Analyse réelle et complexe

François Golse,Yves Laszlo,Claude Viterbo

Version provisoire du 22 mars 2010

Chapitre I

Introduction

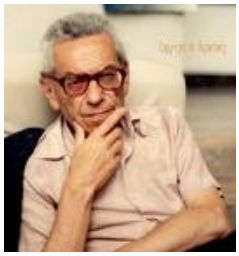


FIGURE I.1 – P. Erdös

« Pourquoi les nombres sont-ils beaux ? Cela revient à se demander pourquoi la neuvième symphonie de Beethoven est belle. Si vous ne voyez pas pourquoi, personne ne pourra vous l'expliquer. Je sais que les nombres sont beaux. S'ils ne sont pas beaux, rien ne l'est. », (P. Hoffman et M. Mashaal, *Erdös. L'homme qui n'aimait que les nombres 1913/1996*, Pour la Science Paris (2000)).



FIGURE I.2 – P. Dirac

« *The steady progress of physics requires for its theoretical formulation a mathematics which get continually more advanced. This is only natural and to be expected. What however was not expected by the scientific workers of the last century was the particular form that the line of advancement of mathematics would take, namely it was expected that mathematics would get more and more complicated, but would rest on a permanent basis of axioms and definitions, while actually the modern physical developments have required a mathematics that continually shifts its foundation and gets more abstract. Non-euclidean geometry and noncommutative algebra, which were at one time were considered to be purely fictions of the mind and pastimes of logical thinkers, have now been found to be very necessary for the description of general facts of the physical world. It seems likely that this process of increasing abstraction will continue in the future and the advance in physics is to be associated with continual modification and generalisation of the axioms at the base of mathematics rather than with a logical development of any one mathematical scheme on a fixed foundation.* », (P. Dirac, Proc. Roy. Soc. 133 (1931)).



FIGURE I.3 – R. Feynman

« *To those who do not know mathematics it is difficult to get across a real feeling as to the beauty, the deepest beauty, of nature ... If you want to learn about nature, to appreciate nature, it is necessary to understand the language that she speaks.* », (R. Feynman, *The Character of Physical Law*, MIT Press (1967))

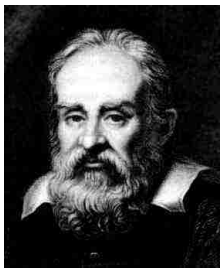


FIGURE I.4 – G. Galilee

« La philosophie est écrite dans cet immense livre que nous tenons toujours ouvert sous nos yeux, je veux dire l'univers. Nous ne pouvons pas le comprendre si nous n'avons pas cherché à l'avance à en apprendre la langue, et à connaître les caractères au moyen desquels il a été écrit. Or il est écrit en langue mathématique, et ses caractères sont des triangles, des cercles et des figures géométriques, sans lesquels il serait impossible à tout homme d'en saisir le sens. Sans eux, nous ne ferons que tournoyer en vain dans un labyrinthe obscur. », (Galileo Galilei, *Il Saggiatore*, Opere, Ricciardi, Milano).

Au travers de ces quelques citations, on s'aperçoit de la multiplicité de points de vue qu'on peut avoir sur l'activité mathématique. Toutes ont droit de cité, ont leurs lettres de noblesse.

Notre parti pris a été dans ce cours de tenter de concilier ces façons de voir en abordant trois thèmes importants, et assez intimement liés, les théories de l'intégration de Lebesgue, des espaces de Hilbert et des fonctions holomorphes d'une variable. Prenant comme point de départ les compétences de mathématiques (et de physique) acquises dans les années précédentes (plus précisément l'*infimum* des programmes de classes préparatoires de MP et de PC), on donne ici les bases de ces théories qui permettront, d'un point de vue pratique, d'aborder avec les armes nécessaires les cours de seconde année et de suivre avec profit ceux de première année (liens Hilbert, Fourier et mécanique quantique, intégration et probaibilités...).

Guidés par ces multiplicités de points de vue, nous avons à la fois montré la puissance et, espérons-le, l'harmonie des constructions [1]. Nous avons donné donc des exemples à la fois purement mathématiques (fonctions holomorphes et topologie du plan, espaces de Hilbert et théorème de Münts-Szasz), mais aussi d'applications des mathématiques à d'autres sciences (transformée de Fourier et diffraction ou encore inégalités de Heisenberg, espaces de Hilbert et échantillonage en théorie du signal, fonctions holomorphes et relations de dispersion).

1. constructions ou découvertes, question de point de vue là encore : « La différence de vues qui existe entre Platon et Aristote au sujet de l'existence des êtres mathématiques pourrait caractériser à elle seule l'essentiel de ce que nous avons à dire. Pour Platon le monde des mathématiques est un monde indépendant, portant en lui-même ses propres lois et supérieur au physique dans sa façon d'être. L'existence des êtres mathématiques est, de ce fait, indépendante de la pensée humaine comme, en général, de toute activité extérieure. Pour Aristote, au contraire, il n'y a pas de monde mathématique en soi ; si l'on en parle, c'est en tant qu'idées abstraites de l'activité humaine, à savoir des constructions des mathématiciens créateurs. » (A. Fraenkel, « Sur la notion d'existence dans les mathématiques », L'Enseignement mathématique, 34 (1935), p. 18-19).

Nous avons cherché à éviter la technicité autant que faire se peut, ce qui nous a conduit à ne pas chercher les énoncés optimaux, encore moins l'exhaustivité. Pour faciliter l'étude du polycopié, on a regroupé à la fin de chaque chapitre les résultats essentiels. Les lecteurs intéressés pour aller plus loin pourront se reporter à la bibliographie ou, mieux, frapper à la porte des bureaux du Centre de mathématiques Laurent Schwartz.

Le but de ce cours aura été atteint si le lecteur en ressort avec la perception de ce que les mathématiques sont à la fois une science riche et autonome, au même titre que la physique,la chimie ou la biologie par exemple et aussi un outil de compréhension du monde « réel »[2].

2. Notion qui mériterait d'être discutée, par exemple à la lumière des paradoxes de Banach-Tarski ou bien du chat de Scrödinger ou bien par exemple des points de vue de Paul Valéry (« Ma main se sent touchée aussi bien qu'elle touche. Réel veut dire cela, rien de plus. », P. Valéry, *Mon Faust*, Œuvres, Pléiade ou de Platon (« Ces hommes sont enchaînés dans une sombre caverne ; ils ne voient ni la véritable lumière, ni la source d'où elle jaillit, ni les choses réelles, mais seulement une faible lueur diffuse dans la caverne et les ombres des choses réelles qui passent devant un grand feu, derrière les hommes : pourtant ils se figurent que les ombres sont des réalités, et, s'ils connaissent l'ordre de succession de ces ombres, ils croient posséder la véritable sagesse. » Platon, *La République*, Flammarion (2002)).

Chapitre II

Compléments de topologie



Henri Lebesgue



Georg Cantor



Émile Borel

1 Introduction

Nous avons adopté dans ce texte le cadre des espaces métriques, introduits au début du siècle dernier par Maurice Fréchet, qui dégagea dans ce contexte l'importance des notions de complétude et de compacité. Ce cadre semblera trop restrictif aux amateurs d'espaces topologiques, filtres et ultrafiltres, et trop général à ceux qui sont habitués aux espaces normés. À l'intention de ces derniers, précisons qu'une bonne partie du cours peut être lu en remplaçant espaces métriques par sous-ensembles d'un espace vectoriel normé. Le lecteur est invité en première lecture à passer ce chapitre (exceptées les sections sur le théorème de Dini (4.2 page 16 et 4.6 page 20) et à ne s'y reporter que si le besoin s'en fait sentir.



Ulisse Dini



Maurice Fréchet

2 Topologie des espaces métriques

2.1 Notion d'espace métrique

Soit X un ensemble.

Définition 2.1.1. Une distance d sur X est une application $\mathrm{d} : X \times X \to \mathbf{R}^+$ vérifiant :
- $\mathrm{d}(x, y) = 0$ si et seulement si $x = y$.
- $\forall x, y \in X,\ \mathrm{d}(x, y) = \mathrm{d}(y, x)$.
- $\forall x, y, z \in X,\ d(x, z) \leq d(x, y) + d(y, z)$ (inégalité triangulaire)

On dit que (X, d) (ou X lorsque le contexte est clair) est un espace métrique.

Exemple 2.1.2. L'exemple fondamental est le celui où X est un sous-ensemble d'un espace vectoriel normé avec pour distance $\mathrm{d}(x, y) = ||x - y||$. Évidemment, un sous-ensemble d'un espace métrique, muni de la distance induite, est un espace métrique.

Le choix des distances ou des normes est important suivant le contexte. Par exemple sur $C^0([0, 1], \mathbf{R})$ espace des vitesses de sa voiture, l'automobiliste choisira

$$||\text{vitesse}|| = \sup(\text{vitesse}) \text{ ou } \sqrt{\int (\text{vitesse})^2},\ \text{vitesse} \in C^0([0, T], \mathbf{R})$$

suivant qu'il est préoccupé par son nombre de points sur son permis ou sa consommation.

On note $B(x, r) = \{y \in X | \mathrm{d}(y, x) < r\}$ la boule ouverte de rayon r centrée en x. L'allure des boules change considérablement si on change les normes. On a dessiné ici les (frontières des) boules pour les distances associées aux normes 1 (en violet) 2 (en bleu) et sup (en vert) sur $\mathbf{R}^2$.

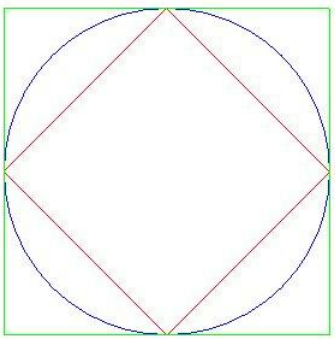


FIGURE II.1 – Boules associées à diverses distances

Exercice 2.1.3. Montrer que $C^0(\mathbf{R},\mathbf{R})$ muni de $\mathrm{d}(f,g) = \min(1, \sup|f-g|)$ est un espace métrique.

2.2 Topologie

Rappelons quelques définitions bien connues dans le cadre des espaces normés.

Définition 2.2.1. Soit (X,d) un espace métrique.

(1) On dit qu'une suite $(x_n)_{n\geq 1}$ est convergente de limite x, si $\lim_{n\to\infty}\mathrm{d}(x_n,x)=0$.

(2) Un sous-ensemble U de X est ouvert si pour tout x dans U, il existe une boule centrée en x, de rayon strictement positif contenue dans U.

(3) On appelle topologie associée à la métrique d l'ensemble des ouverts de X.

(4) Un sous-ensemble F de X est fermé si et seulement si $X-F$ est ouvert

(5) L'adhérence de $Y\subset X$ est définie par

$$\overline{Y} = \bigcap_{\substack{F\text{ fermé}\\ Y\subset F}} F.$$

(6) On appelle voisinage d'un point ou d'un ensemble tout ensemble contenant un ouvert contenant ce point ou cet ensemble.

(7) Une application $f:(X,\mathrm{d})\longrightarrow(X',\mathrm{d}')$ est continue si et seulement si pour tout x_0 et ε strictement positif, il existe α strictement positif tel que

$$\mathrm{d}(x_0,x)<\alpha \Longrightarrow \mathrm{d}'(f(x_0),f(x))\leq\varepsilon.$$

(8) Le support d'une fonction à valeurs réelles ou complexes est le fermé, adhérence des points où elle ne s'annule pas.

Les propriétés suivantes sont classiques.

Propriétés 2.2.2.
- La limite d'une suite convergente est unique (caractère *séparé* des métriques).
- $F\subset X$ est fermé si la limite de toute suite de points de F convergeant dans X est dans F.
- Une réunion d'ouverts (respectivement intersection de fermés) est un ouvert, (respectivement fermé) et une intersection finie d'ouverts (respectivement réunion finie de fermés) est un ouvert (respectivement fermé).
- L'adhérence dans X d'une partie $Y\subset X$ est le plus petit fermé de X contenant Y. C'est l'ensemble des limites de suites de Y convergeant dans X.
- $f: X\to Y$ est continue si et seulement si si pour toute suite $(x_n)_{n\geq 1}$ convergente, $\lim_{n\to\infty} f(x_n)_{n\geq 1} = f(\lim_{n\to\infty} x_n)$ ou encore si si pour tout ouvert (respectivement fermé) Y de X, $f^{-1}(Y)$ est un ouvert (respectivement fermé).

Exemple 2.2.3. Soient $(X,\mathrm{d}),(X',\mathrm{d}')$ des métriques.

(1) Le produit $X\times X'$ muni de la « distance somme » δ définie par

$$\delta((x_1,x_1'),(x_0,x_0')) = \mathrm{d}(x_1,x_0)+\mathrm{d}'(x_1',x_0')$$

est un espace métrique.

(2) L'application $\mathrm{d} : (X \times X, \delta) \to \mathbf{R}$ est continue, puisque d'après l'inégalité triangulaire

$$|\mathrm{d}(x_1, x_2) - \mathrm{d}(y_1, y_2)| \leq \mathrm{d}(x_1, y_1) + \mathrm{d}(x_2, y_2) = \delta((x_1, x_2), (y_1, y_2)).$$

L'application $\mathrm{d}_{x_0} : X \to \mathbf{R}$ définie par $\mathrm{d}_{x_0}(x) = \mathrm{d}(x, x_0)$ est aussi continue, comme restriction de la fonction continue δ à $X \times \{x_0\}$.

(3) Tout ensemble défini par des inégalités strictes sur des fonctions continues est ouvert. Tout ensemble défini par des inégalités larges sur des fonctions continues est fermé. La distance étant continue, la boule ouverte est bien ouverte, et la boule fermée $\bar{B}(x, r) = \{y \in X \mid \mathrm{d}(y, x) \leq r\}$ est bien fermée (image réciproque par d d'un fermé).

Remarque 2.2.4. **Attention :** l'adhérence de la boule ouverte de rayon r n'est pas nécessairement la boule fermée de rayon r. Par exemple sur $\mathbf{Z}$ muni de la distance $\mathrm{d}(x, y) = |x - y|$ la boule ouverte de rayon 1 centrée en 0 se réduit à $\{0\}$, et donc son adhérence aussi, mais la boule fermée de centre 0 et de rayon 1 est l'ensemble $\{-1, 0, 1\}$.

Ce qu'il faut retenir c'est que lorsqu'on parle d'espaces métriques :

- la principale propriété est l'inégalité triangulaire,

- les propriétés d'être ouvert, fermé, dans l'adhérence d'un ensemble, peuvent se vérifier en prenant des suites.

Remarque 2.2.5. Il est légitime de se demander pourquoi on s'encombre de la notion de métrique, alors que la plupart du temps, on n'utilise que des parties d'espaces vectoriels normés. Les espaces métriques constituent un cadre à la fois souple et commode, et il n'est pas plus difficile de manier une métrique qu'une norme.

Exercice 2.2.6. Soit Y une partie d'un espace métrique X. Montrer que les ouverts de Y (pour la distance induite) sont les traces $U \bigcap X$ des ouverts U de X. La topologie correspondante sur Y s'appelle la topologie induite. Montrer que si $V \subset Y$ est ouvert dans Y, en général il n'est pas ouvert dans X. Montrer qu'en revanche si on suppose que Y est ouvert dans X, alors V ouvert dans Y si et seulement si V ouvert dans X. Énoncer et prouver une propriété analogue pour les fermés.

Exercice 2.2.7. Soit E un $\mathbf{R}$-espace vectoriel normé. Montrer que deux normes équivalentes définissent la même topologie. En particulier, si E est de dimension finie, le choix des normes n'a pas d'influence sur la topologie, dite topologie de la norme (cf. 4.4.1 page 18). Montrer qu'il existe une distance sur $\mathbf{R}$ telle que la topologie associée soit la topologie discrète (toutes les parties sont ouvertes). En déduire qu'il existe des topologies métriques sur $\mathbf{R}$ différentes de la topologie de la norme.

Exercice 2.2.8. Trouver deux normes sur $C^0([0, 1])$ définissant des topologies différentes.

2.3 Densité

Définition 2.3.1. Soit X un espace métrique.

- On dit qu'une partie Y de X est dense dans X si $\bar{Y} = X$.
- Une suite extraite de $(x_n)_{n\geq 1}$ est une suite de la forme $x_{\varphi(n)}$ avec $\varphi : \mathbf{N} \to \mathbf{N}$ strictement croissante (on note $\varphi \nearrow\nearrow$).
- Un point $x \in X$ est une valeur d'adhérence de la suite $(x_n)_{n\geq 1}$ si elle est limite d'une suite extraite.

On a immédiatement la caractérisation commode des valeurs d'adhérence de suite.

Propriétés 2.3.2. $x \in X$ est une valeur d'adhérence de $(x_n)_{n\geq 1}$ si et seulement si

$$\forall \epsilon > 0, \mathrm{card}(\{k | x_k \in B(x,\epsilon)\}) = \infty.$$

De même, Y dense dans X ssi

$$\forall x \in X, \forall \epsilon > 0,\ B(x,\epsilon) \cap Y \neq \varnothing.$$

Le lien entre valeurs d'adhérence et adhérence est donné par la proposition suivante.

Proposition 2.3.3. *L'adhérence $\bar{Y}$ de $Y \subset X$ est l'ensemble des valeurs d'adhérence dans X des suites à valeurs dans Y.*

Démonstration. Exercice □

L'ensemble des valeurs d'adhérences d'une suite est un fermé. Plus précisément, on a

Proposition 2.3.4. *L'ensemble des valeurs d'adhérences de la suite $(x_n)_{n\geq 1}$ est*

$$\bigcap_n \overline{\{x_k, |k \geq n\}}.$$

Démonstration. Soit x une valeur d'adhérence. D'après la propriété 2.3.2, on sait que $\mathrm{card}(\{k|x_k \in B(x,\epsilon)\}) = \infty$ ce qui assure que pour tout n et tout $\epsilon > 0$, la boule $B(x,\epsilon)$ rencontre $\{x_k, |k \geq n\}$. D'après *loc. cit.*, ceci assure $y \in \overline{\{x_k, |k \geq n\}}$.

De façon analogue, soit $x \in \bigcap_n \overline{\{x_k, |k \geq n\}}$ et $\epsilon > 0$. Comme $B(x,\epsilon)$ rencontre chacun des $X_n := \{x_k, |k \geq n\}$, on construit (grâce à la décroissance des X_n) par récurrence une fonction strictement croissante φ telle que $x_{\varphi(n)} \in \{x_k, |k \geq n\} \cap B(x,\epsilon)$ ce qui assure que le cardinal de $\{k|x_k \in B(x,\epsilon)\}$ est bien infini. □

Exercice 2.3.5. Montrer que $(x_n)_{n\geq 1}$ converge si et seulement si elle a une unique valeur d'adhérence.

Exercice 2.3.6. Montrer que $\mathbf{Q}$ est dense dans $\mathbf{R}$. Montrer que $\mathbf{R}[t]$ est dense dans $(C^0([0,1]), \sup_{[0,1]})$ mais pas dans $C^0(\mathbf{R})$ muni de $\mathrm{d}(f,g) = \min(1, \sup|f-g|)$

.

Exercice 2.3.7. Pour $Y \subset X$, on pose $\mathrm{d}(x,Y) = \inf_{y\in Y} \mathrm{d}(x,y)$. Montrer l'inégalité $|\mathrm{d}(x,Y) - \mathrm{d}(x',Y)| \leq \mathrm{d}(x,x')$ et en déduire la continuité de $x \to \mathrm{d}(x,Y)$ Montrer que $\mathrm{d}(x,Y) = 0$ si et seulement si $x \in \bar{Y}$. Montrer que les fermés dans les espaces métriques sont exactement les ensembles de zéros de fonction continues réelles.

3 Complétude

3.1 Espaces complets

Il agit d'une notion de première importance en analyse.

Définition 3.1.1. Une suite $(x_n)_{n\geq 1}$ d'un espace métrique (X,d) est de Cauchy si

$$\forall \varepsilon > 0 \ \exists N, \quad \forall n,m \geq N \ ; \quad d(x_n,x_m) < \varepsilon$$

Une suite convergente est de Cauchy. Mais bien qu'une suite de Cauchy soit toujours bornée, elle n'est en général pas convergente. Par exemple, si $X = \mathbf{Q}$ muni de la distance usuelle, la densité de $\mathbf{Q}$ dans $\mathbf{R}$ prouve qu'il existe des suites de rationnels r_n convergeant dans $\mathbf{R}$ vers $\sqrt{2}$. En particulier, la suite r_n est une suite de Cauchy de X, mais n'a pas de limite dans $\mathbf{Q}$ car $\sqrt{2} \notin \mathbf{Q}$.

Lemme 3.1.2. *Une suite de Cauchy $(x_n)_{n\geq 1}$ qui possède une valeur d'adhérence est convergente.*

Démonstration. Soit ℓ une valeur d'adhérence et $\epsilon > 0$. Choisissons M, N tels que

$$\mathrm{d}(x_M,\ell) < \varepsilon/2 \text{ et } \ \forall n,m \geq N \ d(x_n,x_m) < \varepsilon/2.$$

Alors, pour $n \geq \sup(N,M)$, on a $\mathrm{d}(x_n,\ell) < \epsilon$. □

Définition 3.1.3 (Espace complet). Un espace métrique (X,d) est complet si et seulement si toute suite de Cauchy est convergente.

On vérifie sans mal qu'une partie F d'un espace métrique complet X est complète si et seulement si elle est fermée (appliquer la définition de la fermeture utilisant les suites).

Lemme 3.1.4. *Le produit de deux métriques complets $(X_i,\mathrm{d}_i), i = 1,2$ muni de de la distance somme ((1) page 9) est complet.*

Démonstration. Les relations

$$\sup(\mathrm{d}_1((x_1,y_1),\mathrm{d}_2(x_2,y_2)) \leq \mathrm{d}((x_1,x_2),(y_1,y_2)) \leq 2\sup(\mathrm{d}_1((x_1,y_1),\mathrm{d}_2(x_2,y_2)) \tag{3.1.a}$$

prouvent qu'une suite de $X_1 \times X_2$ $((x_{1,n},x_{2,n}))$ est de Cauchy (respectivement converge) si et seulement si les deux suites de X_1, X_2 $(x_{1,n}), (x_{2,n})$ sont de Cauchy (respectivement convergent). Le lemme suit. □

Remarque 3.1.5. **Attention :** la complétude est une propriété qui dépend de la métrique, et pas seulement de la topologie sur X. Par exemple, $\mathbf{R}$ muni de la métrique $\mathrm{d}(x,y) = |e^{-x} - e^{-y}|$ n'est pas complet. En effet la suite $x_n = n$ est de Cauchy pour cette métrique, mais ne converge pas (exercice). Pourtant un ensemble est ouvert pour cette métrique si et seulement s'il l'est pour la métrique usuelle : chaque boule pour l'une de ces deux métriques contient une boule de l'autre.

Ceci suggère la définition suivante

Définition 3.1.6. Deux métriques sur X sont équivalentes si il existe deux réels k, K strictement positifs tels que

$$k\mathrm{d}_1 \leq \mathrm{d}_2 \leq K\mathrm{d}_1$$

Deux métriques équivalentes $\mathrm{d}_1, \mathrm{d}_2$ ont mêmes suites de Cauchy, et sont donc $(X,d_1$ et (X,d_2) simultanément complets ou non complets.

Proposition 3.1.7. *Alors* $\mathbf{R}^n$ *muni de la norme somme* $||(x_i)|| = \sum_i |x_i|$ *est complet.*

Démonstration. D'après (3.1.4 page précédente), il suffit de montrer que $\mathbf{R}$ muni de la valeur absolue est complet et donc de montrer qu'une suite de Cauchy réelle $(x_n)_{n\geq 1}$ a une valeur d'adhérence (3.1.2 page ci-contre). Or on sait (cf. [7, pages 298, 341]), puisque $(x_n)_{n\geq 1}$ est bornée, quela méthode de dichotomie permet de construire deux suites extraites $x_{\varphi(n)} \leq x_{\psi(n)}$ adjacentes, donc convergentes vers $\ell \in \mathbf{R}$ qui est par définition une valeur d'adhérence. □

On verra plus bas (4.4.1 page 18) qu'en dimension finie toutes les normes sont équivalentes. D'après la remarque précédente on a donc

Théorème 3.1.8. *Un* $\mathbf{R}$*-espace vectoriel normé de dimension finie est complet.*

Un espace vectoriel normé complet s'appelle espace de Banach .

Remarque 3.1.9. Notons qu'on peut remplacer $\mathbf{R}$ par $\mathbf{C}$ dans l'énoncé précédent (exercice : pourquoi).

Donnons quelques exemples.

(1) $\mathbf{Q}$ n'est pas complet pas plus que $\mathbf{R}^*$.

(2) Tout espace vectoriel normé de dimension finie sur $\mathbf{R}$ ou $\mathbf{C}$ est complet.

(3) On pose
$$\mathfrak{b}(\mathbf{N}, \mathbf{R}^n) = \{x = (x_k)_{k\geq 0} \mid \sup |x_k| < \infty\}$$
muni de la norme naturelle
$$\|x\| = \sup_{k\in\mathbf{N}} |x_k|$$
Vérifions que $\mathfrak{b}$ est un espace de Banach, c'est-à-dire que $(\mathfrak{b}, \| \bullet \|)$ est complet.
En effet, si $x^\nu = (x_k^\nu)_{k\in\mathbf{N}}$ est une suite de Cauchy de $\mathfrak{b}$, on aura
$$\forall \varepsilon > 0, \quad \exists \nu_0 \quad \text{tel que pour } \mu, \nu \geq \nu_0 \qquad \sup_k |x_k^\nu - x_k^\mu| \leq \varepsilon$$
Mais cela entraîne que pour chaque k la suite $\nu \to x_k^\nu$ est de Cauchy, et donc par complétude de $\mathbf{R}^n$, il existe z_k tel que $\lim_\nu x_k^\nu = z_k$. Si $z = (z_k)_{k\in\mathbf{N}}$, on a $\lim_\nu x^\nu = z$. En effet pour $\nu, \mu \geq \nu_0$ et tout k, on a $|x_k^\nu - x_k^\mu| \leq \varepsilon$ et donc par passage à la limite en μ, on aura , $|x_k^\nu - z_k| \leq \varepsilon$. Donc pour $\nu \geq \nu_0$, on aura $\|x^\nu - z\| \leq \varepsilon$, et donc x^ν converge vers z.

(4) De même, l'espace $\mathfrak{b}_0(\mathbf{N}, \mathbf{R}^n) = \{(x_k)_{k\in\mathbf{N}} \mid \lim_{k\to\infty} x_k = 0\}$ est un sous-espace vectoriel fermé de $\mathfrak{b}(\mathbf{N}, \mathbf{R}^n)$. C'est donc aussi un espace de Banach.

(5) On verra dans le chapitre sur les espaces de Hilbert (exemple (1) page 147), que l'espace $\ell_2(\mathbf{N}, \mathbf{R})$ des suites réelles (x_i) telle que $\sum_i x_i^2 < \infty$ muni de la norme définie par
$$||(x_i)||^2 = \sum_i x_i^2$$
est complet. C'est l'exemple standard d'espace de Hilbert de dimension infinie.

Exercice 3.1.10. Soient X et Y des espaces métriques avec Y complet. Montrer que $C^0(X,Y)$, ensemble des fonctions continues, muni de la distance d_0 définie par

$$\mathrm{d}_0(f,g) = \min\{1, \sup_{x\in X} d(f(x),g(x))\}$$

est un espace métrique complet. En déduire que l'espace des suites réelles,

$$\mathcal{F}(\mathbf{N},\mathbf{R}^n) = \{x = (x_k)_{k\in\mathbf{N}} \mid x_k \in \mathbf{R}\}$$

muni de la métrique

$$\mathrm{d}(x,y) = \min\{1, \sup_{k\in\mathbf{N}} |x_k - y_k|\}$$

est complet.

Exercice 3.1.11 (Critère de complétude des espaces vectoriels normés). Un espace vectoriel normé est complet si et seulement si toute série absolument convergente converge.

Exercice 3.1.12. $C^0([-1,1],\mathbf{R})$ muni de $N_\infty(f) = \sup|f|$ est complet alors que muni de $N_1(f) = \sqrt{\int f^2}$ il ne l'est pas (cf. 1.0.1 page 111). Ceci montre l'importance des normes en dimension infinie.

3.2 Le théorème des fermés emboîtés

Définissons le diamètre d'une partie X d'un métrique par

$$\mathrm{diam}(X) = \sup_{x,y\in X} \mathrm{d}(x,y) \in \bar{\mathbf{R}}^+ = [0,\infty].$$

Théorème 3.2.1 (Théorème des fermés emboîtés)**.** *L'intersection d'une suite décroissantes de fermés non vides F_n d'un espace complet dont le diamètre tend vers* 0 *est réduite à un point.*

Démonstration. Unicité : si

$$x, y \in \bigcap F_n, \text{ on a } \mathrm{diam}(F_n) \geq \mathrm{d}(x,y)$$

et donc $\mathrm{d}(x,y) = 0$.

Existence : on choisit $x_n \in F_n$. Pour tout $p \leq q$, on a

$$x_p, x_q \in F_p \text{ donc } \mathrm{d}(x_p,x_q) \leq \mathrm{diam}(F_p).$$

Comme $\lim \mathrm{diam}(F_p) = 0$, la suite $(x_n)_{n\geq 1}$ est de Cauchy et donc converge vers $x \in X$. Fixons alors p. On a $x = \lim_n x_{n+p}$ de sorte que x est limite d'une suite d'éléments de $F_{n+p} \subset F_p$ ce qui assure $x \in F_p$ car F_p est fermé. □

4 Compacité

4.1 Définitions

La problématique consiste en une question de complexité. Étant donné une famille finie d'ouverts recouvrant un espace métrique X, peut-on en extraire un sous-recouvrement fini ?

Notons que cette propriété n'est pas vraie en général, même pour un ensemble borné. Par exemple le recouvrement de $]0,1[$ par $\bigcup_{n>0}]1/n, 1-1/n[$ n'a pas de sous-recouvrement fini.

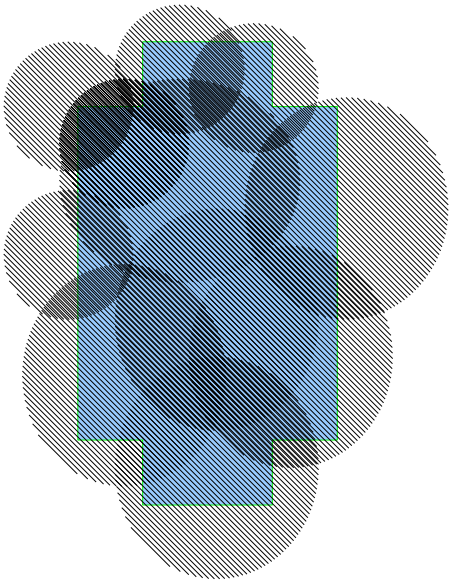


FIGURE II.2 – Un recouvrement fini

Définition 4.1.1. Un espace métrique X est compact si l'une des deux propriétés équivalentes suivantes est vérifiée

1) Si de tout recouvrement ouvert on peut extraire un recouvrement fini
1') Si toute famille de fermés d'intersection vide admet une sous famille finie d'intersection vide.

Pour voir l'équivalence entre 1) et 1'), il suffit de passer au complémentaire.

La proposition suivante découle immédiatement de 1').

Proposition 4.1.2.
- *Une partie d'un compact est compacte si et seulement si elle est fermée.*
- *Si $X \subset Y$ et X compact alors X fermé dans Y.*
- *Une intersection décroissante de compacts non vides est non vide.*
- *L'image continue d'un compact est compacte.*

4.2 Théorème de Dini

Proposition 4.2.1 (Théorème de Dini)**.** *Soit (X, d) un espace métrique. Toute suite décroissante $(f_n)_{n\geq 1}$ de fonctions continues à support compact convergeant en tout point vers 0 converge uniformément vers 0 :* i.e.$\lim_n \sup_X f_n(x) = 0$.

Démonstration. Notons que par hypothèse $f_n \geq 0$ et donc pour montrer la convergence uniforme il suffit de montrer que les $\sup(f_n)$ convergent vers 0. Soit $\epsilon > 0$ et

$$F_n = \{x \in F | f_n(x) \geq \epsilon\}.$$

La suite F_n décroît comme les f_n et F_n est compact puisque fermé dans un compact (le support de f_0 par exemple). Comme $\lim f_n(x) = 0$ pour tout x, l'intersection des F_n est vide. Or, une intersection décroissante de compacts non vides est non vide (4.1.2 page précédente) de sorte qu'il existe N tel que $F_N = \varnothing$. La décroissance des F_n assure $\forall n \geq N\ F_n = \varnothing$ et donc $0 \leq \sup f_n \leq \epsilon$. □

4.3 Caractérisation séquentielle

Théorème 4.3.1 (Théorème de Borel-Lebesgue)**.** *Un espace métrique X est compact si et seulement si toute suite de X a une suite extraite convergente. En particulier un espace compact est complet (cf. 3.1.2 page 12).*

Démonstration. Supposons X compact. On doit montrer que l'intersection décroissante des fermés

$$\bigcap_{n\geq 0} \overline{\{x_k, k \geq n\}}$$

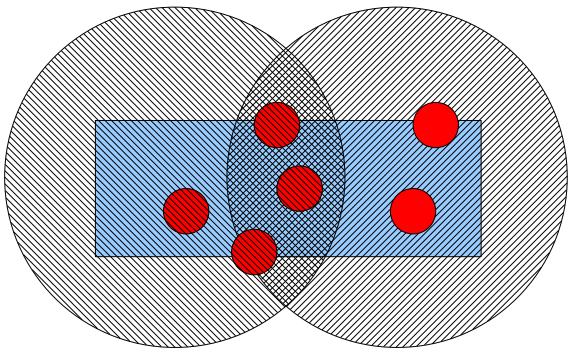
est non vide. Sinon, on aurait

$$\exists N | \bigcap_{n\leq N} \overline{\{x_k, k \geq n\}} = \overline{\{x_k, k \geq N\}} = \varnothing,$$

ce qui contredirait la compacité de X.

Pour montrer la réciproque, on utilise le lemme clef :

Lemme 4.3.2 (Nombre de Lebesgue)**.** *Soit U_i recouvrement ouvert de X. Il existe $\delta > 0$ tel que $\forall x \in X \exists i,\ B(x, \delta) \subset U_i$.*



Boules de Lebesgue d'un recouvrement ouvert d'un compact

Démonstration. Supposons le théorème en défaut. On construit alors une suite $(x_n)_{n\geq 1}$ telle que pour tout n, $B(x_n, 1/n) \not\subset U_i$. Choisissons une valeur d'adhérence x de $(x_n)_{n\geq 1}$ de sorte qu'on peut écrire $x = \lim x_{\varphi(n)}$ où φ est strictement croissante. Comme les U_i recouvrent X, on peut choisir j tel que $x \in U_j$. Comme U_j est ouvert, on peut choisir N tel que $B(x, 1/N) \subset U_j$.

Comme

$$\lim x_{\varphi(n)} = x \text{ et } \lim 1/\varphi(n) = 0, \text{ on a}$$

$$B(x_{\varphi(n)}, 1/\varphi(n)) \subset B(x, 1/N) \subset U_j \text{ si } n \text{ est suffisamment grand}$$

ce qui contredit la définition de x_n. □

On suppose que toute union finie des U_i est $\neq X$. Soit δ un nombre de Lebesgue. Par récurrence, on construit alors une suite $(x_n)_{n\geq 1}$ telle que

$$B(x_n, \delta) \subset U_n \text{ et } x_n \notin U_1, \cdots, U_{n-1}.$$

Puisque pour $m < n$ on a $x_n \notin B(x_m, \delta) \subset U_m$, on en déduit que $\partial(x_m, x_n) \geq \delta$ pour $m \neq n$ et donc x_n n'aurait pas de suite extraite convergente. □

Remarque 4.3.3. Si X est fini, c'est une conséquence du principe des tiroirs qui dit qu'une suite à un nombre fini de valeurs prend une infinité de fois la même valeur !

Corollaire 4.3.4. *Le produit de deux espaces métriques compacts X, Y est encore compact.*

Démonstration. Soit (x_n, y_n) une suite de $X \times Y$. On extrait $(x_{\varphi(n)})$ de x_n convergeant vers x puis $(y_{\varphi(\psi(n))})$ de $y_{\varphi(n)}$ convergeant vers y et on observe grâce à (3.1.a page 12) que $(x_{\varphi(\psi(n))}, y_{\varphi(\psi(n))})$ converge vers (x, y) qui est donc une valeur d'adhérence de (x_n, y_n). □

Corollaire 4.3.5 (Théorème de Bolzano-Weierstrass)**.** *(1) Un sous ensemble de $\mathbf{R}^n$ muni de la norme somme $||(x_i)|| = \sum_i |x_i|$ est compact si et seulement si il est fermé et borné.*

(2) Une fonction réelle continue sur un compact non vide est bornée et atteint ses bornes.

Démonstration. Un fermé d'un compact est compact. Toute boule pour la norme somme étant contenue dans un cube d'arête $2A, A \in \mathbf{R}$ (cf. (3.1.a page 12)), il suffit de prouver que $[-A, A]^n$ est compact pour la norme somme, donc d'après la proposition précédente que $[-A, A]$ est compact pour la valeur absolue ce qui découle de la méthode de dichotomie (cf. la preuve de (3.1.7 page 13)). L'image continue d'un compact étant compacte (4.1.2 page 15), le second point découle du premier : dire qu'une fonction atteint ses bornes résulte de la fermeture de son image. □

Corollaire 4.3.6 (Uniforme continuité sur les compacts)**.** *Soit f une application continue entre espaces métriques $f : X_1 \to X_2$. Si X_1 est compact, alors f est est uniformément continue.*

Démonstration. Soit $\epsilon > 0$ et définissons pour tout $\alpha > 0$

$$K_\alpha = \{(x_1, x_1') \text{ tels que } \mathrm{d}_1(x_1, x_1') \leq \alpha \text{ et } \mathrm{d}_2(f(x_1), f(x_1')) \geq \epsilon\}.$$

Par continuité de la distance, K_α est fermé dans $X_1 \times X_1$ comme image réciproque d'un fermé par une application continue, donc compact (4.3.4 page précédente). La famille des K_α est décroissante d'intersection vide donc il existe $\alpha_0 > 0$ tel que $K_{\alpha_0} = \varnothing$ (4.1.2 page 15). □

4.4 Normes de $\mathbf{R}^N$ et leurs compacts.

Commençons par nous affranchir du choix des normes.

Théorème 4.4.1. *Toutes les normes de $\mathbf{R}^n$ sont équivalentes.*

Démonstration. Soit N une norme et $a = \sup_i N(e_i)$ où e_i est la base canonique de $\mathbf{R}^n$. On a $a > 0$. On note ici N_1 la norme somme

$$N_1((x_1, ..., x_n)) = \sum_i |x_i|.$$

L'inégalité triangulaire donne le sens facile, à savoir

$$(*) \qquad N \leq aN_1.$$

L'inégalité inverse est plus subtile. D'abord, (*) assure que

$$N : (\mathbf{R}^n, N_1) \to (\mathbf{R}, |\bullet|)$$

est continue. Mais alors, (4.3.5 page précédente) assure que la sphère unité S_{n-1} de $(\mathbf{R}^n, N_1)$ est compacte et que N est minorée par un élément de $b \in N(S_{n-1}) \subset \mathbf{R}_+^*$. On a donc

$$\forall x \in S_{n-1},\ N(x) \geq b = nN_1(x).$$

Par homogénéité des normes, on déduit $N \geq bN_1$. □

Corollaire 4.4.2 (Bolzano-Weierstrass)**.** *Pour toute norme sur $\mathbf{R}^n$, les compacts sont les fermés bornés.*

Démonstration. Un compact est fermé et borné (contenu dans un nombre fini de boules). Inversement, si $X \subset \mathbf{R}^n$ est fermé borné, $X \subset [-A, A]^n$ pour A assez grand (équivalence des normes, partie facile). Or, un cube est compact pour la norme somme (4.3.5 page 17) et donc pour N (équivalence des normes, partie subtile). □

Remarque 4.4.3. Attention, si on met une distance et non une norme, c'est faux en général. Exercice : Donner un exemple de distance sur $\mathbf{R}^n$ qui n'est pas équivalente à la distance définie par une norme.



Bernard Bolzano



Karl Weierstrass

Remarque 4.4.4. Attention : dans un Banach de dimension infinie, la boule unité n'est jamais compacte (théorème de Riesz, cf. (3.2.1 page 164) dans le cas Hilbert).

4.5 Limites inf, sup

Rappelons que la la droite numérique achevée est définie par $\bar{\mathbf{R}} = \mathbf{R} \cup \{\pm\infty\}$. On peut la munir d'une distance qui assure que la suite réelle x_n tend vers $\pm\infty$ au sens de la distance de $\bar{\mathbf{R}}$ si et seulement si $\lim x_n = \pm\infty$ au sens usuel. Par exemple, on peut prendre

$$\mathrm{d}(x, y) = |\arctan(x) - \arctan(y)|$$

en posant

$$\arctan(\pm\infty) = \pm\pi/2.$$

Le théorème des accroissements finis assure que pour tout intervalle compact $I \subset \mathbf{R}$, il existe $a, b > 0$ tels que

$$\forall x, y \in \arctan(I),\ a|x - y| \leq |\tan(x) - \tan(y)| \leq b|x - y|.$$

La restriction de d à I est donc équivalente à la valeur absolue. Comme la topologie de $\mathbf{R}$ est engendrée par les intervalles ouverts *bornés*, d induit la topologie usuelle sur $\mathbf{R}$.

Avec cette distance, $\bar{\mathbf{R}}$ est l'image continue de $[-\pi/2, \pi/2]$ par la tangente (prolongée par continuité) et donc est compact. Par construction, toute partie non vide de $\bar{\mathbf{R}}$ admet des bornes supérieure et inférieure.

On déduit que toute suite de $\bar{\mathbf{R}}$ admet au moins une valeur d'adhérence dans $\bar{\mathbf{R}}$. L'ensemble des valeurs d'adhérences d'une suite $(x_n)_{n\geq 1}$ de $\bar{\mathbf{R}}$ étant fermé, on déduit qu'une suite de $\bar{\mathbf{R}}$ admet une plus grande valeur d'adhérence notée $\overline{\lim}\, x_n$ (respectivement une plus petite valeur d'adhérence notée $\underline{\lim}\, x_n$).

À titre d'exercice, le lecteur pourra vérifier les propriétés utiles suivante.

Proposition 4.5.1. *Soit x_n une suite de $\bar{\mathbf{R}}$.*

- *L'inégalité $\underline{\lim}\, x_n \leq \overline{\lim}\, x_n$ est une égalité si et seulement si la suite $(x_n)_{n\geq 1}$ converge dans $\bar{\mathbf{R}}$. On a alors*
$$\lim x_n = \underline{\lim}\, x_n = \overline{\lim}\, x_n.$$
- *La suite réelle x_n converge dans $\mathbf{R}$ vers $\ell \in \mathbf{R}$ si et seulement si $\underline{\lim}\, x_n = \overline{\lim}\, x_n = \ell$.*
- *Les suites*
$$(\inf\{x_k, k \geq n\})_n, (\sup\{x_k, k \geq n\})_n$$
sont monotones de limites respectives
$$\underline{\lim}\, x_n, \overline{\lim}\, x_n.$$

4.6 Applications de la compacité à la théorie de la mesure

Soit Ω ouvert de $\mathbf{R}^N$ et $C_c(\Omega)$ le $\mathbf{R}$-espace vectoriel des fonctions continues (réelles) à support compact et $\Lambda : C_c(\Omega) \to \mathbf{R}$ une forme linéaire positive, ie

$$f \geq 0 \Rightarrow \Lambda f \geq 0.$$

Par exemple l'intégrale sur $[a, b]$ pour $\Omega =]a, b[$ est un telle forme linéaire.

Notons que si Λ est une forme linéaire positive, on a évidemment que $f \leq g$ entraîne que $\Lambda f \leq \Lambda g$.

Proposition 4.6.1 (Application fondamentale de Dini)**.** *Soit $f_n \in C^0(\Omega, \mathbf{R}), n \geq 0$ suite décroissante de fonctions continues à support compact convergeant en tout point vers 0. Alors* $\lim \Lambda f_n = 0$.

Démonstration. Soit $K \subset \Omega$ compact en dehors duquel f_0 est nulle. La décroissance de la suite $(f_n)_{n\geq 1}$, chaque f_n est nulle en dehors de K. Puisque K est inclus dans Ω, la distance $\mathrm{d}(x, \mathbf{R}^N - \Omega)$ de $x \in K$ au fermé $\mathbf{R}^N - \Omega$ est > 0 (exercice 2.3.7 page 11). Par compacité et continuité de la distance (*loc. cit.*et le corollaire 4.3.5 page 17),

$$\exists \epsilon > 0 \text{ tel que } \forall x \in K,\ \mathrm{d}(x, \mathbf{R}^N - \Omega) \geq 2\epsilon$$

et donc

$$(*) \qquad \forall x \in K, \forall y \notin \Omega,\ \mathrm{d}(x, y) \geq 2\epsilon.$$

Comme K, le fermé de $\mathbf{R}^N$

$$K_\epsilon = \{\xi \in \mathbf{R}^N | \mathrm{d}(\xi, K) \leq \epsilon\}$$

est borné, donc compact (Bolzan-Weierstrass (4.4.2 page 18) et $K \subset \Omega$ d'après (*). La fonction continue

$$\xi \mapsto g(\xi) = \sup(1 - \mathrm{d}(\xi, K)/\epsilon, 0)$$

vaut 1 sur K et 0 en dehors du compact $K_\epsilon \subset \Omega$. On a par construction

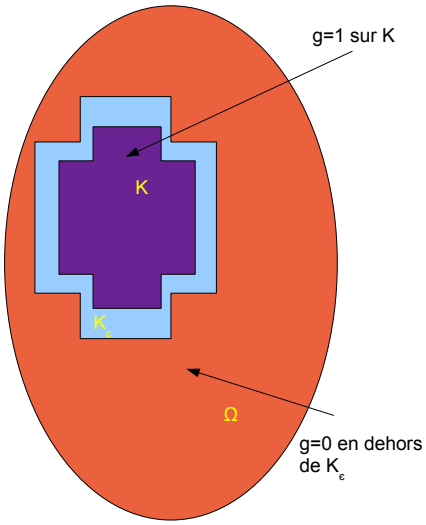


FIGURE II.3 – Un exemple plan de $K \subset K_\epsilon \subset U$

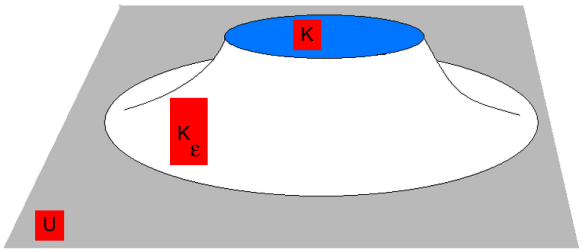


FIGURE II.4 – Graphe d'un g en perspective

$$g \in C_c(\Omega) \text{ et } f_n = gf_n \text{ car } g = 1 \text{ sur support}(f_n)$$

et donc

$$\Lambda f_n = \Lambda g f_n \leq \sup(f_n)\Lambda g \quad \text{par positivité de } \Lambda$$

et on applique Dini (théorème 4.2.1 page 16) pour obtenir que $\sup(f_n)$ tend vers 0. □

Au cours de la démonstration, on a démontré un résultat topologique intéressant au cours de la preuve (connu sous le nom de lemme d'Urysohn) qui sera utilisé en intégration (*stricto sensu* démontrée pour l'espace métrique $\mathbf{R}^N$, mais qui se généralise immédiatement).

Corollaire 4.6.2 ((de la preuve)). *Soit Ω un ouvert d'un métrique contenant un compact K. Alors, il existe $f \in C_c(\Omega)$ qui vaut 1 sur K.*

5 Connexité

Intuitivement, un espace connexe est ≪ fait d'un seul tenant ≫.

Définition 5.0.3. Un espace métrique X, est dit connexe, s'il n'existe pas de sous-ensemble de X autre que $\emptyset$ et X qui soit à la fois ouvert et fermé.

La connexité de X s'exprime aussi en disant qu'il n'existe pas de fonction continue non constante à valeurs dans $\{0,1\}$: en effet si on pouvait écrire $X = U \cup V$ avec U, V ouverts disjoints non vides, la fonction caractéristique de U serait continue et non constante. Inversement, si $f : X \to \{0,1\}$ est continue les ensembles $U = f^{-1}(0), V = f^{-1}(1)$ sont des ouverts dont la réunion est X, non vides si et seulement si f est non constante.

La propriété fondamentale de la connexité est d'être héréditaire par image directe.

Proposition 5.0.4. *Si f est continue de X dans Y et si X est connexe, alors $f(X)$ est connexe.*

Démonstration. Il suffit de l'écrire ! ! Si $f(X) = A \cup B$ avec A, B ouverts disjoints non vides, alors $f^{-1}(A)$ et $f^{-1}(B)$ sont ouverts, disjoints et non vides, et leur réunion est X. Par connexité de X, l'un des deux ensembles $f^{-1}(A)$ ou $f^{-1}(B)$ est vide, et puisque A et B sont contenus dans l'image de X par f, nécessairement A ou B est vide. □

Une propriété plus intuitive est la **connexité par arcs**. Tous les espaces connexes du cours seront connexes par arcs.

Définition 5.0.5. Un espace métrique (X, d) est connexe par arcs, si deux points quelconques peuvent être reliés par un arc continu :

$$\forall x, y \in X \quad \exists\, \gamma \in C^0([0,1], X) \text{ tel que } \quad \gamma(0) = x, \gamma(1) = y$$

La connexité par arcs entraîne de manière évidente la connexité. En effet, si f est à valeurs dans $\{0,1\}$, et si x et y sont reliés par un chemin, le théorème des valeurs intermédiaires impose à f de prendre la même valeur en x et y. Donc f est constante.

Exemple 5.0.6. Un espace vectoriel, un convexe, un ensemble étoilé [1] sont connexes. Une réunion de connexes ayant un point commun est connexe. L'union des sous-ensembles connexes contenant un point s'appelle la *composante connexe* du point. Le cercle, la sphère et le tore sont connexes.

Exercice 5.0.7. Soit X une partie connexe d'un espace métrique. Montrer que toute partie coincée entre X et $\bar{X}$ est connexe.

Exercice 5.0.8. Montrer qu'un ouvert connexe Ω de $\mathbf{R}^n$ est connexe par arcs. Montrer qu'on peut même joindre deux points de Ω par une ligne polygonale.

1. i.e. tel qu'il existe un point tel que tout segment entre ce point et un autre point de X est contenu dans X.

6 Appendice : Quelques mots sur la dénombrabilité

6.1 Sorites

Rappelons qu'on dit qu'un ensemble X est dénombrable s'il est en bijection avec une partie de $\mathbf{N}$, autrement dit, et plus concrètement, si on peut numéroter ses éléments

$$X = \{x_1, \cdots, x_n, \cdots \ 1 \leq n \leq \mathrm{card}(X)\}$$

Notons que toute partie dénombrable **infinie** de $\mathbf{N}$ est en bijection avec $\mathbf{N}$. Si $X \subset \mathbf{N}$ on pose si X non vide

$$x_1 = \inf(X), x_{n+1} = \inf_{x \in X}\{x > x_n\}) \text{ pour } 1 \leq n < \mathrm{card}(X).$$

Proposition 6.1.1. *Une partie d'un dénombrable, l'image d'un dénombrable sont dénombrables.*

Démonstration. Si $f : X = Y$, on pose $y_1 = f(x_1)$ puis

$$N = \inf_p\{p | f(x_p) \neq y_1, \cdots, y_n\} \text{ si } n < \mathrm{card}(f(X)).$$

On pose alors $y_{n+1} = f(x_N)$ qui donne

$$f(X) = \{y_n, 1 \leq n \leq \mathrm{card}(f(X))\}.$$

□

Premières propriétés :

1) Un produit fini de deux dénombrables est dénombrable. Il suffit en effet (récurrence) de prouver que $\mathbf{N}^2$ est dénombrable. Mais $(n, m) \mapsto 2^n 3^m$ est injective et identifie donc $\mathbf{N}^2$ à une partie de $\mathbf{N}$.

2) Une réunion dénombrable $\cup X_n$ d'ensembles dénombrables est dénombrable. Il suffit d'écrire

$$X_n = \{x_{n,i}, i \geq 0\} \text{ et de considérer } (n, i) =\mapsto x_{n,i}.$$

3) $\mathbf{Z} \xrightarrow{\sim} \pm 1 \times \mathbf{N}^* \cup \{0\}$ et $\mathbf{Q} = \cup_{n \in \mathbf{N}^*} \frac{1}{n}\mathbf{Z}$ sont dénombrables.

6.2 Topologie des réels et dénombrabilité

Proposition 6.2.1. *Un ouvert U de $\mathbf{R}$ est réunion dénombrable disjointe d'intervalles ouverts.*

Démonstration. Pour $x \in U$, soit I_x la réunion des intervalles ouverts contenant x : c'est le plus grand intervalle ouvert contenant x. Considérons l'ensemble de parties de U

$$\mathcal{F} = \{I_x, x \in U\}.$$

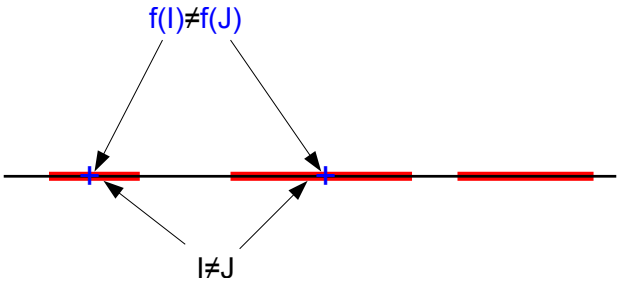
La remarque clef est que deux intervalles de $\mathcal{F}$ sont deux à deux disjoints ou confondus. En effet, si $I_x, I_y \in \mathcal{F}$ ont un point commun, la réunion $I = I_x \cup I_y$ est un intervalle ouvert contenant x et y et donc est contenu à la fois dans I_x et I_y (maximalité) de sorte que $I_x = I_y = I$. On a donc

$$U = \bigsqcup_{I \in \mathcal{F}} I.$$

Montrons que $\mathcal{F}$ est dénombrable. Pour tout $I \in \mathcal{F}$, choisissons un rationnel

$$f(I) \in \mathbf{Q} \cap I$$

(tout intervalle non-vide contient un rationnel, par densité des rationnels dans $\mathbf{R}$). Si $p/q = f(I) = f(J)$, les intervalles I et J ont un point commun p/q et donc sont égaux.



On déduit que f est injective et donc $\mathcal{F}$, comme $\mathbf{Q}$, est dénombrable . □

La non dénombrabilité de $\mathbf{R}$ est bien connue et résultera aussi [2] de l'existence de la mesure de Lebesgue (3 page 98) en fait comme on le verra. Voyons une très jolie preuve due à Cantor, le fondateur de l'arithmétique des cardinaux infinis.

Théorème 6.2.2 (Argument diagonal de Cantor)**.** *$I = [0, 1[$ (et donc $\mathbf{R}$) n'est pas dénombrable.*

Démonstration. On suppose qu'on a énuméré $I = \{x_n, n \geq 0\}$. On écrit le développement (propre) en base 10

$$x_n = \sum_{k \geq 1} a_{k,n} 10^{-k}.$$

On pose $a_n = 2$ si $a_{n,n} = 1$ et $a_n = 1$ sinon ; on a

$$\forall n,\ a_n \neq a_{n,n}.$$

$$\text{On pose } x = \sum_{k \geq 1} a_k 10^{-k} = x_n \text{ développement propre.}$$

Si $\exists n$ tel que $x = x_n$, on a $a_n = a_{n,n}$ par construction, une contradiction.

□

2. Ce qui, soit dit en passant, montre que cette construction de la mesure est non triviale. Si $\mathbf{R}$ était dénombrable, on ne pourrait pas mesurer les longueurs de manière raisonnable...

6.3 Liste des exercices de topologie

Exercice 2.1.3 page 9

Exercice 2.2.6 page 10

Exercice 2.2.7 page 10

Exercice 2.2.8 page 10

Exercice 2.3.5 page 11

Exercice 2.3.6 page 11

Exercice 2.3.7 page 11

Exercice 3.1.12 page 14

Exercice 5.0.8 page 22

Exercice 5.0.7 page 22

Chapitre III

Construction de l'intégrale de Lebesgue

1 Motivations

La notion élémentaire d'intégrale étudiée dans les classes préparatoires est limitée au cadre des fonctions continues par morceaux sur un intervalle de $\mathbf{R}$. C'est tout à fait suffisant tant qu'on veut pouvoir calculer explicitement les intégrales de fonctions élémentaires. Malheureusement, les énoncés d'intégration terme à terme de suites ou de séries de fonctions simplement convergentes sont peu pratiques avec cette notion élémentaire d'intégrale. Or, avec l'importance croissante du calcul scientifique sur ordinateur — probablement l'une des mutations les plus importantes en mathématiques au XXème siècle — il est bien souvent plus utile de savoir approcher une quantité que de savoir la calculer explicitement.

Le but de ce premier chapitre est donc double :

A) définir l'intégrale pour une classe de fonctions plus générale que celle des fonctions continues par morceaux sur un intervalle de $\mathbf{R}$; et

B) disposer de théorèmes d'intégration terme à terme pour des suites de fonctions où il ne soit pas nécessaire de vérifier que la fonction limite est continue par morceaux.

Pour le point B), on aimerait disposer d'un énoncé du type suivant sur les séries de fonctions :

Soit $(u_n)_{n\geq 1}$ suite de fonctions positives "intégrables" sur $\Omega \subset \mathbf{R}^N$; alors

$$\Omega \ni x \mapsto u(x) := \sum_{n\geq 1} u_n$$

est une fonction définie sur Ω à valeurs dans $[0,+\infty]$, et

$$\int_\Omega u(x)dx = \sum_{n\geq 1} \int_\Omega u_n(x)dx \qquad \textit{(égalité dans } [0,+\infty]\textit{)}\,.$$

Voici un énoncé équivalent pour les suites de fonctions :

Soit $(U_n)_{n\geq 1}$ suite de fonctions "intégrables" sur $\Omega \subset \mathbf{R}^N$, croissante au sens où

$$U_0(x) \leq U_1(x) \leq \ldots \leq U_n(x) \leq \ldots \text{ pour tout } x \in \Omega\,;$$

alors

$$\int_\Omega \left(\lim_{n\to+\infty} U_n(x) \right) dx = \lim_{n\to+\infty} \int_\Omega U_n(x)dx \quad \text{(égalité dans } [0,+\infty]\text{)}\,.$$

On passe évidemment d'un énoncé à l'autre en posant

$$U_n(x) = U_0(x) + \sum_{k=1}^{n} u_k(x)\,.$$

Les points A) et B) sont intimement reliés. Pour disposer d'un énoncé d'inté- gration terme à terme de séries de fonctions intégrables convergeant simplement, sans hypothèse sur la somme de la série, il faut pouvoir intégrer des fonctions extrêmement singulières.

Exemple 1.0.1 (Fonction indicatrice des rationnels). Rappelons que $\mathbf{Q}$ est dénombrable, de sorte que $\mathbf{Q} \cap [0,1]$ (sous-ensemble de $\mathbf{Q}$) est également dénombrable. Puisque $\mathbf{Q} \cap [0,1]$ est infini, il existe $\mathbf{N} \ni n \mapsto r_n \in \mathbf{Q} \cap [0,1]$, bijection numérotant tous les rationnels de $[0,1]$. Pour tout $n \in \mathbf{N}$, posons

$$u_n(x) = \begin{cases} 0 & \text{si } x \neq r_n\,, \\ 1 & \text{si } x = r_n\,. \end{cases}$$

La fonction u_n est positive ou nulle et continue par morceaux sur $[0,1]$, mais

$$\sum_{n\geq 0} u_n(x) = \mathbf{1}_{\mathbf{Q}\cap[0,1]}(x) = \begin{cases} 0 & \text{si } x \notin \mathbf{Q} \cap [0,1]\,, \\ 1 & \text{si } x \in \mathbf{Q} \cap [0,1]\,. \end{cases}$$

Or, comme $\mathbf{Q}$ est dense dans $\mathbf{R}$, la fonction $\mathbf{1}_{\mathbf{Q}\cap[0,1]}$ est discontinue en tout point de $[0,1]$. Il n'est donc pas possible de l'intégrer avec la notion élémentaire d'intégrale étudiée en classes préparatoires. Toutefois, l'énoncé B) suggère que

$$\int_0^1 \mathbf{1}_{\mathbf{Q}\cap[0,1]}(x)dx = \sum_{n\geq 0} \int_0^1 u_n(x) = 0\,.$$

Observons toutefois que la somme de Riemann

$$R_n(\mathbf{1}_{\mathbf{Q}\cap[0,1]}) = \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\mathbf{Q}\cap[0,1]} \left(\frac{k}{n} \right) = 1\,.$$

On ne peut donc aboutir à une notion d'intégrale vérifiant la propriété B) ci-dessus en définissant la classe des fonctions intégrables sur $[0,1]$ comme l'ensemble des fonctions $f : [0,1] \to \mathbf{R}$ telles que $R_n(f)$ converge quand $n \to +\infty$ et l'intégrale d'une telle fonction f comme la limite de la suite $R_n(f)$.

Cet exemple suggère que la notion d'intégrale cherchée doit être construite par un procédé radicalement différent de celui permettant d'intégrer les fonctions continues.

2 Rappels sur l'intégration des fonctions continues

La construction de l'intégrale de Lebesgue présentée dans ce cours s'appuie sur la notion usuelle d'intégrale étudiée dans les classes préparatoires, dont elle constitue un prolongement [1]. Commençons donc par quelques rappels et compléments sur cette notion usuelle d'intégrale.

Soient $a < b \in \mathbf{R}$; rappelons que si f est une fonction continue sur le segment $[a, b]$ à valeurs dans $\mathbf{R}$, son intégrale peut être calculée comme limite de sommes de Riemann :

$$\int_a^b f(x)dx = \lim_{q\to+\infty} \frac{1}{q} \sum_{k=0}^{q-1} f\left(a + k\frac{b-a}{q}\right) .$$

Nous aurons besoin de généraliser cette formule au cas de fonctions de plusieurs variables.

Soient $N \geq 1$ et $\Omega \subset \mathbf{R}^N$ ouvert. Dans toute la suite, on notera $C_c(\Omega)$ l'ensemble des fonctions continues sur Ω à valeurs réelles et à support compact dans Ω. Considérons, pour tout $j \in \mathbf{N}$, la somme de Riemann dyadique

$$R_j(f) := 2^{-jN} \sum_{\substack{k\in\mathbf{Z}^N \\ 2^{-j}k\in\Omega}} f\left(2^{-j}k\right) .$$

Comme f est à support compact dans Ω, la somme ci-dessus ne comporte qu'un nombre fini de termes non nuls.

Lemme 2.0.2. Soit $F \in C_c(\mathbf{R}^N)$. Alors la suite de réels indexée par $j \in \mathbf{N}$

$$R_j(F) := 2^{-jN} \sum_{k\in\mathbf{Z}^N} F\left(2^{-j}k\right)$$

converge pour $j \to +\infty$.

Démonstration. De même que les sommes de Riemann sont basées sur une subdivision du segment où l'on intègre la fonction en petits intervalles, nous allons subdiviser l'espace euclidien $\mathbf{R}^N$ en cubes de plus en plus petits.

Soit $C = [0, 1]^N$ le cube unité de $\mathbf{R}^N$. Pour $j \in \mathbf{N}$ et $k \in \mathbf{Z}^N$, on pose $C_{j,k} = \{2^{-j}(k+z) \mid z \in C\}$, qui est un cube fermé de $\mathbf{R}^N$ de côté 2^{-j}, dont l'un des sommet est le point $2^{-j}k$. L'indice j définit donc la taille de la subdivision de $\mathbf{R}^N$ en cubes, et on a

$$\mathbf{R}^N = \bigcup_{k\in\mathbf{Z}^N} C_{j,k} , \quad \text{pour tout } j \in \mathbf{N} .$$

Le volume de $C_{j,k}$ est 2^{-jN}, de sorte que le facteur 2^{-jN} dans la définition de $R_j(F)$ joue un rôle analogue à celui du facteur $1/q$ dans la définition de la somme de Riemann en dimension 1

1. Il existe plusieurs manières de construire l'intégrale de Lebesgue ; celle qui est adoptée ici remonte à P. J. Daniell, *A general form of integral*, Annals of Mathematics **19** (1918), 279–94. La méthode de Daniell est exposée de façon lumineuse dans le chapitre II des *Leçons d'analyse fonctionnelle* de F. Riesz et B. Sz.-Nagy, Gauthier-Villars, Paris, 1968, ou encore, dans un langage plus moderne, dans l'excellent livre de M. Willem [20]. Nous avons suivi ici la présentation concise de M. Willem, en la complétant par de nombreux exemples.

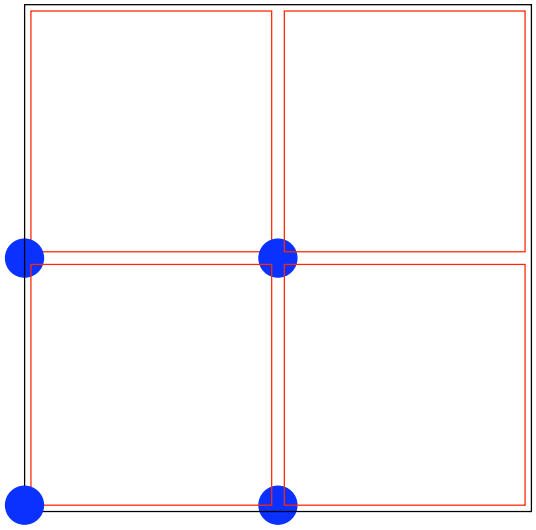


FIGURE III.1 – Le carré noir étant $C_{j,k} \subset \mathbf{R}^2$, les quatre points bleus sont les points de la forme $2^{-j-1}l$ pour $l \in S_k$, et les quatre carrés rouges sont les $C_{j+1,l}$ lorsque l décrit S_k.

rappelée ci-dessus. Posons

$$I_j^-(F) = 2^{-jN} \sum_{k \in \mathbf{Z}^N} \min_{x \in C_{j,k}} F(x) \quad \text{et} \quad I_j^+(F) = 2^{-jN} \sum_{k \in \mathbf{Z}^N} \max_{x \in C_{j,k}} F(x)\,.$$

On a évidemment

$$I_j^-(F) \leq R_j(F) \leq I_j^+(F)\,, \qquad \text{pour tout } j \in \mathbf{N}\,.$$

Soit $S_k = \{l \in \mathbf{Z}^N \,|\, l - 2k \in \{0,1\}^N\}$. Alors

$$C_{j,k} = \bigcup_{l \in S_k} C_{j+1,l}$$

(voir Figure 2). Par conséquent

$$l \in S_k \Rightarrow \min_{x \in C_{j,k}} F(x) \leq \min_{x \in C_{j+1,l}} F(x) \leq \max_{x \in C_{j+1,l}} F(x) \leq \max_{x \in C_{j,k}} F(x)\,,$$

de sorte que

$$\begin{aligned} I_j^-(F) &= 2^{-jN} \sum_{k \in \mathbf{Z}^N} 2^{-N} \sum_{l \in S_k} \min_{x \in C_{j,k}} F(x) \\ &\leq 2^{-jN} \sum_{k \in \mathbf{Z}^N} 2^{-N} \sum_{l \in S_k} \min_{x \in C_{j+1,l}} F(x) = I_{j+1}^-(F)\,. \end{aligned}$$

De même, on montre que $I_j^+(F) \geq I_{j+1}^+(F)$, de sorte que l'on a, pour tout $j \in \mathbf{N}$,

$$I_0^-(F) \leq I_1^-(F) \leq \ldots \leq I_j^-(F) \leq R_j(F) \leq I_j^+(F) \leq \ldots \leq I_1^+(F) \leq I_0^+(F)\,.$$

Enfin, comme F est à support compact, il existe $L \in \mathbf{N}^*$ t.q. $F(x) = 0$ si $x \notin [-L, L]^N$. Donc [2]

$$\begin{aligned} I_j^+(F) - I_j^-(F) &= 2^{-jN} \sum_{k \in \mathbf{Z}^N \cap [-2^j L, 2^j L]^N} \max_{x,y \in C_{j,k}} |F(x) - F(y)| \\ &\leq 2^{-jN} \#(\mathbf{Z}^N \cap [-2^j L, 2^j L]^N) \sup_{|x-y| \leq 2^{-j}\sqrt{N}} |F(x) - F(y)| \\ &= (2L+1)^N \sup_{|x-y| \leq 2^{-j}\sqrt{N}} |F(x) - F(y)| \to 0 \end{aligned}$$

lorsque $j \to +\infty$ puisque F, étant continue à support compact sur $\mathbf{R}^N$, est uniformément continue (d'après le Corollaire 4.3.6 page 18.)

Les suites $(I_j^-(F))_{j \geq 0}$ et $(I_j^+(F))_{j \geq 0}$ sont donc adjacentes, et convergent lorsque $j \to +\infty$ vers une limite commune, ainsi que la suite $(R_j(F))_{j \geq 0}$ qu'elles encadrent. □

Etant donnée $f \in C_c(\Omega)$, on déduit du Lemme 2.0.2 page 29 appliqué à la fonction $F \in C_c(\mathbf{R}^N)$ définie par

$$F(x) := \begin{cases} f(x) & \text{si } x \in \Omega\,, \\ 0 & \text{si } x \notin \Omega\,, \end{cases}$$

que la suite $R_j(f)$ est convergente lorsque $j \to +\infty$.

Définition 2.0.3 (Intégrale des fonctions de $C_c(\Omega)$). Soit $\Omega \subset \mathbf{R}^N$ ouvert et $f \in C_c(\Omega)$. On pose

$$\int_\Omega f(x)dx := \lim_{j \to +\infty} R_j(f)\,.$$

Le lecteur est invité à vérifier par lui-même, à titre d'exercice

a) que cette formule coïncide bien avec la formule d'approximation de l'intégrale par les sommes de Riemann usuelles dans le cas où f est une fonction continue sur $\Omega = \mathbf{R}$ à support dans le segment $[0, 1]$;

b) que, dans le cas où $N = 2$ avec $\Omega = \mathbf{R}^2$ et où f est continue sur Ω et à support dans $K \times K$, où K est un segment de $\mathbf{R}$, cette formule redonne bien la valeur

$$\begin{aligned} \int_{\mathbf{R}^2} f(x)dx &= \iint_{K \times K} f(x_1, x_2)dx_1 dx_2 \\ &= \int_K \left(\int_K f(x_1, x_2)dx_1 \right) dx_2 = \int_K \left(\int_K f(x_1, x_2)dx_2 \right) dx_1\,. \end{aligned}$$

Cette notion d'intégrale que nous venons de définir sur $C_c(\Omega)$ coïncide donc avec la notion d'intégrale multiple étudiée dans les classes préparatoires.

La *seule* propriété de l'intégrale sur $C_c(\Omega)$ que nous allons utiliser pour définir l'intégrale de Lebesgue est la suivante : l'application

$$C_c(\Omega) \ni f \mapsto \int_\Omega f(x)dx \in \mathbf{R}$$

2. Pour la distance euclidienne dans $\mathbf{R}^N$, le diamètre de $[0, 1]^N$ vaut $\sqrt{N}$.

est une forme **R**-linéaire sur le **R**-espace vectoriel $C_c(\Omega)$, **positive**, au sens où

$$(P) \qquad f \in C_c(\Omega) \text{ et } f(x) \geq 0 \text{ pour tout } x \in \Omega \Rightarrow \int_\Omega f(x)dx \geq 0\,.$$

Voici un premier résultat d'intégration terme à terme de suites convergeant simplement, cas particulier de la Proposition 4.6.1 page 20 compte-tenu de la positivité de l'intégrale.

Lemme 2.0.4. Soit $(f_n)_{n\geq 1}$ suite de fonctions de $C_c(\Omega)$ convergeant simplement vers 0 sur Ω et telle que

$$f_1(x) \geq f_2(x) \geq \ldots \geq f_n(x) \geq f_{n+1}(x) \geq \ldots \geq 0 \text{ pour tout } x \in \Omega\,.$$

Alors

$$\lim_{n\to+\infty} \int_\Omega f_n(x)dx = 0\,.$$

3 Définition de l'intégrale de Lebesgue

On va construire l'intégrale de Lebesgue à partir de l'intégrale usuelle en deux étapes :

a) on *impose* l'interversion limite/intégrale pour les suites croissantes de fonctions de $C_c(\Omega)$ dont la suite des intégrales est majorée ; et

b) on *prolonge par linéarité* cette première extension.

Cette construction fournit un **R**-espace vectoriel contenant strictement $C_c(\Omega)$, et une forme linéaire positive définie sur cet espace vectoriel, prolongeant l'intégra- le usuelle sur $C_c(\Omega)$, et vérifiant l'énoncé B) de l'introduction (section 1 page 27).

3.1 Etape 1 : classe $\mathcal{L}^+(\Omega)$

Définition 3.1.1 (Suites fondamentales). Une suite $(f_n)_{n\geq 0}$ de fonctions de $C_c(\Omega)$ est dite fondamentale si elle vérifie

$$f_0(x) \leq f_1(x) \leq \ldots \leq f_n(x) \leq f_{n+1}(x) \leq \ldots \text{ pour tout } x \in \Omega\,,$$

et

$$\sup_{n\geq 0} \int_\Omega f_n(x)dx < +\infty\,.$$

Pour que l'intégrale qu'on cherche à construire vérifie l'énoncé B) de l'introduction (section 1 page 27), cet énoncé doit déjà être vrai pour les suites fondamentales. Ceci motive la définition suivante :

Définition 3.1.2 (Classe $\mathcal{L}^+(\Omega)$). On pose

$$\mathcal{L}^+(\Omega) := \big\{ f : \Omega \to \mathbf{R} \cup \{+\infty\} \mid \text{ il existe } (f_n)_{n\geq 0} \text{ suite fondamentale t.q. } f_n \to f \text{ simplement sur } \Omega \big\}\,.$$

Pour tout $f \in \mathcal{L}^+(\Omega)$, si $(f_n)_{n\geq 0}$ est une suite fondamentale convergeant simplement vers f sur Ω, la suite des intégrales

$$\int_\Omega f_n(x)dx \text{ est croissante est majorée, donc converge vers sa borne supérieure,}$$

et on prolonge l'intégrale à la classe $\mathcal{L}^+(\Omega)$ en posant

$$\int_\Omega f(x)dx = \lim_{n\to+\infty} \int_\Omega f_n(x)dx\,.$$

La définition ci-dessus de l'intégrale de $f \in \mathcal{L}^+(\Omega)$ est indépendante du choix de la suite fondamentale $(f_n)_{n\in\mathbf{N}}$ convergeant simplement vers f sur Ω. En effet :

Lemme 3.1.3. Soient $(f_n)_{n\geq 0}$ et $(g_n)_{n\geq 0}$, suites fondamentales sur Ω ouvert de $\mathbf{R}^N$. Alors

$$\lim_{n\to+\infty} f_n(x) = \lim_{n\to+\infty} g_n(x) \text{ pour tout } x \in \Omega$$
$$\Rightarrow \lim_{n\to+\infty} \int_\Omega f_n(x)dx = \lim_{n\to+\infty} \int_\Omega g(x)dx\,.$$

Avant de donner la démonstration du lemme ci-dessus, rappelons les notions de parties positives et négatives d'un réel, que nous utiliserons systématiquement dans toute la suite.

Parties positives et négatives d'un réel

Pour tout $a \in \mathbf{R}$, on note

$$a^+ = \max(a, 0)\,, \qquad a^- = \max(-a, 0)\,;$$

on a

$$a = a^+ - a^-\,, \qquad |a| = a^+ + a^-\,.$$

Evidemment

$$a^+ = \tfrac{1}{2}(|a| + a)\,, \qquad a^- = \tfrac{1}{2}(|a| - a)\,.$$

Etant donnée une fonction f à valeurs réelles, on notera f^+ la fonction $x \mapsto f(x)^+$ et f^- la fonction $x \mapsto f(x)^-$.

Démonstration. Pour tout $x \in \Omega$, notons

$$f(x) := \lim_{n\to+\infty} f_n(x) \text{ et } g(x) := \lim_{n\to+\infty} g_n(x)\,.$$

Pour $m \in \mathbf{N}$ fixé et pour tout $n \in \mathbf{N}$, on pose $h_n := (f_m - g_n)^+$; évidemment $h_n \in C_c(\Omega)$ car f_m et $g_n \in C_c(\Omega)$.

La suite $(h_n(x))_{n\geq 0}$ est décroissante pour tout $x \in \Omega$, et

$$\lim_{n\to+\infty} h_n(x) = (f_m(x) - g_n(x))^+ \leq (f(x) - g(x))^+ = 0\,, \text{ pour tout } x \in \Omega\,.$$

On déduit alors du Lemme 2.0.4 page 32 que

$$\int_\Omega h_n(x)dx \to 0 \quad \text{lorsque } n \to +\infty\,.$$

Mais puisque

$$f_m(x) \le g_n(x) + h_n(x) \quad \text{pour tout } x \in \Omega\,,$$

on trouve que

$$\int_\Omega f_m(x)dx \le \lim_{n\to+\infty} \int_\Omega (g_n(x) + h_n(x))dx = \lim_{n\to+\infty} \int_\Omega g_n(x)dx\,.$$

Comme ceci vaut pour tout $m \in \mathbf{N}$, on trouve en passant à la limite pour $m \to +\infty$ que

$$\lim_{m\to+\infty} \int_\Omega f_m(x)dx \le \lim_{n\to+\infty} \int_\Omega g_n(x)dx\,.$$

On conclut en échangeant les rôles des suites $(f_n)_{n\ge0}$ et $(g_n)_{n\ge0}$. □

Les fonctions de la classe $\mathcal{L}^+(\Omega)$ ne sont pas continues en général ; toutefois, elles vérifient la propriété suivante, qui est plus faible que la continuité.

Lemme 3.1.4. Soit $f : \Omega \to \mathbf{R} \cup \{+\infty\}$; il y a équivalence entre

a) pour tout $\lambda > 0$, l'ensemble $f^{-1}(]\lambda, +\infty])$ est ouvert dans Ω ;

b) pour toute suite $(x_n)_{n\ge0}$ de points de Ω convergeant vers $\xi \in \Omega$, on a

$$f(\xi) \le \varliminf_{n\to+\infty} f(x_n)\,.$$

Démonstration. Montrons que a) implique b). En effet, soit $(x_n)_{n\ge0}$ suite de points de Ω convergeant vers $\xi \in \Omega$; il existe donc une suite extraite $(x_{n_k})_{k\ge0}$ telle que

$$f(x_{n_k}) \to \phi := \varliminf_{n\to+\infty} f(x_n)\,.$$

Pour $\epsilon > 0$, il existe donc $K(\epsilon) \ge 0$ tel que

$$k \ge K(\epsilon) \Rightarrow x_{n_k} \in f^{-1}(]-\infty, \phi+\epsilon])\,.$$

Or d'après le a), l'ensemble $f^{-1}(]-\infty, \phi+\epsilon])$ est fermé dans Ω ; comme de plus $x_{n_k} \to \xi \in \Omega$ lorsque $k \to +\infty$, on conclut que $\xi \in f^{-1}(]-\infty, \phi+\epsilon])$, c'est-à-dire que $f(\xi) \le \phi + \epsilon$. Comme ceci vaut pour tout $\epsilon > 0$, on conclut que $f(\xi) \le \phi$, qui est l'inégalité du b).

Réciproquement, supposons que b) est vérifiée, et soit $\phi \in \mathbf{R}$. On cherche si $f^{-1}(]-\infty, \phi])$ est fermé dans Ω. Soit donc $(x_n)_{n\ge0}$ suite quelconque de points de $f^{-1}(]-\infty, \phi])$ convergeant vers $\xi \in \Omega$; il suffit de vérifier que l'on a $\xi \in f^{-1}(]-\infty, \phi])$, c'est-à-dire que $f(\xi) \le \phi$. Or, comme par hypothèse $f(x_n) \le \phi$ pour tout $n \ge 0$, d'après le b)

$$f(\xi) \le \varliminf_{n\to+\infty} f(x_n) \le \phi\,,$$

ce qui établit le a). □

Définition 3.1.5 (Fonctions s.c.i.). Une fonction $f : \Omega \to \mathbf{R} \cup \{+\infty\}$ est dite semi-continue inférieurement (s.c.i.) sur Ω si elle vérifie les conditions équivalentes du lemme ci-dessus.

Proposition 3.1.6. Toute fonction $f \in \mathcal{L}^+(\Omega)$ est s.c.i. sur Ω.

Démonstration. Montrons que f vérifie la condition a) du lemme ci-dessus. Soient $\lambda \in \mathbf{R}$, et $(f_n)_{n\geq 0}$ suite fondamentale convergeant simplement vers f sur Ω. Un point $x \in \Omega$ vérifie $f(x) > \lambda$ si et seulement si il existe $n \geq 0$ tel que $f_n(x) > \lambda$; donc on a

$$f^{-1}(]\lambda, +\infty[) = \bigcup_{n\geq 0} f_n^{-1}(]\lambda, +\infty[)\,.$$

Or $f_n^{-1}(]\lambda, +\infty[)$ est ouvert dans Ω pour tout $n \in \mathbf{N}$ car f_n est continue, de sorte que $f^{-1}(]\lambda, +\infty[)$ est ouvert dans Ω comme réunion d'ouverts. □

Voici quelques exemples et contre-exemples de fonctions appartenant à la classe $\mathcal{L}^+(\Omega)$.

Exemple 3.1.7. 1) Pour tout $\alpha > 0$, la fonction

$$f_\alpha : \mathbf{R} \ni x \mapsto \frac{\mathbf{1}_{]0,+\infty[}(x)}{1 + x^\alpha}$$

appartient à $\mathcal{L}^+(\mathbf{R})$ si et seulement si $\alpha > 1$.

2) La fonction $\mathbf{1}_{\mathbf{Q}\cap[0,1]}$ n'appartient pas à $\mathcal{L}^+(\mathbf{R})$.

Le lecteur est invité à vérifier ces deux assertions à titre d'exercice.

Comme on l'a noté, les fonctions de la classe $\mathcal{L}^+(\Omega)$ sont à valeurs dans $\mathbf{R} \cup \{+\infty\}$; pour effectuer des opérations sur ces fonctions, nous aurons besoin des conventions suivantes sur l'arithmétique dans $\mathbf{R} \cup \{+\infty\}$.

ARITHMÉTIQUE DANS $\mathbf{R} \cup \{+\infty\}$

$$\begin{cases} z \in \mathbf{R} \cup \{+\infty\} \Rightarrow z + \infty = +\infty\,, \\ \alpha > 0 \Rightarrow \alpha \times (+\infty) = +\infty\,, \\ 0 \times (+\infty) = 0 \end{cases}$$

Le lecteur habitué à devoir "lever l'indétermination" dans les limites de suites du type $u_n v_n$ sachant que $u_n \to 0$ et $v_n \to +\infty$ lorsque $n \to +\infty$ pourra trouver surprenant que l'on postule a priori que $0 \times (+\infty) = 0$. Cette convention sert à pouvoir calculer une expression du type $\alpha f(x)$ avec $\alpha \geq 0$ et $f \in \mathcal{L}^+(\Omega)$. Si $f(z) = +\infty$ et que $\alpha = 0$, il est naturel de poser $\alpha f(z) = 0$ puisque

$$\alpha f(x) = \alpha \lim_{n\to+\infty} f_n(x)$$

pour tout $x \in \Omega$ et toute suite fondamentale $(f_n)_{n\geq 0}$ sur Ω. Or $f_n \in C_c(\Omega)$ pour tout $n \geq 0$, de sorte que, même si $f(z) = +\infty$, on a $f_n(z) < +\infty$, de sorte que $\alpha f_n(z) = 0$ pour tout $n \geq 0$.

Voici les premières propriétés de $\mathcal{L}^+(\Omega)$ et de l'intégrale définie sur $\mathcal{L}^+(\Omega)$.

Proposition 3.1.8. Soit $\Omega \subset \mathbf{R}^N$ ouvert. Alors

1) $C_c(\Omega) \subset \mathcal{L}^+(\Omega)$;

2) pour tous $\alpha, \beta \geq 0$ et $f, g \in \mathcal{L}^+(\Omega)$, on a $\alpha f + \beta g \in \mathcal{L}^+(\Omega)$ et

$$\int_\Omega (\alpha f(x) + \beta g(x))dx = \alpha \int_\Omega f(x)dx + \beta \int_\Omega g(x)dx\,;$$

3) pour tous $f, g \in \mathcal{L}^+(\Omega)$, on a $\max(f, g)$ et $\min(f, g) \in \mathcal{L}^+(\Omega)$ et

$$f(x) \leq g(x) \text{ pour tout } x \in \Omega \Rightarrow \int_\Omega f(x)dx \leq \int_\Omega g(x)dx\,.$$

Tous ces énoncés sont triviaux, et leur démonstration est laissée au lecteur.

Par construction de la classe $\mathcal{L}^+(\Omega)$, l'énoncé B) d'interversion intégrale/limi- te de l'introduction (section 1 page 27) est vérifié pour les suites fondamentales. Mais ce n'est pas tout : ce même énoncé vaut encore pour les suites croissantes de fonctions de $\mathcal{L}^+(\Omega)$.

Proposition 3.1.9. Soit $(f_n)_{n\geq 0}$ suite de fonctions de $\mathcal{L}^+(\Omega)$ telle que

$$f_0(x) \leq f_1(x) \leq \ldots \leq f_n(x) \leq f_{n+1}(x) \leq \ldots \text{ pour tout } x \in \Omega\,,$$

et

$$\sup_{n\in\mathbf{N}} \int_\Omega f_n(x)dx < +\infty.$$

Soit $f : \Omega \to \mathbf{R} \cup \{+\infty\}$ définie par

$$f(x) := \lim_{n\to+\infty} f_n(x)\,, \quad \text{pour tout } x \in \Omega\,.$$

Alors $f \in \mathcal{L}^+(\Omega)$ et

$$\lim_{n\to+\infty} \int_\Omega f_n(x)dx = \int_\Omega f(x)dx\,.$$

Dans la mesure où on a construit la classe de fonctions $\mathcal{L}^+(\Omega)$ et défini une notion d'intégrale sur $\mathcal{L}^+(\Omega)$ qui prolonge l'intégrale usuelle définie sur $C_c(\Omega)$ et vérifie l'énoncé B) d'interversion intégrale/limite de la section 1 page 27, on pourrait croire que l'on est arrivé à la généralisation voulue de l'intégrale usuelle.

Ce n'est pourtant pas le cas. D'une part, on a vu dans l'Exemple 3.1.7 page précédente que la fonction $\mathbf{1}_{\mathbf{Q}\cap[0,1]} \notin \mathcal{L}^+(\Omega)$: le prolongement de l'intégrale à $\mathcal{L}^+(\Omega)$ ne suffit pas pour répondre complètement aux problèmes soulevés au début de ce chapitre dans la section 1 page 27.

D'autre part, si $f \in \mathcal{L}^+(\Omega)$, en général $-f \notin \mathcal{L}^+(\Omega)$. (Par exemple, la fonction $f : x \mapsto 1/\sqrt{x}$ appartient à $\mathcal{L}^+(]0,1[)$; mais $-f \notin \mathcal{L}^+(]0,1[)$ puisque $-f(x) \to -\infty$ lorsque $x \to 0^+$, de sorte que $-f$ ne peut être minorée par une fonction appartenant à $C_c(]0,1[)$.) Par conséquent, $\mathcal{L}^+(\Omega)$ n'est pas un $\mathbf{R}$-espace vectoriel. Le prolongement à $\mathcal{L}^+(\Omega)$ de l'intégrale usuelle n'est donc pas une forme linéaire définie sur un $\mathbf{R}$-espace vectoriel contenant $C_c(\Omega)$. Or pour des raisons

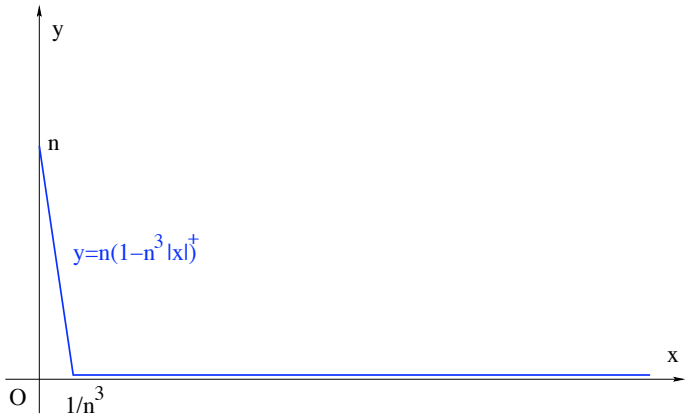


FIGURE III.2 – Graphe de la fonction $x \mapsto n(1-n^3|x|)^+$.

évidentes, il est indispensable que la nouvelle notion d'intégrale que l'on cherche à construire soit une opération linéaire.

La démonstration de la proposition ci-dessus est laissée au lecteur, car nous allons établir bientôt un résultat un peu plus général — cf. Proposition 3.2.5 page 39 ci-dessous.

3.2 Ensembles Lebesgue-négligeables

Le problème qui se pose à nous est de définir maintenant un $\mathbf{R}$-espace vectoriel à partir de $\mathcal{L}^+(\Omega)$.

On se heurte alors à la difficulté suivante : si $f \in \mathcal{L}^+(\Omega)$, alors f peut prendre la valeur $+\infty$. Toute expression de la forme $f-g$ avec $f, g \in \mathcal{L}^+(\Omega)$ risque donc de contenir la "forme indéterminée" $+\infty-\infty$. L'idée clef pour aller plus loin consiste à observer que ceci ne se produit que "très rarement", en un sens que nous allons expliciter.

Définition 3.2.1. Soit Ω ouvert de $\mathbf{R}^N$ et $\mathcal{N} \subset \Omega$. On dira que l'ensemble $\mathcal{N}$ est Lebesgue-négligeable s'il existe $f \in \mathcal{L}^+(\Omega)$ tel que $f(x) = +\infty$ pour tout $x \in \mathcal{N}$.

Voici un premier exemple d'ensemble Lebesgue-négligeable — nous en donnerons d'autres plus loin.

Exemple 3.2.2. Tout singleton de Ω est Lebesgue-négligeable.

Vérifions que $\{0\}$ est Lebesgue-négligeable dans $\mathbf{R}^N$. En effet, soit F définie par

$$F(x) := \sum_{n\geq 1} n(1-n^3|x|)^+; \quad F(x) = +\infty \Leftrightarrow x = 0\,.$$

Comme F est somme d'une série de fonctions positives ou nulles appartenant à $C_c(\mathbf{R}^N)$ et que

$$\sum_{n\geq 1} \int_{\mathbf{R}^N} n(1-n^3|x|)^+ dx = \sum_{n\geq 1} \frac{1}{n^2} < +\infty\,,$$

(voir Figure 3.2 page précédente), on a

$$F \in \mathcal{L}^+(\Omega)\,,$$

d'après la Proposition 3.1.9 page 36.

Définition 3.2.3. Une propriété $\mathcal{P}(x)$ dépendant du point $x \in \Omega$ est dite vraie "presque partout sur Ω" (ce que l'on abrège en "p.p. sur Ω") si

$$\{x \in \Omega \,|\, \mathcal{P}(x) \text{ est fausse } \} \text{ est Lebesgue-négligeable.}$$

La proposition $\mathcal{P}(x)$ peut-être par exemple "$f(x) \geq 0$", où f est une fonction définie sur Ω à valeurs dans $\mathbf{R}$, auquel cas on notera "$f(x) \geq 0$ p.p. en $x \in \Omega$" si l'ensemble des $x \in \Omega$ tels que $f(x) < 0$ est Lebesgue-négligeable.

De même, $\mathcal{P}(x)$ peut-être "$f_n(x) \to f(x)$ lorsque $n \to +\infty$, où $(f_n)_{n\geq 0}$ est une suite de fonctions définies sur Ω et f une fonction définie sur Ω, à valeurs dans $\mathbf{R}$ ou $\mathbf{C}$. Auquel cas on notera "$f_n(x) \to f(x)$ p.p. en $x \in \Omega$ lorsque $n \to +\infty$".

L'intérêt des ensembles Lebesgue-négligeables vient de ce que l'on peut modifier arbitrairement une fonction sur un tel ensemble sans en changer l'intégrale. Voici deux exemples de ce principe, que nous rencontrerons souvent dans la suite de ce chapitre.

D'abord, on peut généraliser le Lemme 3.1.3 page 33 en remplaçant l'hypothèse de convergence simple sur Ω (en tout point de Ω) par la convergence p.p. sur Ω.

Lemme 3.2.4. Soient $(f_n)_{n\geq 0}$ et $(g_n)_{n\geq 0}$, suites fondamentales sur Ω ouvert de $\mathbf{R}^N$. Alors

$$\lim_{n\to+\infty} f_n(x) \leq \lim_{n\to+\infty} g_n(x) \text{ p.p. en } x \in \Omega$$
$$\Rightarrow \lim_{n\to+\infty} \int_\Omega f_n(x)dx \leq \lim_{n\to+\infty} \int_\Omega g_n(x)dx\,.$$

Autrement dit, si $f, g \in \mathcal{L}^+(\Omega)$,

$$f(x) \leq g(x) \text{ p.p. en } x \in \Omega \Rightarrow \int_\Omega f(x)dx \leq \int_\Omega g(x)dx\,.$$

Ce lemme entraîne le Lemme 3.1.3 page 33 en échangeant les rôles des suites $(f_n)_{n\geq 0}$ et $(g_n)_{n\geq 0}$.

Démonstration. Il s'agit d'une variante plus technique de la preuve du Lemme 3.1.3 page 33.

Pour tout $x \in \Omega$, notons

$$f(x) := \lim_{n\to+\infty} f_n(x) \text{ et } g(x) := \lim_{n\to+\infty} g_n(x)\,.$$

Soit $(k_n)_{n\geq 0}$ suite fondamentale convergeant simplement vers $k \in \mathcal{L}^+(\Omega)$; quitte à changer k_n et k_n^+ pour tout $n \in \mathbf{N}$, on peut supposer que $k_n \geq 0$ sur Ω.

Définissons alors, pour tout $n \in \mathbf{N}$, la fonction $l_n := (h_n - k_n)^+$. Comme, pour tout $x \in \Omega$, la suite $(h_n(x))_{n\geq 0}$ est décroissante et la suite $(k_n(x))_{n\geq 0}$ croissante, on a

$$l_0(x) \geq l_1(x) \geq \ldots \geq l_n(x) \geq l_{n+1}(x) \to 0$$

pour tout $x \in \Omega$ lorsque $n \to +\infty$, car $k_n(x) \to +\infty$ pour tous les points $x \in \Omega$ tels que $h_n(x)$ ne tende pas vers 0 lorsque $n \to +\infty$. D'autre part, $l_n \in C_c(\Omega)$ pour tout $n \in \mathbf{N}$. D'après le Lemme 2.0.4 page 32 (basé sur le lemme de Dini),

$$\int_\Omega l_n(x)dx \to 0 \quad \text{ lorsque } n \to +\infty\,.$$

D'autre part,

$$f_m(x) \leq h_n(x) + g_n(x) \leq l_n(x) + k_n(x) + g_n(x) \quad \text{ pour tout } x \in \Omega$$

de sorte que

$$\int_\Omega f_m(x)dx \leq \lim_{n\to+\infty} \int_\Omega (l_n(x) + k_n(x) + g_n(x))dx \leq 0 + \epsilon + \lim_{n\to+\infty} \int_\Omega g_n(x)dx\,.$$

Comme $\epsilon > 0$ est arbitraire, il s'ensuit que

$$\int_\Omega f_m(x)dx \leq \lim_{n\to+\infty} \int_\Omega g_n(x)dx\,.$$

Comme ceci vaut pour tout $m \in \mathbf{N}$ on conclut en passant à la limite dans le membre de gauche de cette dernière inégalité pour $m \to +\infty$. □

Dans le même ordre d'idées, on peut aussi généraliser la Proposition 3.1.9 page 36 en y remplaçant l'hypothèse de convergence simple (partout) sur Ω par la convergence p.p. sur Ω.

Proposition 3.2.5 (Convergence monotone dans $\mathcal{L}^+(\Omega)$). Soit $(f_k)_{k\geq 0}$ suite de fonctions de $\mathcal{L}^+(\Omega)$ telle que

$$f_0(x) \leq f_1(x) \leq \ldots \leq f_k(x) \leq f_{k+1}(x) \leq \ldots \text{ p.p. en } x \in \Omega\,,$$

et

$$\sup_{k\in\mathbf{N}} \int_\Omega f_k(x)dx < +\infty.$$

Alors il existe $f \in \mathcal{L}^+(\Omega)$ tel que

$$f(x) := \lim_{k\to+\infty} f_k(x)\,, \quad \text{ p.p. en } x \in \Omega\,,$$

et

$$\lim_{k\to+\infty} \int_\Omega f_k(x)dx = \int_\Omega f(x)dx\,.$$

Démonstration. Soit $\mathcal{N} \subset \Omega$ négligeable t.q. la suite $(f_n(x))_{n\geq 0}$ soit croissante pour tout $x \in \Omega\backslash\mathcal{N}$. D'autre part, pour tout $k \in \mathbf{N}$, soit $(f_{k,n})_{n\in\mathbf{N}}$, suite fondamentale convergeant simplement vers f_k sur Ω.

Etape 1 : construction de f. Posons

$$F_n := \max(f_{0,n}, \dots, f_{n,n}) \qquad \text{pour tout } n \in \mathbf{N}\,.$$

Evidemment, $F_n \in C_c(\Omega)$; d'autre part la suite $(F_n(x))_{n\geq 0}$ est croissante pour tout $x \in \Omega$ et

$$F_n(x) \leq \max(f_0(x), \dots, f_n(x)) = f_n(x) \text{ pour tout } x \in \Omega \setminus \mathcal{N}\,.$$

D'après le Lemme 3.2.4 page 38,

$$\int_\Omega F_n(x) \leq \int_\Omega f_n(x)dx \leq \sup_{m\in\mathbf{N}} \int_\Omega f_m(x)dx < +\infty$$

de sorte que la suite $(F_n)_{n\geq 0}$ est fondamentale sur Ω. Posons

$$f(x) := \lim_{n\to+\infty} F_n(x) \text{ pour tout } x \in \Omega\,;$$

par construction $f \in \mathcal{L}^+(\Omega)$.

Etape 2 : convergence de la suite $(f_n)_{n\geq 0}$. D'une part, on a vu dans l'étape 1 que

$$F_n(x) \leq f_n(x) \quad \text{ pour tout } x \in \Omega \setminus \mathcal{N}\,;$$

d'autre part, par construction, pour tout $n \geq k \geq 0$, on a

$$f_{k,n}(x) \leq F_n(x) \quad \text{ pour tout } x \in \Omega\,.$$

En passant à la limite dans ces deux inégalités pour $n \to +\infty$, on trouve que

$$f_k(x) \leq f(x) \leq \lim_{n\to+\infty} f_n(x) \quad \text{ pour tout } x \in \Omega \setminus \mathcal{N}\,;$$

puis en passant à la limite lorsque $k \to +\infty$ dans l'inégalité ci-dessus, on conclut que

$$f_k(x) \to f(x) \text{ pour tout } x \in \Omega \setminus \mathcal{N} \text{ lorsque } k \to +\infty\,.$$

Etape 3 : convergence des intégrales. D'après l'étape 2, pour tout $n \geq k \geq 0$

$$f_{k,n}(x) \leq F_n(x) \leq f_n(x) \text{ pour tout } x \in \Omega \setminus \mathcal{N}\,,$$

de sorte que, d'après le Lemme 3.2.4 page 38

$$\int_\Omega f_{k,n}(x)dx \leq \int_\Omega F_n(x)dx \leq \int_\Omega f_n(x)dx\,.$$

Passons à la limite lorsque $n \to +\infty$: comme les suites $(f_{k,n})_{n\geq 0}$ et $(F_n)_{n\geq 0}$ sont fondamentales sur Ω et que

$$\text{la suite } \left(\int_\Omega f_n(x)dx\right)_{n\geq 0} \text{ est croissante et bornée supérieurement,}$$

on trouve que

$$\int_\Omega f_k(x)dx \leq \int_\Omega f(x)dx \leq \lim_{n\to+\infty} \int_\Omega f_n(x)dx$$

d'où le résultat en passant à la limite dans cette dernière inégalité pour $k \to +\infty$. □

Une conséquence extrêmement importante de la convergence monotone dans $\mathcal{L}^+(\Omega)$ est la

Proposition 3.2.6. Toute réunion dénombrable de parties Lebesgue-négligeables est Lebesgue-négligeable.

Démonstration. Soit $(\mathcal{N}_k)_{k\in\mathbf{N}}$, parties négligeables de Ω. Pour chaque $k \in \mathbf{N}$, il existe $f_k \in \mathcal{L}^+(\Omega)$ t.q. $f_k(x) = +\infty$ pour tout $x \in \mathcal{N}_k$. Evidemment $f_k^+ \in \mathcal{L}^+(\Omega)$ pour tout $k \in \mathbf{N}$, d'après la Proposition 3.1.8 page 36 3). Quitte à multiplier f_k^+ par une constante positive bien choisie, on peut supposer que

$$\int_\Omega f_k(x)^+ dx \leq 2^{-k}\,, \quad k \in \mathbf{N}\,.$$

Posons, pour tout $x \in \Omega$ et tout $n \in \mathbf{N}$,

$$F_n(x) = \sum_{k=0}^{n} f_k(x)^+\,.$$

Evidemment $F_n \in \mathcal{L}^+(\Omega)$ pour tout $n \in \mathbf{N}$, et la suite $(F_n(x))_{n\geq 0}$ est croissante pour tout $x \in \Omega$ (comme somme partielle d'une série à termes positifs). D'autre part,

$$\int_\Omega F_n(x)dx = \sum_{k=0}^{n} \int_\Omega f_k(x)^+ dx \leq \sum_{k\geq 0} 2^{-k} = 2\,,$$

pour tout $n \in \mathbf{N}$. Par convergence monotone

$$F(x) := \sum_{k\geq 0} f_k(x)^+ \text{ définit une fonction } F \in \mathcal{L}^+(\Omega)$$

et

$$x \in \bigcup_{k\geq 0} \mathcal{N}_k \Rightarrow F(x) = +\infty\,.$$

Donc

$$\bigcup_{k\geq 0} \mathcal{N}_k \text{ est négligeable.}$$

□

NB. L'hypothèse de dénombrabilité est fondamentale : une réunion non dénom- brable de parties négligeables de Ω n'est pas forcément négligeable en général. (Le lecteur est invité à donner un contre-exemple à titre d'exercice.)

Exemple 3.2.7. Toute partie dénombrable $\mathcal{D}$ de $\mathbf{R}^N$ est Lebesgue-négligeable. En effet,

$$\mathcal{D} = \bigcup_{x\in\mathcal{D}} \{x\}$$

est réunion dénombrable de singletons, qui sont Lebesgue-négligeables (Exemple 3.2.2 page 37.) D'après la Proposition 3.2.6, $\mathcal{D}$ est Lebesgue-négligeable.

Ainsi $\mathbf{Q}$ est Lebesgue-négligeable dans $\mathbf{R}$ et, plus généralement, $\mathbf{Q}^N$ est Lebesgue-négligeable dans $\mathbf{R}^N$ pour tout $N \geq 1$.

Exemple 3.2.8. Un ouvert non vide $\Omega \subset \mathbf{R}^N$ n'est pas Lebesgue-négligeable dans $\mathbf{R}^N$.

En effet, si tel était le cas, il existerait une suite fondamentale $(g_n)_{n\geq 0}$ telle que $g_n(x) \to +\infty$ pour tout $x \in \Omega$. Soit $f \in C_c(\Omega)$ et $k >\geq 1$; comme on a

$$k|f(x)| \leq \lim_{n\to+\infty} g_n(x) = +\infty \text{ pour tout } x \in \Omega$$

on déduirait du Lemme 3.2.4 page 38 que

$$k \int_\Omega |f(x)|dx \leq \lim_{n\to+\infty} \int_\Omega g_n(x)dx < +\infty\,.$$

En passant à la limite pour $k \to +\infty$, on aurait

$$\int_\Omega |f(x)|dx = 0 \quad \text{d'où } f(x) = 0 \text{ pour tout } x \in \Omega\,.$$

Comme f est arbitraire dans $C_c(\Omega)$, ceci entraînerait que $C_c(\Omega) = \{0\}$, ce qui est impossible puisque $\Omega \neq \varnothing$.

En rassemblant les divers résultats ci-dessus concernant les ensembles Lebesgue-négligeables, on retrouve très facilement l'énoncé suivant.

Proposition 3.2.9. Tout ouvert non vide de $\mathbf{R}^N$ (avec $N \geq 1$) est non dénombrable.

Démonstration. Soit Ω ouvert non vide de $\mathbf{R}$; si Ω était dénombrable, il serait Lebesgue-négligeable (Exemple 3.2.7 page précédente). Or c'est impossible d'après l'Exemple 3.2.8. □

On espère que la simplicité de cette preuve, comparée à l'argument diagonal particulièrement astucieux de Cantor, convaincra le lecteur de la puissance de la théorie de l'intégration de Lebesgue — et tout particulièrement de la convergence monotone qui en est le cœur.

Nous reviendrons plus loin sur cette notion d'ensemble négligeable, mais nous en savons maintenant assez pour achever la construction de l'intégrale de Lebesgue.

3.3 Etape 2 : fonctions sommables

Cette deuxième et dernière étape de la construction de l'intégrale de Lebesgue consiste à passer de la notion d'intégrale sur la classe $\mathcal{L}^+(\Omega)$ à une forme linéaire sur un $\mathbf{R}$-espace vectoriel contenant $\mathcal{L}^+(\Omega)$ prolongeant cette intégrale.

Définition 3.3.1. Une fonction f définie p.p. sur Ω et à valeurs dans $\mathbf{R}$ est sommable (ou Lebesgue-intégrable) s'il existe $g, h \in \mathcal{L}^+(\Omega)$ t.q.

$$f(x) = g(x) - h(x) \text{ p.p. en } x \in \Omega\,.$$

On définit alors l'intégrale de Lebesgue de f sur Ω par la formule

$$\int_\Omega f(x)dx = \int_\Omega g(x)dx - \int_\Omega h(x)dx\,.$$

On note $\mathcal{L}^1(\Omega)$ l'ensemble des fonctions sommables sur Ω à valeurs réelles.

Exemple 3.3.2. La fonction $\mathbf{1_Q}$ appartient à $\mathcal{L}^1(\mathbf{R})$ puisque $\mathbf{1_Q}(x) = 0$ p.p. en $x \in \mathbf{R}$.

Vérifions que cette définition de l'intégrale d'une fonction sommable f est indépendante du choix de la décomposition $f = g - h$ avec $g, h \in \mathcal{L}^+(\Omega)$. En effet, supposons que $f \in \mathcal{L}^1(\Omega)$ satisfait

$$f(x) = g_1(x) - h_1(x) = g_2(x) - h_2(x) \text{ p.p. en } x \in \Omega\,,$$

avec $g_1, g_2, h_1, h_2 \in \mathcal{L}^+(\Omega)$. Alors

$$g_1(x) + h_2(x) = g_2(x) + h_1(x) \text{ p.p. en } x \in \Omega\,,$$

et comme $g_1 + h_2$ et $g_2 + h_1$ appartiennent à $\mathcal{L}^+(\Omega)$ d'après la Proposition 3.1.8 page 36 2), on déduit du Lemme 3.2.4 page 38 que

$$\begin{aligned}\int_\Omega g_1(x)dx + \int_\Omega h_2(x)dx &= \int_\Omega (g_1 + h_2)(x)dx \\ &= \int_\Omega (g_2 + h_1)(x)dx = \int_\Omega g_2(x)dx + \int_\Omega h_1(x)dx\,.\end{aligned}$$

d'où

$$\int_\Omega g_1(x)dx - \int_\Omega h_1(x)dx = \int_\Omega g_2(x)dx - \int_\Omega h_2(x)dx\,,$$

ce qui montre que la valeur de l'intégrale de f sur Ω ne dépend pas de la décomposition $g_1 - h_1$ ou $g_2 - h_2$ choisie pour la définir.

Précisons un peu la décomposition de fonctions sommables en différence de fonctions de la classe $\mathcal{L}^+$.

Lemme 3.3.3. Soient Ω ouvert de $\mathbf{R}^N$ et $f \in \mathcal{L}^1(\Omega)$. Pour tout $\epsilon > 0$, il existe $u, v \in \mathcal{L}^+(\Omega)$ telles que

$$f(x) = u(x) - v(x) \text{ p.p. en } x \in \Omega$$

avec

$$v(x) \geq 0 \text{ pour tout } x \in \Omega\,, \quad \text{ et } \int_\Omega v(x)dx < \epsilon\,.$$

Démonstration. Comme $f \in \mathcal{L}^1(\Omega)$, il existe $g, h \in \mathcal{L}^+(\Omega)$ telles que $f(x) = g(x) - h(x)$ p.p. sur Ω. Soient $(g_n)_{n\geq 0}$ et $(h_n)_{n\geq 0}$, suites fondamentales convergeant simplement vers g et h sur Ω. En particulier,

$$\int_\Omega h_n(x)dx \to \int_\Omega h(x)dx$$

lorsque $n \to +\infty$. Il existe donc $m > 0$ tel que

$$n \geq m \Rightarrow \int_\Omega (h(x) - h_n(x))dx < \epsilon\,.$$

Posons alors $u = g + h_m$ et $v = h - h_m$; évidemment $u \in \mathcal{L}^+(\Omega)$ et

$$v(x) \geq 0 \text{ pour tout } x \in \Omega\,, \quad \text{ tandis que } \int_\Omega v(x)dx < \epsilon\,.$$

Enfin, la suite $(v_n)_{n\geq 0}$ définie par $v_n(x) = h_n(x) - h_m(x)$ est évidemment fondamentale sur Ω puisque,

$$\begin{aligned} v_0(x) = h_0(x) - h_m(x) \leq \ldots \leq v_n(x) = h_n(x) - h_m(x) \\ \leq v_{n+1}(x) = h_{n+1}(x) - h_m(x) \leq \ldots \end{aligned}$$

et que, pour tout $n \geq 0$, l'on a

$$\int_\Omega v_n(x)dx = \int_\Omega h_n(x)dx - \int_\Omega h_m(x)dx \leq \int_\Omega h(x)dx - \int_\Omega h_m(x)dx < +\infty\,.$$

Comme $v_n \to v$ simplement sur Ω, on en déduit que $v \in \mathcal{L}^+(\Omega)$. Enfin $f(x) = u(x) - v(x)$ p.p. en $x \in \Omega$. □

Ce lemme prouve que, bien que la classe $\mathcal{L}^+(\Omega)$ ne soit pas le terme de la construction de l'intégrale de Lebesgue, les fonctions sommables sur Ω peuvent être approchées d'aussi près que l'on veut par des fonctions de $\mathcal{L}^+(\Omega)$. Nous reviendrons plus loin sur ce résultat.

Dorénavant, nous aurons à effectuer des opérations élémentaires sur des fonctions définies seulement p.p. sur Ω.

Opérations sur les fonctions définies p.p.

Soient $(f_n)_{n\in\mathcal{D}}$ une suite indexée par un ensemble dénombrable $\mathcal{D}$ de fonctions définies p.p. sur Ω, et une fonction Φ définie sur $\mathbf{C}^{\mathcal{D}}$ à valeurs dans $\mathbf{C}^{\mathcal{D}}$. La fonction $x \mapsto \Phi((f_n(x))_{n\in\mathcal{D}})$ est **définie p.p.** sur Ω. En effet, par hypothèse, pour tout $n \in \mathcal{D}$, il existe $\mathcal{N}_n$ Lebesgue-négligeable dans Ω tel que la fonction f_n soit définie sur $\Omega \setminus \mathcal{N}_n$, de sorte que la fonction $x \mapsto \Phi((f_n(x))_{n\in\mathcal{D}})$ est définie sur

$$\Omega \setminus \bigcup_{n\in\mathcal{D}} \mathcal{N}_n$$

c'est à dire p.p. sur Ω, puisque

$$\bigcup_{n\in\mathcal{D}} \mathcal{N}_n \text{ est Lebesgue-négligeable}$$

en tant que réunion dénombrable d'ensembles Lebesgue-négligeables. Par exemple, on peut avoir $\mathcal{D} = \{1, 2\}$ et $\Phi(f_1, f_2) = f_1 f_2$ ou $\Phi(f_1, f_2) = f_1 + f_2$. Ainsi les opérations élémentaires mettant en jeu une famille dénombrable de fonctions mesurables définies p.p. sur Ω fournissent une fonction définie p.p. sur Ω.

Voici les premières propriétés de cette nouvelle notion d'intégrale.

Proposition 3.3.4 (Propriétés de l'intégrale de Lebesgue). Soit Ω ouvert de $\mathbf{R}^N$. Alors

1) l'ensemble $\mathcal{L}^1(\Omega)$ muni de l'addition des fonctions définies p.p. sur Ω et de la multiplication par les scalaires réels est un $\mathbf{R}$-espace vectoriel et

$$\mathcal{L}^1(\Omega) \ni f \mapsto \int_\Omega f(x)dx$$

une forme $\mathbf{R}$-linéaire ;

2) on a les inclusions $C_c(\Omega) \subset \mathcal{L}^+(\Omega) \subset \mathcal{L}^1(\Omega)$ et l'intégrale de Lebesgue définie sur $\mathcal{L}^1(\Omega)$ coïncide avec l'intégrale usuelle sur $C_c(\Omega)$;

3) si $f, g \in \mathcal{L}^1(\Omega)$ et $f(x) \geq g(x)$ p.p. en $x \in \Omega$, alors

$$\int_\Omega f(x)dx \geq \int_\Omega g(x)dx\,;$$

4) en particulier, pour tout $f \in \mathcal{L}^1(\Omega)$, on a $|f| \in \mathcal{L}^1(\Omega)$ et

$$\left|\int_\Omega f(x)dx\right| \leq \int_\Omega |f(x)|dx\,;$$

5) si $f, g \in \mathcal{L}^1(\Omega)$, alors les fonctions $x \mapsto \max(f(x), g(x))$ et $x \mapsto \min(f(x), g(x))$, définies p.p. sur Ω, appartiennent à $\mathcal{L}^1(\Omega)$.

Démonstration. Les énoncés 1) et 2) sont triviaux, ainsi que l'énoncé 3), qui découle de la définition des fonctions sommables et du Lemme 3.2.4 page 38.

Quant à l'énoncé 4), écrivons que $f(x) = g(x) - h(x)$ p.p. en $x \in \Omega$, avec $g, h \in \mathcal{L}^+(\Omega)$. Alors

$$|f(x)| = |g(x) - h(x)| = \max(g(x), h(x)) - \min(g(x), h(x))$$

Comme g et h appartiennent à $\mathcal{L}^+(\Omega)$, les fonction $x \mapsto \max(g(x), h(x))$ et $x \mapsto \min(g(x), h(x))$ appartiennent à $\mathcal{L}^+(\Omega)$ (voir Proposition 3.1.8 page 36 3) ; et donc $|f| \in \mathcal{L}^1(\Omega)$. Puis $f(x) \leq |f(x)|$ p.p. en $x \in \Omega$, de sorte que, d'après l'énoncé 3)

$$\int_\Omega f(x)dx \leq \int_\Omega |f(x)|dx\,.$$

En appliquant cette inégalité à $-f$, on trouve que

$$\left|\int_\Omega f(x)dx\right| = \sup\left(\int_\Omega f(x)dx, -\int_\Omega f(x)dx\right) \leq \int_\Omega |f(x)|dx\,.$$

L'énoncé 5) découle du 4) en observant que

$$\begin{aligned}\max(f, g) &= g + (f - g)^+ = g + \tfrac{1}{2}|f - g| + \tfrac{1}{2}(f - g)\,,\\ \min(f, g) &= g + (f - g)^- = g + \tfrac{1}{2}|f - g| - \tfrac{1}{2}(f - g)\,,\end{aligned}$$

□

On aura évidemment besoin d'intégrer les fonctions à valeurs complexes.

Définition 3.3.5. On dit qu'une fonction f définie p.p. sur Ω à valeurs dans $\mathbf{C}$ est sommable si les fonctions à valeurs réelles $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ sont sommables ; auquel cas, on pose

$$\int_\Omega f(x)dx := \int_\Omega \mathrm{Re}(f)(x)dx + i\int_\Omega \mathrm{Im}(f)(x)dx\,.$$

On note $\mathcal{L}^1(\Omega; \mathbf{C})$ l'ensemble des fonctions sommables sur Ω à valeurs complexes.

Voici quelques propriétés de l'intégration des fonctions sommables à valeurs complexes.

Proposition 3.3.6 (Intégration des fonctions à valeurs complexes). Soit Ω ouvert de $\mathbf{R}^N$. Alors

1) l'ensemble $\mathcal{L}^1(\Omega;\mathbf{C})$ muni de l'addition des fonctions à valeurs complexes définies p.p. sur Ω et de la multiplication par les scalaires complexes est un $\mathbf{C}$-espace vectoriel et

$$\mathcal{L}^1(\Omega;\mathbf{C}) \ni f \mapsto \int_\Omega f(x)dx$$

une forme $\mathbf{C}$-linéaire ;

2) on a $C_c(\Omega;\mathbf{C}) \subset \mathcal{L}^1(\Omega;\mathbf{C})$, et l'intégrale de Lebesgue coïncide avec l'intégrale usuelle sur $C_c(\Omega;\mathbf{C})$;

3) pour tout $f \in \mathcal{L}^1(\Omega;\mathbf{C})$, on a $|f| \in \mathcal{L}^1(\Omega;\mathbf{R})$ et

$$\left|\int_\Omega f(x)dx\right| \leq \int_\Omega |f(x)|dx\,.$$

Démonstration. Les énoncés 1) et 2) sont triviaux.

Quant à l'énoncé 3), choisissons $\omega \in \mathbf{C}$ tel que $|\omega| = 1$ et

$$\omega \int_\Omega f(x)dx = \left|\int_\Omega f(x)dx\right|\,.$$

Alors

$$\begin{aligned}\left|\int_\Omega f(x)dx\right| = \omega\int_\Omega f(x)dx &= \int_\Omega \mathrm{Re}(\omega f(x))dx + i\int_\Omega \mathrm{Im}(\omega f(x))dx\\ &= \int_\Omega \mathrm{Re}(\omega f(x))dx\,.\end{aligned}$$

Puis, d'après les énoncés 3) et 4) de la Proposition 3.3.4 page 44,

$$\int_\Omega \mathrm{Re}(\omega f(x))dx \leq \int_\Omega |\,\mathrm{Re}(\omega f(x))|dx \leq \int_\Omega |\omega f(x)|dx = \int_\Omega |f(x)|dx\,,$$

d'où le résultat. □

Remarque 3.3.7. On prendra bien garde au fait suivant : une fonction f n'est sommable — c'est-à-dire intégrable au sens de Lebesgue — que si son module (ou sa valeur absolue, si f est à valeurs réelles) est également intégrable. Dans des formules du type

$$\int_0^{+\infty} \frac{\sin x}{x}dx = \frac{\pi}{2}\,,$$

le membre de gauche n'est pas une intégrale de Lebesgue, et la fonction

$$x \mapsto \frac{\sin x}{x}$$

n'est pas sommable sur $\mathbf{R}_+$, puisque

$$\lim_{R\to+\infty}\int_0^R \left|\frac{\sin x}{x}\right| dx = +\infty\,.$$

On rappelle que la formule

$$\int_0^{+\infty} \frac{\sin x}{x} dx = \frac{\pi}{2}\,,$$

signifie simplement que la fonction de la variable $R > 0$ définie par

$$R \mapsto \int_0^R \frac{\sin x}{x} dx$$

vérifie

$$\lim_{R\to+\infty} \int_0^R \frac{\sin x}{x} dx = \frac{\pi}{2}\,.$$

La limite au membre de gauche ci-dessus est notée

$$\int_0^{+\infty} \frac{\sin x}{x} dx\,,$$

ce qui introduit une confusion avec les notations

$$\int_{\mathbf{R}_+} f(x)dx \quad \text{ou} \int_0^{+\infty} f(x)dx$$

que l'on utilise indifféremment pour désigner l'intégrale de Lebesgue de f sur $\mathbf{R}_+$ lorsque f est une fonction sommable sur $\mathbf{R}_+$.

Il est regrettable que l'on utilise ainsi la même notation pour désigner deux objets aussi différents mathématiquement ; malheureusement, tous les mathémati- ciens du monde pratiquent cet abus de notation depuis un siècle, de sorte qu'il n'est pas envisageable de faire autrement.

3.4 Fonctions sommables et ensembles Lebesgue-négligeables.

Terminons cette présentation de l'intégrale de Lebesgue avec quelques remarques sur la relation entre les ensembles négligeables et les fonctions sommables.

Une fonction sommable sur Ω n'est définie que p.p. sur Ω, pourtant son intégrale est bien définie de manière unique. Autrement dit, modifier une fonction sommable sur un ensemble négligeable ne change pas son intégrale. Ainsi, avec la définition de Lebesgue de l'intégrale

$$\phi(x) = 0 \text{ p.p. sur } \Omega \Rightarrow \int_\Omega \phi(x)dx = 0\,.$$

(Il suffit de décomposer ϕ sous la forme $\phi(x) = g(x) - g(x)$ p.p. en $x \in \Omega$.)

Exemple : comme $\mathbf{Q}$ est négligeable dans $\mathbf{R}$, la fonction $\mathbf{1}_\mathbf{Q}(x) = 0$ p.p. en $x \in \mathbf{R}$, de sorte que la définition de l'intégrale de Lebesgue montre que

$$\int_\mathbf{R} \mathbf{1}_{\mathbf{Q}\cap[0,1]}(x)dx = 0\,.$$

Cette définition confirme le calcul de cette intégrale effectué dans l'introduction de ce chapitre par intégration terme à terme d'une série de fonctions positives continues par morceaux.

Cet énoncé admet une sorte de réciproque extrêmement importante pour la suite de la théorie de Lebesgue de l'intégrale.

Théorème 3.4.1. Soit Ω ouvert de $\mathbf{R}^N$. Si $f \in \mathcal{L}^1(\Omega; \mathbf{C})$, alors

$$\int_\Omega |f(x)|dx = 0 \Rightarrow f(x) = 0 \text{ p.p. en } x \in \Omega\,.$$

On comparera cette énoncé avec celui que l'on connaît dans le cas de l'intégrale usuelle des fonctions continues : si $f \in C([a,b]; \mathbf{C})$

$$\int_a^b |f(x)|dx = 0 \Rightarrow f(x) = 0 \text{ pour tout } x \in [a,b]\,.$$

Démonstration. Comme $f \in \mathcal{L}^1(\Omega; \mathbf{C})$, son module $|f| \in \mathcal{L}^1(\Omega)$. D'après le Lemme 3.3.3 page 43, pour tout $\epsilon > 0$, il existe $u, v \in \mathcal{L}^+(\Omega)$ telles que

$$|f| = u - v \text{ p.p. sur } \Omega\,, \text{ avec } v \geq 0 \text{ sur } \Omega \text{ et } \int_\Omega v(x)dx < \epsilon\,.$$

Donc u vérifie

$$u = |f| + v \geq 0 \text{ p.p. sur } \Omega \quad \text{ et } \int_\Omega u(x)dx = \int_\Omega v(x)dx < \epsilon\,.$$

Effectuons la décomposition ci-dessus pour $\epsilon = 2^{-n}$ lorsque n décrit $\mathbf{N}$: on construit ainsi une suite $(u_n)_{n\geq 0}$ telle que, pour tout $n \in \mathbf{N}$, l'on ait

$$u_n \in \mathcal{L}^+(\Omega)\,, \quad u_n \geq |f(x)| \geq 0 \text{ p.p. sur } \Omega \quad \text{ et } \int_\Omega u_n(x)dx < 2^{-n}\,.$$

Par convergence monotone

$$U = \sum_{n\geq 0} 2^{n/2} u_n \in \mathcal{L}^+(\Omega)$$

puisque

$$\sum_{n\geq 0} \int_\Omega 2^{n/2} u_n(x)dx < \sum_{n\geq 0} 2^{-n/2} = \frac{\sqrt{2}}{\sqrt{2}-1} < +\infty\,.$$

Or, pour tout $n \geq 0$, on a $u_n(x) \geq |f(x)|$ p.p. en $x \in \Omega$, de sorte que

$$U(x) \geq |f(x)| \sum_{n\geq 0} 2^{n/2} = +\infty \quad \text{ dès que } |f(x)| > 0\,,$$

d'où la conclusion. □

Par construction, une fonction sommable peut prendre des valeurs infinies. Il suffit en effet de considérer le cas d'une fonction $f \in \mathcal{L}^+(\Omega)$, qui peut prendre la valeur $+\infty$, ou celui de la fonction $-f$ qui peut, elle, prendre la valeur $-\infty$. Mais ceci n'arrive que rarement.

Proposition 3.4.2. Soient Ω ouvert de $\mathbf{R}^N$ et $f \in \mathcal{L}^1(\Omega; \mathbf{C})$. Alors $|f(x)| < +\infty$ p.p. en $x \in \Omega$.

On a déjà vu plus haut (Exemple 3.2.8 page 42) qu'aucun ouvert non vide de $\mathbf{R}^N$ n'est Lebesgue-négligeable. Il n'existe donc pas de fonction $\phi \in \mathcal{L}^+(\Omega)$ telle que $\phi(x) = +\infty$ pour tout $x \in \Omega$. La proposition ci-dessus généralise ce résultat.

Démonstration. Comme $f \in \mathcal{L}^1(\Omega;\mathbf{C})$, son module $|f| \in \mathcal{L}^1(\Omega)$. Il existe donc $u, v \in \mathcal{L}^+(\Omega)$ telles que $|f| = u - v$ p.p. sur Ω. Soit $(v_n)_{n\geq 0}$ suite fondamentale convergeant simplement vers v sur Ω. Evidemment, $v \geq v_0$ sur Ω, de sorte que $u = |f| + v \geq |f| + v_0$ p.p. sur Ω, c'est-à-dire qu'il existe $\mathcal{N} \subset \Omega$ Lebesgue-négligeable t.q.

$$u(x) \geq |f(x)| + v_0(x) \qquad \text{pour } x \in \Omega \setminus \mathcal{N}\,.$$

Comme la fonction $v_0 \in C_c(\Omega)$, elle est bornée sur Ω, de sorte que

$$\{x \in \Omega \setminus \mathcal{N} \,|\, |f(x)| = +\infty\} \subset \{x \in \Omega \,|\, u(x) = +\infty\}$$

qui est Lebesgue-négligeable par définition. □

4 Le théorème de convergence dominée

L'intégrale de Lebesgue que nous venons de construire conduit à des résultats d'interversion intégrale/limite pour des suites de fonctions convergeant simplement (ou même presque partout) qui sont d'une facilité d'utilisation considérable. En particulier, ces résultats n'exigent aucune hypothèse sur la fonction limite — mal connue en pratique dans la plupart des cas. Le résultat le plus utile est le théorème de convergence dominée de Lebesgue, qui s'énonce comme suit.

Théorème 4.0.3 (Convergence dominée). Soient $\Omega \subset \mathbf{R}^N$ ouvert et une suite $(f_n)_{n\in\mathbf{N}}$ de fonctions $\in \mathcal{L}^1(\Omega;\mathbf{C})$. Supposons que

$$f_n(x) \to f(x) \text{ p.p. en } x \in \Omega$$

et qu'il existe $F \in \mathcal{L}^1(\Omega;\mathbf{R})$ t.q.

$$|f_n(x)| \leq F(x) \text{ p.p. en } x \in \Omega\,.$$

Alors la fonction limite $f \in \mathcal{L}^1(\Omega;\mathbf{C})$ et on a

$$\int_\Omega f_n(x)dx \to \int_\Omega f(x)dx$$

lorsque $n \to +\infty$

La démonstration du théorème de convergence dominée s'articule en trois étapes, dont les deux premières correspondent à des énoncés importants pour certaines applications.

Comme la monotonie est à la base de la construction de l'intégrale de Lebesgue, commençons un énoncé de convergence monotone dans le cadre des fonctions sommables.

Proposition 4.0.4 (Convergence monotone dans $\mathcal{L}^1$). Soit $(f_n)_{n\geq 0}$ suite de fonctions de $\mathcal{L}^1(\Omega)$ telle que

a) la suite de nombres $(f_n(x))_{n\geq 0}$ est croissante p.p. en $x \in \Omega$, et

b) la suite des intégrales des fonctions f_n vérifie

$$\sup_{n\geq 0} \int_\Omega f_n(x)dx < +\infty\,.$$

Alors il existe $f \in \mathcal{L}^1(\Omega)$ telle que $f_n(x) \to f(x)$ p.p. en $x \in \Omega$ lorsque $n \to +\infty$, et

$$\int_\Omega f_n(x)dx \to \int_\Omega f(x)dx\,.$$

Démonstration. Pour tout $n \in \mathbf{N}$, posons $u_n(x) := f_{n+1}(x) - f_n(x) \geq 0$ p.p. en $x \in \Omega$; comme $u_n \in \mathcal{L}^1(\Omega)$ pour tout $n \in \mathbf{N}$, il existe, d'après le Lemme 3.3.3 page 43 une décomposition de u_n sous la forme $u_n(x) = v_n(x) - w_n(x)$ p.p. en $x \in \Omega$ avec $v_n, w_n \in \mathcal{L}^+(\Omega)$ et vérifiant en outre

$$w_n(x) \geq 0 \text{ pour tout } x \in \Omega\,, \quad \text{et } \int_\Omega w_n(x)dx < 2^{-n}\,.$$

On a donc $v_n(x) = u_n(x) + w_n(x) \geq 0$ p.p. en $x \in \Omega$ pour tout $n \geq 0$; d'autre part

$$\begin{cases} \displaystyle\sum_{k=0}^{n} \int_\Omega w_k(x)dx \leq 2\,, \\ \displaystyle\sum_{k=0}^{n} \int_\Omega v_k(x)dx \leq 2 - \int_\Omega f_0(x)dx + \sup_{n\geq 0} \int_\Omega f_n(x)dx < +\infty\,. \end{cases}$$

D'après la Proposition 3.2.5 page 39,

$$\sum_{k=0}^{n} v_k(x) \to V(x) \quad \text{ et } \sum_{k=0}^{n} w_k(x) \to W(x) \text{ p.p. en } x \in \Omega$$

avec $V, W \in \mathcal{L}^+(\Omega)$ et

$$\begin{cases} \displaystyle\int_\Omega V(x)dx = \sum_{n\geq 0} \int_\Omega v_n(x)dx\,, \\ \displaystyle\int_\Omega W(x)dx = \sum_{n\geq 0} \int_\Omega w_n(x)dx\,. \end{cases}$$

Définissons

$$f := f_0 + V - W \quad \text{ p.p. sur } \Omega\,;$$

comme $f_0 \in \mathcal{L}^1(\Omega)$ et $V, W \in \mathcal{L}^+(\Omega)$, il s'ensuit que $f \in \mathcal{L}^1(\Omega)$.

D'autre part

$$\begin{aligned} f_n(x) &= f_0(x) + \sum_{k=0}^{n} v_k(x) - \sum_{k=0}^{n} w_k(x) \\ &\to f_0(x) + V(x) - W(x) = f(x) \text{ p.p. en } x \in \Omega \end{aligned}$$

lorsque $n \to +\infty$.

Enfin

$$\begin{aligned} \int_\Omega f_n(x)dx &= \int_\Omega f_0(x)dx + \sum_{k=0}^{n} \int_\Omega v_k(x) - \sum_{k=0}^{n} \int_\Omega w_k(x)dx \\ &\to \int_\Omega f_0(x)dx + \int_\Omega V(x)dx - \int_\Omega W(x)dx = \int_\Omega f(x)dx \end{aligned}$$

lorsque $n \to +\infty$, ce qu'il fallait démontrer. □

Si on abandonne l'hypothèse de monotonie sur la suite de fonctions considérées, l'interversion limite/intégrale n'est plus légitime en toute généralité. Toutefois, on dispose d'une inégalité, fort utile dans la pratique.

Lemme 4.0.5 (de Fatou). Soit $(f_n)_{n\geq 0}$ suite de fonctions de $\mathcal{L}^1(\Omega)$ telle que

a) pour tout $n \in \mathbf{N}$, on a $f_n(x) \geq 0$ p.p. en $x \in \Omega$, et

b) la suite des intégrales des fonctions f_n vérifie

$$\sup_{n\geq 0} \int_\Omega f_n(x)dx < +\infty\,.$$

Alors la fonction

$$\varliminf_{n\to+\infty} f_n \in \mathcal{L}^1(\Omega)$$

et on a

$$\int_\Omega \varliminf_{n\to+\infty} f_n(x)dx \leq \varliminf_{n\to+\infty} \int_\Omega f_n(x)dx\,.$$

Démonstration. Pour $m \geq k \geq 0$, on pose $f_{k,m} := \min(f_k, \ldots, f_m)$, qui est une fonction définie p.p. sur Ω. Evidemment, $f_{k,m} \in \mathcal{L}^1(\Omega)$, d'après la Proposition 3.3.4 page 44 5). Fixons k et faisons $m \to +\infty$: on voit alors que, pour tout $k \geq 0$ fixé,

$$\begin{cases} f_{k,m}(x) \to \phi_k(x) := \inf\limits_{n\geq k} f_n(x) \text{ p.p. en } x \in \Omega \text{ lorsque } m \to +\infty\,, \\ \text{la suite } (f_k(x) - f_{k,m}(x))_{m\geq k} \text{ est croissante p.p. en } x \in \Omega\,, \\ \int_\Omega (f_k - f_{k,m})(x)dx \leq \int_\Omega f_k(x)dx < +\infty\,. \end{cases}$$

D'après la Proposition 4.0.4 page 49 appliquée à la suite $(f_k - f_{k,m})_{m\geq k}$, la fonction $\phi_k \in \mathcal{L}^1(\Omega)$ et

$$\begin{aligned} \int_\Omega \phi_k(x)dx &= \lim_{m\to+\infty} \int_\Omega f_{k,m}(x)dx \\ &\leq \lim_{m\to+\infty} \min_{k\leq n\leq m} \int_\Omega f_n(x)dx = \inf_{n\geq k} \int_\Omega f_n(x)dx\,. \end{aligned}$$

D'autre part,

$$\begin{cases} \phi_k(x) \to \varliminf\limits_{n\to+\infty} f_n(x) \text{ p.p. en } x \in \Omega\,, \\ \text{la suite } (\phi_k(x))_{k\geq 0} \text{ est croissante p.p. en } x \in \Omega\,, \\ \int_\Omega \phi_k(x)dx \leq \sup\limits_{n\geq 0} \int_\Omega f_n(x)dx < +\infty\,. \end{cases}$$

Appliquant de nouveau la Proposition 4.0.4 page 49, cette fois à la suite $(\phi_k)_{k\geq 0}$, on trouve que la fonction

$$\varliminf_{n\to+\infty} f_n \in \mathcal{L}^1(\Omega)\,,$$

et que

$$\int_\Omega \phi_k(x)dx \to \int_\Omega \varliminf_{n\to+\infty} f_n(x)dx$$

lorsque $k \to +\infty$. De plus, en passant à la limite pour $k \to +\infty$ dans l'inégalité

$$\int_\Omega \phi_k(x)dx \leq \inf_{n\geq k} \int_\Omega f_n(x)dx\,,$$

on conclut que

$$\int_\Omega \varliminf_{n\to+\infty} f_n(x)dx \leq \varliminf_{n\to+\infty} \int_\Omega f_n(x)dx\,.$$

□

Démonstration du théorème de convergence dominée. Quitte à décomposer f sous la forme $f = \mathrm{Re}(f) + i\,\mathrm{Im}(f)$, on ne restreint pas la généralité en supposant que f est à valeurs réelles.

Posons $g_n = F - f_n$ pour tout $n \in \mathbf{N}$; par hypothèse $g_n \in \mathcal{L}^1(\Omega)$ comme différence de deux fonctions sommables, et on a $g_n \geq 0$ p.p. sur Ω d'après l'hypothèse de domination. Enfin

$$\int_\Omega g_n(x)dx \leq 2\int_\Omega F(x)dx < +\infty\,.$$

D'après le lemme de Fatou

$$\varliminf_{n\to+\infty} g_n = F - f \in \mathcal{L}^1(\Omega) \Rightarrow f \in \mathcal{L}^1(\Omega)\,.$$

Posons ensuite $h_n = 2F - |f_n - f|$ pour tout $n \in \mathbf{N}$; comme F, f_n et $f \in \mathcal{L}^1(\Omega)$, la fonction $h_n \in \mathcal{L}^1(\Omega)$ (cf. Proposition 3.3.4 page 44 4)) et on a $h_n \geq 0$ p.p. sur Ω d'après l'hypothèse de domination. Enfin

$$\int_\Omega h_n(x)dx \leq 2\int_\Omega F(x)dx < +\infty\,.$$

D'après le lemme de Fatou

$$\begin{aligned} 2\int_\Omega F(x)dx &= \int_\Omega \varliminf_{n\to+\infty} h_n(x)dx \leq \varliminf_{n\to+\infty} \int_\Omega h_n(x)dx \\ &= 2\int_\Omega F(x)dx - \varlimsup_{n\to+\infty} \int_\Omega |f_n(x) - f(x)|dx \end{aligned}$$

d'où

$$0 = \varlimsup_{n\to+\infty} \int_\Omega |f_n(x) - f(x)|dx \geq \varlimsup_{n\to+\infty} \left| \int_\Omega f_n(x)dx - \int_\Omega f(x)dx \right|,$$

ce qui achève la démonstration. □

Concluons cette section avec quelques remarques sur le théorème de convergence dominée.

Remarque 4.0.6. L'hypothèse de domination est absolument essentielle dans le théorème de convergence dominée. En l'absence de cette hypothèse, les deux phénomènes suivants peuvent se produire :

Concentration : prendre $\Omega = \mathbf{R}$ et poser

$$f_n(x) = n(1 - n|x|)^+\,, \quad x \in \mathbf{R}\,,\ n \in \mathbf{N}\,.$$

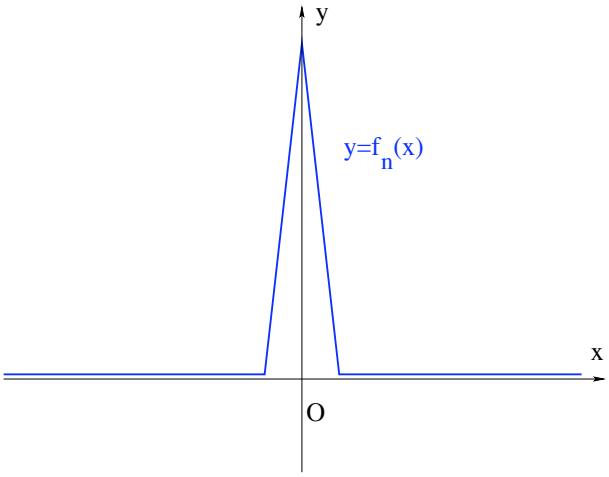


FIGURE III.3 – Phénomène de concentration : graphe de f_n.

Evidemment $f_n \in C_c(\mathbf{R})$ — la continuité est triviale ; d'autre part, pour tout $n \in \mathbf{N}$, $|x| \geq 1 \Rightarrow f_n(x) = 0$. De plus

$$\lim_{n\to+\infty} f_n(x) = \begin{cases} 0 & \text{si } x \neq 0\,, \\ +\infty & \text{si } x = 0\,. \end{cases}$$

Donc $f_n(x) \to 0$ p.p. en $x \in \mathbf{R}$. Pourtant

$$\int_{\mathbf{R}} f_n(x)dx = 1 \text{ pour tout } n \geq 1\,.$$

On n'a donc pas

$$\lim_{n\to+\infty} \int_{\mathbf{R}} f_n(x)dx = \int_{\mathbf{R}} \lim_{n\to+\infty} f_n(x)dx\,.$$

Ce qui se passe dans cet exemple est que les fonctions positives f_n peuvent être vues comme des densités de masse, et que toute la masse contenue dans les fonctions f_n se concentre en 0. Ce phénomène ne peut être détecté par la convergence p.p., car le singleton $\{0\}$ est un ensemble négligeable. Pour aller plus loin dans la compréhension de ce phénomène, on aura besoin de la théorie des distributions — ou à tout le moins de la théorie des mesures de Radon, qui en sont un cas particulier.

Evanescence : prendre $\Omega = \mathbf{R}$, choisir $\phi \in C_c(\mathbf{R})$ — par exemple, prendre $\phi(x) = (1 - |x|)^+$ pour tout $x \in \mathbf{R}$. On définit alors

$$\phi_n(x) = \phi(x - n)\,, \quad x \in \mathbf{R}\,,\ n \in \mathbf{N}\,.$$

De nouveau, la suite ϕ_n converge simplement vers 0 sur $\mathbf{R}$; pourtant

$$\int_{\mathbf{R}} \phi_n(x)dx = \int_{\mathbf{R}} \phi(x - n)dx = \int_{\mathbf{R}} \phi(y)dy = 1 \text{ pour tout } n \in \mathbf{N}\,.$$

On n'a donc pas

$$\lim_{n\to+\infty} \int_{\mathbf{R}} \phi_n(x)dx = \int_{\mathbf{R}} \lim_{n\to+\infty} \phi_n(x)dx\,.$$

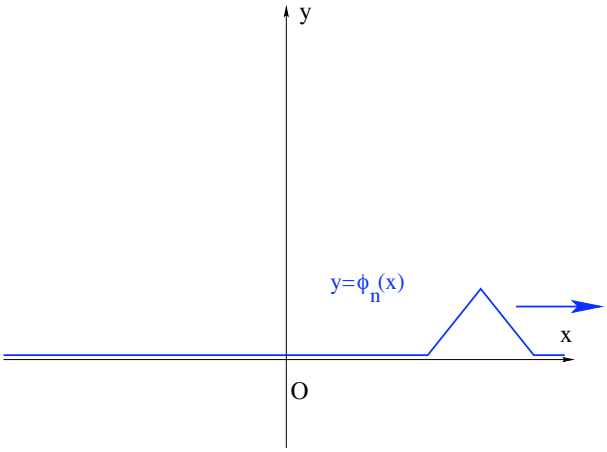


FIGURE III.4 – Phénomène d'évanescence : graphe de ϕ_n.

Avec la même analogie que précédemment, dans cet exemple, toute la masse contenue dans les fonctions ϕ_n part à l'infini, de sorte qu'on ne peut la détecter par convergence simple — puisque la convergence simple ne montre ce qui se passe que localement.

5 Caractérisation géométrique des ensembles Lebesgue-négligeables

Jusqu'ici, nous avons défini les ensembles négligeables à partir de la classe $\mathcal{L}^+$ — ou encore, de façon équivalente, à partir de la notion de suite fondamentale. Ce point de vue est très efficace pour la construction de l'intégrale et pour en établir les principales propriétés. Malheureusement, il n'est pas très pratique pour vérifier qu'un ensemble donné est négligeable.

Dans cette section, nous allons donner une autre caractérisation des ensembles Lebesgue-négligeables, de nature plus géométrique.

Notation pour les cubes de $\mathbf{R}^N$: étant donnés $a = (a_1, \ldots, a_N) \in \mathbf{R}^N$ et $r > 0$, on notera

$$C(a,r) :=]a_1 - r, a_1 + r[\times \ldots \times]a_N - r, a_N + r[$$

l'(hyper)cube ouvert de centre $a \in \mathbf{R}^N$ et de côté $2r$. On notera le "volume " du cube $C(a,r)$

$$|C(a,r)| := \text{ volume du cube } C(a,r) = (2r)^N .$$

Voici donc un critère permettant de vérifier si une partie de $\mathbf{R}^N$ est Lebesgue-négligeable.

Théorème 5.0.7. Un sous-ensemble $\mathcal{N}$ de $\mathbf{R}^N$ est Lebesgue-négligeable si et seulement si, pour tout $\epsilon > 0$, il existe une famille dénombrable (C_n) de cubes de $\mathbf{R}^N$ t.q.

$$\mathcal{N} \subset \bigcup_n C_n , \quad \text{et} \quad \sum_n |C_n| \leq \epsilon .$$

Démonstration. On fera la démonstration dans le seul cas où la dimension de l'espace $N = 1$.

Supposons que $\mathcal{N} \subset \mathbf{R}$ est négligeable. Soit $f \in \mathcal{L}^+(\mathbf{R})$ t.q. $f(x) = +\infty$ pour tout $x \in \mathcal{N}$, et soit $n > 0$ quelconque ; alors

$$\mathcal{N} \subset f^{-1}(\{+\infty\}) \subset f^{-1}(]n, +\infty]) =: \Omega(n)\,.$$

Comme $f \in \mathcal{L}^+(\mathbf{R})$, elle est s.c.i. sur $\mathbf{R}$ (cf. Proposition 3.1.6 page 35), de sorte que $\Omega(n)$ est un ouvert de $\mathbf{R}$.

D'après la Proposition 6.2.1 page 23, $\Omega(n)$ est la réunion d'une famille dénombrable $(I_k(n))$ d'intervalles ouverts disjoints, de sorte que

$$\mathcal{N} \subset \Omega(n) = \bigcup_k I_k(n)\,.$$

Comme $f(x) > n$ pour tout $x \in \Omega(n)$, pour tout ensenmble fini F d'indices k, on a

$$\begin{aligned} 0 &\leq \int_{\mathbf{R}} \mathbf{1}_{\bigcup_{k\in F} I_k(n)}(x)(f(x) - n)dx \\ &= \int_{\mathbf{R}} \mathbf{1}_{\bigcup_{k\in F} I_k(n)}(x)f(x)dx - n\int_{\mathbf{R}} \mathbf{1}_{\bigcup_{k\in F} I_k(n)}(x)dx \\ &= \int_{\mathbf{R}} \mathbf{1}_{\bigcup_{k\in F} I_k(n)}(x)f(x)dx - n\sum_{k\in F} |I_k(n)| \end{aligned}$$

— on remarquera que la fonction

$$\mathbf{1}_{\bigcup_{k\in F} I_k(n)} \text{ est continue par morceaux sur } \mathbf{R}$$

de sorte que le calcul

$$\int_{\mathbf{R}} \mathbf{1}_{\bigcup_{k\in F} I_k(n)}(x)dx = \sum_{k\in F} |I_k(n)|$$

relève de la théorie usuelle (de Riemann) de l'intégration. Au total, comme $f \geq 0$ sur $\mathbf{R}$, on a

$$n\sum_{k\in F} |I_k(n)| \leq \int_{\mathbf{R}} \mathbf{1}_{\bigcup_{k\in F} I_k(n)}(x)f(x)dx \leq \int_{\mathbf{R}} f(x)dx$$

pour tout en semble fini d'indices k, on en déduit que

$$\sum_k |I_k(n)| \leq \frac{1}{n}\int_{\mathbf{R}} f(x)dx\,.$$

Etant donné $\epsilon > 0$, en choisissant $n > 0$ assez grand pour que

$$\frac{1}{n}\int_{\mathbf{R}} f(x)dx < \epsilon\,,$$

on voit que l'on peut recouvrir $\mathcal{N}$ par une réunion dénombrable d'intervalles ouverts (les $I_k(n)$), dont la somme des longueurs est inférieure à ϵ.

Réciproquement, soit $\mathcal{N} \subset \mathbf{R}$, et supposons que, pour tout $n \in \mathbf{N}^*$, il existe une suite $(J_k(n))_{k\in\mathbf{N}^*}$ d'intervalles ouverts de $\mathbf{R}$ t.q.

$$\mathcal{N} \subset \bigcup_{k\geq 1} J_k(n) \quad \text{et} \quad \sum_{k\geq 1} |J_k(n)| < \tfrac{1}{2^n}\,.$$

On vérifie alors sans difficulté que la fonction

$$f = \sum_{n,k\geq 1} 2^{n/2} \mathbf{1}_{J_k(n)} \in \mathcal{L}^+(\mathbf{R})$$

avec

$$\int_{\mathbf{R}} f(x)dx = \sum_{n,k\geq 1} 2^{n/2}|J_k(n)| \leq \sum_{n\geq 1} 2^{-n/2} = \tfrac{1}{\sqrt{2}-1}\,,$$

et que

$$f(x) = \sum_{n\geq 1} 2^{n/2} = +\infty \text{ pour tout } x \in \mathcal{N}\,.$$

□

Exemple 5.0.8. On retrouve très simplement grâce au Théorème 5.0.7 page 54 le fait qu'un ouvert non vide de $\mathbf{R}^N$ n'est jamais Lebesgue-négligeable — déjà établi plus haut, cf. Exemple 3.2.8 page 42. Plus généralement, une partie de $\mathbf{R}^N$ d'intérieur non vide n'est jamais négligeable.

En effet, d'après le Théorème 5.0.7 page 54, un cube ouvert non vide de $\mathbf{R}^N$ étant de volume V strictement positif, il n'est pas possible de le recouvrir par une suite de cubes dont la somme des volumes resterait inférieure à $V/2$.

Exemple 5.0.9. Une droite affine dans $\mathbf{R}^N$ avec $N > 1$, plus généralement un sous-espace affine de dimension $d < N$ dans $\mathbf{R}^N$ sont des parties négligeables de $\mathbf{R}^N$.

Montrons par exemple qu'un segment I de longueur 1 est un ensemble négligeable dans le plan euclidien $\mathbf{R}^2$. Etant donné $n > 1$, on découpe I en n segments de longueur $1/n$, notés $I_1, \ldots, I_n$. Le segment I est donc recouvert par la réunion des disques $D_1, \ldots, D_n$ de diamètres respectifs $I_1, \ldots, I_n$. Pour tout $j = 1, \ldots, n$, soit C_j le plus petit carré de côtés parallèles aux axes contenant D_j. Un tel carré est de côté $\leq 1/n$. Donc on a

$$I \subset C_1 \cup \ldots \cup C_n\,, \quad \text{et } |C_1| + \ldots + |C_n| = n \cdot \left(\frac{1}{n}\right)^2 = \frac{1}{n}\,.$$

Ainsi, pour tout $\epsilon > 0$, le segment I peut être recouvert par $n = [1/\epsilon] + 1$ carrés dont la somme des surfaces reste inférieure à ϵ.

Les autre énoncés de l'exemple ci-dessus sont laissés au lecteur à titre d'exercices.

Les différents exemples d'ensembles négligeables que nous avons rencontrés jusqu'ici étaient relativement simples. Toutefois, il existe de nombreux ensembles particulièrement compliqués qui sont de mesure nulle.

Exemple 5.0.10 (Ensemble de Cantor). On pose $K_0 = [0,1] \subset \mathbf{R}$. Pour tous $a < b \in \mathbf{R}$, soit $\mathcal{T}[a,b] = [a,b]\backslash]\frac{a+b}{2} - \frac{b-a}{3}, \frac{a+b}{2} - \frac{b-a}{3}[$. On définit $K_1 = \mathcal{T}K_0$, puis, par récurrence,

$$K_{n+1} = \bigcup_{1\leq k\leq 2^n} \mathcal{T} J_{n,k}$$

sachant que $J_{n,1}, \ldots, J_{n,2^n}$ sont des segments disjoints tels que

$$K_n = \bigcup_{1\leq k\leq 2^n} J_{n,k}\,.$$



FIGURE III.5 – Les 6 premières étapes de la construction de l'ensemble triadique de Cantor

L'ensemble triadique de Cantor est

$$K = \bigcap_{n \geq 0} K_n \, .$$

Voir sur la figure 5.0.10 les premières étapes de la construction de l'ensemble triadique de Cantor.

C'est un compact de $\mathbf{R}$, totalement discontinu (autrement dit, la composante connexe de tout point x de K est le singleton $\{x\}$), et qui est un ensemble parfait (autrement dit sans point isolé.)

On montre sans peine que K est négligeable — on vérifie en particulier que

$$\sum_{k=1}^{2^n} |J_{n,k}| = \frac{2^n}{3^n} \, .$$



FIGURE III.6 – Le triangle de Sierpiński. A chaque étape, on enlève, dans chaque triangle équilatéral noir, le triangle blanc dont les sommets sont les milieux des côtés du triangle noir.

Exemple 5.0.11 (Triangle de Sierpiński). Le triangle de Sierpinski est l'analogue de l'ensemble triadique de Cantor dans le plan $\mathbf{R}^2$. On a représenté sur la figure 5 les premières étapes de sa construction.

Le triangle de Sierpiński est un exemple d'ensemble négligeable dans le plan euclidien $\mathbf{R}^2$.

En fait, l'ensemble triadique de Cantor, et le triangle de Sierpiński sont deux exemples d'ensemble "fractal", c'est-à-dire de "dimension non entière [3]". On trouve que l'ensemble de

3. Etant donné $0 < \epsilon \ll 1$, on trouve sans peine qu'il faut environ $O(1/\epsilon^N)$ petits (hyper)cubes de côté ϵ pour recouvrir l'hypercube $[0,1]^N \subset \mathbf{R}^N$. La dimension N apparaît donc comme exposant de $1/\epsilon$ dans le nombre de petits (hyper)cubes de côté ϵ nécessaires pour recouvrir $[0,1]^N$. En général, étant donné K compact de $\mathbf{R}^N$, soit $\mathcal{N}(K,\epsilon)$ = la borne inférieure du nombre d'hypercubes de côté $\epsilon \ll 1$ nécessaires pour recouvrir K. On définit alors la dimension de Minkowski de K comme la limite (si elle existe)

$$\dim_{Minkowski}(K) := \lim_{\epsilon \to 0} \frac{\ln \mathcal{N}(K,\epsilon)}{\ln(1/\epsilon)} \, .$$

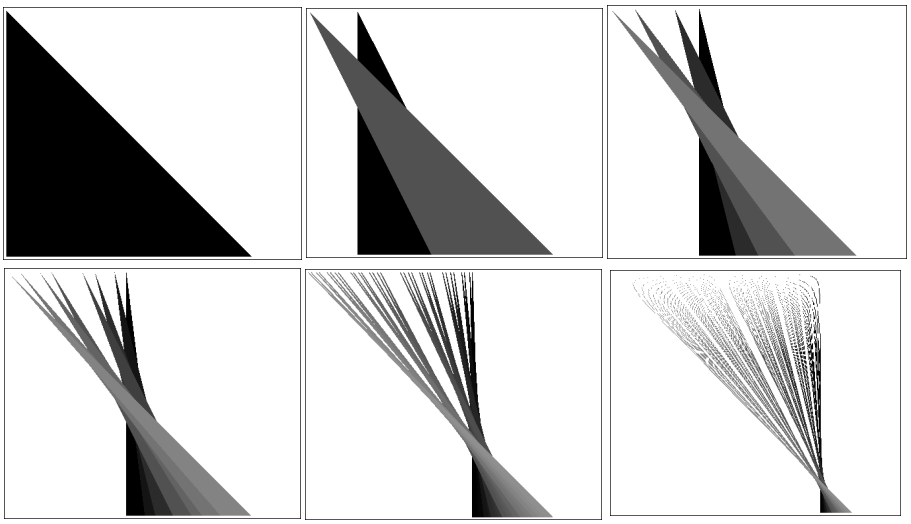


FIGURE III.7 – Premières étapes de la construction de l'ensemble de Besicovich dans le plan. L'ensemble final contient un segment de longueur 1 dans toutes les directions appartenant à l'octant défini par le triangle noir de départ.

Cantor est de dimension $\ln 2/\ln 3 < 1$ et que le triangle de Sierpiński est de dimension $\ln 3/\ln 2 < 2$. Dans les deux cas, l'ensemble considéré est de dimension est strictement inférieure à celle de l'espace ambiant. Tout ensemble fractal de dimension strictement inférieure à la dimension de l'espace ambiant est Lebesgue-négligeable.

Voici un dernier exemple d'ensemble négligeable "pathologique".

Exemple 5.0.12 (Ensemble de Besicovich). Un ensemble de Kakeya dans $\mathbf{R}^N$ est un ensemble qui contient au moins un segment de longueur 1 dans chaque direction. Evidemment, un disque de diamètre 1, un triangle équilatéral de côté $2/\sqrt{3}$ sont des ensembles de Kakeya dans le plan euclidien. Le problème intéres- sant est de construire des ensembles de Kakeya aussi petits que possible.Besicovich fut le premier à construire dans le plan euclidien $\mathbf{R}^2$ un ensemble de Kakeya négligeable. Voici, sur la figure 5, les premières étapes de la construction de l'ensemble de Besicovich.

Terminons cette section avec quelques remarques de nature topologique.

Dire d'un ensemble qu'il est négligeable est une façon de dire qu'il est "petit" ou "raréfié" — en un certain sens. Dire d'un ensemble qu'il est dense dans $\mathbf{R}^N$ signifie qu'on le rencontre au voisinage de tout point. Examinons plus en détail les rapports entre ces deux notions.

D'abord, une partie de $\mathbf{R}$ ou de $\mathbf{R}^N$ peut être à la fois négligeable et dense.

Exemple 5.0.13. Par exemple $\mathbf{Q}$ est négligeable et dense dans $\mathbf{R}$; de même, $\mathbf{Q}^N$ est négligeable et dense dans $\mathbf{R}^N$.

Pour l'ensemble de Cantor, on remarque qu'il est recouvert par K_n, c'est-à-dire par une réunion de 2^n segments de longueur $1/3^n$, de sorte que $\mathcal{N}(K, 3^{-n}) \leq 2^n$. Donc $\dim_{Minkoswki}(K) \leq \frac{\ln 2}{\ln 3}$.

D'autre part, on a vu plus haut (Exemple 5.0.8 page 56) qu'un ouvert non vide de $\mathbf{R}^N$, ou plus généralement une partie d'intérieur non vide ne peut pas être négligeable dans $\mathbf{R}^N$. Par passage au complémentaire, on aboutit au résultat suivant :

Proposition 5.0.14. Soit Ω ouvert de $\mathbf{R}^N$. Alors le complémentaire d'une partie négligeable de Ω est dense dans Ω.

6 Fonctions mesurables

Dans la théorie de l'intégration que nous venons de construire, les fonctions sommables (autrement dit intégrables au sens de Lebesgue) ne sont pas définies partout mais seulement presque partout — c'est-à-dire sur le complémentaire d'un ensemble négligeable.

La notion de convergence simple n'est donc pas adaptée dans ce cadre, et doit être remplacée par la convergence presque partout. Nous sommes donc naturellement amenés à étudier la classe des limites presque partout de suites de fonctions continues à support compact sur un ouvert de $\mathbf{R}$.

Définition 6.0.15. Soit $\Omega \subset \mathbf{R}^N$ ouvert. Une fonction $f : \Omega \to [-\infty, +\infty]$ ou $\mathbf{C}$ est mesurable s'il existe une suite $(f_n)_{n\in\mathbf{N}}$ de fonctions continues à support compact dans Ω et à valeurs dans $\mathbf{R}$ ou $\mathbf{C}$ t.q. $f_n(x) \to f(x)$ p.p. en $x \in \Omega$.

Voici les premières propriétés des fonctions mesurables.

Proposition 6.0.16. Soit Ω ouvert de $\mathbf{R}^N$.

1) toute fonction continue sur Ω à valeurs dans $\mathbf{R}$ ou $\mathbf{C}$ est mesurable ;

2) toute fonction continue par morceaux sur un intervalle $I \subset \mathbf{R}$ est mesurable ;

3) toute fonction appartenant à $\mathcal{L}^+(\Omega)$ est mesurable ;

4) si $f_1, \dots, f_n : \Omega \to \mathbf{C}$ sont mesurables et $\Phi : \mathbf{C}^n \to \mathbf{C}$ continue, alors la fonction

$$\Phi(f_1, \dots, f_n) : \Omega \ni x \mapsto \Phi(f_1(x), \dots, f_n(x))$$

est mesurable sur Ω.

Tous ces énoncés sont essentiellement triviaux et laissés au lecteur à titre d'exercice. (Pour ce qui est de l'énoncé 4), commencer par le démontrer dans le cas où $\Phi(0, \dots, 0) = 0$, puis vérifier que si ϕ est une fonction mesurable sur Ω à valeurs dans $\mathbf{C}$, il en va de même pour $\phi + C$, pour tout $C \in \mathbf{C}$.)

Voici quelques applications importantes de l'énoncé 4).

Corollaire 6.0.17. Soit Ω ouvert de $\mathbf{R}^N$. Alors

a) si $f, g : \Omega \to \mathbf{C}$ sont mesurables, alors les fonctions $f + g$ et fg sont mesurables sur Ω ;

b) toute fonction appartenant à $\mathcal{L}^1(\Omega, \mathbf{C})$ est mesurable sur Ω ;

c) si $f : \Omega \to \mathbf{C}$ est mesurable, alors $|f|$ et $\overline{f}$ sont mesurables sur Ω ;

d) si $f, g : \Omega \to \mathbf{R}$ sont mesurables, alors les fonctions f^+, f^-, et plus généralement $\max(f, g)$ et $\min(f, g)$ sont mesurables.

La notion de fonction mesurable fournit un moyen très simple de vérifier qu'une fonction est sommable, à savoir le théorème de comparaison ci-dessous. Nous y reviendrons plus loin.

Proposition 6.0.18. Soit Ω ouvert de $\mathbf{R}^N$. Soit u, fonction mesurable sur Ω à valeurs complexes. Supposons qu'il existe $f \in \mathcal{L}^1(\Omega)$ t.q. $|u(x)| \leq f(x)$ p.p. en $x \in \Omega$. Alors $u \in \mathcal{L}^1(\Omega)$.

Démonstration. Quitte à décomposer u sous la forme $u = \mathrm{Re}(u) + i\,\mathrm{Im}(u)$, on peut supposer que u est à valeurs réelles. Comme u est mesurable sur Ω, il existe une suite $(u_n)_{n\geq 0}$ de fonctions appartenant à $C_c(\Omega)$ telle que $u_n(x) \to u(x)$ p.p. en $x \in \Omega$ lorsque $n \to +\infty$. Posons

$$v_n(x) = \max(\min(u_n(x), f(x)), -f(x)), \quad \text{p.p. en } x \in \Omega$$

et pour tout $n \in \mathbf{N}$. Par construction

$$v_n(x) \to u(x) \text{ p.p. en } x \in \Omega \text{ lorsque } n \to +\infty .$$

D'après la Proposition 3.3.4 page 44 5), $v_n \in \mathcal{L}^1(\Omega)$ pour tout $n \in \mathbf{N}$; puis

$$v_n(x) \to u(x) \text{ p.p. en } x \in \Omega \text{ lorsque } n \to +\infty ,$$

et

$$|v_n(x)| \leq f(x) \text{ p.p. en } x \in \Omega \text{ et pour tout } n \in \mathbf{N} .$$

Le théorème de convergence dominée entraîne alors que $u \in \mathcal{L}^1(\Omega)$. □

Un résultat extrêmement important est que la classe des fonctions mesurables est stable par convergence presque partout.

Théorème 6.0.19. Soient Ω ouvert de $\mathbf{R}^N$, et une suite de fonctions $f_n : \Omega \to \mathbf{C}$ mesurables pour tout $n \geq 0$ telle que

$$f_n(x) \to f(x) \text{ p.p. sur } \Omega \text{ lorsque } n \to +\infty .$$

Alors la fonction f définie p.p. sur Ω et à valeurs dans $\mathbf{C}$ est mesurable.

Démonstration. Nous allons démontrer ce résultat en plusieurs étapes. D'abord, grâce aux énoncés a) et b) du corollaire précédent, on peut supposer que les fonctions f_n et f sont toutes à valeurs réelles.

Etape 1 : Soit $(u_n)_{n\geq 1}$ suite de fonctions de $\mathcal{L}^+(\Omega)$. Supposons que

$$\begin{cases} u_1(x) \leq u_2(x) \leq \ldots \leq u_n(x) \leq u_{n+1}(x) \leq \ldots \text{ pour tout } x \in \Omega , \\ u_n(x) \to u(x) \in \mathbf{R} \text{ p.p. en } x \in \Omega \text{ lorsque } n \to +\infty . \end{cases}$$

Montrons que u est une fonction mesurable sur Ω.

En effet, pour chaque $k \geq 1$, il existe $(u_{k,n})_{n\geq 1}$, suite fondamentale convergeant simplement vers u_k sur Ω. Posons

$$v_n(x) = \max(u_{1,n}(x), \ldots, u_{n,n}(x))\,;$$

alors la suite $(v_n(x))_{n\geq 1}$ est croissante pour tout $x \in \Omega$ et converge vers une limite notée $v(x) \in \mathbf{R} \cup \{+\infty\}$. Evidemment v est mesurable sur Ω puisque $v_n \in C_c(\Omega)$. Comme

$$v_n(x) \leq \max(u_1(x), \ldots, u_n(x)) = u_n(x)\,,$$

alors

$$v(x) \leq u(x) \text{ p.p. en } x \in \Omega\,.$$

En particulier $v(x) < +\infty$ p.p. en $x \in \Omega$.

D'autre part, pour $k \leq n$, on a

$$u_{k,n}(x) \leq v_n(x) \text{ pour tout } x \in \Omega\,;$$

en passant à la limite pour $n \to +\infty$, on trouve que

$$u_k(x) \leq v(x) \text{ pour tout } x \in \Omega\,.$$

Puis, en passant à la limite en $k \to +\infty$, on trouve que

$$u(x) \leq v(x) \text{ p.p. en } x \in \Omega\,.$$

Au total, $u(x) = v(x)$ p.p. en $x \in \Omega$, et comme v est mesurable sur Ω, il en est de même pour u.

Etape 2 : Soit $(u_n)_{n\geq 1}$ suite de fonctions de $\mathcal{L}^1(\Omega)$. Supposons que

$$\begin{cases} u_1(x) \leq u_2(x) \leq \ldots \leq u_n(x) \leq u_{n+1}(x) \leq \ldots \text{ pour tout } x \in \Omega\,, \\ u_n(x) \to u(x) \in \mathbf{R} \text{ p.p. en } x \in \Omega \text{ lorsque } n \to +\infty\,. \end{cases}$$

Montrons que u est une fonction mesurable sur Ω.

En effet, d'après le Lemme 3.3.3 page 43, pour tout $n \geq 1$, il existe $v_n, w_n \in \mathcal{L}^+(\Omega)$ telles que

$$0 \leq u_n(x) - u_{n-1}(x) = v_n(x) - w_n(x) \text{ p.p. en } x \in \Omega$$

tandis que

$$w_n(x) \geq 0 \text{ pour tout } x \in \Omega\,, \quad \text{ et } \int_\Omega w_n(x)dx < 2^{-n}\,.$$

En particulier, pour tout $n \geq 1$, on a

$$v_n(x) = u_n(x) - u_{n-1}(x) + w_n(x) \geq 0 \text{ p.p. en } x \in \Omega$$

de sorte que les suites

$$W_n(x) = \sum_{k=1}^{n} w_n(x) \quad \text{ et } V_n(x) = \sum_{k=1}^{n} v_n(x)$$

sont croissantes p.p. en $x \in \Omega$. Alors $W_n(x) \to W(x)$ pour tout $x \in \Omega$ lorsque $n \to +\infty$, et comme

$$\sum_{k\geq 1} \int_\Omega w_k(x)dx < \sum_{k\geq 1} 2^{-k} = 1\,,$$

la fonction $W \in \mathcal{L}^+(\Omega)$ d'après la Proposition 4.0.4 page 49.

D'autre part, $V_n(x) \to V(x)$ p.p. en $x \in \Omega$ lorsque $n \to +\infty$, et la suite $(V_n(x))_{n\geq 1}$ est croissante p.p en $x \in \Omega$. D'après l'étape 1, V est mesurable sur Ω. Au total, $u = V - W + u_0$ est mesurable comme somme de fonctions mesurables sur Ω.

Etape 3 : Soit $(u_n)_{n\geq 1}$ suite de fonctions mesurables sur Ω t.q.

$$\begin{cases} u_n(x) \geq 0 \text{ p.p. en } x \in \Omega \text{ et pour tout } n \geq 1\,, \\ u_n(x) \to u(x) \in \mathbf{R} \text{ p.p. en } x \in \Omega \text{ lorsque } n \to +\infty\,. \end{cases}$$

Montrons que u est mesurable sur Ω.

Pour tout $n \geq 1$, il existe une suite $(u_{n,k})_{k\geq 1}$ de fonctions de $C_c(\Omega)$ t.q.

$$\begin{cases} u_{n,k}(x) \to u_n(x) \text{ p.p. en } x \in \Omega \text{ lorsque } k \to +\infty\,, \\ u_{n,k}(x) \geq 0 \text{ p.p. en } x \in \Omega \text{ pour tout } n, k \geq 1\,, \\ U_{n,k} := \displaystyle\int_\Omega u_{n,k}(x)dx > 0 \text{ pour tout } n, k \geq 1\,. \end{cases}$$

D'après la Proposition 3.2.5 page 39

$$v(x) := \sum_{n\geq 1} \sum_{k\geq n} \frac{u_{n,k}(x)}{2^k U_{n,k}}$$

définit une fonction $v \in \mathcal{L}^+(\Omega)$.

Soit alors $m \in \mathbf{N}^*$. D'après la Proposition 6.0.18 page 60, la fonction $w_{m,n} := \min(u_n, mf)$ définie p.p. sur Ω appartient à $\mathcal{L}^1(\Omega)$. D'autre part, pour tout $m \geq 1$ fixé,

$$\begin{cases} w_{m,n}(x) \to w_m(x) := \min(u(x), mf(x)) \text{ p.p. en } x \in \Omega \text{ lorsque } n \to +\infty\,, \\ 0 \leq w_{m,n}(x) \leq mf(x) \text{ p.p. en } x \in \Omega\,. \end{cases}$$

Donc d'après le théorème de convergence dominée, pour tout $m \geq 1$, la fonction $w_m \in \mathcal{L}^1(\Omega)$.

De plus

$$0 \leq w_1(x) \leq \ldots \leq w_m(x) \leq w_{m+1}(x) \to u(x) \text{ p.p. en } x \in \Omega \text{ lorsque } m \to +\infty\,.$$

D'après le résultat de l'étape 2, la fonction u est mesurable sur Ω.

Pour le cas général, on décompose chaque terme de la suite $(u_n)_{n\geq 1}$ sous la forme

$$u_n = \mathrm{Re}(u_n)^+ + \mathrm{Re}(u_n)^- + i\,\mathrm{Im}(u_n)^+ + i\,\mathrm{Im}(u_n)^-$$

et on applique le résultat de l'étape 3, ainsi que le Corollaire 6.0.17 page 59 pour conclure. □

Le Corollaire 6.0.17 page 59 montre que la propriété de mesurabilité est stable par les opérations usuelles, sauf pour ce qui est du passage à l'inverse. Examinons ce cas plus en détail.

Corollaire 6.0.20. Soit Ω ouvert de $\mathbf{R}^N$, et $f : \Omega \to \mathbf{C}$ mesurable t.q. $f(x) \neq 0$ p.p. en $x \in \Omega$. Alors la fonction $x \mapsto 1/f(x)$, définie p.p. sur Ω, est mesurable sur Ω.

Démonstration. Il suffit en effet d'observer que, lorsque $n \to +\infty$,

$$f_n(x) := \frac{\overline{f(x)}}{\frac{1}{n} + |f(x)|^2} \to \frac{1}{f(x)} \text{ pour tout } x \in \Omega \text{ t.q. } f(x) \neq 0\,.$$

Or, pour tout $n \geq 1$, la fonction f_n est mesurable sur Ω d'après la Proposition 6.0.16 page 59 4), appliquée à la fonction continue sur $\mathbf{C}$

$$\Phi : z \mapsto \frac{\overline{z}}{\frac{1}{n} + |z|^2}\,.$$

On conclut alors grâce au Théorème 6.0.19 page 60. □

Une fonction mesurable peut-être extrêmement singulière ; en fait, bien qu'elle soit limite p.p. d'une suite de fonctions continues, une fonction mesurable peut très bien n'être continue nulle part.

Exemple 6.0.21. La fonction indicatrice $\mathbf{1}_{\mathbf{Q}\cap[0,1]}$ est mesurable sur $\mathbf{R}$.

En résumé, toutes les fonctions que l'on construit à partir des fonctions continues par une succession finie d'opérations classiques (addition, multiplication, composition, inversion...), ou par une suite dénombrable de telles opérations et par passage à la limite p.p. sont mesurables.

En fait, il est impossible de construire une fonction non mesurable par un algorithme fini ou dénombrable mettant en jeu des opérations continues. On sait qu'il existe des fonctions non mesurables sur $\mathbf{R}$, mais leur existence est basée sur l'axiome du choix [4]

4. Cet axiome peut se formuler ainsi : pour tout ensemble $I \neq \varnothing$, et toute famille $(X_i)_{i\in I}$ d'ensembles indexée par I

$$X_i \neq \varnothing \text{ pour tout } i \in I \Rightarrow \prod_{i\in I} X_i \neq \varnothing\,.$$

Autrement dit, il existe $(x_i)_{i\in I}$ avec $x_i \in I$ pour tout $i \in I$. Lorsque I est fini, l'implication ci-dessus n'est pas mystérieuse : en identifiant I à $\{1,\dots,n\}$, comme $X_1 \neq \varnothing$, on peut choisir $x_1 \in X_1$, puis comme $X_2 \neq \varnothing$, on peut choisir $x_2 \in X_2$ et ainsi de suite jusqu'à X_n. A la fin, on a donc $(x_1,\dots,x_n) \in X_1 \times \dots \times X_n$. Dans le cas dénombrable, si $I = \mathbf{N}$, on construit une suite $(x_n)_{n\in\mathbf{N}}$ comme ci-dessus, par récurrence (traduction mathématique de la locution "et ainsi de suite..."), sauf que rien ne garantit qu'on ait pu terminer la construction en temps fini, et donc qu'on dispose un jour de l'élément $(x_n)_{n\in\mathbf{N}} \in \prod_{n\in\mathbf{N}} X_n$. Mais on peut imaginer qu'à chaque étape on mette par exemple deux fois moins de temps à choisir qu'à l'étape précédente : le temps nécessaire pour effectuer tous les choix serait fini dans un tel cas. Un autre cas favorable est celui où I n'est pas forcément dénombrable, mais où tous les X_i sont égaux à un même ensemble X. Alors trouver un élément $(x_i)_{i\in I}$ du produit cartésien $\prod_{i\in I} X_i$ revient à trouver une application $f : I \to X$ et à poser $x_i = f(i)$ pour tout $i \in I$. Un choix évident consiste donc à prendre $x^* \in X$ et à considérer l'application constante $f : I \ni i \mapsto f(i) = x^* \in X$. Mais si I n'est pas dénombrable et que les ensembles X_i n'ont rien aucun rapport entre eux, il n'existe pas de moyen constructif d'obtenir un élément $(x_i)_{i\in I} \in \prod_{i\in I} X_i$ .On postule donc que $\prod_{i\in I} X_i$ n'est pas vide, ce qui garantit l'existence d'un tel $(x_i)_{i\in I}$, sans donner d'algorithme pour le construire. Autrement dit, l'axiome du choix affirme la possibilité de choisir *simultanément* un élément dans une infinité quelconque d'ensembles non vides. L'axiome du choix a des conséquences trop utiles en mathématiques pour qu'on songe à s'en passer — par exemple, c'est cet axiome qui

7 Intégration et fonctions mesurables

Jusqu'ici, on a construit l'intégrale sur l'espace $\mathcal{L}^1$ des fonctions sommables, avec un théorème d'interversion intégrale/limite (le théorème de convergence dominée) qui correspond bien à ce que l'on cherchait — voir notamment les points A) et B) de l'introduction (section 1 page 27.)

Cette construction laisse toutefois plusieurs questions en suspens. Par exemple, comment vérifie-t-on en pratique qu'une fonction est sommable si tel est le cas, comment peut-on calculer son intégrale de Lebesgue ?

Nous allons donner quelques éléments de réponse à ces diverses questions pratiques, en nous basant en particulier sur la notion de fonction mesurable que nous venons justement d'introduire.

Commençons par la première question :

I) Étant donnée une fonction f définie p.p. sur Ω ouvert de $\mathbf{R}^N$, comment vérifie-t-on si $f \in \mathcal{L}^1(\Omega)$?

En pratique, on n'essaiera pas d'écrire la fonction f sous la forme $f = g - h$ avec $g, h \in \mathcal{L}^+(\Omega)$: la construction d'une telle décomposition n'a en effet rien d'intuitif.

Mais d'abord, pour que $f \in \mathcal{L}^1(\Omega)$, il faut qu'elle soit mesurable sur Ω : voir Corollaire 6.0.17 page 59. Les remarques de la fin de la section précédente montrent qu'il ne s'agit pas d'une hypothèse très contraignante. Une fonction définie p.p. sur Ω et qui n'est pas construite par un procédé faisant intervenir l'axiome du choix — ce qui passe difficilement inaperçu — est forcément mesurable.

Supposons donc la fonction f mesurable sur Ω, et cherchons quelle(s) condition(s) supplémentaire(s) f doit vérifier pour être sommable sur Ω. L'idée est de procéder comme dans le cas des séries numériques : pour vérifier qu'une série est absolument convergente, on compare la valeur absolue (ou le module) de son terme général à celui d'une série à termes positifs de référence dont on sait qu'elle est convergente. Or nous avons énoncé plus haut un théorème de comparaison qui répond précisément à cette question (Proposition 6.0.18 page 60.)

Toutefois, cet énoncé n'est pas entièrement satisfaisant pour l'esprit, dans la mesure où, pour vérifier que la fonction u est sommable sur Ω, il faut déjà savoir que la fonction f, à laquelle on compare $|u|$, l'est elle-même.

En réalité, on peut modifier très légèrement l'énoncé de cette proposition pour arriver à l'énoncé suivant — qui lui est d'ailleurs équivalent.

Théorème 7.0.22 (Théorème de comparaison). Soient Ω ouvert de $\mathbf{R}^N$ et $f : \Omega \to [-\infty, +\infty]$ mesurable. Alors

$$f \in \mathcal{L}^1(\Omega) \Leftrightarrow \text{ il existe } F \in \mathcal{L}^+(\Omega) \text{ t.q. } |f(x)| \leq F(x) \text{ p.p. en } x \in \Omega\,.$$

garantit que tout espace vectoriel, même de dimension infinie, admet une base. Mais, en contre-partie, l'axiome du choix entraîne l'existence de "monstres" mathématiques — à tout le moins, d'objets mathématiques peu intuitifs — comme par exemple des fonctions non mesurables sur $\mathbf{R}^N$. On verra dans le chapitre suivant une conséquence encore plus dérangeante de ce même axiome.

Démonstration. Comme $\mathcal{L}^+(\Omega) \subset \mathcal{L}^1(\Omega)$, la condition de comparaison est évidemment suffisante d'après la Proposition 6.0.18 page 60.

Réciproquement, si $f \in \mathcal{L}^1(\Omega)$, on sait, d'après la Proposition 3.3.4 page 44 4), que $|f| \in \mathcal{L}^1(\Omega)$. Il existe donc deux fonctions $g, h \in \mathcal{L}^+(\Omega)$ telles que

$$|f(x)| = g(x) - h(x) = g(x)^+ + h(x)^- - (g(x)^- + h(x)^+) \text{ p.p. en } x \in \Omega\,.$$

Posons $F(x) = g(x)^+ + h(x)^-$ pour tout $x \in \Omega$; d'après la Proposition 3.1.8 page 36 2-3), on a $F \in \mathcal{L}^+(\Omega)$ et ce qui précède montre que $|f(x)| \leq F(x)$ p.p. en $x \in \Omega$. □

L'intérêt de la Proposition ci-dessus est qu'il est très facile de vérifier en pratique s'il existe $F \in \mathcal{L}^+(\Omega)$ telle que $|f(x)| \leq F(x)$ p.p. en $x \in \Omega$, dans la mesure ou il suffit de chercher si l'on peut construire une suite $(F_n)_{n\geq 0}$ fondamentale sur Ω telle que

$$\lim_{n\to+\infty} F_n(x) \geq |f(x)|\,.$$

Dans le même ordre d'idées, pour vérifier la condition de domination dans le théorème de convergence dominée (Théorème 4.0.3 page 49) on cherchera en pratique une fonction $F \in \mathcal{L}^+(\Omega)$.

Remarque 7.0.23. Comparons l'énoncé de la Proposition 6.0.18 page 60 (ou du Théorème 7.0.22 page ci-contre) avec le théorème de convergence dominée de Lebesgue (Théorème 4.0.3 page 49.)

Dans la Proposition 6.0.18 page 60 ou le Théorème 7.0.22 page précédente, comme f est mesurable sur Ω, il existe une suite $(f_n)_{n\geq 0}$ de fonctions de $C_c(\Omega)$ telle que $f_n(x) \to f(x)$ p.p. en $x \in \Omega$. (En particulier $f_n \in \mathcal{L}^1(\Omega)$ pour tout $n \in \mathbf{N}$ car $C_c(\Omega) \subset \mathcal{L}^1(\Omega)$.)

Mais comme, dans la Proposition 6.0.18 page 60 ou le Théorème 7.0.22 page précédente, la fonction $F \in \mathcal{L}^1(\Omega)$ ne domine que la limite f, et pas tous les termes de la suite comme dans le théorème de convergence dominée, on a seulement $f \in \mathcal{L}^1(\Omega)$ sans pouvoir conclure que

$$\int_\Omega f_n(x)dx \to \int_\Omega f(x)dx$$

lorsque $n \to +\infty$. Ce n'est pas gênant , dans la mesure où la Proposition 6.0.18 page 60 ou le Théorème 7.0.22 page précédente servent uniquement à vérifier que f est sommable, et pas à calculer son intégrale.

On peut même aller plus loin dans le cas des fonctions mesurables à valeurs positives ou nulles — c'est-à-dire dans $[0, +\infty]$.

Définition 7.0.24 (Intégration des fonctions mesurables positives). Soient Ω ouvert de $\mathbf{R}^N$, et $f : \Omega \to [0, +\infty]$ mesurable définie p.p. sur Ω. Alors

• soit $f \in \mathcal{L}^1(\Omega)$;

• soit $f \notin \mathcal{L}^1(\Omega)$, et on pose alors

$$\int_\Omega f(x)dx = +\infty\,.$$

Autrement dit, d'après le Théorème 7.0.22 page 64,

soit il existe une fonction $F \in \mathcal{L}^+(\Omega)$ telle que $f(x) \leq F(x)$ p.p. en $x \in \Omega$,

soit $\displaystyle\int_\Omega f(x)dx = +\infty$.

Avec cette convention, on aboutit à un énoncé d'intégration terme à terme des séries de fonctions à termes positifs qui est d'une simplicité maximale — exactement comme l'énoncé B) de l'introduction (section 1 page 27.)

Théorème 7.0.25 (Théorème de convergence monotone). Soient Ω ouvert de $\mathbf{R}^N$ et $(f_n)_{n\geq 0}$ suite de fonctions mesurables sur Ω à valeurs dans $[0, +\infty]$. Alors

$$\int_\Omega \left(\sum_{n=0}^{\infty} f_n(x) \right) dx = \sum_{n=0}^{\infty} \int_\Omega f_n(x)dx \quad \text{dans } [0, +\infty]\,.$$

Démonstration. Posons

$$F_n(x) = \sum_{k=0}^{n} f_k(x)\,, \quad \text{p.p. en } x \in \Omega \text{ pour tout } n \in \mathbf{N}\,,$$

ainsi que

$$F(x) = \sum_{k=0}^{\infty} f_k(x)\,, \quad \text{p.p. en } x \in \Omega\,.$$

Evidemment, la série ci-dessus converge dans $[0, +\infty]$, et définit p.p. sur Ω une fonction F mesurable sur Ω.

Si

$$\sum_{n=0}^{\infty} \int_\Omega f_n(x)dx < +\infty\,,$$

on déduit de la Proposition 4.0.4 page 49 que $F \in \mathcal{L}^1(\Omega)$ et que

$$\int_\Omega F(x)dx = \sum_{n=0}^{\infty} \int_\Omega f_n(x)dx\,.$$

Si au contraire

$$\sum_{n=0}^{\infty} \int_\Omega f_n(x)dx = +\infty\,,$$

on a forcément

$$\int_\Omega F(x)dx = +\infty = \sum_{n=0}^{\infty} \int_\Omega f_n(x)dx\,.$$

En effet, si on avait

$$\int_\Omega F(x)dx < +\infty$$

alors $F \in \mathcal{L}^1(\Omega)$ et on aurait

$$\begin{cases} F_n(x) \to F(x) \text{ p.p. en } x \in \Omega \text{ lorsque } n \to +\infty\,, \\ 0 \leq F_n(x) \leq F(x) \text{ p.p. en } x \in \Omega \text{ pour tout } n \geq 0\,. \end{cases}$$

Par convergence dominée (Théorème 4.0.3 page 49) on en déduirait que

$$\int_\Omega F_n(x)dx = \sum_{k=0}^{n} \int_\Omega f_k(x)dx \to \int_\Omega F(x)dx$$

lorsque $n \to +\infty$, c'est-à-dire que

$$\sum_{k\geq 0} \int_\Omega f_k(x)dx = \int_\Omega F(x)dx < +\infty\,,$$

ce qui contredit l'hypothèse de départ.

□

8 Ce qu'il faut retenir

Même si on a suivi dans ce chapitre l'une des voies les plus directe pour effectuer la construction de l'intégrale de Lebesgue, les différents arguments que nous avons présentés sont agencés de manière à s'enchaîner dans une succession logique, qui n'est pas forcément le plus naturel lorsqu'on cherche à *utiliser* la théorie — au lieu de la construire.

Nous allons donc donner dans cette section une présentation *purement descriptive* de l'intégrale de Lebesgue, qui correspond à la façon dont on l'utilise dans la pratique.

Ensembles négligeables

Une partie de $\mathbf{R}^N$ est négligeable si, pour tout $\epsilon > 0$, on peut la recouvrir par une réunion dénombrable d'hypercubes dont la somme des volumes est inférieure à ϵ.

Tout singleton de $\mathbf{R}^N$ est Lebesgue-négligeable, tout sous-espace affine de dimension $< N$ de $\mathbf{R}^N$ est Lebesgue-négligeable.

Toute réunion dénombrable d'ensembles Lebesgue-négligeables est Lebesgue-négligeable.

Une propriété $\mathcal{P}(x)$ dépendant de $x \in \Omega \subset \mathbf{R}^N$ est dite vraie presque partout si l'ensemble des x pour lesquels $\mathcal{P}(x)$ est fausse est Lebesgue-négligeable dans $\mathbf{R}^N$.

Exemples : la propriété $\mathcal{P}(x)$ peut être l'énoncé "$f_n(x) \to f(x)$ lorsque $n \to +\infty$", où $(f_n)_{n\geq 0}$ est une suite de fonctions définies sur Ω, auquel cas on dira "$f_n(x) \to f(x)$ p.p. en $x \in \Omega$ lorsque $n \to +\infty$" ; ou encore "$f(x) = 0$", où f est une fonction définie sur Ω, auquel cas on dira "$f(x) = 0$ p.p. en $x \in \Omega$".

Fonctions mesurables

Une fonction définie p.p. sur $\Omega \subset \mathbf{R}^N$ ouvert à valeurs dans $\mathbf{R}$ ou $\mathbf{C}$ est mesurable si elle est limite p.p. sur Ω d'une suite de fonctions continues à supports compacts dans Ω.

La somme et le produit de deux fonctions mesurables sur Ω à valeurs dans $\mathbf{R}$ ou $\mathbf{C}$ sont mesurables

Si $f_1, \ldots, f_n$ sont mesurables sur Ω à valeurs dans $\mathbf{R}$ ou $\mathbf{C}$, et si Φ est une application continue de $\mathbf{R}^N$ ou $\mathbf{C}^N$ dans $\mathbf{R}$ ou $\mathbf{C}$, la fonction composée $x \mapsto \Phi(f_1(x), \ldots, , f_n(x))$ est mesurable sur Ω.

La limite p.p. d'une suite de fonctions mesurables sur Ω est mesurable sur Ω ; si f est une fonction mesurable sur Ω à valeurs dans $\mathbf{C}$ telle que $f(x) \neq 0$ p.p. sur Ω, la fonction $1/f$ est également mesurable sur Ω.

Intégrale des fonctions mesurables positives

Il existe un unique prolongement de l'intégrale usuelle définie sur $C_c(\Omega)$ à l'ensemble des fonctions mesurables positives p.p. sur Ω vérifiant le

Théorème de convergence monotone Soit $(f_n)_{n\geq 0}$ suite de fonctions mesurables positives p.p. sur Ω. Supposons que

$$f_0(x) \leq f_1(x) \leq \ldots \leq f_n(x) \leq \ldots \text{ p.p. en } x \in \Omega\,.$$

Alors

$$\int_\Omega \left(\lim_{n\to+\infty} f_n(x) \right) dx = \lim_{n\to+\infty} \int_\Omega f_n(x)dx\,.$$

Soit f une fonction mesurable positive sur Ω, alors

$$\int_\Omega f(x)dx < +\infty$$

si et seulement si il existe une suite $(\phi_n)_{n\geq 0}$ de fonctions continues à supports compacts dans Ω telle que

$$\begin{cases} \phi_0(x) \leq \phi_1(x) \leq \ldots \leq \phi_n(x) \leq \ldots \text{ pour tout } x \in \Omega\,, \\ f(x) \leq \lim_{n\to+\infty} \phi_n(x) \text{ p.p. en } x \in \Omega\,, \\ \sup_{n\geq 0} \int_\Omega \phi_n(x)dx < +\infty\,. \end{cases}$$

Si tel n'est pas le cas, on posera

$$\int_\Omega f(x)dx = +\infty\,.$$

Autrement dit, l'intégrale d'une fonction mesurable positive définie p.p. sur Ω est un élément de $[0, +\infty]$.

Fonctions sommables

Une fonction f définie p.p. sur Ω et à valeurs dans $\mathbf{R}$ est sommable sur Ω si f est mesurable sur Ω et que de plus

$$\int_\Omega |f(x)|dx < +\infty\,.$$

L'ensemble $\mathcal{L}^1(\Omega)$ des fonctions sommables sur Ω est un espace vectoriel sur $\mathbf{R}$ pour l'addition des fonctions définies p.p. sur Ω et la multiplication de ces fonctions par les scalaires, et on définit le prolongement de l'intégrale usuelle à $\mathcal{L}^1(\Omega)$ par la formule

$$\int_\Omega f(x)dx = \int_\Omega f^+(x)dx - \int_\Omega f^-(x)dx$$

où les fonctions $f^+ = \max(f, 0)$ et $f^- = \max(-f, 0)$ sont toutes les deux mesurables et positives sur Ω. Le prolongement de l'intégrale ainsi défini, appelé intégrale de Lebesgue, définit une forme linéaire sur $\mathcal{L}^1(\Omega)$. Elle vérifie le

Théorème de convergence dominée Soit $(f_n)_{n\geq 0}$ suite de fonctions mesurables définies p.p. sur Ω. Supposons que

$$f_n(x) \to f(x) \text{ p.p. en } x \in \Omega$$

lorsque $n \to +\infty$, et qu'il existe $F \in \mathcal{L}^1(\Omega)$ telle que

$$|f_n(x)| \leq F(x) \text{ p.p. en } x \in \Omega \text{ pour tout } n \in \mathbf{N}\,.$$

Alors

$$\int_\Omega f_n(x)dx \to \int_\Omega f(x)dx\,.$$

Enfin, on utilisera souvent les deux résultats élémentaires suivants :

a) Toute fonction f sommable sur Ω vérifie $|f(x)| < +\infty$ p.p. en $x \in \Omega$.

b) Soit $f \in \mathcal{L}^1(\Omega)$; alors

$$\int_\Omega |f(x)|dx = 0 \Leftrightarrow f(x) = 0 \text{ p.p. en } x \in \Omega\,.$$

Chapitre IV

Propriétés de l'intégrale de Lebesgue

On a construit l'intégrale de Lebesgue à partir de l'intégrale usuelle des fonctions continues par morceaux par un procédé de "prolongement par continuité de l'intégrale" basé sur la convergence monotone. Comme nous l'avons souligné à maintes reprises au cours du chapitre précédent, l'intégrale de Lebesgue n'est rien d'autre qu'un prolongement de l'intégrale usuelle — c'est-à-dire qu'elle coïncide avec l'intégrale usuelle sur les fonctions continues à support compact dans un ouvert de $\mathbf{R}^N$, ou continues par morceaux sur un segment de $\mathbf{R}$.

Mais, pour pouvoir manipuler l'intégrale de Lebesgue sur des fonctions sommables générales, nous allons devoir en étudier les principales propriétés — dont certaines sont déjà familières au lecteur dans le cadre de la théorie usuelle de l'intégration.

On commencera donc dans ce chapitre par une étude des principales inégalités entre intégrales basées sur la convexité, inégalités constituant une vaste généralisation de l'inégalité de Cauchy-Schwarz déjà connue du lecteur.

Puis on étendra à l'intégrale de Lebesgue la plupart des formules bien connues sur l'intégrale usuelle, comme par exemple

a) l'interversion de l'ordre des intégrations dans les intégrales multiples, et

b) la formule du changement de variables dans les intégrales multiples.

Ce chapitre est une première occasion de mettre en œuvre les notions présen- tées au chapitre précédent lors de la construction de l'intégrale de Lebesgue. Les raisonnements sous-jacents méritent d'être étudiés en détail par le lecteur souhaitant maîtriser cette nouvelle notion d'intégrale.

1 Inégalités intégrales

Les différentes inégalités intégrales que nous présentons dans cette section sont d'un usage constant aussi bien en analyse qu'en calcul des probabilités. On s'en sert notamment pour vérifier

qu'une fonction mesurable vérifiant certaines propriétés est sommable.

1.1 Inégalité de Jensen

La plus fondamentale de toutes les inégalités intégrales basées sur la convexité est l'inégalité de Jensen.

INÉGALITÉ DE JENSEN

Soient f, g deux fonctions mesurables définies p.p. sur Ω ouvert de $\mathbf{R}^N$ et à valeurs réelles, et soit $\Phi : \mathbf{R} \to \mathbf{R}$ convexe. Supposons que

$$g \geq 0 \text{ p.p. sur } \Omega \quad \text{ et } \quad \int_\Omega g(x)dx = 1\,,$$

tandis que

$$fg \text{ et } \Phi(f)g \in \mathcal{L}^1(\Omega)\,.$$

Alors

$$\Phi\left(\int_\Omega f(x)g(x)dx\right) \leq \int_\Omega \Phi(f(x))g(x)dx\,.$$

Commençons par rappeler une propriété fondamentale des fonctions convexes sur $\mathbf{R}$.

Lemme 1.1.1 (Fonctions convexes sur $\mathbf{R}$). Soit I intervalle ouvert de $\mathbf{R}$ et soit $F : I \to \mathbf{R}$ convexe. Alors

a) F est continue et localement lipschitzienne sur I ;

b) F admet en tout point de l'intervalle I une dérivée à droite F'_d et une dérivée à gauche F'_g ;

c) les fonctions F'_d et F'_g sont croissantes sur I, et vérifient $F'_g \leq F'_d$ sur I ;

d) pour tout $x_0 \in I$ et tout $a \in [F'_g(x_0), F'_d(x_0)]$, on a

$$F(x) \geq F(x_0) + a(x - x_0)\,, \qquad \text{pour tout } x \in I\,.$$

Lorsque F est dérivable en x_0 (c'est-à-dire lorsque $F'_g(x_0) = F'_d(x_0)$), la propriété signifie que le graphe de la fonction convexe F reste au-dessus de sa tangente au point d'abscisse x_0 sur tout l'intervalle I. C'est cette propriété que nous allons utiliser pour établir l'inégalité de Jensen.

Démonstration. Partons de l'inégalité de convexité

$$F((1-\theta)x_1 + \theta x_2) \leq (1-\theta)F(x_1) + \theta F(x_2)$$

vérifiée pour tout $x_1, x_2 \in I$ et tout $\theta \in [0, 1]$.

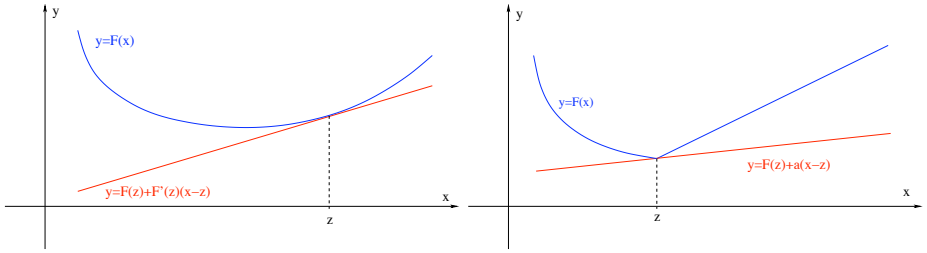


FIGURE IV.1 – A gauche : le graphe de la fonction convexe F reste au-dessus de sa tangente au point d'abscisse z ; à droite, le graphe de la fonction convexe F reste au dessus de la droite de pente $a \in [F'_g(z), F'_d(z)]$ passant par le point d'abscisse $(z, F(z))$.

Si $x < z < y$ sont trois points de I, en prenant $x_1 = x$, $x_2 = y$ et $\theta = \frac{z-x}{y-x}$, on trouve que

$$\frac{F(z) - F(x)}{z - x} \leq \frac{F(y) - F(x)}{y - x}$$

de sorte que la fonction $z \mapsto \frac{F(z)-F(x)}{z-x}$ est croissante sur $I \cap]x, +\infty[$. Elle admet donc une limite à droite en x, qui est $F'_d(x)$ par définition. On démontre de même que la fonction $z \mapsto \frac{F(y)-F(z)}{y-z}$ est croissante sur $I \cap]-\infty, y[$. Elle a donc une limite à gauche en y, qui est $F'_g(y)$ par définition. Enfin, en écrivant l'inégalité de convexité sous la forme

$$(1 - \theta)(F(z) - F(x)) \leq \theta(F(y) - F(z))$$

on trouve que

$$\frac{F(z) - F(x)}{z - x} \leq \frac{F(y) - F(z)}{y - z}$$

et en passant à la limite dans cette inégalité pour $x \nearrow z$ et $y \searrow z$, on trouve que $F'_g(z) \leq F'_d(z)$, ce qui établit le b) et la deuxième moitié du c).

Réécrivons cette dernière inégalité, cette fois avec quatre points successifs $x < t < z < y$, en raisonnant successivement avec les triplets de points $x < t < z$, puis $t < z < y$: on trouve que

$$F'_d(x) \leq \frac{F(t) - F(x)}{t - x} \leq \frac{F(z) - F(t)}{z - t} \leq \frac{F(y) - F(z)}{y - z} \leq F'_g(y)\,.$$

En particulier, F est lipschitzienne sur $[x, y]$ puisque

$$|F(z) - F(t)| \leq \max(|F'_d(x)|, |F'_g(y)|)|z - t|$$

et donc, comme $x < y$ sont arbitraires dans l'intervalle ouvert I, continue et localement lipschitzienne sur I, ce qui établit le a).

Puis en passant à la limite pour $z \searrow t$, puis pour $t \nearrow z$ dans l'avant-dernière chaîne d'inégalités, on trouve que

$$F'_d(x) \leq \frac{F(t) - F(x)}{t - x} \leq F'_d(t)\,, \qquad \text{et} \quad F'_g(z) \leq \frac{F(y) - F(z)}{y - z} \leq F'_g(y)\,,$$

d'où l'on tire le fait que F'_d et F'_g sont toutes les deux croissantes sur I, ce qui termine le c).

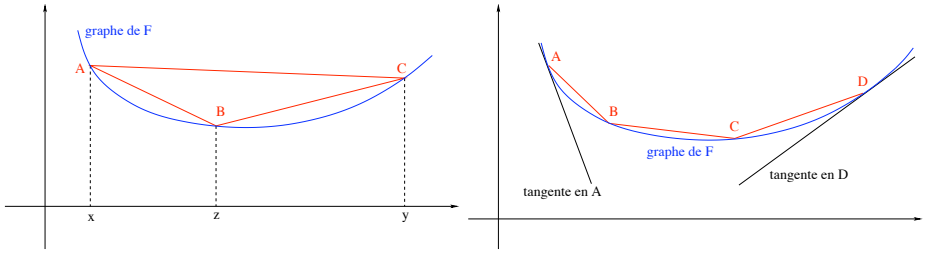


FIGURE IV.2 – A gauche : les pentes des segments $[A,B]$, $[A,C]$ et $[B,C]$ sont ordonnées de façon croissante ; à droite : les pentes de la tangente au point A, des segments $[A,B]$, $[B,C]$, $[C,D]$ et de la tangente en D sont ordonnées de manière croissante.

On en tire aussi les deux inégalités

$$F(t) \geq F(x) + F'_d(x)(t-x)\,, \quad F(z) \geq F(y) + F'_g(y)(z-y)\,,$$

que l'on peut réécrire sous la forme

$$\begin{aligned} F(x) &\geq F(x_0) + F'_d(x_0)(x-x_0) \quad \text{si } x_0 < x \in I\,,\\ F(x) &\geq F(x_0) + F'_g(x_0)(x-x_0) \quad \text{si } x_0 > x \in I\,, \end{aligned}$$

qui se combinent pour donner le d). □

Une fois que l'on dispose de la propriété d) du lemme ci-dessus, la démonstration de l'inégalité de Jensen est presqu'immédiate.

Démonstration. Posons

$$m = \int_\Omega f(x)g(x)dx$$

et appliquons la propriété d) du lemme au cas où $I = \mathbf{R}$, $F = \Phi$ et $x_0 = m$. On trouve alors que

$$\Phi(f(x)) - \Phi(m) - \Phi'_d(m)(f(x)-m) \geq 0 \text{ p.p. en } x \in \Omega\,,$$

de sorte que, en multipliant chaque membre de cette inégalité par $g(x) \geq 0$ p.p. en $x \in \Omega$, on trouve que

$$\Phi(f(x))g(x) - \Phi(m)g(x) - \Phi'_d(m)(f(x)g(x) - mg(x)) \geq 0 \text{ p.p. en } x \in \Omega\,.$$

Par hypothèse, le membre de gauche de cette inégalité est une fonction sommable de x définie p.p. sur Ω (comme combinaison linéaire des fonctions sommables $\Phi(f)g$, g et fg) et à valeurs positives. Son intégrale est donc un nombre positif et on a

$$\begin{aligned} 0 &\leq \int_\Omega \left(\Phi(f(x))g(x) - \Phi(m)g(x) - \Phi'_d(m)(f(x)g(x) - mg(x))\right) dx\\ &= \int_\Omega \Phi(f(x))g(x)dx - \Phi(m)\int_\Omega g(x)dx - \Phi'_d(m)\left(m - m\int_\Omega g(x)dx\right)\\ &= \int_\Omega \Phi(f(x))g(x)dx - \Phi(m) \end{aligned}$$

puisque

$$\int_\Omega g(x)dx = 1\,,$$

ce qui donne l'inégalité cherchée. □

1.2 Inégalité de Hölder

L'inégalité de Hölder est une généralisation de l'inégalité de Cauchy-Schwartz. On s'en sert le plus souvent pour montrer qu'un produit fg de fonctions mesurables est sommable, pourvu que des puissances bien choisies de f et de g soient sommables.

Inégalité de Hölder

Soient $\Omega \subset \mathbf{R}^N$ ouvert, $f, g : \Omega \to \mathbf{R}$ mesurables, et soient $p, p' > 1$ t.q. $\frac{1}{p} + \frac{1}{p'} = 1$. Alors

$$\int_\Omega |f(x)g(x)|dx \leq \left(\int_\Omega |f(x)|^p dx\right)^{1/p} \left(\int_\Omega |g(x)|^{p'} dx\right)^{1/p'}$$

Lemme 1.2.1 (Une inégalité auxiliaire). Soit $p > 1$ et $p' = \frac{p}{p-1}$, de sorte que $\frac{1}{p} + \frac{1}{p'} = 1$. Alors, pour tous $X, Y > 0$, on a

$$XY \leq \frac{X^p}{p} + \frac{Y^{p'}}{p'}\,.$$

La démonstration de ce lemme consiste simplement à étudier la fonction

$$]0, +\infty[\ni X \mapsto XY - \frac{X^p}{p}\,.$$

Cette fonction est de classe C^∞, strictement concave, tend vers 0 pour $X \to 0^+$, et tend vers $-\infty$ pour $X \to +\infty$; elle atteint son maximum sur $]0, +\infty[$, en l'unique point qui annule sa dérivée $X \mapsto Y - X^{p-1}$ sur $]0, +\infty[$, soit $X := Y^{1/(p-1)}$. Ce maximum vaut donc

$$Y \cdot Y^{1/(p-1)} - \frac{Y^{p/(p-1)}}{p} = \frac{Y^{p'}}{p'}\,,$$

d'où

$$XY - \frac{X^p}{p} \leq \frac{Y^{p'}}{p'}\,, \quad \text{pour tout } X > 0\,.$$

Démonstration de l'inégalité de Hölder. Supposons pour commencer que

$$I := \left(\int_\Omega |f(x)|^p dx\right)^{1/p} \in]0, +\infty[\quad \text{et que} \quad J := \left(\int_\Omega |g(x)|^{p'} dx\right)^{1/p'} \in]0, +\infty[\,.$$

Ecrivons l'inégalité du lemme pour $X = |f(x)|/I$ et $Y = |g(x)|/J$: on aboutit à

$$\frac{|f(x)||g(x)|}{IJ} \leq \frac{1}{p}\frac{|f(x)|^p}{I^p} + \frac{1}{p'}\frac{|g(x)|^{p'}}{J^{p'}}\,, \quad \text{p.p. en } x \in \Omega\,.$$

Chaque membre de cette inégalité est une fonction mesurable positive de la fonction x ; d'autre part, le membre de droite est sommable comme combinaison linéaire des fonctions $|f|^p$ et $|g|^{p'}$ supposées sommables pour l'instant. D'après le théorème de comparaison du chapitre III page 27 (Théorème 7.0.22 page 64), on conclut que $|f||g|$ est également sommable sur Ω et que

$$\begin{aligned}\frac{1}{IJ}\int_\Omega |f(x)||g(x)|dx &\le \frac{1}{p}\frac{1}{I^p}\int_\Omega |f(x)|^p dx + \frac{1}{p'}\frac{1}{J^{p'}}\int_\Omega |g(x)|^{p'}dx \\ &= \frac{1}{p}\frac{1}{I^p}I^p + \frac{1}{p'}\frac{1}{J^{p'}}J^{p'} = \frac{1}{p}+\frac{1}{p'} = 1\end{aligned}$$

d'où l'on tire que

$$\int_\Omega |f(x)||g(x)|dx \le IJ\,,$$

ce qui est l'inégalité cherchée.

Si on a seulement $I > 0$ et $J > 0$ et que l'un de ces deux nombres vaut $+\infty$, le membre de droite de l'inégalité de Hölder vaut $+\infty$ et donc l'inégalité est trivialement vérifiée.

Si l'un des deux nombres I ou J vaut 0, disons $I = 0$, alors $f = 0$ p.p. sur Ω, de sorte que $|f||g| = 0$ p.p. sur Ω, et le membre de gauche de l'inégalité vaut 0, de sorte que l'inégalité est encore vraie. □

Lorsque $p = 2$, l'inégalité de Hölder se réduit à l'inégalité de Cauchy-Schwarz, que le lecteur connaît déjà.

1.3 Inégalité de Minkowski

L'inégalité de Minkowski joue également un rôle considérable en analyse. Nous la retrouverons au chapitre VI page 111, où elle nous servira à définir certains espaces de fonctions généralisant l'espace $\mathcal{L}^1(\Omega)$ des fonctions sommables sur Ω.

Inégalité de Minkowski

Soient $\Omega \subset \mathbf{R}^N$ ouvert, $f, g : \Omega \to \mathbf{R}$ mesurables, et $p \in [1, +\infty[$. Alors

$$\left(\int_\Omega |f(x)+g(x)|^p dx\right)^{1/p} \le \left(\int_\Omega |f(x)|^p dx\right)^{1/p} + \left(\int_\Omega |g(x)|^p dx\right)^{1/p}$$

Démonstration. Le cas $p = 1$ résulte immédiatement de l'inégalité triangulaire

$$|f(x) + g(x)| \le |f(x)| + |g(x)|\,, \qquad \text{p.p. en } x \in \Omega\,.$$

Supposons donc que $p > 1$; écrivons que

$$\begin{aligned}\int_\Omega |f(x)+g(x)|^p dx &\le \int_\Omega |f(x)||f(x)+g(x)|^{p-1}dx \\ &\quad + \int_\Omega |g(x)||f(x)+g(x)|^{p-1}dx = A + B\,.\end{aligned}$$

D'après l'inégalité de Hölder

$$\begin{aligned} A &\leq \left(\int_\Omega |f(x)|^p dx\right)^{1/p} \left(\int_\Omega |f(x)+g(x)|^{(p-1)p'} dx\right)^{1/p'} \\ &= \left(\int_\Omega |f(x)|^p dx\right)^{1/p} \left(\int_\Omega |f(x)+g(x)|^p dx\right)^{1-1/p} \end{aligned}$$

puisque $p' = \frac{p}{p-1}$ et $\frac{1}{p} + \frac{1}{p'} = 1$.

De même

$$B \leq \left(\int_\Omega |g(x)|^p dx\right)^{1/p} \left(\int_\Omega |f(x)+g(x)|^p dx\right)^{1-1/p},$$

de sorte que

$$\begin{aligned} \int_\Omega |f(x)+g(x)|^p dx &\leq A + B \\ &\leq \left(\int_\Omega |f(x)+g(x)|^p dx\right)^{1-1/p} \left(\left(\int_\Omega |f(x)|^p dx\right)^{1/p} + \left(\int_\Omega |g(x)|^p dx\right)^{1/p}\right). \end{aligned}$$

Alors, ou bien

$$\int_\Omega |f(x)+g(x)|^p dx = 0$$

et l'inégalité de Minkowski est évidente, ou bien

$$\int_\Omega |f(x)+g(x)|^p dx \neq 0$$

auquel cas elle découle de la dernière inégalité ci-dessus dont on divise chaque membre par

$$\left(\int_\Omega |f(x)+g(x)|^p dx\right)^{1-1/p} > 0\,.$$

□

2 Intégrales multiples

Commençons par rappeler le résultat classique sur l'interversion de l'ordre des intégrations successives dans une intégrale double.

Soit $f \in C(K \times K)$, où K est un segment de $\mathbf{R}$. Alors

a) les fonctions

$$x \mapsto \int_K f(x,y)dy\,, \quad \text{et } y \mapsto \int_K f(x,y)dx$$

sont continues sur K, et

b) on a

$$\int_K \left(\int_K f(x,y)dy \right) dx = \int_K \left(\int_K f(x,y)dx \right) dy \,,$$

c) enfin, la valeur commune obtenue au b) coïncide avec l'intégrale de f sur $K \times K$ obtenue par la définition de l'intégrale des fonctions continues à support compact donnée au chapitre III page 27, c'est-à-dire que

$$\iint_{K\times K} f(x,y)dxdy = \int_K \left(\int_K f(x,y)dy \right) dx = \int_K \left(\int_K f(x,y)dx \right) dy \,.$$

Plus généralement (par exemple en appliquant le résultat ci-dessus de façon itérative), on arrive au résultat suivant.

Soient $\Omega_1 \subset \mathbf{R}^{N_1}$ et $\Omega_2 \subset \mathbf{R}^{N_2}$ ouverts, et soit $f \in C_c(\Omega_1 \times \Omega_2)$. Alors

$$\Omega_1 \ni x_1 \mapsto \int_{\Omega_2} f(x_1,x_2)dx_2 \text{ appartient à } C_c(\Omega_1)\,,$$

et

$$\begin{aligned} \iint_{\Omega_1\times\Omega_2} f(x_1,x_2)dx_1dx_2 &= \int_{\Omega_1} \left(\int_{\Omega_2} f(x_1,x_2)dx_2 \right) dx_1 \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} f(x_1,x_2)dx_1 \right) dx_2 \,. \end{aligned}$$

(en échangeant les rôles des variables x_1 et x_2.)

Le but de cette section est d'étendre ce résultat au cadre de l'intégrale de Lebesgue. Le résultat principal dans cette direction est le

Théorème 2.0.1 (de Fubini). Soient Ω_1 ouvert de $\mathbf{R}^{N_1}$ et Ω_2 ouvert de $\mathbf{R}^{N_2}$. Soit $f \in \mathcal{L}^1(\Omega_1 \times \Omega_2; \mathbf{C})$. Alors

a) p.p. en $x_2 \in \Omega_2$, la fonction $x_1 \mapsto f(x_1,x_2)$ appartient à $\mathcal{L}^1(\Omega_1;\mathbf{C})$;

b) la fonction $x_2 \mapsto \displaystyle\int_{\Omega_1} f(x_1,x_2)dx_1$ définie p.p. sur Ω_1 appartient à $\mathcal{L}^1(\Omega_2;\mathbf{C})$;

c) enfin, on a l'égalité

$$\begin{aligned} \iint_{\Omega_1\times\Omega_2} f(x_1,x_2)dx_1dx_2 &= \int_{\Omega_1} \left(\int_{\Omega_2} f(x_1,x_2)dx_2 \right) dx_1 \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} f(x_1,x_2)dx_1 \right) dx_2 \,, \end{aligned}$$

(en échangeant les rôles des variables x_1 et x_2.)

Nous allons démontrer le théorème de Fubini pour les fonctions sommables en supposant connu le cas des fonctions continues à support compact. En fait, le point crucial est la généralisation au cas où $f \in \mathcal{L}^+(\Omega_1 \times \Omega_2)$, que l'on démontre par un argument de convergence monotone à partir des suites fondamentales.

Lemme 2.0.2. Soit $f \in \mathcal{L}^+(\Omega_1 \times \Omega_2)$. Alors

a) p.p. en $x_2 \in \Omega_2$, la fonction $x_1 \mapsto f(x_1, x_2)$ appartient à $\mathcal{L}^+(\Omega_1)$;

b) la fonction $x_2 \mapsto \displaystyle\int_\Omega f(x_1, x_2)dx_1$ appartient à $\mathcal{L}^+(\Omega_2)$;

c) enfin, on a

$$\iint_{\Omega_1\times\Omega_2} f(x_1, x_2)dx_1dx_2 = \int_{\Omega_1} \left(\int_{\Omega_2} f(x_1, x_2)dx_2 \right) dx_1 .$$

Démonstration. Soit $(f_n)_{n\geq 0}$ suite fondamentale convergeant simplement vers f sur $\Omega_1 \times \Omega_2$.

Etape 1 : Posons

$$F_n(x_2) := \int_{\Omega_1} f_n(x_1, x_2)dx_1 , \quad x_2 \in \Omega_2 , \quad n \geq 0 .$$

Comme $f_n \in C_c(\Omega_1 \times \Omega_2)$, la fonction $F_n \in C_c(\Omega_2)$, d'après le résultat sur les fonctions continues à support compact rappelé avant l'énoncé du théorème de Fubini. De plus

$$f_0 \leq \ldots \leq f_n \leq f_{n+1} \leq \ldots \text{ sur } \Omega_1 \times \Omega_2$$

implique que

$$F_0 \leq \ldots \leq F_n \leq F_{n+1} \leq \ldots \text{ sur } \Omega_2 .$$

Enfin

$$\begin{aligned} \int_{\Omega_2} F_n(x_2)dx_2 &= \iint_{\Omega_1\times\Omega_2} f_n(x_1, x_2)dx_1dx_2 \\ &\leq \iint_{\Omega_1\times\Omega_2} f(x_1, x_2)dx_1dx_2 < +\infty , \end{aligned}$$

de sorte que la suite $(F_n)_{n\geq 0}$ est fondamentale sur Ω_2. Nommons F sa limite simple sur Ω_2 : par définition

$$F \in \mathcal{L}^+(\Omega_2) \text{ et } \int_{\Omega_2} F(x_2)dx_2 = \lim_{n\to+\infty} \int_{\Omega_2} F_n(x_2)dx_2 .$$

Etape 2 : Soit $\mathcal{N} := \{x_2 \in \Omega_2 \,|\, F(x_2) = +\infty\}$; comme $F \in \mathcal{L}^+(\Omega_2)$, l'ensemble $\mathcal{N}$ est négligeable. Pour $x_2 \in \Omega_2 \setminus \mathcal{N}$, la suite de fonctions $x_1 \mapsto f_n(x_1, x_2)$ est fondamentale sur Ω_1 puisque

$$f_0(\cdot, x_2) \leq \ldots \leq f_n(\cdot, x_2) \leq f_{n+1}(\cdot, x_2) \leq \ldots \text{ sur } \Omega_1$$

et

$$\int_{\Omega_2} f_n(x_1, x_2)dx_1 = F_n(x_2) \leq F(x_2) < \infty , \quad n \geq 0 .$$

Comme $f_n(\cdot, x_2) \to f(\cdot, x_2)$ simplement sur Ω_1 pour tout $x_2 \in \Omega_2$, on en déduit que $f(\cdot, x_2) \in \mathcal{L}^+(\Omega_1)$ pour tout $x_2 \in \Omega_2 \setminus \mathcal{N}$, c'est-à-dire p.p. en $x_2 \in \Omega_2$, ce qui établit le a).

Etape 3 : De plus, par convergence monotone

$$\int_{\Omega_1} f(x_1, x_2)dx_1 = \lim_{n\to+\infty} \int_{\Omega_1} f_n(x_1, x_2)dx_1 = \lim_{n\to+\infty} F_n(x_2) = F(x_2) .$$

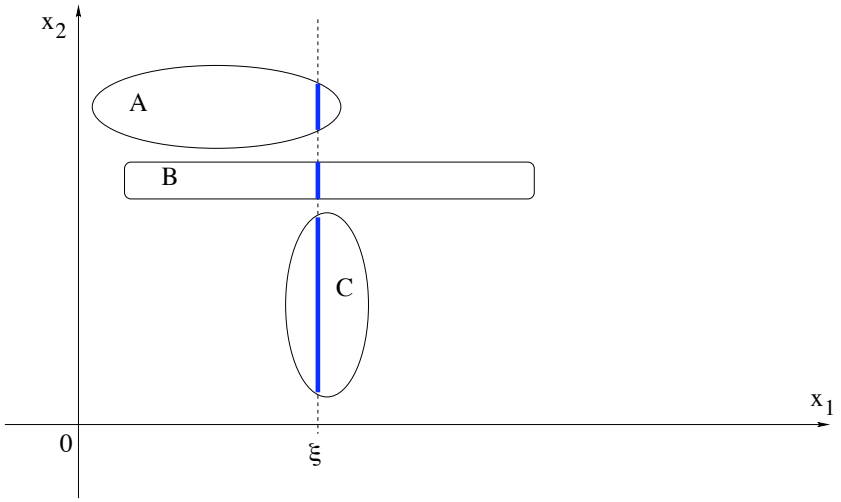


FIGURE IV.3 – La fibre de $A \cup B \cup C$ au-dessus de ξ (en trait bleu).

Or on a vu plus haut que $F \in \mathcal{L}^+(\Omega_2)$, de sorte que la fonction

$$x_2 \mapsto \int_{\Omega_2} f(x_1, x_2) dx_1$$

appartient à $\mathcal{L}^+(\Omega_2)$, ce qui établit le b).

Etape 4 : Enfin

$$\begin{aligned}
\iint_{\Omega_1 \times \Omega_2} f(x_1, x_2) dx_1 dx_2 &= \lim_{n \to +\infty} \iint_{\Omega_1 \times \Omega_2} f_n(x_1, x_2) dx_1 dx_2 \\
&= \lim_{n \to +\infty} \int_{\Omega_2} \left(\int_{\Omega_1} f_n(x_1, x_2) dx_1 \right) dx_2 \\
&= \lim_{n \to +\infty} \int_{\Omega_2} F_n(x_2) dx_2 \\
&= \int_{\Omega_2} F(x_2) dx_2 = \int_{\Omega_2} \left(\int_{\Omega_1} f(x_1, x_2) dx_1 \right) dx_2 \,.
\end{aligned}$$

En effet la première égalité découle du fait que $(f_n)_{n \geq 0}$ est une suite fondamentale convergeant simplement vers f sur $\Omega_1 \times \Omega_2$, la deuxième de l'interversion de l'ordre d'intégration dans le cas (déjà connu) d'intégrales multiples de fonctions continues à support compact, la quatrième de l'étape 1 ci-dessus et la cinquième de l'égalité obtenue dans l'étape 3 ci-dessus. □

Une conséquence importante du théorème de Fubini pour les fonctions de la classe $\mathcal{L}^+$ est le lemme ci-dessous portant sur les fibres d'ensembles négligeables dans l'espace produit $\Omega_1 \times \Omega_2$.

Lemme 2.0.3 (Fibres des ensembles négligeables). Soit $\mathcal{N} \subset \Omega_1 \times \Omega_2$ négligeable. Pour $x_1 \in \Omega_1$ on note

$$\mathcal{N}_{x_1} := \{x_2 \in \Omega_2 \,|\, (x_1, x_2) \in \mathcal{N}\} \subset \Omega_2$$

la fibre de $\mathcal{N}$ au-dessus de x_1 (voir Figure 2 page ci-contre.) Alors, p.p. en $x_1 \in \Omega_1$,

$$\mathcal{N}_{x_1} \text{ est négligeable dans } \Omega_2 \,.$$

Démonstration. Soit $f \in \mathcal{L}^+(\Omega_1 \times \Omega_2)$ telle que $f(x_1, x_2) = +\infty$ pour tout $(x_1, x_2) \in \mathcal{N}$.

D'après le Lemme 2.0.2 page 79 a) (en échangeant les rôles des variables x_1 et x_2), il existe $\mathcal{N}_1 \subset \Omega_1$ négligeable tel que, pour $x_1 \in \Omega_1 \setminus \mathcal{N}_1$, la fonction $f(x_1, \cdot) \in \mathcal{L}^+(\Omega_2)$. Or

$$\mathcal{N}_{x_1} \subset f(x_1, \cdot)^{-1}(\{+\infty\}) \,,$$

donc $\mathcal{N}_{x_1}$ est négligeable dans Ω_2 pour tout $x_1 \in \Omega \setminus \mathcal{N}_1$, c'est-à-dire p.p. en $x_1 \in \Omega_1$. □

On prendra bien garde au fait que les fibres $\mathcal{N}_{x_1}$ de l'ensemble négligeable $\mathcal{N}$ ne sont négligeables dans Ω_2 que p.p. en $x_1 \in \Omega_1$, et pas pour tout $x_1 \in \Omega_1$ en général ; le lecteur est invité a trouver un contre-exemple à titre d'exercice.

Démonstration du théorème de Fubini. Comme $f \in \mathcal{L}^1(\Omega_1 \times \Omega_2)$, il existe $g, h \in \mathcal{L}^+(\Omega_1 \times \Omega_2)$ telles que $f = g - h$ p.p. sur $\Omega_1 \times \Omega_2$. Autrement dit, il existe $\mathcal{N} \subset \Omega_1 \times \Omega_2$ négligeable tel que $f = g - h$ sur $\Omega_1 \times \Omega_2 \setminus \mathcal{N}$.

Appliquons le Lemme 2.0.2 page 79 a) aux fonctions g et h : il existe $\mathcal{N}_2' \subset \Omega_2$ négligeable tel que les fonctions $g(\cdot, x_2)$ et $h(\cdot, x_2)$ appartiennent à $\mathcal{L}^+(\Omega_1)$ pour tout $x_2 \in \Omega_2 \setminus \mathcal{N}_2'$; d'après le Lemme 2.0.3 page précédente, il existe $\mathcal{N}_2 \subset \Omega_2$ négligeable tel que la fibre $\mathcal{N}_{x_2}$ soit négligeable dans Ω_1 pour tout $x_2 \in \Omega_2 \setminus \mathcal{N}_2$. Alors, pour tout $x_2 \in \Omega_2 \setminus (\mathcal{N}_2' \cup \mathcal{N}_2)$, on a

$$f(\cdot, x_2) = g(\cdot, x_2) - h(\cdot, x_2) \in \mathcal{L}^1(\Omega_1) \,,$$

ce qui établit le a).

D'après le Lemme 2.0.2 page 79 b), les fonctions

$$\Omega_2 \setminus (\mathcal{N}_2' \cup \mathcal{N}_2) \ni x_2 \mapsto \int_{\Omega_1} g(x_1, x_2) dx_1$$

et

$$\Omega_2 \setminus (\mathcal{N}_2' \cup \mathcal{N}_2) \ni x_2 \mapsto \int_{\Omega_1} h(x_1, x_2) dx_1$$

appartiennent à $\mathcal{L}^+(\Omega_2)$, de sorte que la fonction

$$\Omega_2 \setminus (\mathcal{N}_2' \cup \mathcal{N}_2) \ni x_2 \mapsto \int_{\Omega_1} g(x_1, x_2) dx_1 - \int_{\Omega_1} h(x_1, x_2) dx_1 = \int_{\Omega_1} f(x_1, x_2) dx_1$$

appartient à $\mathcal{L}^1(\Omega_2)$, ce qui établit le b).

Enfin, d'après le c) du Lemme 2.0.2 page 79, on trouve que

$$\begin{aligned}\iint_{\Omega_1\times\Omega_2} & f(x_1,x_2)dx_1dx_2 \\ &= \iint_{\Omega_1\times\Omega_2} g(x_1,x_2)dx_1dx_2 - \iint_{\Omega_1\times\Omega_2} h(x_1,x_2)dx_1dx_2 \\ &= \int_{\Omega_2}\left(\int_{\Omega_1} g(x_1,x_2)dx_1\right)dx_2 - \int_{\Omega_2}\left(\int_{\Omega_1} h(x_1,x_2)dx_1\right)dx_2 \\ &= \int_{\Omega_2}\left(\int_{\Omega_1} f(x_1,x_2)dx_1\right)dx_2\,,\end{aligned}$$

l'égalité restante étant obtenue en échangeant les variables x_1 et x_2. □

Comme pour le théorème de convergence monotone, le cas des fonctions mesurables positives donne lieu à un énoncé d'une simplicité maximale, qui est de plus d'une grande utilité en pratique.

Théorème 2.0.4 (de Tonelli). Soit $f : \Omega_1 \times \Omega_2 \to [0,+\infty]$ mesurable. Alors

a) p.p. en $x_2 \in \Omega_2$, la fonction $\Omega_1 \ni x_1 \mapsto f(x_1,x_2) \in [0,+\infty]$ est mesurable ;

b) la fonction $x_2 \mapsto \displaystyle\int_{\Omega_1} f(x_1,x_2)dx_1 \in [0,+\infty]$ définie p.p. sur Ω_2 est mesurable ;

c) enfin, on a l'égalité dans $[0,+\infty]$

$$\begin{aligned}\iint_{\Omega_1\times\Omega_2} f(x_1,x_2)dx_1dx_2 &= \int_{\Omega_1}\left(\int_{\Omega_2} f(x_1,x_2)dx_2\right)dx_1 \\ &= \int_{\Omega_2}\left(\int_{\Omega_1} f(x_1,x_2)dx_1\right)dx_2\,.\end{aligned}$$

Démonstration. Introduisons la suite de fonctions appartenant à $\mathcal{L}^1(\Omega_1\times\Omega_2)$ définie par la formule

$$F_n(x_1,x_2) := n\mathbf{1}_{(B(0,n)\times B(0,n))\cap(\Omega_1\times\Omega_2)}(x_1,x_2)\,.$$

Posons alors $f_n := \min(f,F_n)$ pour tout $n \in \mathbf{N}$. Par construction, f_n est, pour tout $n \in \mathbf{N}$, une fonction mesurable définie p.p. sur $\Omega_1\times\Omega_2$ et à valeurs dans $\mathbf{R}_+$; notons $\mathcal{N}' \subset \Omega_1\times\Omega_2$ l'ensemble négligeable en dehors duquel f est définie. Alors $f_n \le F_n$ sur $\Omega_1\times\Omega_2 \setminus \mathcal{N}'$, de sorte que, d'après le théorème de comparaison (Théorème 7.0.22 page 64), $f_n \in \mathcal{L}^1(\Omega_1\times\Omega_2)$.

Appliquons le théorème de Fubini à la fonction f_n : il existe donc $\mathcal{N}_{2,n} \subset \Omega_2$ négligeable tel que la fonction $f_n(\cdot,x_2) \in \mathcal{L}^1(\Omega_2)$ pour tout $x_2 \in \Omega_2 \setminus \mathcal{N}_{2,n}$, et tel que la fonction

$$\Omega_2 \setminus \mathcal{N}_{2,n} \ni x_2 \mapsto \int_{\Omega_1} f_n(x_1,x_2)dx_1$$

appartienne à $\mathcal{L}^1(\Omega_1)$.

Posons

$$\mathcal{N}_2 = \bigcup_{n\ge 0} \mathcal{N}_{2,n}\,;$$

l'ensemble $\mathcal{N}_2 \subset \Omega_2$ est négligeable comme réunion dénombrable d'ensembles négligeables. D'après le Lemme 2.0.3 page 80, il existe $\mathcal{N}_2' \subset \Omega_2$ négligeable tel que, pour tout $x_2 \in \Omega_2 \setminus \mathcal{N}_2'$, la fibre $\mathcal{N}'_{x_2}$ de l'ensemble négligeable $\mathcal{N}'$ en dehors duquel la fonction f est définie soit négligeable dans Ω_1.

Alors, pour tout $x_2 \in \Omega_2 \setminus (\mathcal{N}_2 \cup \mathcal{N}_2')$, la suite de fonctions $f_n(\cdot, x_2)$ converge simplement vers $f(\cdot, x_2)$ sur Ω_1, par construction : donc, $f(\cdot, x_2)$ est mesurable sur Ω_1 pour tout $x_2 \in \Omega_2 \setminus (\mathcal{N}_2 \cup \mathcal{N}_2')$ (comme limite p.p. d'une suite de fonctions mesurables sur Ω_1), c'est-à-dire p.p. en $x_2 \in \Omega_2$. Ceci établit le point a) du théorème de Tonelli.

De plus, toujours par construction

$$f_0 \leq f_1 \leq \ldots \leq f_n \leq f_{n+1} \text{ sur } \Omega_1 \times \Omega_2 \setminus \mathcal{N}'$$

de sorte que

$$f_0(\cdot, x_2) \leq f_1(\cdot, x_2) \leq \ldots \leq f_n(\cdot, x_2) \leq f_{n+1}(\cdot, x_2) \text{ sur } \Omega_1$$

pour tout $x_2 \in \Omega_2 \setminus \mathcal{N}_2'$. Alors, par convergence monotone

$$\int_{\Omega_1} f_n(x_1, x_2) dx_1 \to \int_{\Omega_1} f(x_1, x_2) dx_1$$

pour tout $x_2 \in \Omega_2 \setminus (\mathcal{N}_2 \cup \mathcal{N}_2')$ lorsque $n \to +\infty$. Ainsi la fonction

$$x_2 \mapsto \int_{\Omega_1} f(x_1, x_2) dx_1$$

définie sur $\Omega_2 \setminus (\mathcal{N}_2 \cup \mathcal{N}_2')$, donc p.p. sur Ω_2, est mesurable comme limite p.p. d'une suite de fonctions sommables sur Ω_2, d'où le point b) du théorème de Tonelli.

Enfin, pour tout $n \geq 0$, on a

$$\iint_{\Omega_1 \times \Omega_2} f_n(x_1, x_2) dx_1 dx_2 = \int_{\Omega_2} \left(\int_{\Omega_1} f_n(x_1, x_2) dx_1 \right) dx_2$$

d'après le point c) du théorème de Fubini appliqué à $f_n \in \mathcal{L}^1(\Omega_1 \times \Omega_2)$. On a vu que, par construction

$$f_0 \leq f_1 \leq \ldots \leq f_n \leq f_{n+1} \to f \text{ p.p. sur } \Omega_1 \times \Omega_2 \,,$$

et que

$$\begin{aligned} \int_{\Omega_1} f_0(x_1, x_2) dx_1 \leq \ldots \leq \int_{\Omega_1} f_n(x_1, x_2) dx_1 \\ \leq \int_{\Omega_1} f_{n+1}(x_1, x_2) dx_1 \to \int_{\Omega_1} f(x_1, x_2) dx_1 \end{aligned}$$

p.p. en $x_2 \in \Omega_2$ lorsque $n \to +\infty$. Donc, par convergence monotone, on obtient l'égalité suivante dans $[0, +\infty]$:

$$\begin{aligned} \iint_{\Omega_1 \times \Omega_2} f(x_1, x_2) dx_1 dx_2 = \lim_{n \to +\infty} \iint_{\Omega_1 \times \Omega_2} f_n(x_1, x_2) dx_1 dx_2 \\ = \lim_{n \to +\infty} \int_{\Omega_2} \left(\int_{\Omega_1} f_n(x_1, x_2) dx_1 \right) dx_2 = \int_{\Omega_2} \left(\int_{\Omega_1} f(x_1, x_2) dx_1 \right) dx_2 \,. \end{aligned}$$

On obtient l'autre égalité du point c) en échangeant les rôles des variables x_1 et x_2. □

On pourrait croire que le théorème de Tonelli est un cas particulier du théorème de Fubini, puisqu'il ne porte que sur les fonctions à valeurs positives. L'essentiel de ce théorème tient cependant au fait que l'égalité du point c) vaut dans $[0,+\infty]$ — le théorème de Fubini spécialisé aux fonctions positives ne donnerait la même égalité que dans $[0,+\infty[$, car il ne porte que sur les fonctions sommables. C'est pourquoi le théorème de Tonelli est d'une grande utilité en pratique, car il permet très souvent de répondre à la question suivante :

Comment vérifier qu'une fonction mesurable donnée est sommable sur $\Omega_1 \times \Omega_2$?

Etant donnée une fonction $f : \Omega_1 \times \Omega_2 \to \mathbf{C}$ mesurable, il peut arriver que l'on sache calculer "facilement"

$$I = \int_{\Omega_1} \left(\int_{\Omega_2} |f(x_1,x_2)| dx_2 \right) dx_1$$

ou

$$J = \int_{\Omega_2} \left(\int_{\Omega_1} |f(x_1,x_2)| dx_1 \right) dx_2 \,.$$

D'après le théorème de Tonelli appliqué à $|f|$,

$$\begin{aligned} I \text{ ou } J < +\infty \Rightarrow & \iint_{\Omega_1 \times \Omega_2} |f(x_1,x_2)| dx_1 dx_2 = I = J < +\infty \\ & \Rightarrow f \in \mathcal{L}^1(\Omega_1 \times \Omega_2; \mathbf{C}) \,. \end{aligned}$$

Par le même raisonnement

$$\begin{aligned} I \text{ ou } J = +\infty \Rightarrow & \iint_{\Omega_1 \times \Omega_2} |f(x_1,x_2)| dx_1 dx_2 = I = J = +\infty \\ & \Rightarrow f \notin \mathcal{L}^1(\Omega_1 \times \Omega_2; \mathbf{C}) \,. \end{aligned}$$

3 Changement de variables

Commençons par rappeler la formule du changement de variables dans les intégrales multiples pour des fonctions continues, sous la forme déjà connue du lecteur.

Soient Ω_1, Ω_2 deux ouverts de $\mathbf{R}^N$, et un C^1-difféomorphisme $\Phi : \Omega_1 \to \Omega_2$. Rappelons que ceci signifie que Φ est une bijection de classe C^1 sur Ω_1 dont l'inverse Φ^{-1} est de classe C^1 sur Ω_2.

En pratique, on cherche évite de vérifier que Φ^{-1} est de classe C^1 sur Ω_2, car il n'est pas toujours possible de calculer explicitement l'inverse Φ^{-1}.

On utilise alors la caractérisation suivante : l'application $\Phi : \Omega_1 \to \Omega_2$ est un C^1-difféomorphisme si et seulement si

$$\Phi \text{ est une bijection de classe } C^1 \text{ sur } \Omega_1 \text{ et } \det J\Phi(x) \neq 0 \text{ pour tout } x \in \Omega_1 \,,$$

en notant $J\Phi(x)$ la matrice jacobienne de Φ au point $x \in \Omega_1$.

Voici la formule du changement de variables dans le cadre de l'intégrale usuelle des fonctions continues.

Pour tout $f \in C_c(\Omega_2)$, alors $f \circ \Phi \in C_c(\Omega_1)$ et

$$\int_{\Omega_2} f(y)dy = \int_{\Omega_1} f(\Phi(x))|\det J\Phi(x)|dx\,.$$

C'est cette formule que nous allons généraliser au cas des fonctions sommables.

Théorème 3.0.5 (Formule du changement de variables). Soit $f \in \mathcal{L}^1(\Omega_2)$. Alors la fonction $f \circ \Phi|\det J\Phi| \in \mathcal{L}^1(\Omega_1)$ et on a la formule

$$\int_{\Omega_2} f(y)dy = \int_{\Omega_1} f(\Phi(x))|\det J\Phi(x)|dx\,.$$

Comme dans le cas du théorème de Fubini, nous allons commencer par établir cette formule pour les fonctions de $\mathcal{L}^+(\Omega)$.

Lemme 3.0.6. Soit $f \in \mathcal{L}^+(\Omega_2)$. Alors la fonction $f \circ \Phi|\det J\Phi| \in \mathcal{L}^+(\Omega_1)$ et on a la formule

$$\int_{\Omega_2} f(y)dy = \int_{\Omega_1} f(\Phi(x))|\det J\Phi(x)|dx\,.$$

Démonstration. Soit $(f_n)_{n\geq 0}$ suite fondamentale convergeant simplement vers f sur Ω_2.

Alors, pour tout $n \in \mathbf{N}$, la fonction $f_n \circ \Phi \in C_c(\Omega_1)$ puisque $f_n \in C_c(\Omega_2)$; de plus

$$f_0 \leq \ldots \leq f_n \leq f_{n+1} \leq \ldots \text{ sur } \Omega_2$$

implique évidemment que

$$f_0 \circ \Phi|\det J\Phi| \leq \ldots \leq f_n \circ \Phi|\det J\Phi| \leq f_{n+1} \circ \Phi|\det J\Phi| \leq \ldots \text{ sur } \Omega_1\,.$$

Ensuite, en appliquant la formule du changement de variables à la fonction continue f_n, on trouve que, pour tout $n \geq 0$

$$\int_{\Omega_1} f_n(\Phi(x))|\det J\Phi(x)|dx = \int_{\Omega_2} f_n(y)dy \leq \int_{\Omega_2} f(x)dx < +\infty\,.$$

Par conséquent la suite $(f_n \circ \Phi|\det \Phi|)_{n\geq 0}$ est fondamentale sur Ω_1, et converge simplement sur Ω_1 vers $f \circ \Phi|\det J\Phi|$ qui appartient donc à $\mathcal{L}^+(\Omega_1)$.

Enfin, par définition de l'intégrale sur $\mathcal{L}^+$, on trouve que

$$\begin{aligned}\int_{\Omega_2} f(y)dy &= \lim_{n\to+\infty} \int_{\Omega_2} f_n(y)dy \\ &= \lim_{n\to+\infty} \int_{\Omega_1} f_n(\Phi(x))|\det J\Phi(x)|dx = \int_{\Omega_1} f(\Phi(x))|\det J\Phi(x)|dx\,.\end{aligned}$$

□

Une conséquence immédiate du lemme ci-dessus est le fait que les C^1 difféomor- phismes transforment les ensembles Lebesgue-négligeables en ensembles Lebesgue-négligeables.

Lemme 3.0.7. Si $\mathcal{N} \subset \Omega_2$ est négligeable, alors $\Phi^{-1}(\mathcal{N}) \subset \Omega_1$ est négligeable.

Démonstration. Si $\mathcal{N} \subset \Omega_2$ est négligeable, il existe $f \in \mathcal{L}^+(\Omega_2)$ telle que $f(y) = +\infty$ pour tout $y \in \mathcal{N}$. Donc la fonction $f \circ \Phi$ vérifie $f \circ \Phi(x) = +\infty$ pour tout $x \in \Phi^{-1}(\mathcal{N})$. Comme $\det J\Phi(x) \neq 0$ pour tout $x \in \Omega_1$, on en déduit que $f \circ \Phi(x) |\det J\Phi(x)| = +\infty$ pour tout $x \in \Omega_1$. Or la fonction $f \circ \Phi |\det J\Phi| \in \mathcal{L}^+(\Omega_1)$ d'après le Lemme 3.0.6 page précédente : donc $\Phi^{-1}(\mathcal{N})$ est négligeable. □

Donnons pour terminer la preuve du théorème de changement de variables.

Démonstration du théorème. Soit $f \in \mathcal{L}^1(\Omega_2)$: il existe donc $\mathcal{N} \subset \Omega_2$ négligeable et $g, h \in \mathcal{L}^+(\Omega_2)$ tels que $f = g - h$ sur $\Omega_2 \setminus \mathcal{N}$. D'après le Lemme 3.0.6 page précédente, les fonctions $g \circ \Phi |\det J\Phi|$ et $h \circ \Phi |\det J\Phi|$ appartiennent à $\mathcal{L}^+(\Omega_1)$, et

$$f \circ \Phi |\det J\Phi| = g \circ \Phi |\det J\Phi| - h \circ \Phi |\det J\Phi| \text{ sur } \Omega_1 \setminus \Phi^{-1}(\mathcal{N})$$

c'est-à-dire p.p. sur Ω_1, puisque $\Phi^{-1}(\mathcal{N})$ est négligeable d'après le Lemme 3.0.7 page précédente. Par conséquent, $f \circ \Phi |\det J\Phi| \in \mathcal{L}^1(\Omega_1)$.

Enfin, en utilisant à nouveau le Lemme 3.0.6 page précédente,

$$\begin{aligned}\int_{\Omega_1} f \circ \Phi(x) |\det J\Phi(x)| dx &= \int_{\Omega_1} g \circ \Phi(x) |\det J\Phi(x)| dx \\ &\quad - \int_{\Omega_1} h \circ \Phi(x) |\det J\Phi(x)| dx \\ &= \int_{\Omega_2} g(y) dy - \int_{\Omega_2} h(y) dy = \int_{\Omega_2} f(y) dy \,.\end{aligned}$$

□

Terminons cette présentation de la formule du changement de variables dans l'intégrale de Lebesgue avec quelques exemples que l'on rencontre fréquemment.

Exemple 1 : changements de variables affines

Soit [1] $A \in GL_n(\mathbf{R})$ et $b \in \mathbf{R}^N$; on pose $\Phi(x) = Ax + b$. Alors Φ est de classe C^1 ; de plus Φ est évidemment une application bijective puisque A est une matrice inversible. On a $J\Phi(x) = A$ pour $x \in \mathbf{R}^N$, et $|\det J\Phi(x)| = |\det(A)| \neq 0$ puisque $A \in GL_n(\mathbf{R})$, de sorte que Φ est bien un C^1-difféomorphisme (dont on peut d'ailleurs calculer l'inverse : on trouve $\Phi^{-1}(y) = A^{-1}(y - b)$.)

D'après le théorème du changement de variables, si $f \in \mathcal{L}^1(\mathbf{R}^N)$, la fonction $x \mapsto f(Ax + b)$ appartient à $\mathcal{L}^1(\mathbf{R}^N)$ et on a

$$\int_{\mathbf{R}^N} f(Ax + b) dx = \frac{1}{|\det(A)|} \int_{\mathbf{R}^N} f(y) dy \,.$$

Cas des isométries affines : si $AA^T = A^T A = I$, alors Φ est une isométrie affine, et

$$\int_{\mathbf{R}^N} f(Ax + b) dx = \int_{\mathbf{R}^N} f(y) dy \,.$$

1. On note $GL_n(\mathbf{R})$ l'ensemble des matrices carrées inversibles à n lignes et colonnes et à coefficients réels.

Cas des homothéties : si $A = \lambda I$ avec $\lambda \neq 1$, alors Φ est une homothétie (de rapport λ et de centre $\frac{1}{1-\lambda}b$), et

$$\int_{\mathbf{R}^N} f(\lambda x + b)dx = \frac{1}{|\lambda|^N} \int_{\mathbf{R}^N} f(y)dy\,.$$

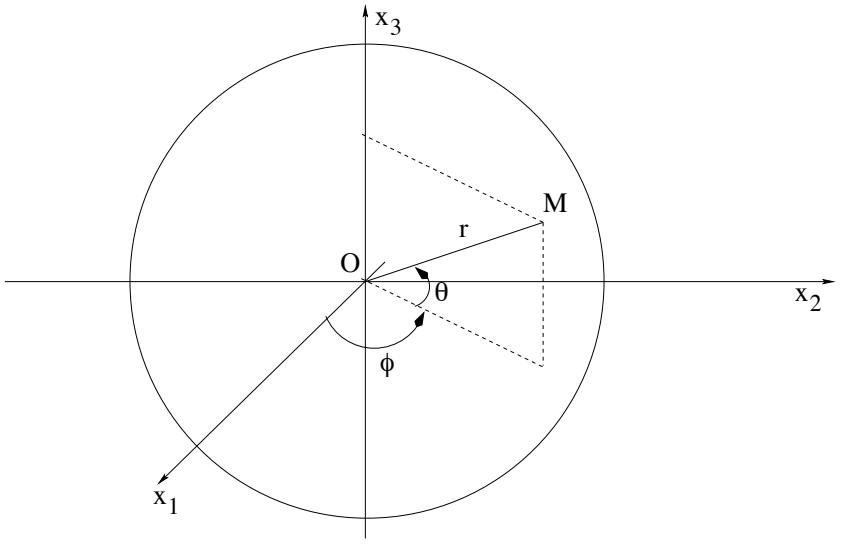


FIGURE IV.4 – Les coordonnées sphériques (r, ϕ, θ) du point $M \in \mathbf{R}^3$

Exemple 2 : coordonnées sphériques

Soit

$$\begin{array}{rcl} \Phi : \mathbf{R}_+^* \times]-\pi, \pi[\times]-\frac{\pi}{2}, \frac{\pi}{2}[& \longrightarrow & \mathbf{R}^3 \setminus \mathcal{P} \\ (r,\phi,\theta) & \mapsto & (r\cos\phi\cos\theta, r\sin\phi\cos\theta, r\sin\theta) \end{array}$$

où $\mathcal{P} := \{x \in \mathbf{R}^3 \,|\, x_1 \leq 0\}$, demi-plan affine, fermé dans $\mathbf{R}^3$.

On a $|\det J\Phi(r, \phi, \theta)| = r^2 \sin\theta$. Donc, pour $g \in \mathcal{L}^1(\mathbf{R}^3 \setminus \mathcal{P})$,

$$\int_{\mathbf{R}^3\setminus\mathcal{P}} g(x)dx = \int_0^{+\infty} \int_{-\pi}^{\pi} \int_{-\pi/2}^{\pi/2} g(\Phi(r, \phi, \theta))r^2 \sin\theta dr d\phi d\theta\,.$$

En fait, comme $\mathcal{P}$ est négligeable dans $\mathbf{R}^3$, on peut prolonger g n'importe comment sur $\mathcal{P}$ sans en changer l'intégrale. Ainsi, la formule du passage en coordonnées sphériques vaut pour tout $f \in \mathcal{L}^1(\mathbf{R}^3)$, et s'écrit

$$\int_{\mathbf{R}^3} f(x)dx = \int_0^{+\infty} \int_{-\pi}^{\pi} \int_{-\pi/2}^{\pi/2} f(\Phi(r, \phi, \theta))r^2 \sin\theta dr d\phi d\theta\,.$$

4 Ce qu'il faut retenir

Les résultats principaux de ce chapitre sont d'une utilité considérable en intégration, et il suffit pour l'essentiel d'être capable de les utiliser sans forcément en connaître la démonstration.

Inégalités intégrales

Inégalité de Jensen

Soient f, g deux fonctions mesurables définies p.p. sur Ω ouvert de $\mathbf{R}^N$ et à valeurs réelles, et soit $\Phi : \mathbf{R} \to \mathbf{R}$ convexe. Supposons que

$$g \geq 0 \text{ p.p. sur } \Omega \quad \text{ et } \quad \int_\Omega g(x)dx = 1\,,$$

tandis que

$$fg \text{ et } \Phi(f)g \in \mathcal{L}^1(\Omega)\,.$$

Alors

$$\Phi\left(\int_\Omega f(x)g(x)dx\right) \leq \int_\Omega \Phi(f(x))g(x)dx\,.$$

Inégalité de Hölder

Soient $\Omega \subset \mathbf{R}^N$ ouvert, $f, g : \Omega \to \mathbf{R}$ mesurables, et soient $p, p' > 1$ t.q. $\frac{1}{p} + \frac{1}{p'} = 1$. Alors

$$\int_\Omega |f(x)g(x)|dx \leq \left(\int_\Omega |f(x)|^p dx\right)^{1/p} \left(\int_\Omega |g(x)|^{p'} dx\right)^{1/p'}$$

Inégalité de Minkowski

Soient $\Omega \subset \mathbf{R}^N$ ouvert, $f, g : \Omega \to \mathbf{R}$ mesurables, et $p \in [1, +\infty[$. Alors

$$\left(\int_\Omega |f(x)+g(x)|^p dx\right)^{1/p} \leq \left(\int_\Omega |f(x)|^p dx\right)^{1/p} + \left(\int_\Omega |g(x)|^p dx\right)^{1/p}$$

Principaux théorèmes du calcul intégral

Théorème de Fubini.

Soient Ω_1 ouvert de $\mathbf{R}^{N_1}$ et Ω_2 ouvert de $\mathbf{R}^{N_2}$. Soit $f \in \mathcal{L}^1(\Omega_1 \times \Omega_2; \mathbf{C})$. Alors

a) p.p. en $x_2 \in \Omega_2$, la fonction $x_1 \mapsto f(x_1, x_2)$ appartient à $\mathcal{L}^1(\Omega_1; \mathbf{C})$;

b) la fonction $x_2 \mapsto \displaystyle\int_{\Omega_1} f(x_1, x_2)dx_1$ définie p.p. sur Ω_1 appartient à $\mathcal{L}^1(\Omega_2; \mathbf{C})$;

c) enfin, on a l'égalité

$$\iint_{\Omega_1\times\Omega_2} f(x_1,x_2)dx_1dx_2 = \int_{\Omega_1}\left(\int_{\Omega_2} f(x_1,x_2)dx_2\right)dx_1 \\ = \int_{\Omega_2}\left(\int_{\Omega_1} f(x_1,x_2)dx_1\right)dx_2\,,$$

(en échangeant les rôles des variables x_1 et x_2.)

Théorème de Tonelli.

Soit $f : \Omega_1 \times \Omega_2 \to [0,+\infty]$ mesurable. Alors

a) p.p. en $x_2 \in \Omega_2$, la fonction $\Omega_1 \ni x_1 \mapsto f(x_1,x_2) \in [0,+\infty]$ est mesurable ;

b) la fonction $x_2 \mapsto \displaystyle\int_{\Omega_1} f(x_1,x_2)dx_1 \in [0,+\infty]$ définie p.p. sur Ω_2 est mesurable ;

c) enfin, on a l'égalité dans $[0,+\infty]$

$$\iint_{\Omega_1\times\Omega_2} f(x_1,x_2)dx_1dx_2 = \int_{\Omega_1}\left(\int_{\Omega_2} f(x_1,x_2)dx_2\right)dx_1 \\ = \int_{\Omega_2}\left(\int_{\Omega_1} f(x_1,x_2)dx_1\right)dx_2\,.$$

Formule du changement de variables

Soit $f \in \mathcal{L}^1(\Omega_2)$. Alors la fonction $f\circ\Phi|\det J\Phi| \in \mathcal{L}^1(\Omega_1)$ et on a la formule

$$\int_{\Omega_2} f(y)dy = \int_{\Omega_1} f(\Phi(x))|\det J\Phi(x)|dx\,.$$

Chapitre V

Mesure des ensembles

1 Introduction

On sait que la notion usuelle d'intégrale s'interprète géométriquement en disant que, pour toute fonction $f \in C([a,b])$ telle que $f \geq 0$,

$$\int_a^b f(x)dx = \text{ aire de } \left\{ (x,y) \in \mathbf{R}^2 \;\middle|\; \begin{array}{l} a \leq x \leq b \\ 0 \leq y \leq f(x) \end{array} \right\}.$$

Historiquement, c'est d'ailleurs cette interprétation géométrique de l'intégrale qui a motivé les travaux de B. Cavalieri, G. Roberval, E. Torricelli et B. Pascal sur la question au début du XVIIème siècle [1]

On a vu que l'intégrale de Lebesgue permet d'intégrer des fonctions extrême- ment singulières — par exemple discontinues en tout point de $[a,b]$. Pour de telles fonctions, il n'est pas clair que le membre de droite de cette égalité, à savoir l'aire de la portion de plan comprise entre l'axe des abscisses et le graphe de la fonction, soit une quantité bien définie du point de vue géométrique.

Ceci suggère de retourner la question : puisque l'intégrale de Lebesgue permet d'intégrer des fonctions très générales, au lieu d'essayer de de définir l'intégra- le comme une l'aire d'une portion de plan, il vaut mieux utiliser l'intégrale qu'on a déjà construite pour définir — et éventuellement calculer — les surfaces ou les volumes de sous–ensembles de l'espace euclidien $\mathbf{R}^N$.

La première chose à étudier est d'ailleurs la classe des sous-ensembles de l'espace euclidien dont on peut ainsi définir la surface ou le volume. Cette question est loin d'être évidente, comme le montre le paradoxe ci-dessous.

Le paradoxe de Banach-Tarski (1924) [2]

1. En particulier, B. Pascal avait expliqué dans son *Traité de la roulette* comment calculer la surface du plan limitée par l'axe des abscisses et une arche de cycloïde.

2. Le lecteur intéressé par la preuve de cet énoncé pourra consulter l'article original

S. Banach, A. Tarski, "Sur la décomposition des ensembles de points en parties respectivement congruentes" ;

Soient A et B boules ouvertes non vides de $\mathbf{R}^N$ avec $N \geq 3$. Il existe alors $n \in \mathbf{N}^*$, et deux partitions finies des boules A et B :

$$A = A_1 \cup \ldots \cup A_n\,, \qquad A_i \cap A_j = \varnothing \text{ si } i \neq j\,,$$
$$B = B_1 \cup \ldots \cup B_n\,, \qquad B_i \cap B_j = \varnothing \text{ si } i \neq j\,,$$

ainsi que des isométries affines T_i de l'espace euclidien $\mathbf{R}^3$ telles que $T(A_i) = B_i$, pour $1 \leq i \leq n$.

Cet énoncé montre par exemple qu'on peut découper une orange en un nombre fini de morceaux, puis réassembler ces morceaux après déplacement pour obtenir une boule de même rayon que Jupiter.

Ceci est évidemment anti-intuitif ; en effet, on a envie de dire qu'étant données les deux partitions de A et de B comme ci-dessus, on devrait avoir d'une part

$$\mathrm{vol}(A) = \sum_{i=1}^{n} \mathrm{vol}(A_i)\,, \quad \mathrm{vol}(B) = \sum_{i=1}^{n} \mathrm{vol}(B_i)\,.$$

D'autre part, puisque les applications T_i sont des isométries affines, elles prése- rvent les angles et les longueurs, de sorte qu'elles transforment un cube en un cube de même volume. En décomposant chaque ensemble A_i en une réunion de cubes adjacents, à une erreur arbitrairement petite près, on arriverait au fait que

$$\mathrm{vol}(B_i) = \mathrm{vol}(T_i A) = \mathrm{vol}(A_i)\,, \quad 1 \leq i \leq n\,,$$

et donc, grâce aux deux égalités ci-dessus, on conclurait que

$$\mathrm{vol}(A) = \mathrm{vol}(B)\,.$$

Or le théorème de Banach-Tarski affirme l'existence des partitions finies de A et B, ainsi que des isométries affines T_i pour $i = 1, \ldots, n$ sans aucune autre hypothèse que le fait que A et B soient non vides.

Une première remarque avant d'aller plus loin : la démonstration de Banach-Tarski utilise l'axiome du choix (dont nous avons déjà parlé au chapitre III page 27.) Autant dire qu'elle est fondamentalement non constructive, et qu'il est donc impossible d'en dessiner les morceaux.

D'autre part, le calcul de volume ci-dessus n'utilise que les deux arguments suivants :

a) le volume d'une réunion finie d'ensembles disjoints est la somme de leurs volumes, et

b) le volume d'un ensemble est invariant par les isométries affines.

Il n'est évidemment pas question de remettre en cause ces deux assertions. Nous verrons dans la suite de ce chapitre que la bonne notion de volume satisfait à une propriété plus forte

Fundamenta Mathematicae **6** (1924), 244-277.

L'article est disponible sur le web, à l'adresse `http ://matwbn.icm.edu.pl/ksiazki/fm/fm6/fm6127.pdf`. Bien qu'écrit en excellent français, cet article utilise certaines notations vieillies (comme par exemple $A \times B$ pour désigner ce que l'on note aujourd'hui $A \cap B$...) qui peuvent en rendre la lecture malaisée. On pourra consulter également une présentation de ce résultat dans le langage mathématique d'aujourd'hui à l'adresse `http ://www.umpa.ens-lyon.fr/JME/Vol1Num1/artAReissman/artAReissman.pdf`.

que a) ; d'autre part, il n'est pas envisageable de définir une notion de volume qui ne serait pas invariante par les déplacements de l'espace euclidien ou par les symétries orthogonales.

La seule explication possible de ce paradoxe est donc qu'il est impossible de définir le volume des ensembles A_i et B_i pour $i = 1, \dots, n$. Par conséquent, nous allons donc commencer notre étude par la définition des ensembles dont on peut calculer le volume.

2 Ensembles mesurables

Commençons par rappeler la notation pour les fonctions indicatrices : étant donnée une partie A de $\mathbf{R}^N$, la fonction indicatrice de A dans $\mathbf{R}^N$ est

$$\mathbf{1}_A : x \mapsto \begin{cases} 1 \text{ si } x \in A\,, \\ 0 \text{ si } x \notin A\,. \end{cases}$$

Puisqu'une partie de $\mathbf{R}^N$ est complètement déterminée par sa fonction indicatrice, il est naturel d'introduire la définition suivante. (En réalité, le bien-fondé de cette terminologie n'apparaîtra que dans la section suivante.)

Définition 2.0.8 (Ensembles mesurables). Un sous-ensemble $A \subset \mathbf{R}^N$ est mesurable si et seulement si sa fonction indicatrice $\mathbf{1}_A$ est mesurable sur $\mathbf{R}^N$.

Voici les premières propriétés des ensembles mesurables.

Proposition 2.0.9. Soit $N \in \mathbf{N}^*$. Alors

a) $\varnothing$ et $\mathbf{R}^N$ sont mesurables ;

b) si A et $B \subset \mathbf{R}^N$ sont mesurables, $A \setminus B$ est mesurable ;

c) si $(A_n)_{n\in\mathbf{N}}$ est une famille **dénombrable** de parties mesurables de $\mathbf{R}^N$, alors leur réunion et leur intersection

$$\bigcup_{n\in\mathbf{N}} A_n \quad \text{et} \quad \bigcap_{n\in\mathbf{N}} A_n \text{ sont mesurables.}$$

(L'énoncé c) contient évidemment le cas des réunions finies, en prenant tous les A_n vides, à l'exception d'un nombre fini d'entre eux.)

Etant donné un ensemble Ω quelconque, on appelle *tribu* sur Ω un sous-ensemble $\mathcal{F}$ de l'ensemble $\mathcal{P}(\Omega)$ des parties de Ω vérifiant

a) $\varnothing \in \mathcal{F}$;

b) $A \in \mathcal{F} \Rightarrow \Omega \setminus A \in \mathcal{F}$;

c) si $(A_n)_{n\in\mathbf{N}}$ est une suite d'éléments de $\mathcal{F}$, alors

$$\bigcup_{n\in\mathbf{N}} A_n \in \mathcal{F}\,.$$

Ainsi, la proposition ci-dessus exprime le fait que les ensembles mesurables de $\mathbf{R}^N$ forment une tribu de $\mathbf{R}^N$. Nous reviendrons plus loin sur cette propriété.

Démonstration. Le a) est trivial car $\mathbf{1}_\varnothing = 0$ et $\mathbf{1}_{\mathbf{R}^N} = 1$ sont mesurables.

Pour le b), observons que

$$\mathbf{1}_{A\setminus B} = \mathbf{1}_{A\cap(\mathbf{R}^N\setminus B)} = \mathbf{1}_A\mathbf{1}_{\mathbf{R}^N\setminus B} = \mathbf{1}_A(1-\mathbf{1}_B)$$

est mesurable si $\mathbf{1}_A$ et $\mathbf{1}_B$ sont mesurables.

Démontrons le point c). Posons

$$\mathcal{A}_n = \bigcap_{k=0}^{n} A_k\,.$$

Alors

$$\mathbf{1}_{\mathcal{A}_n} = \prod_{k=0}^{n} \mathbf{1}_{A_k}\,;$$

d'autre part, comme $0 \leq \mathbf{1}_{A_n} \leq 1$, on a

$$\mathbf{1}_{\mathcal{A}_0} \geq \mathbf{1}_{\mathcal{A}_1} \geq \ldots \geq \mathbf{1}_{\mathcal{A}_n} \geq \mathbf{1}_{\mathcal{A}_{n+1}} \geq \ldots \geq 0\,.$$

Par conséquent, pour tout $x \in \mathbf{R}^N$, la suite décroissante de réels positifs $(\mathbf{1}_{\mathcal{A}_n}(x))_{n\geq 0}$ converge vers sa borne inférieure lorsque $n \to +\infty$, et

$$\inf_{n\geq 0} \mathbf{1}_{\mathcal{A}_n}(x) = \mathbf{1}_{\bigcap_{k\geq 0} A_k}(x)\,.$$

En effet

$$x \in \bigcap_{k\geq 0} A_k \Rightarrow x \in \mathcal{A}_n \text{ pour tout } n \in \mathbf{N}$$

de sorte que

$$\inf_{n\geq 0} \mathbf{1}_{\mathcal{A}_n}(x) \geq \mathbf{1}_{\bigcap_{k\geq 0} A_k}(x)\,.$$

Inversement, supposons que

$$\inf_{n\geq 0} \mathbf{1}_{\mathcal{A}_n}(x) = 1;$$

comme la suite des $\mathbf{1}_{\mathcal{A}_n}(x)$ est décroissante, cela signifie que $x \in \mathcal{A}_n \subset A_n$ pour tout $n \geq 0$, de sorte que

$$x \in \bigcap_{k\geq 0} A_k\,.$$

Par conséquent

$$\inf_{n\geq 0} \mathbf{1}_{\mathcal{A}_n}(x) \leq \mathbf{1}_{\bigcap_{k\geq 0} A_k}(x)\,.$$

Au total, on a donc montré que, lorsque $n \to +\infty$

$$\mathbf{1}_{\mathcal{A}_n} \to \mathbf{1}_{\bigcap_{k\geq 0} A_k} = \prod_{k\geq 0} \mathbf{1}_{A_k}$$

simplement sur $\mathbf{R}^N$, ce qui montre que

$$\mathbf{1}_{\bigcap_{k\geq 0} A_k}$$

est mesurable comme limite simple d'une suite de fonctions mesurables.

L'énoncé sur les réunions dénombrable se déduit de l'énoncé sur les intersections dénombrables par passage au complémentaire, c'est-à-dire en écrivant que

$$\bigcup_{n\geq 0} A_n = \mathbf{R}^N \setminus \bigcap_{n\geq 0} (\mathbf{R}^N \setminus A_n)$$

et en appliquant le b). □

Exemples d'ensembles mesurables :

a) tout intervalle de $\mathbf{R}$ est une partie mesurable de $\mathbf{R}$;

b) tout pavé de $\mathbf{R}^N$ est une partie mesurable de $\mathbf{R}^N$;

c) tout ouvert de $\mathbf{R}^N$, tout fermé de $\mathbf{R}^N$ est une partie mesurable de $\mathbf{R}^N$;

d) toute partie négligeable de $\mathbf{R}^N$ est mesurable dans $\mathbf{R}^N$.

Démonstration. Démontrons le c). Il suffit de vérifier que tout fermé de $\mathbf{R}^N$ est mesurable, le cas des ouverts en découle alors par passage au complémentaire (cf. le b) de la proposition ci-dessus.)

D'autre part, tout fermé de $\mathbf{R}^N$ est une réunion dénombrable de compacts : en effet, étant donné $F \subset \mathbf{R}^N$, on a

$$F = \bigcup_{n\geq 0} (F \cap \overline{B(0,n)}),$$

et $F \cap \overline{B(0,n)}$ est compact pour tout $n \geq 0$ comme fermé borné de $\mathbf{R}^N$. Il suffit donc de vérifier que tout compact de $\mathbf{R}^N$ est mesurable.

Soient donc $K \subset \mathbf{R}^N$ compact, et, pour tout $n \in \mathbf{N}^*$, la fonction f_n définie par $f_n(x) := (1 - n\,\mathrm{dist}(x,K))_+$. Rappelons que la fonction $\mathbf{R}^N \ni x \mapsto \mathrm{dist}(x,K)$ est continue car, pour $x, y \in \mathbf{R}^N$,

$$|\,\mathrm{dist}(x,K) - \mathrm{dist}(y,K)| \leq \mathrm{dist}(x,y)$$

d'après l'inégalité triangulaire. Donc la fonction f_n est continue pour tout $n \geq 1$ comme composée de fonctions continues. De plus, pour tout $n \geq 1$, si $\mathrm{dist}(x,K) > 1$, on a $f_n(x) = 0$. Par conséquent, $f_n \in C_c(\mathbf{R}^N)$ pour tout $n \geq 1$. Enfin, comme K est fermé, $\mathrm{dist}(x,K) = 0$ si et seulement si $x \in K$. Par conséquent, la suite $(f_n)_{n\geq 1}$ converge simplement vers la fonction indicatrice $\mathbf{1}_K$ sur $\mathbf{R}^N$. On en déduit que $\mathbf{1}_K$ est une fonction mesurable comme limite simple d'une suite de fonctions de $C_c(\mathbf{R}^N)$, et donc que K est une partie mesurable de $\mathbf{R}^N$.

Pour le d), observons que, si $\mathcal{N} \subset \mathbf{R}^N$ est négligeable, alors $\mathbf{1}_{\mathcal{N}} = 0$ p.p. sur $\mathbf{R}^N$ est évidemment mesurable comme limite p.p. de la suite dont tous les termes sont la fonction nulle.

Les points a) et b) sont maintenant évidents. Si P est un pavé fermé de $\mathbf{R}^N$, il est mesurable d'après le c). Si P' est un pavé quelconque, il est de la forme

$$P' = \overline{P'} \setminus B$$

où B est une partie de la frontière de P'. Or la frontière de P' est Lebesgue-négligeable (comme réunion finie de parties d'hyperplans affines de $\mathbf{R}^N$.) Donc B est également négligeable, et donc mesurable d'après le d). On en déduit que P' est mesurable grâce à la propriété b) de la proposition précédente. Ceci établit le b) ; le point a) est le cas particulier du b) pour $N = 1$. □

Remarque : La Proposition 2.0.9 page 93 montre qu'on ne peut pas *construire* de partie non mesurable de $\mathbf{R}^N$ par un algorithme, c'est-à-dire une suite dénombrable d'opérations élémentaires (réunion, intersection, passage au complémentaire...) à partir de sous-ensembles élémentaires (pavés, par exemple).

Mais il *existe* des parties non mesurables de $\mathbf{R}^N$ (par exemple les ensembles A_i et B_i des partitions du paradoxe de Banach-Tarski). Toutefois, pour fabriquer une partie non mesurable de $\mathbf{R}^N$, il est nécessaire de recourir à l'axiome du choix — l'existence de tels ensembles ne peut pas découler d'un procédé constructif.

Terminons cette présentation rapide de la notion d'ensemble mesurable avec une généralisation de la notion d'intégration sur un ouvert de $\mathbf{R}^N$.

Pour $\Omega \subset \mathbf{R}^N$ ouvert et $f \in \mathcal{L}^1(\Omega; \mathbf{C})$, le prolongement de f

$$F\colon x \mapsto \begin{cases} f(x) & \text{si } x \in \Omega \\ 0 & \text{si } x \notin \Omega \end{cases}$$

appartient à $\mathcal{L}^1(\mathbf{R}^N; \mathbf{C})$ et

$$\int_\Omega f(x)dx = \int_{\mathbf{R}^N} F(x)dx\,.$$

(Démontrer cette égalité ne présente aucune difficulté : en effet, elle est vraie pour $f \in C_c(\Omega)$, puisqu'alors $F \in C_c(\mathbf{R}^N)$, et qu'il s'agit d'une égalité classique entre intégrales au sens usuel. Puis on montre cette égalité pour des fonctions de $\mathcal{L}^+(\Omega)$, puisqu'elle vaut pour tous les termes des suites fondamentales. Enfin, on la généralise à $\mathcal{L}^1(\Omega)$ par linéarité.)

Cette identité suggère alors la définition suivante.

Définition 2.0.10 (Intégration sur les ensembles mesurables). Soient $A \subset \mathbf{R}^N$ mesurable et f fonction à valeurs complexes définie p.p. sur A. On dira que $f \in \mathcal{L}^1(A; \mathbf{C})$ si son prolongement

$$F\colon x \mapsto \begin{cases} f(x) & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases}$$

appartient à $\mathcal{L}^1(\mathbf{R}^N; \mathbf{C})$, auquel cas

$$\int_A f(x)dx := \int_{\mathbf{R}^N} F(x)dx$$

La relation de Chasles bien connue sur les intégrales usuelles, à savoir

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx$$

pour tout $f \in C(\mathbf{R})$, se généralise alors de la façon suivante : pour $A, B \subset \mathbf{R}^N$ mesurables, et pour $f \in \mathcal{L}^1(A \cup B; \mathbf{C})$, on a

$$\int_{A\cup B} f(x)dx = \int_A f(x)dx + \int_B f(x)dx\,.$$

Jusqu'ici, nous avons défini les ensembles mesurables à partir des fonctions mesurables. Inversement, on peut caractériser les fonctions mesurables en vérifiant que certains ensembles sont mesurables.

Proposition 2.0.11 (Fonctions/ensembles mesurables). Soit Ω ouvert de $\mathbf{R}^N$ et soit $f : \Omega \to \mathbf{R}$ définie p.p. ; alors

$$\begin{aligned} f \text{ mesurable } &\Leftrightarrow f^{-1}(]\lambda, +\infty[) \text{ mesurable pour tout } \lambda \in \mathbf{R} \\ &\Leftrightarrow f^{-1}([\lambda, +\infty[) \text{ mesurable pour tout } \lambda \in \mathbf{R} \\ &\Leftrightarrow f^{-1}(I) \text{ mesurable pour tout intervalle } I \subset \mathbf{R} \end{aligned}$$

Démonstration. Il suffit évidemment de vérifier la première équivalence. En effet

$$f^{-1}([\lambda, +\infty[) = \bigcap_{n\geq 1} f^{-1}(]\lambda - \tfrac{1}{n}, +\infty[)$$

et

$$f^{-1}(]\lambda, +\infty[) = \bigcup_{n\geq 1} f^{-1}([\lambda + \tfrac{1}{n}, +\infty[)\,,$$

ce qui montre que les deux premières conditions sont équivalentes.

Tout intervalle I de $\mathbf{R}$ est de la forme $I = A \setminus B$, où A et B sont de la forme $[\lambda, +\infty[$ ou $]\lambda, +\infty[$. La troisième condition est donc évidemment équivalente aux deux premières puisque $f^{-1}(A \setminus B) = f^{-1}(A) \setminus f^{-1}(B)$.

Vérifions par exemple la première équivalence.

Si f est mesurable, définissons, pour tout $n \geq 1$ et p.p. en $x \in \mathbf{R}^N$,

$$u_n(x) = n\left(\min\left(f(x), \lambda + \frac{1}{n}\right) - \min(f(x), \lambda)\right)\,;$$

(voir sur la figure 2 page suivante la signification géométrique de cette formule.)

Pour tout $n \geq 1$, la fonction u_n est mesurable comme composée de la fonction continue sur $\mathbf{R}$

$$y \mapsto n\left(\min\left(y, \lambda + \frac{1}{n}\right) - \min(y, \lambda)\right)$$

et de la fonction mesurable f. Comme $u_n(x) \to \mathbf{1}_{]\lambda,+\infty[}(f(x))$ lorsque $n \to +\infty$ pour tout $x \in \Omega$ tel que $f(x)$ soit défini, c'est-à-dire p.p. en $x \in \Omega$, on en déduit que la fonction $\mathbf{1}_{]\lambda,+\infty[}(f(x))$ est mesurable comme limite p.p. d'une suite de fonctions mesurables. Or

$$\mathbf{1}_{]\lambda,+\infty[} \circ f = \mathbf{1}_{f^{-1}(]\lambda,+\infty[)}\,;$$

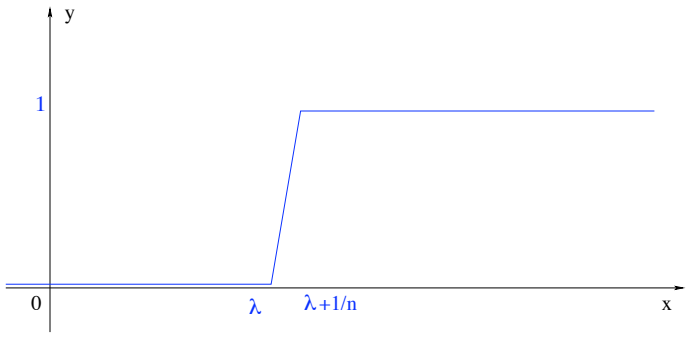


FIGURE V.1 – Graphe de la fonction $x \mapsto n\left(\min\left(x, \lambda + \frac{1}{n}\right) - \min(x, \lambda)\right)$

comme cette fonction est mesurable, on en déduit que l'ensemble $f^{-1}(]\lambda, +\infty[)$ est mesurable.

Réciproquement supposons que $f^{-1}(]\lambda, +\infty[)$ est mesurable pour tout $\lambda \in \mathbf{R}$. Montrons que f est une fonction mesurable.

Supposons pour commencer que f est à valeurs dans $\mathbf{R}_+$. Pour $n \geq 1$, on pose

$$f_n(x) = \sum_{k=0}^{4^n} \frac{k}{2^n} \mathbf{1}_{f^{-1}([2^{-n}k, 2^{-n}(k+1)[)}(x)\,.$$

Evidemment, pour tout $n \geq 1$, la fonction f_n est mesurable comme combinaison linéaires des fonctions mesurables

$$\mathbf{1}_{f^{-1}([2^{-n}k, 2^{-n}(k+1)[)} = \mathbf{1}_{f^{-1}([2^{-n}k, +\infty[)} - \mathbf{1}_{f^{-1}([2^{-n}(k+1), +\infty[)}\,.$$

Pour tout $n \geq 1$ assez grand pour que $f(x) < 2^n$, on a $|f(x) - f_n(x)| \leq \frac{1}{2^n}$. Par conséquent, $f_n \to f$ simplement sur Ω, de sorte que f est mesurable comme limite simple de la suite de fonctions mesurables $(f_n)_{n \geq 1}$.

Dans le cas général, on pose $F_n = (f + n)_+$ et $f_n = F_n - n = \max(f, -n)$; évidemment

$$F_n^{-1}(]\lambda, +\infty[) = \begin{cases} f^{-1}(]\lambda - n, +\infty[) & \text{si } \lambda \geq n \\ \Omega & \text{si } \lambda < n \end{cases}$$

de sorte que F_n vérifie la même condition que f ; de plus F_n est positive par construction, donc F_n est mesurable pour tout $n \geq 0$, ainsi que f_n. Comme $f(x) > -\infty$ p.p. sur Ω par hypothèse, $f_n \to f$ p.p. sur Ω, de sorte que f est mesurable comme limite p.p. d'une suite de fonctions mesurables. □

3 Mesure de Lebesgue

Définition 3.0.12 (Mesure de Lebesgue). Pour tout $A \subset \mathbf{R}^N$ mesurable, la mesure de Lebesgue de A est

$$|A| = \int_{\mathbf{R}^N} \mathbf{1}_A(x) dx\,.$$

Cette définition éclaire la section précédente. En effet, la partie A de $\mathbf{R}^N$ est mesurable si et seulement si la fonction $\mathbf{1}_A$ est mesurable. Comme $\mathbf{1}_A$ est une fonction positive, son intégrale est bien définie comme élément de $[0, +\infty]$.

Proposition 3.0.13. La mesure de Lebesgue vérifie les propriétés suivantes :

a) la mesure de Lebesgue d'un pavé de $\mathbf{R}^N$ est son volume :

$$|(a_1, b_1) \times \ldots \times (a_N, b_N)| = \prod_{k=1}^{N} (b_k - a_k)_+$$

— en particulier $|\varnothing| = 0$;

b) pour tout $A \subset \mathbf{R}^N$ mesurable et pour toute isométrie affine T de $\mathbf{R}^N$,

$$|T(A)| = |A| ;$$

c) soit (A_n) suite **dénombrable** de parties mesurables de $\mathbf{R}^N$; alors

$$A_k \cap A_l = \varnothing \text{ pour tous } k, l \quad \Rightarrow \left| \bigcup_n A_n \right| = \sum_n |A_n| ;$$

d) $A \subset B \subset \mathbf{R}^N$ mesurables $\Rightarrow |A| \leq |B|$;

e) soit (B_n) suite **dénombrable** de parties mesurables de $\mathbf{R}^N$

$$\left| \bigcup_n B_n \right| \leq \sum_n |B_n| .$$

Comme nous l'avons annoncé dans l'introduction de ce chapitre, la mesure de Lebesgue — qui généralise la notion usuelle de volume — est invariante par les isométries affines (propriété b)), et vérifie la propriété c) dite d'*additivité dénombrable*, qui généralise le fait intuitivement évident que le volume de la réunion de deux ensembles disjoints est la somme de leurs voumes. Cette généralisation au cas de réunions dénombrables est particulièrement subtile — voir en particulier la Proposition 3.2.9 page 42 et sa démonstration.

Démonstration. Dans le cas $N = 1$, la fonction $\mathbf{1}_{(a,b)}$ est continue par morceaux, le calcul intégral usuel nous apprend que

$$\int_{\mathbf{R}} \mathbf{1}_{(a,b)}(x) dx = \begin{cases} \displaystyle\int_a^b dx = b - a & \text{si } b \geq a , \\ 0 & \text{si } a > b . \end{cases}$$

Le cas $N > 1$ découle du cas $N =$ en appliquant le théorème de Tonelli. Ceci démontre le a).

Le b) découle de la formule du changement de variables, et du fait que le déterminant jacobien d'une isométrie affine vaut ± 1.

Quant au point c), on observe pour commencer que

$$A_k \cap A_l = \varnothing \text{ pour tous } k, l \quad \Rightarrow \mathbf{1}_{\bigcup_n A_n} = \sum_n \mathbf{1}_{A_n} .$$

Noter que, comme les ensembles A_n sont disjoints, pour tout $x \in \mathbf{R}^N$, il n'y a qu'un seul terme non nul dans la série

$$\sum_n \mathbf{1}_{A_n}(x)$$

et d'ailleurs x appartient au plus à l'un des ensembles A_n : la formule ci-dessus est donc triviale.

La formule du c) découle alors de la linéarité de l'intégrale dans le cas d'une réunion finie, et du théorème de convergence monotone dans le cas d'une réunion infinie.

Le point d) est une conséquence du point c) : en effet, $B = A \cup (B \setminus A)$ est une réunion disjointe, de sorte que

$$|B| = |A| + |B \setminus A| \geq |A| .$$

Pour ce qui est du point e), supposons que la suite (B_n) est infinie et indexée par $\mathbf{N}$ — si ce n'est pas le cas, on complètera le nombre fini d'ensembles B_n en rajoutant l'ensemble vide une infinité dénombrable de fois.

Définissons $A_0 = B_0$, $A_1 = B_1 \setminus B_0$, et plus généralement

$$A_n = B_n \setminus (B_0 \cup \ldots \cup B_{n-1}) , \quad n \geq 1 .$$

Alors

$$\bigcup_{n\geq 0} A_n = \bigcup_{n\geq 0} B_n ;$$

de plus

$$A_n \cap A_m \subset A_n \cap B_m = \varnothing$$

si $n > m$. Donc, d'après le c) et le d)

$$\left|\bigcup_{n\geq 0} B_n\right| = \left|\bigcup_{n\geq 0} A_n\right| = \sum_{n\geq 0} |A_n| \leq \sum_{n\geq 0} |B_n|$$

puisque $A_n \subset B_n$ pour tout n. □

La mesure de Lebesgue permet donc de calculer le volume d'ensembles tellement généraux (les parties mesurables de $\mathbf{R}^N$) qu'il n'est pas possible d'en *construire* un dont on ne sache pas définir le volume — c'est-à-dire un ensemble non mesurable ; comme on l'a dit plus haut, l'existence de tels ensembles nécessite l'emploi de l'axiome du choix, de sorte qu'il est impossible de les *construire* par une suite dénombrable d'opérations élémentaires.

Exemple : mesure de Lebesgue des ouverts de R

Un ouvert de $\mathbf{R}$ peut être géométriquement assez compliqué — on peut penser par exemple au complémentaire de l'ensemble de Cantor (rappelons que l'ensemble de Cantor est compact). Néanmoins la complexité d'un ouvert est relativement limitée. Rappelons en effet que

Tout ouvert $\Omega \subset \mathbf{R}$ est réunion d'une suite dénombrable d'intervalles ouverts non vides $]a_n, b_n[$ deux à deux disjoints — les composantes connexes des rationnels de Ω.

On déduit de cet énoncé le calcul de la mesure de Lebesgue d'un ouvert Ω quelconque de $\mathbf{R}$: on écrit la décomposition ci-dessus de Ω, à savoir

$$\Omega = \bigcup_n]a_n, b_n[\quad \text{où } a_n < b_n \text{ et } \;]a_m, b_m[\cap]a_n, b_n[= \varnothing \text{ si } m \neq n\,,$$

et par additivité dénombrable de la mesure de Lebesgue

$$|\Omega| = \sum_n (b_n - a_n)\,.$$

4 Fonction de répartition complémentaire

On a vu dans la section 2 page 93 que l'image réciproque d'un intervalle par une fonction mesurable définie p.p. et à valeurs dans $\mathbf{R}$ est une partie mesurable de Ω. Nous allons expliquer dans cette section comment la mesure des ensembles ainsi obtenus est reliée à l'intégrale de f sur Ω.

Définition 4.0.14 (Fonction de répartition complémentaire). Soient Ω un ouvert de $\mathbf{R}^N$, et $f : \Omega \to \mathbf{R}_+$ une fonction mesurable. La fonction de répartition complémentaire de f est la fonction $\rho_f : \mathbf{R}_+ \to [0, +\infty]$ définie par

$$\rho_f(\lambda) := |f^{-1}(]\lambda, +\infty[)| = |\{x \in \Omega \,|\, f(x) > \lambda\}|\,.$$

Evidemment, la fonction ρ_f est définie pour tout $\lambda \geq 0$, puisque, f étant mesurable, l'ensemble $f^{-1}(]\lambda, +\infty[)$ est mesurable pour tout $\lambda \geq 0$.

Par construction, la fonction ρ_f est évidemment décroissante au sens large, puisque

$$\begin{aligned} \lambda' \geq \lambda &\Rightarrow f^{-1}(]\lambda', +\infty[) \subset f^{-1}(]\lambda, +\infty[) \\ &\Rightarrow \rho_f(\lambda') = |f^{-1}(]\lambda', +\infty[)| \leq |f^{-1}(]\lambda, +\infty[)| = \rho_f(\lambda)\,. \end{aligned}$$

Voici une première relation entre la fonction de répartition complémentaire d'une fonction et son intégrale de Lebesgue.

Proposition 4.0.15 (Inégalité de Markov). Soient Ω ouvert de $\mathbf{R}^N$ et $f \in \mathcal{L}^1(\Omega)$. Alors, pour tout $\lambda > 0$, on a

$$\rho_{|f|}(\lambda) = |\{x \in \Omega \,|\, |f(x)| \geq \lambda\}| \leq \frac{1}{\lambda} \int_\Omega |f(x)| dx\,.$$

Démonstration. En effet

$$\begin{aligned} \int_\Omega |f(x)| dx &\geq \int_\Omega |f(x)| \mathbf{1}_{[\lambda, +\infty]}(f(x)) dx \\ &\geq \lambda \int_\Omega \mathbf{1}_{[\lambda, +\infty]}(f(x)) dx = \lambda |f^{-1}([\lambda, +\infty])| = \lambda \rho_{|f|}(\lambda)\,, \end{aligned}$$

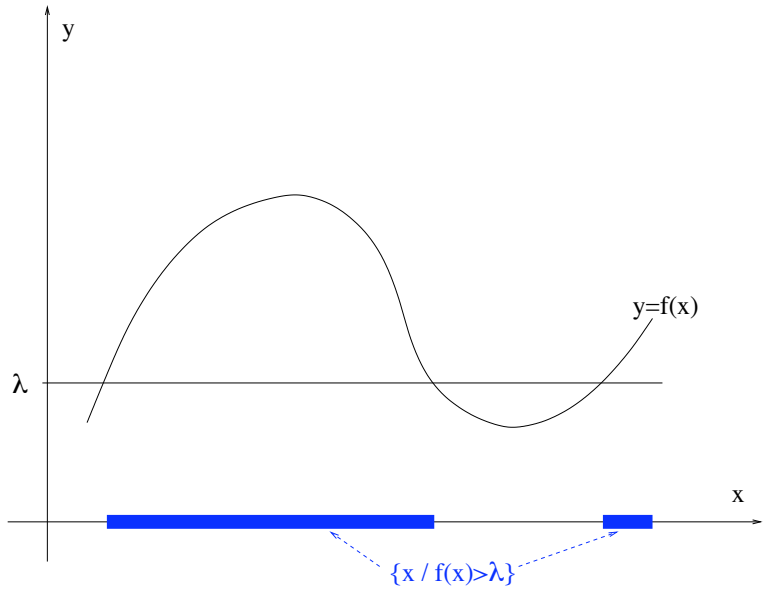


FIGURE V.2 – Fonction de répartition complémentaire

puisque

$$\mathbf{1}_{[\lambda,+\infty]}(f(x)) = \mathbf{1}_{f^{-1}([\lambda,+\infty])}(x)$$

pour tout $x \in \Omega$ tel que $f(x)$ soit défini, c'est-à-dire p.p. en $x \in \Omega$. □

L'inégalité de Markov a de nombreuses conséquences importantes. En particulier, elle montre que les ensembles négligeables sont les ensembles de mesure nulle.

Théorème 4.0.16 (Mesure et ensembles négligeables). Soient $A \subset \mathbf{R}^N$ et f fonction mesurable positive définie sur Ω ouvert de $\mathbf{R}^N$. Alors

1) A est Lebesgue-négligeable $\Leftrightarrow |A| = 0$;

2) si $\displaystyle\int_\Omega f(x)dx = 0$, alors $|\{x \in \Omega \,|\, f(x) \neq 0\}| = 0$;

3) si $\displaystyle\int_\Omega f(x)dx < +\infty$, alors $|\{x \in \Omega \,|\, f(x) = +\infty\}| = 0$.

Démonstration. Etablissons le 3) : pour tout $n \in \mathbf{N}^*$, l'inégalité de Markov entraîne que

$$|\{x \in \Omega \,|\, |f(x)| = +\infty\}| \leq \rho_{|f|}(n) \leq \frac{1}{n}\int_\Omega |f(x)|dx\,;$$

on conclut en faisant tendre n vers $+\infty$.

Pour le 2),

$$\{x \in \Omega \,|\, f(x) \neq 0\} = \bigcup_{n\geq 1} \{x \in \Omega \,|\, f(x) > \tfrac{1}{n}\}\,;$$

or d'après l'inégalité de Markov

$$|\{x \in \Omega \,|\, f(x) > \tfrac{1}{n}\}| \leq n \int_\Omega f(x)dx = 0\,,$$

de sorte que

$$|\{x \in \Omega \,|\, f(x) \neq 0\}| \leq \sum_{n\geq 1} \{x \in \Omega \,|\, f(x) > \tfrac{1}{n}\} = 0$$

d'après la Proposition 3.0.13 page 99 e).

Terminons avec le 1).

Si A est Lebesgue-négligeable, il existe $f \in \mathcal{L}^+(\mathbf{R}^N)$ telle que $f(x) = +\infty$ pour tout $x \in A$. Or

$$A \subset f^{-1}(\{+\infty\}) \subset \{x \in \mathbf{R}^N \,|\, |f(x)| = +\infty\}$$

et comme $f \in \mathcal{L}^+(\mathbf{R}^N) \subset \mathcal{L}^1(\mathbf{R}^N)$, d'après le 3), on a

$$|A| \leq |\{x \in \mathbf{R}^N \,|\, |f(x)| = +\infty\}| = 0\,.$$

Inversement, si $|A| = 0$, la fonction positive $\mathbf{1}_A$ est sommable et vérifie

$$\int_{\mathbf{R}^N} \mathbf{1}_A(x)dx = |A| = 0\,.$$

D'après le Théorème 3.4.1 page 48, on conclut que $\mathbf{1}_A = 0$ p.p. sur $\mathbf{R}^N$, ce qui équivaut précisément à dire que A est Lebesgue-négligeable. □

La propriété la plus importante de la fonction de répartition complémentaire est le principe de Cavalieri, qui fut l'un des pionniers du calcul intégral au XVIIème siècle.

Théorème 4.0.17 (Principe de Cavalieri). Soient $f : \Omega \to [0, +\infty]$ mesurable, et ρ_f sa fonction de répartition complémentaire. Alors, pour tout $\Phi : \mathbf{R}_+ \to \mathbf{R}_+$ croissante de classe C^1 vérifiant $\Phi(0) = 0$, on a

$$\int_\Omega \Phi(f(x))dx = \int_{\mathbf{R}_+} \Phi'(\lambda)\rho_f(\lambda)d\lambda$$

En particulier, pour $\Phi(\lambda) = \lambda$

$$\int_\Omega f(x)dx = \int_{\mathbf{R}_+} \rho_f(\lambda)d\lambda\,.$$

Avant de donner la preuve du théorème, une remarque s'impose. Les deux membres de la formule de Cavalieri, dans le cas particulier où Φ est l'identité, n'ont pas du tout le même statut. En effet, le membre de gauche est une intégrale de Lebesgue, puisqu'on ne suppose rien d'autre sur la fonction f que sa mesurabilité. En revanche, comme la fonction ρ_f est monotone — décroissante au sens large, en l'espèce — on peut l'intégrer par le procédé de Riemann [3].

3. Si $\phi : \mathbf{R}_+ \to \mathbf{R}_+$ est une fonction décroissante au sens large,

$$\int_{\mathbf{R}_+} \phi(t)dt = \sup_{0<t_0<t_1<\dots} \sum_{k\geq 0} (t_{k+1} - t_k)\phi(t_{k+1})\,,$$

le membre de droite de cette égalité étant évidemment une somme de Riemann.

Autrement dit, toute intégrale de Lebesgue peut se ramener à une intégrale de Riemann... une fois qu'on a défini la mesure de Lebesgue. Nous reviendrons plus loin sur cette remarque, qui est lourde de conséquences.

Le cas particulier ci-dessus du principe de Cavalieri suggère une autre remarque importante.

L'intégrale usuelle, dont la définition est due à Riemann (et même, dans le cas des fonctions continues, à Cauchy) consiste à définir, pour $f \in C([a,b])$,

$$\int_a^b f(x)dx = \lim_{n\to+\infty} \frac{b-a}{n} \sum_{k=1}^{n} f\left(a + k\frac{b-a}{n}\right) .$$

L'idée clé faisant fonctionner cette méthode est que la continuité de f permet de garantir que l'oscillation de f sur chacun des intervalles $[a + k\frac{b-a}{n}, a + (k+1)\frac{b-a}{n}]$ est uniformément petite [4] lorsque $n \to +\infty$, de sorte qu'on peut assimiler f à une constante sur chacun de ces intervalles lorsque leur longueur tend vers 0.

Autrement dit, le coeur de cet argument est que la subdivision de l'intervalle $[a,b]$ en petits intervalles de longueur $\frac{b-a}{n}$, jointe à la continuité de f, permet de regrouper ensemble des valeurs arbitrairement proches de la fonction f.

Si l'on veut définir une notion d'intégrale permettant d'intégrer des fonctions très singulières, par exemples discontinues en tout point, comme la fonction indicatrice des rationnels, l'idée naturelle proposée par Lebesgue consiste à subdiviser non pas l'intervalle sur lequel f est définie, mais l'ensemble des valeurs prises par la fonction f. En effet, en procédant ainsi, on peut regrouper sans aucune les valeurs voisines de la fonction f sans aucune hypothèse de continuité — il ne reste plus qu'à évaluer la longueur de l'ensemble des points x dont l'image par f reste à une distance au plus $\epsilon > 0$ d'une valeur donnée prise par la fonction f.

Le principe de Cavalieri, dans sa formulation ci-dessus, explique justement comment l'intégrale d'une fonction sommable quelconque s'exprime comme une intégrale usuelle de Riemann à partir du moment où on intègre sur l'ensemble des valeurs prises par la fonction f.

Donnons à présent la démonstration du principe de Cavalieri.

Démonstration. L'idée clé consiste à écrire

$$\Phi(f(x)) = \int_0^{f(x)} \Phi'(\lambda)d\lambda \quad \text{p.p. en } x \in \Omega .$$

Comme f est mesurable et Φ de classe C^1 croissante sur $\mathbf{R}_+$, la fonction

$$\Omega \times \mathbf{R}_+ \ni (x,\lambda) \mapsto \mathbf{1}_{[0,f(x)[}(\lambda)\Phi'(\lambda) \in \mathbf{R}_+$$

4. En effet, f étant continue sur le compact $[a,b]$, elle y est uniforément continue, d'après le Corollaire 4.3.6 page 18.

est mesurable. D'après le théorème de Tonelli,

$$\begin{aligned}\int_\Omega \Phi(f(x))dx &= \int_\Omega \left(\int_0^{f(x)} \Phi'(\lambda)d\lambda \right) dx \\ &= \int_\Omega \left(\int_{\mathbf{R}_+} \mathbf{1}_{[0,f(x)[}(\lambda)\Phi'(\lambda)d\lambda \right) dx \\ &= \iint_{\Omega\times\mathbf{R}_+} \mathbf{1}_{[0,f(x)[}(\lambda)\Phi'(\lambda)dxd\lambda\,.\end{aligned}$$

Puis, pour tout $\lambda \geq 0$,

$$\mathbf{1}_{[0,f(x)[}(\lambda) = \mathbf{1}_{]\lambda,+\infty[}(f(x)) \text{ pour tout } \lambda \in \mathbf{R}_+ \text{ p.p. en } x \in \Omega\,,$$

de sorte que, toujours d'après le théorème de Tonelli

$$\begin{aligned}\int_\Omega \Phi(f(x))dx &= \iint_{\Omega\times\mathbf{R}_+} \mathbf{1}_{[0,f(x)[}(\lambda)\Phi'(\lambda)dxd\lambda \\ &= \iint_{\Omega\times\mathbf{R}_+} \mathbf{1}_{]\lambda,+\infty[}(f(x))\Phi'(\lambda)dxd\lambda \\ &= \int_{\mathbf{R}_+} \Phi'(\lambda) \left(\int_\Omega \mathbf{1}_{]\lambda,+\infty[}(f(x))dx \right) d\lambda \\ &= \int_{\mathbf{R}_+} \Phi'(\lambda)|\{x \in \Omega \text{ tq. } f(x) > \lambda\}|d\lambda \\ &= \int_{\mathbf{R}_+} \Phi'(\lambda)\rho_f(\lambda)d\lambda\,.\end{aligned}$$

□

5 Généralisations

5.1 Mesures de Radon

Au lieu de partir de l'intégrale usuelle sur $C_c(\Omega)$, on peut se donner une forme linéaire $\mu : C_c(\Omega) \to \mathbf{R}$ positive [5], au sens où

$$f \in C_c(\Omega) \text{ et } f \geq 0 \Rightarrow \langle \mu, f \rangle \geq 0\,.$$

Le procédé de **prolongement par monotonie** qui nous a permis de constru- ire l'intégrale de Lebesgue s'applique de même et aboutit à l'espace vectoriel $\mathcal{L}^1(\Omega;\mu)$ des fonctions μ-sommables, et à un prolongement à $\mathcal{L}^1(\Omega;\mu)$ de la forme linéaire μ, noté

$$\mathcal{L}^1(\Omega;\mu) \ni f \mapsto \int_\Omega f(x)d\mu(x) \in \mathbf{R}$$

5. On utilisera la notation $\langle \mu, \phi \rangle$ pour désigner l'image de la fonction $\phi \in C_c(\Omega)$ par la forme linéaire μ.

par analogie avec l'intégrale de Lebesgue — c'est-à-dire que $C_c(\Omega) \subset \mathcal{L}^1(\Omega;\mu)$ et

$$f \in C_c(\Omega) \Rightarrow \int_\Omega f(x)d\mu(x) = \langle \mu; f \rangle \,.$$

De même que dans le cas de l'intégrale de Lebesgue, le prolongement de la forme linéaire ainsi obtenu vérifie les théorèmes de convergence monotone et dominée.

On prendra bien garde au fait que, dans ce cas, les ensembles μ-négligeables ne sont plus forcément les mêmes que pour l'intégrale usuelle. Ils doivent être définis au moyen des suites fondamentales pour la forme linéaire μ, suivant la Définition 3.2.1 page 37, et surtout pas en utilisant la caractérisation géométrique du Théorème 5.0.7 page 54, qui ne vaut que pour les ensembles Lebesgue-négligeables.

On définit de même une notion de convergence d'une suite de fonctions μ-p.p. (convergence simple sur le complémentaire d'un ensemble μ-négligeable), ainsi qu'une notion de fonction μ-mesurable (limite μ-p.p. d'une suite de fonctions de $C_c(\Omega)$.)

On définit ensuite la μ-mesure d'un ensemble μ-mesurable (c'est-à-dire d'un ensemble $A \subset \Omega$ dont la fonction indicatrice $\mathbf{1}_A$ est une fonction μ-mesurable) par la formule

$$\mu(A) := \int_\Omega \mathbf{1}_A(x)d\mu(x) \,.$$

On vérifie alors que

$$A \text{ est } \mu\text{-négligeable } \Leftrightarrow \mu(A) = 0 \,.$$

La clef de toute cette construction est la Proposition 4.6.1 page 20, qui est une variante du lemme de Dini.

La terminologie est la suivante : on appelle "mesure de Radon sur Ω" une forme linéaire positive μ définie sur $C_c(\Omega)$ comme ci-dessus ; son prolongement à la classe des fonctions μ-sommables est appelé "intégrale par rapport à la mesure μ". Cette terminologie est malheureuse, car elle conduit à une confusion entre mesure et intégrale. En effet, dans le formalisme de l'"intégration abstraite", utilisé systématiquement en calcul des probabilités, on appelle "mesure sur Ω" une application qui associe une valeur réelle à certains sous-ensembles de Ω, et "intégrale par rapport à une mesure sur Ω" une application qui associe une valeur réelle à certaines fonctions définies sur Ω.

Ainsi, l'intégrale usuelle définie sur $C_c(\Omega)$ est un exemple de mesure de Radon ; mais ce que l'on appelle "mesure de Lebesgue sur Ω" est l'application définie sur l'ensemble des parties mesurables de Ω par $A \mapsto |A| \in [0, +\infty]$, que l'on distingue de l'"intégrale de Lebesgue", qui est le prolongement de l'intégrale usuelle à la classe $\mathcal{L}^1(\Omega)$ des fonctions sommables sur Ω.

Malheureusement le terme de "mesure de Radon" confondant les notions d'intégrale (des fonctions) et de mesure (des ensembles) est consacré par l'usage dans le monde entier depuis plus d'un demi-siècle, de sorte qu'il n'est pas envisable d'en chercher une meilleure.

Voici quelques exemples de mesures de Radon :

Exemples

1) Masse de Dirac : à $x_0 \in \Omega$, on associe la forme linéaire δ_{x_0} sur $C_c(\Omega)$ définie par

$$\langle \delta_{x_0}, f \rangle = f(x_0)\,.$$

On remarquera que toute partie de Ω est δ_{x_0}-mesurable, et que, pour tout $A \subset \Omega$,

$$\delta_{x_0}(A) = \begin{cases} 1 \text{ si } x_0 \in A\,, \\ 0 \text{ si } x_0 \notin A\,. \end{cases}$$

Autrement dit, les ensembles δ_{x_0}-négligeables sont tous les sous-ensembles de Ω ne contenant pas x_0. Ces ensembles ne sont évidemment pas Lebesgue-négligea- bles en général. Cet exemple montre donc clairement que la notion d'ensemble négligeable dépend très fortement de la mesure de Radon considérée.

2) Mesure de comptage : à tout ensemble $A \subset \Omega$ discret et dénombrable, on associe la forme linéaire sur $C_c(\Omega)$ définie par

$$N_A = \sum_{a \in A} \delta_a\,,$$

c'est-à-dire

$$\langle N_A, f \rangle = \sum_{a \in A} f(a) =: \int_\Omega f(x) dN_A(x)\,.$$

La théorie des séries numériques coïncide donc avec la théorie de l'intégration par rapport à la mesure de comptage sur $\mathbf{R}$, associée au sous-ensemble $\mathbf{N}$ et discret dénombrable dans $\mathbf{R}$

Le point de vue des mesures de Radon est un cas particulier de la théorie des distributions, qui sera étudiée dans la seconde partie du cours MAT431 de deuxième année.

5.2 Mesure et intégration "abstraite"

Au lieu de partir d'une mesure de Radon, c'est-à-dire d'une forme linéaire positive μ sur $C_c(\Omega)$ où Ω est un ouvert de $\mathbf{R}^N$, donnons-nous maintenant un ensemble Ω **quelconque** — pas forcément un ouvert de $\mathbf{R}^N$, ni même muni d'une topologie.

On a vu dans la Proposition 2.0.9 page 93 que les parties Lebesgue-mesurables d'un ouvert Ω forment une tribu sur Ω — c'est-à-dire $\mathcal{F} \subset \mathcal{P}(\Omega)$ contenant $\varnothing$, et stable par passage au complémentaire et par réunion dénombrable.

Inversement, donnons nous une tribu $\mathcal{F}$ quelconque de parties de Ω. On appellera "sous-ensembles $\mathcal{F}$-mesurables" les parties de Ω appartenant à $\mathcal{F}$.

On appellera "mesure positive sur Ω" une application $m : \mathcal{F} \to [0, +\infty]$ dénombrablement additive, c'est-à-dire vérifiant la propriété suivante

pour toute suite dénombrable (A_n) d'éléments de $\mathcal{F}$

$$A_m \cap A_n = \varnothing \text{ pour tous } m \neq n \Rightarrow \mu\left(\bigcup_n A_n\right) = \sum_n \mu(A_n)$$

Une fonction $f : \Omega \to [0, +\infty]$ sera dite $\mathcal{F}$-mesurable si et seulement si, pour tout intervalle $I \subset [0, +\infty]$, son image réciproque par f est $\mathcal{F}$-mesurable, soit $f^{-1}(I) \in \mathcal{F}$.

A partir de là, on **définit** l'intégrale de f sur Ω par rapport à la mesure m grâce à la formule de Cavalieri :

$$\int_\Omega f(x)dm(x) := \int_{\mathbf{R}_+} m(\{x \in \Omega \,|\, f(x) > \lambda\})d\lambda \,.$$

Comme nous l'avons déjà évoqué plus haut, tout l'intérêt de cette définition est que membre de droite est une intégrale au sens habituel — c'est-à-dire au sens de Riemann — car c'est l'**intégrale d'une fonction décroissante** au sens large sur $\mathbf{R}_+$. (Voir la note en bas de page suivant l'énoncé du principe de Cavalieri.)

Cette présentation originale de l'intégration abstraite est due à E. De Giorgi (qui l'enseignait à la Scuola Normale Superiore de Pise au début des années 1980) : voir l'appendice du livre [2] pour plus de détails sur cette présentation. On trouvera dans les références [16] et [19] une présentation plus classique mais très détaillée de l'intégration abstraite, ainsi que dans l'appendice de [5].

C'est ce point de vue de l'intégration abstraite qui a permis à A.N. Kolmogorov de proposer, en 1933, une théorie axiomatique parfaitement rigoureuse des probabilités, sujet qui avait passionné les plus grands mathématiciens (Pascal, Laplace, Gauss, pour ne citer qu'eux), mais qui était resté quelque peu à l'écart des grandes avancées en mathématiques jusque là. A partir de là, les probabilités sont devenues l'un des domaines les plus dynamiques des mathématiques et jouent un rôle important aussi bien dans leurs branches les plus fondamentales (théorie des matrices aléatoires, analyse en dimension infinie) que dans leur interaction avec la physique (mécanique statistique, théorie des champs) ainsi que dans les applications (contrôle des systèmes, économie, finance...)

6 Ce qu'il faut retenir

Un sous-ensemble $A \subset \mathbf{R}^N$ est mesurable si et seulement si sa fonction indicatrice $\mathbf{1}_A$ est mesurable.

L'ensemble des parties mesurables de $\mathbf{R}^N$ forme une tribu sur $\mathbf{R}^N$ — c'est à dire un ensemble de parties de $\mathbf{R}^N$ contenant $\varnothing$ et stable par passage au complémentaire et par réunion **dénombrable**. Tout ouvert et tout fermé de $\mathbf{R}^N$ est mesurable.

Pour $\Omega \subset \mathbf{R}^N$ ouvert non vide et $A \subset \Omega$ mesurable, on définit

$$\int_A f(x)dx = \int_\Omega \mathbf{1}_A(x)f(x)dx \,.$$

Si A est une partie mesurable de $\mathbf{R}^N$, on définit sa mesure de Lebesgue par la formule

$$|A| := \int_{\mathbf{R}^N} \mathbf{1}_A(x)dx \,.$$

La mesure de Lebesgue est une fonction d'ensemble croissante :

$$A \subset B \subset \mathbf{R}^N \text{ mesurables } \Rightarrow |A| \leq |B| \,.$$

Les parties Lebesgue négligeables de $\mathbf{R}^N$ sont les parties mesurables de mesure de Lebesgue nulle.

Fonction de répartition complémentaire.

Soient Ω un ouvert de $\mathbf{R}^N$, et $f : \Omega \to \mathbf{R}_+$ une fonction mesurable. La fonction de répartition complémentaire de f est la fonction $\rho_f : \mathbf{R}_+ \to [0, +\infty]$ définie par

$$\rho_f(\lambda) := |f^{-1}(]\lambda, +\infty[)| = |\{x \in \Omega \,|\, f(x) > \lambda\}| \,.$$

La fonction de répartition complémentaire d'une fonction sommable vérifie les propriétés suivantes.

Inégalité de Markov

Soient Ω ouvert de $\mathbf{R}^N$ et $f \in \mathcal{L}^1(\Omega)$. Alors, pour tout $\lambda > 0$, on a

$$\rho_{|f|}(\lambda) = |\{x \in \Omega \,|\, |f(x)| \geq \lambda\}| \leq \frac{1}{\lambda} \int_\Omega |f(x)| dx \,.$$

Principe de Cavalieri Soient $f : \Omega \to [0, +\infty]$ mesurable, et ρ_f sa fonction de répartition complémentaire. Alors, pour tout $\Phi : \mathbf{R}_+ \to \mathbf{R}_+$ croissante de classe C^1 vérifiant $\Phi(0) = 0$, on a

$$\int_\Omega \Phi(f(x)) dx = \int_{\mathbf{R}_+} \Phi'(\lambda) \rho_f(\lambda) d\lambda$$

Chapitre VI

Espaces de Lebesgue

1 Motivations

A partir de $C_c(\Omega)$, on a construit dans le chapitre III page 27 l'intégrale de Lebesgue et l'espace vectoriel $\mathcal{L}^1(\Omega)$ des fonctions sommables sur Ω ouvert de $\mathbf{R}^N$. On a vu également dans ce même chapitre que l'une des motivations principales pour la construction de l'intégrale de Lebesgue et de l'espace $\mathcal{L}^1(\Omega)$ des fonctions sommables était l'obtention de théorèmes d'une grande souplesse sur l'intégration terme à terme des séries de fonctions — c'est-à-dire sur l'interversion intégrale/limite pour les suites de fonctions simplement convergentes.

Pour aller plus loin dans l'étude de la convergence des suites de fonctions et de leurs intégrales, il est naturel d'essayer d'appliquer à cet espace $\mathcal{L}^1(\Omega)$ des fonctions sommables les méthodes de la topologie, c'est-à-dire de considérer les fonctions sommables comme les points d'un espace vectoriel normé. C'est précisément ce programme que l'on va mettre en œuvre dans le présent chapitre.

Sur $C_c(\Omega)$, la convergence en moyenne des suites ou séries de fonctions est définie par la norme

$$N_1(f) = \int_\Omega |f(x)|dx\,.$$

Malheureusement, l'espace vectoriel $C_c(\Omega)$ muni de la norme N_1 n'est pas complet.

Exemple 1.0.1. Pour tout $n \geq 1$, on définit $f_n : [-1,1] \to \mathbf{R}$ comme la fonction impaire telle que

$$f_n(x) = \min(1, (n+1)x)\,, \quad x \in [0,1]\,.$$

Pour tout $n \geq 1$, la fonction f_n est continue sur $[-1,1]$; d'autre part, pour $n \geq m > 1$, on a

$$\begin{aligned} N_1(f_n - f_m) &= \int_{-1}^{1} |f_n(x) - f_m(x)|dx \\ &= 2\int_0^1 |f_n(x) - f_m(x)|dx = \frac{1}{m+1} - \frac{1}{n+1}\,. \end{aligned}$$

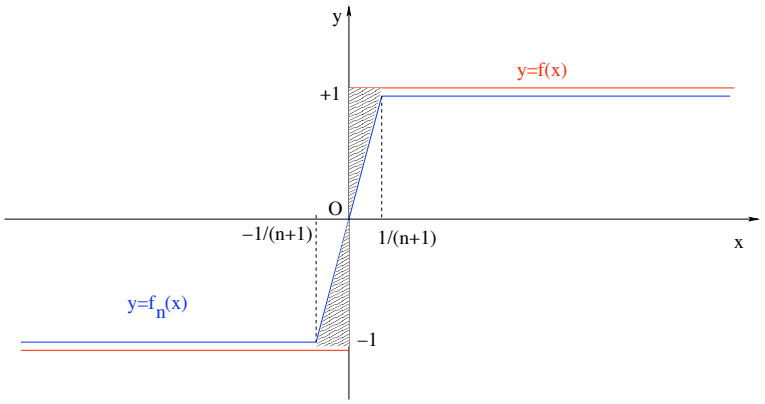


FIGURE VI.1 – Graphes de f_n et de f ; la distance de f_n à f est l'aire de la surface hachurée.

La suite $(f_n)_{n\geq 1}$ est donc de Cauchy, puisque

$$m, n \geq R \Rightarrow N_1(f_n - f_m) \leq \frac{1}{R+1} \to 0 \text{ lorsque } R \to +\infty\,.$$

Mais la suite $(f_n)_{n\geq 1}$ converge en moyenne vers la fonction

$$f : [-1,1] \ni x \mapsto f(x) := \begin{cases} +1 \text{ si } x > 0\,, \\ -1 \text{ si } x < 0\,. \end{cases}$$

Or f n'est pas continue sur $[-1,1]$, puisque

$$\int_{-1}^{1} |f_n(x) - f(x)| dx = 2\int_0^{1/(n+1)} |1-(n+1)x| = \frac{1}{n+1}\,.$$

Il est à peu près inconcevable d'envisager l'analyse dans un espace vectoriel normé qui ne serait pas complet — c'est d'ailleurs une motivation supplémen- taire pour développer une théorie de l'intégration plus puissante que celle de l'intégrale usuelle de Riemann sur $C_c(\Omega)$.

Le but de ce chapitre est donc de construire un espace complet à partir de $\mathcal{L}^1(\Omega)$ et de la norme N_1 — ce qui permet notamment d'appliquer aux séries de fonctions le critère "convergence normale $\Rightarrow$ convergence" (analogue au critère "convergence absolue $\Rightarrow$ convergence" pour les séries numériques à termes réels ou complexes.)

Ce que l'on va réaliser dans ce chapitre est donc l'analogue pour la théorie de l'intégration de la construction du corps $\mathbf{R}$ des réels à partir du corps $\mathbf{Q}$ des rationnels.

2 L'espace de Lebesgue L^1

Soit $\Omega \subset \mathbf{R}^N$ ouvert ; l'idée naturelle est de prolonger la norme N_1 de la convergence en moyenne à l'espace $\mathcal{L}^1(\Omega)$ ou $\mathcal{L}^1(\Omega;\mathbf{C})$.

Considérons donc l'application

$$N_1 : \mathcal{L}^1(\Omega;\mathbf{C}) \ni f \mapsto N_1(f) := \int_\Omega |f(x)|dx \in \mathbf{R}_+ .$$

Cette application est une *semi-norme* — c'est-à-dire qu'elle vérifie

(a) $N_1(\lambda f) = |\lambda| N_1(f)$ pour tout $f \in \mathcal{L}^1(\Omega;\mathbf{C})$ et tout $\lambda \in \mathbf{C}$, et

(b) $N_1(f+g) \leq N_1(f) + N_1(g)$ pour tout $f, g \in \mathcal{L}^1(\Omega;\mathbf{C})$.

Mais N_1 *n'est pas une norme sur* $\mathcal{L}^1(\Omega;\mathbf{C})$: en effet, d'après le Théorème 3.4.1 page 48

$$f \in \mathcal{L}^1(\Omega;\mathbf{C}) \text{ et } N_1(f) = 0 \Leftrightarrow f(x) = 0 \text{ p.p. en } x \in \Omega .$$

Donc $N_1(f) = 0$ *n'implique pas* que $f = 0$ *partout* sur Ω.

Cette difficulté n'existe pas sur $C_c(\Omega)$, car $N_1(f) = 0$ implique que $f(x) = 0$ p.p. en $x \in \Omega$, de sorte que, si f est continue sur Ω, on a $f = 0$ *partout* sur Ω. Ce résultat est classique, mais en voici une preuve basée sur le Théorème 3.4.1 page 48 et la Proposition 5.0.14 page 59.

En effet, dire que $f(x) = 0$ p.p. en $x \in \Omega$, c'est dire que $\{x \in \Omega \,|\, f(x) \neq 0\}$ est négligeable. Le complémentaire de cet ensemble, soit $\{x \in \Omega \,|\, f(x) = 0\}$ est donc dense dans Ω — voir Proposition 5.0.14 page 59 — et comme cet ensemble est fermé dans Ω si f est continue (comme image réciproque du fermé $\{0\}$ par la fonction continue f), on conclut que cet ensemble est Ω tout entier, de sorte que $f = 0$ sur Ω.

L'idée clé pour pallier cette difficulté consiste à identifier deux fonctions sommables égales p.p. sur Ω. C'est tout à fait naturel dans le contexte de l'intégration, puisque deux fonctions sommables sur Ω égales p.p. sur Ω ont même intégrale.

Voici comment on effectue cette identification dans le langage mathématique : pour $f, g \in \mathcal{L}^1(\Omega;\mathbf{C})$, on notera

$$f \simeq g \quad \text{ si et seulement si } f(x) - g(x) = 0 \text{ p.p. en } x \in \Omega .$$

Observons que les opérations élémentaires sur $\mathcal{L}^1(\Omega)$ sont compatibles avec l'identification des fonctions égales p.p. sur Ω :

1) pour tous $\alpha, \beta \in \mathbf{C}$ et $f_1, f_2, g_1, g_2 \in \mathcal{L}^1(\Omega;\mathbf{C})$

$$f_1 \simeq g_1 \text{ et } f_2 \simeq g_2 \Rightarrow \alpha f_1 + \beta f_2 \simeq \alpha g_1 + \beta g_2 .$$

De plus l'intégrale de Lebesgue et la semi-norme N_1 sont également compatibles avec l'identification des fonctions égales p.p. :

2) pour tout $f, g \in \mathcal{L}^1(\Omega;\mathbf{C})$

$$f \simeq g \Rightarrow \int_\Omega f(x)dx = \int_\Omega g(x)dx \text{ et } N_1(f) = N_1(g) .$$

Définition 2.0.2 (Espace de Lebesgue L^1). Soit Ω ouvert de $\mathbf{R}^N$. Pour tout $f \in \mathcal{L}^1(\Omega)$ (resp. $\mathcal{L}^1(\Omega;\mathbf{C})$), on note

$$[f] := \{\phi \in \mathcal{L}^1(\Omega;\mathbf{C}) \,|\, \phi \simeq f\}$$

la classe d'équivalence de f. L'espace de Lebesgue $L^1(\Omega)$ (ou $L^1(\Omega;\mathbf{C})$) est défini comme suit :

$$L^1(\Omega) \text{ resp. } L^1(\Omega;\mathbf{C}) := \{[f] \,|\, f \in \mathcal{L}^1(\Omega) \text{ resp. } \mathcal{L}^1(\Omega;\mathbf{C})\}\,.$$

L'espace de Lebesgue $L^1(\Omega)$ (resp. $L^1(\Omega;\mathbf{C})$) est un $\mathbf{R}$-espace vectoriel (resp. un $\mathbf{C}$-espace vectoriel) pour l'addition des (classes d'équivalences de) fonctions et la multiplication par les scalaires réels (resp. complexes), définies comme suit :

pour tout $\alpha \in \mathbf{C}$ et $f, g \in \mathcal{L}^1(\Omega;\mathbf{C})$, on pose

$$[f] + [g] := [f+g] \quad \text{ et } \alpha[f] := [\alpha f]\,.$$

On vérifie grâce à la propriété 1) ci-dessus que ces définitions sont indépendantes des éléments choisis dans les classes $[f]$ et $[g]$.

L'intégrale de Lebesgue définit une forme $\mathbf{R}$-linéaire (resp. $\mathbf{C}$-linéaire) sur $L^1(\Omega)$ (resp. $L^1(\Omega;\mathbf{C})$), comme suit :

$$L^1(\Omega) \ni [f] \mapsto \int_\Omega [f](x)dx := \int_\Omega f(x)dx \in \mathbf{R} \quad (\text{ resp. } \mathbf{C}\,.)$$

La propriété 2) ci-dessus montre que cette définition est indépendante de l'élément choisi dans la classe $[f]$.

La semi-norme N_1 sur $\mathcal{L}^1(\Omega)$ (resp. $\mathcal{L}^1(\Omega;\mathbf{C})$) définit une semi-norme sur $L^1(\Omega)$ (resp. $L^1(\Omega;\mathbf{C})$), par la formule

$$N_1([f]) := N_1(f) = \int_\Omega |f(x)|dx\,.$$

Il s'agit bien d'une semi-norme, car pour tout $[f], [g] \in L^1(\Omega)$ (resp. $L^1(\Omega;\mathbf{C})$) et $\alpha \in \mathbf{R}$ (resp. $\mathbf{C}$), on a

$$N_1(\alpha[f]) = |\alpha| N_1([f]) \quad \text{ et } N_1([f]+[g]) \le N_1([f]) + N_1([g])\,.$$

Mais l'avantage de cette construction est que la semi-norme N_1 ainsi définie sur $L^1(\Omega)$ (resp. $L^1(\Omega);\mathbf{C})$) est une norme, car

$$N_1([f]) = 0 \Leftrightarrow \int_\Omega |f(x)|dx = 0 \Leftrightarrow f = 0 \text{ p.p. sur } \Omega \Leftrightarrow f \simeq 0 \Leftrightarrow [f] = 0\,.$$

Notation : il est d'usage de noter la norme N_1 ainsi définie sur l'espace de Lebesgue $L^1(\Omega)$ (resp. $L^1(\Omega;\mathbf{C})$)

$$N_1(f) =: \|f\|_1 \text{ ou } \|g\|_{L^1} \text{ ou } \|f\|_{L^1(\Omega)} \text{ ou encore } \|f\|_{L^1(\Omega;\mathbf{C})}$$

selon le degré d'ambigüité du contexte.

Convention de langage : Dorénavant, on identifiera toujours l'élément $[f]$ de $L^1(\Omega)$ avec n'importe laquelle des fonctions de $[f]$ — par exemple f elle-même. (Aucun mathématicien ne pense à un élément de $L^1(\Omega)$ ou de $L^1(\Omega;\mathbf{C})$ comme à une classe d'équivalence de fonctions.)

On pense donc à un élément de $L^1(\Omega)$ ou de $L^1(\Omega;\mathbf{C})$ comme à une fonction f définie p.p. sur Ω et mesurable.

Mais quand on écrit que

$$f = g \text{ dans } L^1(\Omega)\,,$$

cela signifie que

$$f(x) = g(x) \text{ p.p. sur } \Omega\,.$$

Autrement dit, la définition rigoureuse de $L^1(\Omega)$ ou de $L^1(\Omega;\mathbf{C})$ consiste à considérer comme éléments des classes d'équivalences de fonctions, en conservant la signification habituelle de l'égalité — en se souvenant qu'ici, il s'agit précisément d'une égalité entre classes d'équivalence.

Intuitivement, on préfère penser aux éléments de $L^1(\Omega)$ ou de $L^1(\Omega;\mathbf{C})$ comme à des fonctions définies p.p. sur Ω, mais remplacer la notion habituelle d'égalité entre fonctions définies partout sur Ω par l'égalité p.p. sur Ω entre fonctions définies seulement p.p. sur Ω. Ces deux façons de voir l'espace de Lebesgue sont évidemment rigoureusement équivalentes. Une des raisons pour laquelle la définition usuelle de l'espace de Lebesgue fait intervenir l'égalité au sens usuel entre des classes d'équivalences de fonctions est que le concept d'égalité est à ce point fondamental en mathématiques que les mathématiciens ont probablement quelque scrupule à en modifier le sens, fût-ce par commodité, à l'intérieur de telle ou telle théorie particulière.

Ce que l'on a gagné avec cette construction est évidemment que $L^1(\Omega)$ (resp. $L^1(\Omega;\mathbf{C})$) muni de la norme $\|\cdot\|_{L^1}$ est un espace vectoriel normé dont on va pouvoir étudier les propriétés topologiques.

L'inconvénient est que, étant donné $x_0 \in \Omega$ fixé et $f \in L^1(\Omega)$, on ne pourra plus *jamais* parler de $f(x_0)$ — en effet, f est définie seulement p.p. sur Ω et $\{x_0\}$ est négligeable. Donc la valeur $f(x_0)$ peut être n'importe quoi, ou encore n'être même pas définie.

Mais ce n'est pas très grave dans le contexte de l'intégration, où l'on s'intéresse seulement à des *quantités intégrées* comme

$$\int_\Omega f(x)dx \quad \text{ ou } \quad \|f\|_{L^1} = \int_\Omega |f(x)|dx\,.$$

D'ailleurs, en physique, on pense souvent à une fonction comme à une collection de mesures d'une certaine grandeur — par exemple la pression, ou la température dans un fluide, ou encore les composantes d'un champ électromagnétique — mesures que l'on effectuerait en tout point de l'espace.

Ceci est toutefois une vue de l'esprit, car un appareil de mesure ne fournit jamais qu'une valeur moyenne locale de la quantité mesurée — comme la pression dans un fluide par exemple. C'est à dire qu'au lieu de fournir la pression $p(x)$ en tout point $x \in \mathbf{R}^3$, ce que l'on peut mesurer est plutôt une quantité du type

$$\frac{1}{|A|}\int_A p(x)dx$$

pour tout cube A de l'espace $\mathbf{R}^3$ — par exemple.

C'est d'ailleurs ce point de vue qui prévaut en mécanique quantique. Les quantités (énergie, impulsion...) relatives à un système que l'on mesure en mécani que quantique, quantités nommées les "observables", sont des valeurs moyennes d'une fonction par rapport au carré du module de la fonction d'onde dans l'espace des positions ou celui des impulsions — voir, pour plus de détails, le cours PHYS311.

3 Propriétés topologiques de L^1

3.1 Complétude de L^1

L'énorme avantage qu'offrent l'intégrale de Lebesgue et l'espace de Lebesgue L^1 sur les théories antérieures de l'intégration tient au résultat suivant.

Théorème 3.1.1 (Complétude de L^1). Soit Ω ouvert de $\mathbf{R}^N$. Alors l'espace $L^1(\Omega)$ (resp. $L^1(\Omega;\mathbf{C})$) muni de la norme $\|\cdot\|_{L^1}$ est complet.

Pour faire sentir au lecteur tout ce que l'on a gagné avec la construction de l'intégrale de Lebesgue, disons qu'avec la notion usuelle d'intégrale définie sur les fonctions continues, c'est un peu comme si on se limitait aux opérations usuelles dans le corps $\mathbf{Q}$ des rationnels. Ainsi, chaque fois que l'on serait en présence d'une série absolument convergente de rationnels, on devrait calculer sa limite pour vérifier qu'elle est rationnelle. (C'est essentiellement ce que réclament les théorèmes limites de l'intégration présentés dans les classes préparatoires.) De la sorte, on s'interdirait de considérer des suites de rationnels convergeant vers $\sqrt{2}$; de même, l'usage des nombres π ou e serait rigoureusement impossible — ce qui aurait des répercussions fort gênantes, par exemple dans la théorie des équations différentielles.

Evidemment, ces quelques remarques n'ont pas pour objet de diminuer l'importance considérable des premières théories de l'intégration, qui commencent avec les travaux de B. Cavalieri, G. Roberval, E. Torricelli, B. Pascal, puis d'I. Newton, de G. Leibnitz, et enfin d'A. Cauchy et de B. Riemann au XIXème siècle. L'avènement du calcul différentiel et intégral a été un progrès retentissant dans l'histoire des sciences, d'une portée absolument exceptionnelle, car il a transformé radicalement les moyens dont disposait l'homme pour comprendre et prédire le monde qui l'entoure.

En réalité, nous allons démontrer un résultat plus précis que le théorème ci-dessus.

Théorème 3.1.2 (Réciproque de la convergence dominée). Soient Ω ouvert de $\mathbf{R}^N$, et $(f_n)_{n\geq 0}$ suite de Cauchy de $L^1(\Omega;\mathbf{C})$. Il existe alors une suite extraite (f_{n_k}) vérifiant les hypothèses du théorème de convergence dominée. C'est-à-dire qu'il existe $f, F \in L^1(\Omega;\mathbf{C})$ t.q.

a) $f_{n_k} \to f$ p.p. sur Ω lorsque $n_k \to +\infty$, et

b) $|f_{n_k}| \leq F$ p.p. sur Ω pour tout $n_k \geq 0$.

En particulier $\|f_n - f\|_{L^1} \to 0$ lorsque $n \to +\infty$ — autrement dit, toute la suite $(f_n)_{n\geq 0}$ est convergente vers f dans $L^1(\Omega;\mathbf{C})$.

Outre le fait que ce dernier théorème entraîne la complétude de l'espace de Lebesgue $L^1(\Omega;\mathbf{C})$, dont nous avons déjà souligné l'importance, il constitue une quasi-réciproque du théorème de convergence dominée. En d'autres termes, on ne peut pas espérer de théorème de passage à la limite sous le signe somme de portée beaucoup plus générale que le théorème de convergence dominée.

Démonstration. Comme $(f_n)_{n\geq 0}$ est une suite de Cauchy dans $L^1(\Omega;\mathbf{C})$, pour tout $\epsilon > 0$, il existe $N(\epsilon) \geq 0$ t.q.

$$m, n \geq N(\epsilon) \Rightarrow \|f_m - f_n\|_{L^1(\Omega)} < \epsilon\,.$$

Choisissons $\epsilon = \frac{1}{2}$, et posons $n_1 = N(\frac{1}{2})$: alors

$$m, n \geq n_1 \Rightarrow \|f_m - f_n\|_{L^1(\Omega)} < \tfrac{1}{2}\,.$$

Puis, pour $\epsilon = \frac{1}{4}$, on pose $n_2 = \max(n_1 + 1, N(\frac{1}{4}))$, de sorte que

$$m, n \geq n_2 \Rightarrow \|f_m - f_n\|_{L^1(\Omega)} < \tfrac{1}{4}\,.$$

Supposons construits $0 \leq n_1 < n_2 < \ldots < n_k$ tels que

$$m, n \geq n_k \Rightarrow \|f_m - f_n\|_{L^1(\Omega)} < \tfrac{1}{2^k}\,.$$

Alors, pour $\epsilon = \frac{1}{2^{k+1}}$, on pose $n_{k+1} = \max(n_k+1, N(\frac{1}{2^{k+1}}))$; ainsi

$$m, n \geq n_{k+1} \Rightarrow \|f_m - f_n\|_{L^1(\Omega)} < \tfrac{1}{2^{k+1}}\,.$$

Par récurrence, on obtient ainsi une suite infinie d'entiers

$$0 \leq n_1 < \ldots < n_k < n_{k+1} < \ldots\,,$$

et par construction

$$\|f_{n_k} - f_{n_{k+1}}\|_{L^1(\Omega)} < \tfrac{1}{2^k}\,.$$

Posons, p.p. en $x \in \Omega$,

$$\Phi(x) = \sum_{k\geq 1} |f_{n_{k+1}}(x) - f_{n_k}(x)| \in [0, +\infty]\,.$$

La fonction Φ est mesurable (comme somme d'une série de terme général mesurable, d'après le Théorème 6.0.19 page 60), à valeurs dans $[0, +\infty]$; d'autre part, par convergence monotone

$$\begin{aligned}\int_\Omega \Phi(x)dx &= \sum_{k\geq 1} \int_\Omega |f_{n_{k+1}}(x) - f_{n_k}(x)|dx \\ &= \sum_{k\geq 1} \|f_{n_k} - f_{n_{k+1}}\|_{L^1(\Omega)} \leq \sum_{k\geq 1} \frac{1}{2^k} = 1\,,\end{aligned}$$

de sorte que $\Phi \in L^1(\Omega)$. D'après la Proposition 3.4.2 page 48, il existe donc $\mathcal{N} \subset \Omega$ négligeable t.q. $\Phi(x) < +\infty$ pour tout $x \in \Omega \setminus \mathcal{N}$.

Considérons alors la série

$$\phi(x) := \sum_{k\geq 1} \left(f_{n_{k+1}}(x) - f_{n_k}(x)\right) .$$

D'après ce qui précède, pour tout $x \in \Omega \setminus \mathcal{N}$, la série $\phi(x)$ est absolument convergente dans $\mathbf{C}$ (qui est complet) et donc convergente. Par conséquent

$$f_{n_k}(x) = f_{n_1}(x) + \sum_{l=1}^{k-1} \left(f_{n_{k+1}} - f_{n_k}\right)(x) \to (f_{n_1} + \phi)(x) =: f(x)$$

p.p. en $x \in \Omega \setminus \mathcal{N}$, donc p.p. en $x \in \Omega$.

Enfin

$$|f_{n_k}(x)| \leq |f_{n_1}(x)| + \Phi(x) \text{ p.p. en } x \in \Omega \text{ pour tout } k \geq 1 ,$$

et la fonction

$$F := |f_{n_1}| + \Phi \in L^1(\Omega) .$$

La suite extraite $(f_{n_k})_{k\geq 1}$ vérifie donc les hypothèses du théorème de convergence dominée.

En particulier

$$\begin{cases} |f_{n_k}(x) - f(x)| \to 0 \text{ p.p. en } x \in \Omega \text{ pour } k \to +\infty , \\ |f_{n_k}(x) - f(x)| = \sum_{l>k} |f_{n_{l+1}}(x) - f_{n_l}(x)| \\ \qquad\qquad\qquad \leq \Phi(x) \text{ p.p. en } x \in \Omega \text{ pour } k \geq 1 . \end{cases}$$

Par convergence dominée, lorsque $k \to +\infty$,

$$\|f_{n_k} - f\|_{L^1(\Omega)} = \int_\Omega |f_{n_k}(x) - f(x)| dx \to 0 .$$

Dans la suite de Cauchy $(f_n)_{n\geq 0}$, on a construit la suite extraite $(f_{n_k})_{k\geq 1}$ qui est convergente dans $L^1(\Omega; \mathbf{C})$ et vérifie de plus les hypothèses du théorème de convergence dominée.

Or si une suite de Cauchy dans un espace métrique admet une sous-suite convergente, la suite toute entière est convergente. Donc la suite de Cauchy $(f_n)_{n\geq 0}$ est convergente ; comme ceci vaut pour n'importe quelle suite de Cauchy dans $L^1(\Omega; \mathbf{C})$, on en déduit que $L^1(\Omega; \mathbf{C})$ est complet. □

Le théorème ci-dessus implique en particulier que toute suite convergente de $L^1(\Omega; \mathbf{C})$ admet une sous-suite convergeant p.p. sur Ω. A nouveau ce résultat est à peu près optimal, dans la mesure où une suite convergente dans $L^1(\Omega; \mathbf{C})$ n'est pas toujours toute entière convergente p.p.. (Le lecteur est invité à chercher un contre-exemple, à titre d'exercice.)

3.2 Densité de C_c dans L^1

A partir de $C_c(\Omega)$, nous avons construit l'espace de Lebesgue $L^1(\Omega)$, ainsi que l'intégrale de Lebesgue qui est une forme linéaire positive sur $L^1(\Omega)$, lequel est complet pour la norme $\|\cdot\|_{L^1}$.

On peut se demander si ce prolongement est "minimal" — c'est-à-dire le plus "économique possible".

Une réponse à cette question est fournie par le

Théorème 3.2.1 (Densité de C_c dans L^1). L'espace $C_c(\Omega;\mathbf{C})$ s'identifie à un sous-espace dense de $L^1(\Omega;\mathbf{C})$. Autrement dit, pour tout $\epsilon > 0$ et toute fonction f sommable sur Ω à valeurs complexes, il existe $\phi \in C_c(\Omega)$ telle que

$$\|f - \phi\|_{L^1} < \epsilon\,.$$

Démonstration. Montrons d'abord que $C_c(\Omega)$ s'identifie à un sous-espace de $L^1(\Omega)$. Pour cela, considérons l'application

$$C_c(\Omega;\mathbf{C}) \ni \phi \mapsto [\phi| \in L^1(\Omega;\mathbf{C})\,;$$

cette application est évidemment $\mathbf{C}$-linéaire ; montrons qu'elle est injective. En effet, dire qu'une fonction $\phi \in C_c(\Omega;\mathbf{C})$ appartient à son noyau, c'est dire que $[\phi] = 0$, ou encore, de façon équivalente, que $\phi = 0$ p.p. sur Ω. L'ensemble

$$\{x \in \Omega \,|\, \phi(x) = 0\}$$

est évidemment fermé dans Ω (comme image réciproque du fermé $\{0\}$ par l'application continue ϕ) ; de plus il est dense dans Ω (comme complémentaire d'un ensemble Lebesgue négligeable : voir Proposition 5.0.14 page 59.) Par conséquent, $C_c(\Omega)$ est isomorphe à son image par l'application linéaire injective ci-dessus, qui est un sous-espace vectoriel de $L^1(\Omega;\mathbf{C})$, et auquel on identifiera désormais $C_c(\Omega)$.

Démontrons la densité de $C_c(\Omega;\mathbf{C})$ dans $L^1(\Omega;\mathbf{C})$. Soit donc une classe $[f] \in L^1(\Omega;\mathbf{C})$, et soit $f \in \mathcal{L}^1(\Omega;\mathbf{C})$ une fonction sommable de cette classe. Quitte à considérer séparément $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$, on peut supposer que f est à valeurs réelles.

Par définition de $\mathcal{L}^1(\Omega)$, il existe donc $g, h \in \mathcal{L}^+(\Omega)$ telles que $f = g - h$ p.p. sur Ω. Soient $(g_n)_{n\geq 0}$ et $(h_n)_{n\geq 0}$, suites fondamentales sur Ω convergeant respectivement vers g et h. On a donc

$$g_0(x) \leq \ldots \leq g_n(x) \leq g_{n+1}(x) \to g(x) \text{ pour tout } x \in \Omega\,,$$
$$h_0(x) \leq \ldots \leq h_n(x) \leq h_{n+1}(x) \to g(x) \text{ pour tout } x \in \Omega\,,$$

et

$$\int_\Omega |g(x) - g_n(x)|dx = \int_\Omega g(x)dx - \int_\Omega g_n(x)dx \to 0\,,$$
$$\int_\Omega |h(x) - h_n(x)|dx = \int_\Omega h(x)dx - \int_\Omega h_n(x)dx \to 0\,,$$

pour $n \to +\infty$. Posons alors, $f_n = g_n - h_n$ pour tout $n \geq 0$; évidemment $f_n \in C_c(\Omega)$ pour tout $n \geq 0$ puisque $g_n, h_n \in C_c(\Omega)$; de plus

$$\int_\Omega |f(x) - f_n(x)|dx \leq \int_\Omega |g(x) - g_n(x)|dx + \int_\Omega |h(x) - h_n(x)|dx \to 0$$

lorsque $n \to +\infty$. Etant donné $\epsilon > 0$ arbitrairement petit, il existe donc $N_\epsilon > 0$ assez grand pour que

$$m \geq N_\epsilon \Rightarrow \int_\Omega |f(x) - f_n(x)|dx < \epsilon\,;$$

en particulier

$$\|[f] - [f_{N_\epsilon}\|_{L^1} = \int_\Omega |f(x) - f_{N_\epsilon}(x)|dx < \epsilon\,.$$

□

La densité de $C_c(\Omega)$ dans $L^1(\Omega)$, bien qu'étant une conséquence directe de la définition même de la notion de fonction sommable, a de nombreuses conséquences très importantes. En voici une que nous utiliserons à plusieurs reprises en analyse de Fourier.

Corollaire 3.2.2 (Continuité L^1 des translations). Pour tout fonction f sommable sur $\mathbf{R}^N$ à valeurs dans $\mathbf{R}$ ou dans $\mathbf{C}$,

$$\lim_{|y|\to 0} \int_{\mathbf{R}^N} |f(x-y) - f(x)|dx = 0\,.$$

En général, une fonction sommable n'est pas continue — penser à la fonction indicatrice $\mathbf{1}_{\mathbf{Q}\cap[0,1]}$ des rationnels du segment $[0,1]$, déjà rencontrée au chapitre III page 27, qui est sommable car nulle p.p., mais discontinue en tout point de $[0,1]$. La propriété ci-dessus est pourtant une sorte de *continuité intégrale* — plus précisément, la continuité de l'action des translations sur L^1 — qui est la trace résiduelle de la continuité des fonctions à partir desquelles on a construit l'espace des fonctions sommables — ou l'espace de Lebesgue L^1.

Démonstration. Soit $\epsilon > 0$; par densité de $C_c(\mathbf{R}^N;\mathbf{C})$ dans $L^1(\mathbf{R}^N;\mathbf{C})$, il existe $\phi \in C_c(\mathbf{R};\mathbf{C})$ telle que

$$\|f - \phi\|_{L^1} < \epsilon\,.$$

Soit $R > 0$ tq. l'on ait $\mathrm{supp}(\phi) \subset \overline{B(0,R)}$. Alors

$$\begin{aligned}\int_{\mathbf{R}^N} |f(x-y) - f(x)|dx \leq &\int_{\mathbf{R}} |f(x-y) - \phi(x-y)|dx \\ &+ \int_{\mathbf{R}^N} |\phi(x-y) - \phi(x)|dx \\ &+ \int_{\mathbf{R}^N} |\phi(x) - f(x)|dx = I + J + K\,.\end{aligned}$$

Effectuons le changement de variables $x \mapsto x + z$: on trouve que $I = K < \epsilon$ par choix de ϕ, de sorte que

$$\int_{\mathbf{R}^N} |f(x-y) - f(x)|dx < 2\epsilon + \int_{\mathbf{R}} |\phi(x-y) - \phi(x)|dx\,.$$

Enfin

$$\begin{cases} |\phi(x-y_n) - \phi(x)| \to 0 \text{ pour tout } x \in \mathbf{R}^N\,, \\ |\phi(x-y_n) - \phi(x)| \leq 2\sup_{y\in\mathbf{R}} |\phi(y)|\mathbf{1}_{\overline{B(0,R+1)}}(x)\,, \end{cases}$$

pour toute suite $(y_n)_{n\geq 0}$ de $\mathbf{R}^N$ telle que $y_n \to 0$ lorsque $n \to +\infty$, et pour n assez grand de sorte que $|y_n| < 1$. Par convergence dominée

$$\int_{\mathbf{R}^N} |\phi(x - y_n) - \phi(x)|dx \to 0 \quad \text{lorsque } n \to +\infty\,.$$

La suite $(y_n)_{n\geq 0}$ étant arbitraire, ceci implique que

$$\lim_{|y|\to 0} \int_{\mathbf{R}^N} |\phi(x - y) - \phi(x)|dx = 0\,.$$

Au total, étant donné $\epsilon > 0$, il existe donc $\delta > 0$ tel que

$$|y| < \delta \Rightarrow \int_{\mathbf{R}^N} |f(x - y) - f(x)|dx < 3\epsilon\,.$$

Comme $\epsilon > 0$ peut être choisi arbitrairement petit, on conclut que

$$\int_{\mathbf{R}^N} |f(x - y) - f(x)|dx \to 0 \quad \text{lorsque } |y| \to 0\,.$$

□

4 Généralisations

L'espace de Lebesgue $L^1(\Omega)$ ou $L^1(\Omega; \mathbf{C})$ est un cas particulier d'une famille d'espaces d'une importance fondamentale en analyse.

L'inégalité de Minkowski étudiée au chapitre IV page 71 montre que la quantité

$$N_p(f) := \left(\int_\Omega |f(x)|^p dx\right)^{1/p}$$

définit pour tout $p \in [1, +\infty[$ une norme sur $C_c(\Omega)$.

Définition 4.0.3 (Espace $\mathcal{L}^p$). Soit Ω ouvert de $\mathbf{R}^N$, et soit $p \in]1, +\infty[$; on désigne par $\mathcal{L}^p(\Omega)$ (resp. $\mathcal{L}^p(\Omega; \mathbf{C})$) l'ensemble des fonctions mesurables définies p.p. sur Ω à valeurs réelles (complexes) vérifiant la condition

$$\int_\Omega |f(x)|^p dx < +\infty\,.$$

On vérifie sans peine, grâce à l'inégalité de Minkowski, que $\mathcal{L}^p(\Omega)$ (resp. $\mathcal{L}^p(\Omega; \mathbf{C})$) est un $\mathbf{R}$-espace vectoriel (resp. un $\mathbf{C}$-espace vectoriel) pour l'addition et la multiplication par les scalaires des fonctions définies p.p. sur Ω, et que N_p est une semi-norme sur $\mathcal{L}^p(\Omega)$ ou $\mathcal{L}^p(\Omega; \mathbf{C})$, mais pas une norme, pour la même raison que N_1 n'est pas une norme sur $\mathcal{L}^1(\Omega)$ ou $\mathcal{L}^1(\Omega; \mathbf{C})$.

On notera en particulier que, pour $p = 2$, la norme N_2 définit la convergence en moyenne quadratique des suites de fonctions dans $C_c(\Omega)$ ou $C_c(\Omega; \mathbf{C})$; dans ce cas, l'espace $C_c(\Omega)$ ou $C_c(\Omega; \mathbf{C})$ muni de la norme N_2 est un exemple d'espace préhilbertien, non complet. (On laisse au lecteur le soin de construire une suite de Cauchy de $C([-1, 1])$ pour la norme N_2 qui ne converge pas en moyenne quadratique vers une fonction appartenant à $C_c([-1, 1])$.)

Comme dans le cas $p = 1$, on construit l'espace obtenu en identifiant les fonctions de $\mathcal{L}^p(\Omega; \mathbf{C})$ égales p.p. sur Ω.

Définition 4.0.4 (Espace de Lebesgue L^p). Soit Ω ouvert de $\mathbf{R}^N$, et soit $p \in]1, +\infty[$. Pour $f \in \mathcal{L}^p(\Omega)$ ou $\mathcal{L}^p(\Omega; \mathbf{C})$, on note $[f]$ l'esemble de toutes les fonctions mesurables définies p.p. sur Ω égales à f p.p. sur Ω, et on pose

$$L^p(\Omega) \text{ ou } L^p(\Omega; \mathbf{C}) = \{[f] \,|\, f \in \mathcal{L}^p(\Omega) \text{ ou } \mathcal{L}^p(\Omega; \mathbf{C})\}\,.$$

L'addition des éléments de $L^p(\Omega; \mathbf{C})$ et la multiplication par les scalaires est définie comme dans le cas $p = 1$: pour tous $[f], [g] \in L^p(\Omega; \mathbf{C})$ et tout $\alpha \in \mathbf{C}$, on pose

$$[f] + [g] := [f + g]\,, \qquad \alpha[f] := [\alpha f]\,.$$

Enfin la semi-norme N_p induit sur $L^p(\Omega; \mathbf{C})$ l'application

$$\|[f]\|_{L^p} := N_p(f) = \left(\int_\Omega |f(x)|^p dx\right)^{1/p},$$

qui est une norme.

Comme dans le cas $p = 1$, le fait que l'application $L^p(\Omega; \mathbf{C}) \ni [f] \mapsto \|[f]\|_{L^p}$ soit une norme sur $L^p(\Omega; \mathbf{C})$ explique pourquoi il vaut mieux considérer des classes de fonctions mesurables égales p.p. plutôt que des fonctions isolées. Et de même que dans le cas $p = 1$, la topologie induite sur $L^p(\Omega; \mathbf{C})$ vérifie les propriétés résumées dans le théorème suivant.

Théorème 4.0.5. Soit Ω ouvert de $\mathbf{R}^N$ et $p \in]1, +\infty[$. Alors

1) l'espace $L^p(\Omega; \mathbf{C})$ muni de la norme $\|\cdot\|_{L^p}$ est complet ;

2) l'espace $C_c(\Omega; \mathbf{C})$ s'identifie à un sous-espace vectoriel dense dans $L^p(\Omega; \mathbf{C})$;

3) pour tout $f \in \mathcal{L}^p(\Omega; \mathbf{C})$ on a

$$\lim_{|y| to 0} \int_\Omega |f(x - y) - f(x)|^p dx = 0\,.$$

Démonstration du point 1). Soit $(f_n)_{n\geq 0}$ suite de Cauchy dans $L^p(\Omega; \mathbf{C})$. Construisons comme dans le démonstration du Théorème 3.1.1 une suite infinie d'entiers $0 \leq n_1 < n_2 < \ldots < n_k < n_{k+1} < \ldots$ telle que

$$\|f_{n_k} - f_{n_{k+1}}\|_{L^p(\Omega)} < \tfrac{1}{2^k}\,.$$

Définissons la fonction mesurable Φ comme dans la démonstration du Théorème 3.1.1. D'après l'inégalité de Minkowski, pour tout $k \geq 1$,

$$\left\| \sum_{l=1}^{k-1} |f_{n_l} - f_{n_{l+1}}| \right\|_{L^p(\Omega)} \leq \sum_{l=1}^{k-1} \|f_{n_k} - f_{n_{k+1}}\|_{L^p(\Omega)} < 1\,,$$

c'est-à-dire

$$\int_\Omega \left(\sum_{l=1}^{k-1} |f_{n_l} - f_{n_{l+1}}|(x) \right)^p dx \leq 1\,.$$

D'après le lemme de Fatou, en passant à la limite pour $k \to +\infty$, on trouve que

$$\begin{aligned}\int_\Omega |\Phi(x)|^p dx &= \int_\Omega \lim_{k\to+\infty} \left(\sum_{l=1}^{k-1} |f_{n_l} - f_{n_{l+1}}|(x)\right)^p dx \\ &\leq \varliminf_{k\to+\infty} \int_\Omega \left(\sum_{l=1}^{k-1} |f_{n_l} - f_{n_{l+1}}|(x)\right)^p dx \leq 1\,.\end{aligned}$$

Comme Φ^p est sommable sur Ω, il existe $\mathcal{N} \subset \Omega$ Lebesgue-négligeable tel que $\Phi(x) < +\infty$ pour tout $x \in \Omega \setminus \mathcal{N}$.

Définissons alors la fonction mesurable ϕ comme dans la démonstration du Théorème 3.1.1. La série

$$\phi(x) := \sum_{k\geq 1} (f_{n_{k+1}} - f_{n_k})(x)$$

est absolument convergente, donc convergente dans $\mathbf{C}$ pour tout $x \in \Omega \setminus \mathcal{N}$, c'est-à-dire p.p. en $x \in \Omega$; elle définit donc une fonction mesurable comme limite p.p. sur Ω d'une suite de fonctions mesurables.

De plus, l'inégalité triangulaire pour le module entraîne que

$$|\phi(x)| \leq \Phi(x)\,, \quad x \in \Omega \setminus \mathcal{N}$$

de sorte que

$$\int_\Omega |\phi(x)|^p dx \leq \int_\Omega \Phi(x)^p dx < +\infty\,.$$

Donc $\phi \in \mathcal{L}^p(\Omega; \mathbf{C})$.

Posons $f := f_{n_1} + \phi$; évidemment $f \in \mathcal{L}^p(\Omega; \mathbf{C})$ comme somme de deux fonctions de $\mathcal{L}^p(\Omega; \mathbf{C})$.

Ensuite

$$\begin{aligned}\left(\int_\Omega |f_{n_m} - f_{n_k}|^p(x) dx\right)^{1/p} &\leq \sum_{l=k}^{m-1} \|f_{n_{l+1}} - f_{n_l}\|_{L^p(\Omega)} \\ &\leq \sum_{l=k}^{m-1} \frac{1}{2^l} \leq \frac{1}{2^{k-1}}\end{aligned}$$

de sorte que, d'après le lemme de Fatou

$$\begin{aligned}\int_\Omega |f - f_{n_k}|^p(x) dx &= \int_\Omega \lim_{m\to+\infty} |f_{n_m} - f_{n_k}|^p(x) dx \\ &\leq \varliminf_{m\to+\infty} \int_\Omega |f_{n_m} - f_{n_k}|^p(x) dx \leq \frac{1}{2^{k-1}}\,.\end{aligned}$$

Autrement dit, la suite extraite $(f_{n_k})_{k\geq 1}$ converge vers f en norme L^p.

Enfin, comme la suite $(f_n)_{n\geq 0}$ est de Cauchy, on en déduit qu'elle converge vers f en norme L^p. □

Le cas $p = 2$ est d'une importance considérable en analyse, en physique (tout particulièrement en mécanique quantique), et en algèbre (théorie des représentations) ; le lecteur vérifiera sans peine que la norme $\|\cdot\|_{L^2}$ découle du produit scalaire hermitien défini par

$$([\phi]|[\psi])_{L^2} := \int_\Omega \overline{\phi(x)}\psi(x)dx$$

pour tous $\phi, \psi \in \mathcal{L}^2(\Omega; \mathbf{C})$. L'espace $L^2(\Omega; \mathbf{C})$ est donc préhilbertien et complet, c'est-à-dire que c'est un espace de Hilbert. En fait, l'espace L^2 est le prototype des espaces de Hilbert, dont la théorie sera étudiée en détail dans une autre partie de ce cours. Les espaces de Hilbert, et notamment l'espace L^2, jouent un rôle absolument fondamental en mécanique quantique — analogue à celui de l'espace euclidien $\mathbf{R}^3$ en mécanique classique.

Concluons ce chapitre avec une dernière remarque. Il est possible de définir une notion d'espace $\mathcal{L}^\infty$ — et L^∞ par le même mécanisme de passage aux classes de fonctions mesurables égales p.p., ainsi qu'une norme $\|\cdot\|_{L^\infty}$.

On peut définir l'espace $\mathcal{L}^\infty(\Omega; \mathbf{C})$ comme suit. Pour toute fonction f mesurable à valeurs complexes définie p.p. sur Ω ouvert de $\mathbf{R}^N$, et pour tout $p \in [1, \infty[$, on considère l'application $\mathbf{C}$-linéaire

$$\mathcal{L}^p(\Omega; \mathbf{C}) \ni \phi \mapsto f\phi$$

à valeurs dans l'espace des fonctions mesurables complexes définies p.p. sur Ω. On démontre alors qu'il y a équivalence entre les conditions suivantes :

a) il existe $p \in [1, \infty[$ tel que l'application linéaire $\phi \mapsto f\phi$ ci-dessus soit continue de $\mathcal{L}^p(\Omega; \mathbf{C})$ dans lui-même ;

b) pour tout $p \in [1, \infty[$, l'application linéaire $\phi \mapsto f\phi$ ci-dessus soit continue de $\mathcal{L}^p(\Omega; \mathbf{C})$ dans lui-même.

On définit $\mathcal{L}^\infty(\Omega; \mathbf{C})$ comme l'ensemble des fonctions mesurables complexes définies p.p. sur Ω vérifiant a) ou b), et $L^\infty(\Omega; \mathbf{C})$ comme l'ensemble des classes de fonctions de $\mathcal{L}^\infty(\Omega; \mathbf{C})$ égales p.p. sur Ω. Pour tout élément $[f]$ de $L^\infty(\Omega; \mathbf{C})$, on définit $\|f\|_{L^\infty}$ comme étant la norme de l'application linéaire

$$\mathcal{L}^p(\Omega; \mathbf{C}) \ni \phi \mapsto f\phi \in \mathcal{L}^p(\Omega; \mathbf{C}),$$

laquelle est évidemment indépendante du représentant f de $[f]$, mais aussi — ce qui est moins évident — du nombre p.

On montre que l'espace L^∞ ainsi obtenu muni de la norme $\|\cdot\|_{L^\infty}$ est complet, et que C_c s'identifie à un sous-espace vectoriel de cet espace L^∞, qui cette fois *n'est pas dense* dans L^∞. La norme $\|\cdot\|_{L^\infty}$ restreinte à C_c y définit en effet la topologie de la convergence uniforme sur Ω des suites de fonctions. On rappelle que la limite uniforme d'une suite de fonctions continues est une fonction continue, or il existe dans $\mathcal{L}^\infty(\Omega; \mathbf{C})$ des fonctions qui ne sont pas continues — par exemple la fonction indicatrice $\mathbf{1}_{\mathbf{Q}\cap[0,1]}$ appartient à $\mathcal{L}^\infty(\mathbf{R})$.

En conséquence, la propriété 3) du théorème ci-dessus est fausse pour l'espace L^∞.

L'étude plus détaillée de l'espace L^∞ pourra être poursuivie dans les petites classes.

5 Ce qu'il faut retenir

Pour tout ouvert $\Omega \subset \mathbf{R}^N$ non vide, l'espace de Lebesgue $L^1(\Omega;\mathbf{C})$ est l'ensemble des classes de fonctions sommables sur Ω à valeurs dans $\mathbf{C}$ égales p.p. sur Ω. C'est un espace vectoriel sur $\mathbf{C}$ pour l'addition des fonctions définies p.p. Ω, et la multiplication par les scalaires complexes.

L'application

$$L^1(\Omega;\mathbf{C}) \ni [f] \mapsto \|[f]\|_{L^1(\Omega)} := \int_\Omega |f(x)|dx \in \mathbf{R}_+$$

définit une norme sur $L^1(\Omega;\mathbf{C})$.

Théorème

1) L'espace $L^1(\Omega;\mathbf{C})$ muni de la norme $\|\cdot\|_{L^1}$ est complet.

2) Toute suite convergente dans $L^1(\Omega;\mathbf{C})$ admet une sous-suite convergeant p.p. sur Ω vers la même limite.

3) L'espace $C_c(\Omega;\mathbf{C})$ s'identifie à un sous-espace vectoriel dense de $L^1(\Omega;\mathbf{C})$.

4) Pour toute fonction f sommable sur $\mathbf{R}^N$ à valeurs dans $\mathbf{C}$, on a

$$\lim_{|y|\to 0} \int_{\mathbf{R}^N} |f(x-y) - f(x)|dx = 0\,.$$

Chapitre VII

La transformation de Fourier sur L^1

1 Motivations

L'analyse de Fourier — et ses diverses généralisations — jouent un rôle de tout premier plan dans diverses branches des mathématiques (analyse des équations aux dérivées partielles, théorie du signal, analyse d'images, représentations des groupes, arithmétique...) et de la physique (optique, acoustique, mécanique quantique...)

Le lecteur a déjà rencontré l'analyse de Fourier des fonctions périodiques — c'est-à-dire la théorie des séries de Fourier — supposées assez régulières (par exemple de classe C^1 par morceaux). Mais les idées fondamentales de l'analyse de Fourier s'appliquent aussi à l'étude de fonctions plus générales, pas forcément continues ou périodiques.

On va s'intéresser tout particulièrement dans ce chapitre à l'action de la transformation de Fourier sur les fonctions intégrables sur l'espace euclidien $\mathbf{R}^N$, en insistant sur l'apport de la théorie de Lebesgue de l'intégration à l'analyse de Fourier.

Il existe de très nombreuses motivations pour étudier la transformation de Fourier ; certaines sont internes aux mathématiques, d'autres ont pour origine certains modèles mathématiques de la physique [1].

Considérons par exemple le cas de la diffraction de Fraunhofer. Rappelons qu'il s'agit de la diffraction d'une onde plane par une pupille — autrement dit une ouverture — observée sur un écran plan situé très loin de cette pupille. Plus précisément, la diffraction de Fraunhofer correspond au régime où le diamètre d de la pupille, sa distance r à l'écran (plan) d'observation et la longueur d'onde λ de l'onde incidente vérifient la condition asymptotique $d^2 \ll \lambda r$.

Alors, d'après la théorie de C. Huygens et A. Fresnel, chaque élément de surface dS sur la pupille se comporte comme une source ponctuelle émettant une onde sphérique d'amplitude proportionnelle à AdS, où A est l'amplitude de l'onde incidente au centre de dS. La figure de diffraction observée sur l'écran est produite par les interférences entre les ondes issues de ces

1. Historiquement, J. Fourier introduisit les séries de Fourier pour résoudre l'équation de la chaleur, qui régit l'évolution de la température dans un milieu soumis à l'influence de sources de chaleur. Voir F. Golse, *Distributions, analyse de Fourier et équations aux dérivées partielles*, MAT431.

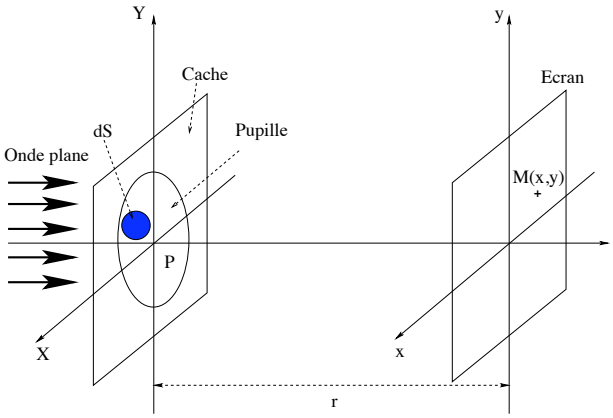


FIGURE VII.1 – La diffraction de Fraunhofer

différentes sources secondaires.

L'onde au point M de l'écran de coordonnées (x, y) émise par l'élément de surface dS de la pupille P centré au point de coordonnées (X, Y) vaut donc, en supposant l'onde plane incidente de phase nulle sur la pupille,

$$\text{Const.} AdS \frac{e^{i\frac{2\pi}{\lambda}\sqrt{r^2+(x-X)^2+(y-Y)^2}}}{\sqrt{r^2+(x-X)^2+(y-Y)^2}} .$$

(Le dénominateur correspond au facteur d'amortissement par dispersion dans la formule des potentiels retardés, et l'argument de l'exponentielle complexe au déphasage par différence de longueur de marche entre le point de coordonnées (X, Y) sur la pupille et le point M de coordonnées (x, y) de l'écran.)

La somme de ces contributions intégrée sur la surface de la pupille P vaut donc

$$\text{Const.} A \iint_P \frac{e^{i\frac{2\pi}{\lambda}\sqrt{r^2+(x-X)^2+(y-Y)^2}}}{\sqrt{r^2+(x-X)^2+(y-Y)^2}} dX dY .$$

On fait alors l'approximation suivante :

$$\begin{aligned}
\frac{e^{i\frac{2\pi}{\lambda}\sqrt{r^2+(x-X)^2+(y-Y)^2}}}{\sqrt{r^2+(x-X)^2+(y-Y)^2}} &\simeq \frac{1}{r}\exp\left(i\frac{2\pi r}{\lambda}\left(1+\frac{(x-X)^2+(y-Y)^2}{2r^2}\right)\right) \\
&\simeq \frac{1}{r}\exp\left(i\frac{2\pi r}{\lambda}\left(1+\frac{x^2+y^2}{2r^2}+O\left(\frac{d^2}{r^2}\right)\right)\right)\exp\left(i\frac{2\pi r}{\lambda}\left(-\frac{xX+yY}{r^2}\right)\right) \\
&\simeq \frac{1}{r}\exp\left(i\frac{2\pi}{\lambda r}\left(r^2+\tfrac{1}{2}(x^2+y^2)\right)\right)\exp\left(-i\frac{2\pi}{\lambda r}\left(xX+yY\right)\right)
\end{aligned}$$

sous l'hypothèse que $d^2 \ll \lambda r$. Dans le membre de droite ci-dessus, le facteur

$$\frac{1}{r}\exp\left(i\frac{2\pi}{\lambda r}\left(r^2+\tfrac{1}{2}(x^2+y^2)\right)\right)$$

est indépendant des variables d'intégration (X, Y) décrivant la surface de la pupille, de sorte que la somme des contributions des ondes élémentaires rayonnées par chaque élément de surface sur la pupille vaut en première approximation, au point M de coordonnées (x, y) sur l'écran,

$$\text{Const.}\frac{A}{r}\exp\left(i\frac{2\pi}{\lambda r}\left(r^2+\tfrac{1}{2}(x^2+y^2)\right)\right)\iint_P \exp\left(-i\frac{2\pi}{\lambda r}\left(xX+yY\right)\right)dXdY\,.$$

L'amplitude totale de l'onde diffractée par la pupille P au point M de coordonnées (x, y) de l'écran est donc proportionnelle à

$$\left|\iint_{\mathbf{R}^2}\mathbf{1}_P(X,Y)e^{-i\frac{2\pi}{\lambda r}(xX+yY)}dXdY\right|\,.$$

L'intégrale

$$\iint_{\mathbf{R}^2}\mathbf{1}_P(X,Y)e^{-i\frac{2\pi}{\lambda r}(xX+yY)}dXdY$$

est un exemple d'intégrale de Fourier — plus précisément, c'est la transformée de Fourier de la fonction indicatrice de la pupille.

Voici quelques exemples de figures de diffraction.

Exemple 1 : diffraction par une pupille carrée.

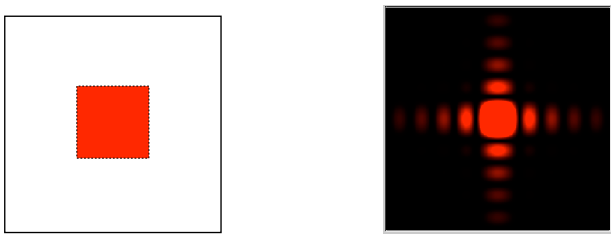


FIGURE VII.2 – A gauche, pupille carrée ; à droite, l'image diffractée sur l'écran

Exemple 2 : diffraction par une pupille circulaire

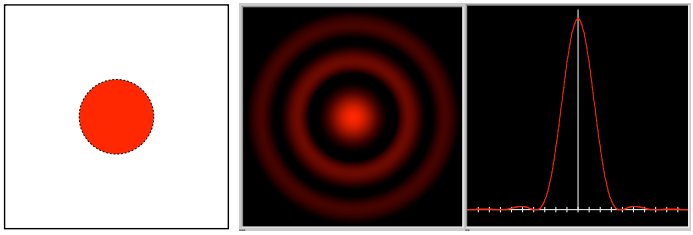


FIGURE VII.3 – A gauche, pupille circulaire ; au centre, son image diffractée ; à droite, graphe de l'intensité lumineuse en fonction de la distance au centre

L'étude des figures de diffraction suggère le problème suivant : peut-on retrouver la forme de la pupille P à partir de la fonction

$$(x,y) \mapsto \iint_{\mathbf{R}^2} \mathbf{1}_P(X,Y) e^{-i\frac{2\pi}{\lambda r}(xX+yY)} dX dY ?$$

Cette question sera le point de départ de notre étude de la transformation de Fourier.

On trouvera d'autres applications de la théorie de Fourier dans les textes remarquables de Dym-McKean [8], Malliavin-Airault [16] et Lieb-Loss [15].

2 Propriétés de la transformation de Fourier

Définition 2.0.6 (Transformation de Fourier sur L^1). A tout $f \in L^1(\mathbf{R}^N;\mathbf{C})$ la transformation de Fourier $\mathcal{F}$ associe la fonction $\hat{f} = \mathcal{F}f$ définie sur $\mathbf{R}^N$ par la formule

$$\hat{f}(\xi) = \int_{\mathbf{R}^N} e^{-i\xi\cdot x} f(x) dx\,, \quad \xi \in \mathbf{R}^N\,.$$

Il y a différentes manières de normaliser la transformation de Fourier ; on a adoptée ici celle qu'utilisent en général les spécialistes d'équations aux dérivées partielles [2].

Remarquons une différence importante entre f et sa transformée $\mathcal{F}f$: l'élé- ment $f \in L^1(\mathbf{R}^N;\mathbf{C})$ est une classe d'équivalence de fonctions ; elle s'identifie à une fonction définie p.p. sur $\mathbf{R}^N$, tandis que $\hat{f}$ est une fonction définie *en tout point* de $\mathbf{R}^N$: en effet, si $f \in L^1(\mathbf{R}^N;\mathbf{C})$, alors $x \mapsto e^{-i\xi\cdot x} f(x)$ est, pour tout $\xi \in \mathbf{R}^N$, une fonction mesurable définie p.p. sur $\mathbf{R}^N$, qui est sommable car

$$\int_{\mathbf{R}^N} \left| e^{-i\xi\cdot x} f(x) \right| dx = \int_{\mathbf{R}^N} |f(x)| dx = \|f\|_{L^1} < +\infty\,.$$

Plus précisément :

Théorème 2.0.7 (de Riemann-Lebesgue). Pour tout $f \in L^1(\mathbf{R}^N;\mathbf{C})$, sa transformée de Fourier $\hat{f} = \mathcal{F}f$ vérifie

2. Le lecteur notera la différence avec la normalisation utilisée dans le cours PHYS311.

a) $\hat{f} \in C(\mathbf{R}^N;\mathbf{C})$ et $|\hat{f}(\xi)| \leq \|f\|_{L^1}$ pour tout $\xi \in \mathbf{R}^N$;

b) $\hat{f}(\xi) \to 0$ lorsque $|\xi| \to +\infty$.

Démonstration. Pour tout $\xi \in \mathbf{R}^N$, on vient de voir que

$$|\hat{f}(\xi)| = \left| \int_{\mathbf{R}^N} e^{-i\xi\cdot x} f(x) dx \right| \leq \int_{\mathbf{R}^N} \left| e^{-i\xi\cdot x} f(x) \right| dx = \int_{\mathbf{R}^N} |f(x)| dx = \|f\|_{L^1} .$$

Montrons que la fonction $\hat{f}$ est continue sur $\mathbf{R}^N$. En effet, soit $(\xi_n)_{n\geq 0}$ suite quelconque de $\mathbf{R}^N$ convergeant vers $\xi \in \mathbf{R}^N$. Alors on a

$$\begin{cases} e^{-i\xi_n\cdot x} f(x) \to e^{-i\xi\cdot x} f(x) \text{ p.p. en } x \in \mathbf{R}^N \text{ lorsque } n \to +\infty , \\ \\ |e^{-i\xi_n\cdot x} f(x)| \leq |f(x)| \text{ p.p. en } x \in \mathbf{R}^N \text{ pour tout } n \geq 0 . \end{cases}$$

Comme $f \in L^1(\mathbf{R}^N;\mathbf{C})$, alors $|f| \in L^1(\mathbf{R}^N)$ et on déduit du théorème de convergence dominée que

$$\hat{f}(\xi_n) = \int_{\mathbf{R}^N} e^{-i\xi_n\cdot x} f(x) dx \to \int_{\mathbf{R}^N} e^{-i\xi\cdot x} f(x) dx = \hat{f}(\xi)$$

lorsque $n \to +\infty$. Comme la suite $(\xi_n)_{n\geq 0}$ convergeant vers ξ est arbitraire, ceci entraîne que $\hat{f}$ est continue sur $\mathbf{R}^N$, et établit donc le point a).

Démontrons le point b). Pour $\xi \neq 0$, on a

$$\begin{aligned} \hat{f}(\xi) &= \int_{\mathbf{R}^N} e^{-i\xi\cdot x} f(x) dx \\ &= -\int_{\mathbf{R}} e^{-i\xi\cdot\left(x+\frac{\pi}{|\xi|^2}\xi\right)} f(x) dx = -\int_{\mathbf{R}^N} e^{-i\xi\cdot y} f\left(y - \tfrac{\pi}{|\xi|^2}\xi\right) dy , \end{aligned}$$

grâce au changement de variables $y = x + \frac{\pi}{|\xi|^2}\xi$. Ainsi

$$\hat{f}(\xi) = \tfrac{1}{2} \int_{\mathbf{R}^N} e^{-i\xi\cdot x} \left(f(x) - f\left(x - \tfrac{\pi}{|\xi|^2}\xi\right)\right) dx ,$$

de sorte que

$$|\hat{f}(\xi)| \leq \tfrac{1}{2} \int_{\mathbf{R}^N} \left| f(x) - f\left(x - \tfrac{\pi}{|\xi|^2}\xi\right)\right| dx = \left\| f - f\left(\cdot - \tfrac{\pi}{|\xi|^2}\xi\right)\right\|_{L^1} \to 0$$

lorsque $|\xi| \to +\infty$, d'après le Corollaire 3.2.2 page 120 du Théorème 3.2.1 page 119 de densité de $C_c(\mathbf{R}^N;\mathbf{C})$ dans $L^1(\mathbf{R}^N;\mathbf{C})$. □

On sait que les coefficients de Fourier d'une fonction périodique sur $\mathbf{R}$ décroissent d'autant plus rapidement que la fonction est régulière — plus précisé- ment, si u est une fonction continue sur $\mathbf{R}$ et périodique de période un, on a l'implication

$$u \text{ de classe } C^m \Rightarrow \hat{u}(k) = \int_0^1 e^{i2\pi kx} u(x) dx = O\left(\frac{1}{k^m}\right) \text{ pour } |k| \to +\infty .$$

Il en va de même pour la transformation de Fourier, comme le montre l'énoncé suivant.

Théorème 2.0.8 (Transformation de Fourier et dérivation). Soit f appartenant à $L^1(\mathbf{R}^N;\mathbf{C})$.

a)Si $|x|f \in L^1(\mathbf{R}^N;\mathbf{C})$, alors $\hat{f} \in C^1(\mathbf{R}^N;\mathbf{C})$ et on a, pour tout $k = 1,\dots,N$

$$\frac{\partial \hat{f}}{\partial \xi_k}(\xi) = \int_{\mathbf{R}^N} -ix_k e^{-i\xi\cdot x} f(x)dx = -i\widehat{x_k f}(\xi)\,.$$

b) Si $f \in C^1(\mathbf{R}^N;\mathbf{C})$ et $\frac{\partial f}{\partial x_k} \in L^1(\mathbf{R}^N;\mathbf{C})$, alors, pour tout $k = 1,\dots,N$

$$\widehat{\frac{\partial f}{\partial x_k}}(\xi) = \int_{\mathbf{R}^N} e^{-i\xi\cdot x}\frac{\partial f}{\partial x_k}(x)dx = i\xi_k \hat{f}(\xi)\,.$$

Cet énoncé est l'analogue, dans le cadre de la transformation de Fourier, de ce que nous avons rappelé à propos des coefficients de Fourier des fonctions périodiques. Il implique en effet que

a) plus une fonction décroît vite à l'infini, plus sa transformée de Fourier est régulière. Soit $f \in C(\mathbf{R}^N;\mathbf{C})$; on a l'implication suivante :

$$\text{il existe } \epsilon > 0 \text{ t.q. } f(x) = O(|x|^{-N-k-\epsilon}) \Rightarrow \hat{f} \in C^k(\mathbf{R}^N)\,.$$

(Appliquer successivement le théorème ci-dessus à $\hat{f}$ et à ses dérivées partielles d'ordre $\leq k-1$.)

b) plus une fonction est régulière, plus sa transformée de Fourier décroît vite à l'infini. Soit $f \in C^k(\mathbf{R};\mathbf{C})$;

$$f^{(m)} \in L^1(\mathbf{R}) \text{ pour tout } 0 \leq m \leq k \Rightarrow \hat{f}(\xi) = o(|\xi|^{-k})$$

(Appliquer successivement le théorème ci-dessus à f, puis à f',..., jusqu'à f^{k-1}.)

Ces remarques montrent que la transformation de Fourier $\mathcal{F}$ échange régularité et décroissance à l'infini.

Passons à la démonstration du Théorème 2.0.8. Elle repose sur le théorème de dérivation sous le signe somme, que nous présentons ci-dessous.

Théorème 2.0.9 (Dérivation sous le signe somme). Soient I intervalle ouvert de $\mathbf{R}$ et Ω ouvert de $\mathbf{R}^N$; soit $f : I\times\Omega \ni (t,x) \mapsto f(t,x) \in \mathbf{C}$ vérifiant

$$\begin{cases} f(t,\cdot) \in \mathcal{L}^1(\Omega;\mathbf{C}) \text{ pour tout } t\in I, \\ f(\cdot,x) \text{ dérivable sur } I \text{ p.p. en } x \in \Omega\,. \end{cases}$$

S'il existe $\Phi \in \mathcal{L}^1(\Omega)$ tel que, p.p. en $x \in \Omega$, l'on ait

$$\left|\frac{\partial f}{\partial t}(t,x)\right| \leq \Phi(x) \quad \text{pour tout } t \in I\,,$$

alors la fonction

$$I \ni t \mapsto F(t) = \int_\Omega f(t,x)dx \text{ est dérivable sur } I\,,$$

et

$$F'(t) = \int_\Omega \frac{\partial f}{\partial t}(t,x)dx \text{ pour tout } t \in I\,.$$

De plus, si $f(\cdot,x) \in C^1(I;\mathbf{C})$ p.p. en $x \in \Omega$, alors la fonction $F \in C^1(I;\mathbf{C})$.

Démonstration. Soit $(h_n)_{n\geq 0}$ suite de $\mathbf{R}$ t.q. $h_n \to 0$ lorsque $n \to +\infty$, et soit $t_0 \in I$; étudions la suite

$$\frac{F(t_0+h_n)-F(t_0)}{h_n} = \int_\Omega \frac{f(t_0+h_n,x)-f(t_0,x)}{h_n} dx\,.$$

Par hypothèse, il existe $\mathcal{N} \subset \Omega$ négligeable tel que

$$\frac{f(t_0+h_n,x)-f(t_0,x)}{h_n} \to \frac{\partial f}{\partial t}(t_0,x) \text{ pour tout } x \in \Omega \setminus \mathcal{N}$$

lorsque $n \to +\infty$. D'autre part,

$$\left|\frac{f(t_0+h_n,x)-f(t_0,x)}{h_n}\right| \leq \Phi(x) \text{ pour tout } x \in \Omega \setminus \mathcal{N}$$

d'après le théorème des accroissements finis, pour tout $n \in \mathbf{N}$ assez grand pour que $t_0+h_n \in I$, de sorte que le segment $[t_0, t_0+h_n]$ (ou $[t_0+h_n, t_0]$ selon que h_n est positif ou négatif) soit inclus dans I. Par convergence dominée

$$\frac{F(t_0+h_n)-F(t_0)}{h_n} = \int_\Omega \frac{f(t_0+h_n,x)-f(t_0,x)}{h_n} dx \to \int_\Omega \frac{\partial f}{\partial t}(t_0,x)dx$$

lorsque $n \to +\infty$. Comme la suite $(h_n)_{n\geq 0}$ tendant vers 0 est arbitraire, ainsi que $t_0 \in I$, on en déduit que F est dérivable sur I et que

$$F'(t) = \int_\Omega \frac{\partial f}{\partial t}(t,x)dx \quad \text{ pour tout } t \in I\,.$$

De plus, supposons qu'il existe $\mathcal{N}' \subset \Omega$ négligeable t.q. $f(\cdot,x) \in C^1(I)$ pour tout $x \in \Omega \setminus \mathcal{N}'$. Soit $t^* \in I$ et $(t_n)_{n\geq 0}$ suite quelconque de I convergeant vers t^* lorsque $n \to +\infty$. Alors

$$\begin{cases} \frac{\partial f}{\partial t}(t_n,x) \to \frac{\partial f}{\partial t}(t^*,x) \text{ pour tout } x \in \Omega \setminus \mathcal{N}'\,, \\ \left|\frac{\partial f}{\partial t}(t_n,x)\right| \leq \Phi(x) \text{ pour tout } x \in \Omega \setminus \mathcal{N}' \text{ et } n \geq 0\,. \end{cases}$$

Par convergence dominée

$$F'(t_n) = \int_\Omega \frac{\partial f}{\partial t}(t_n,x)dx \to \int_\Omega \frac{\partial f}{\partial t}(t^*,x)dx = F'(t^*)$$

lorsque $n \to +\infty$. Comme ceci vaut pour toute suite $(t_n)_{n\geq 0}$ convergeant vers t^*, on en déduit que F' est continue en t^* ; d'autre part, t^* étant un point arbitraire de I, il s'ensuit que $F' \in C^1(I)$. □

Appliquons maintenant le théorème de dérivation sous le signe somme à la transformation de Fourier.

Démonstration. Soit $\mathcal{N} \subset \Omega$ négligeable t.q. $|f(x)| < +\infty$ pour tout $x \in \mathbf{R}^N \setminus \mathcal{N}$; alors la fonction

$$\phi : \mathbf{R}^N \times \mathbf{R}^N \ni (\xi,x) \mapsto e^{-i\xi\cdot x} f(x) \in \mathbf{C}$$

vérifie

$$\begin{cases} \phi(\cdot, x) \in C^1(\mathbf{R}^N; \mathbf{C}) \text{ pour tout } x \in \mathbf{R}^N \setminus \mathcal{N}\,, \\ \phi(\xi, \cdot) \in L^1(\Omega; \mathbf{C}) \text{ pour tout } \xi \in \mathbf{R}^N\,. \end{cases}$$

D'autre part,

$$\left| \frac{\partial\, \phi}{\partial\, \xi_k}(\xi, x) \right| = |x_k||f(x)| \leq |x||f(x)| \text{ pour tout } x \in \mathbf{R}^N \setminus \mathcal{N} \text{ et } \xi \in \mathbf{R}^N\,.$$

Par théorème de dérivation sous le signe somme, on conclut que $\hat{f} \in C^1(\mathbf{R}^N)$ et que

$$\frac{\partial\, \hat{f}}{\partial\, \xi_k}(\xi) = \int_{\mathbf{R}^N} -ix_k e^{-i\xi\cdot x} f(x) dx\,,$$

pour tout $x \in \mathbf{R}^N$, ce qui établit le a).

Pour simplifier les notations, nous allons traiter le seul cas où l'ordre de dérivation $k = 1$ dans le point b). Comme f est de classe C^1 sur $\mathbf{R}^N$, une intégration par parties montre que

$$\int_a^b e^{-i\xi\cdot x} \frac{\partial\, f}{\partial\, x_1}(x) dx_1 = \Big[e^{-i\xi\cdot x} f(x) \Big]_{x_1=a}^{x_1=b} + i\xi_1 \int_a^b e^{-i\xi\cdot x} f(x) dx_1\,.$$

Notons $x' = (x_2, \ldots, x_N)$; comme f et $\frac{\partial\, f}{\partial\, x_1}$ sont sommables sur $\mathbf{R}^N$, on déduit du théorème de Fubini que

$$\begin{aligned} \int_{\mathbf{R}^{N-1}} \left(\int_a^b e^{-i\xi\cdot x} \frac{\partial\, f}{\partial\, x_1}(x) dx_1 \right) dx' &= \int_{\mathbf{R}^N} e^{-i\xi\cdot x} \frac{\partial\, f}{\partial\, x_1}(x) \mathbf{1}_{[a,b]}(x_1) dx \\ \int_{\mathbf{R}^{N-1}} \left(\int_a^b e^{-i\xi\cdot x} f(x) dx_1 \right) dx' &= \int_{\mathbf{R}^N} e^{-i\xi\cdot x} f(x) \mathbf{1}_{[a,b]}(x_1) dx \end{aligned}$$

Comme $\mathbf{1}_{[a,b]}(x_1) \to 1$ lorsque $a \to -\infty$ et $b \to +\infty$, et que f et $\frac{\partial\, f}{\partial\, x_1}$ sont sommables sur $\mathbf{R}^N$, on déduit du théorème de convergence dominée et des deux égalités ci-dessus que

$$\begin{aligned} \int_{\mathbf{R}^{N-1}} \left(\int_{a_n}^{b_n} e^{-i\xi\cdot x} \frac{\partial\, f}{\partial\, x_1}(x) dx_1 \right) dx' &\to \int_{\mathbf{R}^N} e^{-i\xi\cdot x} \frac{\partial\, f}{\partial\, x_1}(x) dx \\ \int_{\mathbf{R}^{N-1}} \left(\int_{a_n}^{b_n} e^{-i\xi\cdot x} f(x) dx_1 \right) dx' &\to \int_{\mathbf{R}^N} e^{-i\xi\cdot x} f(x) dx \end{aligned}$$

pour toutes les suites $(a_n)_{n\geq 0}$ et $(b_n)_{n\geq 0}$ convergeant respectivement vers $-\infty$ et $+\infty$.

D'autre part, comme f est de classe C^1 et sommable sur $\mathbf{R}^N$, la fonction

$$x_1 \mapsto \int_{\mathbf{R}^{N-1}} e^{-i\xi\cdot x} f(x_1, x') dx'$$

est définie continue et sommable sur $\mathbf{R}$. Il existe donc deux suites $(a_n)_{n\geq 0}$ et $(b_n)_{n\geq 0}$ telles que $a_n \to -\infty$ et $b_n \to +\infty$ lorsque $n \to +\infty$ et

$$\begin{aligned} \int_{\mathbf{R}^{N-1}} e^{-i\xi\cdot x} f(a_n, x') dx' &\to 0 \\ \int_{\mathbf{R}^{N-1}} e^{-i\xi\cdot x} f(b_n, x') dx' &\to 0 \end{aligned}$$

lorsque $n \to +\infty$.

En injectant ces deux suites $(a_n)_{n\geq 0}$ et $(b_n)_{n\geq 0}$ dans l'identité

$$\int_{a_n}^{b_n} e^{-i\xi\cdot x}\frac{\partial f}{\partial x_1}(x)dx_1 = \Big[e^{-i\xi\cdot x}f(x)\Big]_{x_1=a_n}^{x_1=b_n} + i\xi_1\int_{a_n}^{b_n} e^{-i\xi\cdot x}f(x)dx_1$$

que l'on intègre sur $\mathbf{R}^{N-1}$ par rapport à x', et en passant à la limite pour $n \to +\infty$, on aboutit à l'égalité

$$\int_{\mathbf{R}^N} e^{-i\xi\cdot x}\frac{\partial f}{\partial x_1}(x)dx = i\xi_1\int_{\mathbf{R}^N} e^{-i\xi\cdot x}f(x)dx\,.$$

□

3 Inversion de la transformation de Fourier

Commençons par un calcul explicite de transformée de Fourier, qui est d'une grande utilité dans différentes branches des mathématiques, et en particulier dans le calcul des probabilités.

Définition 3.0.10 (Densité gaussienne). Pour tout $a > 0$, on appelle densité gaussienne sur $\mathbf{R}^N$ de matrice de covariance aI la fonction

$$G_a(x) = \frac{1}{(2\pi a)^{N/2}}e^{-|x|^2/2a}\,, \quad x \in \mathbf{R}^N\,.$$

En particulier, pour tout $a > 0$, la fonction G_a est continue sur $\mathbf{R}^N$ et vérifie

$$G_a(x) > 0 \text{ pour tout } x \in \mathbf{R}^N\,, \quad \text{et } \int_{\mathbf{R}^N} G_a(x)dx = 1\,.$$

Autrement dit, G_a est une densité de probabilité sur $\mathbf{R}^N$.

Proposition 3.0.11 (Transformée de Fourier des gaussiennes). Pour tout $a > 0$, on a

$$\widehat{G_a}(\xi) = e^{-a|\xi|^2/2}\,, \quad \xi \in \mathbf{R}^N\,.$$

Autrement dit, pour tout $a > 0$, on a

$$\widehat{G_a} = \left(\frac{2\pi}{a}\right)^{N/2} G_{1/a}\,.$$

Cette proposition montre que, sur les densités gaussiennes, la transformation de Fourier consiste à changer la matrice de covariance en son inverse — abstraction faite du facteur de normalisation $\left(\frac{2\pi}{a}\right)^{N/2}$.

Démonstration. Commençons par traiter le cas $N = 1$.

Pour tout $a > 0$, la fonction G_a appartient à $C^\infty(\mathbf{R})$ et vérifie

$$G_a'(x) + \tfrac{1}{a}xG_a(x) = 0\,, \quad x \in \mathbf{R}\,.$$

Comme les fonctions G_a, xG_a et $G_a' \in \mathcal{L}^1(\mathbf{R})$, on déduit du Théorème 2.0.8 page 132 que la transformée de Fourier $\widehat{G_a} \in C^1(\mathbf{R})$, et que

$$\widehat{G_a'}(\xi) = i\xi\widehat{G_a}(\xi)\,, \quad \left(\widehat{G_a}\right)'(\xi) = -i\widehat{xG_a}(\xi)\,.$$

On déduit de ces deux identités, et de l'égalité ci-dessus vérifiée par G_a et G_a', que

$$\xi\widehat{G_a}(\xi) + \tfrac{1}{a}\left(\widehat{G_a}\right)'(\xi) = 0\,, \quad \xi \in \mathbf{R}^N\,.$$

Cette égalité est une équation différentielle satisfaite par $\widehat{G_a}$, équation différentielle dont la solution générale est de la forme

$$\widehat{G_a}(\xi) = Ce^{-a\xi^2/2}$$

où C est une constante restant à déterminer. Or

$$C = \widehat{G_a}(0) = \int_{\mathbf{R}} G_a(x)dx = 1\,.$$

Dans le cas général où $N \geq 2$, on a

$$\begin{aligned}\int_{\mathbf{R}^N} e^{-i\xi\cdot x}\tfrac{1}{(2\pi a)^{N/2}}e^{-|x|^2/2a}dx &= \int_{\mathbf{R}^N}\prod_{k=1}^N e^{-i\xi_k x_k}\tfrac{1}{\sqrt{2\pi a}}e^{-x_k^2/2a}dx\\ &= \prod_{k=1}^N\int_{\mathbf{R}} e^{-i\xi_k x_k}\tfrac{1}{\sqrt{2\pi a}}e^{-x_k^2/2a}dx_k\\ &= \prod_{k=1}^N e^{-a\xi_k^2/2} = e^{-a|\xi|^2}\,.\end{aligned}$$

□

Le calcul de la transformée de Fourier des densités gaussiennes joue un rôle important dans le théorème d'inversion de la transformation de Fourier dans le cadre L^1, que nous énonçons maintenant.

Théorème 3.0.12 (d'inversion de Fourier). Soit $f \in L^1(\mathbf{R}^N;\mathbf{C})$ t.q. $\hat{f} \in L^1(\mathbf{R}^N;\mathbf{C})$. Alors, p.p. en $x \in \mathbf{R}^N$

$$f(x) = \tfrac{1}{(2\pi)^N}\int_{\mathbf{R}^N} e^{+ix\cdot\xi}\hat{f}(\xi)d\xi\,.$$

Une remarque sur l'énoncé de la formule d'inversion : on observera que le membre de droite, à savoir la fonction

$$x \mapsto \tfrac{1}{(2\pi)^N}\int_{\mathbf{R}^N} e^{+ix\cdot\xi}\hat{f}(\xi)d\xi$$

est continue sur $\mathbf{R}^N$ et même tend vers 0 lorsque $|x| \to +\infty$ d'après le théorème de Riemann-Lebesgue, puisque cette fonction est précisément $\frac{1}{(2\pi)^N}\mathcal{F}\hat{f}\circ(-\mathrm{Id}_{\mathbf{R}^N})$ et que $\hat{f} \in L^1(\mathbf{R}^N;\mathbf{C})$ par hypothèse. Mais l'égalité qu'exprime le théorème d'inversion n'a de sens que p.p. en $x \in \mathbf{R}^N$, puisque f est un élément de $L^1(\mathbf{R}^N)$, c'est-à-dire que f est identifiée à une fonction sommable définie seulement p.p. sur $\mathbf{R}^N$.

Ainsi, le théorème d'inversion ci-dessus ne vaut que pour des fonctions f sommables, p.p. égales à une fonction continue sur $\mathbf{R}^N$ qui tend vers 0 à l'infini.

Avant de donner la démonstration proprement dite de ce théorème, qui pourrait paraître un peu mystérieuse, nous allons commencer par en expliquer l'idée principale.

La formule d'inversion que l'on cherche à établir s'écrit

$$f(x) = \tfrac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{+ix\cdot\xi} \left(\int_{\mathbf{R}^N} e^{-i\xi\cdot y} f(y) dy \right) d\xi \,.$$

Il semble naturel d'essayer d'échanger l'ordre des intégrations par rapport aux variables y et ξ :

$$f(x) = \int_{\mathbf{R}^N} f(y) \left(\int_{\mathbf{R}^N} e^{+i(x-y)\cdot\xi} \tfrac{d\xi}{(2\pi)^N} \right) dy \,,$$

ce qui conduit à la formule

$$f(x) = \int_{\mathbf{R}^N} K(x-y) f(y) dy \,,$$

avec

$$K(z) = \tfrac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{-iz\cdot\xi} d\xi \,.$$

Malheureusement, ce raisonnement semble conduire à une impasse.

En effet, d'une part on ne sait pas calculer explicitement l'intégrale définissant la fonction K, et d'ailleurs, cette intégrale n'a aucun sens dans la théorie de Lebesgue, car la fonction $\xi \mapsto e^{-iz\cdot\xi}$ n'est sommable sur $\mathbf{R}^N$ pour aucune valeur de $z \in \mathbf{R}^N$, puisque

$$\int_{\mathbf{R}^N} \left| e^{-iz\cdot\xi} \right| d\xi = \int_{\mathbf{R}^N} d\xi = +\infty$$

pour tout $z \in \mathbf{R}^N$.

D'autre part l'interversion des intégrales par rapport aux variables ξ et y ci-dessus ne peut pas se justifier grâce au théorème de Fubini puisque la fonction

$$(y, \xi) \mapsto e^{i\xi\cdot(x-y)} f(y)$$

n'est pas sommable sur $\mathbf{R}^N \times \mathbf{R}^N$, car

$$\begin{aligned} \iint_{\mathbf{R}^N \times \mathbf{R}^N} \left| e^{i\xi\cdot(x-y)} f(y) \right| d\xi dy &= \iint_{\mathbf{R}^N \times \mathbf{R}^N} |f(y)| \, d\xi dy \\ &= \int_{\mathbf{R}^N} d\xi \int_{\mathbf{R}^N} |f(y)| dy = +\infty \end{aligned}$$

sauf si $f(y) = 0$ p.p. en $y \in \mathbf{R}^N$.

Ces deux obstacles disparaissent en même temps si on rend l'intégrande dans $K(z)$ sommable en le multipliant par un poids gaussien, c'est-à-dire si on pose, pour tout $\epsilon > 0$ et tout $z \in \mathbf{R}^N$,

$$K_\epsilon(z) = \tfrac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{-iz\cdot\xi} e^{-\frac{1}{2}\epsilon^2|\xi|^2} d\xi \,.$$

D'une part, on sait calculer explicitement K_ϵ pour tout $\epsilon > 0$ grâce à la Proposition 3.0.11 page 135.

D'autre part, pour tout $\epsilon > 0$, l'interversion d'intégrales ci-dessus devient licite, et on a

$$\begin{aligned}
&\tfrac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{+ix\cdot\xi - \frac{1}{2}\epsilon^2|\xi|^2} \hat{f}(\xi) d\xi \\
&\quad = \tfrac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{+ix\cdot\xi - \frac{1}{2}\epsilon^2|\xi|^2} \left(\int_{\mathbf{R}^N} e^{-i\xi\cdot y} f(y) dy \right) d\xi \\
&\quad = \int_{\mathbf{R}^N} f(y) \left(\tfrac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{i(x-y)\cdot\xi} e^{-\frac{1}{2}\epsilon^2|\xi|^2} d\xi \right) dy \\
&\quad = \int_{\mathbf{R}^N} f(y) K_\epsilon(x-y) dy
\end{aligned}$$

La démonstration du théorème va consister à montrer que le membre de gauche converge vers

$$\tfrac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{+ix\cdot\xi} \hat{f}(\xi) d\xi$$

pour tout $x \in \mathbf{R}^N$, et que le membre de droite converge p.p. en $x \in \mathbf{R}^N$ vers $f(x)$, à extraction d'une sous-suite près lorsque $\epsilon \to 0$.

Curieusement, l'interversion non justifiée d'intégrales effectuée ci-dessus, qui semblait aboutir à une impasse, nous a pourtant mis sur la voie de la démonstration correcte [3].

Démonstration. Etape 1 : D'après la Proposition 3.0.11 page 135, pour tout $\epsilon > 0$,

$$\begin{aligned}
G_\epsilon(z) = \frac{1}{(2\pi\epsilon)^{N/2}} \hat{G}_{1/\epsilon}(z) &= \frac{1}{(2\pi\epsilon)^{N/2}} \int_{\mathbf{R}^N} e^{i\xi\cdot z} \frac{e^{-\frac{1}{2}\epsilon^2|\xi|^2}}{(2\pi/\epsilon)^{N/2}} d\xi \\
&= \int_{\mathbf{R}^N} e^{i\xi\cdot z - \frac{1}{2}\epsilon^2|\xi|^2} \tfrac{d\xi}{(2\pi)^N}, \quad z \in \mathbf{R}^N .
\end{aligned}$$

On en déduit l'égalité

$$\begin{aligned}
\int_{\mathbf{R}^N} f(y) G_{\epsilon^2}(x-y) dy &= \int_{\mathbf{R}^N} f(y) \left(\int_{\mathbf{R}^N} e^{i\xi\cdot(x-y) - \frac{1}{2}\epsilon^2|\xi|^2} \tfrac{d\xi}{(2\pi)^N} \right) dy \\
= \int_{\mathbf{R}^N} e^{i\xi\cdot x - \frac{1}{2}\epsilon^2|\xi|^2} \left(\int_{\mathbf{R}^N} e^{-i\xi\cdot y} f(y) dy \right) \tfrac{d\xi}{(2\pi)^N} &= \int_{\mathbf{R}^N} e^{i\xi\cdot x - \frac{1}{2}\epsilon^2|\xi|^2} \hat{f}(\xi) \tfrac{d\xi}{(2\pi)^N} .
\end{aligned}$$

3. En fait, ce schéma de démonstration, et notamment l'expression

$$\text{“}K(z) = \tfrac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{i\xi\cdot z} d\xi''$$

ont un sens dans la théorie des distributions, vaste généralisation de la théorie des fonctions, qui dépasse le cadre de ce cours, mais sera étudiée dans le cours MAT431 de deuxième année. Dans cette théorie, on montre que $K = \delta_0$, à savoir la masse de Dirac en 0, introduite dans la section 5 page 105, et la formule d'inversion de Fourier se réduit alors à l'identité

$$f(x) = \text{“} \int_{\mathbf{R}^N} f(y) K(x-y) dy = \int_{\mathbf{R}^N} f(x-y) K(y) dy\text{”} = \langle \delta_0, f(x-\cdot) \rangle$$

qui découle de la définition même de la masse de Dirac.

En effet, l'interversion des intégrales en y et ξ (la seconde égalité ci-desus) est justifiée grâce au théorème de Fubini, car

$$(y,\xi) \mapsto f(y)e^{i\xi(x-y)-\frac{1}{2}\epsilon^2|\xi|^2} \in L^1(\mathbf{R}^N \times \mathbf{R}^N)\,,$$

puisque

$$\iint_{\mathbf{R}^N\times\mathbf{R}^N} \left| f(y)e^{i\xi\cdot(x-y)-\frac{1}{2}\epsilon^2|\xi|^2} \right| dyd\xi = \int_{\mathbf{R}^N} |f(y)|dy \int_{\mathbf{R}^N} e^{-\frac{1}{2}\epsilon^2|\xi|^2} d\xi < +\infty\,.$$

Etape 2 : Passons à la limite en $\epsilon \to 0^+$ dans le membre de droite : pour tout $x \in \mathbf{R}^N$

$$\begin{cases} e^{i\xi\cdot x-\frac{1}{2}\epsilon^2|\xi|^2}\hat{f}(\xi) \to e^{i\xi\cdot x}\hat{f}(\xi) \text{ p.p. en } \xi \in \mathbf{R}^N \\ \left| e^{i\xi\cdot x-\frac{1}{2}\epsilon^2|\xi|^2}\hat{f}(\xi) \right| \le |\hat{f}(\xi)| \text{ p.p. en } \xi \in \mathbf{R}^N \text{ pour tout } \epsilon > 0 \end{cases}$$

Par convergence dominée (puisque $\hat{f} \in L^1(\mathbf{R}^N;\mathbf{C})$)

$$\int_{\mathbf{R}} e^{i\xi\cdot x-\frac{1}{2}\epsilon^2|\xi|^2}\hat{f}(\xi)\tfrac{d\xi}{(2\pi)^N} \to \int_{\mathbf{R}} e^{i\xi\cdot x}\hat{f}(\xi)\tfrac{d\xi}{(2\pi)^N}$$

pour tout $x \in \mathbf{R}^N$ lorsque $\epsilon \to 0$.

Etape 3 : Etudions le membre de gauche

$$\int_{\mathbf{R}^N} f(y)G_{\epsilon^2}(x-y)dy = \int_{\mathbf{R}^N} f(x-\epsilon z)e^{-|z|^2/2}\tfrac{dz}{(2\pi)^{N/2}}$$

grâce au changement de variables $y = x - \epsilon z$.

Comme

$$\int_{\mathbf{R}^N} e^{-|z|^2/2}\tfrac{dz}{(2\pi)^{N/2}} = 1\,,$$

on voit que

$$\begin{aligned} &\int_{\mathbf{R}^N} \left| \int_{\mathbf{R}^N} f(y)G_{\epsilon^2}(x-y)dy - f(x) \right| dx \\ &\quad = \int_{\mathbf{R}^N} \left| \int_{\mathbf{R}^N} (f(x-\epsilon z)-f(x))e^{-|z|^2/2}\tfrac{dz}{(2\pi)^{N/2}} \right| dx \\ &\quad \le \int_{\mathbf{R}^N} e^{-|z|^2/2} \left(\int_{\mathbf{R}^N} |f(x-\epsilon z)-f(x)|dx \right) \tfrac{dz}{(2\pi)^{N/2}}\,. \end{aligned}$$

D'après le Corollaire 3.2.2 page 120 du théorème de densité de $C_c(\mathbf{R}^N;\mathbf{C})$ dans $L^1(\mathbf{R}^N;\mathbf{C})$, on a, pour toute suite $(\epsilon_n)_{n\ge 0}$ convergeant vers 0,

$$\begin{cases} \displaystyle\int_{\mathbf{R}^N} |f(x-\epsilon_n z)-f(x)|dx \to 0 \text{ pour tout } z \in \mathbf{R} \text{ quand } n \to +\infty\,, \\ \displaystyle e^{-|z|^2/2}\int_{\mathbf{R}^N} |f(x-\epsilon_n z)-f(x)|dx \le 2\|f\|_{L^1(\mathbf{R}^N)}e^{-|z|^2/2}\,, \end{cases}$$

de sorte que, par convergence dominée

$$\int_{\mathbf{R}^N} \left| \int_{\mathbf{R}^N} f(y) G_{\epsilon_n^2}(x-y)dy - f(x) \right| dx \to 0$$

quand $n \to +\infty$. Comme la suite $(\epsilon_n)_{n\geq 0}$ convergeant vers 0 est arbitraire, on en conclut que

$$\int_{\mathbf{R}^N} \left| \int_{\mathbf{R}^N} f(y) G_{\epsilon^2}(x-y)dy - f(x) \right| dx \to 0$$

quand $\epsilon \to 0$.

Etape 4 : pour tout $x \in \mathbf{R}^N$ et tout $\epsilon > 0$, on a l'égalité

$$\int_{\mathbf{R}^N} f(y) G_{\epsilon^2}(x-y)dy = \int_{\mathbf{R}^N} e^{i\xi\cdot x - \frac{1}{2}\epsilon^2|\xi|^2} \hat{f}(\xi) \tfrac{d\xi}{(2\pi)^N} .$$

D'après l'étape 2, pour tout x et lorsque $\epsilon \to 0$,

$$\text{membre de droite } \to \int_{\mathbf{R}^N} e^{i\xi\cdot x} \hat{f}(\xi) \tfrac{d\xi}{(2\pi)^N}$$

D'après l'étape 3

$$\text{membre de gauche } \to f$$

dans $L^1(\mathbf{R}^N;\mathbf{C})$ quand $\epsilon \to 0$, et donc p.p. sur $\mathbf{R}^N$, quitte à extraire une sous-suite $\epsilon_k \to 0$, d'après le Théorème 3.1.2 page 116.

Donc

$$f(x) = \int_{\mathbf{R}^N} e^{i\xi\cdot x} \hat{f}(\xi) \tfrac{d\xi}{(2\pi)^N} \text{ p.p. en } x \in \mathbf{R}^N .$$

□

4 Bilan et remarques finales

Commençons par un énoncé qui synthétise les propriétés de la transformation de Fourier que nous venons d'établir.

Théorème 4.0.13. La transformation de Fourier est une application linéaire continue injective

$$\mathcal{F} : L^1(\mathbf{R}^N;\mathbf{C}) \to C_0(\mathbf{R}^N;\mathbf{C})$$

de norme $\|\mathcal{F}\| \leq 1$, où on a noté $C_0(\mathbf{R}^N;\mathbf{C})$ l'ensemble des fonctions continues sur $\mathbf{R}^N$ tendant vers 0 à l'infini, muni de la norme de la convergence uniforme — c'est-à-dire que

$$\|\phi\|_{C_0(\mathbf{R}^N;\mathbf{C})} = \sup_{x\in\mathbf{R}^N} |\phi(x)| .$$

Que $\mathcal{F}$ soit une application linéaire continue vérifiant $\|\mathcal{F}\| \leq 1$ découle du théorème de Riemann-Lebesgue. Que $\mathcal{F}$ soit injective découle du théorème d'inversion. En effet, si $f \in L^1(\mathbf{R}^N;\mathbf{C})$ vérifie $\mathcal{F}f = 0$, on a évidemment $\mathcal{F}f \in L^1(\mathbf{R}^N;\mathbf{C})$ de sorte que la formule d'inversion s'applique et donne $f = 0$ p.p. sur $\mathbf{R}^N$. Ceci montre que $\mathrm{Ker}(\mathcal{F}) = \{0\}$, d'où l'injectivité de $\mathcal{F}$.

En revanche, il faut savoir que $\mathcal{F}$ n'est pas surjective de $L^1(\mathbf{R}^N;\mathbf{C})$ dans $C_0(\mathbf{R}^N;\mathbf{C})$ — on établit ce point par un argument de nature topologique sur les espaces de Banach (c'est-à-dire les espaces vectoriels normés complets.)

Même si la théorie de la transformation de Fourier que nous venons de présenter semble fonctionner de façon particulièrement agréable dans le cadre de l'espace de Lebesgue L^1, il faut être conscient que ce cadre est beaucoup trop restrictif du point de vue des applications.

En voici une première indication. Si on revient au cas de la diffraction de Fraunhofer dont nous nous sommes servis pour motiver notre étude, l'image diffractée par la pupille P, à savoir

$$\widehat{\mathbf{1}_P}(\xi) = \int_{\mathbf{R}^2} \mathbf{1}_P(x) e^{-i\xi\cdot x} dx$$

n'est jamais sommable en ξ, car $\mathbf{1}_P$ n'est jamais p.p. égale à une fonction continue tendant vers 0 à l'infini [4] — sauf si elle est p.p. nulle, ce qui correspond au cas trivial où P est négligeable (c'est-à-dire en pratique où il n'y a pas d'ouverture.)

Le théorème d'inversion ci-dessus n'est donc pas le cadre mathématique adéquat pour reconstituer la forme de la pupille P à partir de l'image diffractée.

Cette difficulté sera levée dans la suite du cours consacrée à l'analyse hilbertienne, car on verra que la transformation de Fourier est une bijection linéaire isométrique de l'espace $L^2(\mathbf{R}^N;\mathbf{C})$ des fonctions de carré sommable sur lui-même.

Le lecteur n'aura pas manqué de remarquer l'analogie frappante entre les formules de la théorie des séries de Fourier et celles qui sont relatives à la transformation de Fourier dans L^1. En effet

a) pour toute fonction $f \in L^1(\mathbf{R};\mathbf{C})$ telle que $\hat{f} \in L^1(\mathbf{R};\mathbf{C})$, on définit la transformation de Fourier

$$f \mapsto \hat{f}\,, \quad \text{où } \hat{f}(\xi) = \int_{\mathbf{R}} e^{-i\xi x} f(x) dx$$

et on reconstruit f à partir de $\hat{f}$ par la formule d'inversion

$$f(x) = \tfrac{1}{2\pi} \int_{\mathbf{R}} e^{ix\xi} \hat{f}(\xi) d\xi \quad \text{p.p. en } x \in \mathbf{R}^N\,;$$

b) pour toute fonction $F \in C^1(\mathbf{R};\mathbf{C})$ 2π-périodique, on définit la suite (indexée par $\mathbf{Z}$) de ses coefficients de Fourier

$$F \mapsto (\hat{F}(k))_{k\in\mathbf{Z}}\,, \quad \text{où } \hat{F}(k) = \int_{-\pi}^{\pi} e^{-ikx} F(x) dx\,,$$

4. Le lecteur est invité à rédiger une preuve de cette remarque, à titre d'exercice.

et on reconstruit F à partir de la suite $(\hat{F}(k))_{k\in\mathbf{Z}}$ par la formule

$$F(x) = \tfrac{1}{2\pi}\sum_{k\in\mathbf{Z}} \hat{f}(k)e^{ikx}\,.$$

Cependant, on ne peut pas dire que la transformation de Fourier soit plus générale que les séries de Fourier car toute fonction périodique sommable sur $\mathbf{R}$ est p.p. nulle.

Toutefois, il existe bien un cadre général englobant à la fois la théorie des séries de Fourier et de la transformation de Fourier : c'est la théorie des distributions, qui a emergé vers la fin des années 1940, et qui est due à Laurent Schwartz (après les travaux de plusieurs précurseurs comme Jean Leray ou Sergei Sobolev dans les annees 1930.) Cette théorie dépasse malheureusement le cadre de ce cours, mais sera présentée en deuxième année (cours MAT431).

Terminons ce chapitre en revenant à la formule donnant la transformation de Fourier des densités gaussiennes, dans le cas particulier $N = 1$ (pour simplifier les calculs) à savoir

$$\mathcal{F}G_a = \sqrt{\frac{2\pi}{a}}G_{1/a}\,,$$

où on rappelle que, pour tout $a > 0$,

$$G_a(x) = \frac{1}{\sqrt{2\pi a}}e^{-|x|^2/2a}\,.$$

On sait que G_a est une densité de probabilité sur $\mathbf{R}$, c'est-à-dire que

$$G_a \geq 0 \text{ sur } \mathbf{R}\,, \quad \text{ et } \int_{\mathbf{R}} G_a(x)dx = 1\,,$$

et que la moyenne et la variance de la densité de probabilité G_a sont respectivement 0 et a, c'est-à-dire que

$$\int_{\mathbf{R}} xG_a(x)dx = 0\,, \quad \int_{\mathbf{R}} x^2G_a(x)dx = a\,.$$

(Voir le cours de S. Méléard *"Aléatoire"*.)

Ces formules montrent que plus $a > 0$ est petit, plus la densité de probabilité G_a est concentrée sur sa valeur moyenne, soit 0. La formule donnant $\hat{G}_a$ montre alors que plus $a > 0$ est petit, moins la transformée de Fourier $\hat{G}_a$ est concentrée sur sa valeur moyenne (également nulle.) Autrement dit, plus G_a est localisée et moins $\hat{G}_a$ est localisée, et vice-versa.

Ce comportement remarquable est à rapprocher de la dualité onde-corpuscule en mécanique quantique, et de la représentation de la fonction d'onde d'une particule quantique dans l'espace des positions ou dans l'espace des impulsions. A une fonction $\phi \equiv \phi(x)$ de la variable x de position — typiquement, la fonction d'onde d'une particule — la transformée de Fourier associe la fonction $\hat{\phi} \equiv \hat{\phi}(\xi)$ de la variable d'impulsion ξ — en choisissant un système d'unités convenable dans lequel la constante de Planck vaut 1. On passe donc de la représentation des états quantiques en variable de position à la représentation duale en variable d'impulsion par la transformation de Fourier.

Le fait que la densité gaussienne G_a et sa transformée de Fourier $\hat{G}_a$ soient plus ou moins localisées lorsque $a > 0$ tend vers 0 en décroissant évoque évidemment le principe d'incertitude de Heisenberg. Plus précisément, on vérifie sans peine que

$$\int_{\mathbf{R}} x^2 |G_a(x)|^2 dx \cdot \int_{\mathbf{R}} \xi^2 |\widehat{G_a}(\xi)|^2 \tfrac{d\xi}{2\pi} = \tfrac{1}{4} \left(\int_{\mathbf{R}} |G_a(x)|^2 dx \right)^2 ,$$

alors qu'une formulation mathématique du principe d'incertitude de Heisenberg est l'inégalité

$$\int_{\mathbf{R}} x^2 |\phi(x)|^2 dx \cdot \int_{\mathbf{R}} \xi^2 |\hat{\phi}(\xi)|^2 \tfrac{d\xi}{2\pi} \geq \tfrac{1}{4} \left(\int_{\mathbf{R}} |\phi(x)|^2 dx \right)^2 ,$$

pour toute fonction $\phi \in C^1(\mathbf{R}; \mathbf{C})$ telle que

$$\int_{\mathbf{R}} (x^2 |\phi(x)|^2 + |\phi'(x)|^2) dx < \infty ;$$

— on reviendra sur ce point dans la partie du cours consacrée à l'analyse hilbertienne.

Ainsi les densités gaussiennes présentent la propriété remarquable de réaliser l'égalité dans l'inégalité traduisant le principe d'incertitude de Heisenberg.

5 Ce qu'il faut retenir

Définition de la transformation de Fourier.

A tout $f \in L^1(\mathbf{R}^N; \mathbf{C})$ la transformation de Fourier $\mathcal{F}$ associe la fonction $\hat{f} = \mathcal{F} f$ définie sur $\mathbf{R}^N$ par la formule

$$\hat{f}(\xi) = \int_{\mathbf{R}^N} e^{-i\xi \cdot x} f(x) dx , \quad \xi \in \mathbf{R}^N .$$

Transformée de Fourier des gaussiennes

Pour tout $a > 0$, on appelle densité gaussienne sur $\mathbf{R}^N$ de matrice de covariance aI la fonction

$$G_a(x) = \frac{1}{(2\pi a)^{N/2}} e^{-|x|^2/2a} , \quad x \in \mathbf{R}^N .$$

Alors

$$\widehat{G_a}(\xi) = e^{-a|\xi|^2/2} , \quad \xi \in \mathbf{R}^N .$$

Théorème de Riemann-Lebesgue.

Pour tout $f \in L^1(\mathbf{R}^N; \mathbf{C})$, sa transformée de Fourier $\hat{f} = \mathcal{F} f$ vérifie

a) $\hat{f} \in C(\mathbf{R}^N; \mathbf{C})$ et $|\hat{f}(\xi)| \leq \|f\|_{L^1}$ pour tout $\xi \in \mathbf{R}^N$;

b) $\hat{f}(\xi) \to 0$ lorsque $|\xi| \to +\infty$.

Transformation de Fourier et dérivation

Soit f appartenant à $L^1(\mathbf{R}^N;\mathbf{C})$.

a)Si $|x|f \in L^1(\mathbf{R}^N;\mathbf{C})$, alors $\hat{f} \in C^1(\mathbf{R}^N;\mathbf{C})$ et on a, pour tout $k = 1,\dots,N$

$$\frac{\partial \hat{f}}{\partial \xi_k}(\xi) = \int_{\mathbf{R}^N} -ix_k e^{-i\xi\cdot x} f(x)dx = -i\widehat{x_k f}(\xi)\,.$$

b) Si $f \in C^1(\mathbf{R}^N;\mathbf{C})$ et $\frac{\partial f}{\partial x_k} \in L^1(\mathbf{R}^N;\mathbf{C})$, alors, pour tout $k = 1,\dots,N$

$$\widehat{\frac{\partial f}{\partial x_k}}(\xi) = \int_{\mathbf{R}^N} e^{-i\xi\cdot x}\frac{\partial f}{\partial x_k}(x)dx = i\xi_k \hat{f}(\xi)\,.$$

Théorème d'inversion de Fourier.

Soit $f \in L^1(\mathbf{R}^N;\mathbf{C})$ t.q. $\hat{f} \in L^1(\mathbf{R}^N;\mathbf{C})$. Alors, p.p. en $x \in \mathbf{R}^N$

$$f(x) = \tfrac{1}{(2\pi)^N}\int_{\mathbf{R}^N} e^{+ix\cdot\xi}\hat{f}(\xi)d\xi\,.$$

Chapitre VIII

Espaces de Hilbert, théorie de base



FIGURE VIII.1 – David Hilbert, 1862-1943.

La théorie des espaces de Hilbert est la théorie la plus simple des espaces de dimension infinie. C'est la généralisation des espaces hermitiens (ou euclidiens) de dimension finie. Le cas le plus général, celui des espaces de Banach, c'est-à-dire des espaces vectoriels normés complets – la norme n'étant pas nécessairement induite par un produit scalaire– en est une généralisation. Les démonstrations des théorèmes dans ce cadre plus général sont plus difficiles, mais le cadre des espaces Hilbertiens permet déjà de résoudre de nombreux problèmes concrets. C'est de plus le cadre naturel de la mécanique quantique.

1 Définitions, premières propriétés

Dans ce chapitre tous les espaces vectoriels seront supposés définis sur $\mathbf{C}$. Un espace vectoriel normé est un espace vectoriel muni d'une norme. Rappelons qu'une norme $\|\bullet\|$ est une application de E dans $\mathbf{R}_+$, vérifiant les propriétés suivantes

(1) $\|x\| = 0$ si et seulement si $x = 0$

(2) $\|\lambda x\| = |\lambda| \cdot \|x\|$ pour tous $\lambda \in \mathbf{C}$ et $x \in E$

(3) $\|x + y\| \leq \|x\| + \|y\|$ pour tous $x, y \in E$

Dans la suite on dit qu'une application f entre deux $\mathbf{C}$-espaces vectoriels E et F est anti-linéaire si elle vérifie $f(\lambda x + \mu y) = \bar{\lambda} f(x) + \bar{\mu} f(y)$ quels que soient $(\lambda, \mu) \in \mathbf{C}^2$ et $x, y \in E$. On appelle **forme sesquilinéaire** hermitienne, une application B de $E \times E$ dans $\mathbf{C}$ telle que $x \longrightarrow B(x, y)$ est anti-linéaire et $y \longrightarrow B(x, y)$ est linéaire.

Définition 1.0.1. On appelle forme hermitienne une forme sesquilinéaire, notée $\langle \bullet, \bullet \rangle$ telle que $\langle x, y \rangle = \overline{\langle y, x \rangle}$ (en particulier $\langle x, x \rangle$ est réel).

On dit que la forme hermitienne est définie positive, si $\langle x, x \rangle > 0$ pour $x \neq 0$. On appelle une telle forme **produit hermitien**. Il induit alors une norme sur E en posant $\|x\| = (\langle x, x \rangle)^{1/2}$.

On appelle espace préhilbertien un espace vectoriel muni d'un produit hermitien, et de la norme associée. C'est un espace métrique pour la distance $d(x, y) = \|x - y\|$.

Exemple 1.0.2. Sur $\mathbf{C}^n$, la forme $\langle x, y \rangle = \sum_{j=1}^{n} \bar{x}_j y_j$ définit un produit hermitienne. Sur un espace vectoriel de dimension finie, tout produit hermitien se ramène à cette forme dans une base adaptée (cf. exercice (2) page 186).

Propriétés du produit hermitien :

On a

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + \langle x, y \rangle + \langle y, x \rangle = \|x\|^2 + \|y\|^2 + 2\,\mathrm{Re}(\langle x, y \rangle)$$

d'où on tire les égalités et inégalités suivantes :

(Inégalité de Cauchy-Schwarz)

$$|\langle x, y \rangle| \leq \|x\| \|y\|$$

(Inégalité triangulaire)

$$\|x + y\| \leq \|x\| + \|y\|$$

(Théorème de Pythagore) $\mathrm{Re}\langle x, y\rangle = 0 \Longrightarrow \|x+y\|^2 = \|x\|^2 + \|y\|^2$

(Égalité du parallélogramme)

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$$

(Égalité de la médiane)

$$\|x-a\|^2 + \|x-b\|^2 = 2\left\|x - \frac{a+b}{2}\right\|^2 + 2\left\|\frac{a-b}{2}\right\|^2$$

(Formule de polarisation)

$$\mathrm{Re}\langle x, y\rangle = \frac{\left(\|x+y\|^2 - \|x-y\|^2\right)}{4}$$
$$\mathrm{Im}\langle x, y\rangle = \frac{\left(\|x+iy\|^2 - \|x-iy\|^2\right)}{4}$$

Définition 1.0.3. Soit $(H, \langle,\rangle)$ un espace vectoriel sur $\mathbf{C}$ muni d'une forme hermitienne définie positive (i.e. un espace préhilbertien). On dit que $(H, \langle,\rangle)$ est un espace de Hilbert, si muni de la métrique associée à $\|\bullet\|$, c'est un espace complet.

Comme tout fermé dans un espace complet est complet, tout sous-espace vectoriel **fermé** d'un espace de Hilbert est lui-même un espace de Hilbert muni de la restriction du produit hermitien. On prendra garde à ce que dans un espace de Hilbert, il existe des sous-espaces qui ne sont pas fermés (par exemple des sous-espaces propres et denses, cf. Exemples 2.0.11 page 150).

Remarque 1.0.4. Si E est un espace vectoriel normé sur $\mathbf{C}$ quelconque, on peut se demander si sa norme provient d'une norme Hilbertienne. La réponse est donnée par la **formule de polarisation** : $\|\bullet\|$ est une norme hermitienne si et seulement si

$$(*) \qquad B(x,y) = \frac{\left(\|x+y\|^2 - \|x-y\|^2\right)}{4} - i\frac{\left(\|x+iy\|^2 - \|x-iy\|^2\right)}{4}$$

est une forme hermitienne. : si c'est le cas, on a bien $B(x,x) = \|x\|^2$.

Exemples 1.0.5. (1) Tout espace préhilbertien de dimension finie est un espace de Hilbert, car sur un espace de dimension finie, toutes les normes sont équivalentes, et rendent donc l'espace complet.

(2) L'espace $\ell^2(\mathbf{N})$ des suites $(x_n)_{n\geq 0}$ telle que $\sum_{n=0}^{+\infty} x_n^2 < +\infty$ muni de $\langle x, y\rangle = \sum_{n=0}^{+\infty} \bar{x}_n y_n$. Montrons que cet espace est complet. Si $(\mathbf{x}_k)_{k\geq 1}$ est une suite de Cauchy de suites $\mathbf{x}_k = (x_{n,k})_{n\geq 0}$ de $\ell^2(\mathbf{N})$, chacune des suites des coordonnées $(x_{n,k})_{k\geq 1}$ est de Cauchy (car $|x_{n,k} -$

$x_{n,l}|^2 \leq \|x_k - x_l\|^2$) et donc converge vers un réel z_n. On prétend maintenant que $\mathbf{z} = (z_n)_{n\geq 0}$ est la limite des $(\mathbf{x}_k)_{k\geq 1}$. En effet pour tout n, on a que

$$\sum_{j=0}^{n} |x_{j,k} - z_j|^2 = \lim_{q\to\infty} \sum_{j=0}^{n} |x_{j,k} - x_{j,q}|^2 \leq \sup_{q\geq k} \|\mathbf{x}_k - \mathbf{x}_q\|^2$$

Comme le terme de droite est indépendant de n, on obtient en faisant tendre n vers l'infini $\|\mathbf{x}_k - \mathbf{z}\| \leq \sup_{q\geq k} \|\mathbf{x}_k - \mathbf{x}_q\|^2$ qui tend vers 0 avec k car la suite $(\mathbf{x}_k)_{k\geq 1}$ est de Cauchy. On en déduit que $(\mathbf{x}_k)_{k\geq 1}$ tend vers $\mathbf{z}$.

(3) Pour un réel positif k, on définit l'espace $H^k(\mathbf{N})$ des suites $(x_n)_{n\geq 1}$ telle que $\sum_{n=0}^{+\infty} n^{2k}|x_n|^2 + |x_0|^2 < +\infty$ muni de $\langle x, y\rangle = \sum_{n=1}^{+\infty} n^{2k}\bar{x}_n y_n + \bar{x}_0 y_0$ (démonstration analogue au cas précédent)

(4) L'espace $L^2(\mathbf{R}, \mathbf{C})$ des fonctions de carré intégrable sur $\mathbf{R}$ muni de $\langle f, g\rangle = \int_{\mathbf{R}} \bar{f}(t)g(t)d\mu$ (voir le cours d'intégration (cf. le théorème 4.0.5 page 122).

(5) L'espace $L^2(\mathbf{S}^1, \mathbf{C})$ des fonctions périodiques de période 2π, c'est-à-dire telles que $f(x + 2\pi) = f(x)$ p.p. et telles que $\int_0^{2\pi} f(t)^2 dt < +\infty$, muni de $\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} \bar{f}(t)g(t)d\mu$. Cet espace est canoniquement isomorphe à l'espace $L^2([0, 2\pi], \mathbf{C})$, car toute fonction sur $[0, 2\pi]$ coincide p.p. avec une unique fonction 2π-périodique. Et l'application qui associe à $x \mapsto f(x)$ la fonction $x \mapsto f(2\pi x)$ réalise une isomorphisme entre $L^2([0, 2\pi], \mathbf{C})$ et $L^2([0, 1], \mathbf{C})$.
On utilisera dans la suite indifféremment les espaces $L^2(\mathbf{S}^1, \mathbf{C})$, $L^2([0, 2\pi], \mathbf{C})$, $L^2([0, 1], \mathbf{C})$, sachant que ces espaces s'identifient entre eux de manière naturelle.

(6) L'espace des fonctions holomorphes [1] sur $\mathbf{C}$, telles que en notant $x = \mathrm{Re}(z), y = \mathrm{Im}(z)$ $\iint_{\mathbf{C}} |f(z)|^2 e^{-\frac{|z|^2}{2}} dxdy < \infty$ muni du produit hermitien $\langle f, g\rangle = \iint_{\mathbf{C}} \bar{f}(z)g(z)e^{-\frac{|z|^2}{2}} dxdy = \iint \overline{f}(x + iy)g(x + iy)e^{-\frac{(x^2+y^2)}{2}} dxdy$.

Si E est un espace préhilbertien, on peut le compléter [2] en un espace Hilbertien dont E est un sous-espace dense. On considère l'ensemble Cauchy(E) des suites de Cauchy de E, muni de la relation d'équivalence

$$(x_n)_{n\geq 1} \simeq (y_n)_{n\geq 1} \Longleftrightarrow \lim_{n\to\infty} \|x_n - y_n\| = 0$$

On appelle H l'ensemble des classes d'équivalence [3] déléments de Cauchy(E). Si $(x_n)_{n\geq 1} \in$ Cauchy(E), on pose $[(x_n)_{n\geq 1}] = \{(y_n)_{n\geq 1} \mid (y_n)_{n\geq 1} \simeq (x_n)_{n\geq 1}\}$. C'est clairement un espace vectoriel et on le munit de la forme hermitienne

$$\langle (x_n)_{n\geq 1}, (y_n)_{n\geq 1}\rangle = \lim_{n\to\infty} \langle x_n, y_n\rangle$$

L'inégalité

$$|\langle x_n, y_n\rangle - \langle x_m, y_m\rangle| \leq \|x_n - x_m\| \cdot \|y_n\| + \|x_m\| \cdot \|y_n - y_m\|$$

montre que $\langle x_n, y_n\rangle$ est une suite de Cauchy, et que sa limite ne dépend que de la classe d'équivalence des suites $(x_n)_{n\geq 1}$ et $(y_n)_{n\geq 1}$. On envoie E dans H en associant à x élément

1. c'est-à-dire des fonctions qui sont somme d'une série entière convergeant sur $\mathbf{C}$ tout entier (cf. 3.2.4 page 203 sur les fonctions holomorphes).

2. Cette construction est en fait valable pour n'importe quel espace métrique.

3. Nous sommes en train de construire le quotient d'un ensemble par une relation d'équivalence, comme cela a été fait dans un cadre légèrement différent lors de la construction de l'espace L^1.

de E la suite constante égale à x. On vérifie sans difficulté que c'est une isométrie. La densité de E dans H se démontre comme suit. Soit $\varepsilon > 0$ un réel, et $(x_n)_{n\geq 1}$ une suite de $Cauchy(E)$. On choisit N assez grand pour que $|x_n - x_m| \leq \varepsilon$ pour $m, n \geq N$. Alors la suite constante égale à x_N est à distance au plus ε de cette suite de Cauchy.

On a donc démontré

Proposition 1.0.6. *Tout espace préhilbertien E admet un complété, H dans lequel il est dense. Inversement, si E est un sous-espace vectoriel dense d'un espace de Hilbert, le complété de E est isomorphe à H.*

Remarque 1.0.7. On peut se référer à l'exercice (5) page 186 pour une construction alternative.

Exemple 1.0.8. On considère l'espace des fonctions L^2 sur $\mathbf{R}$ pour la mesure $e^{-\frac{x^2}{2}}dx$, i.e. le complété de l'espace préhilbertien des fonctions continues vérifiant $\int_{\mathbf{R}} |f(x)|^2 e^{-\frac{x^2}{2}}dx < +\infty$ muni de la forme hermitienne

$$\langle f, g\rangle = \int_{\mathbf{R}} \bar{f}(x)g(x)e^{-\frac{x^2}{2}}dx$$

On peut bien entendu remplacer $e^{-\frac{x^2}{2}}$ par n'importe quelle fonction $w(x)$ strictement positive sur $\mathbf{R}$. On note $L^2(\mathbf{R}, w(x)dx)$ cet espace, et on l'appelle espace L^2 à poids $w(x)$.

Soient H_1, H_2 deux espaces de Hilbert munis de $\langle,\rangle_1, \langle,\rangle_2$. Alors $H_1 \oplus H_2 = \{(x_1, x_2) \mid x_1 \in H_1, x_2 \in H_2\}$ muni de $\langle (x_1, x_2), (y_1, y_2)\rangle = \langle x_1, y_1\rangle_1 + \langle x_2, y_2\rangle_2$ est un espace de Hilbert appelé somme directe de H_1 et H_2.

Exercice 1.0.9. Exercice vérifier que $H_1 \oplus H_2$ est complet.

Définition 1.0.10. Soient H_1, H_2 deux espaces préhilbertiens. Une application linéaire $u : H_1 \longrightarrow H_2$ est un isomorphisme d'espaces de Hilbert (ou un isomorphisme isométrique) si et seulement si elle est **bijective**, et pour tout $x \in H_1$ on a

$$\|u(x)\| = \|x\|$$

Notons que la formule de polarisation entraîne que si u est une isométrie, elle préserve le produit hermitien : quels que soient x, y dans H_1, on a

$$\langle u(x), u(y)\rangle = \langle x, y\rangle$$

Notons aussi qu'en dimension infinie, la bijectivité de u ne résulte pas de la conservation de la norme. Par exemple l'application $u : \ell^2(\mathbf{N}) \longrightarrow \ell^2(\mathbf{N})$ donnée par

$$u((x_1, x_2, ..., x_n, ..)) = (0, x_1, x_2, ..., x_{n-1}, ...)$$

préserve la norme mais ce n'est pas une isométrie. C'est aussi un exemple d'injection non surjective. De même, il existe des surjections non injectives, comme l'application

$$u((x_1, x_2, ..., x_n, ..)) = (x_2, x_3, ..., x_{n+1}, ...)$$

2 Sous-espaces fermés, applications continues

Sauf mention contraire, H sera dans la suite un espace de Hilbert.

Soit F un sous-espace vectoriel d'un espace de Hilbert H. Contrairement au cas de la dimension finie, F n'est pas nécessairement fermé.

Exemples 2.0.11. (1) Soit E_0 le sous-espace des suites de $\ell^2(\mathbf{N})$ nulles à partir d'un certain rang. Ce sous-espace est dense, car quel que soit $x = (x_n)_{n\geq 0}$ et quel que soit $\varepsilon > 0$ il existe N tel que $\sum_{n>N} |x_n|^2 < \varepsilon$. Il en résulte que si $\bar{x}_n = x_n$ pour $n \geq N$ et $\bar{x}_n = 0$ pour $n \geq N+1$, on aura $\|x - \bar{x}\| \leq \varepsilon$. Bien entendu il existe des suites de $\ell^2(\mathbf{N})$ qui ont une infinité de termes non nuls, et le sous-espace E_0 est donc distinct de $\ell^2(\mathbf{N})$. On verra une autre démonstration basée sur le théorème de Riesz en 2.2.10 page 155

(2) On sait que l'espace vectoriel $C_c(\mathbf{R}, \mathbf{C})$ des fonctions à support compact de $\mathbf{R}$ est dense dans $L^2(\mathbf{R}, \mathbf{C})$ (voir le cours d'intégration pour L^1, en déduire que c'est encore vrai pour L^2). Comme il existe des fonctions L^2 qui ne sont pas continues, le sous-espace $C_c(\mathbf{R}, \mathbf{C})$ n'est donc pas fermé dans $L^2(\mathbf{R}, \mathbf{C})$.

Remarque 2.0.12. Si F est un sous-espace vectoriel d'un espace de Hilbert H, alors $\overline{F}$, adhérence de F est un sous-espace vectoriel fermé de H.

Définition 2.0.13. On note $F^\perp = \{x \in H \mid \forall y \in F, \langle x, y\rangle = 0\}$ l'orthogonal de F. On note H^* l'ensemble des formes linéaires **continues** sur H que l'on appelle **dual topologique**[4] de H.

Proposition 2.0.14. *(1) Pour tout sous-espace vectoriel F de H, le sous-espace $F^\perp$ est fermé.*

(2) $F^\perp = (\overline{F})^\perp$

Démonstration. En effet si x_n est une suite de H convergeant vers x, vérifiant $\langle x_n, y\rangle = 0$, on a

$$|\langle x, y\rangle| = |\langle x, y\rangle - \langle x_n, y\rangle| \leq \|x_n - x\|\|y\|$$

qui tend vers 0. On en déduit que $\langle x, y\rangle = 0$. Donc si $x_n \in F^\perp$ on aura $x \in F^\perp$ et $F^\perp$ est donc fermé ce qui démontre (1). De manière analogue, si $y \in F^\perp$ et $x_n \in F$ a pour limite $x \in \overline{F}$, on voit que $y \in (\overline{F})^\perp$, d'où $(F)^\perp \subset (\overline{F})^\perp$. Comme l'inclusion inverse[5] résulte de $F \subset \overline{F}$, on en déduit (2). □

De même qu'il existe des sous-espaces qui ne sont pas fermés, il existe des applications linéaires qui ne sont pas continues. Le critère suivant permet de trouver celles qui le sont.

Proposition 2.0.15. *Une application linéaire u entre espaces de Hilbert est continue si et seulement si il existe une constante C telle que $\|u(x)\| \leq C\|x\|$.*

Démonstration. En effet, soit $x \in H$ et $\varepsilon > 0$. Comme

$$\|u(x) - u(y)\| = \|u(x - y)\| \leq C\|x - y\|$$

4. pour le distinguer de l'espace de toutes les formes linéaires. Il arrive lorsque le contexte est clair, que l'on omette l'adjectif « topologique ».

5. Rappelons que $F \subset G$ entraîne que $G^\perp \subset F^\perp$.

l'inégalité $\|y-x\| \leq \varepsilon/C$ entraîne $\|u(x)-u(y)\| \leq \varepsilon$. La propriété énoncée entraîne donc la continuité. Inversement si u est continue, on écrit la continuité en 0 : il existe $\eta > 0$ tel que $\|x\| < \eta \Longrightarrow \|u(x)\| \leq 1$ et par homogénéité,

$$\|u(x)\| = \frac{\|x\|}{\eta}\|u\left(\frac{\eta}{\|x\|}x\right) \leq \frac{1}{\eta}\|x\|$$

□

Remarque 2.0.16. Ce résultat reste vrai pour tout application entre espaces vectoriels normés.

Si E, F sont des espaces de Hilbert, on note $\mathcal{L}(E,F)$ l'espace des applications linéaires continues de E dans F, et E^* l'espace $\mathcal{L}(E,\mathbf{C})$ des fonctions linéaires continues de E dans $\mathbf{C}$.

L'espace E^* est un espace de Hilbert, pour la norme

$$\|u\| = \sup_{\|x\|\leq 1} |u(x)|$$

Que ce soit un espace vectoriel normé est clair. Que ce soit un espace de Hilbert résultera du théorème de Riesz (théorème 2.2.1 page 153). Mais en général, si F est de dimension infinine, l'espace $\mathcal{L}(E,F)$ muni de la norme $\|u\| = \sup_{\|x\|\leq 1} \|u(x)\|$ n'est pas un espace de Hilbert. Bien entendu, à partir d'une application continue, u, on construit facilement des espaces fermés : les $\ker(u-\lambda I)$ par exemple. Mais attention, même si u est continue, $\mathrm{Im}(u)$ n'est pas nécessairement fermé.

Exemples 2.0.17.

(1) Si g est une fonction continue bornée, l'application M_g de $L^2([0,1],\mathbf{C})$ dans lui-même, qui envoie $f(t)$ sur $g(t)f(t)$ est continue. En effet, $\|M_g f\| \leq \sup_{t\in[0,1]} |g(t)|\|f\|$. On laisse le lecteur vérifier qu'il suffit que g soit mesurable et appartienne à $L^\infty([0,1],\mathbf{C})$ pour que M_g soit continue. Si c'est le cas sa norme est $\|g\|_\infty$.

(2) L'application $I : L^2([0,1],\mathbf{C}) \longrightarrow L^2([0,1],\mathbf{C})$ définie par $I(f)(t) = \int_0^t f(s)ds$ est continue. En effet, soit χ_t la fonction indicatrice de $[0,T]$. Alors $I(f)(t) = \int_0^1 \chi_t(s)f(s)ds$ et donc $|I(f)(t)| \leq \int_0^1 |f(s)|ds$. On peut écrire

$$\|I(f)\|_{L^2}^2 = \int_0^1 |I(f)(t)|^2 dt \leq \int_0^1 \left(\int_0^1 |f(s)|ds\right)^2 dt \leq$$
$$\int_0^1 |f(s)|^2 ds = \|f\|_{L^2}^2$$

I est donc continue.

(3) Soit K une fonction de $L^2([0,1]^2,\mathbf{C})$ et $F_K : L^2([0,1]) \longrightarrow L^2([0,1]^2)$ l'application $F_K(f)(s,t) = \int_0^1 K(s,t)f(s)$. Cette application est continue, car si on pose $k(t) = \int_0^1 |K(s,t)|^2 ds$, on a d'après le théorème de Fubini (2.0.1 page 78) que $\int_0^1 k(t)dt = \int_{[0,1]^2} |K(s,t)|^2 dsdt$ est fini et donc

$$\|F_K(f)\|^2 \leq \int_0^1 \left(\int_0^1 |K(s,t)||f(s)|ds\right)^2 dt \leq$$
$$\int_0^1 k(t)\|f\|^2 dt = \left(\int_{[0,1]^2} |K(s,t)|^2 dt\right)\|f\|^2$$

Notons que l'exemple précédent est un cas particulier de celui-ci car $K(s,t) = \chi_t(s)$ est bien dans $L^2([0,1]^2,\mathbf{C})$.

(4) Soit $u : \ell^2(\mathbf{N}^*) \longrightarrow \ell^2(\mathbf{N}^*)$ donnée par $u((x_n)_{n\geq 1}) = (\frac{x_n}{n})_{n\geq 1}$. La norme de $u(x)$ est clairement inférieure à celle de x. C'est donc une application continue. L'image de u contient les suites nulles à partir d'un certain rang, donc elle est dense. Mais elle n'est pas fermée : la suite $(\frac{1}{n\sqrt{n}})_{n\geq 1}$ est dans $\ell^2(\mathbf{N}^*)$ mais n'est pas dans l'image de u.

2.1 Théorèmes fondamentaux : théorème de la projection, théorèmes de représentations.

Commençons par

Théorème 2.1.1 (Théorème de la projection)**.** *Soit E un espace de Hilbert et F un sous-espace fermé de E. Il existe une unique application $P_F : E \longrightarrow F$, telle que*

$$\|x - P_F(x)\| = d(x, F) = \inf_{y\in F} \|x - y\|$$

De plus $P_F(x)$ est l'unique élément de F vérifiant cette égalité et les vecteurs $P_F(x)$ et $(x-P_F(x))$ sont orthogonaux. Enfin P_F est 1-Lipschitzienne (donc **continue***) :*

$$\|P_F(x) - P_F(y)\| \leq \|x - y\|$$

Démonstration. Soit une suite $y_n \in F$ telle que $\lim_n \|x - y_n\| = d(x, F)$. Montrons que $(y_n)_{n\geq 1}$ est une suite de Cauchy dans F. Pour tout $\varepsilon > 0$ il existe $N(\varepsilon)$ tel que pour $n \geq N(\varepsilon)$ on aura $\|x-y_n\| \leq d(x, F)+\varepsilon$. Nous allons utiliser l'égalité du parallélogramme et le fait que $\frac{y_m+y_n}{2} \in F$ et donc $\|x - \frac{y_m+y_n}{2}\| \geq d(x, F)$: si $m, n \geq N(\varepsilon)$ on a

$$\begin{aligned}\|y_n - y_m\|^2 &= 2(\|y_n - x\|^2 + \|y_m - x\|^2) - 4(\|x - \frac{y_n + y_m}{2}\|^2)\\ &\leq 4(d(x, F) + \varepsilon)^2 - 4d(x, F)^2\end{aligned}$$

Par ailleurs, si on avait deux points y_1, y_2 réalisant le minimum de la distance, l'égalité du parallélogramme nous dit que

$$\begin{aligned}\|y_1 - y_2\|^2 &= 2(\|y_1 - x\|^2 + \|y_2 - x\|^2) - 4(\|x - \frac{y_1 + y_2}{2}\|^2)\\ &\leq 4d(x, F)^2 - 4d(x, F)^2 = 0\end{aligned}$$

Montrons que $x - P_F(x) \in F^\perp$. Mais si $y \in F$, $u = P_F(x)$ et $t \in \mathbf{C}$ on a

$$\|x - u + ty\|^2 = \|x - u\|^2 + 2\,\mathrm{Re}(t\langle x - u, y\rangle) + |t|^2\|y\|^2$$

Si $\langle x - u, y\rangle \neq 0$ en choisissant t petit en module et d'argument égal à celui de $\overline{\langle x - u, y\rangle}$, on trouve un élément $u + ty$ de F tel que $\|x - u + ty\|^2 < \|x - u\|^2$, ce qui contredit la minimalité de $\|x - u\|^2$. On a donc $\langle x - u, y\rangle = 0$ pour tout y dans F, soit $x - P_F(x) \in F^\perp$.

Reste enfin à vérifier que P_F est 1-Lipschitizienne. Or cela résulte immédiatement du théorème de Pythagore

$$\|x - y\|^2 = \|(x - P_F(x)) - (y - P_F(y))\|^2 + \|P_F(x) - P_F(y)\|^2$$

□

L'exemple le plus élémentaire de sous-espace fermé est celui de sous-espace de dimension finie. Si F est engendré par les vecteurs $e_1, ..., e_n$ formant une famille orthonormée, c'est-à-dire que $\langle e_i, e_j \rangle = \delta_i^j$, la projection P_F peut s'écrire

$$P_F(x) = \sum_{j=1}^{n} \langle x, e_j \rangle e_j$$

Il suffit pour cela de vérifier que $x - P_F(x) \in F^\perp$ ou encore que $x - P_F(x)$ est orthogonal à tous les e_j, ce qui résulte du calcul suivant

$$\langle x - P_F(x), e_j \rangle = \langle x, e_j \rangle - \langle x, e_j \rangle = 0$$

Corollaire 2.1.2. *Si F est fermé, on a la décomposition $H = F \oplus F^\perp$. Par conséquent si F est fermé, $(F^\perp)^\perp = F$. Plus généralement si F est un sous-espace de H, non nécessairement fermé, $(F^\perp)^\perp = \overline{F}$.*

Démonstration. Le premier énoncé n'est qu'une reformulation du théorème de la projection : $x = P_F(x) + (x - P_F(x)$ avec $x - P_F(x) \in F^\perp$ et de ce que $F \cap F^\perp = \{0\}$. Le second énoncé résulte de ce que par définition $F \subset (F^\perp)^\perp$. Si l'inclusion était stricte, il existerait en vertu de la décomposition démontrée, un élement x non nul de $(F^\perp)^\perp$ contenu dans $F^\perp$. Mais nécessairement $\langle x, x \rangle = 0$ d'où $x = 0$. Enfin, dans le cas général, on a vu dans la proposition 2.0.14 page 150 que $F^\perp = (\overline{F})^\perp$, donc $(F^\perp)^\perp = ((\overline{F})^\perp)^\perp = \bar{F}$. □

Il résulte du théorème précédent que $F \subset E$ fermé si et seulement si il existe un espace vectoriel normé G et une application linéaire continue $u : E \longrightarrow G$ telle que $F = \ker(u)$: la suffisance est évidente, la nécessité s'obtient en prenant $G = E$ et $u = \mathrm{Id} - p_F$.

2.2 Théorèmes de représentation, théorème de Hahn-Banach et critère de densité

Tirons les conséquences des résultats précédents, en commençant par le théorème de représentation de Riesz.

Théorème 2.2.1 (Théorème de représentation de Riesz)**.** *Soit u une forme linéaire* **continue** *sur un espace de Hilbert. Il existe un unique vecteur $a_u \in E$ tel que*

$$\forall x \in E \ , \ u(x) = \langle a_u, x \rangle$$

L'application $a \longrightarrow \langle a, \bullet \rangle$ qui envoie H sur H^* est donc bijective. C'est un isomorphisme isométrique.

Démonstration. Comme u est continue, l'espace $F = \ker(u)$ est fermé, donc $H = F \oplus F^\perp$. De plus $F^\perp$ est de dimension un, car si b_u est unitaire dans $F^\perp$ et x dans $F^\perp$, on a $u(x - u(x)b_u) = 0$ donc $x - u(x)b_u$ est dans $F \cap F^\perp$ donc $x = u(x)b_u$. Maintenant, si $a_u = u(b_u)b_u$, on a $\langle a_u, x \rangle = u(x)$. En effet, si $x \in F$ les deux termes sont nuls, si $x = \lambda b_u$ les deux termes valent $\lambda u(b_u)$. Par linéarité les deux termes sont égaux sur $F \oplus F^\perp = H$.

Nous savons maintenant que l'application $a \longrightarrow \langle a, \bullet \rangle$ est surjective. Elle est évidemment injective, car si $a \neq 0$ $\langle a, \bullet \rangle$ est non nulle sur a.

Vérifions que cette application est une isométrie. La norme que nous avons défini sur H^* est $\|u\| = \sup_{\|x\| \leq 1} |u(x)|$. Si $u(x) = \langle a, x \rangle$, on obtient

$$\|u\| = \sup_{\|x\| \leq 1} |\langle a, x \rangle| \leq \sup_{\|x\| \leq 1} \|a\| \cdot \|x\|$$

Comme pour $a \neq 0$ l'égalité est atteinte si $x = \frac{a}{\|a\|}$, on voit que $\|u\| = \|a\|$ et l'application ci-dessus est une isométrie. □

Il résulte en particulier du théorème de représentation de Riesz que H^* est un espace de Hilbert, ce qui n'était pas évident *a priori*. Le produit hermitien sur H^* est défini en posant $\langle\langle u, v \rangle\rangle = \langle a_u, a_v \rangle$. Notons aussi que la notation des physiciens en **bra** et **kets** correspons exactement à cette identification de H et H^* : le produit hermitien étant noté $\langle x \mid y \rangle$ le bra $\langle x \mid$ correspond à l'élément de H^* que nous avons noté $\langle x, \bullet \rangle$ et $\mid y \rangle$ au vecteur ordinaire x. Si u est une application linéaire, $\langle x \mid u \mid y \rangle$ correspond donc à $\langle x, u(y) \rangle$.

L'extension suivante du théorème de Riesz aux formes sesquilinéaires sera utile pour l'étude des équations de Sturm-Liouville.

Définition 2.2.2 (Critère de continuité pour les formes sesquilinéaires). On dit qu'une forme sesquilinéaire $B(u, v)$ est continue sur H si il existe une constante C telle que

$$\forall u, v \in H \quad |B(u, v)| \leq C\|u\|\|v\|$$

Remarque 2.2.3. Il résulte aisément de la formule de polarisation que si B est hermitienne, il suffit de vérifier que $B(u, u) \leq C\|u\|^2$ pour conclure à la continuité de B.

On a alors

Théorème 2.2.4 (Théorème de représentation de Riesz (version sequilinéaire))**.** *Soit B une forme sesquilinéaire continue. Il existe alors une application linéaire unique $L : H \longrightarrow H$, continue, telle que*

$$\forall u, v \in H, \quad B(u, v) = \langle Lu, v \rangle$$

Démonstration. D'après l' hypothèse de continuité, pour chaque u fixé la forme linéaire $v \longrightarrow B(u, v)$ est continue. Par le théorème de Riesz, il existe Lu unique tel que $B(u, v) = \langle Lu, v \rangle$. Il nous faut montrer que L est linéaire et continue. La linéarité résulte immédiatement de l'unicité et de la sesquilinéarité de B. Pour la continuité, nous avons par hypothèse

$$\|Lu\|^2 = \langle Lu, Lu \rangle = B(u, Lu) \leq C\|u\|\|Lu\|$$

On en déduit

$$\|Lu\| \leq C\|u\|$$

et donc la continuité de L. □

Exercices 2.2.5. (1) Montrer que s'il existe une constante C_0 strictement positive telle que B vérifie pour tout u l'inégalité $B(u, u) \geq C_0\|u\|^2$ alors l'application L du théorème précédent est inversible et d'inverse continu.

(2) Montrer que si B est hermitienne, alors $L^* = L$ (on dit que L est hermitienne).

Une conséquence de ce théorème est la continuité de l'adjoint.

Définition 2.2.6. Soit $u : E \longrightarrow F$ un opérateur continu. On appelle adjoint de u et note u^*,un opérateur vérifiant

$$\forall (x, y) \in E \times F \;\; \langle u(x), y\rangle = \langle x, u^* y\rangle$$

Notons que si u possède un adjoint, il résulte de l'injectivité de l'application $a \longrightarrow \langle a, \bullet\rangle$ que cet adjoint est unique.

Corollaire 2.2.7. *Si u est un opérateur continu $u : E \longrightarrow F$ entre espaces de Hilbert, son adjoint est bien défini et c'est encore un opérateur continu.*

Démonstration. En effet, la forme sesquilinéaire $(x, y) \longrightarrow \langle x, u(y)\rangle$ est continue, vu que

$$|\langle x, u(y)\rangle| \leq \|x\|\|u(y)\| \leq \|x\|\|u\|\|y\|$$

Il existe donc, d'après le théorème de Riesz pour les formes sesquiinéaires un unique opérateur u^*, continu, tel que $\langle u^*(x), y\rangle = \langle x, u(y)\rangle$. □

Passons maintenant au théorème de Hahn-Banach, qui nous donnera un critère de densité bien utile dans la section suivante.

Théorème 2.2.8 (Hahn-Banach dans un Hilbert)**.** *Soit F un sous espace d'un espace de Hilbert. Si $x_0 \notin \overline{F}$ alors il existe une forme linéaire continue u telle que $u(F) = 0$ et $u(x_0) = 1$*

Démonstration. Soit $W = F^\perp = (\overline{F})^\perp$. On a $x_0 = p_W(x_0) + (x - p_W(x))$ où $x - p_W(x) \in W^\perp = \overline{F}$. Comme $p_W(x_0) \neq 0$ (sinon $x_0 \in \overline{F}$), on peut définir $a = \frac{p_W(x_0)}{|p_W(x_0)|^2} \in F^\perp = \overline{F}^\perp$. Alors $\langle a, x_0\rangle = 1$ et $\langle a, x\rangle = 0$ pour $x \in \overline{F}$. □

Le résultat suivant a de nombreuses applications à des démonstrations de densité, et en particulier sera très utile pour montrer qu'une suite de vecteurs forme une base Hilbertienne.

Corollaire 2.2.9 (Critère de densité)**.** *Soit $F \subset H$ un sous-espace vectoriel. Alors F est dense si et seulement si il n'existe pas d'élément $a \neq 0$ dans H tel que $\forall x \in F, \langle a, x\rangle = 0$, c'est-à -dire si $F^\perp = \{0\}$.*

Exemple 2.2.10. Soit E_0 l'ensemble des suites de $\ell^2(\mathbf{N})$ nulles à partir d'un certain rang, et e_n la suite dont tous les termes sont nuls sauf le n-ième égal à 1. Si $a = (a_n)_{n\geq 0}$ est orthogonal à E_0, on a $\langle a, e_n\rangle = a_n = 0$, donc $a = 0$. On en conclut que E_0 est dense dans $\ell^2(\mathbf{N})$.

3 Bases hilbertiennes

Soit $(e_n)_{n\geq 1}$ une famille dénombrable de vecteurs d'un espace de Hilbert E.

Définition 3.0.11. On dit que $(e_n)_{n\geq 1}$ est une base Hilbertienne et on note $H = \oplus_n \mathbf{C} e_n$ si

(1) $\forall i, j, i \neq j, e_i \perp e_j$ et $\|e_j\| = 1$

(2) l'ensemble $\{\sum_{j=1}^{q} x_j e_j\}$ des combinaisons linéaires finies de e_j est dense dans E

Remarque 3.0.12. Attention, une « base Hilbertienne » n'est pas une base (on dit parfois « base algébrique » pour éviter toute confusion) car pour une base algébrique, tout élément est combinaison linéaire **finie** d'éléments de la base. Dans la suite, sauf mention contraire, on réservera la mot de **base** pour des bases Hilbertiennes.

Théorème 3.0.13. *Si $(e_n)_{n\geq 1}$ est une base Hilbertienne de H, tout élément x de H s'écrit de manière unique comme série convergente dans H $x = \sum_n x_n e_n$ avec $x_n = \langle x, e_n\rangle \in \mathbf{C}$. Réciproquement toute somme $x = \sum_n x_n e_n$ telle que $\sum_n |x_n|^2 < +\infty$ converge dans H. De plus*

(égalité de Parseval) $$\|x\|^2 = \sum_n |x_n|^2 = \sum_n |\langle x, e_n\rangle|^2$$

Démonstration. Soit G le sous-espace de H constitué des éléments qui s'écrivent comme sommes finies des e_n. À la fois la décomposition et l'égalité de Parseval sont vraies pour les éléments de G, puisqu'elles le sont dans un espace, somme finie des $\mathbf{C}e_n$, et que tout élément de G est contenu dans un tel sous-espace. Montrons alors l'existence de la décomposition. Soit $G_n = \oplus_{j=1}^{n} \mathbf{C}e_n$ et P_n la projection orthogonale sur G_n. On affirme que $\lim_n P_n(x) = x$, puisque $P_n(x)$ est le point de G_n le plus proche de x et que $G = \bigcup_n G_n$ est dense dans H. On écrit alors $x_1 e_1 = P_1(x)$ et $x_n e_n = P_n(x) - P_{n-1}(x) \in \mathbf{C}e_n$, d'où $P_n(x) = \sum_{j=1}^{n} x_j e_j$ et la série converge donc vers x. Comme les x_j sont deux à deux orthogonaux, on a $\|P_n(x)\|^2 = \sum_{j=1}^{n} |x_j|^2$ par le théorème de Pythagore. Comme la norme d'une limite est la limite de la norme, on obtient la formule de Parseval

$$\|x\|^2 = \sum_{j=1}^{+\infty} |x_j|^2$$

Il suffit de vérifier l'unicité pour $x = 0$. Mais si $0 = \lim_n \sum_{j=1}^{n} x_k$ on a $0 = \langle x_k, 0\rangle = \|x_k\|^2$, avec les x_k deux à deux orthogonaux, les x_k sont donc tous nuls. La formule de décomposition pour une base Hilbertienne résulte de ce que $w_n = \sum_{j=1}^{n} \langle x, e_j\rangle e_j \in G_n$ et $x - w_n \perp G_n$. Enfin, pour montrer qu'une série $\sum_n x_n e_n$ converge dans H si $\sum_n |x_n|^2 < +\infty$, il suffit de remarquer que $\| \sum_{q=m}^{n} x_q e_q\|^2 \leq \sum_{q=m}^{n} |x_q|^2$ d'après le théorème de Pythagore, et donc que la suite $\sum_{j=1}^{n} |x_n|^2$ est convergente, donc de Cauchy, et il s'ensuit que la suite des sommes partielles $s_n = \sum_{j=1}^{n} x_j e_j$ est de Cauchy.

□

Remarque 3.0.14. Notons que la série $\sum_n x_n e_n$ n'est pas en général absolument convergente : $\sum_n |x_n|^2$ converge, mais il n'y a pas de raison que $\sum_n |x_n|$ converge.

3.1 Orthonormalisation de Schmidt, familles totales, inégalité de Bessel

Le procédé **d'orthonormalisation de Schmidt** qui permet – en dimension finie– à partir d'une famille de vecteurs engendrant l'espace de construire une base, s'applique aussi dans un espace de Hilbert. Plus précisément soit $(f_n)_{n\geq 1}$ une famille de vecteurs que l'on peut supposer, quitte à en oublier éventuellement certains, libre. On note F_n l'espace engendré par $f_1, ..., f_n$, et

on suppose avoir construit une base orthonormée de F_n, $(e_1, ..; , e_n)$. On prend alors pour e_{n+1} un vecteur de norme un dans $F_n^\perp \cap F_{n+1}$ qui est un espace de dimension $\dim(F_{n+1}) - \dim(F_n) = 1$. La famille (e_n) est donc telle que toute somme finie des f_n s'écrit comme somme finie des e_n. Si le sous-espace vectoriel $\bigcup_n F_n$ est dense, on aura la densité des combinaisons linéaires des e_n, et donc $(e_n)_{n\geq 1}$ forme une base Hilbertienne.

Soit alors $(e_n)_{n\geq 1}$ un ensemble dénombrable de vecteurs de H. On dit que c'est un **famille totale** si l'espace engendré par les e_n est dense dans H. Un système orthonormé forme une base Hilbertienne si et seulement si il forme une famille totale, ou encore si et seulement si la formule de Parseval est vérifiée. En effet on a le lemme suivant

Lemme 3.1.1. *Soit $(e_n)_{n\geq 1}$ une famille orthonormale de vecteurs de H. Alors $\sum_{n=1}^{+\infty}\langle x, e_n\rangle e_n$ est exactement la projection orthogonale de x sur la fermeture F, de l'espace engendré par les e_n.*

Démonstration. Il nous faut d'abord vérifier que $\sum_{n=1}^{+\infty}\langle x, e_n\rangle e_n$ définit bien un élément de H, c'est-à -dire vérifier la convergence de la série $\sum_{n=1}^{+\infty}|\langle x, e_n\rangle|^2$. Or si F_n est l'espace engendré par $e_1, ..., e_n$ nous avons vu que la projection P_n sur F_n est bien donnée par $\sum_{j=1}^{n}\langle x, e_j\rangle e_j$ et donc $\|P_n(x)\| \leq \|x\|$ d'où

$$\sum_{j=1}^{n}|\langle x, e_j\rangle|^2 \leq \|x\|^2$$

et donc la série $\sum_n |\langle x, e_n\rangle|^2$ converge. L'élément $y = \sum_{n=1}^{+\infty}\langle x, e_n\rangle e_n$ appartient à $\overline{F}$, et $\langle x - y, e_j\rangle = 0$ pour tout j, donc par unicité de la décomposition, $y = P_{\overline{F}}(x)$. □

Il en résulte que

$$\|x\|^2 = \|x - P_{\overline{F}}(x)\|^2 + \|P_{\overline{F}}(x)\|^2 =$$
$$\|x - P_{\overline{F}}(x)\|^2 + \|\sum_n \langle x, e_n\rangle e_n\|^2 = \|x - P_{\overline{F}}(x)\|^2 + \sum_n |\langle x, e_n\rangle|^2$$

On en déduit d'une part l'inégalité de Bessel

$$\sum_n |\langle x, e_n\rangle|^2 \leq \|x\|^2$$

et que l'égalité de Parseval équivaut à $x = P_{\overline{F}}(x)$ quel que soit x et donc que F est dense dans H ou encore que la familel $(e_n)_{n\geq 1}$ forme une base Hilbertienne.

Enfin, notons qu'une famille orthonormée de vecteurs $(f_n)_{n\geq 1}$, peut toujours se compléter en une base Hilbertienne. Soit en effet F l'adhérence de l'espace engendré par les f_n, c'est un sous-espace fermé. D'après le Corollaire 2.1.2 page 153 on a une décomposition $H = F \oplus F^\perp$. Prenant une base Hilbertienne $(g_m)_{m\geq 1}$ de $F^\perp$ en la juxtaposant à $(f_n)_{n\geq 1}$ on obtient [6] une base Hilbertienne de H.

Exemples 3.1.2. (1) (Séries de Fourier dans $L^2(\mathbf{S}^1, \mathbf{C})$) Soit $E = L^2(\mathbf{S}^1, \mathbf{C})$ muni de la forme hermitienne $\langle f, g\rangle = \frac{1}{2\pi}\int_{\mathbf{S}^1}\overline{f}(t)g(t)dt$. La famille $(e_n)_{n\in\mathbf{Z}}$ où $e_n(t) = e^{int}$, est une famille orthonormée : $\langle e_n, e_m\rangle = \delta_{m,n}$. Dans ce cas, $\langle f, e_n\rangle = \frac{1}{2\pi}\int_{\mathbf{S}^1} f(t)e^{-int}dt = c_n(f)$ est le

6. par exemple on pose $e_{2n} = f_n, e_{2n-1} = g_n$.

n-ième coefficient de Fourier. On montre (cf. appendice 4 page 167) que la fermeture de l'espace engendré par les e_p est dense dans $L^2(\mathbf{S}^1, \mathbf{C})$. Il en résulte que la série de Fourier de f converge pour la norme L^2 vers f, et les coefficients vérifient l'égalité de Parseval

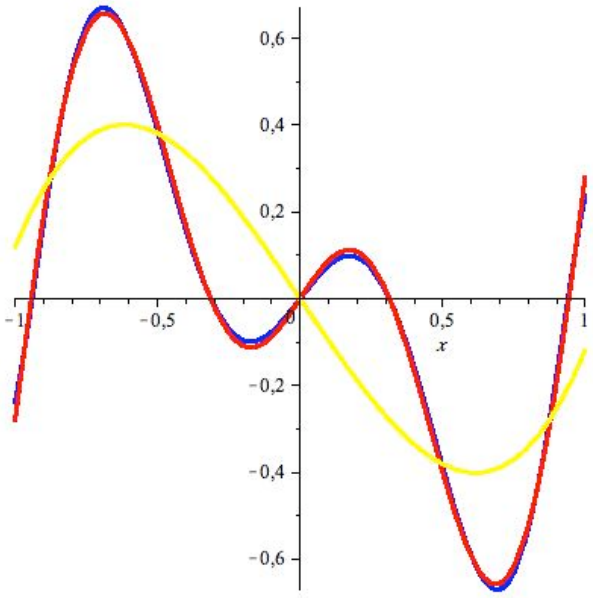
$$\sum_{j=-\infty}^{+\infty} |c_j(f)|^2 = \frac{1}{2\pi} \int_{\mathbf{S}^1} |f(t)|^2$$

Inversement si $(a_n)_{n\in\mathbf{Z}}$ est une suite telle que $\sum_{n\in\mathbf{Z}} |a_n|^2 < +\infty$, la série de Fourier $\sum_n a_n e^{int}$ converge dans $L^2(\mathbf{S}^1, \mathbf{C})$.

(2) Soit w une fonction continue sur un intervalle borné $[a, b]$ et $H = L^2([a, b], w(x)dx)$ l'espace des fonctions L^2 à valeurs complexes pour la mesure $w(x)dx$, c'est à dire que le produit hermitien est donné par $\langle f, g\rangle = \int_a^b \bar{f}(t)g(t)w(t)dt$. Le théorème de Stone-Weierstrass (cf. exercice 4.0.9 page 168) montre que les polynômes forment un sous espace dense de $C^0([a, b], \mathbf{C})$ pour la norme C^0 qui est plus forte que la norme L^2. Comme d'après le cours d'intégration (théorème 3.2.1 page 119) le sous-espace $C^0([a, b], \mathbf{C})$ est dense dans L^2 (on l'a montré pour $w \equiv 1$, le cas général est laissé en exercice). On en conclut que les polynômes forment un sous-espace dense de $L^2([a, b], w(x)dx)$. L'orthonormalisation de Schmidt de la base (non orthogonale) $(1, t, t^2, ..., t^n, ...)$, fournit des familles de polynômes remarquables. Par exemple si $w = 1$ et $a = -1, b = 1$ on obtient les polynômes de Legendre.

(3) Si dans l'exemple ci-dessus on prend $a = -\infty, b = +\infty$ et $w(x) = e^{-x^2}$, la densité des polynômes est encore vraie dans $L^2(\mathbf{R}, e^{-x^2}dx)$. Le procédé d'orthonormalisation de Schmidt fournit alors les polynômes d'Hermite ; si $w(x) = e^{-x}$ $a = 0, b = +\infty$ les polynômes de Laguerre [7]. Comme dans le cas de Fourier, toute fonction de $L^2([-1, 1], \mathbf{C})$ a une unique décomposition en somme de polynômes de Legendre, de même pour une fonction de $L^2([0, +\infty[, e^{-x}dx)$ en somme de polynômes de Laguerre ou une fonction de $L^2(\mathbf{R}, e^{-x^2}dx)$ en somme de polynômes d'Hermite.

7. Nous n'avons pas montré que les polynômes forment un sous espace dense de $L^2(\mathbf{R}, e^{-x^2})$ (cf. exercice (11) page 187), densité qui serait fausse sur $L^2(\mathbf{R}, \mathbf{C})$ (exercice).



Décomposition de $cos(5x)\exp(-x)$ (en rouge) en polynômes de Legendre (jaune $n = 5$, bleu $n = 9$) .

Les ondelettes de Haar

On note $\varphi = \chi_{[0,1/2[} - \chi_{[1/2,1[}$ et $\varphi_{n,k}(x) = 2^{n/2}\varphi(2^n x - k)$ où χ_I désigne la fonction indicatrice de l'intervalle I et $n, k \in \mathbf{N}$. On a donc $\varphi_{1,0} = \varphi$ et pose $\varphi_{0,0} = \chi_{[0,1]}$. Nous allons montrer que la fonction constante $1 = \varphi_{0,0}$ et les $\varphi_{n,k}$ pour $0 \leq k < 2^n$, forment une base Hilbertienne de $L^2([0,1], \mathbf{C})$.

Tout d'abord, il est facile de voir que $\langle \varphi_{n,k}, \varphi_{p,j} \rangle = 0$ si $(n,k) \neq (m,j)$ et vaut 1 sinon. En effet il est clair que $\varphi_{n,k}$ est constante sur les intervalles de le forme $[\frac{r}{2^n}, \frac{r+1}{2^n}]$ avec r entier, et n'est non nulle que sur un des intervalles $[\frac{r}{2^{n-1}}, \frac{r+1}{2^{n-1}}]$ (celui donné par $r = k$). Alors si $m > n$ on remarque que $\varphi_{p,j}$ est nulle en dehors de l'intervalle $[\frac{j}{2^{p-1}}, \frac{j+1}{2^{p-1}}]$ sur lequel $\varphi_{n,k}$ est constante. Donc,

$$\int_0^1 \varphi_{n,k}(t)\varphi_{p,j}(t)dt$$

s'annule, car cela revient à intégrer $\varphi_{p,j}$ sur un intervalle hors duquel elle est nulle. Comme les $\varphi_{p,j}$ sont d'intégrale nulle sur $[0,1]$, cette intégrale est nulle. Rest à traiter le cas où $m = n$. mais alors si $j \neq k$, les supports de $\varphi_{n,j}$ et $\varphi_{n,k}$ sont disjoints et l'intégrale de leur prduit est nulle.

Maintenant d'après le corollaire 2.2.9 page 155, il suffit de montrer que si $f \in L^2([0,1], \mathbf{C})$ est orthogonale à toutes les fonctions précédentes, elle est nulle p.p. Nous allons montrer par récurrence sur n que

$$\int_{\frac{k}{2^n}}^{\frac{(k+1)}{2^n}} f(x)dx = 0$$

pour $0 \leq k < 2^n$. Comme $\langle f, 1 \rangle = 0$, on a $\int_0^1 f(t)dt = 0$.

Montrons par récurrence sur m que si E_m est l'espace engendré par les $\varphi_{n,k}$ pour $0 \leq k < 2^n \leq 2^m$, alors E_m contient les $\chi_{[\frac{j}{2^m},\frac{j+1}{2^m}[}$ pour $0 \leq j < 2^m$.

En effet, $1 - \varphi_{1,0} = \chi_{[1/2,1[}$ et $1 + \varphi_{1,0} = \chi_{[0,1/2[}$, donc E_1 contient $\chi_{[0,1[}, \chi_{[0,1/2[}, \chi_{[1/2,1[}$. Supposons l'assertion vraie à l'ordre m et vérifions la à l'ordre $m+1$. On vérifie sans peine les quatre formules suivantes

$$\begin{cases} \varphi_{n+1,k}(x) = \sqrt{2}\varphi_{n,k}(2x) \\ \varphi_{n+1,k+2^n}(x) = \sqrt{2}\varphi_{n,k}(2x-1) \end{cases}$$

et pour $0 \leq j < 2^m$

$$\begin{cases} \chi_{[\frac{j}{2^{m+1}},\frac{j+1}{2^{m+1}}[} = \chi_{[\frac{j}{2^m},\frac{j+1}{2^m}[}(2x) \\ \chi_{[\frac{j+2^m}{2^{m+1}},\frac{j+2^m+1}{2^{m+1}}[} = \chi_{[\frac{j}{2^m},\frac{j+1}{2^m}[}(2x-1) \end{cases}$$

On voit alors que pour $0 \leq l < 2^{m+1}$ si $l < 2^m$, on écrit $\chi_{[\frac{l}{2^m},\frac{l+1}{2^m}[}$ comme combinaison linéaire des $\varphi_{p,k}$ pour $0 \leq k < 2^p \leq 2^m$ et le changement de x en $2x$ remplace $\chi_{[\frac{l}{2^m},\frac{l+1}{2^m}[}$ par $\chi_{[\frac{l}{2^m},\frac{l+1}{2^m}[}$ et les $\varphi_{p,k}$ par les $\varphi_{p+1,k}$. Si $2^m \leq l \leq 2^{m+1}$, on écrit $l = j + 2^m$ avec $0 \leq j < 2^m$ et le changement de x en $2x-1$ fournit l'écriture de $\chi_{[\frac{l}{2^m},\frac{l+1}{2^m}[}$ en fonction des $\varphi_{p+1,k+2^p}$.

Maintenant si f est orthogonale à tous les E_m, pour $m \in \mathbf{N}$, il résulte de ce qui précède que $\langle f, \chi_{[\frac{j}{2^m},\frac{j+1}{2^m}[}\rangle = 0$, c'est-à-dire que

$$\int_{\frac{j}{2^m}}^{\frac{(j+1)}{2^m}} f(x)dx = 0$$

pour $0 \leq j < 2^m$.

Nous pouvons maintenant conclure, car si $f \in L^2([0,1], \mathbf{C})$, la fonction $x \to \int_0^x f(t)dt$ est continue d'après le théorème de Cauchy-Schwarz : on écrit $|\int_x^y f(s)ds| \leq \left(\int_x^y |f(s)|^2 ds\right)^{1/2} (|x-y|)^{1/2}$. Mais cette fonction prend la même valeur en tous les points de la forme $\frac{k}{2^m}$ pour $0 \leq j < 2^m$ et tout $m \in \mathbf{N}$. Comme ces points forment un sous-ensemble dense de $[0,1]$ on conclut que $\int_0^x f(t)dt$ est constante sur $[0,1]$ donc identiquement nulle (car elle est nulle en $x = 0$).

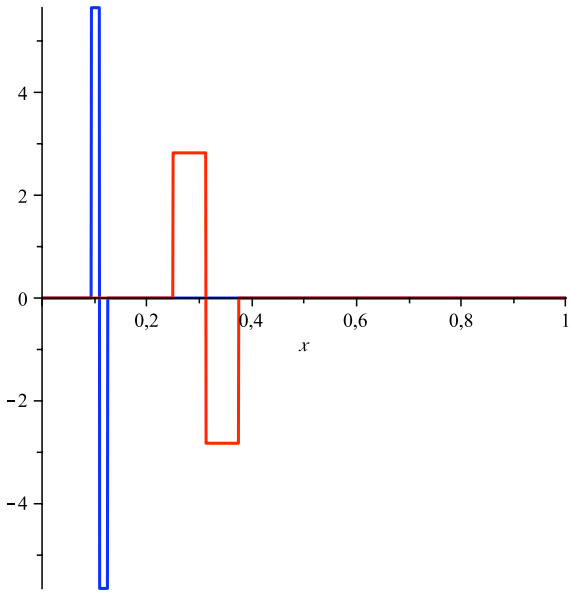
On appelle plus généralement base d'ondelettes [8] une base Hilbertienne constituée de fonctions du type $2^{n/2}\psi(2^n x - k)$.

Intermède culturel : convergence des séries de Fourier

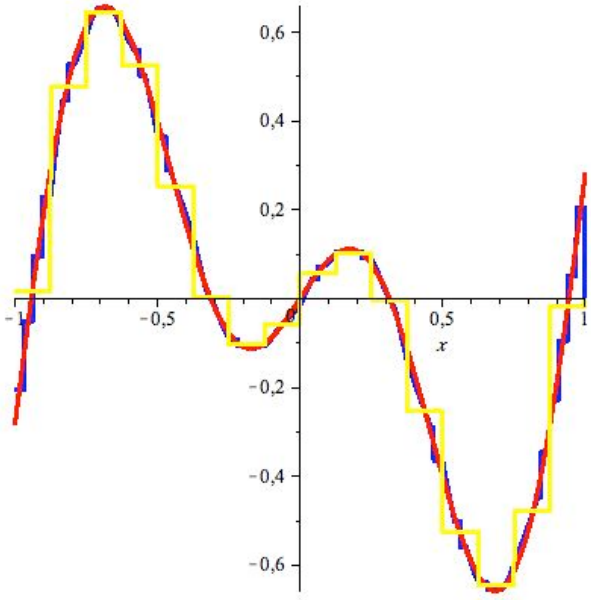
La question de la convergence de la série de Fourier de f, $\sum_{j=0}^{+\infty} c_n(f)e^{int}$ vers f date de Dirichlet. Il existe de nombreux théorèmes positifs (Dirichlet, Féjer, etc) qui affirment la convergence uniforme des sommes partielles $\sum_{p=-n}^{n} c_p(f)e^{ipt}$ vers f, pour lesquels nous renvoyons au cours de prépa, ou à des ouvrages plus avancés [9]. Mais même si f est continue, en général la série de Fourier de f ne converge pas uniformément vers f. Ce phénomène n'a plus lieu si f est

8. Voir Daubechies, Ten lectures on Wavelets, CBMS regional conference series in Applied Mathematics, SIAM, Philadelphia, 1992, ou le cours de Stéphane Mallat.

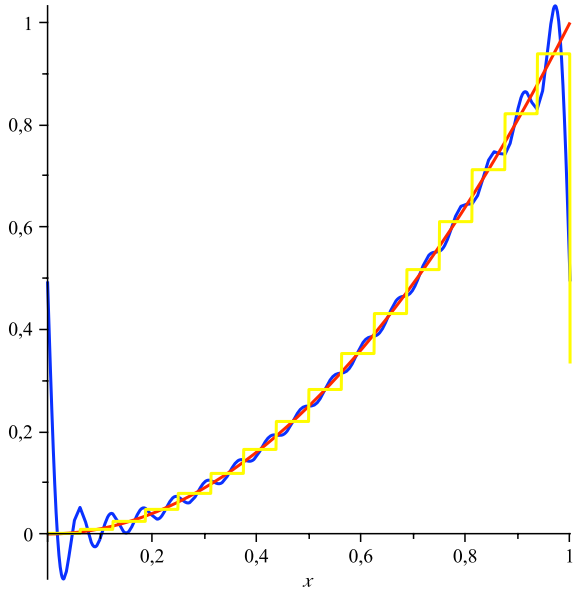
9. Comme par exemple [12]



Les ondelettes de Haar $\varphi_{5,3}$ (en bleu) et $\varphi_{3,2}$ (en rouge).



Décomposition en ondelettes de Haar de la fonction $x\cos(5x)$ sur $[0,1]$ avec 16 et 64 coefficients.



Décomposition en Fourier (en bleu) et en ondelettes de Haar (en jaune) de x^2 (en rouge) sur $[0,1]$ avec 16 coefficients.

Hölder[10]*. Cependant une suite qui converge L^2 a une sous-suite convergeant p.p, et on peut donc trouver une suite extraite des sommes partielles convergeant p.p. Un résultat célèbre et difficile de Carleson montre que pour f dans $L^2(\mathbf{S}^1)$, la suite $\sum_{-n}^{n} c_n(f)e^{int}$ converge p.p. vers $f(t)$. Un résultat de Kahane et Katznelson (1965) affirme qu'étant donné un ensemble quelconque de mesure nulle sur $\mathbf{S}^1$, il existe une fonction continue dont la série de Fourier ne converge pas sur E.*

Définition 3.1.3. On dit qu'un espace de Hilbert est séparable, s'il possède un ensemble dénombrable dense.

Théorème 3.1.4. *Tout espace de Hilbert séparable possède une base Hilbertienne.*

Démonstration. C'est le procédé d'orthonormalisation de Schmidt (voir remarque 3.1 page 156) appliqué à une suite de vecteurs formant un sous-ensemble dense. Plus précisément, soit x_n une suite dense dans H, et F_n le sev engendré par $x_1, ..., x_n$. Les F_n forment une suite croissante (au sens large), et $\dim(F_{n+1}) \leq \dim(F_n)+1$. On renumérote alors les F_n pour que ces espaces soient tous distincts (i.e. F_n étant défini, on note F_{n+1} le premier espace de la liste distinct de F_n . On aura alors $\dim(F_n) = n$ et on prend pour e_n un générateur de norme un de $F_{n-1}^{\perp} \cap F_n$ qui est de dimension un. Alors (e_n) est une base Hilbertienne, puisqu'il forme une famille orthonormale, et que les combinaisons linéaires finies des e_n engendrent $\bigcup_n F_n$ qui est dense par hypothèse. □

Remarque 3.1.5. Notons que si H possède une base Hilbertienne, il est séparable. En effet, l'ensemble $\{x = \sum_{j=1}^{n} a_j e_j \mid a_j \in \mathbf{Q}, n \geq 1\}$ est dénombrable [11] et dense.

Tous les espaces usuels, et tous les espaces considérés dans ce cours sont séparables : $\ell^2(\mathbf{N})$, $L^2([0,1],\mathbf{C})$, $L^2(\mathbf{R},\mathbf{C})$.... sont séparables. . C'est clair pour $\ell^2(\mathbf{N})$ qui possède la base Hilber-

10. i.e. il existe $\alpha > 0$ et une constante C tels que pour tout x, y, on a $|f(x) - f(y)| \leq C|x-y|^{\alpha}$.
11. Comme réunion dénombrable des $\mathbf{Q}^n$ qui sont dénombrables.

tienne donnée par $(e_n)_{n\geq 1}$) où e_n est la suite identiquement nulle à l'exception du n-ème terme, égal à 1. De même pour $L^2(\mathbf{S}^1, \mathbf{C})$ qui possède la base $e_n(t) = e^{int}$ pour $n \in \mathbf{Z}$. La situation est un peu plus délicate pour $L^2(\mathbf{R}, \mathbf{C})$.

Exercice 3.1.6. Montrer que $L^2(\mathbf{R}, \mathbf{C})$ est séparable [12]. On pourra

procéder de la manière suivante

(1) Montrer que $L^2([-n, n], \mathbf{C})$ s'injecte continûment dans $L^2(\mathbf{R}, \mathbf{C})$ de manière « naturelle ». On le considère alors comme un sous-espace de $L^2(\mathbf{R}, \mathbf{C})$.

(2) Montrer que
$$\bigcup_{n\geq 1} L^2([-n, n], \mathbf{C})$$
est dense dans $L^2(\mathbf{R}, \mathbf{C})$.

(3) En utilisant le fait que $L^2([-n, n], \mathbf{C})$ est séparable, en déduire que $L^2(\mathbf{R}, \mathbf{C})$ est séparable.

Corollaire 3.1.7. *Deux espaces Hilbertiens séparables sont isomorphes.*

Démonstration. On va montrer que H est isomorphe à $\ell^2(\mathbf{N})$. Soit $(e_n)_{n\geq 1}$ une base Hilbertienne de H.
A un élément $x = (x_n)_{n\geq 0}$ de $\ell^2(\mathbf{N})$ on associe $u(x) = \sum_{j=0}^{\infty} x_n e_n$, élément de H. Cette application u de $\ell^2(\mathbf{N})$ dans H vérifie $\|u(x)\| = \|x\|$ d'après l'égalité de Parseval. Elle est donc continue et injective. Montrons qu'elle est surjective d'inverse continue. Or on sait par définition des bases Hilbertiennes que si y est dans H il s'écrit de manière unique comme $y = \sum_{j=0}^{\infty} x_n e_n$ où $(x_n)_{n\geq 0}$ est dans $\ell^2(\mathbf{N})$. Ceci prouve que $y = u(x)$. L'égalité des normes de x et y entraîne alors la continuité de cette application. □

Notons que choisir un tel isomorphisme équivaut à choisir une base Hilbertienne. Si $(f_n)_{n\geq 0}$ est la base canonique de $\ell^2(\mathbf{N})$, et u est un tel isomorphisme, $(u(f_n))_{n\geq 1}$ est une base Hilbertienne de H.

Proposition 3.1.8. *Un sous-espace fermé d'un espace de Hilbert séparable est séparable.*

Démonstration. Soit $X = \{x_n \mid n \geq 1\}$ un sous-ensemble dénombrable dense de H. Alors, $P_F(X)$ est dense dans F. En effet, si $y \in F$, pour tout $\varepsilon > 0$ il existe n tel que $\|y - x_n\| \leq \varepsilon$. Mais P_F étant 1-Lipschitizienne, et égale à l'identité sur F,
$$\|P_F(x_n) - y\| = \|P_F(x_n) - P_F(y)\| \leq \|x_n - y\| \leq \varepsilon$$
ce qui montre la densité de P_F. □

Corollaire 3.1.9. *Si E est un espace de Hilbert de dimension infinie, il n'existe pas de mesure non triviale sur E, invariante par translation.*

Démonstration. En effet, $B(0, 1)$ contient une infinité de boules de rayon $1/4$, les $B(\frac{e_j}{2}, \frac{1}{4})$, donc si $\mu(B(0, 1))$ est finie, $\mu(B(0, 1/4)) = 0$, et comme H est réunion dénombrable de boules de rayon $1/4$, donc $\mu(H) = 0$. Il en résulte que pour tout $A \subset H$ $\mu(A) = 0$. □

12. On pourrait bien entendu utiliser l'existence d'une base Hilbertienne donnée par la dernière question de l'exercice 3.1 page 159

Remarque 3.1.10. Soit $(e_\alpha)_{\alpha\in A}$ une famille orthonormée de vecteurs d'un espace séparable. Alors A est dénombrable. En effet, pour chaque vecteur x, on vérifie par l'inégalité de Bessel que

$$\|x\|^2 \geq \sum_{\alpha\in A} |\langle x, e_\alpha\rangle|^2$$

puisque ce sera vrai pour toute sous-somme finie. Mais une somme convergente de réels positifs est nécéssairement dénombrable [13]. On en déduit que pour chaque $x \in H$, seul un nombre dénombrable des $\langle x, e_\alpha\rangle$ sont non nuls. Mais alors si X est un sous-ensemble dénombrable dense de H, l'ensemble $\{\alpha \in A \mid \exists n \langle x_n, \mathrm{e}_\alpha\rangle \neq 0\}$ est dénombrable comme réunion dénombrable d'ensembles dénombrables. Par sensité on en conclut que $\{\alpha \mid \exists x \in H \langle x, e_\alpha\rangle \neq 0\}$ est dénombrable. Mais comme $\langle e_\alpha, e_\alpha\rangle = 1 \neq 0$ cet ensemble est égal à A qui est donc lui-même dénombrable.

3.2 Compacité et non-compacité dans les espaces de Hilbert. Convergence faible.

Commençons par une conséquence de l'existence d'une base Hilbertienne [14].

Corollaire 3.2.1 (théorème de Riesz)**.** *Si H est de dimension infinie, sa boule unité $B = \{x \in E \mid \|x\| \leq 1\}$ n'est pas compacte.*

Démonstration. Supposons que la suite $(e_n)_{n\geq 1}$ ait une sous-suite convergente de limite z. Alors, d'une part pour tout vecteur x on a $\lim_{n\to\infty}\langle x, e_n\rangle = 0$ puisque $\sum_n |\langle x, e_n\rangle|^2$ est convergente. On déduit par continuité du produit scalaire que si z est une valeur d'adhérence de la suite, nécessairement $z = 0$. D'autre part $\|e_n\| = 1$ pour tout n, donc par continuité de la norme, on doit avoir $\|z\| = 1$, d'où une contradiction. □

Il est parfois souhaitable de récupérer une forme de compacité pour la boule unité. Mais il faut bien entendu changer de topologie. Nous introduisons alors la définition suivante.

Définition 3.2.2. On dit que la suite $(x_n)_{n\geq 0}$ « converge faiblement » vers $\bar{x}$ si et seulement si

$$\forall w \in E \quad \lim_n \langle x_n, v\rangle = \langle x_\infty, v\rangle$$

On note souvent $x_n \rightharpoonup \bar{x}$.

Comme son nom l'indique, la convergence faible est une condition plus faible que la convergence ordinaire, souvent appelée « convergence forte » lorsqu'on veut éviter toute ambiguité. En effet, si x_n converge fortement vers x, $|\langle z, x_n\rangle - \langle z, x\rangle| \leq \|z\|\|x_n - x\|$ qui tend donc vers 0 lorsque n tend vers l'infini. Plus délicat, les ensembles fermés pour la convergence forte ne le sont pas nécessairement pour la convergence faible. Par exemple la sphère unité est fermée pour la convergence forte (toute limite forte de vecteurs de norme un est de norme un) mais pas pour la convergence faible : la suite $(e_n)_{n\geq 1}$ est une suite de vecteurs de norme 1 qui n'a pas de limite forte, mais dont la limite faible est 0. En effet, quel que soit x, il s'écrit $x = \sum_{j=1}^{+\infty} x_n e_n$ avec $\sum_{j=1} x_n^2 < +\infty$ donc $\lim_n x_n = 0$. Mais $\lim_n x_n = \lim_n \langle x, e_n\rangle$, donc la limite faible de la

13. En effet, si $\sum_{\alpha\in A} x_\alpha$ (avec $x_\alpha \geq 0$) est convergente, alors l'ensemble des x_α non nuls est dénombrable. car l'ensemble $\{\alpha \mid x_\alpha \geq 1/n\}$ est fini, et $\{\alpha \mid x_\alpha \neq 0\} = \bigcup_{n\in\mathbf{N}^*}\{\alpha \mid x_\alpha \geq 1/n\}$.

14. Ce théroème reste vrai dans un espace de Banach, mais sa démonstration est plus compliquée.

suite $(e_n)_{n\geq 1}$ est bien 0. Une autre difficulté vient de ce que la convergence faible n'est pas une convergence associée à une distance (cf. exercice (18) page 189), mais seulement à une topologie. Néanmoins si une suite possède une limite, celle-ci est unique, car deux vecteurs distincts définissent des formes linéaires distinctes. Nous pouvons alors énoncer

Corollaire 3.2.3. *Si E est de dimension infinie, toute suite bornée $(x_k)_{k\geq 1}$ possède une sous-suite « convergeant faiblement » vers $\bar{x}$. Si $\bar{x}$ et la limite d'une telle sous-suite, on a l'inégalité $\|\bar{x}\| \leq \limsup \|x_k\|$.*

Démonstration. Notons que si la suite $(x_k)_{k\geq 1}$ converge faiblement vers $\bar{x}$, et on écrit les coordonnées dans une base Hilbertienne, $x_k = \sum_{j=1}^{+\infty} x_{j,k} e_j$, on doit avoir la convergence de chacune des coordonnés, $x_{j,k}$ vers $\bar{x}_j$ puisque $x_{j,k} = \langle x_k, e_j\rangle$. Notons aussi que la suite étant bornée, chacune des coordonnées reste bornée.

On commence alors par chercher une sous-suite telle que chacune des coordonnées converge, en utlisant le procédé d'extraction diagonale (voir le Chapitre de Topologie théorème 6.2.2 page 24 page 24). On extrait d'abord une sous-suite convergente de la suite borné des nombres complexes $x_{1,k}$ dont l'existence est assurée par le théorème de Bolzano-Weierstrass (Chapitre de Topologie, corollaire 4.3.5 page 17, page 17). On note $x_{1,\varphi_1(k)}$ cette sous-suite, et $\bar{x}_1$ sa limite. Puis une sous-suite convergente de $x_{2,\varphi_1(k)}$, $x_{2,\varphi_1\varphi_2(k)}$, etc.... On obtient des limites, $\bar{x}_1, \bar{x}_2, ...\bar{x}_k, ...$ Montrons d'abord que $\sum_{j=1}^{+\infty} |\bar{x}_j|^2$ est finie, ce qui permet de définir $\bar{x} = \sum_{j=1}^{+\infty} \bar{x}_j e_j$.

Pour simplifier les notations, on va encore noter $(X_k)_{k\geq 1}$ la suite extraite diagonale $X_k = x_{\varphi_1 \circ ... \circ \varphi_k(k)}$. Puisque pour tout j, on a $|\bar{x}_j| = \lim_{k\to\infty} |X_{j,k}|$ on a

$$\sum_{j=1}^{N} |\bar{x}_j|^2 = \lim_{k\to\infty} \sum_{j=1}^{N} |X_{k,j}|^2 \leq \lim_{k\to\infty} \|X_k\|^2 \leq \limsup_{k\to\infty} \|x_k\|$$

En prenant la limite lorsque N tend vers l'infini du membre de gauche, on obtient

$$\sum_{j=1}^{+\infty} |\bar{x}_j|^2 \leq \limsup_{k\to\infty} \|x_k\|^2$$

et donc $\bar{x} = \sum_{j=1}^{+\infty} \bar{x}_j e_j$ définit un élément de H qui vérifie

$$\|\bar{x}\|^2 \leq \limsup_{k\to\infty} \|x_k\|^2$$

Enfin la suite $(X_k)_{k\geq 1}$ converge faiblement vers $\bar{x} = \sum_{j=1}^{+\infty} \bar{x}_j e_j$. En effet, il est clair que $\langle X_k, e_j\rangle = x_{\varphi_1\varphi_2....\varphi_k(k),j}$ converge vers $\langle \bar{x}, e_j\rangle = \bar{x}_j$ vu que pour $k \geq j$, la suite $X_k = x_{\varphi_1\varphi_2....\varphi_k(k)}$ est extraite de $x_{\varphi_1\varphi_2....\varphi_j(k)}$ et a donc même limite.

Il en résulte que pour toute combinaison linéaire (finie) z, d'éléments de la base $(e_j)_{j\geq 1}$, on a $\lim_k \langle X_k, z\rangle = \langle \bar{x}, z\rangle$. Les combinaisons linéaires finies d'éléments de la base étant denses, si $z \in H$, quel que soit ε strictement positif, on peut trouver z_N combinaison linéaire finie des e_k tel que $\|z - z_N\| \leq \frac{\varepsilon}{4(\limsup_{k\to\infty} \|x_k\|)}$. Pour k assez grand, nous avons montré que $|\langle X_k, z_N\rangle - \langle \bar{x}, z_N\rangle| \leq \varepsilon/2$. On a alors

$$|\langle X_k - \bar{x}, z\rangle| \leq |\langle X_k - \bar{x}, z_N\rangle| + |\langle X_k - \bar{x}, z - z_N\rangle| \leq \varepsilon/2 + 2\left(\limsup_{k\to\infty} \|x_k\|\right)\varepsilon/4 \leq \varepsilon$$

□

Remarques 3.2.4. (1) Une suite convergeant faiblement est nécessairement bornée. Ce résultat, pour lequel on réfère à l'exercice (17) page 188, n'est pas évident, contrairement au cas de la convergence forte.

(2) Les applications linéaires continues (pour la convergence forte) préservent la convergence faible : si x_n converge faiblement vers x, alors $A(x_n)$ converge faiblement vers $A(x)$. En effet $\langle z, A(x_n)\rangle = \langle A^*(z), x_n\rangle$ converge par hypothèse vers $\langle A^*(z), x\rangle = \langle z, Ax\rangle$. En conclusion, quel que soit z, $\langle z, A(x_n)\rangle$ converge vers $\langle z, A(x)\rangle$ et donc $A(x_n)$ converge faiblement vers $A(x)$.

L'exemple de la suite $(e_n)_{n\geq 1}$ qui converge faiblement vers 0 montre que l'inégalité du théorème n'est pas, en géneral, une égalité.

Exercice 3.2.5. Montrer que si la suite $(x_n)_{n\geq 1}$ converge faiblement vers x et que $\lim_{n\to+\infty} \|x_n\| = \|x\|$ alors x_n converge fortement vers x. Indication : calculer $\|x_n - x\|^2$.

Revenons maintenant à la compacité pour la topologie usuelle d'un espace de Hilbert. Bien sûr tout fermé borné contenu dans un sous-espace de dimension finie d'un espace de Hilbert est compact. Mais il existe des compacts qui ne sont pas contenus dans un sous-espace de dimension finie. Un exemple fondamental est le « cube de Hilbert »

Proposition 3.2.6 (Compacité du cube de Hilbert)**.** *Le sous-ensemble de $\ell^2(\mathbf{N}^*)$ donné par*

$$K = \left\{ \sum_{n=1}^{+\infty} x_n e_n \mid |x_n| \leq \frac{1}{n} \right\}$$

est compact.

Démonstration. Notons d'abord que la somme $\sum_{n=1}^{+\infty} \frac{1}{n^2}$ étant convergente, la condition $|x_n| \leq \frac{1}{n}$ définit bien un élément de $\ell^2(\mathbf{N}^*)$. Soit alors une suite $(x_k)_{k\geq 1}$ d'éléments du cube de Hilbert. D'après le théorème a une limite faible, $\bar{x}$. En appliquant la définition de la limite faible aux éléments de la base Hilbertienne, on voit immédiatement que $\bar{x} \in K$. Montrons que x_k converge au sens usuel vers $\bar{x}$. Soit en effet $\varepsilon > 0$ et q assez grand pour que $\sum_{j\geq N} \frac{1}{j^2} \leq \varepsilon/8$. Alors choisissons N assez grand pour que si $1 \leq j \leq q$ on ait $|x_{k,j} - \bar{x}_j|^2 \leq \frac{\varepsilon}{2q}$. On a alors pour $k \geq N$

$$\|x_k - \bar{x}\|^2 \leq \sum_{j=1}^{q} |x_{k,j} - \bar{x}_j|^2 + \sum_{j\geq q+1} |x_{k,j} - \bar{x}_j|^2$$

Le premier terme de la partie droite de l'inégalité est majoré par $q\frac{\varepsilon}{2q}$, le second par $2\sum_{j\geq q+1} \left(|x_{k,j}|^2 + |\bar{x}_j|^2\right) \leq 4\sum_{j\geq q+1} \frac{1}{j^2} \leq \varepsilon/2$. On en déduit que $\|x_k - \bar{x}\|^2 \leq \varepsilon$, ce qui conclut la démonstration. □

4 Appendice : convergence des sommes de Féjer et complétude de la famille $(e^{int})_{n\in\mathbf{Z}}$ sur $L^2(\mathbf{S}^1,\mathbf{C})$

Soit $f(t)$ une fonction continue sur $\mathbf{S}^1$, que l'on identifie à une fonction 2π-périodique sur $\mathbf{R}$, et $c_n(f)=\frac{1}{2\pi}\int_{\mathbf{S}^1}e^{-int}f(t)dt$ ses coefficients de Fourier. On pose $s_n(f)(s)=\sum_{p=-n}^{n}c_p(f)e^{ips}$ la somme partielle de la série de Fourier de f, et σ_n la moyenne de Césaro de cette suite, à savoir

$$\sigma_n(f)(s)=\frac{1}{n+1}\sum_{q=0}^{n}s_q(f)(s)$$

On voit immédiatment que les s_q et donc les σ_n sont dans l'espace engendré par les e_p. De plus $\sigma_n(f)(t)=\int_{\mathbf{S}^1}f(t)F_n(s-t)dt=f\star F_n=\int_{\mathbf{S}^1}f(s-t)F_n(t)dt$ où

$$F_n(t)=\frac{1}{n+1}\sum_{q=0}^{n}D_q(t)$$

avec

$$D_q(t)=\sum_{p=-n}^{n}e^{ipt}=e^{-iqt}\frac{\left(e^{i(2q+1)t}-1\right)}{(e^{it}-1)}=\frac{\sin((q+\frac{1}{2})t)}{\sin(t/2)}$$

et donc

$$F_n(t)=\frac{1}{(n+1)\sin(t/2)}\sum_{q=0}^{}\sin((q+\frac{1}{2})t)=\frac{1}{(n+1)\sin(t/2)}\operatorname{Im}\left(\sum_{q=0}^{n}e^{i(q+1/2)t}\right)=$$
$$\frac{1}{(n+1)\sin(t/2)}\operatorname{Im}\left(e^{it/2}\frac{e^{i(n+1)t)}-1}{(e^{it}-1)}\right)=$$
$$\frac{1}{(n+1)\sin(t/2)}\operatorname{Im}\left(\frac{\sin((n+1/2)t)}{\sin(t/2)}e^{i(n+1/2)t}\right)=\frac{1}{(n+1)}\frac{\sin((n+1/2)t)^2}{\sin(t/2)^2}$$

La fonction $F_n(t)$ est appelée noyau de Féjer. Elle vérifie les propriétés suivantes :

(1) $\int_{\mathbf{S}^1}F_n(t)dt=1$

(2) $F_n(t)\geq 0$

(3) Pour tout $\delta>0$ on a $\lim_{n\to\infty}\int_{|t|\geq\delta}F_n(t)dt=0$

La première propriété résulte de ce que $\int_{\mathbf{S}^1}D_q(t)dt=1$, la seconde de la formule explicite précédente, et la troisième de ce que sur $|t|\geq\delta$, on a $F_n(t)\leq\frac{1}{(n+1)\sin(\delta/2)^2}$.

Alors $|f(s)-\sigma_n(f)(s)|=|\int_{-\pi}^{\pi}(f(s)-f(s-t))F_n(t)dt|$. On découpe cette intégrale en une partie sur $[-\delta,\delta]$ où δ est choisi assez petit pour que $|f(s)-f(s-t)|\leq\varepsilon$ pour tout s et tout t

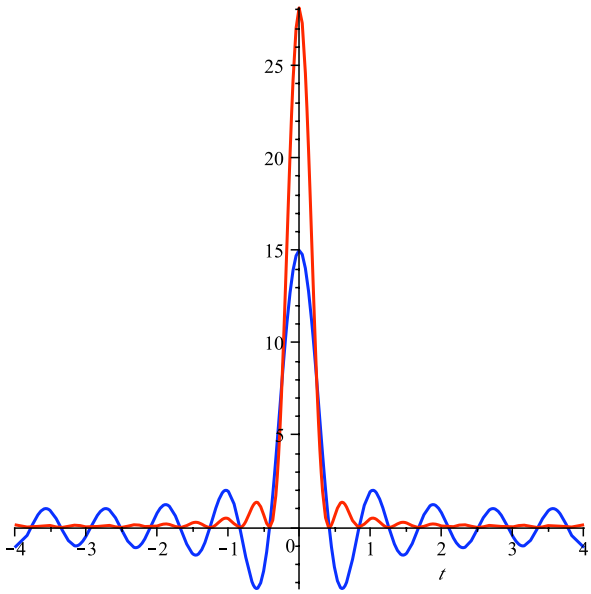


FIGURE VIII.2 – Les noyaux de Féjer (en rouge) et Dirichlet (en bleu) pour $n = 7$.

inférieur en module à δ. L'existence d'une tel δ résulte de la continuité uniforme de f sur $[-\pi, \pi]$. Alors

$$|\int_{-\pi}^{\pi}(f(s)-f(s-t))F_n(t)dt| \leq \int_{|t|\geq\delta} 2\left(\sup_{x\in[-\pi,\pi]}|f(x)|\right)F_n(t)dt + \int_{-\delta}^{\delta}\varepsilon F_n(t)dt$$

Notons que dans la dernière expression nous avons remplacé $|F_n(t)$ par $F_n(t)$ en utilisant la positivité de F_n. Mais $\int_{-\delta}^{\delta}\varepsilon F_n(t)dt \leq \varepsilon\int_{\mathbf{S}^1}F_n(t)dt = \varepsilon$. On en déduit que pourvu que n soit assez grand pour que $\int_{|t|\geq\delta}F_n(t)dt \leq \frac{\varepsilon}{2\|f\|}$ on a

$$|f(s)-\sigma_n(f)(s)| \leq 2\varepsilon$$

Ceci montre

Proposition 4.0.7. *Si f est une fonction continue sur $\mathbf{S}^1$, alors les sommes de Féjer convergent uniformément vers f.*

Remarque 4.0.8. Notons qu'il n'est pas vrai, en général, que les sommes partielles $s_n(f)$ convergent vers f, sauf si on suppose que f est par exemple C^1. Comme on l'a vu le noyau de Féjer est positif, ce qui n'est pas le cas du noyau de Dirichlet, et cette positivité joue un rôle crucial dans la preuve.

Exercice 4.0.9. Soit f une fonction continue sur $[0, \pi]$. Utiliser le théorème de Féjer pour montrer que f est limite uniforme de polynômes (théorème de Stone-Weierstrass).

(a) Montrer que l'on peut étendre f en une fonction continue paire sur $[-\pi, \pi]$, et poser $g(t) = f(\cos(t))$.

(b) Montrer que l'on peut approcher $g(t)$ par une suite de polynômes trigonométriques de la forme $\sum_{j=0}^{n} a_n \cos(nt)$.

(c) Utiliser le fait que $\cos(nt) = T_n(\cos(t))$ où T_n est le n-ième polynôme de Tchebychev [15] et conclure.

Corollaire 4.0.10. *Si $f \in L^2(\mathbf{S}^1,\mathbf{C})$ les sommes partielles de la série de Fourier de f convergent dans L^2 vers f. En d'autres termes posant $e_p(t) = e^{ipt}$, on a*

$$\|f - \sum_{p=-N}^{N} c_p(f)e_p\|_2$$

tend vers 0

Notons que les sommes de Féjer ont disparu du corollaire.

Démonstration. D'après le théorème de Riesz, et le théorème 3.0.13 page 156, il suffit de montrer que si $f \in L^2$ est orthogonale à tous les e_p, elle est nulle presque partout. Or si on avait une telle f non nulle, on peut par homothétie se ramener au cas où $\|f\| = 1$. Le noyau de Féjer nous donne la densité (L^2)de $C^0([-\pi,\pi],\mathbf{C})$ dans $L^2([-\pi,\pi],\mathbf{C})$, on peut trouver g continue telle que $\|g - f\| \leq \varepsilon$. Mais alors, utilisant le fait que F_n est dans l'espace engendré par les e_p et donc $G_n = g \star F_n$ aussi, on a $\langle f, G_n\rangle = 0$, donc

$$|\langle g, G_n\rangle| = |\langle f - g, G_n\rangle| \leq \|f - g\|\|G_n\| \leq \varepsilon 2\|g\|$$

puisque G_n converge uniformément vers g sur $[-\pi,\pi]$. Mais par cette même convergence uniforme, G_n converge vers g dans $L^2([-\pi,\pi],\mathbf{C})$ et donc $\langle g, G_n\rangle$ converge vers $\|g\|^2$. On en déduit que $\|g\|^2 \leq 2\varepsilon\|g\|$ et puisque $1 - \varepsilon \leq \|g\| \leq 1 + \varepsilon$, que $(1-\varepsilon)^2 \leq 2\varepsilon(1+\varepsilon)$ ce qui est impossible pour ε assez petit. □

Notons que d'autres démonstrations utilisent la convergence uniforme des sommes partielles de Fourier, en demandant plus de régularité à f, par exemple que f soit C^1. On conclut de la même manière, en utilisant la densité de l'espace des fonctions C^1 pour la norme L^2 (cf. exercice (14) page 188).

4.1 Ce qu'il faut retenir

Un espace de Hilbert est un espace préhilbertien complet

Inégalités et égalité fondamentales pour le produit hermitien

Les sous-espaces fermés vectoriel d'un Hilbert sont des Hilbert

Exemples de sous-espaces non fermés

Formules d'orthogonalité classiques (i.e. celles de la dim finie) plus $(F^\perp)^\perp = \overline{F}$

Théorème de la projection.

15. Les polynômes de Tchebychev forment d'ailleurs une famille orthogonale pour $\frac{1}{\sqrt{1-x^2}}$, ce qui résulte par un changement de variable de l'orthogonalité des $\cos(nx)$.

Formule pour la projection sur l'adhérence de l'espace engendré par une famille orthonormée

Théorème de Riesz (cas linéaire et sesquilinéaire)

Critère de densité-Hahn-Banach

Critère de continuité d'un opérateur. Existence et continuité de l'adjoint

Bases Hilbertiennes : définition, théorème d'existence.

Si $(e_n)_{n\geq 1}$ est une base Hilbertienne, la suite $\sum_{n=1}^{\infty} x_n e_n$ converge ssi $\sum_{n=1}^{\infty} |x_n|^2$ converge.

Formule de Parseval

$$\|x\|^2 = \sum_{n=1}^{\infty} \langle x, e_n \rangle^2$$

Les $(e^{int})_{n\in\mathbf{Z}}$ forment une base Hilbertienne de $L^2(S^1, \mathbf{C})$. Théorème de Féjer

Procédé d'orthogonalisation de Schmidt, formule de Bessel.

Résultats de compacité et non-compacité :

Non-compacité de la boule unité (thèorème de Riesz)

Convergence faible et interprétation comme "convergence coordonnée par coordonnée"

Compacité séquentielle de la boule unité pour la convergence faible

Compacité du cube de Hilbert

Chapitre IX

Opérateurs dans un espace de Hilbert

1 Extension d'opérateurs, transformation de Fourier L^2 et applications

1.1 Extension d'opérateurs et transformation de Fourier dans L^2

On considère deux espaces de Hilbert $(H, \langle , \rangle_H)$ et $(E, \langle , \rangle_E)$. Lorsqu'aucune confusion n'est possible sur l'espace ambiant, on note simplement $\langle , \rangle$ le produit hermitien. On souhaite prolonger une application définie sur un sous-espace dense de H à valeurs dans E.

Théorème 1.1.1. *Soit $u : F \longrightarrow E$ une application linéaire définie sur un sous-espace dense F de H. Soit $N_F(u) = \sup_{x\in F, \|x\|=1} \|u(x)\|$. Si $N_F(u) < +\infty$, alors u se prolonge en une application linéaire continue de H dans E telle que $\|u\| = N_F(u)$.*

Démonstration. Si $x \in H$ est limite d'une suite x_n d''éléments de F, on pose $u(x) = \lim_{n\to\infty} u(x_n)$. Comme $\|u(x_n) - u(x_m)\| \leq N_F(u)\|x_n - x_m\|$ et que la suite x_n est de Cauchy, on en déduit que $u(x_n)$ est de Cauchy, et donc possède une limite. Par le même argument, cette limite ne dépend pas du choix de la suite, et

$$\|u(x)\| = \lim_{n\to\infty} \|u(x_n)\| \leq N_F(u) \lim_{n\to\infty} \|x_n\| = N_F(u)\|x\|$$

□

Le théorème précédent permet d'étendre la transformation de Fourier à l'espace des fonctions L^2.

Théorème 1.1.2. *Soit Φ la transformée de Fourier normalisée, définie par*

$$\Phi(f)(\xi) = \frac{1}{\sqrt{2\pi}} \widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} e^{-ix\xi} f(x)dx$$

sur l'espace $L^1(\mathbf{R},\mathbf{C})$. *Alors* Φ *s'étend en une isométrie bijective de* $L^2(\mathbf{R},\mathbf{C})$ *dans lui-même. Il en est de même pour* $\Psi(f)=\frac{1}{\sqrt{2\pi}}\int_{\mathbf{R}} e^{ix\xi}f(x)dx$ *et son extension vérifie* $\Phi\circ\Psi=\Psi\circ\Phi=Id$.

Démonstration. On a montré (cf. théorème 3.0.12 page 136) que posant $\Psi=\frac{1}{\sqrt{2\pi}}\int_{\mathbf{R}} e^{ix\xi}f(x)dx$, on a $\Phi\Psi=\Psi\Phi=\mathrm{Id}$ si $f,\hat{f}\in L^1(\mathbf{R},\mathbf{C})$. Ceci est en particulier vérifié si $f\in C_0^2$, car alors $f''\in L^1$ et d'après le théorème de Riemann-Lebesgue (cf. théorème 2.0.7 page 130), sa transformée de Fourier qui est $\xi^2\hat{f}(\xi)$ tend vers 0 lorsque ξ tend vers l'infini. Il en résulte que $\hat{f}(\xi)\leq\frac{C}{|\xi|^2}$ qui est bien dans L^1. Rappelons que $C_0^2(\mathbf{R},\mathbf{C})$ est dense dans l'esapce $L^2(\mathbf{R},\mathbf{C})$. Par ailleurs, Ψ est « l'adjoint de Φ » pour le produit scalaire $L^2(\mathbf{R},\mathbf{C})$, c'est-à-dire que

$$<\Phi f,g>=<f,\Psi g>$$

En effet, comme $|f(x)||g(\xi)|$ est dans $L^1(\mathbf{R}^2,\mathbf{C})$, le théorème de Fubini (cf. théorème 2.0.1 page 78) permet d'intervertir les intégralesdans le calcul suivant

$$<\Phi f,g>=\frac{1}{\sqrt{2\pi}}\int_{\mathbf{R}}\left(\int_{\mathbf{R}} e^{-ix\xi}\bar{f}(x)dx\right)g(\xi)d\xi=\frac{1}{\sqrt{2\pi}}\int_{\mathbf{R}}\int_{\mathbf{R}}\bar{f}(x)g(\xi)e^{-ix\xi}dxd\xi=$$
$$\frac{1}{\sqrt{2\pi}}\int_{\mathbf{R}}\bar{f}(x)\int_{\mathbf{R}} g(\xi)e^{-ix\xi}dxd\xi=<f,\Psi g>$$

On en déduit

$$\langle\Phi f,\Phi g\rangle=\langle f,\Psi\Phi g\rangle=\langle f,g\rangle$$

et de manière analogue $\langle\Psi f,\Psi g\rangle=\langle f,g\rangle$

On en déduit que Φ et Ψ s'étendent à l'espace $L^2(\mathbf{R},\mathbf{C})$, que les extensions sont des isométries, et que $\Phi\circ\Psi=Id$, car C_0^2 est dense dans $L^2(\mathbf{R},\mathbf{C})$ et une identité vraie sur un sous-espace dense est vraie partout. □

Ce que nous venons de faire avec une variable s'étend sans difficulté au cadre de plusieurs variables. Ceci permet en particulier de résoudre le problème de la diffraction de Fraunhofer, mentionnée dans la section 1 page 127. En effet, si on connaît $\Phi(1_P)$, où 1_P est la fonction indicatrice de la pupille, qui est un ensemble borné (et mesurable) du plan et puisque cette fonction est dans $L^2(\mathbf{R}^2,\mathbf{C})$, on peut écrire $\Psi(\Phi(1_P))=1_P$: en appliquant Ψ (qui est la transformée de Fourier à un coefficient près) à $\widehat{1_P}$ on retrouve 1_P.

1.2 Inégalité de Heisenberg.

Cette inégalité a de nombreuses conséquences de la physique quantique à la théorie du signal. Elle affirme en effet qu'un signal ne peut être localisé à la fois en temps et en fréquence, ou qu'une particule ne peut avoir simultanément une position et une quantité de mouvement bien déterminées (cf. le cours de mécanique quantique), et explicite cette restriction de manière quantitative.

Proposition 1.2.1. *Si f est une fonction dans $C_c^1(\mathbf{R},\mathbf{C})$. On a*

$$\left(\int_{\mathbf{R}} x^2|f(x)|^2dx\right)^{1/2}\left(\int_{\mathbf{R}}|\frac{d}{dx}f(x)|^2d\xi\right)^{1/2}\geq\frac{1}{2}\|f\|_{L^2}^2$$

Cette inégalité est équivalente à l'inégalité suivante, vérifiée par toute fonction de $L^2(\mathbf{R},\mathbf{C})$

$$\left(\int_{\mathbf{R}} x^2|f(x)|^2dx\right)^{1/2}\left(\int_{\mathbf{R}}\xi^2|\hat{f}(\xi)|^2d\xi\right)^{1/2}\geq\sqrt{\frac{\pi}{2}}\|f\|_{L^2}^2$$

ou encore

$$\|xf\|_{L^2}\|\xi\hat{f}\|_{L^2}\geq\sqrt{\frac{\pi}{2}}\|f\|_{L^2}^2$$

Bien entendu les termes de gauche ne sont finis que si $xf(x)$ et $\xi\hat{f}(\xi)$ sont dans $L^2(\mathbf{R})$. Notons que $\xi\hat{f}(\xi)$ est dans $L^2(\mathbf{R})$ si et seulement si sa transformée de Fourier y est, c'est-à-dire $\frac{d}{dx}f(x)$ est dans L^2.

Démonstration de l'inégalité de Heisenberg. Note : nous ne démontrons ici que le cas où f est dans $C_c^1(\mathbf{R},\mathbf{C})$, l'extension au cas où xf et $\xi\hat{f}$ sont dans $L^2(\mathbf{R},\mathbf{C})$ étant laissée en exercice (cf. exercice (21) page 189). La démonstration est une version hermitienne de Cauchy-Schwarz. Plus généralement, si A,B sont des opérateurs hermitiens (i.e. tels que $A^*=A, B^*=B$), on pose $C=i(AB-BA)$ qui est aussi hermitien. On développe alors le trinôme positif suivant

$$0\leq\langle Ax+itBx,Ax+itBx\rangle=\|Ax\|^2+it\left[\langle Bx,Ax\rangle+\langle Ax,Bx\rangle\right]+t^2\|Bx\|^2=$$
$$\|Ax\|^2+t\langle i(AB-BA)x,x\rangle+t^2\|Bx\|^2$$

D'où en écrivant la négativité de son disciminant

$$\langle Cx,x\rangle\leq 4\|Ax\|^2\|Bx\|^2$$

Prenant $A(f)=xf$ et $B(f)=i\frac{d}{dx}f$ il faut faire attention au fait que A,B ne sont pas définis sur tout l'espace $L^2(\mathbf{R},\mathbf{C})$. On doit vérifier que

$$\langle A(f),B(f)\rangle=\int_{\mathbf{R}}x\overline{f}(x)i\frac{d}{dx}f(x)dx=\int_{\mathbf{R}}\overline{i\frac{d}{dx}(xf(x))}f(x)dx=\langle BA(f),f\rangle$$

ce qui résulte d'une intégration par parties et que

$$\langle B(f),A(f)\rangle=\int_{\mathbf{R}}\frac{d}{dx}\overline{f}(x)xf(x)dx=\int_{\mathbf{R}}(x\frac{d}{dx}(\overline{f}(x)))f(x)dx=\langle AB(f),f\rangle$$

le on aura $BA(f)=i\frac{d}{dx}(xf)=if+ix\frac{d}{dx}(f)$ alors que $AB(f)=ix\frac{d}{dx}f$ et $i(AB-BA)(f)=f$ et $\langle Cf,f\rangle=\|f\|_{L^2}^2$. Ceci démontre l'inégalité si f est suffisamment régulière, par exemple si f est de classe C^1 et f,f' sont dasn L^2. Si de plus f est dans C_0^1 et telle que f et sa dérivée soient dans $L^2\cap L^1$, alors, d'après le théorème 2.0.8 page 132 (cf. page 132) on peut écrire $\widehat{f'}(\xi)=\xi\widehat{f}(\xi)$ et la seconde inégalité est alors équivalente à la première. Cette seconde inégalité s'étend au cas géneral par densité : si $f_n\longrightarrow f$ dans L^2 et la convergence a lieu presque partout, alors $\widehat{f_n}$ converge vers $\widehat{f}$ dans L^2, donc quitte à extraire une sous-suite, $\widehat{f_n}$ converge p.p. vers $\widehat{f}$. On en déduit d'une part que $\|f_n\|\longrightarrow\|f\|$ et $\lim_{n\to\infty}\|xf_n\|=$ □

On voit que l'inégalité ne devient égalité que si $Af + itBf$ s'annule pour un certain t, soit si $xf(x) - t\frac{d}{dx}f(x) = 0$ soit $f(x) = e^{\frac{-x^2}{2t}}$.

L'interprétation en termes de concentration est la suivante : si $\int_{\mathbf{R}} |f(x)|^2 dx = 1$ la quantité $\int_{\mathbf{R}} x^2 f(x)^2 dx$ est d'autant plus petite que f est concentrée au voisinage de 0. Si on souhaite mesurer la concentration de f au voisinage de q et celle de ses fréquences au voisinage de p, on regarde les quantités

$\left(\int_{\mathbf{R}} (x-q)^2 |f(x)|^2 dx\right)$ et $\left(\int_{\mathbf{R}} (\xi - p)^2 |\hat{f}(\xi)|^2 d\xi\right)$ qui satisfont la même inégalité que dans le cas $q = 0, p = 0$:

$$\|(x-q)f\|_{L^2} \|(\xi - p)\hat{f}\|_{L^2} \geq \frac{\sqrt{2\pi}}{4} \|f\|_{L^2}^2$$

et le terme de gauche est minimal lorsque $f_{q,p}(x) = e^{\frac{-(x-q)^2}{2t}} e^{ipx}$. Les physiciens appellent les fonctions $f_{q,p}$ états cohérents. Ce sont les fonctions qui réalisent l'égalité dans l'inégalité de Heisenberg. Elles représentent une particule aussi localisée que possible dans l'espace au voisinage de q et en fréquence au voisinage de p.

La décomposition de Fourier écrit une fonction comme somme des $e^{i\xi x}$ qui sont parfaitement localisées en fréquence (égale à ξ), mais pas du tout en espace. Il est souvent utile d'avoir une décomposition qui tient compte de la localisation spatiale des fonctions. Une méthode utlisée avant les ondelettes consiste en une transformation de Fourier « avec fenêtre » dont le principe de base est le suivant : on restreint $f(x)$ à $[2\pi n, 2\pi(n+1)]$ et on décompose en série de Fourier. L'un des problèmes est l'ajout artificiel de hautes fréquences du à la troncature. Les ondelettes permettent, elles, de conserver une information qui tient compte à la fois de la localisation spatiale et fréquentielle.

1.3 Formule de Poisson

Soit f une fonction de $L^1(\mathbf{R}, \mathbf{C})$. On lui associe la fonction $g(t) = \sum_{n\in\mathbf{Z}} f(t + 2\pi n)$ qui est périodique de période 2π. Elle n'a aucune raison d'être dans $L^1(\mathbf{R}, \mathbf{C})$ mais c'est une fonction de $L^1(\mathbf{S}^1, \mathbf{C})$, grâce au théorème de convergence monotone qui nous permet d'intervertir somme et intégrale dans le calcul suivant :

$$\int_0^{2\pi} |\sum_{n\in\mathbf{Z}} f(t+2\pi n)| dt \leq \int_0^{2\pi} \sum_{n\in\mathbf{Z}} |f(t+2\pi n)| dt = \sum_{n\in\mathbf{Z}} \int_0^{2\pi} |f(t)| dt =$$
$$\sum_{n\in\mathbf{Z}} \int_{2n\pi}^{2(n+1)\pi} |f(t)| dt = \int_{\mathbf{R}} |f(t)| dt$$

On peut alors considérer la transformée de Fourier de f d'une part, et les coefficients de Fourier de g de l'autre. On s'attend bien sûr à un lien entre ces deux objets. Utilisons encore le théorème de convergence monoitone, on a

$$c_k(g) = \frac{1}{2\pi} \int_0^{2\pi} \sum_n f(t+2\pi n) e^{-ikt} dt = \frac{1}{2\pi} \sum_n \int_0^{2\pi} f(t+2\pi n) e^{-ikt} dt =$$
$$\frac{1}{2\pi} \sum_n \int_{2\pi n}^{2\pi(n+1)} f(t) e^{-ikt} e^{i2\pi kn} dt = \frac{1}{2\pi} \int_{\mathbf{R}} f(t) e^{-ikt} dt = \frac{1}{2\pi} \hat{f}(k)$$

Comme la formule $g(t) = \sum_{k\in\mathbf{Z}} c_k(g)e^{ikt}$ est valable par exemple si g est dans L^2, la convergence étant dans ce même espace (et non pas ponctuelle), on obtient formellement que

$$\sum_{n\in\mathbf{Z}} f(t+2\pi n) = \frac{1}{2\pi}\sum_k \hat{f}(k)e^{ikt}$$

Le résultat sera vrai par exemple si le terme de gauche est dans $L^2([0,2\pi],\mathbf{C})$, et la convergence des sommes doit être comprise dans $L^2([0,2\pi],\mathbf{C})$. On peut alors montrer

Proposition 1.3.1. *Soit f une fonction telle qu'il existe une constante C telle que $|f(t)| \leq \frac{C}{1+|t|^2}$ pour tout $t \in \mathbf{R}$. Alors on a*

$$\sum_{n\in\mathbf{Z}} f(t+2\pi n) = \frac{1}{2\pi}\sum_k \hat{f}(k)e^{ikt}$$

la convergence des séries étant entendue dans $L^2([0,2\pi],\mathbf{C})$.

Démonstration. En effet il suffit de vérifier que g est dans $L^2(\mathbf{S}^1,\mathbf{C})$. Or

$$\int_0^{2\pi} |f(t)|^2 dt = \int_0^{2\pi} \sum_{m,n\in\mathbf{Z}} \overline{f}(t+2\pi m)f(t+2\pi n)dt \leq \sum_{n\in\mathbf{N}}\sum_{m\in\mathbf{Z}} \int_{2\pi m}^{2\pi(m+1)} |f(t)||f(t+2\pi(m-n))|dt =$$
$$\sum_{q\in\mathbf{Z}} \int_{\mathbf{R}} |\overline{f}(t)||f(t+2\pi q)|$$

Nous savons que

$$\int_{|t|\geq n} |f(t)|^2 dt \leq C\int_{|t|\geq n} \frac{dt}{(1+|t|^2)^2} \leq \frac{C'}{n^3}$$

Or

$$\int_{\mathbf{R}} |f(t)||f(t+2\pi q)|dt = \int_{|t|\leq 2\pi n} |f(t)||f(t+2\pi q)|dt + \int_{|t|\geq 2\pi n} |f(t)||f(t+2\pi q)|dt =$$
$$\int_{|t-2\pi q|\leq 2\pi n} |f(t-2\pi q)||f(t)|dt + \int_{|t|\geq 2\pi n} |f(t)||f(t+2\pi q)|dt \leq$$
$$\int_{|t|\geq 2\pi|q-n|} |f(t-2\pi q)||f(t)|dt + \int_{|t|\geq 2\pi n} |f(t)||f(t+2\pi q)|dt$$

donc si $n \simeq |q|/2$ cette somme se majore par

$$2\left(\int_{|t|\geq \pi q} |f(t)|^2 dt\right)^{1/2}\left(\int_{\mathbf{R}} |f(t)|^2 dt\right)^{1/2}$$

soit $2\frac{\sqrt{C}'}{q^{3/2}}\|f\|_{L^2}$ qui est le terme général d'une série absolument convergente. On en conclut que g est dans $L^2([0,2\pi],\mathbf{C})$. Elle est donc égale dans$L^2([0,2\pi],\mathbf{C})$ à la somme de sa série de Fourier, ce qui démontre la proposition. □

Nous voudrions maintenant pouvoir utiliser cette égalité en $t = 0$. Pour que cela ait un sens, il faut au moins que g soit continue et qu'elle soit limite en tout point de sa série de Fourier. Par exemple, c'est le cas si f est de classe C^2 et g dans L^2.

Corollaire 1.3.2 (Formule de Poisson)**.** *Si f est de classe C^2 et qu'il existe une constante C telle que $|f(t)| \leq \frac{C}{1+|t|^2}$ pour tout $t \in \mathbf{R}$, on a*

$$\sum_{n\in\mathbf{Z}} f(2\pi n) = \frac{1}{2\pi}\sum_{k\in\mathbf{Z}} \hat{f}(k)$$

Démonstration. En effet lorsque f est de classe C^2, $\xi^2\hat{f}(\xi)$ tend vers 0 d'après le théorème de Riemann-Lebesgue (cf. théorème 2.0.7). Il en résulte que la série $\sum_{k\in\mathbf{Z}} \hat{f}(k)e^{ikt}$ est absolument convergente, et sa somme représente donc une fonction continue. De même l'inégalité $|f(t)| \leq \frac{C}{1+|t|^2}$ entraîne que la série de fonctions de terme général $f(t+2\pi n)$ est absolument convergente sur tout compact, et converge donc vers la fonction $g(t)$ qui est donc continue.

La proposition précédente affirme alors l'égalite dans $L^2([0,2\pi],\mathbf{C})$ de deux fonctions continues. Il en résulte que les fonctions sont égales en tout point. □

En remplaçant $f(x)$ par $f_a(x) = f(ax)$ on a $\widehat{f_a}(\xi) = \frac{1}{a}\widehat{f}(\frac{\xi}{a})$. La formule précédente s'écrit alors

Corollaire 1.3.3. *Sous les hypothèses du corollaire 1.3.2, on a*

$$\sum_{n\in\mathbf{Z}} f(2\pi na) = \frac{1}{2\pi a}\sum_{k\in\mathbf{Z}} \hat{f}(k)$$

Application : Théorème d'échantillonnage de Shannon-Nyquist.

Soit f de classe C^2 et de support contenu dans $]-\pi,\pi[$. Elle vérifie bien entendu les hypothèses permettant d'écrire la formule de Poisson, et, on a pour $t \in [-\pi,\pi]$

$$g(t) = f(t) = \frac{1}{2\pi}\sum_{k} \hat{f}(k)e^{ikt}$$

Il suffit donc de connaître les $\hat{f}(k)$ pour reconstituer f. Ceci n'est évidememment pas surprenant, car on peut considérer f comme une fonction 2π périodique, et la formule n'est rien d'autre que la formule d'inversion de Fourier. Inversement, en remplaçant f par $\hat{f}$ et en utilisant la formule d'inversion de Fourier, on obtient

$$\sum_{n\in\mathbf{Z}} \hat{f}(t+2\pi n) = \sum_{k\in\mathbf{Z}} f(k)e^{-ikt}$$

si $\hat{f}(t)$ est à support dans $]-\pi,\pi[$, on a pour

$$\forall t \in]-\pi,\pi[\ \ \hat{f}(t) = \sum_{n\in\mathbf{Z}} \hat{f}(t+2\pi n) = \sum_{k\in\mathbf{Z}} f(k)e^{-ikt}$$

et la connaissance des $f(k)$ permet de reconstituer $\hat{f}$ et donc f.

En d'autres termes si on a un signal dont les fréquences sont contenues dans $]-\pi,\pi[$, l'échantillonage $f(k)$ permet de reconstituer le signal. Plus généralement, en appliquant la formule ci-dessus à $f_a(t) = f(ta)$ et en notant que $\hat{f}_a(\xi) = \frac{1}{a} f(\xi/a)$, si on a un signal de fréquence contenues dans $]-\pi/a, \pi/a[$, l'échantillonage $f(ka)$ permet de reconstituer le signal par la formule

$$\hat{f}\left(\frac{\xi}{a}\right) = a \sum_{k\in\mathbf{Z}} f(ka) e^{-ika\xi}$$

2 Introduction à la théorie spectrale

Dans ce chapitre, tous les espaces de Hilbert sont supposés séparables.

On se demande s'il existe une théorie semblable à celle des valeurs propres et vecteurs propres en dimension finie. On rappelle que si A est un opérateur hermitien d'un espace de **dimension finie**, il possède une base orthonormée de vecteurs propres.

Cependant, en dimension infinie un endomorphisme d'un espace de Hilbert peut ne pas être inversible tout en étant injectif. Cela nous amène à distinguer deux objets

Définition 2.0.4. Soit A un endomorphisme de l'espace de Hilbert H. On appelle **spectre** de A, noté sp(A), l'ensemble des λ tels que $A - \lambda \operatorname{Id}_H$ soit inversible. On appelle valeur propres de A, note vp(A) l'ensemble des valeurs propres de A, ensemble des λ tel que $A - \lambda \operatorname{Id}_H$ ne soit pas injectif – ou encore il existe $x \neq 0$ tel que $Ax = \lambda x$.

Bien entendu vp(A) $\subset$ sp(A) et l'exemple ci-dessous montre que l'égalité n'a pas nécessairement lieu.

Exemple 2.0.5. Soit A l'opérateur sur $L^2(\mathbf{S}^1, \mathbf{C})$ défini par $A(f)(t) = f(t)\cos(t)$ et f un vecteur propre. Alors $A(f)(t) = \lambda f(t)$ signifie que $\cos(t)f(t) = \lambda f(t)$, d'où $f = 0$ *p.p.* et donc $f \equiv 0$ dans $L^2(\mathbf{S}^1, \mathbf{C})$. Notons que A est hermitien, c'est-à -dire auto-adjoint : $\langle Af, g\rangle = \langle f, Ag\rangle$.

Nous cherchons alors une classe d'opérateurs pour lesquels la situation est proche de celle de la dimension finie.

Opérateurs hermitien : on dit que $A \in \mathcal{L}(H, H)$ est hermitien (resp. antihermitien) si

$$\forall x, y \in H \;\; \langle Ax, y\rangle = \langle x, Ay\rangle \;\; (\text{resp.} \langle Ax, y\rangle = -\langle x, Ay\rangle)$$

Rappelons que les valeurs propres d'un opérateur hermitien sont nécessairement réelles [1], car $\langle Ax, x\rangle = \langle x, Ax\rangle = \overline{\langle Ax, x\rangle}$ est réel, et si $Ax = \lambda x$, on aura $\langle Ax, x\rangle = \lambda \|x\|^2$, donc $\lambda \in \mathbf{R}$.

Opérateurs compacts : On dit que A est compact si $A(B(0,1))$ est contenu dans un sous-ensemble compact de H.

Exemples 2.0.6. (1) Si P_F est l'opérateur de projection sur F, c'est un opérateur hermitien. D'après le théorème de Riesz (théorème 3.2.1 page 164), il est compact si et seulement si F est de dimension finie, vu que $P_F(B(0,1)) \supset B(0,1) \cap F$. D'après l'exercice (22) page 190,

1. Le même argument montre que si A est anti-hermitien, ses valeurs propres sont imaginaires pures.

un opérateur est compact si et seulement si il est limite, pour la norme définie sur $\mathcal{L}(H,H)$ en norme, d'opérateurs de rang fini.

(2) Une limite en norme d'opérateurs compacts est un opérateur compact. (cf. exercice (19) page 189).

(3) Si F est une application linéaire continue et A un opérateur compact, alors $F \circ A$ et $A \circ F$ sont compacts. En effet, supposons $A(B(0,1))$ contenu dans le compact K. Comme F envoie $B(0,1)$ dans $B(0,\|F\|)$ on a $AF(B(0,1)) \subset A(B(0,\|F\|))$ est contenu dans $\|F\| \cdot K$ qui est aussi compact. De manière analogue, $F(A(B(0,1)) \subset F(K)$ et la continuité de F entraîne la compacité de $F(K)$.

(4) Si L est l'opérateur associé par le théorème de Riesz à une forme sesquilinéaire B, alors L est hermitien si et seulement si B est hermitienne.

(5) L'application $f \longrightarrow I(f)$ définie par $I(f)(t) = \int_0^t f(s)ds$ sur $L^2([0,1],\mathbf{C})$ est compact. En effet sur les $e_n = e^{int}$, on a $I(e_n) = \frac{1}{in}e_n$. on en déduit que I envoie l'ensemble des $(a_n)_{n\geq 1}$ telles que $\sum_{n\geq 1}|a_n|^2 \leq 1$ dans l'ensemble des suites $(b_n)_{n\geq 1}$ telles que $\sum_{n\geq 1} n^2|b_n|^2 \leq 1$. Cet ensemble est contenu dans le cube de Hilbert qui est compact (cf. proposition 3.2.6 page 166).

(6) Pour $k \geq 1$ l'inclusion de H^k dans L^2 est compacte (même argument qu'en (5)).

(7) L'application F_K de $L^2([0,1],\mathbf{C})$ dans lui-même, définie par $f \longrightarrow \int_0^1 K(s,t)f(s)ds$ (pour K continu) est compacte (voir le problème 3.1 page 180 pour la démonstration).

(8) Un opérateur compact n'est pas inversible (cf. exercice (26) page 190), donc $0 \in \mathrm{sp(A)}$.

Nous pouvons maintenant énoncer la généralisation du théorème spectral au cas d'un espace de Hilbert :

Théorème 2.0.7. *Soit A un opérateur compact hermitien d'un espace de Hilbert de dimension infinie, H, dans lui-même. Il existe alors une base Hilbertienne de H, (e_n) telle que $Ae_n = \lambda_n e_n$. Les valeurs propres λ_n sont réelles de multiplicités finies, et tendent vers 0. On a toujours $0 \in \mathrm{sp(A)}$. Si 0 n'est pas valeur propre, $\mathrm{vp(A)} = \mathrm{sp(A)} \setminus \{0\}$, sinon $\mathrm{vp(A)} = \mathrm{sp(A)}$.*

Démonstration. On vérifie comme en dimension finie que deux sous-espaces propres sont orthogonaux, et que l'orthogonal d'un espace invariant est aussi invariant. En effet, si $Ax = \lambda x$ et $Ay = \mu y$ on a $\lambda\langle x,y\rangle = \langle Ax,y\rangle = \langle x,Ay\rangle = \mu\langle x,y\rangle$, et donc si $\lambda \neq \mu$, on a $\langle x,y\rangle = 0$. Enfin si E est invariant, et $x \in E^\perp$, on aura pour tout y dans E, $\langle Ax,y\rangle = \langle x,Ay\rangle = 0$ puisque $Ay \in E$. On en déduit que Ax est aussi dans $E^\perp$, ce qui montre que $E^\perp$ est invariant .

Soit alors F l'adhérence de la somme de tous les espaces propres (i.e. l'adhérence des combinaisons linéaires finies de vecteurs propres). Par définition, c'est un sous-espace vectoriel fermé de H et donc un espace de Hilbert. On prétend que $F^\perp = 0$. Comme $F^\perp$, est invariant par A, et par définition $A_{|F^\perp}$ ne possède pas de sous-espace propre, le lemme suivant permettra de conclure :

Lemme 2.0.8. *Soit A hermitien et compact défini sur un espace de Hilbert, E, différent de $\{0\}$. Alors $\lambda = \sup_{|x|=1}\langle Ax,x\rangle$ et $\mu = \inf_{|x|=1}\langle Ax,x\rangle$ sont valeurs propres de A.*

Démonstration. Montrons tout d'abord que le supremum est atteint. Soit $(x_n)_{n\geq 1}$ une suite d'éléments de norme 1 tels que $< Ax_n, x_n > \longrightarrow \lambda$.

Par compacité de A, on peut extraire une sous suite telle que Ax_n converge vers z. Comme

$$|\langle Ax_n - z, x_n\rangle| \leq \|Ax_n - z\| \cdot \|x_n\|$$

on en déduit que $< z, x_n >$ converge vers λ. Quitte à re-extraire une sous-suite, on peut supposer que x_n converge faiblement vers x_∞. Alors Ax_n converge faiblement [2] vers Ax_∞ et fortement vers z, ce qui entraîne que $Ax_\infty = z$, et donc $< Ax_\infty x_\infty >= \lambda$. Or si $< Ax, x >$ restreint à la sphère atteint son maximum en $\overline{x}$, on a par le théorème des extrémas liés [3] que pour tout $u \in (\mathbf{C}x_\infty)^\perp$, $< Ax_\infty, u >= 0$. On en déduit que $Ax_\infty = \lambda x_\infty$. Le même argument montre que μ est aussi une valeur propre de A. □

La démonstration du théorème résulte immédiatement de ce que par hypothèse, A n'a pas de valeur propre sur $F^\perp$. On en déduit que $F^\perp = 0$.

Montrons que l'espace propre associé à une valeur propre λ non nulle est de dimension finie. En effet soit $E_\lambda = \mathrm{Ker}(A - \lambda\,\mathrm{Id})$, et considérons plus généralement une somme d'espaces propres correspondant à une valeur propre de module supérieur à ε, que l'on note E. On a clairemnt que $A(B(0,1) \supset \varepsilon B(0,1) \cap E$ et donc par le théorème de non-compacité de Riesz, E doit être de dimension finie. On en déduit que les espaces propres sont de dimensions finies, et que la suite des valeurs propres tend vers 0. □

On dit qu'un opérateur auto-adjoint est positif si $\langle Ax, x\rangle \geq 0$ pour tout x dans H.

Proposition 2.0.9 (Caractérisation variationnelle des valeurs propres)**.** *Si l'opérateur A est compact et positif sur l'espace de Hilbert H, on a*

$$\lambda_k = \sup_{\dim(T)=k} \inf_{x\in T, \|x\|=1} \|\langle A(x), x\rangle\|$$

Démonstration. En effet, les valeurs propres $(\lambda_k)_{k\geq 1}$ sont nécessairement positives et peuvent s'écrire comme une suite décroissante tendant vers 0. On a alors une base Hilbertienne $(e_k)_{k\geq 1}$ de H, telle que $Ae_k = \lambda_k e_k$. Soit T_0 l'espace engendré par $(e_1, ..., e_k)$, on a $\|A(x_1e_1 + ... + x_ke_k)\|^2 = x_1^2\lambda_1 + ... + x_k^2\lambda_k \geq (x_1^2 + ... + x_k^2)\lambda_k$. L'expression de droite de l'énoncé est donc supérieure à λ_k. De même sur $S_0 = T_0^\perp \oplus \mathbf{C}e_k$ on a $|\langle Ax, x\rangle| \leq \lambda_k |x|^2$. Maintenant tout sous-espace de dimension k doit rencontrer S_0, car si P est la restriction à T_0 de la projection orthogonale sur $S_0^\perp$, qui est de dimension $k-1$, a nécessairement un noyau. Par définition un élément de ce noyau est un élement x de $T_0 \cap S_0$. On a alors trouvé un vecteur x de T tel que $\langle Ax, x\rangle \leq \lambda_k \|x\|^2$. L'expression de droite de l'énoncé est donc majorée par λ_k. Elle lui est donc égale. □

Exercice 2.0.10. Montrer que l'on a de manière analogue

$$\lambda_k = \inf_{\dim(T)=k} \sup_{x\in T^\perp, \|x\|=1} \|\langle A(x), x\rangle\|$$

2. En fait fortement d'après l'exercice (23) page 190.

3. En voici une démonstration directe : $x + tu$ est de norme $\sqrt{1+t^2}$ donc si le maximum de $\langle Ax, x\rangle$ sur la sphère unité est atteint en x_0, on aura $\langle A(x_0 + tu), (x_0 + tu)\rangle \leq (1+t^2)\langle Ax_0, x_0\rangle$. En développant et prenant t de module suffisamment petit, on obtient $\langle Ax, u\rangle = 0$ (cf. la démonstration du théorème de la prjection (2.1.1 page 152 pour le même calcul).

3 Compléments : Opérateur à noyau, Équations de Sturm-Liouville

3.1 Problème I : Opérateurs à noyau.

Soit $K(x,y)$ une fonction de $L^2([0,1]^2,\mathbf{C})$. On considère l'opérateur

$$f \longrightarrow \int_0^1 K(x,y)f(y)dy$$

de $L^2([0,1],\mathbf{C})$ dans lui-même. On a vu que cet opérateur est continu (voir (3) page 151 de l'exemple 2.0.17 page 151 page 151).

(1) Montrer que l'opérateur F_K est auto-adjoint si et seulement si $K(x,y)=\overline{K}(y,x)$.
On veut montrer que l'opérateur F_K est compact.

(2) Soit f_n une suite bornée de $L^2([0,1],\mathbf{C})$. Quitte à extraire une sous-suite, on peut supposer qu'elle converge faiblement vers f. Montrer en utilisant le théorème de Fubini, que pour presque tout x, on a

$$\int_0^1 K(x,y)f_n(y)dy = \langle K(x,\bullet), f_n(y)\rangle \longrightarrow \langle K(x,\bullet), f\rangle = \int_0^1 K(x,y)f(y)dy$$

(3) En déduire que $F_K f_n$ est une suite qui converge presque partout. Démontrer l'inégalité

$$|F_k f(x)| \leq\leq k(x)|||f||$$

et le fait que $k(x)\in L^2$, utiliser le théorème de convergence dominée pour conclure que $F_K f_n$ converge dans $L^2([0,1],\mathbf{C})$ vers $F_K f$.

(4) En déduire la proposition suivante
Proposition 3.1.1. *Il existe une suite (φ_n,λ_n) de solutions de*

$$\int_0^1 K(s,t)\varphi(t)dt = \lambda\varphi(s)$$

telle que φ_n soit une base Hilbertienne, et λ_n une suite de réels tendant vers 0. On a

$$K(s,t) = \sum_{n=1}^{+\infty} \lambda_n\varphi_n(s)\overline{\varphi}_n(t)$$

(5) Montrer que

$$\int_{[0,1]^2} |K(s,t)|^2 dsdt = \sum_{j=1}^{\infty} \lambda_n^2$$

(6) En déduire que les λ_n vérifient $\sum_{j=1}^{\infty}\lambda_n^2 < +\infty$, condition plus forte que la simple convergence vers 0 des λ_n.

3.2 Équation de Sturm-Liouville.

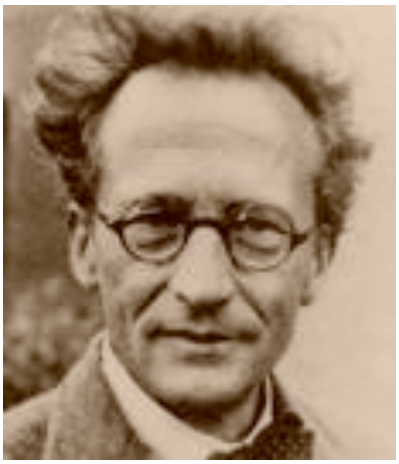




FIGURE IX.1 – Erwin Schrödinger (1887-1961) Joseph Liouville (1809-1882)

Les résultats de la section précédente permettent de résoudre les équations de type Sturm-Liouville.

Soit $q(t)$ une fonction continue sur $[0,1]$ à valeurs réelles. On cherche les couples (u,λ) tels que λ soit un nombre complexe et u une solution non identiquement nulle de l'équation

$$\text{(SLq)} \qquad \begin{cases} -u''(t)+q(t)u(t)=\lambda u(t) \\ u(0)=u(1)=0 \end{cases}$$

avec conditions aux limites $u(0)=0, u(1)=0$.

De telles équations apparaissent dans de nombreux domaines de la science. Par exemple, on peut la voir comme une équation de Schrödinger stationnaire pour une particule quantique dans un champ dérivant du potentiel $q(t)$.

Intermède culturel : Equation de Schrödinger et de Sturm-Liouville. (cf. le cours de Mécanique quantique [3], chapitre 2) *Soit $\psi(x,t)$ la fonction d'onde d'une particule quantique soumise à un potentiel $V(x)$ dépendant d'une variable unidimensionnelle x. Alors cette fonction d'onde vérifie l'équation de Schrödinger*

$$-i\hbar\frac{\partial}{\partial t}\psi(x,t)=-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x,t)+V(x)\psi(x,t)=0$$

Si on cherche $\psi(x,t)$ sous la forme $\psi(x,t)=\sum_{j=1}^{n}\psi_n(x)e^{i\frac{E_n}{\hbar}t}$, les ψ_n doivent vérifier l'équation

$$\text{(Schrödinger stationnaire)} \qquad -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi_n(x)+V(x)\psi_n(x)=E_n\psi_n(x)$$

les E_n sont les niveaux d'énergie d'une particule soumise au potentiel V. Une autre manière de voir ce problème est de considérer que l'on cherche une base Hibertienne de $L^2(\mathbf{R},\mathbf{C})$ telle que si on décompose ψ sur cette base, l'équation de Schrödinger prenne une forme simple. Or quelle

que soit la base Hilbertienne ψ_n on peut écrire

$$\psi(x,t) = \sum_{j=1}^{n} \psi_n(x) a_n(t)$$

et l'équation s'écrit alors

$$\left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi_n(x) + V(x)\psi_n(x) \right) a_n(t) = i\hbar a_n'(t)\psi_n(x)$$

Si les ψ_n vérifient l'équation de Schrödinger stationnaire, cette équation se ramène à

$$E_n a_n(t) = i\hbar a_n'(t)$$

soit

$$a_n(t) = Ce^{i\frac{E_n}{\hbar}t}$$

Si on imagine la particule contenue dans une « boîte » c'est-à-dire que l'on ajoute un potentiel nul sur $[0,1]$ et infini hors de cet intervalle, cela revient à remplacer V par

$$W(x) = \begin{cases} V(x) & si\ x \in [0,1] \\ +\infty & si\ x \notin [0,1] \end{cases}$$

On admettra que cela est équivalent à ajouter les conditions au bord $\psi_n(0) = \psi_n(1) = 0$ à l'équation précédente. On est donc ramené à un problème de Sturm-Liouville.

Théorème 3.2.1. *Soit q une fonction continue sur $[0,1]$. Les solutions (u_n, λ_n) de l'équation de Sturm-Liouville* (SLq) *forment un ensemble dénombrable. La suite λ_n est une suite de réels tendant vers $+\infty$ et les u_n forment une base Hilbertienne de $L^2([0,1], \mathbf{C})$ telle que pour tout n, u_n est de classe C^2 et solution de*

$$-u_n''(t) + q(t)u_n(t) = \lambda_n u(t)$$

avec conditions aux limites $u_n(0) = u_n(1) = 0$

Remarque 3.2.2. Les λ étant réels, on peut supposer les u_n réels, car si u est solution de l'équation de Sturm-Liouville, il en est de même pour sa partie réelle et imaginaire.

Démonstration. Nous allons supposer $1 \leq q(t) \leq A$, ce qui ne restreint pas la géneéralité du problème, car ajouter une constante à q revient à décaler λ de la même constante : si (u, λ) est solution de (SLq) , alors $(u, \lambda + C)$ est solution du même problème avec $q(t)$ remplacé par $q(t) + C$.

On pourrait considérer les espaces de fonctions $H^1([0,1], \mathbf{C})$ constitué des fonctions $L^2([0,1], \mathbf{C})$ ayant des dérivées [4] dans L^2 . Il est cependant ici plus rapide de travailler directement sur les dérivées, et de retrouver les fonctions en utilisant l'opérateur I de $L^2([0,1], \mathbf{C})$ dans lui-même défini par $I(f)(s) = \int_0^t f(s)ds$. On pose donc $L_0^2([0,1], \mathbf{C}) = \{f \in L^2 \mid \int_0^1 f(s)ds = 0\}$.

4. L'espace $H^1([0,1], \mathbf{C})$ se définit rigoureusement par la propriété des coefficients de Fourier $\sum_j (1 + j^2)|c_j(f)|^2 < +\infty$. On peut aussi poser $H_0^1([0,1], \mathbf{C}) = I(L^2([0,1], \mathbf{C}))$, et au moins formellement, $H_0^1([0,1], \mathbf{C}))$ est constitué des fonctions de $L^2([0,1], \mathbf{C}))$ à dérivée dans $L^2([0,1], \mathbf{C}))$ et s'annnulant en 0 et 1

C'est un sous-espace fermé de $L^2([0,1],\mathbf{C})$, et donc un espace de Hilbert pour le produit hermitien de L^2.

On pose alors

$$\langle \varphi, \psi \rangle_q = \langle \varphi, \psi \rangle + \langle qI(\varphi), I(\psi) \rangle$$

où q est identifié à l'opérateur de multiplication par $q(t)$, qui est continu sur $L^2([0,1],\mathbf{C})$ car q est borné (cf. exemple (1) page 151, page 151). On notera $\|\varphi\|_q^2 = \langle \varphi, \varphi \rangle_q$. Faisons quelques remarques simples :

(1) $\langle \bullet, \bullet \rangle_q$ définit une forme hermitienne et les normes $\|\varphi\|_q$ et $\|\varphi\|$ sont équivalentes. En effet de l'inégalité $1 \leq q$ résulte l'inégalité $\|\varphi\| \leq \|\varphi\|_q$. La continuité de I sur $L^2([0,1],\mathbf{C})$ et le fait que q soit bornée entraînent l'existence d'une constante C telle que $\|qI(\varphi)\| \leq C\|\varphi\|$ et donc $\|\varphi\|_q^2 \leq \|\varphi\|^2 + C\|\varphi\|\|\varphi\| \leq (C+1)\|\varphi\|^2$.

(2) Soient $\varphi, \psi \in L^2([0,1],\mathbf{C})$. On pose

$$B(\varphi, \psi) = \langle I(\varphi), I(\psi) \rangle$$

La forme hermitienne B est continue par rapport à $\langle , \rangle_q$, car I est continue sur $L^2([0,1],\mathbf{C}$, et il existe donc une constante C telle que $B(\varphi,\varphi) = \|I\varphi\|^2 \leq C\|\varphi\|^2 \leq C\|\varphi\|_q^2$.

(3) D'après l'exemple (5) page 178 de 2.0.6 page 177, I est compact, et d'après le théorème de Cauchy-Schwarz son image est dans $C^0([0,1],\mathbf{C})$. En effet, on écrit $|\int_x^y f(s)ds| \leq \left(\int_x^y |f(s)|^2 ds\right)^{1/2} (|x-y|)^{1/2}$. On en déduit que l'image de I est contenue dans $C^0([0,1],\mathbf{C})$.

(4) L'opérateur I est anti-hermitien sur $L^2([0,1],\mathbf{C})$. En effet, en intégrant par parties, et supposant f, g continues, on a $[I(f)]'(t) = f(t)$ d'où en utilisant que pour f dans $L^2([0,1],\mathbf{C})$ $I(f)(1) = I(f)(0) = 0$

$$\langle I(f), g \rangle = \int_0^1 I(f)(t)g(t)dt = [I(f)(1)I(g)(1) - I(f)(0)I(g)(0)] - \int_0^1 f(t)I(g)(t)dt = \langle f, I(g) \rangle$$

Par densité de $C^0([0,1],\mathbf{C})$ dans $L^2([0,1],\mathbf{C})$, la propriété ci-dessus s'étend à tout $L^2([0,1],\mathbf{C})$.

(5) Si $f = I(\varphi)$ et $g = I(\psi)$ sont de classe C^1 on aurait

$$\langle \varphi, \psi \rangle_q = \int_0^1 [\bar{f}'(t)g'(t) + q(t)\bar{f}(t)g(t)]dt$$

Soit alors $L : L_0^2([0,1],\mathbf{C}) \longrightarrow L_0^2([0,1],\mathbf{C})$ l'opérateur donné par le théorème de représentation de Riesz (théorème III-Riesz-sesquilineaire), défini par

$$B(\varphi, \psi) = \langle L\varphi, \psi \rangle_q$$

Comme B est hermitienne et positive, il en est de même pour L. Reste à montrer que l'opérateur L est compact. Or la formule définissant L se récrit, en utilisant le seul produit scalaire L^2 et la remarque (4) sous la forme

$$\langle -I(I(\varphi)), \psi\rangle = \langle L(\varphi), \psi\rangle - \langle I(qI(L(\varphi))), \psi\rangle$$

Puisque cela est vrai pour tout ψ de moyenne nulle, il existe une constante c telle que

$$-I(I(\varphi)) = L(\varphi) - I(qI(L(\varphi))) + c$$

Bien entendu la constante c est uniquement déterminée par la nécessité que $L(\varphi)$ soit de moyenne nulle [5].

Il résulte immédiatement de cette formule que L est compact. En effet soit φ_n une suite de fonctions de L^2_0 de norme inférieure à 1. La compacité de I entraîne l'existence d'une sous-suite telle que $I^2(\varphi_n)$ et $-I(qI(L(\varphi_n)))$ convergent (car L étant continue, $L(\varphi_n)$ est borné, donc son image par I_0 a une sous suite convergente). Mais alors $L(\varphi_n)$ projection orthogonale sur $L^2_0([0,1],\mathbf{C})$ de $-I^2(\varphi_n) + I(qI(L(\varphi_n))$ converge aussi.

En conclusion L est compact, hermitien, positif, et donc possède d'après le théorème spectral, une base Hilbertienne pour $\langle , \rangle_q$ de vecteurs propres, $L\varphi_n = \alpha_n\varphi_n$ où les α_n sont positifs et tendent vers 0.

Il reste à démontrer

Lemme 3.2.3. *Si $\varphi \in L^2_0([0,1],\mathbf{C})$ on a $L\varphi = \alpha\varphi$ si et seulement si $u = I(\varphi)$ est de classe C^2 et vérifie $-u''(t) + q(t)u(t) = \frac{1}{\alpha}u(t)$ et $u(0) = u(1) = 0$. Inversement si u est C^2 et vérifie l'équation précédente, $u'(t) = \varphi(t) \in L^2_0([0,1],\mathbf{C})$ est fonction propre de* (SLq).

Démonstration. Notons que dire que $u = I(\varphi)$ vérifie $u(0) = u(1) = 0$ est équivalent à dire que $\varphi \in L^2_0([0,1],\mathbf{C})$. Supposons $L(\varphi) = \alpha\varphi$, on a alors

$$-I^2(\varphi) = \alpha\varphi - \alpha I(qI(\varphi)) + c \qquad \text{(SLq-I)}$$

Montrons que $u \in C^2([0,1],\mathbf{C})$. En effet, $I(\varphi) \in C^0([0,1],\mathbf{C})$ donc $I^2(\varphi) \in C^1([0,1],\mathbf{C})$ et $qI(\varphi) \in C^0([0,1],\mathbf{C})$ donc $I(qI(\varphi) \in C^1([0,1],\mathbf{C})$. Il en résulte, en utilisant l'expression ci-dessus, et le fait que $\alpha \neq 0$ que $\varphi \in C^1([0,1],\mathbf{C})$. Mais alors, $u = I(\varphi) \in C^2$, et il est légitime de dériver (SLq-I)

$$-u(t) = \alpha(u''(t) - q(t)u(t))$$

Il est facile de remonter les égalités, et voir que si φ est C^1 et solution de (SLq), elle vérifie $L\varphi = \frac{1}{\lambda}\varphi$. □

On peut alors conclure la démonstration du théorème. Soit en effet φ_n une base orthonormée de vecteurs propres de L de valeur propre α_n. D'après le lemme, $(\varphi_n, \frac{1}{\alpha_n})$ sont les solutions de (SLq). Comme par ailleurs $B(\varphi_n,\varphi_m) = 0$ si $m \neq n$, cela signifie $\langle I(\varphi_n), I(\varphi_m)\rangle = 0$ si $m \neq n$, soit $\langle u_m, u_n\rangle = 0$ pour $m \neq n$. Il n'y a pas de raison pour que les u_m soient unitaires en norme L^2, mais on peut toujours les normaliser. □

5. On pourrait aussi dire que $L(\varphi)$ est la projection orthogonale sur $L^2_0([0,1],\mathbf{C})$ de $-I(I(\varphi)) + I(qI(L(\varphi)))$.

Remarques 3.2.4. (1) De nombreux systèmes orthogonaux sont obtenus comme fonctions propres de systèmes d'équations différentielles. Cependant, il s'agit souvent d'équations d'apparence un peu plus compliquées que celles que nous présentons, à savoir du type

$$\frac{d}{dx}(p(x)u'(x)) + q(x)u(x) = \lambda w(x)u(x)$$

La théorie est la même que ci-dessus, mais il faut remplacer $\langle f, g\rangle_q$ par $\langle f, g\rangle_{p,q} = \int_0^1 [p(t)f'(t)g'(t)+q(t)f(t)g(t)]dt$ et $B(f, g) = \int_0^1 \bar{f}(t)g(t)dt$ par $B_w(f, g) = \int_0^1 w(t)\bar{f}(t)g(t)dt$. Pourvu que w et p soient continues et strictement positives, la théorie n'est pas modifiée.

Les fonctions $\sin(nt)$ sont les fonctions propres correspondant à $q \equiv 0$, et donc solutions de $u'' + \lambda u = 0$, dont les valeurs propres sont les $\lambda_n = n^2$

L'équation $((1 - x^2)y')' + cy = 0$ correspond aux polynômes de Legendre, qui forment un système orthogonal pour $L^2([-1, 1], \mathbf{C})$ muni de $\langle f, g\rangle = \int_{-1}^1 \bar{f}(x)g(x)dx$.

L'équation $xy'' + (1 - x)y' + cy = 0$ correspond aux polynômes de Laguerre, qui forment un système orthogonal pour $L^2([0, +\infty], e^{-x})$ muni de $\langle f, g\rangle = \int_0^{+\infty} \bar{f}(x)g(x)e^{-x}dx$.

(2) L'opérateur qui apparaît naturellement dans le problème de Sturm-Liouville est $F : u \longrightarrow -u''(t) + q(t)u(t)$, mais cet opérateur n'est pas bien défini [6] sur L^2. C'est un exemple d'opérateur non borné. De tels opérateurs apparaissent naturellement en physique (par exemple l'opérateur de position $f(x) \longrightarrow xf(x)$ ou l'opérateur d'impulsion $f(x) \longrightarrow f'(x)$. Ici nous avons contourné le problème en remplaçant F par F^{-1} qui est compact, et dont les valeurs propres sont les inverses des valeurs propres de F.

(3) La démonstration ci-dessus, bien que dans un cadre élémentaire, met en place les principes généraux permettant de résoudre les équations aux dérivées partielles. Pour cela, il faut en général, travailler dans des espaces de Banach-du moins lorsqu'on aborde des équations non-linéaires- plutôt que Hilbert. Mais les principaux outils sont les mêmes : théorème de Riesz, utilisation de la transformation de Fourier, obtention de solutions faibles peu régulières, amélioration de leur régularité.D'autres outils comme la théorie des distributions seront vus en 2ème année.

Soit F l'opérateur $F(h)(t) = -h''(t) + q(t)h(t)$, nous avons montré que $L(f) = h$ si et seulement si $F(h) = f$. En d'autres termes, $F \circ L(f) = f$, du moins si f est suffisamment régulier. Il est alors facile de résoudre le système

$$-u''(t) + q(t)u(t) = f(t)$$

avec conditions aux limites $u(0) = 0, u(1) = 0$. En effet, on a $f(t) = \sum_n f_n\varphi_n(t)$ donc $u(t) = (Lf)(t) = \sum_n \frac{1}{\lambda_n} f_n\varphi_n(t)$. Cela peut se réécrire en utilisant que $f_n = B(\varphi_n, f) = \int_0^1 f(s)\varphi_n(s)ds$ et en posant $K(s, t) = \sum_n \frac{1}{\lambda_n}\varphi_n(s)\varphi_n(t)$, $u(t) = \int_0^1 K(s, t)f(s)ds$

3.3 Remarque finale : les espaces de Banach

Un espace de Banach est un espace vectoriel normé complet. Certains des résultats démontrés ici sont encore vrais dans les espaces de Banach, mais leur démonstration est souvent plus compliquée. Par exemple, le théorème de Hahn-Banach, la non compacité de la boule unité en

6. ni sur $H_0^1([0, 1], \mathbf{C})$, bien que L soit une application $H_0^1([0, 1], \mathbf{C}) \longrightarrow L^2([0, 1], \mathbf{C})$

dimension infine. La compacité de la boule unité pour la topologie faible reste vraie pour une large classe d'espaces de Banach (mais pas tous, par exemple $L^1(\mathbf{R}, \mathbf{C})$ ne vérifie pas cette propriété.). Mais on ne peut parler de décomposition orthogonale, ni de base hilbertienne. Cependant les espaces de Banach sont les espaces fonctionnels modernes les plus utilisés. Ils apparaissent à la fois par leur propriétés géométriques et leur propriétés analytiques, tout particulièrement dans l'étude des équations aux dérivés partielles.

3.4 Ce qu'il faut retenir

Extension de la transformation de Fourier au cas L^2

Inégalité de Heisenberg

Formule de Poisson

Opérateurs compacts. Opérateurs hermitiens.

Compacité de l'opérateur d'intégration, des opérateurs à noyau

Théorème spectral.

4 Exercices

(1) Démontrer les égalités et inégalités sur les formes hermitiennes contenues dans l'encadré page 146.

(2) Soit h une forme hermitienne sur un espace vectoriel de dimension finie H. On veut montrer qu'il existe une base dans laquelle $h(x, y) = \sum_{j=1}^{n} \bar{x}_j y_j$
   (a) Montrer le résultat si $\dim(H) = 1$.
   (b) Montrer que si H est de dimension $n \geq 1$ il existe v tel que $h(v, v) = 1$. Montrer que $(\mathbf{C}v)^\perp = \{x \in H \mid h(x, v) = 0\}$ est un espace de dimension $n-1$, et que la restriction de h à $(\mathbf{C}v)^\perp$ est une forme hermitienne.
   (c) Conclure en utilisant un raisonnement par récurrence.

(3) Démontrer que si H est de dimension finie, et F un sous-espace, $\dim(F) + \dim(F^\perp) = \dim(H)$.

(4) Soit C un convexe fermé dans un espace de Hilbert. Montrer que l'analogue du théorème de la projection reste vrai : pour tout x dans H il existe $u \in C$ unique tel que $\|x-u\| = d(x, C)$ et que l'on a $\langle x - u, u - v\rangle \leq 0$ quel que soit v dans C.

(5) Montrer que si E est un espace préhilbertien, E^* espace des formes linéaires continues sur E s'identifie au complété de E. Pour cela on pourra montrer que
   (a) l'application $a \mapsto \langle a, \bullet\rangle$ fournit une injection continue de E dans E^*
   (b) E^* est complet
   (c) E est dense dans E^*
   (d) la norme sur E^* induit la norme usuelle sur E, et comme cette dernière est induite par une forme hermitienne, il en est de même pour celle sur E^*

(6) Quel est le complété de l'espace préhilbertien $C_0^0(\mathbf{R})$ espace des fonctions à support compact sur $\mathbf{R}$ pour $\langle f, g\rangle = \int_0^1 \bar{f}(t)g(t)dt$?

(7) (« Frames ») Soit e_k une famille de vecteurs d'un espace de Hilbert H.
On cherche une condition nécessaire et suffisante pour qu'il existe une famille v_k de vecteurs telle que pour tout x on ait

$$\sum_{j=1}^{\infty} \langle v_k, x\rangle e_k$$

(a) Montrer qu'une telle condition est donnée par l'existence de constantes positives A, B telle que pour tout vecteur x on a

$$A\|x\|^2 \leq \sum_{j=1}^{\infty} \langle x, e_k\rangle^2 \leq b\|x\|^2$$

(b) On dit qu'un « frame » est serré si $A = B = 1$. Donner un exemple de « frame » serré qui ne soit pas une base Hilbertienne. Montrer que si de plus e_k est de norme 1 pour tout k, alors le « frame » est une base Hilbertienne.

(8) Montrer que la construction du complété, faite pour un espace de Hilbert, peut se faire pour tout espace vectoriel normé, et même pour tout espace métrique. Quel est le complété de l'espace vectoriel normé $C_0^0(\mathbf{R})$ pour $\|f\| = \int_0^1 |f(t)|dt$

(9) Montrer que si les u_n sont des réels positifs tendant vers 0, le produit $\prod_{j=1}^{+\infty}(1-u_n)$ converge vers un réel non nul si et seulement si la série $\sum_{j=1}^{+\infty} u_n$ diverge.
Indication : Montrer que $\log(\prod_{j=1}^{n}(1-u_k)) \simeq -\sum_{j=1}^{n} u_k$

(10) Soit H un espace de Hilbert, et $(e_n)_{n\geq 1}$ une base Hilbertienne. Soit $(a_n)_{n\geq 1}$ une suite à termes réels positifs, telle que $\sum_0^{+\infty} |a_n|^2$ converge. Montrer que l'ensemble $\{x \in H \mid x = \sum_0^{+\infty} x_n e_n, |x_n| \leq a_n\}$ est un compact de H.

(11) Soit l'espace de Hilbert $H = L^2(\mathbf{R}, e^{-x^2}dx)$. On veut montrer que les polynômes forment un sous-espace vectoriel dense de H.

(a) Soit $F(z) = \int_{\mathbf{R}} f(t)e^{zt-t^2}dt$. Montrer que F est holomorphe et à un développement en série entière

$$F(z) = \sum_{j=0}^{+\infty} \frac{z^j}{j!} \int_{\mathbf{R}} f(t)e^{-t^2}dt$$

(b) On suppose que f est orthogonale à l'espace engendré par les polynômes. En déduire que $F \equiv 0$ et donc que $F(ix) = 0$ quel que soit x.

(c) Montrer que $F(ix)$ est la transformée de Fourier de $f(t)e^{-t^2}$ et en déduire que $f = 0$.

(d) Conclure.

(12) Soit l'espace de Hilbert $H = L^2(\mathbf{R}, e^{-x^2}dx)$. On considère les polynômes de Hermite $H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n}(e^{-x^2})$. et les fonctions de Hermite $\psi_n(x) = e^{\frac{-x^2}{2}} H_n(x)$.

(a) Montrer que H_n est un polynôme de degré n et qu'il est orthogonal (pour le produit hermitien de H) à l'espace engendré par les polynômes de degré inférieur ou égal à $n-1$.

(b) Montrer que $H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x)$ et que $\frac{d}{dx}H_n(x) = 2nH_{n-1}(x)$

(c) Montrer que $(\frac{d}{dx}+x)\psi_n(x) = 2n\psi_{n-1}$ et que $(-\frac{d}{dx}+x)\psi_n = \psi_{n+1}$

(d) Montrer enfin que $(-\frac{d^2}{dx^2}+x^2)\psi_n(x) = (2n+1)\psi_n(x)$

(e) Montrer que ψ_n est une fonction propre pour la transformée de Fourier (on pourra calculer l'équation différentielle satisfaite par $\widehat{\psi_n}$).

(13) Montrer que la construction du complété de E dans l'exercice (5) page 186, on a en fait $H = E^*$. Indication : utiliser que $H = H^*$.

(14) On veut montrer que si f est de classe C^1 sur $\mathbf{S}^1$, les sommes partielles de Fourier convergent uniformément vers f. On pose $D_n(t) = \sum_{p=-n}^{n} e^{ipt} = \frac{\sin(n+1/2)t}{\sin(t/2)}$

(a) Montrer que si f est de classe C^0, $|c_n(f)| \leq \frac{2\|f\|_{C^0}}{n}$

(b) Montrer que la somme partielle de Fourier est donnée par $s_n(f)(s) = \int_{\mathbf{S}^1} f(t)D_n(t-s)dt = \int_{\mathbf{S}^1} f(t-s)D_n(t)dt$

(c) Montrer que $f(s) - s_n(f)(s)$ est donnée par $\int_{\mathbf{S}^1} \frac{f(s-t)-f(s)}{\sin(t/2)} \sin((n+1/2)t)dt$ et que ceci s'écrit comme somme de deux coefficients de Fourier de $\frac{f(s-t)-f(s)}{\sin(t/2)} = g_t(s)$

(d) Montrer que si f est de classe C^1, g_t est continue et $\|g_t\|_{C^0}$ est borné uniformément en t.

(e) Conclure

(15) Montrer que $H^1(\mathbf{S}^1, \mathbf{C})$ est dense dans $L^2(\mathbf{S}^1, \mathbf{C})$ et que $L^2(\mathbf{R}, \mathbf{C})$ est dense dans $L^2(\mathbf{R}, e^{-x^2}dx)$

(16) (Exercice à faire après le théorème des résidus)

(a) Montrer que la transformé de Fourier de $e^{-a|x|}$ est $\frac{2a}{a^2+4\pi^2n^2}$

(b) Appliquer la formule de Poisson, pour démontrer que $\sum_{n\in\mathbf{Z}} e^{-a|n|} = \sum_{n\in\mathbf{Z}} \frac{2a}{a^2+4\pi^2n^2}$

(c) Calculer explicitement $\sum_{n\in\mathbf{Z}} e^{-a|n|}$. En faisant tendre a vers 0 en déduire que $\sum_{n\geq 1} \frac{1}{n^2} = \frac{\pi^2}{6}$.

(17) Soit H un espace de Hilbert. On veut montrer que si la suite $(x_k)_{k\geq 1}$ converge faiblement dans H, alors elle est bornée. Supposons en effet que ce ne soit pas le cas : alors $x_k \rightharpoonup \bar{x}$ et x_k tend vers l'infini. Quitte à soustraire $\bar{x}$ on peut même supposer $x_k \rightharpoonup 0$ et $|x_k|$ tend vers l'infini. On suppose donc que $x_k \rightharpoonup 0$ mais que sa norme tend vers $+\infty$.

(a) On veut construire par récurrence une base Hilbertienne (éventuellement d'un sous-espace de H) (e_k) vérifiant la condition suivante : pour tout N il existe $k(N) > k(N-1)$ tel que $|\langle x_{k(N)}, e_N\rangle| > C_N$ où $\lim_N C_N = +\infty$.

Supposons la base construite jusqu'au N-ième terme et construisons là au terme suivant. Soit $F_N =< e_1, ..., e_N >$ comme le produit scalaire des x_k avec $e_1, ..., e_N$ tend vers 0, la projection de x_k sur $F_N^{\perp}$ pour k assez grand sera de norme supérieure à $N+1$. Soit $C_{N+1}e_{N+1}$ la projection d'un tel x_k, que nous notons $x_{k(N)}$, où e_{N+1} est de norme 1 et $C_{N+1} > N+1$.

(b) Montrer que $z = \sum_{p=1}^{+\infty} \frac{1}{C_N} e_N$ est convergente dans H et que $\langle z, x_k\rangle$ ne tend pas vers 0.

Il s'agit d'un cas particulier du théorème de Banach-Steinhaus, qui affirme que si L_n est une suite d'applications linéaires continues entre espaces de Banach, soit elle est uniformément bornée, c'est-à-dire qu'il existe C tel que $\|L_n\| \leq C$, soit il existe x tel que $\|L_n x\| \longrightarrow +\infty$.

(18) Montrer que la convergence faible n'est pas induite par une métrique, c'est-à-dire qu'il n'existe pas de métrique sur H telle que

$$x_n \rightharpoonup x \Longleftrightarrow d(x_n, x) \longrightarrow 0$$

Indication : Montrer d'abord que si c'était le cas, nous aurions $\varepsilon_n = d(e_n, 0) \longrightarrow 0$. Mais alors $\sqrt{\varepsilon}_n^{-1} e_n$ tendrait aussi vers 0, et cela signifierait que pour toute suite de réels $(x_k)_{k\geq 1})$ telle que $\sum_{k=1}^{\infty} x_k^2 < +\infty$ on a $\sqrt{\varepsilon}_k^{-1} x_k \longrightarrow 0$. Trouver un contre-exemple à cette dernière assertion.

(19) Soit K un compact d'un espace de Hilbert. Montrer que pour tout $\varepsilon > 0$ il existe un sous-espace F de dimension finie de H tel que $\sup_{x\in K} \|x - P_F(x)\| = \sup_{x\in K} d(x, F) \leq \varepsilon$.

Indication : utiliser un recouvrement fini de K par des boules de rayon ε.

Montrer qu'inversement si F est un fermé borné tel que pour tout $\varepsilon > 0$ il existe un sous-espace de dimension finie tel que pour tout x de F, $d(x, F) \leq \varepsilon$, alors F est compact.

En déduire que si A est un opérateur compact, pour tout $\varepsilon > 0$ il existe un opérateur de rang fini, A_ε tel que $\|A - A_\varepsilon\| \leq \varepsilon$.

Indication : prendre pour $K = A(B(0, 1))$ construire F comme dans la première question et considérer $P_F \circ A$. En déduire que A est limite d'opérateurs de rang fini.

(20) Soit $\mathcal{S}$ l'espace des fonctions de $C^\infty(\mathbf{R}, \mathbf{C})$ à décroissance rapide (appelé aussi « espace de Schwartz »), c'est-à-dire vérifiant pour tous p, q entiers positifs

$$\lim_{|x|\to+\infty} |x^p \frac{d^q}{dx^q} f(x)| = 0$$

Cet espace est un sous-espace de $L^2(\mathbf{R}, \mathbf{C})$ (et aussi de $L^p(\mathbf{R}, \mathbf{C})$ quel que soit l'entier p), car une fonction de $\mathcal{S}$ vérifie en particulier $|f(x)| \leq \frac{C}{(1+|x|^2)}$ pour une certaine constante C. C'est donc une fonction de carré intégrable.

(a) En utilisant le fait que la transformation de Fourier échange dérivation et multiplication par $i\xi$ montrer que la transformée de Fourier d'une fonction de $\mathcal{S}$ est encore dans $\mathcal{S}$. On

utilisera le fait que $\xi^p \frac{d^q}{d\xi^q} \widehat{f}(\xi)$ est, à une constante multiplicative près, la transformée de Fourier de $\frac{d^q}{dx^p}(x^q f(x))$, et on utiliser le théorème de Riemann-Lebesgue.

(b) Montrer que le même argument montre que la transformation de Fourier inverse envoie aussi $\mathcal{S}$ dans lui-même.

(c) En déduire que la transformation de Fourier réalise une bijection de $\mathcal{S}$ dans lui-même. L'injectivité résulte bien entendu de l'injectivité dans le cas de $L^2(\mathbf{R}, \mathbf{C})$. Pour la surjectivité, on pourra utiliser le fait que la composée de la transformation de Fourier et de la transformation de Fourier inverse soit l'identité.

(21) Soit

$$W := \left\{ f \in L^2(\mathbf{R}, \mathbf{C}) \mid \int_{\mathbf{R}} (1 + x^2)|f(x)|^2 dx + \int_{\mathbf{R}} \xi^2 |\widehat{f}(\xi)|^2 d\xi \right\}$$

muni du produit hermitien

$$\langle f, g\rangle_W = \int_{\mathbf{R}} (1 + x^2)\overline{f}(x)g(x)dx + \int_{\mathbf{R}} \xi^2 \overline{\widehat{f}(\xi)}\widehat{g}(\xi)d\xi$$

(a) Montrer que si f, g sont C^2, on a

$$\langle f, g\rangle_W = \langle (1+x^2)f - f'', g\rangle$$

où $\langle \bullet, \bullet\rangle$ est le produit hermitien L^2.

(b) Montrer que si $H_n(x)$ est le n-ième polynôme de Hermite (cf. exercice (12) page 187), la famille $\psi_n(x) = C_n e^{-\frac{x^2}{2}} H_n(x)$ est une base Hilbertienne de $L^2(\mathbf{R}, \mathbf{C})$, où les C_n sont des constantes de normalisation que l'on ne calculera pas.
On rappelle que les ψ_n vérifient l'équation

$$\psi_n''(x) + (2n + 1 - x^2)\psi_n(x) = 0$$

et forment une base Hilbertienne de W.

(c) Montrer que les fonctions ψ_n sont dans $\mathcal{S}$, et que $\mathcal{S}$ est donc dense dans W.

(d) Déduire de ce qui précède que l'inégalité de Heisenberg est vérifiée pour toute fonction de W, et donc pour toute fonction de L^2 (on remarquera que si une fonction f de $L^2(\mathbf{R}, \mathbf{C})$ n'est pas dans W, l'un des facteurs du terme de gauche de l'inégalité est infini, et que l'autre n'est jamais nul sauf si $f = 0$)

(22) Montrer en utilisant l'exercice (19) page précédente que A est un opérateur compact si et seulement si il est limite d'opérateurs de rang fini.

(23) Montrer qu'un opérateur est compact si et seulement si toute suite faiblement convergente est envoyée sur une suite convergente (au sens usuel).

(24) Montrer que la somme de deux opérateurs compacts est un opérateur compact.

(25) Soit l'opérateur compact $T_K : f \longrightarrow \int_0^1 K(s,t)f(s)ds$. Trouver une suite d'opérateurs de rang fini approchant T_K. On pourra discrétiser et approcher l'intégrale par $\frac{1}{N}\sum_{j=1}^{N} f(j/N)K(j/N, t)$.

(26) Montrer qu'un opérateur compact n'est jamais inversible.

(27) (a) On veut montrer que l'espace $H^1([0, 2\pi], \mathbf{C})$ est inclus dans $C^0([0, 2\pi], \mathbf{C})$. Pour cela, montrer que si les coefficients de Fourier de f, $c_k(f)$ vérifient $c_0(f)^2 + \sum_{k=-\infty}^{\infty} k^2 c_k(f)^2 < +\infty$, la série de Fourier de f, $\sum_{k=-\infty}^{\infty} c_k(f)e^{ikt}$ est normalement convergente, et a donc pour limite une fonction continue.

(b) En déduire que l'application $ev_t : H^1([0, 2\pi], \mathbf{C}) \longrightarrow \mathbf{C}$ donné par $f \mapsto f(t)$ est bien définie et continue.

(c) Montrer que le sous-espace $H^1_0([0, 2\pi], \mathbf{C})$ de H^1 constitué des éléments s'annulant en $t = 0$ et $t = 2\pi$ est bien défini et qu'il est fermé.

(28) Soit E l'espace de Hilbert des fonctions $L^2([0, 1]^2, \mathbf{C})$ telles que $K(s,t) = \overline{K}(t,s)$. Soit F les sous-espace constitué par les combinaisons linéaires de fonctions de la forme $\overline{f}(s)f(t)$. Montrer que F est dense dans E.

(29) (Zéros des fonctions propres) Soit φ_1 la fonction propre correspondant à la plus petite valeur propre d'un opérateur de Sturm-Liouville. Montrer que φ_1 ne s'annule pas sur $]0, 1[$.

– En utilisant le théorème de Cauchy-Lipschitz de la théorie des équations différentielles linéaires, montrer que si $\varphi_1(t_0) = 0$ alors $\varphi_1'(t_0) \neq 0$.
– Montrer que φ_1 réalise le minimum de $\int_0^1 [|f'(t)|^2 + q(t)|f(t)|^2]dt$ parmi les fonctions telles que $\|f\|_{L^2} = 1$. En déduire que si φ_1 est fonction propre, alors $|\varphi_1|$ aussi.

– Conclure en utilisant le fait qu'une fonction propre est de classe C^1 (en fait même C^2).

(30) (Zéros des fonctions propres-suite) Soit φ_k la fonction propre associée à la k-ième valeur propre. Montrer en utilisant la relation $\int_0^1 \bar{\varphi}_1(t)\varphi_2(t)dt = 0$ que φ_2 possède au moins un zéro sur $]0,1[$.

Chapitre X

Fonctions holomorphes d'une variable complexe

1 Introduction

L'objet de ce chapitre est d'introduire le lecteur à la théorie des fonctions holomorphes d'une variable complexe.



FIGURE X.1 – Augustin Louis Cauchy, un des fondateurs de la théorie des fonctions holomorphes

Le but de ce chapitre est de présenter au lecteur la théorie des fonctions holomorphes. Comme on le verra rapidement, cette notion coïncide avec celle de fonction développable en série entière sur $\mathbf{C}$. Ces fonctions sont omniprésentes, en mathématiques comme en physique (cf.[14, paragraphe 82] ou les références de la note 5 page 218). L'une des raisons de leur présence en physique est que les fonctions harmoniques, c'est-à-dire les fonctions de Laplacien nul, sont localement les parties réelles des fonctions holomorphes, et que les potentiels apparaissant en physique satisfont bien souvent cette propriété.

Nous n'avons recherché ni l'exhaustivité ni les hypothèses minimales dans les théorèmes. Par exemple, il est possible de définir les fonctions holomorphes comme les fonctions dérivables au sens complexe, sans faire d'hypothèse de continuité des dérivées. Le théorème de Morera, fort

utile pour des applications fines, affirme que cela ne change rien aux objets ainsi définis. On verra dans les cours de mathématiques ultérieurs (MAT 431) que l'on peut même s'affranchir totalement des hypothèses de régularité dans le cadre de la théorie des distributions.

Le lecteur intéressé est encouragé à consulter l'abondante littérature [6], [1], [19], [13]. Pour la théorie à plusieurs variables, qui ont des propriétés surprenantes on consultera [17] et [10].

Si ce cours avait pu se poursuivre, nous aurions entraîné le lecteur vers les merveilles de la géométrie complexe, qu'elle soit différentielle ou algébrique, monde des fonctions holomorphes de plusieurs variables par excellence qu'a largement ouvert l'immense mathématicien Bernhard Riemann (cf. (8.0.14 page 224)). C'est un sujet d'intense activité à l'heure actuelle, objet de l'un des problèmes Clay par la « conjecture de Hodge » dont la solution récompenserait son auteur de bien plus que du million de dollars promis par la fondation : c'est une question centrale du domaine qui atteste s'il le fallait la vitalité du sujet.



FIGURE X.2 – William Hodge

2 Notion d'holomorphie-Premières propriétés

2.1 Notations et rappels sur les fonctions de deux variables

On identifiera $\mathbf{C}$ à $\mathbf{R}^2$ grâce à l'application

$$z \mapsto x + iy,\ x = \mathrm{Re}(z), y = \mathrm{Im}(z)$$

(l'écriture $z = x + iy$ sous-entendra $x, y \in \mathbf{R}$). Le symbole Ω désignera un ouvert non vide de $\mathbf{C}$. Ainsi, une fonction à valeurs complexes $f : \Omega \longrightarrow \mathbf{C}$ est de classe C^1 si la fonction de deux variables, qu'on notera

$$\tilde{f} : (x, y) \mapsto f(x + iy)$$

est C^1 sur Ω. Par définition, ceci signifie qu'elle admet des dérivées partielles

$$\frac{\partial f}{\partial x}(z_0) := \frac{\partial \tilde{f}}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(z_0) := \frac{\partial \tilde{f}}{\partial y}(x_0, y_0)$$

continues en tout point $z_0 = x_0 + iy_0 \in \Omega$. La formule de Taylor fournit alors un développement limité à l'ordre 1 en $z_0 \in \Omega$ de la forme

(2.1.a) $$f(z_0+z)=f(z_0)+x\frac{\partial f}{\partial x}(z_0)+y\frac{\partial f}{\partial y}(z_0)+o(|z|) \text{ avec } z=x+iy.$$

Il est très commode de réécrire ce développement limité uniquement en termes des variables $z,\bar{z}$.

2.2 Les opérateurs $\frac{\partial}{\partial z}, \frac{\partial}{\partial \bar{z}}$

Définition 2.2.1. Soit f une fonction sur Ω admettant des dérivées partielles et $z_0\in\Omega$. On pose

(2.2.a) $$\frac{\partial f}{\partial z}:=\frac{1}{2}(\frac{\partial f}{\partial x}-i\frac{\partial f}{\partial y}) \text{ et } \frac{\partial f}{\partial \bar{z}}:=\frac{1}{2}(\frac{\partial f}{\partial x}+i\frac{\partial f}{\partial y}).$$

Par linéarité des dérivées partielles, on a

(2.2.b) $$\overline{\frac{\partial f}{\partial z}}=\frac{\partial \bar{f}}{\partial \bar{z}}.$$

Par linéarité encore, on obtient immédiatement le formulaire habituel

(2.2.c) $$\partial(f+g)=\partial f+\partial g, \partial(fg)=f(\partial g)+(\partial f)g, \partial(1/f)=-\frac{\partial f}{f^2}.$$

pour $\partial=\frac{\partial}{\partial z},\frac{\partial}{\partial \bar{z}}$.

Pour la composition, il s'agit de dériver un composé de fonctions de deux variable.

Lemme 2.2.2. *On a*

$$\partial(f\circ g)=\partial g\times\frac{\partial f}{\partial z}\circ g+\partial\bar{g}\times\frac{\partial f}{\partial \bar{z}}\circ g$$

pour $\partial=\frac{\partial}{\partial z},\frac{\partial}{\partial \bar{z}}$.

Démonstration. On écrit

$$f\circ g=\tilde{f}(\frac{g+\bar{g}}{2},\frac{g-\bar{g}}{2i})$$

et on dérive la fonction composée de manière habituelle

$$\begin{aligned}\partial(f\circ g) &= \partial(\frac{g+\bar{g}}{2})\times\frac{\partial\tilde{f}}{\partial x}\circ g+\partial(\frac{g-\bar{g}}{2i})\times\frac{\partial\tilde{f}}{\partial y}\circ g\\ &= \partial g\times\frac{1}{2}(\frac{\partial\tilde{f}}{\partial x}\circ g-i\frac{\partial\tilde{f}}{\partial y}\circ g)+\partial\bar{g}\times\frac{1}{2}(\frac{\partial\tilde{f}}{\partial x}\circ g+i\frac{\partial\tilde{f}}{\partial y}\circ g)\\ &= \partial g\times\frac{\partial f}{\partial z}\circ g+\partial\bar{g}\times\frac{\partial f}{\partial\bar{z}}\circ g.\end{aligned}$$

□

On réécrit alors (2.1.a page précédente) avec ces opérateurs et on obtient après un calcul élémentaire

$$f(z_0+z) = f(z_0) + z\frac{\partial f}{\partial z}(z_0) + \bar{z}\frac{\partial f}{\partial \bar{z}}(z_0) + o(|z|). \tag{2.2.d}$$

2.3 Holomorphie : définition et stabilité

Définition 2.3.1 (Condition de Cauchy complexe). Une fonction $f : \Omega \longrightarrow \mathbf{C}$ de classe C^1 est holomorphe sur Ω si

$$\forall z_0 \in \Omega,\ \frac{\partial f}{\partial \bar{z}}(z_0) = 0.$$

Dans ce cas, on note $f'(z_0) := \frac{\partial f}{\partial z}(z_0)$ la **dérivée** de f. On note $\mathrm{Hol}(\Omega)$ l'ensemble des fonctions holomorphes sur Ω.

D'après la formule de Taylor (2.2.d), une fonction de classe C^1 sur Ω est holomorphe si elle admet un développement limité en tout $z_0 \in \Omega$ de la forme

$$f(z_0+z) = f(z_0) + \lambda.z + o(|z|) \tag{2.3.a}$$

(avec alors $\lambda = \frac{\partial f}{\partial z}(z_0)$). Autrement dit, f de classe C^1 est holomorphe si et seulement si la limite

$$\lim_{z \longrightarrow 0} \frac{f(z_0+z) - f(z_0)}{z}. \tag{2.3.b}$$

existe en tout point, **i.e.**si f est $\mathbf{C}$-dérivable.

Les propriétés de stabilité par les opérations usuelles des fonctions C^1 et les formules (2.2.c page précédente) et (2.2.2 page précédente) donnent, comme dans le cas des fonctions de variable réelles, la proposition usuelle

Proposition 2.3.2. *Les combinaisons linéaires, produits et compositions de fonctions holomorphes sont holomorphes. Si f est holomorphe, alors $1/f$ est holomorphe sur l'ouvert où elle est non nulle. Soit $f \in \mathrm{Hol}(\Omega)$ et $g \in \mathrm{Hol}(\Omega')$. On suppose $g(\Omega') \subset \Omega$. Alors, $f \circ g \in \mathrm{Hol}(\Omega)$.*

Bien entendu, les dérivées complexes se calculent avec les formules habituelles.

Exercice 2.3.3. Soit f une fonction mesurable à support dans $\mathbf{R}^+$ et appartenant à $L^\infty(\mathbf{R}^+)$. Sa transformée de Fourier est dans $\mathrm{Hol}(\mathbf{H}^-)$, $\mathbf{H}^- = \{\mathrm{z}|\ \mathrm{Im(z)} < 0\}$.

Exercice 2.3.4. Une série entière définit une fonction $f(z) = \sum_{n\geq 0} a_n z^n$ holomorphe sur son disque ouvert de convergence, tandis que sa conjuguée $\overline{f(z)} = \sum \bar{a}_n \bar{z}^n$ ne l'est pas dès que f est non-constante.

Remarque 2.3.5. L'exercice précédent montre que bon nombre des fonctions de la variable complexe usuelles sont holomorphes : les fractions rationnelles, les fonctions construites à base d'exponentielles comme les fonctions trigonométriques ou hyperboliques etc... Mais attention, si f holomorphe, $|f|$ ou $\bar{f}$ ne sont essentiellement jamais holomorphes : si f est non holomorphe,

$$\frac{\partial f}{\partial z} = 0 \not\Rightarrow f \text{ constante.}$$

Exercice 2.3.6. Supposons Ω connexe (par arcs). Montrer que f et $\bar{f}$ sont holomorphes si et seulement si f est constante. Montrer que si $f \in \mathrm{Hol}(\Omega)$ vérifie $f' = 0$, alors f est constante. Plus généralement, caractériser les fonction C^1 telles que $\frac{\partial f}{\partial \bar{z}} = 0$.

2.4 Les conditions de Cauchy réelles

Soit $f \in C^1(\Omega)$ et écrivons

$$f = P + iQ,\ P, Q \text{ réelles.}$$

On a

$$\frac{\partial f}{\partial \bar{z}} = \frac{1}{2}(\frac{\partial P}{\partial x} - \frac{\partial Q}{\partial y}) + \frac{i}{2}(\frac{\partial P}{\partial y} + \frac{\partial Q}{\partial x}).$$

On a donc la proposition

Proposition 2.4.1 (Conditions de Cauchy réelles)**.** *Avec les notations précédentes, $f = P + iQ \in C^1(\Omega)$ est holomorphe si et seulement si*

$$\frac{\partial P}{\partial x} = \frac{\partial Q}{\partial y} \text{ et } \frac{\partial P}{\partial y} = -\frac{\partial Q}{\partial x}.$$

.

Rappelons la caractérisation des similitudes directes du plan complexe. Ce sont les applications $\mathbf{R}$-linéaires de $\mathbf{C} = \mathbf{R}^2$ qui conservent les angles orientés. Sous forme complexe, ce sont les transformations $z \mapsto \lambda z$ où λ est le rapport de la similitude. Dans la base $(1, i)$ de $\mathbf{C} = \mathbf{R}^2$), ce sont les matrices de la forme

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}, a, b \in \mathbf{R}.$$

On passe de l'un à l'autre des points de vue grâce à la relation $\lambda = a + ib$.

Proposition 2.4.2. *$f \in C^1(\Omega)$ est holomorphe si et seulement si sa matrice jacobienne en tout point $z_0 \in \Omega$ est une similitude directe de rapport $f'(z_0)$ du plan $\mathbf{C} = \mathbf{R}^2$.*

Démonstration. Vue comme fonction de deux variables à valeurs dans $\mathbf{C} = \mathbf{R}^2$, les conditions de Cauchy équivalent à ce que la matrice jacobienne

$$J_{z_0}(f) = \begin{pmatrix} \frac{\partial P}{\partial x} & \frac{\partial P}{\partial y} \\ \frac{\partial Q}{\partial x} & \frac{\partial Q}{\partial y} \end{pmatrix}$$

soit une similitude directe. Dans ce cas, son rapport de similitude λ est

$$\frac{\partial P}{\partial x} + i\frac{\partial Q}{\partial x} \overset{Cauchy}{=} \frac{1}{2}(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}) + \frac{i}{2}(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}) = f'(z_0),$$

ce qu'on aurait pu déduire directement du développement limité (2.3.a page précédente). □

Corollaire 2.4.3. *Soit $f \in \mathrm{Hol}(\Omega)$ et z_0 tel que $f'(z_0) \neq 0$. Alors, il existe un ouvert $\Omega_1 \subset \Omega$ contenant z_0 tel que*

1) $\Omega_2 := f(\Omega_1)$ est ouvert dans $\mathbf{C}$.

*2) La restriction de f à Ω_1 est un biholomorphisme de Ω_1 **sur** Ω_2, c'est-à-dire une bijection Ω_1 sur Ω_2 dont l'inverse g est holomorphe.*

3) En particulier, « le logarithme complexe » est bien défini et holomorphe au voisinage de tout point de $\mathbf{C}^*$.

Démonstration. Comme la jacobienne J_{z_0} de f en z_0 est une similitude de rapport non nul, c'est un isomorphisme. D'après le théorème d'inversion locale, il existe un ouvert $\Omega_1 \subset \Omega$ contenant z_0 tel que

1) $\Omega_2 := f(\Omega_1)$ est ouvert dans $\mathbf{C}$.

2) La restriction de f à Ω_1 est un C^1-difféomorphisme de Ω_1 **sur** Ω_2, c'est-à-dire une bijection Ω_1 sur Ω_2 dont l'inverse g est C^1.

En particulier, pour tout $z \in \Omega_1$, la jacobienne de g en $f(z)$ est l'inverse de la jacobienne de f en z (ce qui impose d'ailleurs $f'(z) \neq 0$ pour tout $z \in \Omega_1$). C'est donc une similitude (de rapport $1/f'(z)$) ce qui prouve l'holomorphie de g d'après la proposition précédente.

Le 3) découle de 2) appliqué à $f = \exp$. □

Corollaire 2.4.4 (du corollaire)**.** *On a* $\exp(\mathbf{C}) = \mathbf{C}^*$.

Démonstration. En effet, comme la dérivée de l'exponentielle ne s'annule jamais, l'image I est un voisinage de chacun de ses points donc est ouverte dans $\mathbf{C}^*$. Mais la propriété $\exp(z_1 + z_2) = \exp(z_1)\exp(z_2)$ montre que I est un sous-groupe de $\mathbf{C}^*$. Donc si $z \in \mathbf{C}^* - I$, on a évidemment $Iz \subset\subset \mathbf{C}^* - I$ car si on avait $u \in Iz \cap I$ on peut écrire $uz = v$ avec $u, v \in I$ et donc $z = u^{-1}v \in I$

Mais Iz est ouvert comme I de sorte que le complémentaire $\mathbf{C}^* - I = \bigcup_{z \in \mathbf{C}\setminus I} Iz$ est ouvert comme réunion d'ouverts.Son complémentaire I est donc fermé et. ainsi I est ouvert et fermé dans $\mathbf{C}^*$ (et non vide). Comme $\mathbf{C}^*$ est connexe par arcs donc connexe, on doit avoir $I = \mathbf{C}^*$. □

On reviendra plus bas sur les question de définition globales du logarithme complexe (cf. 6.4 page 223).

2.5 Transformations conformes

Soit $f \in \mathrm{Hol}(\Omega)$. On s'intéresse au comportement de f au voisinage d'un point z_0 tel que $f'(z_0) \neq 0$.

Remarque 2.5.1. On verra plus tard (6.1.3 page 220) que, pourvu que f ne soit pas localement constante, les points d'annulation de f' sont isolés et que cette hypothèse n'est donc pas très restrictive.

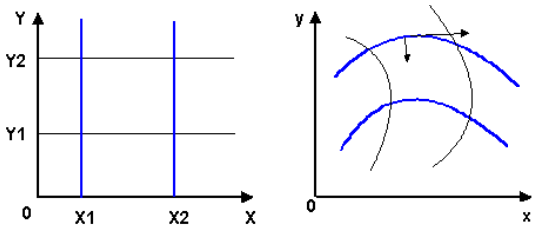
D'après le corollaire 2.4.3 page précédente, on peut supposer que f est un biholomorphisme de Ω sur son image Ω', d'inverse g qui est donc holomorphe. Considérons deux arcs C^1

$$\gamma_1, \gamma_2 : \mathbf{R} \longrightarrow \mathbf{C}$$

passant par z_0, à $t = 0$. On suppose ces arcs réguliers au voisinage de $t = 0$, c'est-à-dire ayant des tangentes $\gamma_j'(0) \neq 0$ en $z_0 = \gamma_j(0)$ (on peut supposer γ_j injectif pour $|t|$ assez petit d'après le théorème d'inversion locale), $j = 1, 2$. Pour $j = 1, 2$, les arcs images $f \circ \gamma_j$ ont pour tangente

$$J_{z_0}(f)(\gamma_j'(0)) = f'(z_0)\gamma_j'(0)$$

en $t = 0$ de sorte que les angles des tangentes à γ_j en z_0 sont les mêmes que ceux des arcs images en $f(z_0)$. On fait référence à cette propriété des fonctions holomorphes de conservation des angles en disant que ce sont des **transformations conformes**.



Si donc on regarde l'image par une application holomorphe du « grillage » formée des droites verticales et horizontales du plan par une application holomorphe (disons sur un domaine où la dérivée ne s'annule pas), on obtient un « grillage » de courbes aux tangentes orthogonales. Si on observe que les lignes de niveau

$$\mathrm{Re}(f) = cnte, \mathrm{Im}(f) = cnte'$$

sont les images des droites horizontales, verticales

$$\mathrm{Re}(z) = cnte, \mathrm{Im}(z) = cnte'$$

par l'inverse g de f, le résultat précédent montre que ces lignes de niveau sont des courbes à tangentes orthogonales aux points où elles se coupent. Dans le cas $g = \tanh$, on a

$$f = \mathrm{argtanh} = \frac{1}{2}\ln(\frac{z-1}{z+1})$$

et on obtient le dessin suivant.

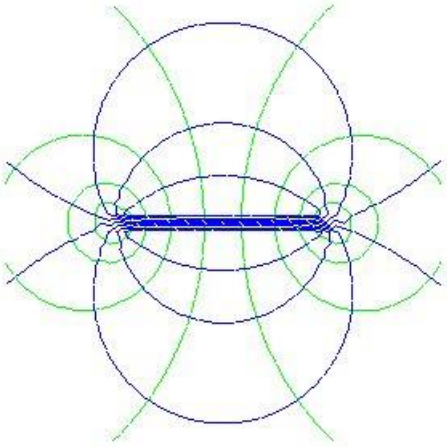


FIGURE X.3 – lignes de niveau tanh

Le lecteur féru d'électromagnétisme reconnaîtra les équipotentielles (en vert) et le lignes du champ magnétique (en bleu) crées par deux fils infinis parallèles dans un plan normal parcourus par des courants d'intensités égales et constantes de sens opposés. Cette propriété bien connue d'orthogonalité entre lignes de champs et équipotentielles , dans le cas plan, reflète la nature conforme des applications holomorphes. On développera plus bas cette observation qui est intimement liée à la nature **harmonique** du potentiel et au lien entre fonctions harmoniques réelles et les parties réelles de fonctions holomorphes (cf. 7 page 223).

3 Formule de Cauchy pour un cercle

Notons $C(z,r)$ le cercle de centre z et de rayon r.

3.1 Énoncé et preuve

Théorème 3.1.1. *Soit $\bar{D}(\omega,R) \subset \Omega$ un disque compact de bord le cercle $\Gamma = \partial\bar{D} = C(\omega,R)$, d'intérieur le disque ouvert $D = \bar{D} - \Gamma$ et $f \in \mathrm{Hol}(\Omega)$. Alors,*

$$\forall z_0 \in D,\ f(z_0) = \frac{1}{2i\pi}\int_\Gamma \frac{f(z)}{z-z_0}\mathrm{d}z.$$

Démonstration. Le théorème équivaut aux deux formules suivantes

$$\int_\Gamma \frac{f(z)-f(z_0)}{z-z_0}\mathrm{d}z = 0 \text{ et } \int_\Gamma \frac{\mathrm{d}z}{z-z_0} = 2i\pi.$$

– Calcul de la seconde intégrale.

On pose $z = \omega + R\,\mathrm{e}^{i\theta}$ puis

$$\int_\Gamma \frac{\mathrm{d}z}{z-z_0} = \int_0^{2\pi} \frac{iR\,\mathrm{e}^{i\theta}\,\mathrm{d}\theta}{\omega + R\,\mathrm{e}^{i\theta} - z_0} = i\int_0^{2\pi} \frac{\mathrm{d}\theta}{1+\frac{\omega-z_0}{R}\,\mathrm{e}^{-i\theta}}.$$

Comme $\left|\frac{\omega-z_0}{R}\right| < 1$, on a un développement en série normalement convergente sur $[0, 2\pi]$

$$\frac{1}{1+\frac{\omega-z_0}{R}\,\mathrm{e}^{-i\theta}} = \sum_{n=0}^{\infty}(\frac{\omega-z_0}{R})^n\,\mathrm{e}^{-in\theta}$$

qu'on peut donc intégrer terme à terme pour trouver

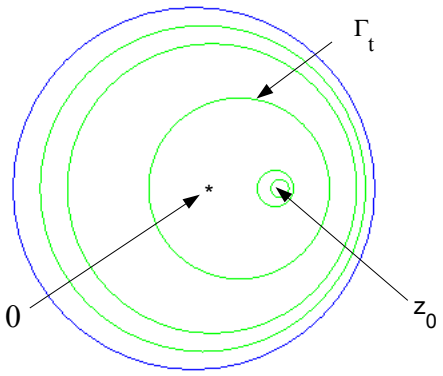
$$i\int_0^{2\pi}\frac{\mathrm{d}\theta}{1+\frac{\omega-z_0}{R}\,\mathrm{e}^{-i\theta}} = i\sum_{n=0}^{\infty}(\frac{\omega-z_0}{R})^n\int_0^{2\pi}\mathrm{e}^{-in\theta}\,\mathrm{d}\theta = 2i\pi,$$

ce qu'on voulait.

– Calcul de la première intégrale.
Quitte à changer z en $\omega + Rz$ (avec $D = D(\omega, R)$), on peut supposer $D = D(0, 1)$ de sorte que $|z_0| < 1$. On pose

$$I(t) = \int_{\Gamma_t}\frac{f(z)-f(z_0)}{z-z_0}\mathrm{d}z \text{ avec } t \in [0,1] \text{ et } \Gamma_t = C(tz_0, 1-t)$$

(déformation continue de $\Gamma = C(0, 1)$ sur z_0).



On pose donc

$$z = tz_0 + (1-t)\,\mathrm{e}^{i\theta}, \theta \in [0, 2\pi]$$

$$\text{On a } I(t) = \int_0^{2\pi}\frac{f(tz_0+(1-t)\,\mathrm{e}^{i\theta})-f(z_0)}{tz_0+(1-t)\,\mathrm{e}^{i\theta}-z_0}\mathrm{d}(tz_0+(1-t)\,\mathrm{e}^{i\theta})$$

$$= i\int_0^{2\pi}\frac{f(tz_0+(1-t)\,\mathrm{e}^{i\theta})-f(z_0)}{\mathrm{e}^{i\theta}-z_0}\,\mathrm{e}^{i\theta}\,\mathrm{d}\theta = \int_0^{2\pi}F(t,\theta)\mathrm{d}\theta.$$

Comme $\frac{f(z)-f(z_0)}{z-z_0} \in C^0(\Omega)\bigcap \mathrm{Hol}(\Omega - \mathrm{z}_0)$, on a

$$F \in C^0([0,1]\times[0,2\pi])\bigcap C^1(]0,1]\times[0,2\pi])$$

et donc $I \in C^0([0,1])\bigcap C^1(]0,1])$ et la dérivée de I s'obtient en dérivant sous le signe somme :

$$I'(t) = \int_0^{2\pi}\frac{\partial F(t,\theta)}{\partial t}\mathrm{d}\theta.$$

$$\text{Or } F(t,\theta) = i\frac{f(tz_0 + (1-t)\,\mathrm{e}^{i\theta}) - f(z_0)}{\mathrm{e}^{i\theta} - z_0}\,\mathrm{e}^{i\theta},$$

$$\text{donc } \frac{\partial F(t,\theta)}{\partial t} = if'(tz_0 + (1-t)\,\mathrm{e}^{i\theta})\,\mathrm{e}^{i\theta} = \frac{\partial F(t,\theta)}{\partial \theta}$$

et

$$I'(t) = \int_0^{2\pi} \frac{\partial F(t,\theta)}{\partial \theta}\mathrm{d}\theta = F(t,2\pi) - F(t,0) = 0.$$

I est donc constante sur $]0,1]$, donc sur $[0,1]$ (continuité) et $I(1) = I(0) = 0$.□

□

Les conséquences de la formule de Cauchy sont TRÈS nombreuses.

3.2 Holomorphie et analycité

Définition 3.2.1. Un fonction sur Ω est analytique si pour tout $z_0 \in \Omega$ $f(z)$ admet un développement en série entière

$$f(z) = \sum_{n\geq 0} a_n(z - z_0)^n \text{ convergeant pour } |z - z_0| \text{ assez petit.}$$

Théorème 3.2.2. *Une fonction est holomorphe si et seulement si elle est analytique.*

En particulier, les fonctions holomorphes sont de classe C^∞, alors que leur définition ne leur demandait que d'être C^1.

Démonstration. Comme la série dérivée d'une série entière convergente $\sum a_n(z - z_0)^n$ a même rayon de convergence R que la série initiale, la série des dérivées

$$\sum \partial a_n z^n = \sum n\partial a_n z^{n-1}$$

avec $\partial = \partial_x, -i\partial_y$ est normalement convergente de sorte que les dérivées partielles de $\sum a_n(z - z_0)^n$ s'obtiennent par dérivation terme à terme et sont continues. De plus, on a

$$\frac{\partial \sum a_n(z - z_0)^n}{\partial \bar{z}} = \sum a_n \frac{\partial (z - z_0)^n}{\partial \bar{z}} = 0$$

pour z dans le disque de convergence. Ainsi, les fonctions analytiques sont holomorphes.

Inversement, supposons f analytique sur Ω. Soit $z_0 \in D(\omega, R), 0 < R < \infty$ avec $\bar{D} \subset \Omega$. On note Γ le cercle bordant $\bar{D}$.

$$\begin{aligned} f(z_0) &\overset{Cauchy}{=} \frac{1}{2i\pi}\int_{\partial\bar{D}} \frac{f(z)}{z-z_0}\mathrm{d}z \text{ (en posant } z=\omega+R\,\mathrm{e}^{i\theta}) \\ &= \frac{1}{2\pi}\int_0^{2\pi} \frac{f(\omega+R\,\mathrm{e}^{i\theta})}{R\,\mathrm{e}^{i\theta}+\omega-z_0} R\,\mathrm{e}^{i\theta}\,\mathrm{d}\theta \\ &= \frac{1}{2\pi}\int_0^{2\pi} \frac{f(\omega+R\,\mathrm{e}^{i\theta})}{1-\frac{z_0-\omega}{R}\,\mathrm{e}^{-i\theta}}\mathrm{d}\theta \\ &= \frac{1}{2\pi}\int_0^{2\pi} \sum (\frac{z_0-\omega}{R})^n f(\omega+R\,\mathrm{e}^{i\theta})\mathrm{e}^{-in\theta}\mathrm{d}\theta. \end{aligned}$$

Les majorations

$$\left|(\frac{z_0-\omega}{R})^n f(\omega+R\,\mathrm{e}^{i\theta})\mathrm{e}^{-in\theta}\right| \leq \frac{|z_0-\omega|^n}{R^n}\sup_\Gamma |f| \text{ et } \frac{|z_0-\omega|}{R}<1$$

donnent la convergence normale permettant d'intervertir $\int$ et $\sum$ pour obtenir

$$\text{(3.2.a)} \qquad f(z_0)=\sum_{n\geq 0} a_n(z_0-\omega)^n \text{ avec } a_nR^n=\frac{1}{2\pi}\int_0^{2\pi} f(\omega+R\,\mathrm{e}^{i\theta})\,\mathrm{e}^{-in\theta}\,\mathrm{d}\theta$$

ce qui termine la preuve. □

Remarque 3.2.3. Les propriétés de dérivabilité terme à terme des séries entières sur leurs disques de convergence utilisées dans la preuve précédente assurent que f a des dérivés (complexes) à tout ordre et que $a_n = f^{(n)}(z_0)/n!$ pour tout n. Autrement dit, f est analytique si et seulement si f est somme de sa série de Taylor en tout point.

La preuve précédente donne en fait le résultat plus précis suivant

Corollaire 3.2.4 (Estimées de Cauchy)**.** *Avec les notations précédentes, $f(z_0)$ admet le développement en série entière pour $z_0 \in D(\omega, R)$*

$$f(z_0)=\sum_{n\geq 0}\frac{f^{(n)}(\omega)}{n!}(z_0-\omega)^n$$

où

$$f^{(n)}(z_0)=\frac{n!}{2\pi}\int_{C(\omega,r)}\frac{f(z)}{(z-z_0)^{n+1}}dz$$

et en particulier

$$\text{avec } |f^{(n)}(\omega)|\leq \frac{n!\sup_\Gamma|f|}{R^n} \text{ (estimées de Cauchy).}$$

En particulier, toute fonction holomorphe sur un disque ouvert y est développable en série entière.

Démonstration. C'est une conséquence immédiate de (3.2.a) et de la formule habituelle $a_n = f^{(n)}(\omega)/n!$. □

On en déduit alors le résultat très important suivant, dont une variante globale plus subtile sera donnée plus bas (6.3.1 page 222).

Corollaire 3.2.5 (Existence locale de primitives)**.** *Toute fonction holomorphe admet localement une primitive. Plus précisément, si f est holomorphe sur un disque ouvert D, il existe une primitive $F \in \mathrm{Hol}(\Omega)$, **i.e.** telle que $F' = \Omega$.*

Démonstration. Seul le deuxième point est à prouver. Or, d'après le corollaire précédent, il suffit de remarquer que toute série entière convergeant sur D admet une primitive convergeant sur D. □

4 Séries de fonctions holomorphes

Théorème 4.0.6. *Soit $\sum f_n$ série de fonctions holomorphes sur Ω convergeant normalement sur tout disque compact $\bar{D} \subset \Omega$. Alors, $\sum f_n'$ converge normalement sur tout disque compact $\bar{D} \subset \Omega$ et $\sum_{n\geq 0} f_n$ holomorphe de dérivée $\sum_{n\geq 0} f_n'$.*

Démonstration. Tout disque compact $\overline{D(\omega, r)} \subset \Omega$ est contenu dans $\bar{D} = \overline{D(\omega, R)} \subset \Omega$ pour $R > r$ convenable, car la distance de $\overline{D(\omega, r)}$ au complémentaire de Ω est strictement positive [1]. Comme . Soit $\omega \in \bar{D}$ et $C = \partial\overline{D(\omega, r)}$ et $f \in \mathrm{Hol}(\Omega)$.

On a (Cauchy)

$$f(\omega) = \frac{1}{2i\pi} \int_C \frac{f(z)}{z - \omega} \mathrm{d}z.$$

Comme

$$\frac{\partial g}{\partial x} \text{ et } \frac{\partial g}{\partial y} \in C^0(C \times \bar{D}), \text{ où } g = \frac{f(z)}{z - \omega},$$

on peut dériver sous le signe somme et on trouve

$$f'(\omega) = -\frac{1}{2i\pi} \int_C \frac{f(z)}{(z - \omega)^2} \mathrm{d}z.$$

Comme $|z - \omega| \geq R - r$, on a

$$\forall \omega \in \bar{D},\ |f'(\omega)| \leq \frac{\max \sup_C |f|}{2\pi(R - r)^2}.$$

La convergence normale de $\sum f_n$ sur $\bar{D}$ entraîne donc celle des f_n' et on peut dériver terme à terme par rapport à x et y pour obtenir le théorème. □

1. La distance $d(x, \mathbf{C} - \Omega)$ est continue donc atteint son minimum sur le compact $\overline{D(\omega, r)}$. Comme cette distance ne s'annule pas, ce minimum est strictement positif. Si on le note ε, on aura $D(\omega, r + \frac{\varepsilon}{2}) \subset \Omega$

5 La formule des résidus

On va donner une puissante formule d'intégration calculant la circulation [2] de la forme $f(z)\mathrm{d}z, f \in \mathrm{Hol}(\Omega)$ le long d'une courbe fermée.

5.1 Rappel sur la circulation des 1-formes différentielles

Définition 5.1.1 (Rappel). Un chemin γ chemin (orienté) de Ω est une application continue $\gamma : [0,1] \longrightarrow \Omega$. Le chemin opposé est $\gamma^{\mathrm{opp}} : t \longrightarrow \gamma(1-t)$. On dira que γ est un chemin régulier si γ est C^1 par morceaux. On dira dans ce cours que le chemin γ est un lacet s'il est régulier et si de plus $\gamma(0) = \gamma(1)$.

Soit alors

$$(P(x,y), Q(x,y)) : \mathbf{R}^2 \longrightarrow \mathbf{C}$$

une application continue. Si γ est un lacet de $\mathbf{C} = \mathbf{R}^2$, on a défini la circulation de la **1-forme** [3]. $P\mathrm{d}x + Q\mathrm{d}y$ le long de $\gamma = \gamma_1 + i\gamma_2$ par la formule

$$\int_\gamma P\mathrm{d}x + Q\mathrm{d}y := \int_0^1 P(\gamma(t))\gamma_1'(t)\mathrm{d}t + \int_0^1 Q(\gamma(t))\gamma_2'(t)\mathrm{d}t$$

avec γ_1, γ_2 parties réelles et imaginaires de γ.

On sait que la circulation est invariante par changement de paramétrage croissant. Ceci permet, ce qu'on fera sans plus de précision, de considérer la circulation d'une 1-forme différentielle le long d'un arc régulier orienté sans préciser le paramétrage. La circulation change en son opposé par inversion du sens de parcours :

(5.1.a)
$$\int_\gamma = -\int_{\gamma^{\mathrm{opp}}}$$

Nous nous intéresserons ici uniquement à des formes du type

$$f(z)\mathrm{d}z := (P + iQ)(\mathrm{d}x + i\mathrm{d}y)$$

en posant comme plus haut $P = \mathrm{Re}(f), Q = \mathrm{Im}(f)$. La circulation le long de γ est alors par définition

(5.1.b)
$$\int_\gamma f(z)\mathrm{d}z = \int_0^1 f(\gamma(t))\gamma'(t)\mathrm{d}t.$$

2. Cette notion a largement été utilisée en classe préparatoire, que ce soit en physique (cf. par exemple le théorème d'Ampère sur la circulation du champ magnétique le long d'un contour fermé) et en mathématique (cf. par exemple la formule de Green-Riemann sur l'intégrale curviligne des formes différentielles).

3. Dans ce cours l'expression $Pdx + Qdy$ a un sens formel. On peut l'identifier au champ de vecteurs du plan (P, Q)

5.2 Énoncé de la formule

Avant d'énoncer la formule générale des résidus qui demandera d'assez longs développements, montrons le résultat suivant, qui en est la clef.

Théorème 5.2.1. *Soit $f \in \mathrm{Hol}(\Omega)$ et γ un lacet de Ω. Si f admet une primitive holomorphe F sur Ω, alors*

$$\int_\gamma f(z)\mathrm{d}z = 0.$$

Démonstration. Commençons par un calcul (l'idée générale est de remplacer les variables réelles (x, y) par les variables complexes $z, \bar{z}$ de sorte que les fonctions $\mathrm{Re}(f), \mathrm{Im}(f)$ sont remplacées dans la formule de composition des dérivées par les fonctions $f, \bar{f}$).

Lemme 5.2.2. *Soit $f : \Omega \longrightarrow \mathbf{C}$ une fonction $\mathbf{C}^1$ et γ un lacet de Ω. On a la formule*

$$\frac{d}{dt} f(\gamma(t)) = \gamma'(t)\frac{\partial f}{\partial z}(\gamma(t)) + \bar{\gamma}'(t)\frac{\partial f}{\partial \bar{z}}(\gamma(t)).$$

En particulier, si f holomorphe, on a

$$\frac{d}{dt} f(\gamma(t)) = f'(\gamma(t))\gamma'(t).$$

Démonstration. C'est une variation laissée au lecteur de la formule (2.2.2 page 195). □

On a donc

$$\int_\gamma f(z)\mathrm{d}z = \int_0^1 F'(\gamma(t))\gamma'(t)\mathrm{d}t = [F(\gamma(t))]_0^1 = 0.$$

□

Énonçons maintenant la formule générale

Théorème 5.2.3 (Formule des résidus)**.** *Soit f méromorphe sur Ω simplement connexe (par ex. convexe) ayant un ensemble fini de pôles S et γ lacet régulier dans $\Omega - S$. Alors,*

$$\frac{1}{2i\pi}\int_\gamma f(z)\mathrm{d}z = \sum_{s\in S} \mathrm{res}_s(f)\,\mathrm{ind}_s(\gamma).$$

Avant de définir les notions apparaissant dans l'énoncé disons que les fonctions holomorphes sont méromorphes mais n'ont aucun pôle par définition. Donc, le théorème précédent assure que si Ω est simplement connexe, l'intégrale $\int_\gamma f(z)\mathrm{d}z$ est nulle le long de tout lacet γ, ce qui est une généralisation de la formule de Cauchy pour un cercle (un disque étant convexe donc simplement connexe comme on l'a dit plus haut). On verra par ailleurs que toutes les fonctions holomorphes admettent des primitives sur les ouverts simplement connexes (6.3.1 page 222). Donc, le théorème des résidus est compatible avec (5.2.1), ce qui somme toute est assez heureux.

5.3 Fonctions méromorphes, résidu

Définition 5.3.1. Une fonction $f : \Omega \longrightarrow \mathbf{C}$ a un pôle en $z_0 \in \Omega$ si $f \in \mathrm{Hol(D} - \{\mathrm{z_0}\})$ où $D \subset \Omega$ est un disque ouvert contenant z_0 et s'il existe $N \in \mathbf{N}$ tel que $(z - z_0)^N f$ soit bornée sur D. On dit alors f est méromorphe au voisinage de z_0.Si f est méromorphe au voisinage de tout point de Ω, on dit que f est méromorphe sur Ω.

Les fonctions holomorphes sont évidemment méromorphes. Si f, g sont méromorphes, $f + g, fg$ le sont aussi.

Proposition 5.3.2. *Le point z_0 de Ω est un pôle de f si et seulement si f a un développement de Laurent près de z_0*

$$f(z) = \sum_{n \geq -N} a_n (z - z_0)^n$$

sur une couronne assez mince $\{z$ tels que $0 < |z - z_0| < R\}, R << 1$. Ce développement est alors unique. De plus, on a convergence normale de la série sur toute couronne $C_{\epsilon,r} = \{z$ tels que $0 < \epsilon < |z - z_0| < r \leq R\}$. Enfin, les pôles d'une fonction méromorphe sont isolés.

Définition 5.3.3. Avec les notations précédentes, on définit le résidu $\mathrm{res}_{z_0}(f)$ de f en z_0 par

$$\mathrm{res}_{z_0}(f) = a_{-1}.$$

Démonstration. Si z_0 est un pôle de f, $(z - z_0)^N f(z)$ est holomorphe sur $D - \{z_0\}$ et est bornée sur D. Donc, $g(z) = (z - z_0)^{N+2} f(z)$ est C^1 en z_0 (avec $g'(z_0) = 0$) de sorte que $g \in \mathrm{Hol(D)}$. Or on sait que les fonctions holomorphes admettent sur D un unique développement en série entière

$$g(z) = \sum_{n \geq 0} a_n (z - z_0)^n$$

qui converge normalement sur tout compact contenu dans D (3.2.4 page 203). On a alors,

$$f(z) = \sum_{n \geq 0} a_n (z - z_0)^{n-N-2}$$

sur le disque épointé $D - \{z_0\}$. La réciproque est évidente. □

On déduit immédiatement de la proposition précédente la propriété suivante

Corollaire 5.3.4. *L'application res_{z_0} de l'espace des fonctions méromorphes sur Ω vers $\mathbf{C}$ est une forme linéaire.*

Le lien entre circulation et résidu est donné par l'énoncé suivant

Proposition 5.3.5. *Soit z_0 un pôle de la fonction méromorphe f. Alors, pour $\epsilon > 0$ assez petit, on a*

$$\mathrm{res}_{z_0}(f) := \frac{1}{2i\pi} \int_{C(z_0,\epsilon)} f(z) \mathrm{d}z.$$

Démonstration. Si $\epsilon > 0$ est assez petit, le développement de Laurent en z_0 de f

$$f(z) = \sum_{n \geq -N} a_n (z - z_0)^n$$

est normalement convergent sur $C(z_0, \epsilon)$ de sorte qu'on peut intégrer terme à terme. Or on a

$$\int_{C(z_0,\epsilon)} (z - z_0)^n \mathrm{d}z = \int_0^{2\pi} (z_0 - z_0 - \epsilon^n \, \mathrm{e}^{in\theta}) \mathrm{d}(z_0 + \epsilon \, \mathrm{e}^{i\theta}) = i\epsilon^{1+n} \int_0^{2\pi} \mathrm{e}^{i(1+n)\theta} \, \mathrm{d}\theta$$

qui vaut 0 si $n \neq 0$ et $2i\pi$ sinon. Le lemme suit. □

Exercice 5.3.6. Si g est holomorphe et f est méromorphe, montrer que $f \circ g$ méromorphe.

Exercice 5.3.7. Montrer que si f et g sont méromorphes en z_0, alors soit g est nulle dans un voisinage de z_0 soit f/g est méromorphe en z_0 (voir le théorème des zéros isolés (6.1.3 page 220) pour plus de précision sur le lieu d'annulation de g).

Exercice 5.3.8. Montrer que $\tan(z)$ est méromorphe mais que $\sin(1/z)$ ne l'est pas.

5.4 Indice d'un lacet

Soit γ un lacet dans $\mathbf{C}$ et $z_0 \notin \gamma$ (c'est-à-dire $z_0 \notin \gamma([0,1]))$..

Définition 5.4.1. L'indice de γ en z_0 est défini par

$$\mathrm{ind}_{z_0}(\gamma) := \int_\gamma \frac{\mathrm{d}z}{z - z_0} = \int_0^1 \frac{\gamma'(t)}{\gamma(t) - z_0} \mathrm{d}t.$$

On a $\mathrm{ind}_{z_0}(\gamma^{\mathrm{opp}}) = -\mathrm{ind}_{z_0}(\gamma)$ (cf. 5.1.a page 205). On verra plus bas (5.6.1 page 211) que l'indice est invariant par déformation continue, **i.e.**par *homotopie de lacets* (cf. 5.5.1 page 210).

Proposition 5.4.2. *L'indice d'un lacet est à valeurs dans* $2i\pi\mathbf{Z}$.

Démonstration. On pose

$$f(t) = \exp(\int_0^t \frac{\gamma'(u)}{\gamma(u) - z_0} \mathrm{d}u), t \in [0,1].$$

Comme

$$f'/f = (\gamma - z_0)'/\gamma - z_0$$

(sauf au nombre fini de points non C^1 de γ), on a

$$\frac{f(t)}{f(0)} = \frac{\gamma(t) - z_0}{\gamma(0) - z_0}.$$

Comme $f(0) = 1$ et $\gamma(0) = \gamma(1)$, on a $f(1) = 1$ et donc

$$\exp(\mathrm{ind}_{z_0}(\gamma)) = 1.$$

□

Pratiquement, l'indice compte algébriquement le nombre de tours de γ autour de zéro.

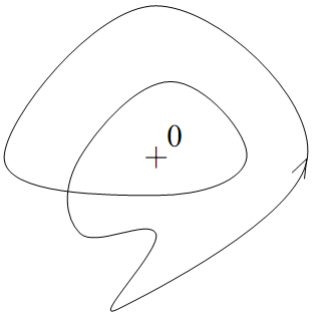
Exercice 5.4.3. Si z_0 est à l'intérieur d'un cercle

$$C : \left\{ \begin{array}{lll} [0,1], & \longrightarrow \mathbf{C} & \\ t & \mapsto & \omega + R\exp(2i\pi t) \end{array} \right.$$

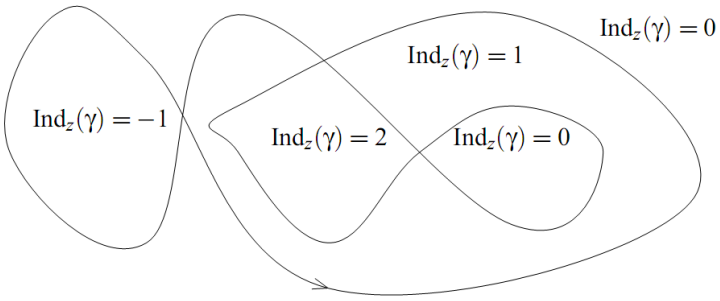
on a $\mathrm{ind}_{z_0}(C) = 1$ [Appliquer par exemple la formule de Cauchy à $f = 1$].

Si z_0 à l'extérieur de C, on a $\mathrm{ind}_{z_0}(C) = 0$ [Développer $1/(z - z_0)$ en puissances de $1/z$ en se ramenant à $C = C(0,1)$ par exemple].

Exercice 5.4.4 (Difficile). Montrer que si $U \subset \mathbf{C} - \gamma([0,1])$ est connexe, l'indice $\mathrm{ind}_z(\gamma)$ est constant quand z décrit U. Supposons $\gamma'(t_0) \neq 0$. Montrer qu'il existe un voisinage U de $\gamma(t_0)$ telles que $U - \gamma$ a deux composantes connexes ouvertes $U^{\pm}$ telles que $|\mathrm{ind}_{z^+}(\gamma) - \mathrm{ind}_{z^-}(\gamma)| = 1$ pour $z^{\pm} \in U^{\pm}$.



Lacet d'indice 2 par rapport à l'origine



Variation de l'indice en fonction de la position de z

5.5 Simple connexité

On formalise ici la notion de déformation continue de chemins déjà rencontrée dans la preuve de la formule de Cauchy.

Définition 5.5.1. – Une **homotopie** entre deux chemins de Ω, γ_0 et γ_1 est une application $\gamma \in C^0([0,1]^2, \Omega)$ telle que

$$\forall t,\ \gamma(0,t) = \gamma_0, \gamma(1,t) = \gamma_1(t).$$

– On dit que γ est une **homotopie de lacets** si de plus $\forall s,\ \gamma(s,0) = \gamma(s,1)$. On dit alors que les lacets $\gamma_i : t \mapsto \gamma_i(t), i = 0,1$ sont homotopes.
– On dit que Ω est **simplement connexe** s'il est connexe et si pour tout lacet γ dans Ω il existe une **homotopie de lacets** avec un lacet constant.

Lemme 5.5.2. *La relation sur les lacets de Ω « il existe une homotopie de lacets entre γ_0 et γ_1 » est une relation d'équivalence.*

Démonstration. La réflexivité est claire. Pour la symétrie, on observe que si γ est une homotopie de lacets entre γ_0 et γ_1, alors, $\gamma(1-s,t)$ est une homotopie de lacets entre γ_1 et γ_0. Passons à la transitivité. On a donc trois lacets $\gamma_1, \gamma_2, \gamma_3$ et des homotopies de lacets $\gamma_{i,i+1}(s,t)$ entre γ_i et $\gamma_{i+1}, i = 1,2$. On pose alors

$$\gamma(s,t) = \gamma_{1,2}(2s,t) \text{ si } s \leq 1/2 \text{ et } \gamma(s,t) = \gamma_{2,3}(2s-1,t) \text{ si } s \geq 1/2$$

qui est une homotopie de lacets entre γ_1 et γ_2. □

Remarque 5.5.3. Donnons-nous un point $z_0 \in \Omega$. On peut définir donc l'ensemble $\pi_1(\Omega, z_0)$ des classes d'équivalence d'homotopie de lacet partant et arrivant en z_0. On peut montrer que la concaténation des lacets définit une structure de groupe, en général compliquée, sur $\pi_1(\Omega)$. C'est le célèbre groupe fondamental de Poincaré de Ω. Comme on le voit, on peut le définir pour des espaces beaucoup plus généraux que Ω. En ces termes, dire que l'ensemble connexe Ω est simplement connexe c'est dire que ce groupe est réduit à l'élément neutre.



FIGURE X.4 – Henri Poincaré

Donnons un exemple simple et utile d'ouvert simplement connexe.

Proposition 5.5.4. *Un ouvert étoilé (par exemple un convexe) est simplement connexe.*

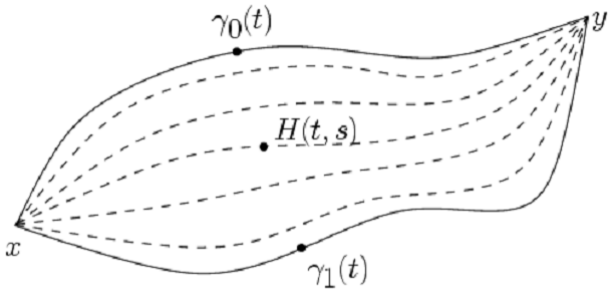
Démonstration. Rappelons que Ω est dit étoilé s'il existe $\omega \in \Omega$ tel que pour tout $z \in \Omega$ le segment $[\omega, z]$ est contenu dans Ω. Soit γ un lacet de Ω. On constate alors que

$$(s,t) \mapsto sz_0 + (1-s)\gamma(t)$$

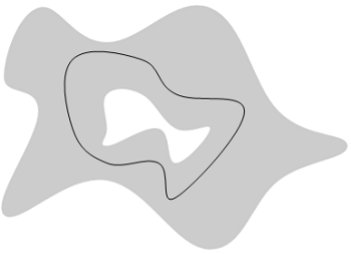
est une homotopie de lacets entre γ et le lacet constant de valeur z_0. □

Exercice 5.5.5 (Difficile). Montrer que $\mathbf{C}^*$ n'est pas simplement connexe.

Cet exercice deviendra clair après le théorème d'homotopie (5.6.1).



Une homotopie



Lacet non homotope au lacet constant

5.6 Invariance par homotopie de l'intégrale sur un lacet d'une fonction holomorphe

Les termes du théorème 5.2.3 page 206 étant définis, passons à sa preuve. C'est une conséquence de la généralisation suivante du théorème de Cauchy.

Théorème 5.6.1. *Soient γ homotopie de lacets entre γ_0, γ_1 réguliers et $f \in \mathrm{Hol}(\Omega)$, alors*

$$\int_{\gamma_1} f(z)\mathrm{d}z = \int_{\gamma_2} f(z)\mathrm{d}z.$$

En particulier, l'indice d'un lacet est invariant par homotopie.

Démonstration. Soit $n > 0$ et $C_{i,j}$ le carré plein $[\frac{i}{n}, \frac{i+1}{n}] \times [\frac{j}{n}, \frac{j+1}{n}]$. On a

$$C = [0,1]^2 = \bigcup_{0 \leq i,j, \leq n-1} C_{i,j}.$$

On note $\partial C, \partial C_{i,j}$ les bords de $C, C_{i,j}, 0 \leq i, j, \leq n-1$ qu'on voit comme des lacets orientés dans le sens trigonométrique. Soit

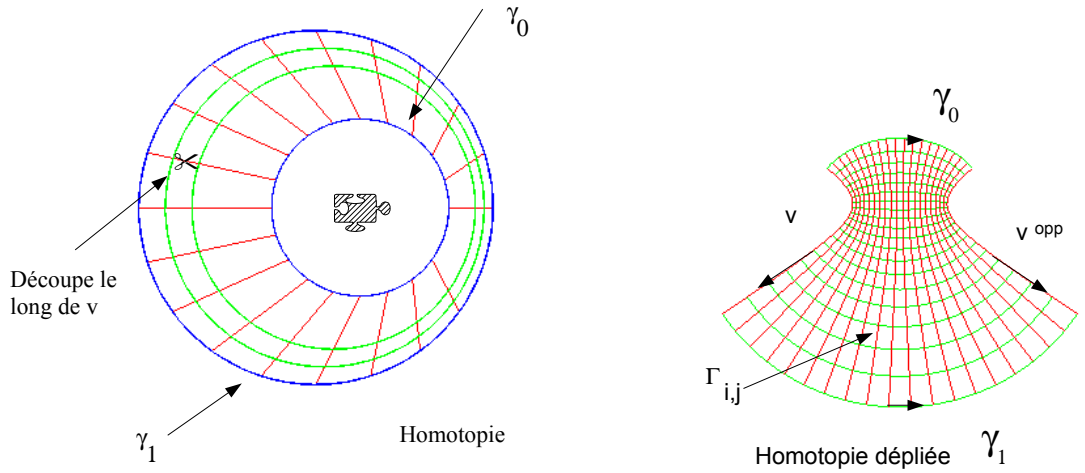
$$\gamma : C \longrightarrow \Omega$$

une homotopie entre γ_0 et γ_1 et v le chemin $s \mapsto \gamma(s,0) = \gamma(s,1)$. On a la décomposition de l'arc image

$$\Gamma = \gamma(\partial C) = \gamma_0 \cup v \cup \gamma_1^{opp} \cup v^{opp}.$$

On déduit la formule pour la circulation de $f(z)\mathrm{d}z$ le long de Γ

$$\int_\Gamma = \int_{\gamma_0} + \int_v - \int_{\gamma_1} - \int_v. \tag{5.6.a}$$



L'idée est maintenant, grâce à la compacité de C, de « grillager » comme sur les dessins précédents la surface $\gamma(C)$ en une réunion finie de surfaces contenues dans des ouverts assez petits dans lesquels $f(z)\mathrm{d}z$ a une primitive (3.2.5 page 204). La circulation sera alors la somme des circulations sur le bords $\Gamma_{i,j}$ de ces surfaces (additivité de la circulation par réunion des chemins) et donc sera nulle (5.2.1 page 206). D'une certaine manière, on pourrait s'arrêter là, le reste n'étant que mise en forme technique de l'idée exposée. Expliquons malgré tout les détails.

On recouvre $\gamma(C)$ par des disques $D_\alpha \subset \Omega$ sur lesquels f a une primitive (3.2.5 page 204). Comme $\gamma(C)$ est compact comme image continue d'un compact, on peut choisir un nombre de Lebesgue $\delta > 0$ du recouvrement (4.3.2 page 16). Par définition, on a donc

$$\forall x \in \gamma(C)\ \exists \alpha \text{ tel que } D(x,\delta) \subset D_\alpha.$$

Comme γ est continue, elle est uniformément continue sur le compact C (4.3.6 page 18). Ceci permet de choisir n assez grand de sorte que pour tout $x, y \in C$ contenus dans un carré de côté $1/n$ on ait $|f(x) - f(y)| < \delta$. On a alors

$$\forall i, j\ \exists \alpha \text{ avec } \gamma(C_{i,j}) \subset D_\alpha.$$

On pourrait être tenté de prendre $\Gamma_{i,j}$ comme étant le chemin image du bord orienté (dans le sens indirect) de $C_{i,j}$. Ceci est correct si on suppose de plus γ de classe C^1 car alors ce chemin

est bien régulier. Mais on n'a pas envie de supposer γ de classe C^1. On va en fait remplacer les chemins intermédiaires par des chemins affines par morceaux.

On commence par remplacer $\gamma(0,t), \gamma(1,t)$ par des chemins affines par morceaux en restant dans les disques D_α dans chaque pas de la subdivision. Pour cela, on définit $\Gamma_{-1,j}$ comme le chemin orienté

$$\gamma_{|\{0\}\times[\frac{i}{n},\frac{i+1}{n}]}$$

suivi de

$$[\gamma(0,\frac{i+1}{n}),\gamma(0,\frac{i}{n})].$$

De même, on définit $\Gamma_{n,j}$ comme le chemin orienté

$$\gamma_{|\{1\}\times[\frac{i}{n},\frac{i+1}{n}]}$$

suivi de

$$[\gamma(1,\frac{i+1}{n}),\gamma(1,\frac{i}{n})].$$

Puis, on grillage en définissant alors simplement $\Gamma_{i,j}, i = 0,\cdots,n-1$ comme le quadrilatère orienté

$$\Gamma_{i,j} = \big(\gamma(\frac{i}{n},\frac{j}{n}),\gamma(\frac{i+1}{n},\frac{j}{n}),\gamma(\frac{i+1}{n},\frac{j+1}{n})),\gamma(\frac{i}{n},\frac{j+1}{n})\big).$$

Le bord Comme le disque D_α est convexe, il contient $\Gamma_{i,j}$. Comme f a une primitive sur D_α, la circulation de $f(z)\mathrm{d}z$ le long du lacet $\Gamma_{i,j}$ est nul. Par construction,et les

$$\Gamma_{i,j}, -1 \leq i,j \leq n$$

grillagent $\gamma(C)$ comme sur le dessin (les bords de deux quadrilatères consécutifs ont un côté commun). Par additivité de la circulation, on a

$$\int_\Gamma f(z)\mathrm{d}z = \sum_{i,j}\int_{\Gamma_{i,j}} f(z)\mathrm{d}z$$

car les contributions des côtés intérieurs se détruisent deux à deux et donc

$$0 = \int_\Gamma f(z)\mathrm{d}z \overset{5.6.a}{=} \int_{\gamma_0} f(z)\mathrm{d}z - \int_{\gamma_1} f(z)\mathrm{d}z.$$

□

Corollaire 5.6.2. *Pour tout lacet γ de Ω simplement connexe et tout $f \in \mathrm{Hol}(\Omega)$, on a*

$$\int_\gamma f(z)\mathrm{d}z = 0.$$

Démonstration. Considérer une homotopie de lacets entre γ et un lacet constant et appliquer le théorème. □

5.7 Preuve de la formule des résidus

Passons à la preuve de la formule des résidus (théorème 5.2.3 page 206), à savoir

$$\frac{1}{2i\pi}\int_\gamma f(z)\mathrm{d}z = \sum_{s\in S}\mathrm{res}_s(f)\,\mathrm{ind}_s(\gamma)$$

pour f méromorphe sur Ω simplement connexe ayant un ensemble fini de pôles S et γ lacet régulier dans $\Omega - S$.

Démonstration du théorème des résidus. Écrivons le développement de Laurent de f en $s \in S$

$$f(z) = \sum_{n\geq -N} a_n(z-s)^n = Q_s(z) + \sum_{n\geq 0} a_n(z-s)^n$$

pour z proche de $s \in S$ avec a_n, s dépendant de s. $f(z) - \sum_s Q_s(z)$ est holomorphe sur Ω simplement connexe donc (corollaire 5.6.2 page précédente de la forme globale de Cauchy)

$$\int_\gamma (f(z) - \sum_s Q_s(z))\mathrm{d}z = 0.$$

Mais comme $(z-s)^n, n \neq -1$ a une primitive au voisinage de γ,

$$\frac{1}{2i\pi}\int_\gamma Q_s(z)\mathrm{d}z = a_{-1}\frac{1}{2i\pi}\int_\gamma \frac{\mathrm{d}z}{z-s} = \mathrm{res}_s(f)\,\mathrm{ind}_s(\gamma).$$

□

Remarque 5.7.1. La preuve montre qu'on n'a pas vraiment besoin de supposer Ω simplement connexe mais simplement de savoir que le lacet γ est homotope à un lacet constant.

5.8 Principe d'utilisation de la formule des résidus pour le calcul d'intégrale

Donnons sans preuve une méthode d'utilisation que le lecteur justifiera au cas par cas.

On se donne un ouvert connexe borné U de $\mathbf{C}$ dont le bord admet une paramétrisation propre, qu'on peut supposer par longueur d'arc (en dehors du nombre fini de points non C^1), à savoir $|\gamma'(t)| = 1$ par un arc régulier γ (c'est-à-dire que la restriction de γ à $[0, \mathrm{long}(\gamma)[$ est injective) [4]. Par exemple, un disque, un demi disque, un triangle...

Ainsi, γ tourne autour de chaque point de U une seule fois. On peut supposer (quitte à renverser le sens de parcours) que γ est orienté dans le sens trigonométrique. Ceci signifie qu'on pousse légèrement γ dans la direction normale $i\gamma'(t)$, on tombe dans U, soit, mathématiquement $\gamma(t) - i\epsilon\gamma'(t)$ est dans U pour ϵ assez petit). Dans ce cas, l'indice des points de U relativement à γ est 1.

En pratique, c'est très facile à justifier en utilisant l'invariance par homotopie de l'indice et en déformant continûment γ sur un cercle par exemple.

4. On dit que γ est une courbe de Jordan régulière, cf. 9.2.2 page 230.

Soit maintenant Ω est un ouvert simplement connexe contenant $\bar{U}$ et f méromorphe sur Ω, sans pôle sur γ. Soit S l'ensemble des pôles de f contenus dans U. C'est un ensemble fini (exercice). La formule des résidus s'écrit alors

(5.8.a) $$\int_\gamma f(z)\mathrm{d}z = 2i\pi \sum_{s\in S} \mathrm{res}_s(f).$$

En fait, le difficile théorème 9.2.2 page 230 donne une preuve générale. Le théorème de Jordan-Schönflies assure même qu'une courbe de Jordan plane borde un domaine U qui est automatiquement simplement connexe, puisque homéomorphe à disque !

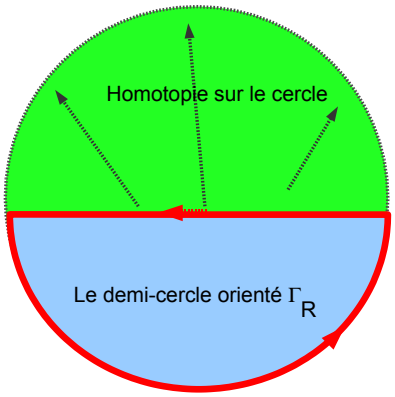
5.9 Exemple d'application à des calculs d'intégrales

La formule des résidus permet de calculer de nombreuses intégrales autrement inaccessibles. Elle est particulièrement utile pour les transformations de Fourier, bien que ce ne soit pas sa seule utilisation. A titre d'exemple, calculons la transformé de Fourier de $\frac{1}{1+x^2}$. Elle est donnée par l'intégrale

$$I(\xi) = \int_{\mathbf{R}} \frac{\exp(-ix\xi)}{1+x^2} dx$$

Supposons d'abord $\xi > 0$ et considérons le contour Γ_R bord du demi-disque inférieur $D_R^- = \{z \in \mathbf{C} \mid \mathrm{Im}(z) < 0, |z| \leq R\}$.

Ce demi-disque orienté est homotope au cercle $C(0,R)$ (pousser continûment $[-R,R]$ sur le demi-cercle supérieur) de sorte que les points intérieurs sont d'indice 1.



Ce bord est constitué du segment réel $[-R,R]$ et de l'arc de cercle inférieur $\{Re^{-i\theta} \mid 0 \leq \theta \leq \pi\}$. On a alors

$$\int_{\Gamma_R} \frac{e^{-iz\xi}}{1+z^2} dz = \int_{-R}^{R} \frac{e^{-ix\xi}}{1+x^2} dx + \int_0^\pi \frac{exp(-iRe^{-i\theta}\xi)}{1+(Re^{i\theta})^2} iRe^{i\theta} d\theta$$

La seconde intégrale s'estime en norme par

$$\int_0^\pi \frac{exp(-R\sin(\theta)\xi)}{R^2-1} Rd\theta = \frac{\pi R}{R^2-1}$$

qui tend vers 0 avec R. On en déduit (5.8.a page précédente)

$$\lim_{R\to+\infty} \int_{-R}^{R} \frac{e^{-ix\xi}}{1+x^2} dx = \lim_{R\to+\infty} \int_{\Gamma_R} \frac{e^{-iz\xi}}{1+z^2} dz = 2i\pi \sum_{z\in D^-} \text{res}\left(\frac{e^{-iz\xi}}{1+z^2}\right)$$

Or $\frac{1}{1+z^2} = \frac{-1}{2i}\left(\frac{1}{z+i} + \frac{1}{z-i}\right)$. Il y a donc deux pôles, mais seul $-i$ est dans D^-, et son résidu est $\frac{-1}{2i}$. on a donc

$$\lim_{R\to+\infty} \int_{-R}^{R} \frac{e^{-ix\xi}}{1+x^2} dx = \pi e^{\xi}$$

Le cas de $\xi < 0$ se ramène à celui-ci, car le changement de variable $y = -x$ montre que $I(-\xi) = I(\xi)$. En conclusion, on a montré $I(\xi) = \pi e^{-|\xi|}$.

5.10 Intermède culturel : Bulles de savon et méromorphie

Nous allons présenter sans preuve (même si la démonstration serait accessible par le lecteur de ce cours disposant de plus de temps) l'apparition des fonctions méromorphes dans un cadre somme toute inattendu, la forme des bulles de savon.

Les bulles de savon minimisent les contraintes : le travail pour lutter contre les forces de pression. Ce travail est proportionnel à la surface (formule de Laplace) : la forme d'une bulle de savon minimise donc l'aire.



FIGURE X.5 – Alfred Enneper



FIGURE X.6 – Leonhard Euler

Enneper et Weierstrass ont montré en 1866 que ces surfaces dites minimales se paramétraient nécessairement (localement) en $z \mapsto \text{Re}(F(z))$ avec

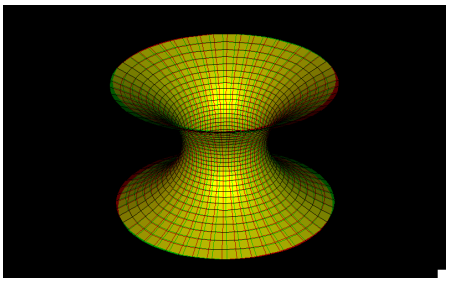
$$x(z) = \mathrm{Re}\int_{[z_0,z]}(1-g(w)^2)f(w)dw, y(z) = \mathrm{Re}\int_{[z_0,z]} i(1+g(w)^2)f(w)dw, h(z) = \mathrm{Re}\int_{[z_0,z]} 2g(w)f(w)dw$$

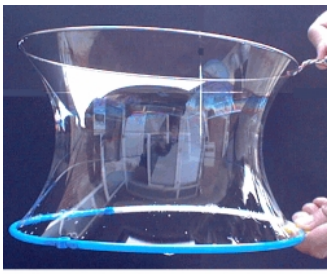
que l'on écrit de manière plus concise $F : \mathbf{C} \to \mathbf{R}^3$

$$F(z) = \mathrm{Re}\int_{[z_0,z]} \big[(1-g(w)^2)f(w), i(1+g(w)^2)f(w), 2g(w)f(w)\big]\mathrm{d}w.$$

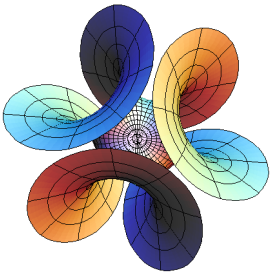
et f **méromorphe** et g **holomorphe**.

Avec $g(w) = w, f(w) = 1/w^2$, on trouve la caténoïde découverte par Euler en 1744, surface minimale s'appuyant sur deux cercles parallèles. L'exemple $g(w) = w, f = 1$ donne la très jolie surface d'Enneper.





Caténoïde



Surface d'Enneper

Maintenant, on aimerait comprendre la question suivante. La caténoïde est une surface minimale s'appuyant sur deux cercles parallèles donnés comme sur la figure. Si les deux cercles ne sont pas trop éloignés, la bulle de savon prend cette forme. Mais si on écarte progressivement les deux cercles vient un moment où la bulle se casse et devient une surface beaucoup plus simple, à savoir deux disques de savon bordant les armatures circulaires (faire l'expérience est très simple et intéressant). En fait, on touche du doigt une aspect intéressant qu'on rencontre souvent dans les problèmes de minimisation du calcul des variations (voir [4]). Précisément, l'application qui à une surface associe son aire admet une différentielle nulle exactement pour les surfaces minimales. Donc, les minima sont à chercher parmi ces points. Toutefois, il se peut qu'on ne trouve que des minima locaux. C'est ce qui se passe ici : au bout d'un moment, l'aire des disques bordant nos cercles est plus petite que la caténoïde qui s'appuie dessus.

Exercice 5.10.1 (Difficile, sujet de réflexion). Trouver en fonction du rayon des cercles d'appui la distance entre les cercles d'appui de la bulle de forme caténoïde à partir de laquelle la bulle se scinde en deux.

Pour des raisons de minimisation de contraintes, on retrouve des surfaces minimales en architecture. Ici le toit du célèbre Arena Stadium de Munich.



FIGURE X.7 – Arena stadium

5.11 Application à la dispersion

Nous allons donner un aperçu d'une application de la formule de Cauchy globale aux phénomènes de dispersion (théorie de la réponse linéaire) connue des physiciens sous le nom de formules de Kramers et Kronig. Bien que découvertes à la fin du XIXème, elles sont d'une importance expérimentale très grande dans la physique la plus moderne et se rencontrent dans des domaines très divers de la physique [5]

Nous illustrerons ici le cas de la propagation des ondes électromagnétiques dans un diélectrique. Les deux ingrédients clefs y apparaissent clairement, l'un physique : le principe de causalité, l'autre mathématique : les propriétés d'holomorphie de la transformée de Fourier (2.3.3 page 196). Pour les détails, le lecteur se reportera au lumineux cours [18, chapitre 4].

Dans un milieu diélectrique homogène sans propriété magnétique, la transformée de Fourier en temps $\hat{E}(x,\omega)$ du champ électrique vérifie (Maxwell)

$$\Delta_x \hat{E} + \frac{\omega^2}{c^2}\epsilon_r(\omega)\hat{E} = 0 \text{ où } \epsilon_r(\omega) = \text{permétivité relative.}$$

En particulier (et de façon essentiellement équivalente), une onde plane de pulsation ω

$$E(x,t) = \exp(i(k.x - \omega t))$$

5. Voir le très joli article de physique « Causalité et relations de Kramers et Kronig » de N. G. van Kampen et F. Lurçat (*Le journal de physique et le radium*, 22 (1961), p. 179-191), le chapitre IX section 46 p. 458-61 du livre *Introduction à la Théorie Quantique des Champs* de N. Bogoliubov, D. Shirkov (Dunod (1960)), le chapitre 15 du livre *High Energy Hadron Physics* de M. Perl (Wiley (1974)). Pour des utilisations les plus récentes, voir par exemple la prépublication de M. Davier (« Evaluation of $\alpha(M_Z^2)$ and $(g-2)_\mu$ » (1998), arXiv :9812.370 (hep-ph)) où ces formules contribuent à étudier le modèle standard.

vérifie

(relation de dispersion) $$k = \frac{\omega}{c}\sqrt{\epsilon_r(\omega)}$$

Or, E est relié à la polarisation P par (linéarité, homogénéité ET causalité [6])

$$P(x,T) = \epsilon_0 \int_0^\infty G(t)E(T-t)\mathrm{d}t$$

où la fonction de Green G, est en général dans $L_1(\mathbf{R})$ à support dans $\mathbf{R}^+$ et ϵ_0 est la permétivité du vide. Le lien physique entre fonction de Green et permétivité étant décrit donné par les formules

$$\epsilon_r(\omega) = \epsilon_0(1+\chi) \text{ avec } \chi = \hat{G}(\omega) = \int_0^\infty G(t)\exp(-it\omega)\mathrm{d}t.$$

Mais on a la relation fondamentale

$$\hat{P} = \chi\hat{E} \text{ (liens Fourier et convolution ou exercice).}$$

En particulier, χ est complexe et holomorphe sur $\mathbf{H}$ (2.3.3 page 196) jamais nulle (physique). Physiquement,

$$\sqrt{\epsilon_r} = n + ia$$

avec $n(\omega)$ indice de réfraction et $a(\omega)$ coefficient d'absorption (dissipation [7]). En intégrant sur un contour convenable contenant $\omega_t = \omega - it, t << 1$ la fonction méromorphe

$$\frac{\chi}{\omega - \omega_t},$$

et en appliquant le théorème des résidus, on obtient (voir PC) la formule de Kramers et Kronig

$$\forall \omega \in \mathbf{R},\ \mathrm{Im}(\chi)(\omega) = -\frac{1}{\pi}\mathcal{PP}\int \frac{\mathrm{Re}(\chi)(x)}{x-\omega}\mathrm{d}x$$

où $\mathcal{PP}$ est la partie principale de Cauchy définie par

$$\mathcal{PP}\int \frac{\mathrm{Re}(\chi)(x)}{x-\omega}\mathrm{d}x = \lim_{\alpha\to 0}\int_{\mathbf{R}-[-\alpha,\alpha]} \frac{\mathrm{Re}(\chi)(x)}{x-\omega}\mathrm{d}x$$

L'importance expérimentale vient de ce que la formule précédente permet de calculer la réfraction en fonction de l'absorption et réciproquement, ce qui par exemple donne accès à l'indice de réfraction dans les milieux très absorbants.

6. On n'intègre que sur $\mathbf{R}^+$ d'après le principe de causalité qui impose que la réponse (qu'on a supposée linéaire d'où la forme de son expression sous forme intégrale) ne dépende que des temps passés

7. C'est la partie imaginaire de k qui donne un amortissement exponentiel dans l'onde plane.

6 Propriétés « globales » des fonctions holomorphes

6.1 Zéros de fonctions holomorphes

Proposition 6.1.1 (Théorème de Liouville)**.** *Soit $f \in \mathrm{Hol}(\mathbf{C})$ bornée. Alors, f est constante.*

Démonstration. En effet, on sait f est développable en série entière sur $\mathbf{C}$

$$f = \sum_n a_n z^n$$

avec de plus (estimées de Cauchy)

$$|a_n| \leq \sup |f|/R^n$$

pour tout $R > 0$ (3.2.4 page 203). On fait alors tendre R vers l'infini pour obtenir $a_n = 0$ si $n > 0$. □

Corollaire 6.1.2 (Théorème de d'Alembert-Gauss)**.** *Tout polynôme à coefficients complexes possède une racine dans $\mathbf{C}$. En d'autres termes, le corps $\mathbf{C}$ est algébriquement clos.*

Démonstration. Soit P un polynôme à coefficients complexe. Si pour tout $z \in \mathbf{C}$, $P(z)$ est non nul, alors $1/P$ est holomorphe sur $\mathbf{C}$ comme inverse d'une fonction holomorphe jamais nulle. Or,

$$\lim_{|z|\to\infty} |P(z)| = \infty$$

ce qui assure que $1/P$ est également borné sur $\mathbf{C}$ de sorte que $1/P$ est constant d'après le théorème de Liouville. □



Karl Friedriech Gauss



Jean le Rond d'Alembert

Ainsi les polynômes s'annulent sur $\mathbf{C}$, mais pas très souvent comme on le sait... C'est un phénomène très utile, qui se généralise aux fonctions holomorphes.

Théorème 6.1.3 (Principe des zéros isolés)**.** *Les zéros d'une fonction holomorphe non identiquement nulle sur un ouvert Ω connexe sont isolés,* ***i.e.****si $f(z_0) = 0$ et $0 < |z - z_0| << 1$ alors $f(z) \neq 0$.*

Démonstration. Soit z_0 un zéro de f. L'ensemble

$$U = \{z \in \Omega \text{ tels que } \forall n \geq 0,\ f^{(n)}(z) = 0\}$$

est fermé (comme intersection de fermés) et ouvert (car f est somme de sa série de Taylor au voisinage de tout point z). Comme f est non identiquement nulle, $U \neq \Omega$ de sorte que $U = \varnothing$(connexité). Il existe donc $N > 0$ minimal tel que $f^{(N)}(z_0) \neq 0$. On écrit alors au voisinage de z_0

$$f(z) = (z - z_0)^N (a_N + z \sum_{n>N} a_n (z - z_0)^{n-N})$$

ce qui assure

$$f(z) \sim_{z_0} a_N (z - z_0)^N$$

et donc f ne s'annule pas dans un voisinage épointé de z_0. □

Corollaire 6.1.4 (Principe du prolongement analytique)**.** *Deux fonctions holomorphes sur un ouvert connexe coïncidant sur un ouvert, voire sur un arc non constant, coïncident partout.*

Remarque 6.1.5. La connexité n'entre en jeu que pour avoir un joli énoncé. La preuve permet de conclure que si les zéros ne sont pas isolés, la fonction holomorphe est nulle sur une des composantes connexes (nécessairement ouverte) de Ω.

6.2 Propriété de la moyenne et principe du maximum

Théorème 6.2.1 (Formule de la moyenne)**.** *Soit $f \in \mathrm{Hol}(\Omega)$ et $\overline{D(z_0, R)}$ disque compact contenu dans Ω. Alors,*

$$f(z_0) = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + R\,\mathrm{e}^{i\theta}) \mathrm{d}\theta.$$

Démonstration. On a

$$f(z_0) \overset{Cauchy}{=} \frac{1}{2i\pi} \int_{\partial D(z_0,R)} \frac{f(z)}{z - z_0} \mathrm{d}z$$

$$= \frac{1}{2i\pi} \int_0^{2\pi} \frac{f(z_0 + R\,\mathrm{e}^{i\theta})}{R\,\mathrm{e}^{i\theta}} iR\,\mathrm{e}^{i\theta}\,\mathrm{d}\theta = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + R\,\mathrm{e}^{i\theta}) \mathrm{d}\theta.$$

□

Corollaire 6.2.2.

$$|f(z_0)| \leq \mathrm{Max}_\theta |f(z_0 + R\,\mathrm{e}^{i\theta})|.$$

Autrement dit, le maximum de $|f|$ est atteint sur le bord du disque. L'énoncé est beaucoup plus général et constitue un point important de la théorie.

Théorème 6.2.3 (Principe du maximum)**.** *Soit $f \in \mathrm{Hol}(\Omega)$. Si le module $|f|$ a un maximum local en $z_0 \in \Omega$ alors f est constante au voisinage z_0, donc partout si Ω est connexe (zéros isolés).*

Démonstration. Soit $\overline{D(0,R)} \subset \Omega$ et $r \leq R$. Soit

$$c_n = \frac{1}{2\pi}\int_0^{2\pi} f(z_0 + r\,\mathrm{e}^{i\theta})\,\mathrm{e}^{-in\theta}\,\mathrm{d}\theta$$

le n-ième coefficient de Fourier $f_r(\theta) = f(z_0 + r\,\mathrm{e}^{i\theta})$.

Supposons que

$$|c_0| = \Big|\frac{1}{2\pi}\int_0^{2\pi} f(z_0 + r\,e^{i\theta})\Big| \overset{moyenne}{=} |f(z_0)| = \mathrm{Max}_{\overline{D(z_0,R)}}|f(z)|.$$

Alors, d'après l'égalité de Parseval (cf. exemple 3.1.2 page 157), on a

$$|c_0|^2 \leq \sum |c_n|^2 \overset{\text{Parseval}}{=} \frac{1}{2\pi}\int_0^{2\pi} |f(z_0 + r\,e^{i\theta})|^2 \mathrm{d}\theta \leq \frac{1}{2\pi}\int_0^{2\pi} |c_0|^2 \mathrm{d}\theta = |c_0|^2.$$

et donc $c_n = 0, n \neq 0$ **i.e.**$||f_r - c_0||_2 = 0,\ \forall r < R$ ce qui entraîne $f = c_0$ sur $D(z_0, R)$ par continuité. □

Remarque 6.2.4. La preuve n'utilise que continuité et formule de la moyenne : elle s'adaptera sans modification aux fonctions harmoniques.

Exercice 6.2.5. Montrer que si f holomorphe sur Ω connexe est à valeurs réelles, alors f constante (considérer $\exp(if)$).

6.3 Existence globale de primitives

Théorème 6.3.1. *Toute fonction holomorphe f définie sur un ouvert Ω ouvert simplement connexe admet une primitive holomorphe (unique à un scalaire près.)*

Démonstration. Soit $z_0 \in \mathbf{C}$. L'ensemble des points pouvant être reliés à z_0 par un chemin régulier (même polygonal) est ouvert et fermé : par connexité, c'est Ω. Pour tout $\omega \in \Omega$, choisissons Γ_ω régulier reliant z_0 à ω et posons

$$F(\omega) = \int_{\Gamma_\omega} f(z)\mathrm{d}z.$$

Si Γ'_ω est un autre choix, $\Gamma'_\omega \cup \Gamma_\omega^{opp}$ est un lacet, homotope à zéro par hypothèse. D'après l'invariance par homotopie de la circulation (5.6.2 page 213), on a

$$\int_{\Gamma_\omega} f(z)\mathrm{d}z = \int_{\Gamma'_\omega} f(z)\mathrm{d}z$$

et F est bien donc définie. Puisque le choix du chemin est indifférent pour le calcul de F, on a pour $D(\omega, r) \subset \Omega$ la formule

$$\frac{F(\omega+h) - F(\omega)}{h} = \int_{[\omega,\omega+h]} f(z)\mathrm{d}z = \int_0^1 f(\omega + th)\mathrm{d}t \overset{h\to 0}{\to} f(\omega)$$

ce qui assure que F est $\mathbf{C}$-dérivable et $F' = f$. Grâce aux propriétés de dérivabilité des intégrales, cette formule prouve également que F est C^1 de sorte que F est bien holomorphe. □

Cet énoncé permet de construire de nombreuses fonctions holomorphes.

6.4 Application au logarithme

On se donne dans cette section $h \in \mathrm{Hol}(\Omega, \mathbf{C}^*)$ de sorte que $h'/h \in \mathrm{Hol}(\Omega)$.

Définition 6.4.1. On appelle logarithme de h et on note $\ln(h)$ toute primitive holomorphe, lorsqu'elle existe, de h'/h.

Un logarithme complexe sur un ouvert connexe est donc bien défini à addition d'un scalaire près. On a le corollaire immédiat du théorème 6.3.1 page ci-contre d'existence globale de primitives.

Proposition 6.4.2. *Si Ω est simplement connexe, alors h admet un logarithme complexe.*

On pose alors

$$h(z)^\alpha = \exp(\alpha \ln(h(z)), \alpha \in \mathbf{C}.$$

En particulier, on peut définir $\ln(z)$ sur la **coupure** $\Omega = \mathbf{C} - \mathbf{R}^-$ comme **la** primitive de $1/z$ s'annulant en $z = 1$. On l'appelle **la détermination principale** du logarithme. Elle permet par exemple de définir $\sqrt{z}$.

Exercice 6.4.3. Soit $\ln(z)$ la détermination principale du logarithme. Montrer

$$\ln(r\,\mathrm{e}^{i\theta}) = \ln(r) + i\theta, \theta \in]-\pi, \pi[.$$

Montrer qu'il n'existe pas de prolongement continu du logarithme à $\mathbf{C}^*$ [Sinon on aurait $\int_{C(0,1)} \frac{dz}{z}(z)\mathrm{d}z = 0$ d'après le théorème 5.2.1 page 206.].

Exercice 6.4.4. Montrer les propriétés

$$\ln(ab) = \ln(a) + \ln(b) + C_1$$

et les formules

$$\ln(\exp(c)) = c + C_2, \exp(\ln(a)) = C_3 a$$

pour tout $a, b \in \mathrm{Hol}(\Omega, \mathbf{C}^*), \mathrm{c} \in \mathrm{Hol}(\Omega)$ pour des constantes convenables C_1, C_2, C_3 de $\mathbf{C}^*$ (Ω est supposé simplement connexe).

7 Fonctions harmoniques

Définition 7.0.5. La fonction $f : \Omega \to \mathbf{C}$ est dite *harmonique* si

$$\Delta(f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial x^2} = 4\frac{\partial}{\partial z}\frac{\partial f}{\partial \bar{z}} = 4\frac{\partial}{\partial \bar{z}}\frac{\partial f}{\partial z} = 0.$$

Par exemple, si f est holomorphe, $f, \bar{f}, \mathrm{Re}(f), \mathrm{Im}(f)$ sont harmoniques. Ce sont essentiellement les seuls exemples.

Théorème 7.0.6. *Toute fonction harmonique h sur Ω simplement connexe s'écrit sous la forme $f + \bar{g}$ avec $f, g \in \mathrm{Hol}(\Omega)$.*

Démonstration. On a $\frac{\partial}{\partial \bar{z}}\frac{\partial h}{\partial z} = 0$ donc $\frac{\partial h}{\partial z}$ holomorphe. Soit f primitive holomorphe de $\frac{\partial h}{\partial z}$. On a

$$0 = \overline{\frac{\partial h}{\partial z} - \frac{\partial f}{\partial z}} = \overline{\frac{\partial (h-f)}{\partial z}} = \frac{\partial (\overline{h-f})}{\partial \bar{z}}$$

donc $\overline{h-f} = g \in \mathrm{Hol}(\Omega)$. □

Avec les notations de la preuve, si h est de plus réelle, on a $h = \mathrm{Re}(f+g)/2$:

Corollaire 7.0.7. *Une fonction harmonique réelle sur un ouvert simplement connexe Ω est la partie réelle d'une fonction holomorphe, unique à $i\mathbf{R}$ près.*

En appliquant la formule de la moyenne à f et g, on obtient par linéarité (cf. la remarque 6.2.4 page 222)

Corollaire 7.0.8. *Une fonction harmonique sur un ouvert quelconque vérifie la propriété de la moyenne et donc également le principe du maximum.*

Remarque 7.0.9. L'importance du principe du maximum vient notamment du fait qu'il donne bien souvent l'unicité de fonctions harmoniques vérifiant des conditions au bord fixées (cf. 7.0.12 et les applications plus bas (8) par exemple).

Un certain nombres de résultats s'adaptent du cas holomorphe au cas harmonique. À titre d'exercice, on prouvera les résultats utiles suivants.

Exercice 7.0.10. Calculer le laplacien en coordonnées polaires. En déduire que les fonctions $a\ln(|z|) + b, a, b \in \mathbf{C}$ sont les seules fonctions harmoniques à symétrie de révolution. Donner des illustrations physiques de ce résultat.

Exercice 7.0.11. Une fonction harmonique bornée sur $\mathbf{C}$ est constante.

Exercice 7.0.12. Si $f \in L^\infty(\bar{D})$ est harmonique sur $D = D(0,1)$ et a un partie réelle nulle pp sur ∂D, alors $f \equiv 0$ [Écrire $\mathrm{Re}(f) = \mathrm{Re}(g)$ avec $g = \sum a_n z_n \in \mathrm{Hol}(\Omega)$ et comparer les coefficients de Fourier de $\theta \mapsto \mathrm{Re}(f(r\,\mathrm{e}^{i\theta}))$ et $a_n r^n$ lorsque r tend vers 1].

Exercice 7.0.13. Montrer que $z \mapsto \frac{z-i}{z+i}$ est un biholomorphisme $D = D(0,1) \xrightarrow{\sim} \mathbf{H}$. En déduire que si $f \in L^\infty(\bar{H}, \mathbf{R})$ est harmonique sur $\mathbf{H}$ et nulle p.p. sur $\mathbf{R} = \partial\mathbf{H}$ alors f est identiquement nulle.

8 Application physiques -idéalisées-

Si f, g, h sont holomorphes, $f \circ h + \bar{g} \circ h$ est harmonique : le composé d'une fonction harmonique et d'une transformation holomorphe est harmonique. Or, on a le (difficile et profond) théorème de Riemann :

Théorème 8.0.14 (Riemann)**.** *Tout ouvert simplement connexe de $\mathbf{C}$ est biholomorphe à $\mathbf{C}$ ou au demi-plan de Poincaré $\mathbf{H} = \{z | \mathrm{Im}(z) > 0\}$.*



FIGURE X.8 – Bernhard Riemann

Ceci permet bien souvent de se ramener au demi-plan ou au plan par transformation conforme comme dans l'exercice précédent. Précisément, si on cherche f harmonique sur Ω simplement connexe vérifiant des conditions aux limites $\Lambda(\omega)$, on trouve un biholomorphisme h de Ω sur $\mathbf{C}$ ou $\mathbf{H}$ et on cherche ϕ sur $\mathbf{H}$ ou Ω harmonique vérifiant $\Lambda \circ h(z)$. Or, nombre de régimes stationnaires sont définis par l'équation $\Delta(f) = 0$: les champs magnétiques ou électriques en électrostatique (Maxwell), le potentiel des vitesses d'un fluide parfait, le champ de gravitation classique...

Exercice 8.0.15. Montrer que $\mathbf{H}$ et $\mathbf{C}$ ne sont pas biholomorphes.

8.1 Écoulement d'un fluide

Considérons un fluide parfait irrotationnel et incompressible s'écoulant sans frottement dans un domaine du plan Ω simplement connexe. Le champ des vitesses $\overrightarrow{v}$ satisfait donc $\mathrm{rot}(\overrightarrow{v}) = 0$ (irrotationnel), donc s'écrit $\overrightarrow{v} = -\nabla(\Phi)$. Comme il est à divergence nulle (incompressibilité), on a

$$-\mathrm{div}(\nabla(\Phi)) = -\Delta(\Phi) = 0$$

de sorte que

$$v_x = -\partial_x\Phi, v_y = -\partial_y\Phi$$

sont harmoniques avec la composante normale v_y nulle au bord (absence de frottements). Le théorème de Riemann dicte de s'intéresser d'abord à $\Omega = \mathbf{H}$ et à $\overrightarrow{v} \in L^\infty(\bar{\mathbf{H}})$ (la vitesse est bornée pour des raisons physiques !).

Comme v_y est harmonique bornée et nulle sur le bord, elle est identiquement nulle (principe du maximum, cf. exercice 7.0.12 page ci-contre). Ansi, $\partial_y\Phi = -v_y \equiv 0$, **i.e.**$\Phi = \Phi(x)$ et $\partial_x v_x = -\partial_x^2 v_x = -\Delta\Phi = 0$ de sorte $v_x = cnte$ et $\overrightarrow{v} = (cnte, 0)$.

Pour un domaine Ω général, il suffit de trouver une fonction holomorphe, f, envoyant $\mathbf{H}$ sur Ω (comme f est conforme, f préserve la dérivée normale). Par exemple, pour le « coin » d'équation polaire $0 \leq \theta \leq \alpha$, on considère $f(z) = z^{\alpha/\pi}$ ce qui donne comme champ des vitesses - $\mathrm{Re}((v_0 z)^{\alpha/\pi})$.

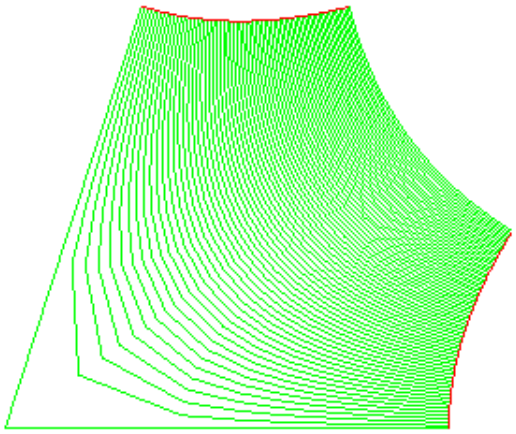


FIGURE X.9 – Écoulement dans un coin

8.2 Propagation de la chaleur

Une gaine rectiligne infinie a une section en couronne (non simplement connexe !) de rayons $R^- < R^+$ et baigne dans des calorimètres à l'intérieur et à l'extérieur à température T^-, T^+. En régime permanent ($\frac{\partial T}{\partial t} = 0$), la température T est harmonique car l'équation de la chaleur s'écrit

$$\frac{\partial T}{\partial t} - \Delta(T) = 0$$

et T vaut T^+, T^- sur les cercles de bords de rayons T^+, T^-. La différence de deux solutions est nulle au bord donc nulle (principe du maximum) : si le problème possède une solution, celle-ci est unique. Les fonctions harmoniques invariantes par rotation (7.0.10 page 224) sont de la forme $a \ln|z| + b$ et donc, si la couronne est centrée en zéro,

$$T_{\text{centrée}} = \frac{T_-}{\ln(R^-) - \ln(R^+)} \ln(\frac{|z|}{R^+}) + \frac{T_+}{\ln(R^+) - \ln(R^-)} \ln(\frac{|z|}{R^-}).$$

Dans le cas général, on remarque qu'on peut centrer une couronne par une homographie f, qui préserve les cercles de bord :la température sera de la forme $T_{\text{centrée}} \circ f^{-1}$.

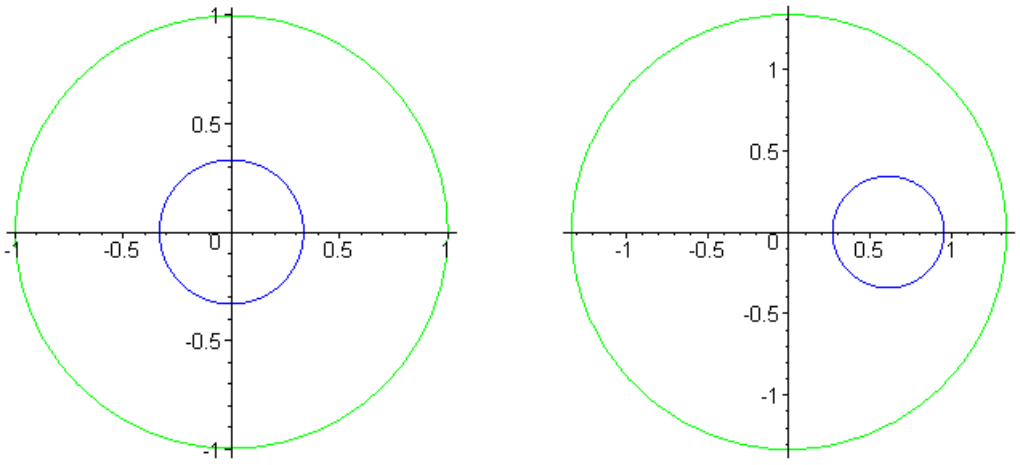


FIGURE X.10 – Transformation conforme de couronnes

$$z \mapsto \frac{5}{3} + \frac{z-3}{z-2}$$

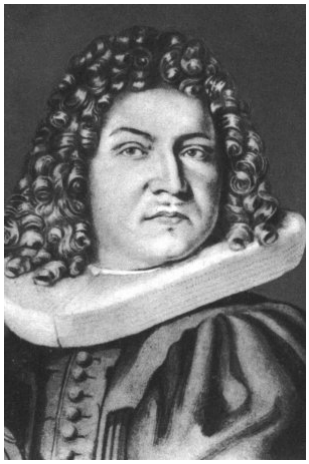
9 Applications mathématiques

Les propriétés des fonctions holomorphes, d'une ou plusieurs variables, ont de multiples applications. En voici une aussi jolie qu'élémentaire. Elle relie les sommes

$$\zeta(2k) = \forall k > 0,\ \sum_{n=1}^{\infty} n^{-2k}, k \in \mathbf{N}^*$$

aux nombres de Bernoulli, qui à des constantes explicites près, sont définis comme les coefficients du développement en série de la cotangente par la formule

$$z\cot(z) =:= \sum_{n=0}^{\infty} \frac{(-1)^n 4^n B_{2n}}{(2n)!} z^{2n},\ |z| < \frac{\pi}{2}.$$



Jakob Bernoulli

Les pôles de $\cot(\pi z)$ sont les entiers. Par périodicité, il suffit pour calculer son développement de Laurent aux pôles de le faire en $z = 0$. Des développements usuels des sinus et cosinus, on tire

$$\cot(\pi z) = \frac{1}{z} + \text{holomorphe.}$$

La série $\sum_{n>0} \frac{1}{z-n} + \frac{1}{z+n}$ converge normalement sur les compacts de $\mathbf{C} - \mathbf{Z}$ de sorte que

$$\frac{1}{z} + \sum_{n>0} \frac{1}{z-n} + \frac{1}{z+n}$$

est méromorphe sur $\mathbf{C}$ (4.0.6 page 204). La fonction

$$g(z) := \cot(\pi z) - (\frac{1}{z} + \sum_{n>0} \frac{1}{z-n} + \frac{1}{z+n})$$

est holomorphe en 0 (on a « tué » le pôle) donc partout car elle est 1-périodique.. La périodicité (et la continuité) assure qu'elle est bornée sur une bande $|\operatorname{Im}(z)| \leq 1$. En dehors de la bande, elle est bornée (calcul direct). Le théorème de Liouville assure qu'elle est constante, donc identiquement nulle puisqu'elle s'annule à l'origine. On a donc

(*Euler*). $$\cot(\pi z) - (\frac{1}{z} + \sum_{n>0} \frac{1}{z-n} + \frac{1}{z+n})$$

En développant en série, on obtient

$$\forall k > 0,\ \sum_{n=1}^{\infty} n^{-2k} = \frac{(2\pi)^{2k} B_{2k}}{2(2k)!}$$

Les nombres de Bernoulli B_n sont des nombres rationnels aux propriétés remarquables. Les premières valeurs sont $B_0 = 1, B_2 = 1/6, B_4 = -1/30 \cdots$

Les valeurs de

$$\forall k > 0,\ \sum_{n=1}^{\infty} n^{-2k-1}$$

restent très mystérieuses même si des conjectures difficiles existent ainsi que des progrès récents (Apéry, Rivoal...).

9.1 Le théorème de Müntz-Szasz

La preuve du résultat suivant applique de façon un peu surprenante la théorie des fonctions holomorphe à la théorie des Hilbert.

Théorème 9.1.1 (Müntz-Szasz)**.** *Soit* (λ_n) *une suite réelle positive et* e_n *la fonction* $e^{i\lambda_n t}$ *définie sur* $[0,1]$*. Si* λ_n *tend vers l'infini et* $\sum_n \frac{1}{\lambda_n} = +\infty$*, alors l'espace engendré par les* e_n *est dense dans* $L^2([0,1], \mathbf{C})$*.*

Démonstration. Sinon il existerait une fonction $f \in L^2([0,1], \mathbf{C})$ telle que pour tout n

$$\int_0^1 f(t)e^{-i\lambda_n t}dt = 0.$$

La fonction $z \longrightarrow F_f(z) = \int_0^1 f(t)e^{-izt}dt$ est définie et holomorphe sur $\mathbf{C}$. Elle vérifie $|F_f(z)| \leq e^{\mathrm{Im}(z)}$ et s'annule en les λ_n. On en déduit que $F_f(\frac{1+iz}{z+i})$ est bornée sur le disque unité (car $z \longrightarrow \frac{1+iz}{z+i}$ envoie $|z| \leq 1$, sur $\mathrm{Im}(z) \leq 0$, cf. l'exercice 7.0.13 page 224) et s'annule pour $\alpha_n = (\frac{1+i\lambda_n}{\lambda_n+i})$. Nous allons montrer que ceci est impossible si F_f est non nul et $\sum_n(1-|\alpha_n|) = +\infty$. En effet

Lemme 9.1.2. *Soit*

$$G_n(z) = \prod_{i=1}^{n} \frac{(\alpha_i - z)}{(1 - \bar{\alpha}_i z)}$$

Si $\sum_{j=1}^{+\infty}(1-|\alpha_i|) = +\infty$*, alors la suite* $|G_n(z)|$ *tend uniformément vers une fonction* g *qui s'annule sur un ensemble ayant* 0 *comme point d'accumulation.*

Remarque 9.1.3. En fait des arguments un peu plus sophistiqués de théorie des fonctions holomorphes montrent que G_n tend uniformément vers 0 sur tout compact contenu dans $D(0,1)$ (exercice).

Démonstration. Tout d'abord si $|z| = 1$, $\left|\frac{(\alpha_i - z)}{(1-\bar{\alpha}_i z)}\right| = \left|\frac{(\alpha_i - z)}{(\bar{z}-\bar{\alpha}_i)}\right| = 1$. Par le principe du maximum, $\left|\frac{(\alpha_i - z)}{(1-\bar{\alpha}_i z)}\right| \leq 1$ si $|z| \leq 1$, donc $|G_n(z)| \leq 1$ sur le disque unité, et la suite $|G_n(z)|$ est décroissante. Supposons d'abord que les α_j n'aient pas de point d'accumulation à l'intérieur du disque. Par le théorème de Dini (4.2.1 page 16), la suite $|G_n(z)|$ converge uniformément vers une fonction $g(z)$ positive ou nulle. Or la condition sur les α_n entraîne que $|G_n(0)| = \prod_{j=1}^{n} |\alpha_j|$ converge vers 0 (cf. exercice (9) page 187). Maintenant la fonction $f_n(z) = -\log(|G_n(z)|)$ est positive, harmonique, et $f_n(0)$ tend vers $+\infty$. Le principe du maximum entraîne que pour tout cercle de rayon $r < 1$ centré en 0, il existe un point $z_n(r)$ tel que $f_n(z_n(r)) \geq f_n(0)$. En extrayant, par compacité du cercle, une suite convergente de $z_n(r)$ de limite $z(r)$ on a $\lim_n G_n(z_n(r)) = g(z(r)) = 0$. On en conclut que tout cercle de rayon r contient un zéro de g. □

Nous pouvons maintenant terminer la démonstration du théorème : $G_n(z)$ est de module borné par 1 sur le disque et s'annule en les mêmes points et au même ordre que F_f, donc la fonction holomorphe $F_f(z)/G_n(z)$ est définie, sans zéro sur $D(0,1)$, bornée par 1 sur S^1. D'après le principe du maximum (théorème 6.2.3 page 221) , $|F_f(z)| \leq |G_n(z)|$ sur $D(0,1)$. Mais comme $\lim_{n\to\infty} |G_n(z)|$ a des zéros ayant un point d'accumulation en 0, la même chose est vraie pour F_f, qui est holomorphe, ce qui entraîne la nullité de cette fonction (d'après le principe des zéros isolés). □

Remarque 9.1.4. Contrairement au cas des bases Hilbertiennes, on voit que si l'on retire une fonction de la famille des $e^{i\lambda_n t}$, celle ci engendre toujours un espace dense. On peut aussi montrer réciproquement que si $\sum_n \frac{1}{\lambda_n} < +\infty$, l'espace engendré par les $e^{i\lambda_n t}$ n'est pas dense (c'est un exercice classique de calcul de la distance d'un point à un sous-espace).

9.2 Application de la notion d'indice au théorème de séparation de Jordan

Nous allons montrer comment la notion d'indice d'un lacet permet de répondre à la question suivante : en combien de régions une courbe fermée sans point double partage-t-il le plan. Intuitivement, on pense que la réponse est 2, l'intérieur et l'extérieur. C'est vrai, mais le dessin suivant montre que ce n'est pas si facile que cela peut paraître : le point rouge est-il extérieur ou intérieur ? Y a-t-il vraiment un intérieur et un extérieur ?

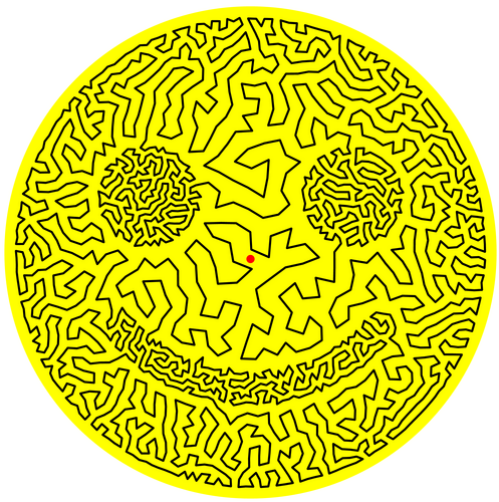


FIGURE X.11 – Quid du point rouge ?

Nous présentons ce thème comme un sujet de réflexion : nous laissons les preuves des lemmes en exercice. Précisons le problème.

Définition 9.2.1. Une courbe de Jordan est une application continue $\gamma : I = [0, L] \longrightarrow \mathbf{C}$ telle que $\gamma(0) = \gamma(L)$ qui est injective sur $[0, L[$. On dit qu'elle est régulière si de plus γ est C^1 et $\gamma'(t) \neq 0$ pour tout t.

Une autre manière de voir est de dire que γ est une injection continue d'un cercle dans $\mathbf{C}$ (exercice : Montrer l'équivalence des deux définitions). L'arc $\Gamma := \gamma(I)$ est compact.

Théorème 9.2.2 (Jordan)**.** *L'ouvert $\Omega = \mathbf{C} - \gamma$ a deux composantes connexes, l'une bornée l'autre non. Si γ est régulier (orientée dans le sens trigonométrique), la composante bornée est constituée des points z tels que* $\mathrm{ind}_\gamma(z) = 1$.

Démonstration. On va juste donner les grandes lignes dans le cas régulier laissant les preuves en exercice. On prolonge γ par L-périodicité. Quitte à changer de paramétrage, on peut supposer

que Γ est paramétré par longueur d'arc, **i.e.**$|\gamma'(t)| = 1$ pour tout t. Quitte à faire une translation et une homotéthie, on peut supposer $\gamma(0) = 0, \gamma'(0) = 1$.

1) On va pousser un peu Γ le long des normales $i\gamma'(t)$. Soit donc $\epsilon \neq 0$ et

$$\gamma_\epsilon(t) = \gamma(t) + i\epsilon\gamma'(t).$$

On note Γ_ϵ les arcs images.

Lemme 9.2.3 (Assez difficile)**.** *Il existe $\alpha > 0$ tel que si $0 < |\epsilon| < \alpha$ est assez petit, Γ_ϵ ne rencontre pas Γ.*

2) Définissons $P^\pm \subset \mathbf{C}$ comme la réunion

$$P^\pm = \bigcup_{0<\epsilon<\alpha} \gamma_{\pm\epsilon}(\mathbf{R}).$$

On observe que P^+ et P^- sont des parties connexes de $\mathbf{C} - \Gamma$ comme images continue du connexe $]0, \alpha[\times\mathbf{R}$.

3) On montre que $\mathbf{C} - \Omega$ a au plus deux composantes connexes en reliant (dans $\mathbf{C} - \Omega$) $z \notin \Gamma$ par un segment à un point de $P^\pm$.

Lemme 9.2.4. *Soit t_0 tel que $\mathrm{d}(z, \Gamma) = \mathrm{d}(z, \gamma(t_0))$. Montrer en calculant la dérivée en t_0 de $|z - \gamma(t)|^2$ que $z - \gamma(t_0)$ est colinéaire à $\gamma'(t_0)$. En déduire qu'il existe un segment $[z, z^\pm]$ contenu dans la normale à Γ en $\gamma(t_0)$ qui est inclus dans $\mathbf{C} - \Gamma$.*

4) On montre qu'il y a au moins deux composantes. Par continuité de l'indice, $z \mapsto \mathrm{ind}_z(\gamma)$ est constant sur les parties connexes de $\mathbf{C} - \Omega$ puisque l'indice est à valeurs entière.

Lemme 9.2.5. *On a $\lim_{\epsilon \longrightarrow 0^+} |\mathrm{ind}_{i\epsilon}(\gamma) - \mathrm{ind}_{-i\epsilon}(\gamma)| = 1$.*

On déduit que l'indice prend au moins deux valeurs et donc que $\mathbf{C} - \Omega$ n'est pas connexe.

5) D'après 3) et 4), $\mathbf{C} - \Omega$ a deux composantes C_N, C_{N+1} distinguées par des indices différents N et $N + 1$. Quitte à changer le sens de γ, on peut supposer $N \geq 0$. Je dis que C_{N+1} est bornée. En effet, une application de la convergence dominée donne

$$\lim_{|z| \longrightarrow \infty} \mathrm{ind}_z(\gamma) = 0.$$

On en déduit que si (z_n) est une suite de C_{N+1} tendant vers l'infini, l'indice de z_n est nul pour n assez grand (c'est un entier) contredisant $N + 1 > 0$. Comme Γ est borné, C_N est nécessairement non bornée. Notons qu'on a prouvé au passage que N est nul. □

L'illustration graphique est alors comme suit : l'intérieur est en rouge et notre point est intérieur.



FIGURE X.12 – Le point rouge est intérieur

Le cas continu est un peu plus délicat. On invite le lecteur à consulter des ouvrages de topologie algébrique (par exemple [9] ou [11])) pour avoir des preuves de type topologie algébriques ouvrant des perspectives très grandes (théorème de Brouwer, d'invariance du domaine...).

On peut montrer mieux, à savoir le théorème de Jordan-Schönflies : il existe un homéomorphisme du plan qui transforme Γ en le cercle usuel (de sorte que l'énoncé sur les composantes devient évident). En particulier, l'extérieur d'une courbe de Jordan est homéomorphe au complémentaire d'un disque et l'intérieur au disque.

La généralisation naturelle en dimension supérieure du théorème de séparation de Jordan est vraie. C'est le théorème de Jordan-Brouwer qui affirme qu'un plongement de la sphère euclidienne sépare $\mathbf{R}^n$ en deux composantes. Mais on doit dans ce genre de problème se méfier de l'intuition. En effet, l'analogue dans l'espace du théorème de Jordan-Schönflies est faux. Le mathématicien John Alexander, après avoir cru prouver que le théorème se généralisait, a donné en 1921 un contre-exemple, la sphère à cornes d'Alexander : son extérieur n'est pas homéomorphe au complémentaire d'une boule ([11, 2B.2]) ! Elle ressemble au dessin suivant.



FIGURE X.13 – Sphère à cornes

10 Récapitulatifs

10.1 Liste des exercices de fonctions holomorphes

Exercice 2.3.4 page 196

Exercice 2.3.6 page 197

Exercice 2.3.3 page 196

Exercice 5.3.6 page 208

Exercice 5.3.7 page 208

Exercice 5.3.8 page 208

Exercice 5.10.1 page 217

Exercice 5.4.3 page 209

Exercice 5.5.5 page 211

Exercice 6.2.5 page 222

Exercice 6.4.3 page 223

Exercice 6.4.4 page 223

Exercice 7.0.11 page 224

Exercice 7.0.12 page 224

Exercice 7.0.13 page 224

Exercice 8.0.15 page 225

Exercice 7.0.10 page 224

10.2 Ce qu'il faut retenir

Une fonction $f = P + iQ : \Omega \longrightarrow \mathbf{C}$ de classe C^1 est holomorphe sur Ω si

$$\frac{\partial f}{\partial \bar{z}}(z_0) := \frac{1}{2}(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y})$$

nul sur Ω. Dans ce cas, on note

$$f' := \frac{\partial f}{\partial z} :=:= \frac{1}{2}(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y})$$

la dérivée (complexe) de f.

Les théorèmes de stabilité pour les opérations usuelles sont vérifiées pour les fonctions holomorphes.

Si f est C^1 sur Ω, les 3 propriétés suivantes sont équivalentes

(1) $f \in \mathrm{Hol}(\Omega)$;

(2) f est $\mathbf{C}$-dérivable ;

(3) $\mathrm{d}f$ est une similitude (f conforme) ;

(4)
$$\frac{\partial P}{\partial x} = \frac{\partial Q}{\partial y} \text{ et } \frac{\partial P}{\partial y} = -\frac{\partial Q}{\partial x}.$$

On a la formule de Cauchy pour un cercle C bordant un disque ouvert D contenu dans $\Omega, f \in \mathrm{Hol}(\Omega)$:

$$\forall z_0 \in D, f(z_0) = \frac{1}{2i\pi} \int_C \frac{f(z)\mathrm{d}z}{z - z_0}.$$

Les fonctions holomorphes sur un disque $D(z_0, R)$ sont exactement les fonctions développables en séries entières

$$\sum_{n \geq 0} a_n (z - z_0)^n$$

de rayon de convergence $\geq R$. Les fonctions holomorphes sur un ouvert sont donc exactement les fonctions analytiques. Elles sont sommes de leur série de Taylor en tout point.

Une fonction définie au voisinage de z_0 est méromorphe au voisinage de z_0 si f est holomorphe sur un voisinage épointé de z_0 et s'il existe $N \geq 0$ tel que $(z - z_0)^N f$ est borné. La fonction f admet localement un unique développement de Laurent

$$f(z) = \sum_{n \geq -N} a_n (z - z_0)^n$$

et on pose

$$\mathrm{res}_{z_0}(f) = a_{-1}.$$

Les pôles de f sont isolés.

Soit γ est un chemin régulier dans Ω, paramétrisation propre du bord d'un ouvert borné U. Soit f une fonction méromorphe sur Ω, sans pôle sur γ. Soit S l'ensemble fini de ses pôles dans U. Alors, si Ω est simplement connexe (par exemple étoilé), on a

$$\int_\gamma f(z)\mathrm{d}z = 2i\pi \sum_{s \in S} \mathrm{res}_s(f).$$

Soit $f, g \in \mathrm{Hol}\Omega$.

Principe des zéros isolés. Les zéros de f sont isolés. Autrement, si f et g coïncident sur une infinité de points ayant une valeur d'adhérence dans Ω connexe, elles sont égales.

Propriété de la moyenne. Soit $\bar{D} \subset \Omega$ un disque compact contenue dans Ω. Alors, pour tout $z_0 \in D$, $f(z_0$ est la moyenne de f sur le cercle $\partial\bar{D}$.

Principe du maximum. Si f a un maximum local z_0, alors f est constante sur la composante connexe de Ω contenant z_0.

Les fonctions harmoniques sur Ω sont les fonctions de laplacien nul. Elles s'écrivent $f + \bar{g}$ avec f et g holomorphes, et $\mathrm{Re}(f)$ si on les suppose de plus à valeurs réelles. Elles vérifient le principe du maximum et la formule de la moyenne.

Index

$H^k(\mathbf{N})$, 148
$L^2(\mathbf{S}^1, \mathbf{C})$, 148
$\ell^2(\mathbf{N})$, 147
Égalité de la médiane, 146
Égalité du parallélogramme, 146
Égalite de Parseval, 156
Équation
    de Schrödinger, 181
    de Sturm-Liouville, 181
Équivalence des normes, 18

Additivitédénombrable, 99
Adhérence
    d'une partie, 9
    valeur d', 10
Adjoint d'un opérateur, 155
Analytique
    fonction, 202
Anti-linéaire, 146
Arithmétique dans $[0, +\infty]$, 35

Banach, 13
Base Hilbertienne, 155
Bernoulli
    nombre de, 227, 228
Biholomorphisme, 198
Bolzano-Weierstrass, 18
Borel-Lebesgue, 16
Bulles de savon, 216

Cantor
    argument diagonal de, 24
Caténoïde, 217
Cauchy
    condition de Cauchy complexe, 196
    conditions de Cauchy réelles, 197
    formule de Cauchy pour un cercle, 200
    suite de, 12
Chaleur, 226
Circulation, 205
Classe $\mathcal{L}^+$, 32
Compacité du cube de Hilbert, 166
Compacité séquentielle faible de la boule unité, 165
Compact, 15
    opérateur, 177
Complété d'un espace préhilbertien, 149
Complétude
    de L^1, 116
    de L^p, 122
Complet
    espace métrique, 11
Concentration, 52
Conforme
    transformation, 198
Connexe, 22
    composante, 22
    par arcs, 22
Continuité L^1 des translations, 120
Continuité uniforme, 18
Convergence faible, 164
Coordonnées sphériques, 87
Critère
    de continuité des applications linéaires, 150
    de continuité pour les formes sesquilinéaires, 154
    de densité, 155
Cube de Hilbert, 166

Décomposition en série de Fourier sur $L^2([0, 1], \mathbf{C})$, 157
Décomposition orthogonale d'un espace de Hilbert, 153
Dénombrabilité, 23
Dérivation
    sous le signe somme, 132
Développement de Laurent, 207
Densité, 10
    de C_c dans L^1, 119
    de C_c dans L^p, 122
    gaussienne, 135

Diamètre, 14
Diffraction, 127
Dini, 16
Dispersion, 218
Distance, 8
    somme, 9
Dual topologique d'un espace de Hilbert, 150

Ensemble
    de Besicovich, 58
    de Cantor, 56
Ensembles
    Lebesgue-négligeables, 37, 47, 54
    mesurables, 93
    négligeables, 37
Espace
    L^1, 114
    L^∞, 124
    L^p, 122
    $\mathcal{L}^1$, 42
    $\mathcal{L}^p$, 121
    de Banach, 13
    dense, 10
    métrique, 8
      complet, 11
    métrique connexe, 22
Espace de Hilbert, 147
Espace métrique
    compact, 15
Espace préhilbertien, 146
Espace séparable, 162
Evanescence, 53

Famille totale, 157
Fluide, 225
Fonction
    de répartition complémentaire, 101
Fonction indicatrice
    des rationnels, 28
Fonctions
    convexes, 72
    et ensembles mesurables, 97
    mesurables, 59
    s.c.i., 35
    sommables, 42
Forme hermitienne, 146
Forme linéaire
    positive, 32
Formule
    du changement de variables, 85
Formule de la moyenne, 221
Formule de Poisson, 174

Harmonique
    fonction, 223
Hermitien
    opérateur, 177
    produit, 146
Hermitienne
    forme, 146
Holomorphe
    fonction, 196
Homotopie, 209

Inégalité
    de Hölder, 75
    de Jensen, 72
    de Markov, 101
    de Minkowski, 76
Inégalité de Cauchy-Schwarz, 146
Inégalité de Heisenberg, 172
Inégalité triangulaire, 146
Indice d'un lacet, 208
Intégrale
    de Lebesgue, 42
    des fonctions à valeurs complexes, 45
    des fonctions continues, 31
    des fonctions mesurables positives, 64, 65
    semi-convergente, 46
    sur un ensemble mesurable, 96
Isomorphisme
    entre espaces préhilbertiens, 149
    isométrique, 149

Jacobienne
    matrice, 197

Kramers et Kronig
    formule de, 218

Lemme
    de Fatou, 51
Logarithme complexe, 223

Méromorphe
    fonction, 207
Masse de Dirac, 107
Mesure
    de comptage, 107

de Lebesgue, 98
de Radon, 106

Nombre de Lebesgue, 16
Non-compacité de la boule unité, 164
Norme
L^1, 114
L^p, 122

Ondelettes de Haar, 159
Opérateur
à noyau, 180
compact, 177
hermitien, 177
Opérations sur les fonctions définies p.p., 44
Orthogonal d'un sous-espace, 150
Orthonormalisation de Schmidt, 156

Paradoxe de Banach-Tarski, 91
Parties positives/négatives, 33
Poisson
formule, 174
Préhilbertien, 146
Presque partout (p.p.), 38
Principe d'incertitude de Heisenberg, 143
Principe de Cavalieri, 103
Principe du maximum, 221
Produit hermitien, 146
Prolongement analytique, 221

Résidu, 207
Résidus
formule des, 206

Sesquilinéaire, 146
Simplement connexe, 209
Somme directe d'espaces de Hilbert, 149
Sommes de Riemann, 29
Sommes de Riemann dyadiques, 29
Spectre d'un opérateur, 177
Sturm-Liouville
équation, 181
Suite extraite, 10
Suite fondamentale, 32

Théorème
uniforme continuité, 18
d'existence des bases hilbertiennes, 162
d'inversion de Fourier dans L^1, 136
de Bolzano-Weierstrass, 18
de Borel-Lebesgue, 16
de comparaison, 64
de convergence dominée, 49
de convergence monotone, 66
de convergence monotone dans $\mathcal{L}^+$, 39
de convergence monotone dans $\mathcal{L}^1$, 49
de d'Alembert-Gauss, 220
de Dini, 16
de Féjer, 168
de Fubini, 78
de Hahn-Banach, 155
de la projection, 152
de Liouville, 220
de Müntz-Szasz, 229
de non-compacité de Riesz, 164
de Pythagore, 146
de représentation de Riesz, 153
de représentation de Riesz (version sequilinéaire), 154
de Riemann-Lebesgue, 130
de Riesz-Fischer, 156
de Stone-Weierstrass, 168
de Sturm-Liouville, 182
de Tonelli, 82
spectral, 178
Transformée de Fourier
dans L^2, 171
normalisée, 171
Transformation
de Fourier, 130
de Fourier et dérivation, 132
Triangle de Sierpiński, 57
Tribu, 93

Valeur propre d'un opérateur, 177

Zéros isolés, 220

Bibliographie

[1] Lars V. Ahlfors. *Complex analysis : An introduction of the theory of analytic functions of one complex variable*. Second edition. McGraw-Hill Book Co., New York, 1966.

[2] Luigi Ambrosio and Paolo Tilli. *Topics on Analysis in Metric Spaces*. Oxford University Press, Oxford, New York, 2004.

[3] Jean-Louis Basdevant et Jean Dalibard. *Cours de Mécanique Quantique*. Éditions de l'Ecole polytechnique, Palaiseau, 2002.

[4] Jean-Pierre Bourguignon. *Calcul variationnel*. École polytechnique, Palaiseau, 2007.

[5] Haim Brezis. *Analisi Funzionale. Teoria e applicazioni. Con un'appendice su Integrazione Astratta di Carlo Sbordone*. Liguori Editore, Naples, 1986.

[6] Henri Cartan. *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes*. Avec le concours de Reiji Takahashi. Enseignement des Sciences. Hermann, Paris, 1961.

[7] Claude Deschamps et J. Warusfel, André. *Mathématiques Tout-en-un MPSI, PCSI*. Dunod, Paris, 2010.

[8] Harry Dym and Henry P. McKean. *Fourier series and integrals*. Academic Press, 1972.

[9] Marvin Greenberg and John Harper. *Algebraic Topology :A First Course*. Addison-Wesley, New-York, 1981.

[10] Robert C. Gunning. *Introduction to holomorphic functions of several variables (Function theory I Local theory II Homological theory III)*. Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA, 1990.

[11] Allen Hatcher. *Algebraic Topology*. Cambridge Univ. Press, New-York, 2001.

[12] Yitzhak Katznelson. *Harmonic Analysis*. Dover, Palaiseau, 1976.

[13] Steven G. Krantz. *Complex analysis : the geometric viewpoint*, volume 23 of *Carus Mathematical Monographs*. Mathematical Association of America, Washington, DC, second edition, 2004.

[14] L.D. Landau and E.M. Lifschitz. *Electrodynamics of Continuous media* in *Course of Theoretical Physics*, volume 8. Pergamon Press, Oxford, New-York, Beijing, Frankfurt, 1984.

[15] Elliot Lieb Lieb and Michael Loss. *Analysis*. American Math. Soc., Providence RI, 2001.

[16] Paul Malliavin et Hélène Airault. *Intégration, analyse de Fourier, probabilités, analyse gaussienne*. Masson, Paris, Milan, Barcelone, 1994.

[17] Raghavan Narasimhan. *Several complex variables*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1995. Reprint of the 1971 original.

[18] Jean-Michel Raymond. *Electromagnétisme et relativité*. Cours de magistère ÉNS, P6, P7. http ://www.phys.ens.fr/cours/notes-de-cours/jmr/electromagnetisme.htm.

[19] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.

[20] Michel Willem. Cassini, Paris, 2003.

Table des matières

I Introduction **3**

II Compléments de topologie **7**

1 Introduction 7

2 Topologie des espaces métriques 8

2.1 Notion d'espace métrique 8

2.2 Topologie 9

2.3 Densité 10

3 Complétude 11

3.1 Espaces complets 11

3.2 Le théorème des fermés emboîtés 14

4 Compacité 15

4.1 Définitions 15

4.2 Théorème de Dini 16

4.3 Caractérisation séquentielle 16

4.4 Normes de $\mathbf{R}^N$ et leurs compacts. 18

4.5 Limites inf, sup 19

4.6 Applications de la compacité à la théorie de la mesure 20

5 Connexité 22

6 Appendice : Quelques mots sur la dénombrabilité 23

6.1 Sorites 23

6.2 Topologie des réels et dénombrabilité 23
6.3 Liste des exercices de topologie 25

III Construction de l'intégrale **27**

1 Motivations 27
2 Rappels sur l'intégration des fonctions continues 29
3 Définition de l'intégrale de Lebesgue 32
3.1 Etape 1 : classe $\mathcal{L}^+(\Omega)$ 32
3.2 Ensembles Lebesgue-négligeables 37
3.3 Etape 2 : fonctions sommables 42
3.4 Fonctions sommables et ensembles Lebesgue-négligeables. 47
4 Le théorème de convergence dominée 49
5 Parties négligeables et géométrie 54
6 Fonctions mesurables 59
7 Intégration et fonctions mesurables 64
8 Ce qu'il faut retenir 67

IV Propriétés de l'intégrale **71**

1 Inégalités intégrales 71
1.1 Inégalité de Jensen 72
1.2 Inégalité de Hölder 75
1.3 Inégalité de Minkowski 76
2 Intégrales multiples 77
3 Changement de variables 84
4 Ce qu'il faut retenir 87

V Mesure **91**

1 Introduction 91
2 Ensembles mesurables 93

3 Mesure de Lebesgue 98

4 Fonction de répartition complémentaire 101

5 Généralisations 105

5.1 Mesures de Radon 105

5.2 Mesure et intégration “abstraite” 107

6 Ce qu'il faut retenir 108

VI Espaces de Lebesgue **111**

1 Motivations 111

2 L'espace de Lebesgue L^1 112

3 Propriétés topologiques de L^1 116

3.1 Complétude de L^1 116

3.2 Densité de C_c dans L^1 119

4 Généralisations 121

5 Ce qu'il faut retenir 125

VII Transformation de Fourier sur L^1 **127**

1 Motivations 127

2 Propriétés de la transformation de Fourier 130

3 Inversion de la transformation de Fourier 135

4 Bilan et remarques finales 140

5 Ce qu'il faut retenir 143

VIII Espaces de Hilbert, théorie de base **145**

1 Définitions, premières propriétés 146

Propriétés du produit hermitien : 146

2 Sous-espaces fermés, applications continues 150

2.1 Théorèmes fondamentaux : théorème de la projection, théorèmes de représentations. 152

2.2 Théorèmes de représentation, théorème de Hahn-Banach et critère de densité 153

3 Bases hilbertiennes . . . 155

3.1 Orthonormalisation de Schmidt, familles totales, inégalité de Bessel . . . 156

Les ondelettes de Haar . . . 159

3.2 Compacité et non-compacité dans les espaces de Hilbert. Convergence faible. 164

4 Appendice : convergence des sommes de Féjer et complétude de la famille $(e^{int})_{n\in\mathbf{Z}}$ sur $L^2(\mathbf{S}^1, \mathbf{C})$ . . . 167

4.1 Ce qu'il faut retenir . . . 169

IX Opérateurs dans un espace de Hilbert 171

1 Extension d'opérateurs, transformation de Fourier L^2 et applications . . . 171

1.1 Extension d'opérateurs et transformation de Fourier dans L^2 . . . 171

1.2 Inégalité de Heisenberg. . . . 172

1.3 Formule de Poisson . . . 174

Application : Théorème d'échantillonnage de Shannon-Nyquist. . . . 176

2 Introduction à la théorie spectrale . . . 177

3 Compléments : Opérateur à noyau, Équations de Sturm-Liouville . . . 180

3.1 Problème I : Opérateurs à noyau. . . . 180

3.2 Équation de Sturm-Liouville. . . . 181

3.3 Remarque finale : les espaces de Banach . . . 185

3.4 Ce qu'il faut retenir . . . 186

4 Exercices . . . 186

X Fonctions holomorphes 193

1 Introduction . . . 193

2 Notion d'holomorphie-Premières propriétés . . . 194

2.1 Notations et rappels sur les fonctions de deux variables . . . 194

2.2 Les opérateurs $\frac{\partial}{\partial z}, \frac{\partial}{\partial \bar{z}}$ . . . 195

2.3 Holomorphie : définition et stabilité . . . 196

2.4 Les conditions de Cauchy réelles . . . 197

2.5 Transformations conformes 198

3 Formule de Cauchy pour un cercle 200

3.1 Énoncé et preuve 200

3.2 Holomorphie et analycité 202

4 Séries de fonctions holomorphes 204

5 La formule des résidus 205

5.1 Rappel sur la circulation des 1-formes différentielles 205

5.2 Énoncé de la formule 206

5.3 Fonctions méromorphes, résidu 207

5.4 Indice d'un lacet 208

5.5 Simple connexité 209

5.6 Invariance par homotopie de l'intégrale sur un lacet d'une fonction holomorphe 211

5.7 Preuve de la formule des résidus 214

5.8 Principe d'utilisation de la formule des résidus pour le calcul d'intégrale . 214

5.9 Exemple d'application à des calculs d'intégrales 215

5.10 Intermède culturel : Bulles de savon et méromorphie 216

5.11 Application à la dispersion 218

6 Propriétés ≪ globales ≫ des fonctions holomorphes 220

6.1 Zéros de fonctions holomorphes 220

6.2 Propriété de la moyenne et principe du maximum 221

6.3 Existence globale de primitives 222

6.4 Application au logarithme 223

7 Fonctions harmoniques 223

8 Application physiques -idéalisées- 224

8.1 Écoulement d'un fluide 225

8.2 Propagation de la chaleur 226

9 Applications mathématiques 227

9.1 Le théorème de Müntz-Szasz 229
9.2 Application de la notion d'indice au théorème de séparation de Jordan . . 230
10 Récapitulatifs 233
10.1 Liste des exercices de fonctions holomorphes 233
10.2 Ce qu'il faut retenir 233