty[$ tout entier, mais on peut commencer par l'utiliser sur $]0, 1]$ puisque l'intégrande est borné par $u^{\alpha-1}$, qui est intégrable sur cet intervalle car $\alpha > 0$. Sur l'intervalle $[1, +\infty[$, on procède à une intégration par parties :

$$\int_1^{+\infty} e^{-u/x} e^{iu} u^{\alpha-1} du = \left[\frac{e^{(i-1/x)u}}{i-1/x} u^{\alpha-1} \right]_1^{+\infty} - \frac{\alpha-1}{i-1/x} \int_1^{+\infty} e^{(i-1/x)u} u^{\alpha-2} du$$

On peut appliquer le théorème de convergence dominée à la seconde intégrale car l'intégrande est borné par $u^{\alpha-2}$, qui est intégrable sur l'intervalle $[1, +\infty[$ car $\alpha < 1$. Par conséquent,

$$\int_0^{+\infty} e^{-u/x} e^{iu} u^{\alpha-1} du \xrightarrow{x \rightarrow +\infty} \int_0^1 e^{iu} u^{\alpha-1} du - \frac{e^i}{i} - \frac{\alpha-1}{i} \int_1^{+\infty} e^{iu} u^{\alpha-2} du = J(\alpha).$$

Pour obtenir cette dernière égalité, on lit l'intégration par parties dans l'autre sens en se plaçant d'abord sur des intervalles de la forme $[1, N]$ car l'intégrale est seulement semi-convergente. On en déduit que

$$J(\alpha) = \lim_{x \rightarrow +\infty} \int_0^{+\infty} e^{-u/x} e^{iu} u^{\alpha-1} du = \lim_{x \rightarrow +\infty} x^\alpha I_\alpha(x) = \lim_{x \rightarrow +\infty} \frac{x^\alpha \Gamma(\alpha) e^{i\alpha \arctan x}}{(x^2 + 1)^{\alpha/2}} = \Gamma(\alpha) e^{i\alpha\pi/2}.$$

Exercice 16.7 (Logarithme intégral)

Leçons : 239, 265. Référence : [6].

On considère la fonction *logarithme intégral* définie par $\text{Li}(x) = \int_2^x \frac{dt}{\log t}$ pour tout réel $x \geq 2$. Donner un développement asymptotique à l'infini de Li à n'importe quel ordre.

Corrigé.

En procédant à n intégrations par parties, on trouve pour $x \geq 2$,

$$\text{Li}(x) = \sum_{k=0}^{n-1} \left[\frac{k! t}{(\log t)^{k+1}} \right]_2^x + \int_2^x \frac{n! dt}{(\log t)^{n+1}}.$$

Cependant, dans cette dernière intégrale, l'intégrande est équivalent à la dérivée de $t \mapsto \frac{n! t}{(\log t)^{n+1}}$ quand t tend vers l'infini. Comme l'intégrale diverge, on peut intégrer les équivalents, et conclure que

$$\text{Li}(x) = \sum_{k=0}^n \frac{k! x}{(\log x)^{k+1}} + o_{x \rightarrow +\infty} \left(\frac{x}{(\log x)^{n+1}} \right).$$

Exercice 16.8 (Translations et convergence en moyenne)

Leçon : 239. Référence : [7].

Soit une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ continue et intégrable sur $\mathbb{R}$. On pose

$$I(a) = \int_{-\infty}^{+\infty} |f(t+a) - f(t)| dt.$$

1. Montrer que $I(a) \xrightarrow{a \rightarrow 0} 0$.
2. Montrer que $I(a)$ admet une limite finie quand $a \rightarrow +\infty$, et calculer sa valeur.

Corrigé.

1. L'intégrabilité de f ne fournit pas à elle seule une domination intégrable uniforme des translatées pour a voisin de zéro. Cependant, en utilisant la continuité et l'intégrabilité de f sur $\mathbb{R}$, on écrit pour tout réel $X > 0$,

$$\begin{aligned} I(a) &= \int_{-\infty}^{-X} |f(t+a) - f(t)| dt + \int_{-X}^X |f(t+a) - f(t)| dt + \int_X^{+\infty} |f(t+a) - f(t)| dt \\ &\leq \int_{|t| \geq X} |f(t)| dt + \int_{|t| \geq X} |f(t+a)| dt + 2X \sup_{|t| \leq X} |f(t+a) - f(t)| \\ &\leq \int_{|t| \geq X} |f(t)| dt + \int_{|s| \geq X-|a|} |f(s)| ds + 2X \sup_{|t| \leq X} |f(t+a) - f(t)|. \end{aligned}$$

Par continuité de f , et donc par uniforme continuité sur les compacts, le supremum ci-dessus tend vers 0 quand a tend vers 0, pour tout X fixé. On a donc

$$\limsup_{a \rightarrow 0} I(a) \leq 2 \int_{|t| \geq X} |f(t)| dt \xrightarrow{X \rightarrow +\infty} 0.$$

2. Un simple changement de variable assure que

$$I(a) = \int_{-\infty}^{+\infty} \left| f\left(s + \frac{a}{2}\right) - f\left(s - \frac{a}{2}\right) \right| ds = \int_{\mathbb{R}} |f_{a/2} - f_{-a/2}|,$$

où l'on a posé $f_b(s) = f(s+b)$ pour $s, b \in \mathbb{R}$. Pour $b > 0$, on a

$$\begin{aligned} \left| \int_{\mathbb{R}^+} |f_b - f_{-b}| - \int_{\mathbb{R}} |f| \right| &= \left| \int_{\mathbb{R}^+} |f_b - f_{-b}| - \int_{\mathbb{R}} |f_{-b}| \right| \\ &= \left| \int_{\mathbb{R}^+} (|f_b - f_{-b}| - |f_{-b}|) - \int_{\mathbb{R}^-} |f_{-b}| \right| \\ &\leq \int_{\mathbb{R}^+} ||f_b - f_{-b}| - |f_{-b}|| + \int_{\mathbb{R}^-} |f_{-b}|. \end{aligned}$$

D'après l'inégalité triangulaire, on a $||f_b - f_{-b}| - |f_{-b}|| \leq |f_b|$, ce qui donne

$$\int_{\mathbb{R}^+} ||f_b - f_{-b}| - |f_{-b}|| + \int_{\mathbb{R}^-} |f_{-b}| \leq \int_{\mathbb{R}^+} |f_b| + \int_{\mathbb{R}^-} |f_{-b}| = \int_{|t| \geq b} |f(t)| dt.$$

Par symétrie, on a le même résultat pour l'intégrale sur $\mathbb{R}^-$ de $|f_b - f_{-b}|$. Finalement,

$$\left| I(a) - 2 \int_{\mathbb{R}} |f| \right| \leq 2 \int_{|t| \geq a/2} |f(t)| dt \xrightarrow{a \rightarrow +\infty} 0.$$

On en conclut que $I(a)$ converge vers $2 \int_{\mathbb{R}} |f|$ quand $a \rightarrow +\infty$.

Convexité

Références : [6, 13, 16, 18].

L'enveloppe convexe $\text{conv}(A)$ d'une partie A est l'ensemble de ses combinaisons convexes finies $\sum_{j=1}^m \lambda_j x_j$, où $x_j \in A$, $\lambda_j \geq 0$ et $\sum_j \lambda_j = 1$; on pose $\text{conv}(\emptyset) = \emptyset$. Le barycentre de points x_j affectés de coefficients $a_j \geq 0$, non tous nuls, est $(\sum_j a_j x_j) / (\sum_j a_j)$.

Exercice 17.1 (Autour du théorème de Carathéodory)

Leçon : 253. Références : [6, 13, 18].

1. Soit C un convexe d'un espace vectoriel normé E . Montrer que l'adhérence $\bar{C}$ et l'intérieur $\overset{\circ}{C}$ sont convexes.
2. (Carathéodory) Soit E un $\mathbb{R}$ -espace vectoriel normé de dimension $n \in \mathbb{N}^*$, et soit x le barycentre de p vecteurs $x_1, \dots, x_p \in E$ affectés de coefficients positifs. Montrer qu'il existe $I \subset \{1, \dots, p\}$ de cardinal au plus $n + 1$ tel que x soit le barycentre des points x_i , avec $i \in I$, affectés de coefficients positifs.
3. Soit E un $\mathbb{R}$ -espace vectoriel de dimension finie, et soit A une partie compacte de E . Montrer que l'enveloppe convexe de A est compacte.
4. On suppose maintenant que E est un espace vectoriel normé de dimension infinie. Montrer qu'il existe un compact A de E dont l'enveloppe convexe n'est pas fermée.

Corrigé.

1. Soit $\lambda \in [0, 1]$ et $x, y \in \bar{C}$. Il existe une suite (x_n) dans C qui converge vers x , et de même une suite (y_n) dans C qui converge vers y . Par convexité, le point $\lambda x_n + (1 - \lambda)y_n$ appartient à C pour tout n , et par passage à la limite, on en déduit que le point $\lambda x + (1 - \lambda)y$ appartient à $\bar{C}$. Ainsi, cet ensemble est convexe.
Soit $\lambda \in [0, 1]$ et $x, y \in \overset{\circ}{C}$. Il existe un réel $\rho > 0$ tel que les boules ouvertes de centres respectifs x et y , et de rayon ρ soient incluses dans C . Pour tout vecteur h de norme strictement inférieure à ρ , on a $\lambda x + (1 - \lambda)y + h = \lambda(x + h) + (1 - \lambda)(y + h) \in C$, par convexité de C , puisque cet ensemble contient les points $x + h$ et $y + h$. Ainsi, la boule ouverte de centre $\lambda x + (1 - \lambda)y$ et de rayon ρ est incluse dans C , ce qui montre que $\overset{\circ}{C}$ est convexe.
2. On note Γ l'ensemble des parties $J \subset \{1, \dots, p\}$ telles que x est barycentre des points x_i , avec $i \in J$, affectés de coefficients positifs. On note ensuite q le cardinal minimal d'une telle partie. On veut montrer que $q \leq n + 1$.
On suppose au contraire que $q \geq n + 2$. On a donc, pour une certaine partie J de cardinal q ,

$$x = \sum_{i \in J} \lambda_i x_i \quad \text{avec} \quad \sum_{i \in J} \lambda_i = 1 \text{ et } \lambda_i \geq 0.$$

Comme E est de dimension n , les points x_i , pour $i \in J$, sont affinement liés, c'est-à-dire qu'il existe des coefficients μ_i , pour $i \in J$, non tous nuls tels que

$$\sum_{i \in J} \mu_i x_i = 0 \quad \text{et} \quad \sum_{i \in J} \mu_i = 0.$$

On note α la plus petite valeur prise par λ_i / μ_i , pour $i \in J$ tel que $\mu_i > 0$. De tels indices existent puisque les μ_i sont non tous nuls et de somme nulle. On a $\alpha \geq 0$, et donc $\lambda_i - \alpha \mu_i \geq 0$ pour tout $i \in J$. On peut alors écrire

$$x = \sum_{i \in J} (\lambda_i - \alpha \mu_i) x_i \quad \text{avec} \quad \sum_{i \in J} (\lambda_i - \alpha \mu_i) = 1 \text{ et } \lambda_i - \alpha \mu_i \geq 0.$$

Cependant, l'un des coefficients $\lambda_i - \alpha \mu_i$ est nul, par définition de α . On peut donc retirer l'indice i correspondant de l'ensemble J et garder la même décomposition. Cela contredit donc la minimalité de q . Finalement, $q \leq n + 1$.

3. On note Λ l'ensemble des $(n+1)$ -uplets de réels $\lambda_i \geq 0$ tels que $\lambda_1 + \dots + \lambda_{n+1} = 1$. Il s'agit d'un sous-ensemble fermé du compact $[0, 1]^{n+1}$, donc c'est un compact. De plus, l'enveloppe convexe de A est l'ensemble des barycentres de points de A affectés de coefficients positifs. D'après le théorème de Carathéodory, on peut l'obtenir comme l'image par l'application continue $(\lambda_1, \dots, \lambda_{n+1}, x_1, \dots, x_{n+1}) \mapsto \lambda_1 x_1 + \dots + \lambda_{n+1} x_{n+1}$ du compact $\Lambda \times A^{n+1}$. Il s'agit donc d'un ensemble compact.
4. Choisissons une suite $(u_n)_{n \geq 1}$ de vecteurs unitaires linéairement indépendants dans E , et posons

$$x_n = 2^{-n} u_n - 2^{-n-2} u_{n+1}, \quad A = \{0\} \cup \{x_n : n \geq 1\}.$$

Comme $\|x_n\| \leq 2^{-n} + 2^{-n-2}$, la suite (x_n) converge vers 0 et A est compact. Pour tout $N \geq 1$, le point

$$s_N = 2^{-N} \cdot 0 + \sum_{n=1}^N 2^{-n} x_n = \frac{1}{4} u_1 - 4^{-N-1} u_{N+1}$$

appartient à l'enveloppe convexe de A et converge dans E vers $u_1/4$. Cependant, $u_1/4$ n'appartient même pas à l'espace vectoriel engendré par les x_n . En effet, dans une combinaison linéaire finie non nulle $\sum_{n=1}^N a_n x_n$ avec $a_N \neq 0$, le coefficient de u_{N+1} vaut $-2^{-N-2} a_N \neq 0$, ce qui exclut l'égalité avec $u_1/4$ par indépendance des u_n . L'enveloppe convexe de A n'est donc pas fermée.

Exercice 17.2 (Critères de convexité pour les fonctions)

Leçon : 253. Référence : [6].

- Soit f une application d'un intervalle I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$. Montrer que f est convexe si et seulement si pour tout segment $[a, b] \subset I$ et tout réel μ , l'application définie par $\varphi(x) = f(x) + \mu x$ pour tout $x \in [a, b]$ est majorée sur $[a, b]$ et atteint sa borne supérieure en a ou en b .
- Soit f une application continue d'un intervalle I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$. Montrer que f est convexe si et seulement si pour tous $x \in I$ et $h > 0$ tels que $x - h$ et $x + h$ appartiennent à I , on a $2hf(x) \leq \int_{x-h}^{x+h} f(t) dt$.
- Soit f une application continue d'un intervalle I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$.
 - Montrer que f est convexe si et seulement si

$$\forall x \in \overset{\circ}{I} \quad \forall \varepsilon > 0 \quad \exists h \in]0, \varepsilon[, \quad [x - h, x + h] \subset I \quad f(x) \leq \frac{1}{2} (f(x + h) + f(x - h)).$$

- En déduire que f est convexe si pour tout $x \in \overset{\circ}{I}$,

$$\limsup_{h \rightarrow 0} \frac{1}{h^2} (f(x + h) + f(x - h) - 2f(x)) \geq 0.$$

- Conclure que f est affine si pour tout $x \in \overset{\circ}{I}$,

$$\lim_{h \rightarrow 0} \frac{1}{h^2} (f(x + h) + f(x - h) - 2f(x)) = 0.$$

Corrigé.

- On suppose que f est convexe. Pour tout segment $[a, b] \subset I$ et tout réel μ , la fonction $\varphi(x) = f(x) + \mu x$ est clairement convexe aussi. De plus, pour tout $\lambda \in [0, 1]$, on a par convexité $\varphi(\lambda a + (1 - \lambda)b) \leq \lambda \varphi(a) + (1 - \lambda) \varphi(b) \leq \max\{\varphi(a), \varphi(b)\}$. La fonction φ est donc majorée sur $[a, b]$ et atteint sa borne supérieure en a ou en b . Réciproquement, considérons un intervalle $[a, b]$ inclus dans I et fixons μ de sorte que l'application définie par $\varphi(x) = f(x) + \mu x$ vérifie la condition $\varphi(a) = \varphi(b)$. Cette application est majorée sur $[a, b]$ et atteint sa borne supérieure en a et en b , donc $\varphi(\lambda a + (1 - \lambda)b) \leq \varphi(a) = \varphi(b)$ pour tout $\lambda \in [0, 1]$. En revenant à la fonction f ,

$$\begin{aligned} f(\lambda a + (1 - \lambda)b) &= \varphi(\lambda a + (1 - \lambda)b) - \mu(\lambda a + (1 - \lambda)b) \\ &\leq \lambda \varphi(a) + (1 - \lambda) \varphi(b) - \lambda \mu a - (1 - \lambda) \mu b = \lambda f(a) + (1 - \lambda) f(b). \end{aligned}$$

Ce résultat étant vrai pour tout segment $[a, b]$ inclus dans I , on en déduit que la fonction f est convexe sur I .

2. On suppose que f est convexe, et on considère $x \in I$ et $h > 0$ tels que $x - h$ et $x + h$ appartiennent à I . Par convexité, on a pour tout $t \in [0, h]$, l'inégalité

$$f(x) = f\left(\frac{(x-t) + (x+t)}{2}\right) \leq \frac{1}{2}(f(x-t) + f(x+t)).$$

En intégrant cette inégalité pour $t \in [0, h]$, on obtient

$$2hf(x) \leq 2 \int_0^h f(x) dt \leq \int_0^h (f(x-t) + f(x+t)) dt = \int_{x-h}^{x+h} f(t) dt.$$

Supposons réciproquement que l'inégalité est vérifiée. Si la fonction f n'est pas convexe, alors il existe un segment $[a, b]$ inclus dans I et un réel μ tel que l'application définie par $\varphi(x) = f(x) + \mu x$ n'admet pas son maximum en a ou en b . Comme φ est continue, elle admet son supremum sur $[a, b]$ en un point c , qui est donc distinct de a et b . On note $h = \min\{b - c, c - a\} > 0$. On a donc $\int_{c-h}^{c+h} \varphi(t) dt < \int_{c-h}^{c+h} \varphi(c) dt = 2h\varphi(c)$, avec une inégalité stricte du fait de la continuité de φ . En revenant à la fonction f , on en déduit que $\int_{c-h}^{c+h} f(t) dt < 2hf(c)$, ce qui contredit l'hypothèse. Ainsi, f est convexe.

3. (a). Le sens direct est trivial. Pour la réciproque, on utilise le critère de la première question. Considérons un segment $[a, b]$ inclus dans I et un réel μ . La condition sur f est aussi vérifiée par la fonction définie par $\varphi(x) = f(x) + \mu x$. Il suffit de montrer que φ atteint son supremum en a ou en b . Comme φ est continue sur le segment, elle y atteint son supremum en un point c , que l'on peut supposer le plus petit possible, et que l'on suppose distinct de a et b . On note $\varepsilon > 0$ tel que $a < c - \varepsilon$ et $c + \varepsilon < b$. Par hypothèse, il existe $h \in]0, \varepsilon[$ tel que $\varphi(c) \leq \frac{1}{2}(\varphi(c+h) + \varphi(c-h))$. Par maximalité de $\varphi(c)$, on en déduit que $\varphi(c+h) = \varphi(c-h) = \varphi(c)$, ce qui contredit la minimalité de c . Ainsi, c coïncide avec a ou b . Le critère de la première question est donc vérifié, ce qui prouve que la fonction f est convexe.
- (b). Pour tout réel $\alpha > 0$, on considère la fonction définie sur I par $f_\alpha(x) = f(x) + \alpha x^2$. L'hypothèse faite sur la fonction f implique que pour tous $x \in I$ et $\varepsilon > 0$, il existe $h \in]0, \varepsilon[$ tel que

$$f_\alpha(x+h) + f_\alpha(x-h) - 2f_\alpha(x) = f(x+h) + f(x-h) - 2f(x) + 2\alpha h^2 \geq -\varepsilon h^2 + 2\alpha h^2.$$

En choisissant $\varepsilon < 2\alpha$, on en déduit que $f_\alpha(x) \leq \frac{1}{2}(f_\alpha(x+h) + f_\alpha(x-h))$. La condition de la question précédente est donc satisfaite, ce qui montre que la fonction f_α est convexe. Finalement, comme f est limite simple des fonctions f_α quand $\alpha \rightarrow 0^+$, on en conclut que f est convexe également.

- (c). La condition sur f implique que les deux fonctions f et $-f$ vérifient la condition de la question précédente. On en déduit que f est à la fois convexe et concave. Elle est donc affine.

Exercice 17.3 (Log-convexité de la fonction Gamma)

Leçon : 253. Références : [6, 16].

(sans calcul de dérivée). Soit f une fonction définie sur un intervalle I de $\mathbb{R}$ et à valeurs strictement positives. On dit que f est *logarithmiquement convexe*, ou plus simplement *log-convexe*, si la composée $\log \circ f$ est convexe.

- Montrer que si la fonction f est log-convexe, alors elle est convexe.
- Montrer que f est log-convexe si et seulement si la fonction $x \mapsto f(x)c^x$ est convexe pour tout $c > 0$.
- Montrer que toute combinaison linéaire à coefficients positifs, non tous nuls, de fonctions log-convexes est log-convexe.

4. Pour tout $h > 0$, montrer que la fonction définie pour $x > 0$ par $\Gamma_h(x) = \sum_{n=1}^{\infty} (nh)^{x-1} e^{-nh} h$ est log-convexe.

5. Pour tout $x > 0$, montrer que $\Gamma_h(x) \xrightarrow{h \rightarrow 0^+} \Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$, puis conclure que Γ est log-convexe.

6. (Bohr-Mollerup) Montrer que toute fonction log-convexe $f :]0, +\infty[\rightarrow]0, +\infty[$ vérifiant $f(1) = 1$ et $f(x+1) = x f(x)$ pour tout $x > 0$ est nécessairement la fonction Γ . On pourra montrer au passage l'identité

$$\forall x > 0 \quad \Gamma(x) = \lim_{n \rightarrow \infty} \frac{n! n^x}{x(x+1) \dots (x+n)},$$

de laquelle on déduira la formule de duplication

$$\forall x > 0 \quad \Gamma(x) \Gamma\left(x + \frac{1}{2}\right) = 2^{1-2x} \sqrt{\pi} \Gamma(2x).$$

Corrigé.

- Comme la fonction f est log-convexe sur I , et la fonction log est concave sur $]0, +\infty[$, on a pour tous $a, b \in I$ et $\lambda \in [0, 1]$, les inégalités $\log f(\lambda a + (1-\lambda)b) \leq \lambda \log f(a) + (1-\lambda) \log f(b) \leq \log(\lambda f(a) + (1-\lambda)f(b))$, et on conclut à la convexité de f par croissance du logarithme.
- Pour le sens direct, on observe que pour tout réel $c > 0$, la fonction $x \mapsto \log f(x) + x \log c$ est convexe. La fonction $x \mapsto f(x) c^x$ est donc log-convexe, puis convexe, d'après la question précédente. Pour la réciproque, on considère $a, b \in I$ et $\lambda \in [0, 1]$. Les cas $a = b$ ou $\lambda \in \{0, 1\}$ sont immédiats; on suppose donc $a \neq b$ et $0 < \lambda < 1$. Pour tout $c > 0$, on a $f(\lambda a + (1-\lambda)b) c^{\lambda a + (1-\lambda)b} \leq \lambda f(a) c^a + (1-\lambda) f(b) c^b$, ce qui donne encore

$$f(\lambda a + (1-\lambda)b) \leq \lambda f(a) c^{(1-\lambda)(a-b)} + (1-\lambda) f(b) c^{\lambda(b-a)} = \lambda f(a) \alpha^{1-\lambda} + (1-\lambda) f(b) \alpha^{-\lambda},$$

en posant $\alpha = c^{a-b}$. On fixe les paramètres a, b, λ , et on minimise la borne supérieure sur $c > 0$, ce qui revient à minimiser sur $\alpha > 0$. La fonction correspondante atteint son minimum en $\alpha = f(b)/f(a)$, et vaut $f(a)^\lambda f(b)^{1-\lambda}$ en ce point. Finalement, on obtient $f(\lambda a + (1-\lambda)b) \leq f(a)^\lambda f(b)^{1-\lambda}$, ce qui prouve la log-convexité de f .

- Considérons une combinaison linéaire $\lambda_1 f_1 + \dots + \lambda_n f_n$ à coefficients positifs, non tous nuls, de fonctions log-convexes. Alors, pour tout $c > 0$, les fonctions $x \mapsto f_i(x) c^x$ sont convexes. On en déduit directement que la fonction $x \mapsto (\lambda_1 f_1(x) + \dots + \lambda_n f_n(x)) c^x$ est convexe pour tout $c > 0$, et finalement que $\lambda_1 f_1 + \dots + \lambda_n f_n$ est log-convexe.
- Soit $h > 0$. La série qui définit $\Gamma_h(x)$ est clairement convergente pour tout $x > 0$. De plus, chaque fonction $x \mapsto (nh)^{x-1} e^{-nh} h$ est log-convexe. On déduit de la question précédente que les sommes partielles de la série sont log-convexes. Pour conclure que la somme de la série est elle-même une fonction log-convexe, il suffit d'observer que toute limite simple strictement positive de fonctions log-convexes est log-convexe.
- Fixons $x > 0$ et posons $\varphi(t) = t^{x-1} e^{-t}$, de sorte que $\Gamma_h(x) = \sum_{n=1}^{\infty} \varphi(nh) h$ et $\Gamma(x) = \int_0^{\infty} \varphi(t) dt$. Pour $0 < x \leq 1$, la fonction φ est décroissante, donc par comparaison série intégrale, on a

$$\int_h^{\infty} \varphi(t) dt \leq \Gamma_h(x) \leq \Gamma(x),$$

ce qui montre directement que $\Gamma_h(x)$ tend vers $\Gamma(x)$ quand $h \rightarrow 0^+$. Supposons $x \geq 1$. On observe que

$$\Gamma(x) - \Gamma_h(x) = \sum_{n=1}^{\infty} \int_{(n-1)h}^{nh} (\varphi(t) - \varphi(nh)) dt.$$

Pour tout entier $n \geq 1$, une intégration par parties donne

$$\int_{(n-1)h}^{nh} (\varphi(t) - \varphi(nh)) dt = [(t - (n-1)h)(\varphi(t) - \varphi(nh))]_{(n-1)h}^{nh} - \int_{(n-1)h}^{nh} (t - (n-1)h) \varphi'(t) dt,$$

où le terme entre crochets est nul. Ainsi,

$$\left| \int_{(n-1)h}^{nh} (\varphi(t) - \varphi(nh)) dt \right| \leq h \int_{(n-1)h}^{nh} |\varphi'(t)| dt.$$

Comme la fonction φ' est intégrable sur l'intervalle $]0, +\infty[$, on en déduit que la série précédente converge absolument, et que

$$|\Gamma(x) - \Gamma_h(x)| \leq h \int_0^\infty |\varphi'(t)| dt \xrightarrow{h \rightarrow 0^+} 0.$$

On a donc $\Gamma_h(x) \xrightarrow{h \rightarrow 0^+} \Gamma(x)$ quel que soit $x > 0$. Finalement, la fonction Γ est log-convexe, en tant que limite simple de fonctions log-convexes.

6. On vient de prouver que la fonction Γ est log-convexe. De plus, on a clairement $\Gamma(1) = 1$ et la relation fonctionnelle $\Gamma(x+1) = x\Gamma(x)$ pour tout $x > 0$, par intégration par parties. Il suffit donc de montrer que ces conditions caractérisent une unique fonction. La fonction $\varphi = \log \circ f$ est convexe, nulle au point 1, et vérifie $\varphi(x+1) = \varphi(x) + \log x$ pour tout $x > 0$. En particulier, une récurrence élémentaire montre que $\varphi(n+1) = \log(n!)$ pour tout entier $n \in \mathbb{N}^*$. De plus, par convexité, quand $x \in]0, 1[$, les pentes sur les intervalles $[n, n+1]$, $[n+1, n+1+x]$ et $[n+1, n+2]$ sont croissantes, ce qui donne

$$\log n \leq \frac{\varphi(n+1+x) - \varphi(n+1)}{x} \leq \log(n+1).$$

En appliquant $n+1$ fois la relation fonctionnelle vérifiée par φ , on a $\varphi(n+1+x) = \varphi(x) + \log(x(x+1) \dots (x+n))$. On en déduit que

$$0 \leq \varphi(x) - \log \frac{n! n^x}{x(x+1) \dots (x+n)} \leq x \log \left(1 + \frac{1}{n}\right).$$

Ainsi, pour tout $x \in]0, 1[$, on a nécessairement $f(x) = \lim_{n \rightarrow \infty} \frac{n! n^x}{x(x+1) \dots (x+n)}$. L'équation fonctionnelle sur f implique que cette identité reste vraie quel que soit $x > 0$. Finalement, les conditions portant sur f sont vérifiées par une unique fonction, qui est la fonction Γ . On déduit en outre que $\Gamma(x) = \lim_{n \rightarrow \infty} \frac{n! n^x}{x(x+1) \dots (x+n)}$ pour tout réel $x > 0$. Cela permet de prouver la formule de duplication. En effet, on obtient le produit $\Gamma(x)\Gamma(x+1/2)$ comme limite quand $n \rightarrow +\infty$ de

$$\begin{aligned} & \frac{n! n^x}{x(x+1) \dots (x+n)} \cdot \frac{n! n^{x+1/2}}{(x+1/2)(x+3/2) \dots (x+n+1/2)} = \frac{2^{2n+2}(n!)^2 n^{2x+1/2}}{2x(2x+1) \dots (2x+2n+1)} \\ &= \frac{(2n+1)!(2n+1)^{2x}}{2x(2x+1) \dots (2x+2n+1)} \cdot \left(\frac{n}{2n+1}\right)^{2x} \cdot \frac{2^{2n+2}(n!)^2 \sqrt{n}}{(2n+1)!}. \end{aligned}$$

Le premier facteur converge vers $\Gamma(2x)$, le second vers 2^{-2x} , et le troisième vérifie d'après la formule de Stirling

$$\frac{2^{2n+2}(n!)^2 \sqrt{n}}{(2n+1)!} \underset{n \rightarrow \infty}{\sim} \frac{2^{2n+2} (\sqrt{2\pi n}(n/e)^n)^2 \sqrt{n}}{\sqrt{2\pi(2n+1)}((2n+1)/e)^{2n+1}} = \frac{2^{2n+3} \pi n^{2n+3/2} e^{-2n}}{\sqrt{2\pi(2n+1)} 2^{2n+3/2} e^{-2n-1}} = 2e\sqrt{\pi} \left(1 + \frac{1}{2n}\right)^{-2n-3/2}$$

ce qui converge finalement vers $2\sqrt{\pi}$, et conduit bien à la formule de duplication $\Gamma(x)\Gamma\left(x + \frac{1}{2}\right) = 2^{1-2x} \sqrt{\pi} \Gamma(2x)$.

Séries entières et analyse complexe

Références : [7, 12, 13, 14].

Une fonction *entière* est holomorphe sur tout $\mathbb{C}$. Une fonction *méromorphe* sur un ouvert est holomorphe sauf en des pôles isolés : au voisinage d'un pôle a , elle s'écrit $(z - a)^{-m}g(z)$ avec $m \geq 1$, g holomorphe et $g(a) \neq 0$. La *partie singulière* en ce pôle est la somme des termes de puissances négatives de son développement de Laurent.

Exercice 18.1 (Espace de Bergman)

Leçon : 201. Références : [7, 12].

Soit Ω un ouvert du plan complexe $\mathbb{C}$, et $A^2(\Omega)$ l'espace de Bergman sur Ω formé des fonctions holomorphes sur Ω qui sont aussi de carré intégrable pour la mesure de Lebesgue sur Ω , identifié à un ouvert de $\mathbb{R}^2$. Pour tous $z \in \Omega$ et $f \in A^2(\Omega)$, on pose $\delta_z(f) = f(z)$.

1. Montrer que si $r > 0$ et si le disque fermé $\bar{B}(z, r) = \{w \in \mathbb{C} \mid |w - z| \leq r\}$ est inclus dans Ω , alors pour toute fonction $f \in A^2(\Omega)$, on a $f(z) = \frac{1}{\pi r^2} \int_{\bar{B}(z, r)} f(w) dw$, où dw désigne la mesure de Lebesgue sur $\mathbb{R}^2$.
2. On désigne par $d(z, \Omega^c)$ la distance de z au complémentaire de Ω . Montrer que pour tous $f \in A^2(\Omega)$ et $z \in \Omega$,

$$0 < r < d(z, \Omega^c) \quad \implies \quad |f(z)| \leq \frac{\|f\|_2}{r\sqrt{\pi}}.$$

3. Montrer que $A^2(\Omega)$ muni du produit scalaire usuel de $L^2(\Omega)$ est un espace de Hilbert.
4. Montrer que pour tout $z \in \Omega$, l'application δ_z est une forme linéaire continue sur $A^2(\Omega)$. En déduire qu'il existe une fonction $K_z \in A^2(\Omega)$ telle que $f(z) = \int_{\Omega} f(w) \overline{K_z(w)} dw$ pour toute fonction $f \in A^2(\Omega)$.
5. On définit $K_{\Omega}(w, z) = K_z(w)$ pour tout $(w, z) \in \Omega \times \Omega$.
 - (a). Montrer que $K_{\Omega}(z, w) = \overline{K_{\Omega}(w, z)}$.
 - (b). Soit $(e_n)_{n \in \mathbb{N}}$ une base hilbertienne de $A^2(\Omega)$. Montrer que $K_z = \sum_{n=0}^{\infty} \overline{e_n(z)} e_n$ dans $A^2(\Omega)$.
6. On choisit Ω égal au disque unité ouvert $D = \{z \in \mathbb{C} \mid |z| < 1\}$.

(a). Montrer que les fonctions $e_n : z \mapsto z^n \sqrt{\frac{n+1}{\pi}}$, pour $n \in \mathbb{N}$, forment une base hilbertienne de $A^2(D)$.

(b). Calculer explicitement K_D .

Remarque : Cet exercice montre que l'espace de Bergman $A^2(\Omega)$ est un *espace de Hilbert à noyau reproduisant*, de noyau donné par l'application K_{Ω} .

Corrigé.

1. Comme le disque fermé $\bar{B}(z, r) = \{w \in \mathbb{C} \mid |w - z| \leq r\}$ est inclus dans Ω , on peut écrire le développement en série entière $f(w) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z)}{n!} (w - z)^n$ pour tout point w de ce disque, et il y a convergence normale sur tout le disque. D'après le théorème de Fubini et un passage en coordonnées polaires, on a donc

$$\begin{aligned} \int_{\bar{B}(z, r)} f(w) dw &= \sum_{n=0}^{\infty} \frac{f^{(n)}(z)}{n!} \int_{\bar{B}(z, r)} (w - z)^n dw = \sum_{n=0}^{\infty} \frac{f^{(n)}(z)}{n!} \int_0^{2\pi} \int_0^r \rho^n e^{in\theta} \rho d\rho d\theta \\ &= \sum_{n=0}^{\infty} \frac{f^{(n)}(z)}{n!} \frac{r^{n+2}}{n+2} 2\pi \mathbf{1}_{\{n=0\}} = f(z) \pi r^2, \end{aligned}$$

ce qui conduit à l'expression intégrale désirée pour $f(z)$.

2. Si $0 < r < d(z, \Omega^c)$, alors le disque $\bar{B}(z, r)$ est inclus dans Ω . On a donc, d'après la question précédente et l'inégalité de Cauchy-Schwarz dans $L^2(\Omega)$,

$$|f(z)| \pi r^2 = \left| \int_{\bar{B}(z, r)} f(w) dw \right| \leq \|f\|_2 \sqrt{\int_{\Omega} (\mathbf{1}_{\bar{B}(z, r)})^2 dw} = \|f\|_2 r \sqrt{\pi},$$

ce qui donne la borne demandée.

3. Le sous-espace vectoriel $A^2(\Omega)$ hérite par restriction du produit scalaire de $L^2(\Omega)$. Il reste à montrer la complétude de $A^2(\Omega)$ pour la norme associée. Soit $(f_n)_{n \in \mathbb{N}}$ une suite de Cauchy dans $A^2(\Omega)$, et soit K un compact inclus dans Ω . Comme la distance entre K et Ω^c est strictement positive, il existe $r > 0$ tel que $0 < r < d(z, \Omega^c)$ pour tout $z \in K$. Pour tous $p, q \in \mathbb{N}$, on a donc $f_p - f_q \in A^2(\Omega)$, donc la question précédente montre que

$$\forall z \in K \quad |f_p(z) - f_q(z)| \leq \frac{\|f_p - f_q\|_2}{r \sqrt{\pi}}.$$

Il en ressort que la suite $(f_n)_{n \in \mathbb{N}}$ est de Cauchy au sens de la norme $\|\cdot\|_{\infty}$ sur K . Elle converge donc uniformément sur tout compact inclus dans Ω vers une fonction notée f . Le théorème de Weierstrass implique que f est holomorphe sur Ω . De plus, la fonction f est de carré intégrable donc appartient à $A^2(\Omega)$. En effet, d'après le lemme de Fatou, on a pour n'importe quel $\varepsilon > 0$ et pour $p \in \mathbb{N}$ assez grand,

$$\int_{\Omega} |f_p(w) - f(w)|^2 dw \leq \liminf_{q \rightarrow \infty} \int_{\Omega} |f_p(w) - f_q(w)|^2 dw \leq \varepsilon^2,$$

de sorte que $f_p - f$ est de carré intégrable, et donc f également. Il en résulte en outre que $\|f_p - f\|_2$ converge vers zéro, ce qui montre que la suite $(f_n)_{n \in \mathbb{N}}$ converge vers f aussi dans $A^2(\Omega)$.

4. Pour $z \in \Omega$, il est clair que l'application $\delta_z : A^2(\Omega) \rightarrow \mathbb{C}$ est linéaire. De plus, en notant $r > 0$ tel que $0 < r < d(z, \Omega^c)$, on a $|\delta_z(f)| \leq |f(z)| \leq \frac{\|f\|_2}{r \sqrt{\pi}}$. Ainsi, δ_z est une forme linéaire continue de norme majorée par $\frac{1}{d(z, \Omega^c) \sqrt{\pi}}$. À l'aide du théorème de Riesz, on en déduit que δ_z est représentée via le produit scalaire par une fonction de $A^2(\Omega)$ notée K_z . On a ainsi $f(z) = \int_{\Omega} f(w) \overline{K_z(w)} dw$ pour toute fonction $f \in A^2(\Omega)$.

5. (a). Par symétrie hermitienne du produit scalaire, on a

$$K_{\Omega}(z, w) = K_w(z) = \langle K_w, K_z \rangle = \overline{\langle K_z, K_w \rangle} = \overline{K_z(w)} = \overline{K_{\Omega}(w, z)}.$$

- (b). Pour $z \in \Omega$ fixé, on sait que K_z se développe dans $A^2(\Omega)$ sous la forme $K_z = \sum_{n=0}^{\infty} \langle K_z, e_n \rangle e_n$. Cependant, chaque coefficient s'exprime comme $\langle K_z, e_n \rangle = \overline{\langle e_n, K_z \rangle} = \overline{e_n(z)}$. D'où $K_z = \sum_{n=0}^{\infty} \overline{e_n(z)} e_n$.

6. (a). On vérifie que les fonctions $e_n : z \mapsto a_n z^n$, avec $a_n = \sqrt{\frac{n+1}{\pi}}$, forment un système orthonormé de $A^2(D)$, où D est le disque unité ouvert. En effet, pour $p, q \in \mathbb{N}$, on calcule en passant en coordonnées polaires

$$\langle e_p, e_q \rangle = \int_D a_p a_q w^p \overline{w}^q dw = a_p a_q \int_0^{2\pi} \int_0^1 \rho^{p+q} e^{i(p-q)\theta} \rho d\rho d\theta = 2\pi a_p a_q \frac{\mathbf{1}_{\{p=q\}}}{p+q+2} = \mathbf{1}_{\{p=q\}}.$$

Pour montrer que la famille $(e_n)_{n \in \mathbb{N}}$ est totale, il suffit de montrer que pour toute fonction f de $A^2(D)$, on a $\|f\|_2^2 = \sum_{n=0}^{\infty} |\langle f, e_n \rangle|^2$. En effet, cette identité caractérise l'appartenance à l'adhérence de l'espace vectoriel engendré par les e_n . En écrivant f comme somme de la série entière $\sum f_n z^n$, en passant en coordonnées polaires et en utilisant la formule de Parseval sur chaque cercle de rayon ρ centré à l'origine puis enfin le théorème de Fubini-Tonelli, on obtient

$$\begin{aligned} \|f\|_2^2 &= \int_D |f(w)|^2 dw = \int_0^1 \left(\int_0^{2\pi} |f(\rho e^{i\theta})|^2 d\theta \right) \rho d\rho = \int_0^1 \left(2\pi \sum_{n=0}^{\infty} |f_n|^2 \rho^{2n} \right) \rho d\rho \\ &= 2\pi \sum_{n=0}^{\infty} |f_n|^2 \int_0^1 \rho^{2n+1} d\rho = \pi \sum_{n=0}^{\infty} \frac{|f_n|^2}{n+1}. \end{aligned}$$

Par polarisation, cela implique que pour $f = \sum_n f_n z^n$ et $g = \sum_n g_n z^n$ dans $A^2(D)$, on a $\langle f, g \rangle = \pi \sum_{n=0}^{\infty} \frac{f_n \overline{g_n}}{n+1}$.

En particulier, pour tout $n \in \mathbb{N}$, on obtient $\langle f, e_n \rangle = \pi \frac{f_n a_n}{n+1} = f_n \sqrt{\frac{\pi}{n+1}}$. L'identité recherchée est donc vérifiée pour toute fonction f de $A^2(D)$, ce qui montre que $(e_n)_{n \in \mathbb{N}}$ en est bien une base hilbertienne.

(b). On calcule $K_D(w, z) = \langle K_z, K_w \rangle = \sum_{n=0}^{\infty} \overline{e_n(z)} e_n(w) = \sum_{n=0}^{\infty} \frac{n+1}{\pi} (\overline{z}w)^n = \frac{1}{\pi(1-\overline{z}w)^2}$.

Exercice 18.2 (Composantes connexes dans $\mathbb{C}$)

Leçon : 204.

Soit K un compact non vide de $\mathbb{C}$, et $\Omega = K^c$ son complémentaire.

1. Montrer que Ω admet une unique composante connexe non bornée notée ω_∞ .
2. Montrer que toute composante connexe ω de Ω est ouverte et que sa frontière vérifie $\partial\omega \subset K$.
3. Montrer que Ω admet un nombre au plus dénombrable de composantes connexes.
4. Montrer que la réunion $\widehat{K}$ de K et des composantes connexes bornées de Ω est un compact dont le complémentaire est connexe.

Corrigé.

1. Le compact K est inclus dans le disque fermé $\overline{D}(0, R)$ pour R assez grand. L'ensemble $\omega_R = \{z \in \mathbb{C} \mid |z| > R\}$ est un ouvert inclus dans Ω . De plus, ω_R est connexe, puisque c'est l'image du produit de connexes $]R, +\infty[\times [0, 2\pi[$ par l'application continue $(r, \theta) \mapsto re^{i\theta}$. Comme le point $R+1$ appartient à ω_R , la composante connexe qui contient $R+1$ contient nécessairement ω_R , donc est une composante connexe non bornée de Ω , notée ω_∞ . De plus, toute autre composante connexe ω de Ω est nécessairement incluse dans $\overline{D}(0, R)$, donc est bornée. En effet, dans le cas contraire, il existerait $z \in \omega \cap \omega_R \subset \omega \cap \omega_\infty$, ce qui est impossible puisque les composantes connexes sont disjointes.
2. Toute composante connexe ω de Ω est ouverte, puisque $\mathbb{C}$ est localement connexe. Ainsi, $\partial\omega = \overline{\omega} \setminus \omega$. De plus, ω est un fermé de Ω , de sorte que $\overline{\omega} \cap \Omega = \omega$. Il en ressort que $\partial\omega \cap \Omega = \emptyset$, et donc $\partial\omega \subset K$.
3. Pour tout entier p , on note $\mathcal{C}$ l'ensemble des composantes connexes de Ω . Pour tout $\omega \in \mathcal{C}$, l'intersection $\omega \cap D(0, p)$ est un ouvert. S'il est non vide, il contient un disque $D(x_\omega, r_\omega)$. Ces disques sont disjoints et tous inclus dans $D(0, p)$. En prenant la mesure de Lebesgue, on a en particulier

$$\sum_{\substack{\omega \in \mathcal{C} \\ \omega \cap D(0, p) \neq \emptyset}} \pi r_\omega^2 \leq \pi p^2,$$

ce qui montre que l'ensemble des composantes connexes de Ω qui rencontrent $D(0, p)$ est au plus dénombrable. On en conclut que Ω comporte un nombre au plus dénombrable de composantes connexes.

4. Notons Ω_0 la réunion (éventuellement vide) des composantes connexes bornées de Ω . On a $\Omega = \Omega_0 \sqcup \omega_\infty$, donc $\mathbb{C} = K \sqcup \Omega = K \sqcup \Omega_0 \sqcup \omega_\infty = \widehat{K} \sqcup \omega_\infty$, ce qui montre que $\widehat{K} = \omega_\infty^c$ est fermé. De plus, $\widehat{K} = K \cup \Omega_0 \subset \overline{D}(0, R)$, donc $\widehat{K}$ est borné. Finalement, $\widehat{K}$ est compact, et on sait que son complémentaire $\widehat{K}^c = \omega_\infty$ est connexe.

Exercice 18.3 (Théorème de Runge)

Leçon : 204. Référence : [14].

- Soit K un compact non vide de $\mathbb{C}$. On note $C(K)$ l'algèbre des fonctions continues sur K , muni de la norme $\|\cdot\|_\infty$, ainsi que $P(K)$ l'adhérence de l'ensemble des fonctions polynomiales définies sur K . Montrer que $P(K)$ est une sous-algèbre de $C(K)$.
- On suppose dans cette question que le complémentaire K^c est connexe. Pour tout $a \in K^c$, on définit la fonction $\varphi_a \in C(K)$ par $\varphi_a(z) = \frac{1}{z-a}$ pour tout $z \in K$. On note A l'ensemble des points $a \in K^c$ tels que $\varphi_a \in P(K)$.
 - Montrer que A est non vide.
 - Montrer que A est fermé dans K^c .
 - Montrer que A est ouvert dans K^c .
 - En déduire que pour tout point $a \in K^c$, la fonction φ_a est limite uniforme de fonctions polynomiales sur K .
- Montrer que si B désigne une sous-algèbre fermée de $C(K)$, et ω une composante connexe de K^c , alors l'ensemble $\omega_B = \{a \in \omega \mid \varphi_a \in B\}$ est soit vide, soit égal à ω tout entier.
- On note ω_n les composantes connexes bornées de K^c , et on choisit un point $a_n \in \omega_n$ pour chaque n . Montrer que pour tout point $a \in K^c$, la fonction φ_a est limite uniforme sur K de fractions rationnelles dont les pôles sont contenus dans l'ensemble des a_n .

Corrigé.

- Si P_n et Q_n désignent des fonctions polynomiales définies sur K , qui convergent uniformément vers f et g respectivement, et si $\lambda \in \mathbb{C}$, on a clairement $\lambda P_n + Q_n \xrightarrow{n \rightarrow \infty} \lambda f + g$ et $P_n Q_n \xrightarrow{n \rightarrow \infty} fg$. Ainsi, $P(K)$ est une sous-algèbre de $C(K)$.
- (a). Notons $R = \sup\{|z|, z \in K\}$ et considérons un point tel que $|a| > R$. On a alors

$$\varphi_a(z) = -\frac{1}{a} \frac{1}{1 - z/a} = -\sum_{n=0}^{\infty} \frac{z^n}{a^{n+1}}.$$

Cette série entière converge normalement sur K , puisque $\left| \frac{z^n}{a^{n+1}} \right| \leq \frac{1}{|a|} \left(\frac{R}{|a|} \right)^n$. La fonction φ_a est donc limite uniforme sur K de la suite de fonctions polynomiales $-\sum_{k=0}^n \frac{z^k}{a^{k+1}}$. Ainsi $\varphi_a \in P(K)$, puis $a \in A$.

- Pour tous points $a, b \in K^c$ et tout point $z \in K$ tels que $|a - b| < \delta = d(a, K)$, on a

$$|\varphi_a(z) - \varphi_b(z)| = \left| \frac{1}{z-a} - \frac{1}{z-b} \right| = \frac{|a-b|}{|z-a||z-b|} \leq \frac{|a-b|}{\delta(\delta - |a-b|)}.$$

Considérons une suite de points $a_n \in A$ qui converge vers $a \in K^c$. Pour n assez grand, on a $|a - a_n| \leq \delta/2 < \delta$, de sorte que $\|\varphi_{a_n} - \varphi_a\| \leq \frac{2}{\delta^2} |a - a_n| \xrightarrow{n \rightarrow \infty} 0$. Comme chaque fonction φ_{a_n} appartient à $P(K)$, on en déduit que la fonction φ_a appartient à l'adhérence de $P(K)$, c'est-à-dire à $P(K)$ puisque c'est un fermé.

- Pour $a \in A$ et $\delta = d(a, K)$, montrons que le disque fermé $\overline{D}(a, \delta/2)$ est contenu dans A . Pour $|h| \leq \delta/2$ et $z \in K$, on a

$$\varphi_{a+h}(z) = \frac{1}{z-a-h} = \frac{1}{z-a} \frac{1}{1 - h/(z-a)} = \sum_{n=0}^{\infty} \frac{h^n}{(z-a)^{n+1}} = \sum_{n=0}^{\infty} h^n (\varphi_a(z))^{n+1}.$$

Cette série converge normalement sur K car le terme général vérifie $\left| \frac{h^n}{(z-a)^{n+1}} \right| \leq \frac{|h|^n}{\delta^{n+1}} \leq \delta^{-1} 2^{-n}$. On en déduit que φ_{a+h} est limite uniforme sur K de la suite de fonctions $\sum_{k=0}^n h^k (\varphi_a(z))^{k+1}$, et donc appartient à la sous-algèbre fermée $P(K)$. Ainsi, le point $a + h$ appartient à A dès que $|h| \leq \delta/2$.

- (d). On a montré que A est une partie non vide, fermée et ouverte du connexe K^c . Il en ressort que $A = K^c$. Par conséquent, pour tout point $a \in K^c$, la fonction φ_a appartient à $P(K)$.
3. En reproduisant les preuves ci-dessus, on montre que ω_B est une partie ouverte et fermée dans le connexe ω . Ainsi, ω_B est vide ou égal à ω .
4. On note B l'adhérence dans $C(K)$ des restrictions à K des fractions rationnelles dont les pôles sont contenus dans l'ensemble des a_n . Il s'agit d'une sous-algèbre fermée de $C(K)$. D'après ce qui précède, pour toute composante connexe ω de K^c , l'ensemble ω_B des $a \in \omega$ tels que $\varphi_a \in B$ est vide ou égal à ω . On a vu ci-dessus que si a est de module assez grand, alors φ_a est limite uniforme de fonctions polynomiales sur K . Ainsi, si ω est la composante connexe non bornée, ω_B est non vide. Si au contraire on considère une composante connexe bornée ω_n , alors $\varphi_{a_n} \in B$, et donc $\omega_{n,B}$ est non vide aussi. Finalement,

$$\{a \in K^c \mid \varphi_a \in B\} = \bigsqcup_{\omega} \omega_B = \bigsqcup_{\omega} \omega = K^c.$$

Exercice 18.4 (Enveloppe polynomiale d'un compact de $\mathbb{C}$)

Leçon : 204. Référence : [14].

Soit K un compact non vide de $\mathbb{C}$, et $\hat{K}$ la réunion de K et des composantes connexes bornées du complémentaire K^c .

- Montrer que pour toute fonction polynomiale P et pour tout point $a \in \hat{K}$, on a $|P(a)| \leq \sup_K |P|$.
- Soit $a \in \hat{K}^c$. Montrer qu'il existe une suite de fonctions polynomiales P_n qui converge uniformément sur $\hat{K}$ vers $\varphi_a : z \mapsto \frac{1}{z-a}$. En déduire qu'il existe une suite de fonctions polynomiales Q_n telles que $Q_n(a) = 1$ quel que soit n , et $\sup_K |Q_n| \xrightarrow{n \rightarrow \infty} 0$.
- Montrer que $\hat{K}$ est l'ensemble des points $a \in \mathbb{C}$ tels que $|P(a)| \leq \sup_K |P|$ pour toute fonction polynomiale P . On dit que $\hat{K}$ est l'*enveloppe polynomiale* de K .

Corrigé.

- Soit ω une composante connexe bornée de K^c , et soit $a \in \omega$. On a $|P(a)| \leq \sup_{\partial\omega} |P|$ d'après le principe du maximum. De plus, on sait d'après un exercice précédent que $\partial\omega \subset K$. Ainsi, $|P(a)| \leq \sup_K |P|$, et cette borne est valable pour a dans n'importe quelle composante connexe bornée de K^c , et aussi pour a dans K . La borne est donc valable pour tout a dans $\hat{K}$.
- Comme $\hat{K}$ est un compact de complémentaire connexe, le théorème de Runge assure directement que pour tout $a \in \hat{K}^c$, il existe une suite de fonctions polynomiales P_n qui converge uniformément sur $\hat{K}$ vers φ_a . On pose ensuite $Q_n(z) = 1 - (z-a)P_n(z)$. Il est clair que $Q_n(a) = 1$ quel que soit n , et $\sup_K |Q_n| \xrightarrow{n \rightarrow \infty} 0$.
- On a vu que $|P(a)| \leq \sup_K |P|$ pour tout $a \in \hat{K}$. Pour la réciproquement, supposons que $a \notin \hat{K}$ et considérons les fonctions polynomiales Q_n obtenues ci-dessus. Pour n assez grand, on a $|Q_n(a)| = 1 > \sup_K |Q_n|$. Ainsi, il existe une fonction polynomiale P telle que $|P(a)| > \sup_K |P|$.

Exercice 18.5 (Inégalités de Cauchy renforcées)

Leçons : 243, 245. Référence : [13].

Soit $f(z) = \sum_{n=0}^{\infty} a_n z^n$ une fonction holomorphe dans le disque unité ouvert D . On suppose que $|f(z)| \leq (1-|z|)^{-1}$ pour tout $z \in D$. Montrer que

$$|a_0| \leq 1, \quad \forall n \geq 1, \quad |a_n| \leq \left(1 + \frac{1}{n}\right)^n (n+1) \leq e(n+1).$$

Corrigé.

On a $|a_0| = |f(0)| \leq 1$. Fixons maintenant $n \geq 1$ et $r \in]0, 1[$. D'après les formules de Cauchy, on a

$$|a_n| = \frac{1}{2\pi r^n} \left| \int_0^{2\pi} f(re^{it}) e^{-int} dt \right| \leq \frac{1}{2\pi r^n} \int_0^{2\pi} \frac{dt}{1-r} = \frac{1}{r^n(1-r)}.$$

La fonction $g(r) = r^n(1-r)$ admet pour dérivée $g'(r) = r^{n-1}(n - (n+1)r)$, qui s'annule pour $r = \frac{n}{n+1}$. On voit alors que g admet son maximum en cette valeur. La borne ci-dessus est donc minimale pour cette même valeur de r , et on obtient

$$|a_n| \leq \left(1 + \frac{1}{n}\right)^n (n+1).$$

Pour établir la seconde majoration, on utilise la majoration classique $\ln(1+u) \leq u$ pour écrire que

$$\left(1 + \frac{1}{n}\right)^n = \exp\left(n \ln\left(1 + \frac{1}{n}\right)\right) \leq \exp\left(n \frac{1}{n}\right) \leq e.$$

Exercice 18.6 (Conditions nécessaires de convergence uniforme)

Leçon : 243. Référence : [13].

On considère une série entière $S(z) = \sum_{n=0}^{\infty} a_n z^n$ de rayon de convergence $0 < R < \infty$, dont on note $S_n(z)$ les sommes partielles. On note de plus $D(0, R)$, $\overline{D}(0, R)$ et $\partial D(0, R)$ respectivement le disque ouvert, le disque fermé, et le cercle de centre 0 et de rayon R .

1. Montrer que pour tous entiers $0 \leq p < q$, la fonction $|S_q(z) - S_p(z)|$ admet le même supremum sur les ensembles $D(0, R)$, $\overline{D}(0, R)$ et $\partial D(0, R)$.
2. En déduire l'équivalence de la convergence uniforme de la série entière S sur ces trois ensembles, c'est-à-dire que S converge uniformément sur $D(0, R)$ si et seulement si elle converge uniformément sur $\overline{D}(0, R)$, si et seulement si elle converge uniformément sur $\partial D(0, R)$.
3. Montrer que s'il y a convergence uniforme sur l'un de ces ensembles, alors la série $\sum_{n=0}^{\infty} |a_n|^2 R^{2n}$ converge.

Corrigé.

1. Pour tous entiers $0 \leq p < q$, la fonction $S_q(z) - S_p(z)$ est une fonction polynomiale, donc son module vérifie le principe du maximum. Par conséquent,

$$\sup_{z \in D(0, R)} |S_q(z) - S_p(z)| = \sup_{z \in \overline{D}(0, R)} |S_q(z) - S_p(z)| = \sup_{z \in \partial D(0, R)} |S_q(z) - S_p(z)|.$$

2. L'équivalence de la convergence uniforme sur les trois ensembles $D(0, R)$, $\overline{D}(0, R)$ et $\partial D(0, R)$ découle du résultat de la question précédente et du critère de Cauchy uniforme.

3. La fonction définie par $g(\theta) = \sum_{n=0}^{\infty} a_n R^n e^{in\theta}$ est continue et 2π -périodique sur $\mathbb{R}$, comme limite uniforme de ses

sommes partielles, qui vérifient clairement ces propriétés. De plus, on calcule $\int_0^{2\pi} g(\theta) e^{-in\theta} \frac{d\theta}{2\pi} = a_n R^n$ pour

tout entier $n \geq 0$. L'inégalité de Bessel permet de conclure que $\sum_{n=0}^{\infty} |a_n|^2 R^{2n} \leq \int_0^{2\pi} |g(\theta)|^2 \frac{d\theta}{2\pi} < \infty$. (Remarque :

Il y a égalité, d'après l'identité de Parseval, car les coefficients de Fourier de g d'indices $n < 0$ sont nuls.)

Exercice 18.7 (Comportement au bord du disque de convergence)

Leçon : 243. Référence : [13].

On note D le disque unité ouvert, et $\Gamma = \partial D$ son bord, c'est-à-dire le cercle unité. On considère une série entière $f(z) = \sum_{n=0}^{\infty} a_n z^n$ de rayon de convergence 1. On dit qu'un point $a \in \Gamma$ est *régulier* si f admet un prolongement analytique dans un ensemble de la forme $D \cup D_a$, où D_a est un disque ouvert centré en a . Dans le cas contraire,

a est dit *singulier*. On note respectivement A_r et A_s l'ensemble des points réguliers et l'ensemble des points singuliers de Γ pour la série entière f . Enfin, on dit que Γ est une *coupure* pour la série entière f si tous les points de Γ sont singuliers.

1. Montrer que A_r est un ouvert de Γ . (Ainsi, A_s est un fermé de Γ .)
2. Déterminer les ensembles A_r et A_s pour les séries entières $\sum_{n=0}^{\infty} z^n$ et $\sum_{n=1}^{\infty} \frac{z^n}{n^2}$.
3. Montrer que $A_s \neq \emptyset$. (*Indication* : Montrer que si $A_s = \emptyset$, alors on peut prolonger f en une série entière de rayon de convergence > 1 .)
4. Montrer qu'un point $a \in \Gamma$ est régulier si et seulement si on peut trouver un voisinage V de a dans Γ et des réels $\rho_a > 1/2$ et $C_a > 0$ tels que

$$\forall b \in V \quad \forall k \in \mathbb{N} \quad \frac{|f^{(k)}(b/2)|}{k!} \rho_a^k \leq C_a.$$

5. Montrer qu'un point $a \in \Gamma$ est régulier si et seulement si la série entière $\sum_{k=0}^{\infty} \frac{f^{(k)}(a/2)}{k!} w^k$ admet un rayon de convergence $> 1/2$. Étudier le cas particulier où $f(z) = \sum_{n=0}^{\infty} z^n$.
6. On suppose que $a_n \geq 0$ pour tout entier $n \geq 0$. Montrer que le point 1 est singulier pour la série entière f .
7. On considère une suite d'entiers $\lambda_n \geq 1$ tels que $\lambda_{n+1}/\lambda_n \geq \alpha > 1$ pour tout entier $n \geq 1$, puis une série entière de la forme $\sum_{n=1}^{\infty} a_n z^{\lambda_n}$ de rayon de convergence égal à 1. Montrer que Γ est une coupure pour cette série entière. (C'est le théorème des lacunes de Hadamard.)

Corrigé.

1. Soit $a \in A_r$ un point régulier. Par définition, la série entière f admet un prolongement analytique dans l'ensemble $D \cup D(a, r_0)$, où $D(a, r_0)$ désigne le disque ouvert de centre a et de rayon $r_0 > 0$. On note b un point de Γ tel que $|b - a| < r_0/2$. Alors, f admet un prolongement analytique sur l'ensemble $D \cup D(b, r_0/2)$, puisqu'il est contenu dans $D \cup D(a, r_0)$. On en déduit que b est un point régulier. Finalement, l'ensemble A_r est bien ouvert.
2. Pour la série entière $\sum_{n=0}^{\infty} z^n$, on a clairement $A_r = \Gamma \setminus \{1\}$ et $A_s = \{1\}$. Pour la série entière $F(z) = \sum_{n=1}^{\infty} z^n/n^2$, on a également $A_r = \Gamma \setminus \{1\}$ et $A_s = \{1\}$. En effet, pour $0 < |z| < 1$,

$$F'(z) = \sum_{n=1}^{\infty} \frac{z^{n-1}}{n} = -\frac{\log(1-z)}{z}.$$

La fonction du membre de droite, prolongée par la valeur 1 en 0, est holomorphe sur le domaine simplement connexe $\mathbb{C} \setminus [1, +\infty[$. Sa primitive nulle en 0 prolonge donc F à ce domaine. En revanche, $F'(x) \rightarrow +\infty$ lorsque $x \rightarrow 1^-$, ce qui exclut un prolongement holomorphe au voisinage de 1. La convergence uniforme de la série sur $\overline{D}$ assure la continuité au bord, mais pas la régularité au sens du prolongement analytique.

3. Supposons $A_s = \emptyset$. Alors, pour tout point $a \in \Gamma$, la fonction f admet un prolongement analytique noté f_a à un ensemble de la forme $D \cup D_a$, avec $D_a = D(a, r_a)$ le disque ouvert de centre a et de rayon r_a . On observe que si $D_a \cap D_b$ est non vide, alors $f_a = f_b$ sur $D_a \cap D_b$. Le cas $a = b$ est immédiat. Si $a \neq b$, le point $\frac{r_b a + r_a b}{r_a + r_b}$ appartient à $D_a \cap D_b \cap D$, car $|a - b| < r_a + r_b$ et une combinaison convexe stricte de deux points distincts du cercle unité appartient à D . La fonction f coïncide avec f_a sur $D \cap D_a$, et avec f_b sur $D \cap D_b$, donc on a $f_a = f = f_b$ sur $D_a \cap D_b \cap D$, puis par unicité du prolongement analytique $f_a = f_b$ sur $D_a \cap D_b$. On considère alors l'ouvert $\Omega = \bigcup_{a \in \Gamma} (D \cup D_a)$ et la fonction vérifiant $F(z) = f_a(z)$ pour $z \in D \cup D_a$. La fonction F est correctement définie et holomorphe sur Ω .

On peut supposer les r_a uniformément bornés. Alors, $\mathbb{C} \setminus \Omega$ est un fermé non vide disjoint du disque fermé $\overline{D}$. On note $\delta > 0$ la distance entre $\overline{D}$ et $\mathbb{C} \setminus \Omega$. On a alors $D(0, 1 + \delta) \subset \Omega$. La fonction f étant holomorphe sur

$D(0, 1 + \delta)$, elle admet un développement en série entière dans ce disque. Cependant, comme f coïncide avec $\sum_{n=0}^{\infty} a_n z^n$ dans D , on a par unicité du développement en série entière que la série $\sum_{n=0}^{\infty} a_n z^n$ admet au moins $1 + \delta$ comme rayon de convergence. Cela contredit le fait que son rayon de convergence vaut 1.

4. Supposons l'existence d'un voisinage V de a dans Γ et de réels $\rho_a > 1/2$ et $C_a > 0$ tels que

$$\forall b \in V \quad \forall k \in \mathbb{N} \quad \frac{|f^{(k)}(b/2)|}{k!} \rho_a^k \leq C_a.$$

On considère alors sur l'ouvert $\Omega = D \cup D(a/2, \rho_a)$ la fonction holomorphe F définie par $F(z) = f(z)$ si $z \in D$, et par $F(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a/2)}{k!} (z - a/2)^k$ si $z \in D(a/2, \rho_a)$. La formule de Taylor pour le développement en série entière de f au voisinage de tout point de son disque ouvert de convergence montre que la fonction F est correctement définie. La fonction F conduit donc à un prolongement analytique de f à un ensemble de la forme $D \cup D_a$, ce qui montre que a est régulier.

Réciproquement, supposons que a est régulier et considérons un réel $\delta > 0$ tel que f admette un prolongement analytique g sur l'ensemble $D \cup D(a, 2\delta) = \Omega$. On observe que les compacts $K_\varepsilon = \bigcup_{|\theta| \leq \varepsilon} \overline{D}\left(\frac{a}{2} e^{i\theta}, \frac{1}{2} + \varepsilon\right)$ croissent

avec ε , et que leur intersection vérifie $\bigcap_{\varepsilon > 0} K_\varepsilon = \overline{D}\left(\frac{a}{2}, \frac{1}{2}\right) \subset \Omega$. Comme Ω est ouvert, il contient le compact K_{ε_0} ,

pour un certain $\varepsilon_0 > 0$. On considère finalement le voisinage $V = \{ae^{i\theta}, |\theta| \leq \varepsilon_0\}$ de a . Pour $b \in V$, les inégalités de Cauchy impliquent que

$$\frac{|f^{(k)}(b/2)|}{k!} = \frac{|g^{(k)}(b/2)|}{k!} \leq \sup_{|z-b/2|=1/2+\varepsilon_0} \frac{|g(z)|}{(1/2 + \varepsilon_0)^k} \leq \frac{C_a}{\rho_a^k},$$

où l'on a posé $C_a = \sup_{z \in K_{\varepsilon_0}} |g(z)|$ et $\rho_a = 1/2 + \varepsilon_0$.

5. L'équivalence découle de la preuve menée à la question précédente. Dans le cas où $f(z) = \sum_{n=0}^{\infty} z^n = \frac{1}{1-z}$, on

calcule $f^{(k)}(z) = \frac{k!}{(1-z)^{k+1}}$, ce qui donne $\sum_{k=0}^{\infty} \frac{f^{(k)}(a/2)}{k!} w^k = \sum_{k=0}^{\infty} \frac{w^k}{(1-a/2)^{k+1}}$, qui est une série entière dont le rayon de convergence vaut $|1 - a/2|$. Cette valeur est $> 1/2$ si et seulement si $a \neq 1$, ce qui confirme que 1 est bien le seul point singulier de la série entière f .

6. Supposons que le point 1 est régulier. D'après la question précédente, la série entière $\sum_{k=0}^{\infty} \frac{f^{(k)}(1/2)}{k!} w^k$ admet

un rayon de convergence $> 1/2$. On a donc $\sum_{k=0}^{\infty} \frac{|f^{(k)}(1/2)|}{k!} r_0^k < \infty$ pour un certain $r_0 > 1/2$. En utilisant le développement en série entière des dérivées successives de f , puis le théorème de Fubini-Tonelli (les termes sont positifs), on peut écrire que

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{|f^{(k)}(1/2)|}{k!} r_0^k &= \sum_{k=0}^{\infty} \frac{r_0^k}{k!} \sum_{n=k}^{\infty} n(n-1) \dots (n-k+1) a_n (1/2)^{n-k} \\ &= \sum_{n=0}^{\infty} a_n \sum_{k=0}^n \frac{r_0^k}{k!} n(n-1) \dots (n-k+1) (1/2)^{n-k} \\ &= \sum_{n=0}^{\infty} a_n \sum_{k=0}^n \binom{n}{k} r_0^k (1/2)^{n-k} = \sum_{n=0}^{\infty} a_n (r_0 + 1/2)^n. \end{aligned}$$

Cette dernière série converge donc. Comme $r_0 + 1/2 > 1$, cela contredit le fait que la série entière f admet 1 pour rayon de convergence.

7. L'hypothèse de lacunarité implique l'existence d'un entier $p \geq 1$ tel que $p\lambda_{n+1} > (p+1)\lambda_n$ pour tout $n \geq 1$. On suppose que la somme f de la série entière $\sum_{n=1}^{\infty} a_n z^{\lambda_n}$ admet un prolongement analytique g sur un ensemble de la forme $D \cup D(1, t) = \Omega$, avec $t > 0$. On note $\varphi(z) = \frac{z^p}{2}(1+z)$, et on observe que $\varphi(\overline{D}) \subset \Omega$. Ainsi, on a

$\overline{D}_{\varepsilon_0} \subset \varphi^{-1}(\Omega)$ pour un certain $\varepsilon_0 > 0$, où $\overline{D}_\varepsilon$ désigne le disque fermé de centre 0 et de rayon $1 + \varepsilon$. La fonction $g \circ \varphi$ est holomorphe dans le disque $D(0, R)$, avec $R = 1 + \varepsilon_0$, donc développable en série entière :

$$g\left(\frac{z^p + z^{p+1}}{2}\right) = \sum_{n=0}^{\infty} b_n z^n$$

pour $|z| < R$. Cependant, si $|z| < 1$, on a

$$g\left(\frac{z^p + z^{p+1}}{2}\right) = f\left(\frac{z^p + z^{p+1}}{2}\right) = \sum_{n=1}^{\infty} a_n \left(\frac{z^p + z^{p+1}}{2}\right)^{\lambda_n} = \sum_{n=1}^{\infty} a_n P_n(z).$$

La plus grande puissance de z figurant dans P_n est $(p+1)\lambda_n$. Elle est strictement inférieure à la plus petite puissance de z figurant dans P_{n+1} , qui est $p\lambda_{n+1}$. Chaque puissance de z figure donc dans au plus un terme de la forme $P_n(z)$. Pour tout entier $N \geq 1$, on a donc

$$\sum_{n=1}^N a_n \left(\frac{z^p + z^{p+1}}{2}\right)^{\lambda_n} = \sum_{k=0}^{(p+1)\lambda_N} b_k z^k.$$

Pour $z \in]1, R[$, le membre de droite converge quand $N \rightarrow \infty$ vers $g(w)$, avec $w = \varphi(z) = \frac{z^p + z^{p+1}}{2} > 1$. La série $\sum_{n=1}^{\infty} a_n w^{\lambda_n}$ est donc convergente, ce qui contredit le fait que son rayon de convergence vaut 1. On en conclut que le point 1 est singulier.

Pour en déduire qu'un point quelconque $z_0 \in \Gamma$ est singulier, on applique ce qui précède à la série entière $\sum_{n=1}^{\infty} a_n z_0^{\lambda_n} z^{\lambda_n}$, puisque 1 est singulier pour cette série entière si et seulement si z_0 est singulier pour la série entière de départ $\sum_{n=1}^{\infty} a_n z^{\lambda_n}$.

Exercice 18.8 (Suite de Morse)

Leçon : 243. Référence : [13].

On définit la suite de Morse par $\varepsilon_0 = 1$, puis $\varepsilon_{2n} = \varepsilon_n$ et $\varepsilon_{2n+1} = -\varepsilon_n$, puis la série entière $f(z) = \sum_{n=0}^{\infty} \varepsilon_n z^n$ pour tout z dans D le disque unité ouvert.

1. Montrer que $f(z) = (1 - z)f(z^2)$ pour tout $z \in D$.
2. En déduire que $f(z) = \prod_{n=0}^{\infty} (1 - z^{2^n})$ pour tout $z \in D$.
3. Soit I un arc du cercle unité Γ , soit $u = \exp \frac{2ik\pi}{2^N}$ un point dyadique de cet arc, et soit $r \in [0, 1[$. Montrer la majoration $|f(ru)| \leq 2^N(1 - r^{2^N})$.
4. Montrer que Γ est une coupure pour f .

Corrigé.

On définit la suite de Morse par $\varepsilon_0 = 1$, puis $\varepsilon_{2n} = \varepsilon_n$ et $\varepsilon_{2n+1} = -\varepsilon_n$, puis la série entière $f(z) = \sum_{n=0}^{\infty} \varepsilon_n z^n$ pour tout z dans D le disque unité ouvert.

1. On sépare simplement les termes d'indices pairs et impairs dans la somme de la série, puis on utilise la définition de la suite de Morse : pour tout $z \in D$,

$$f(z) = \sum_{n=0}^{\infty} \varepsilon_{2n} z^{2n} + \sum_{n=0}^{\infty} \varepsilon_{2n+1} z^{2n+1} = \sum_{n=0}^{\infty} \varepsilon_n z^{2n} - \sum_{n=0}^{\infty} \varepsilon_n z^{2n+1} = f(z^2) - z f(z^2) = (1 - z)f(z^2).$$

2. En itérant la relation obtenue à la question précédente, on obtient pour tout entier $N \geq 0$,

$$f(z) = (1 - z)f(z^2) = (1 - z)(1 - z^2)f(z^4) = \dots = f(z^{2^N}) \prod_{n=0}^{N-1} (1 - z^{2^n}).$$

Pour $z \in D$, on a par continuité $f(z^{2^N}) \xrightarrow{N \rightarrow \infty} f(0) = \varepsilon_0 = 1$. D'où la relation $f(z) = \prod_{n=0}^{\infty} (1 - z^{2^n})$.

3. On applique la relation ci-dessus au point ru avec $u = \exp \frac{2ik\pi}{2^N}$ et $r \in [0, 1[$, en observant que $u^{2^n} = 1$ pour tout entier $n \geq N$. Cela donne

$$|f(ru)| = \prod_{n=0}^{\infty} |1 - (ru)^{2^n}| = \left(\prod_{n=0}^{N-1} |1 - (ru)^{2^n}| \right) (1 - r^{2^N}) \left(\prod_{n=N+1}^{\infty} (1 - r^{2^n}) \right) \leq 2^N (1 - r^{2^N}).$$

4. Supposons que $a \in \Gamma$ soit un point régulier pour f , et notons g un prolongement analytique de f sur un ensemble de la forme $D \cup D_a$, où D_a est un disque ouvert centré en a . Un point dyadique $u = \exp \frac{2ik\pi}{2^N}$ de l'arc $\Gamma \cap D_a$ vérifie alors d'après la majoration de la question précédente

$$|g(u)| = \lim_{r \rightarrow 1^-} |f(ru)| \leq \liminf_{r \rightarrow 1^-} 2^N (1 - r^{2^N}) = 0.$$

Par densité des points dyadiques dans l'arc, on en déduit que la fonction g est nulle sur $\Gamma \cap D_a$. Par unicité du prolongement analytique, g est en fait nulle sur $D \cup D_a$. Cela implique que f est nulle sur D , ce qui est faux. Finalement, tous les points de Γ sont singuliers, c'est-à-dire que Γ est une coupure pour f .

Exercice 18.9 (La fonction sinus)

Leçon : 245.

On considère la fonction $\sin$ définie sur $\mathbb{C}$ par la formule $\sin z = \frac{e^{iz} - e^{-iz}}{2i}$.

- Justifier que $\sin$ est une fonction holomorphe sur $\mathbb{C}$.
- Pour quelles valeurs de $z \in \mathbb{C}$ a-t-on $\sin z = 0$?
- Montrer que la fonction définie par $f(z) = \sin \frac{\pi}{1-z}$ est holomorphe sur le disque unité ouvert. Quels sont les zéros de f sur ce disque ? Est-ce contradictoire avec le principe des zéros isolés ?

Corrigé.

- La fonction $\sin$ est holomorphe sur $\mathbb{C}$ comme somme et composée de fonctions holomorphes sur $\mathbb{C}$.
- L'équation $\sin z = 0$ est équivalente à $e^{iz} = e^{-iz}$, soit $e^{2iz} = e^{i0}$, c'est-à-dire $2iz = 0 + 2ik\pi$, avec $k \in \mathbb{Z}$, soit finalement $z = k\pi$, pour $k \in \mathbb{Z}$.
- Comme précédemment, la fonction f est holomorphe sur le disque unité ouvert comme composée de fonctions holomorphes sur le disque unité ouvert. Ses zéros sont les nombres complexes $z \in \mathbb{C}$ pour lesquels il existe un entier $k \in \mathbb{Z}$ tel que $\frac{\pi}{1-z} = k\pi$, c'est-à-dire $z = 1 - \frac{1}{k}$ avec seulement $k \in \mathbb{N}^*$ pour que les zéros appartiennent bien au disque unité ouvert. Il y a en particulier une accumulation de zéros autour du point 1, mais cela ne contredit pas le principe des zéros isolés car la fonction f n'est pas holomorphe sur un voisinage ouvert de 1.

Exercice 18.10 (Annulation dans un compact)

Leçon : 245.

Soit f une fonction holomorphe dans un ouvert connexe V de $\mathbb{C}$, non identiquement nulle, et soit K un compact contenu dans V . On note Z l'ensemble des points $z \in K$ tels que $f(z) = 0$.

- Montrer que Z est un compact de $\mathbb{C}$.
- En déduire que Z est fini.

Corrigé.

- L'ensemble Z est l'intersection du compact K et de l'image réciproque $f^{-1}(\{0\})$, qui est fermée puisque f est continue. Cette intersection est donc un fermé inclus dans un compact, donc est compacte.

2. D'après le principe des zéros isolés, pour tout $z \in Z$, il existe un réel $r_z > 0$ tel que f ne s'annule pas sur le disque ouvert $D(z, r_z)$ ailleurs qu'au point z . De plus, ces disques forment un recouvrement de Z par des ouverts, donc par compacité, on peut extraire un recouvrement de Z par un nombre fini de disques $D(z_p, r_{z_p})$, pour $p \in \{1, \dots, n\}$. Comme f ne s'annule pas ailleurs qu'au centre de chacun de ces disques, on en déduit que Z est en fait l'ensemble $\{z_1, \dots, z_n\}$, donc est un ensemble fini.

Exercice 18.11 (Symétrie)

Leçon : 245.

Soit U un ouvert de $\mathbb{C}$ et f une fonction holomorphe sur U . Soit V l'ensemble des nombres complexes $z \in \mathbb{C}$ dont le conjugué vérifie $\bar{z} \in U$. On pose $g(z) = \overline{f(\bar{z})}$ pour $z \in V$. Montrer que la fonction g est holomorphe sur V .

Corrigé.

Vérifions que la fonction g satisfait les équations de Cauchy-Riemann en tout point de V . Si on pose $z = x + iy$ et $f(z) = P(x, y) + iQ(x, y)$, alors on a

$$g(z) = \overline{f(x - iy)} = P(x, -y) - iQ(x, -y) = u(x, y) + i v(x, y),$$

avec $u(x, y) = P(x, -y)$ et $v(x, y) = -Q(x, -y)$. Alors, on a

$$\frac{\partial u}{\partial x}(x, y) = \frac{\partial P}{\partial x}(x, -y) = \frac{\partial Q}{\partial y}(x, -y) \quad \text{et} \quad \frac{\partial v}{\partial y}(x, y) = -\left(-\frac{\partial Q}{\partial y}(x, -y)\right) = \frac{\partial Q}{\partial y}(x, -y).$$

On a donc bien $\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$ et on peut vérifier de même que $\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$. Donc, g est une fonction holomorphe.

Exercice 18.12 (Existence de primitives)

Leçon : 245.

Soit $z_0 \in \mathbb{C}$ et soit γ un arc de cercle de centre z_0 , de rayon strictement positif, orienté dans le sens trigonométrique. On note $\alpha \in [0, 2\pi]$ l'angle décrit par cet arc de cercle (si $\alpha = 2\pi$, alors on a le cercle entier). Pour $z \in \mathbb{C} \setminus \{z_0\}$, on pose $f(z) = \frac{1}{z - z_0}$.

1. Calculer l'intégrale de f le long de γ .
2. La fonction f admet-elle une primitive dans $\mathbb{C} \setminus \{z_0\}$?

Corrigé.

1. On considère comme paramétrisation de l'arc γ l'application $t \mapsto z_0 + r e^{it}$ définie sur un intervalle de la forme $[\theta, \theta + \alpha]$. On a donc

$$\int_{\gamma} f(z) dz = \int_{\theta}^{\theta+\alpha} \frac{r i e^{it}}{r e^{it}} dt = i\alpha.$$

2. Si on prend pour γ le cercle de centre z_0 et de rayon r , alors si f admettait une primitive F , son intégrale serait nulle. Ce n'est pas le cas ici, donc f n'admet pas de primitive dans $\mathbb{C} \setminus \{z_0\}$.

Exercice 18.13 (Un développement en série)

Leçon : 245.

Le but de cet exercice est de démontrer la formule suivante :

$$\forall z \in \mathbb{C} \setminus \mathbb{Z} \quad \sum_{n \in \mathbb{Z}} \frac{1}{(z - n)^2} = \left(\frac{\pi}{\sin(\pi z)} \right)^2.$$

1. On note $\operatorname{sh} y = (e^y - e^{-y})/2$ et $\operatorname{ch} y = (e^y + e^{-y})/2$ les fonctions sinus et cosinus hyperboliques. Montrer que pour tout $z = x + iy$, on a $|\sin z|^2 = \sin^2 x + \operatorname{sh}^2 y$.
2. Montrer que la série $f(z) = \sum_{n \in \mathbb{Z}} \frac{1}{(z - n)^2}$ converge normalement sur tout compact de $\mathbb{C} \setminus \mathbb{Z}$. En déduire que f définit une fonction méromorphe sur $\mathbb{C}$ dont les pôles sont les points de $\mathbb{Z}$.

3. On pose $g(z) = \left(\frac{\pi}{\sin(\pi z)}\right)^2$. Montrer que f et g ont même partie singulière en 0. En déduire que $h = f - g$ se prolonge en une fonction entière.
4. Montrer que h est bornée sur l'ensemble $\{0 \leq \Re z \leq 1 \text{ et } |\Im z| > 1\}$. En déduire que h est constante puis, en étudiant $\lim_{y \rightarrow +\infty} h(iy)$, montrer que $h = 0$.
5. Montrer que $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.

Corrigé.

1. On écrit $\sin z = \sin(x + iy) = \sin x \cos iy + \cos x \sin iy$, puis on remarque que $\cos iy = \operatorname{ch} y$ et $\sin iy = i \operatorname{sh} y$. On obtient donc $\sin z = \sin x \operatorname{ch} y + i \cos x \operatorname{sh} y$, puis $|\sin z|^2 = \sin^2 x \operatorname{ch}^2 y + \cos^2 x \operatorname{sh}^2 y = \sin^2 x (1 + \operatorname{sh}^2 y) + (1 - \sin^2 x) \operatorname{sh}^2 y = \sin^2 x + \operatorname{sh}^2 y$.
2. Soit K un compact de $\mathbb{C} \setminus \mathbb{Z}$. Choisissons $R > 0$ tel que $|z| \leq R$ sur K . Pour $|n| > 2R$, on a $|z - n| \geq |n|/2$, donc $|z - n|^{-2} \leq 4/n^2$. Les termes restants sont en nombre fini et bornés sur K . La série converge donc normalement sur K et sa somme est holomorphe sur $\mathbb{C} \setminus \mathbb{Z}$. Au voisinage de chaque entier m , la série privée du terme $(z - m)^{-2}$ est holomorphe par le même argument. Ainsi, f est méromorphe sur $\mathbb{C}$, avec un pôle double en chaque entier. Un changement d'indice montre qu'elle est 1-périodique.
3. On écrit $f(z) = \frac{1}{z^2} + u(z)$, avec $u(z) = \sum_{n \neq 0} \frac{1}{(z - n)^2}$. La série définissant u est une série de fonctions holomorphes au voisinage de 0 qui converge uniformément dans un voisinage de 0. La fonction u est donc holomorphe, ce qui montre que la partie singulière de f en 0 est $1/z^2$. Pour montrer que c'est aussi la partie singulière de g en 0, on observe que $z \mapsto g(z) - 1/z^2$ est bornée au voisinage de 0 car

$$\frac{1}{g(z)} = \frac{\sin^2(\pi z)}{\pi^2} = z^2 + O(|z|^4).$$

Comme les fonctions f et g ont même partie singulière en 0, on en déduit que leur différence $h = f - g$ est une fonction holomorphe sur un voisinage épointé de 0, et bornée au voisinage de 0. La singularité de h en 0 est donc effaçable, de sorte que h se prolonge en une fonction holomorphe sur un voisinage de 0. Par 1-périodicité de h , toutes les singularités aux entiers sont donc effaçables, et h se prolonge en une fonction entière.

4. Pour $x \in [0, 1]$ et $|y| > 1$, on a $|\sin z|^2 = \sin^2 x + \operatorname{sh}^2 y$, donc $|\sin \pi z| \geq |\operatorname{sh} \pi y| \geq \operatorname{sh} \pi$. De plus,

$$|f(z)| \leq \sum_{n \in \mathbb{Z}} \frac{1}{1 + (x - n)^2} \leq \sum_{n \leq 0} \frac{1}{1 + n^2} + \sum_{n \geq 1} \frac{1}{1 + (n - 1)^2}.$$

On en déduit que les fonctions g et f , et donc h , sont bornées sur l'ensemble $\{0 \leq \Re z \leq 1 \text{ et } |\Im z| > 1\}$. La fonction h est aussi bornée sur le compact $\{0 \leq \Re z \leq 1 \text{ et } |\Im z| \leq 1\}$, car elle est continue. On en déduit que h est bornée sur la bande $\{0 \leq \Re z \leq 1\}$, et par 1-périodicité, sur $\mathbb{C}$ tout entier. Le théorème de Liouville implique alors que h est constante. Finalement, on observe que

$$|\sin \pi iy| \geq |\operatorname{sh} \pi y| \quad \text{et} \quad |f(iy)| \leq \sum_{n \in \mathbb{Z}} \frac{1}{y^2 + n^2},$$

ce qui implique respectivement que $|g(iy)|$ et $|f(iy)|$ tendent tous deux vers 0 quand $y \rightarrow +\infty$. (Pour la limite de f , on peut utiliser le théorème de convergence dominée.) On en conclut que $|h(iy)| \rightarrow 0$ quand $y \rightarrow +\infty$, et donc que h est la fonction nulle.

5. Comme $h(0) = 0$, on sait que $f(z) - 1/z^2$ et $g(z) - 1/z^2$ admettent la même limite a quand $z \rightarrow 0$. En utilisant la fonction f , on observe que $a = u(0) = 2 \sum_{n=1}^{\infty} \frac{1}{n^2}$. Pour utiliser g , on part de

$$\frac{1}{g(z)} = \frac{\sin^2(\pi z)}{\pi^2} = z^2 - \frac{\pi^2}{3} z^4 + O(|z|^6),$$

ce qui donne finalement $a = \pi^2/3$, puis le résultat voulu en identifiant les deux valeurs ainsi trouvées pour a .

Transformation de Fourier

Références : [12, 14].

Exercice 19.1 (Morphismes complexes de l'algèbre $L^1(\mathbb{R}^d)$)

Leçons : 201, 250. Référence : [14].

1. On munit l'espace de Banach $L^1(\mathbb{R}^d)$ du produit de convolution défini par $f * g(x) = \int_{\mathbb{R}^d} f(x-y) g(y) dy$, qui lui donne donc une structure d'algèbre. Montrer que $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$ pour toutes fonctions $f, g \in L^1(\mathbb{R}^d)$. On dit alors que $L^1(\mathbb{R}^d)$ est une algèbre de Banach.
2. On s'intéresse aux morphismes complexes de l'algèbre $L^1(\mathbb{R}^d)$, c'est-à-dire aux formes linéaires $\varphi : L^1(\mathbb{R}^d) \rightarrow \mathbb{C}$ qui vérifient $\varphi(f * g) = \varphi(f)\varphi(g)$ pour toutes fonctions $f, g \in L^1(\mathbb{R}^d)$. Montrer qu'une telle application est continue, puis qu'il existe une fonction $\beta \in L^\infty(\mathbb{R}^d)$ telle que $\varphi(f) = \int_{\mathbb{R}^d} f(x) \beta(x) dx$ quelle que soit $f \in L^1(\mathbb{R}^d)$.
3. Montrer qu'on peut choisir la fonction β continue et telle que $\beta(x+y) = \beta(x)\beta(y)$ pour tous $x, y \in \mathbb{R}^d$.
4. Montrer que toutes les dérivées partielles de la fonction β existent sur $\mathbb{R}^d$ et vérifient $\partial_j \beta(x) = \beta(x) \partial_j \beta(0)$ pour tous $j \in \{1, \dots, d\}$ et $x \in \mathbb{R}^d$.
5. En déduire que, si le morphisme φ n'est pas identiquement nul, alors il existe un élément $\xi \in \mathbb{R}^d$ tel que $\beta(x) = e^{-i\xi \cdot x}$ pour tout $x \in \mathbb{R}^d$.
6. Conclure que pour tout morphisme complexe φ de l'algèbre $L^1(\mathbb{R}^d)$ non identiquement nul, il existe un unique $\xi \in \mathbb{R}^d$ tel que $\varphi(f) = \hat{f}(\xi)$ pour toute fonction $f \in L^1(\mathbb{R}^d)$. (On précisera alors la formule utilisée pour la transformée de Fourier.)

Corrigé.

1. En utilisant l'inégalité triangulaire pour les intégrales et le théorème de Fubini-Tonelli, on peut écrire que

$$\int_{\mathbb{R}^d} |f * g(x)| dx \leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |f(x-y)| |g(y)| dx dy = \int_{\mathbb{R}^d} |g(y)| \underbrace{\int_{\mathbb{R}^d} |f(x-y)| dx}_{=\|f\|_1} dy = \|f\|_1 \|g\|_1,$$

ce qui donne directement $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$.

2. Supposons par l'absurde l'existence d'une fonction $f_0 \in L^1(\mathbb{R}^d)$ telle que $|\varphi(f_0)| > \|f_0\|_1$, puis considérons la fonction $f = f_0 / \varphi(f_0)$. On a clairement $\|f\|_1 < 1$ et $\varphi(f) = 1$. En particulier, la série $\sum_{n=1}^{\infty} \underbrace{f * \dots * f}_{n \text{ fois}}$ converge normalement. Sa somme, notée g , vérifie l'identité $g - f = f * g$. Par conséquent, $\varphi(g) - \varphi(f) = \varphi(f)\varphi(g)$, ce qui donne $\varphi(g) - 1 = \varphi(g)$, et donc une absurdité. Finalement, le morphisme φ définit bien une forme linéaire continue, de norme inférieure à 1. Il s'agit donc d'un élément du dual de $L^1(\mathbb{R}^d)$, qui s'identifie à $L^\infty(\mathbb{R}^d)$. On en déduit l'existence d'une fonction $\beta \in L^\infty(\mathbb{R}^d)$ telle que $\varphi(f) = \int_{\mathbb{R}^d} f(x) \beta(x) dx$ quelle que soit $f \in L^1(\mathbb{R}^d)$.
3. Pour toutes fonctions $f, g \in L^1(\mathbb{R}^d)$, on sait que $\varphi(f * g) = \varphi(f)\varphi(g)$. D'une part, le membre de gauche s'écrit

$$\varphi(f * g) = \int_{\mathbb{R}^d} f * g(x) \beta(x) dx = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \beta(x) f(x-y) g(y) dx dy = \int_{\mathbb{R}^d} g(y) \varphi(f_y) dy,$$

où f_y désigne la fonction $x \mapsto f(x-y)$. D'autre part, le membre de droite peut s'écrire sous la forme

$$\varphi(f)\varphi(g) = \varphi(f) \int_{\mathbb{R}^d} g(y) \beta(y) dy = \int_{\mathbb{R}^d} g(y) \varphi(f) \beta(y) dy.$$

Cela signifie que la forme linéaire continue $g \mapsto \int_{\mathbb{R}^d} g(y) (\varphi(f_y) - \varphi(f)\beta(y)) dy$ est nulle sur $L^1(\mathbb{R}^d)$, et donc que la fonction $y \mapsto \varphi(f_y) - \varphi(f)\beta(y)$ est nulle presque partout. On suppose qu'on peut travailler avec une fonction $f \in L^1(\mathbb{R}^d)$ telle que $\varphi(f) \neq 0$. En effet, dans le cas contraire, la forme linéaire φ et donc la fonction β sont nulles, et le résultat est trivial. Dès lors, on peut écrire que $\beta(y) = \frac{\varphi(f_y)}{\varphi(f)}$ pour presque tout $y \in \mathbb{R}^d$. Cependant, la fonction $y \mapsto f_y$ est continue de $\mathbb{R}^d$ dans $L^1(\mathbb{R}^d)$, et φ est une forme linéaire continue sur $L^1(\mathbb{R}^d)$. Il en ressort que la fonction $y \mapsto \frac{\varphi(f_y)}{\varphi(f)}$ est continue sur $\mathbb{R}^d$. Ainsi, quitte à modifier β sur un ensemble de mesure nulle, on peut supposer qu'elle est continue. De plus, la fonction β vérifie l'identité $\varphi(f_y) = \varphi(f)\beta(y)$ pour toute fonction $f \in L^1(\mathbb{R}^d)$ et tout point $y \in \mathbb{R}^d$. Par conséquent, pour tous $x, y \in \mathbb{R}^d$,

$$\varphi(f)\beta(x+y) = \varphi(f_{x+y}) = \varphi((f_x)_y) = \varphi(f_x)\beta(y) = \varphi(f)\beta(x)\beta(y).$$

En reprenant à nouveau f telle que $\varphi(f) \neq 0$, on en déduit l'identité $\beta(x+y) = \beta(x)\beta(y)$ pour tous $x, y \in \mathbb{R}^d$.

4. On peut à nouveau supposer que le morphisme φ n'est pas identiquement nul, puisque sinon le résultat est trivial. Ainsi, la fonction β n'est pas identiquement nulle. Comme $\beta(x) = \beta(x)\beta(0)$ pour tout $x \in \mathbb{R}^d$, on a donc $\beta(0) = 1$. Par continuité, il existe $\delta > 0$ tel que $\int_{\|y\|_\infty < \delta} \beta(y) dy = c \neq 0$. D'où, pour tout $x \in \mathbb{R}^d$,

$$c\beta(x) = \int_{\|y\|_\infty < \delta} \beta(x)\beta(y) dy = \int_{\|y\|_\infty < \delta} \beta(x+y) dy = \int_{\|y-x\|_\infty < \delta} \beta(y) dy = \prod_{j=1}^d \int_{x_j-\delta}^{x_j+\delta} \beta(te_j) dt.$$

Ici, $(e_1, \dots, e_d)$ désigne la base canonique de $\mathbb{R}^d$ et $(x_1, \dots, x_d)$ les coordonnées de x dans cette base. On observe que $\beta(y) = \beta(y_1 e_1) \dots \beta(y_d e_d)$ pour tout point $y \in \mathbb{R}^d$. On observe que le membre de droite est dérivable par rapport à chaque variable x_j , car la fonction β est continue. Toutes les dérivées partielles de β existent donc. De plus, en dérivant par rapport à t la relation $\beta(x + te_j) = \beta(x)\beta(te_j)$ et en fixant $t = 0$, on obtient $\partial_j \beta(x) = \beta(x) \partial_j \beta(0)$ pour tous $j \in \{1, \dots, d\}$ et $x \in \mathbb{R}^d$.

5. Posons $\beta_j(t) = \beta(te_j)$. D'après la question précédente, $\beta'_j(t) = \partial_j \beta(0) \beta_j(t)$ et $\beta_j(0) = 1$, donc $\beta_j(t) = \exp(t \partial_j \beta(0))$. La fonction β_j étant bornée sur toute la droite réelle, on a $\Re(\partial_j \beta(0)) = 0$. Il existe donc $\xi_j \in \mathbb{R}$ tel que $\partial_j \beta(0) = -i\xi_j$. En posant $\xi = (\xi_1, \dots, \xi_d)$, on obtient

$$\beta(x) = \prod_{j=1}^d \beta_j(x_j) = e^{-i\xi \cdot x}.$$

6. Finalement, en supposant toujours φ non identiquement nul et en utilisant l'écriture ci-dessus pour la fonction β , on obtient l'existence d'un élément $\xi \in \mathbb{R}^d$ tel que $\varphi(f) = \int_{\mathbb{R}^d} f(x) e^{-i\xi \cdot x} dx = \widehat{f}(\xi)$ pour toute fonction $f \in L^1(\mathbb{R}^d)$, en prenant cette formule comme définition de la transformée de Fourier. L'unicité de ξ pour cette propriété découle du fait que si $\xi \neq \zeta$, alors il existe une fonction $f \in L^1(\mathbb{R}^d)$ telle que $\widehat{f}(\xi) \neq \widehat{f}(\zeta)$. Pour construire une telle fonction, on peut par exemple utiliser la fonction définie par $g(x) = e^{-|x_1| - \dots - |x_d|}$ pour tout $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, dont la transformée de Fourier est la fonction $\widehat{g}(\xi) = \frac{2^d}{(1 + \xi_1^2) \dots (1 + \xi_d^2)}$. Si le produit $(1 + \xi_1^2) \dots (1 + \xi_d^2)$ est différent de $(1 + \zeta_1^2) \dots (1 + \zeta_d^2)$, alors cette fonction g convient. Sinon, on peut utiliser une translatée de la forme $g_\alpha = g(\cdot - \alpha)$, dont la transformée de Fourier vérifie alors $\widehat{g}_\alpha(\xi) = \widehat{g}(\xi) e^{-i\alpha \cdot \xi}$ et de même pour ζ , de sorte que $\widehat{g}_\alpha(\xi) \neq \widehat{g}_\alpha(\zeta)$ si et seulement si $e^{i\alpha \cdot (\xi - \zeta)} \neq 1$, ce qui est vérifié pour $\alpha = \frac{\xi - \zeta}{\|\xi - \zeta\|^2}$.

Exercice 19.2 (Formule sommatoire de Poisson, fonction θ de Jacobi et prolongement de la fonction ζ)

Leçons : 250, 265. Références : [12, 14].

Dans cet exercice, on définit la transformée de Fourier d'une fonction $f \in L^1(\mathbb{R})$ par la formule

$$\widehat{f}(\xi) = \int_{-\infty}^{+\infty} f(x) e^{-2i\pi\xi x} dx.$$

1. Soit f une fonction continue et intégrable sur $\mathbb{R}$. On suppose qu'il existe $M > 0$ et $\alpha > 1$ tels que $|f(x)| \leq M(1 + |x|)^{-\alpha}$ pour tout $x \in \mathbb{R}$. Montrer que la formule $\varphi(x) = \sum_{n=-\infty}^{+\infty} f(x+n)$ définit une fonction continue et 1-périodique sur $\mathbb{R}$.
2. Pour tout $m \in \mathbb{Z}$, exprimer en fonction de $\widehat{f}$ le coefficient de Fourier $c_m(\varphi) = \int_0^1 \varphi(x) e^{-2i\pi mx} dx$.
3. On suppose en outre que $\sum_{n=-\infty}^{+\infty} |\widehat{f}(n)| < \infty$. Écrire le développement en série de Fourier de la fonction φ , puis en déduire la formule sommatoire de Poisson : $\sum_{n=-\infty}^{+\infty} f(n) = \sum_{n=-\infty}^{+\infty} \widehat{f}(n)$.
4. La fonction θ de Jacobi est définie par $\theta(t) = \sum_{n=-\infty}^{+\infty} e^{-\pi n^2 t}$ pour tout réel $t > 0$. Montrer que $\theta(1/t) = \theta(t)\sqrt{t}$ pour tout réel $t > 0$. En déduire une équation fonctionnelle portant sur la fonction $\tilde{\theta}(t) = \sum_{n=1}^{\infty} e^{-\pi n^2 t}$.
5. On rappelle que $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ et $\Gamma(s) = \int_0^{\infty} x^{s-1} e^{-x} dx$, pour tout nombre complexe s de partie réelle strictement supérieure à 1 et 0, respectivement. Pour $\Re s > 1$, montrer que $\zeta(s) = \frac{\pi^{s/2}}{\Gamma(s/2)} I(s)$, où I désigne la fonction définie par $I(s) = \int_0^{\infty} \tilde{\theta}(t) t^{s/2-1} dt$.
6. Montrer que la fonction $J : s \mapsto I(s) - \frac{1}{s(s-1)}$, définie pour $\Re s > 1$, admet un prolongement holomorphe à $\mathbb{C}$.
7. Pour $\Re s > 1$, montrer que $\zeta(s) = \frac{1}{s-1} + \psi(s) + \frac{\pi^{s/2}}{\Gamma(s/2)} J(s)$, où ψ est une fonction holomorphe sur $\mathbb{C}$. En déduire que la fonction ζ admet un prolongement analytique au domaine $\mathbb{C} \setminus \{1\}$.

Corrigé.

1. Si $|x| \leq A$ et $|n| \geq 2A$, alors on a $|x+n| \geq |n| - |x| \geq |n| - A \geq |n|/2$, de sorte que $|f(x+n)| \leq \frac{M}{(1 + |n|/2)^\alpha}$. Comme $\alpha > 1$, il en résulte que la série qui définit la fonction $\varphi(x) = \sum_{n=-\infty}^{+\infty} f(x+n)$ converge normalement sur tout compact de $\mathbb{R}$. Comme la fonction de départ f est continue sur $\mathbb{R}$, on en déduit que φ est continue aussi. Enfin, la 1-périodicité de φ découle de façon élémentaire d'un changement d'indice dans la somme.
2. D'après le théorème de Fubini-Tonelli (que l'on peut bien appliquer par convergence normale sur tout compact, et parce que l'exponentielle complexe est bornée), on a pour tout $m \in \mathbb{Z}$,

$$\begin{aligned} c_m(\varphi) &= \int_0^1 \varphi(x) e^{-2i\pi mx} dx = \int_0^1 \sum_{n=-\infty}^{+\infty} f(x+n) e^{-2i\pi mx} dx = \sum_{n=-\infty}^{+\infty} \int_0^1 f(x+n) e^{-2i\pi mx} dx \\ &= \sum_{n=-\infty}^{+\infty} \int_0^1 f(x+n) e^{-2i\pi m(x+n)} dx = \sum_{n=-\infty}^{+\infty} \int_n^{n+1} f(x) e^{-2i\pi mx} dx = \int_{-\infty}^{+\infty} f(x) e^{-2i\pi mx} dx = \widehat{f}(m). \end{aligned}$$

3. La nouvelle hypothèse implique directement que la série de Fourier de la fonction φ converge normalement.

On a donc $\varphi(x) = \sum_{m=-\infty}^{+\infty} c_m(\varphi) e^{2i\pi mx} = \sum_{m=-\infty}^{+\infty} \widehat{f}(m) e^{2i\pi mx}$ pour tout $x \in \mathbb{R}$. On obtient directement la formule sommatoire de Poisson en prenant $x = 0$ dans cette dernière identité.

4. On applique la formule de Poisson à la fonction $f : x \mapsto e^{-\pi x^2 t}$ pour $t > 0$ fixé, dont la transformée de Fourier est la fonction $\widehat{f} : \xi \mapsto t^{-1/2} e^{-\pi \xi^2 / t}$. Cela donne directement $\theta(t) = t^{-1/2} \theta(1/t)$ pour tout $t > 0$. On observe ensuite que $\theta(t) = 1 + 2\tilde{\theta}(t)$, ce qui donne l'équation fonctionnelle $\tilde{\theta}(t) = \frac{1}{2\sqrt{t}} \left(2\tilde{\theta}\left(\frac{1}{t}\right) + 1 \right) - \frac{1}{2}$.

5. Soit $s \in \mathbb{C}$ tel que $\Re s > 1$. Le théorème de Fubini-Tonelli permet d'écrire que

$$\begin{aligned} I(s) &= \int_0^\infty \tilde{\theta}(t) t^{s/2-1} dt = \int_0^\infty \sum_{n=1}^\infty e^{-\pi n^2 t} t^{s/2-1} dt = \sum_{n=1}^\infty \int_0^\infty t^{s/2-1} e^{-\pi n^2 t} dt \\ &= \sum_{n=1}^\infty \int_0^\infty \frac{u^{s/2-1}}{\pi^{s/2-1} n^{s-2}} e^{-u} \frac{du}{\pi n^2} = \sum_{n=1}^\infty \int_0^\infty \frac{u^{s/2-1} e^{-u}}{\pi^{s/2} n^s} du = \frac{\Gamma(s/2)}{\pi^{s/2}} \zeta(s), \end{aligned}$$

ce qui conduit au résultat demandé.

6. En utilisant l'équation fonctionnelle vérifiée par la fonction $\tilde{\theta}$, on peut écrire que pour $\Re s > 1$,

$$\begin{aligned} I(s) &= \int_0^\infty \tilde{\theta}(t) t^{s/2-1} dt = \int_0^1 \tilde{\theta}(t) t^{s/2-1} dt + \int_1^\infty \tilde{\theta}(t) t^{s/2-1} dt \\ &= \int_1^\infty \tilde{\theta}(t^{-1}) t^{-s/2+1} \frac{dt}{t^2} + \int_1^\infty \tilde{\theta}(t) t^{s/2-1} dt \\ &= \int_1^\infty \frac{1}{2} \left((1 + 2\tilde{\theta}(t))\sqrt{t} - 1 \right) t^{-s/2-1} dt + \int_1^\infty \tilde{\theta}(t) t^{s/2-1} dt \\ &= \frac{1}{2} \int_1^\infty t^{-s/2-1/2} dt - \frac{1}{2} \int_1^\infty t^{-s/2-1} dt + \int_1^\infty \tilde{\theta}(t) (t^{-s/2-1/2} + t^{s/2-1}) dt \\ &= \frac{1}{s-1} - \frac{1}{s} + \int_1^\infty \tilde{\theta}(t) (t^{-s/2-1/2} + t^{s/2-1}) dt = \frac{1}{s(s-1)} + J(s). \end{aligned}$$

Il reste donc à montrer que l'intégrale qui définit la fonction J ci-dessus est holomorphe sur $\mathbb{C}$. Pour cela, on peut utiliser le théorème d'holomorphicité des intégrales à paramètre en se plaçant sur tout compact de $\mathbb{C}$.

7. D'après ce qui précède, on peut écrire que pour $\Re s > 1$,

$$\zeta(s) = \frac{\pi^{s/2}}{\Gamma(s/2)} \left(\frac{1}{s(s-1)} + J(s) \right) = \frac{\pi^{s/2}}{2(s-1)\Gamma(s/2+1)} + \frac{\pi^{s/2}}{\Gamma(s/2)} J(s).$$

Le second terme est une fonction holomorphe sur $\mathbb{C}$, en tant que produit de fonctions holomorphes sur $\mathbb{C}$, à condition de considérer le prolongement holomorphe de $1/\Gamma$ au plan complexe tout entier. Le premier terme peut s'écrire sous la forme

$$\frac{\pi^{s/2}}{2(s-1)\Gamma(s/2+1)} = \frac{1}{s-1} + \psi(s) \quad \text{avec} \quad \psi(s) = \frac{1}{s-1} \left(\frac{\pi^{s/2}}{2\Gamma(s/2+1)} - \frac{\pi^{1/2}}{2\Gamma(3/2)} \right),$$

puisque $\Gamma(3/2) = \Gamma(1/2)/2 = \sqrt{\pi}/2$. On peut donc voir la fonction ψ comme un taux d'accroissement d'une fonction holomorphe sur $\mathbb{C}$, donc il s'agit aussi d'une fonction holomorphe sur $\mathbb{C}$. On peut finalement écrire la fonction ζ sous la forme demandée, et en déduire qu'elle admet un prolongement analytique au domaine $\mathbb{C} \setminus \{1\}$.

Leçons couvertes par les exercices

Cette liste indique les exercices associés aux leçons dans les documents d'origine. Les numéros et les intitulés abrégés sont ceux de ces documents historiques ; ils ne décrivent pas le programme actuel du concours. Les numéros d'exercices et les pages ci-dessous renvoient aux énoncés, immédiatement suivis de leur corrigé.

Leçon 201 — Espaces de fonctions

Exercices : 12.3 (p. 122), 12.4 (p. 123), 19.1 (p. 171), 18.1 (p. 159).

Leçon 204 — Connexité

Exercices : 8.5 (p. 79), 8.6 (p. 80), 18.2 (p. 161), 18.3 (p. 162), 18.4 (p. 163).

Leçon 208 — Espaces vectoriels normés

Exercices : 9.1 (p. 82), 12.2 (p. 120), 12.4 (p. 123), 12.1 (p. 119), 13.1 (p. 131).

Leçon 214 — Inversion locale et fonctions implicites

Exercices : 10.7 (p. 89), 10.8 (p. 90), 10.9 (p. 90), 10.14 (p. 94), 10.15 (p. 95), 10.16 (p. 96).

Leçon 219 — Extrema

Exercices : 10.17 (p. 97), 10.18 (p. 100).

Leçon 223 — Suites numériques

Exercices : 14.1 (p. 133), 14.2 (p. 134), 14.3 (p. 134), 14.4 (p. 135), 14.5 (p. 136), 14.6 (p. 137), 14.7 (p. 139).

Leçon 226 — Suites récurrentes

Exercices : 14.3 (p. 134), 14.4 (p. 135), 14.5 (p. 136), 11.7 (p. 108).

Leçon 236 — Calcul d'intégrales

Exercices : 16.1 (p. 146), 16.2 (p. 147), 16.3 (p. 149), 16.4 (p. 150).

Leçon 239 — Intégrales à paramètre

Exercices : 16.5 (p. 150), 16.1 (p. 146), 16.2 (p. 147), 16.6 (p. 151), 16.7 (p. 152), 16.8 (p. 152).

Leçon 241 — Suites et séries de fonctions

Exercices : 15.3 (p. 143), 15.1 (p. 141), 15.4 (p. 144), 8.4 (p. 79), 15.5 (p. 144), 15.6 (p. 145), 15.7 (p. 145), 15.2 (p. 142).

Leçon 243 — Séries entières

Exercices : 18.5 (p. 163), 18.6 (p. 164), 18.7 (p. 164), 18.8 (p. 167).

Leçon 245 — Analyse complexe

Exercices : 18.9 (p. 168), 18.10 (p. 168), 18.11 (p. 169), 18.12 (p. 169), 18.5 (p. 163), 18.13 (p. 169), 12.5 (p. 125).

Leçon 250 — Transformation de Fourier

Exercices : 16.2 (p. 147), 19.2 (p. 173), 12.4 (p. 123), 19.1 (p. 171).

Leçon 253 — Convexité

Exercices : 8.3 (p. 77), 17.1 (p. 154), 17.2 (p. 155), 17.3 (p. 156).

Leçon 265 — Fonctions usuelles et spéciales

Exercices : 16.7 (p. 152), 12.3 (p. 122), 16.1 (p. 146), 19.2 (p. 173).

Leçon 267 — Courbes

Exercices : 10.9 (p. 90), 15.7 (p. 145).

Bibliographie

- [1] A. Avez, *Calcul différentiel*, Masson, 1997.
- [2] S. Benzoni-Gavage, *Calcul différentiel et équations différentielles. Cours et exercices corrigés*, 2^e édition, Dunod, 2021.
- [3] H. Cartan, *Cours de calcul différentiel avec exercices*, Hermann, 2007.
- [4] F. Rideau, *Exercices de calcul différentiel*, Hermann, 1997.
- [5] F. Rouvière, *Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation*, 4^e édition revue et corrigée, Cassini, 2014.
- [6] X. Gourdon, *Les maths en tête. Analyse*, 3^e édition, Ellipses, 2020.
- [7] F. Hirsch et G. Lacombe, *Éléments d'analyse fonctionnelle. Cours et exercices*, Dunod, 2009.
- [8] H. Brezis, *Analyse fonctionnelle. Théorie et applications*, Dunod, 2020.
- [9] J.-P. Demailly, *Analyse numérique et équations différentielles*, 4^e édition, EDP Sciences, 2016.
- [10] S. Francinou, H. Gianella et S. Nicolas, *Exercices de mathématiques. Oraux X-ENS. Algèbre 3*, 2^e édition, Cassini, 2013.
- [11] S. Francinou, H. Gianella et S. Nicolas, *Exercices de mathématiques. Oraux X-ENS. Analyse 4*, Cassini, 2012.
- [12] J. Bernis et L. Bernis, *Analyse pour l'agrégation de mathématiques. 40 développements*, 2^e édition, Ellipses, 2024.
- [13] H. Queffélec et C. Zuily, *Analyse pour l'agrégation*, 5^e édition, Dunod, 2020.
- [14] W. Rudin, *Analyse réelle et complexe*, 3^e édition, Dunod, 2020.
- [15] W. Rudin, *Analyse fonctionnelle*, Édiscience International, 2000.
- [16] W. Rudin, *Principes d'analyse mathématique. Cours et exercices*, Dunod, 2021.
- [17] P. Billingsley, *Convergence of Probability Measures*, 2^e édition, Wiley, 1999.
- [18] J. Lindenstrauss et L. Tzafriri, *Classical Banach Spaces I and II. Sequence Spaces and Function Spaces*, édition réunie, Springer, 1996 ; publication numérique, 2022. Partie I : *Sequence Spaces*.