

ANALYSE RÉELLE

François Golse, Yves Laszlo, Frank Pacard
et
Claude Viterbo

Avertissement

Dans ce cours, nous aborderons des sujets mathématiques importants et assez intimement liés : la théorie de la mesure et de l'intégration de Lebesgue ; la transformation de Fourier ; la théorie des espaces de Hilbert et une introduction aux méthodes variationnelles. Ce cours vise à fournir aux élèves un socle de compétences solide en analyse fonctionnelle, ce qui leur ouvrira l'accès à plusieurs domaines scientifiques : mathématiques fondamentales, mathématiques appliquées, mécanique, physique théorique, . . . L'objectif est de donner les bases des théories qui permettront, d'un point de vue pratique, d'aborder avec les armes nécessaires les cours de deuxième et de troisième année dispensés à l'École polytechnique, mais aussi de suivre avec profit ceux de première année.

Guidés par la multiplicité des points de vue, nous avons essayé de montrer la puissance et, espérons-le, l'harmonie des constructions. Nous donnerons de nombreux exemples d'applications des résultats démontrés dans ce cours, notamment des exemples issus d'autres disciplines scientifiques.

La première partie de ce cours permettra de mettre en perspective la théorie de la mesure et de l'intégration de H. Lebesgue et la théorie des probabilités développée par A. Kolmogorov. Nous verrons que ces deux branches des mathématiques sont en fait deux facettes d'une même théorie qui sert également de fondement à la théorie de la mesure géométrique développée par H. Federer. Parmi les applications spectaculaires de la théorie de la mesure, nous verrons le théorème de récurrence de Poincaré qui est un des résultats fondamentaux de la théorie des systèmes dynamiques.

Nous illustrerons les résultats concernant la transformation de Fourier par l'étude du problème de la diffraction d'une onde ou encore par l'étude des problèmes d'échantillonnage en théorie du signal (théorème de Nyquist-Shannon).

La deuxième partie de ce cours permettra de mettre en évidence les liens profonds qui unissent la théorie des espaces de Hilbert, l'analyse de Fourier et la mécanique quantique. Les résultats démontrés dans cette deuxième partie du cours seront illustrés par plusieurs exemples issus de la physique. Citons l'analyse du spectre de l'opérateur de Sturm-Liouville, l'étude des orbitales atomiques de l'atome d'hydrogène ou bien l'inégalité de Heisenberg qui figurent parmi les résultats les plus importants de la mécanique quantique.

Les espaces de fonctions construits à partir des espaces de Lebesgue et de Sobolev, qui seront définis dans ce cours, sont omniprésents en analyse, notamment dans la résolution de problèmes variationnels et celle des équations aux dérivées partielles. À titre d'illustration de ces théories nous présenterons une application à la recherche de minimum pour des fonctionnelles définies sur des espaces de fonctions. Les techniques que nous développerons trouvent de nombreuses applications en mécanique (principe de moindre action en mécanique classique et principe du minimum en mécanique des milieux continus), et nous étudierons également l'existence de géodésiques sur des surfaces.

Nous avons cherché à éviter la technicité autant que faire se peut, ce qui nous a conduits à ne pas chercher les énoncés optimaux, encore moins l'exhaustivité. Les lecteurs intéressés pour aller plus loin pourront se reporter à la bibliographie. Ce cours ne fait appel à aucunes connaissances mathématiques particulières autres que celles figurant aux programmes des classes préparatoires. Un amphithéâtre sera néanmoins consacré aux compléments de topologie

(topologie des espaces vectoriels normés et des espaces métriques) qui sont des outils indispensables pour l'étude de la théorie de l'intégration au sens de Lebesgue et de l'étude des espaces de Hilbert.

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.

Table des matières

Chapitre 1. Topologie des espaces métriques	9
1. Notion d'espace métrique	9
2. Espaces vectoriels normés	10
3. Topologie des espaces métriques	15
4. Intérieur, adhérence et densité	18
5. Valeurs d'adhérence	22
6. Continuité des applications	23
7. Connexité	28
Chapitre 2. Quelques mots sur la dénombrabilité	31
1. Définitions et propriétés	31
2. Topologie des réels et dénombrabilité	33
Chapitre 3. Compacité et complétude	35
1. La compacité	35
2. Espaces vectoriels de dimension finie	39
3. Quelques applications	41
4. Espaces métriques complets et espaces de Banach	43
5. Théorème de Riesz	46
6. Espaces d'applications continues	48
7. Topologie de la droite achevée, limite inf et limite sup	51
Chapitre 4. Théorèmes fondamentaux de l'analyse fonctionnelle	53
1. Théorème de Baire	53
2. Théorème de Banach-Steinhaus	54
3. Théorème de point fixe de Banach	55
4. Critères de densité dans $\mathcal{C}(X; \mathbf{K})$	56
5. Théorème d'Ascoli	60
6. Introduction à la résolution des équations différentielles ordinaires	64
Chapitre 5. Construction de l'intégrale de Lebesgue	69
1. Motivation	69
2. Intégration des fonctions continues	71
3. Définition de l'intégrale de Lebesgue	74
4. Théorème de la convergence dominée	88
5. Caractérisation géométrique des ensembles négligeables	93
6. Fonctions mesurables	99
7. Intégration et fonctions mesurables	103
8. Intégration des fonctions à valeurs complexes	105
9. Intégrales paramétriques	106

Chapitre 6. Propriétés de l'intégrale de Lebesgue	109
1. Inégalités intégrales	109
2. Intégrales multiples	114
3. Changements de variables	120
Chapitre 7. Mesure des ensembles	125
1. Introduction	125
2. Ensembles mesurables	126
3. Mesure de Lebesgue	131
4. Fonction de répartition complémentaire	134
5. Mesures de Radon	138
6. Théorie abstraite de la mesure	141
7. Théorème de récurrence de Poincaré	145
8. Dérivabilité et intégration	146
Chapitre 8. Espaces de Lebesgue	155
1. Motivation	155
2. L'espace de Lebesgue L^1	156
3. Propriétés topologiques de L^1	159
4. Généralisations	164
Chapitre 9. La transformation de Fourier sur L^1	169
1. Motivation	169
2. Propriétés de la transformation de Fourier	171
3. Inversion de la transformation de Fourier	174
4. Formule de Poisson	180
5. Bilan et remarques finales	183
Chapitre 10. Espaces de Hilbert	187
1. Définitions et premières propriétés	187
2. Sous-espaces fermés, applications continues	193
3. Théorème de la projection sur un convexe fermé	196
4. Théorèmes de représentation de Riesz	198
5. Bases hilbertiennes	202
6. Espaces de Hilbert séparables	209
7. Opérateurs compacts	211
8. Compacité et convergence faible	213
9. Opérateurs à noyau	217
Chapitre 11. Spectre des opérateurs et problèmes variationnels	221
1. Prolongement d'opérateurs	221
2. Transformation de Fourier dans L^2	221
3. Inégalité de Heisenberg	223
4. Spectre d'un opérateur	226
5. L'espace de Sobolev $H^1(I; \mathbb{C})$	230
6. Spectre de l'opérateur de Sturm-Liouville	237
7. Harmoniques sphériques	241
8. Introduction à la résolution de problèmes variationnels	245
9. Géodésiques sur les graphes	247

	TABLE DES MATIÈRES	7
Bibliographie		251
Index		253

CHAPITRE 1

Topologie des espaces métriques

Nous avons adopté dans ce cours le cadre des *espaces métriques*, qui nous a semblé être un bon compromis entre les espaces topologiques les plus généraux et le cadre plus restreint des espaces vectoriels normés. Les espaces métriques ont été introduits au début du siècle dernier par Maurice Fréchet, qui dégagait dans ce contexte l'importance des notions de « complétude » et de « compacité ». Ce cadre semblera certainement 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 de ce chapitre peut être lue en remplaçant « espaces métriques » par « sous-ensembles d'un espace vectoriel normé ».

Précisons que, dans ce cours, les espaces vectoriels seront toujours des \mathbf{K} -espaces vectoriels où le corps \mathbf{K} est égal à \mathbf{R} ou \mathbf{C} .

1. Notion d'espace métrique

Commençons par définir la notion de *distance* sur un ensemble X .

Définition 1.1. Une application $d : X \times X \rightarrow \mathbf{R}^+$ est une distance sur X si les propriétés suivantes sont vérifiées :

- (i) $d(x, y) = 0$ si et seulement si $x = y$ (propriété de séparation) ;
- (ii) $\forall x, y \in X, d(x, y) = d(y, x)$ (propriété de symétrie) ;
- (iii) $\forall x, y, z \in X, d(x, z) \leq d(x, y) + d(y, z)$ (inégalité triangulaire).

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

Bien évidemment, si (X, d) est un espace métrique et si $Y \subset X$ est un sous-ensemble de X , on peut considérer la restriction de d à Y , ce qui permet de munir Y d'une structure d'espace métrique. On parlera alors de *métrique induite* sur Y .

Exemple 1.1. Sur \mathbf{R} , on dispose de la *distance usuelle* définie par

$$d(x, y) := |x - y|.$$

D'autres distances peuvent être définies sur \mathbf{R} . Par exemple, étant donnée une fonction strictement monotone $f : \mathbf{R} \rightarrow \mathbf{R}$, nous pouvons définir la distance d_f par

$$d_f(x, y) := |f(x) - f(y)|,$$

pour tous $x, y \in \mathbf{R}$. On remarquera que l'hypothèse portant sur la stricte monotonie de f est essentielle pour vérifier que d_f satisfait bien la propriété de séparation (i) dans la définition ci-dessus.

Exemple 1.2. La *distance discrète* sur un ensemble X est définie par

$$d(x, x) := 0,$$

pour tout $x \in X$ et

$$d(x, y) := 1,$$

si $x \neq y$.

Exemple 1.3. Soit (X, d) un espace métrique contenant au moins 2 points. Étant donnés deux sous-ensembles non vides $A, B \subset X$, on note

$$\delta(A, B) := \inf_{a \in A, b \in B} d(a, b).$$

On vérifie que δ n'est pas une distance sur l'ensemble des sous-ensembles non vides de X . En effet, si $x \neq y \in X$ alors $\delta(\{x, y\}, \{x\}) = 0$ alors que $\{x, y\} \neq \{x\}$ ce qui montre que la propriété de séparation (i) n'est pas vérifiée.

Si (X, d) est un espace métrique, pour tout $x \in X$ et pour tout $r > 0$, on note

$$B(x, r) := \{y \in X : d(x, y) < r\},$$

la *boule ouverte de rayon r centrée au point x* . La *boule fermée* de centre $x \in X$ et de rayon $r > 0$ est, quant à elle, définie par

$$B_f(x, r) := \{y \in X : d(x, y) \leq r\}.$$

Exemple 1.4. Sur l'ensemble $X := [0, 1[$, muni de la distance usuelle $d(x, y) := |x - y|$, on vérifie que, pour tout $x \in [0, 1[$ et pour tout $r > 0$, on a

$$B(x, r) = \begin{cases} [0, x + r[& \text{si} & x < r \leq 1 - x, \\]x - r, x + r[& \text{si} & r \leq x \leq 1 - r, \\]x - r, 1[& \text{si} & 1 - x \leq r \leq x, \end{cases}$$

et

$$B_f(x, r) = \begin{cases} [0, x + r] & \text{si} & x \leq r < 1 - x, \\ [x - r, x + r] & \text{si} & r \leq x < 1 - r, \\ [x - r, 1[& \text{si} & 1 - x \leq r \leq x. \end{cases}$$

2. Espaces vectoriels normés

Une classe importante d'espaces métriques est celle où l'ensemble X est un espace vectoriel normé ou bien un sous-ensemble d'un espace vectoriel normé. Commençons par rappeler la :

Définition 1.2. Soit E un \mathbf{K} -espace vectoriel. Une application $\mathcal{N} : E \rightarrow \mathbf{R}^+$ est une norme sur E si les propriétés suivantes sont vérifiées :

- (i) $\mathcal{N}(x) = 0$ si et seulement si $x = 0$;
- (ii) $\forall x \in E$ et $\forall \lambda \in \mathbf{K}$, $\mathcal{N}(\lambda x) = |\lambda| \mathcal{N}(x)$;
- (iii) $\forall x, y \in E$, $\mathcal{N}(x + y) \leq \mathcal{N}(x) + \mathcal{N}(y)$.

Dans le cas où (i) n'est pas vérifiée, on parlera de semi-norme sur E .

Soit (E, \mathcal{N}) un espace vectoriel normé. Si $X = E$ ou plus généralement si $X \subset E$, on peut définir

$$d(x, y) := \mathcal{N}(x - y),$$

pour tous $x, y \in X$. On vérifie facilement, en utilisant les propriétés de la norme, que d est une distance sur X . C'est la *distance associée à la norme \mathcal{N}* .

Voici quelques exemples de normes sur les espaces usuels. On invite le lecteur à vérifier à titre d'exercice, que les applications définies ci-dessous sont bien des normes sur les espaces sur lesquels elles sont définies. Commençons par quelques normes définies sur \mathbf{K}^N .

Exemple 1.5. Sur \mathbf{K}^N , on peut définir les normes

$$\|x\|_1 := \sum_{i=1}^N |x_i|, \quad \|x\|_2 := \left(\sum_{i=1}^N |x_i|^2 \right)^{1/2} \quad \text{et} \quad \|x\|_\infty := \max_{i=1, \dots, N} |x_i|,$$

si $x := (x_1, \dots, x_N)$.

On remarquera que, pour démontrer que $\|\cdot\|_2$ vérifie l'inégalité triangulaire, on utilise l'inégalité de Cauchy-Schwarz

$$\left| \sum_{i=1}^N x_i y_i \right| \leq \left(\sum_{i=1}^N |x_i|^2 \right)^{1/2} \left(\sum_{i=1}^N |y_i|^2 \right)^{1/2},$$

pour tous $x, y \in \mathbf{K}^N$.

Voici maintenant quelques exemples de normes qui sont définies sur les espaces de matrices. Ces normes, appelées *normes matricielles*, sont d'un usage courant, notamment en analyse numérique.

Exemple 1.6. Soit $N \geq 1$. Sur $M_N(\mathbf{K})$, l'espace vectoriel des matrices $N \times N$ à coefficients dans \mathbf{K} , on peut définir les normes

$$\|A\|_1 := \sum_{i,j=1}^N |a_{ij}|, \quad \|A\|_2 := \left(\sum_{i,j=1}^N |a_{ij}|^2 \right)^{1/2} \quad \text{et} \quad \|A\|_\infty := \max_{i,j=1, \dots, N} |a_{ij}|,$$

si $A := (a_{ij})_{1 \leq i, j \leq N}$.

Exemple 1.7. Soit $N \geq 1$. Étant donnée une norme \mathcal{N} sur \mathbf{K}^N , on peut définir sur $M_N(\mathbf{K})$ la *norme matricielle subordonnée à \mathcal{N}* par

$$\|A\| := \sup_{x \in \mathbf{K}^N - \{0\}} \frac{\mathcal{N}(Ax)}{\mathcal{N}(x)}.$$

On laisse le soin au lecteur de vérifier que l'on a là une norme sur $M_N(\mathbf{K})$ et que de plus on dispose des définitions équivalentes

$$\|A\| := \sup_{\substack{x \in \mathbf{K}^N \\ \mathcal{N}(x)=1}} \mathcal{N}(Ax) = \sup_{\substack{x \in \mathbf{K}^N \\ \mathcal{N}(x) \leq 1}} \mathcal{N}(Ax).$$

On vérifie également que

$$\mathcal{N}(Ax) \leq \|A\| \mathcal{N}(x),$$

pour tout $x \in \mathbf{K}^N$ et que l'on a l'inégalité

$$\|AB\| \leq \|A\| \|B\|,$$

pour toutes $A, B \in M_N(\mathbf{K})$.

Enfin, donnons quelques exemples classiques de normes qui sont définies sur des espaces de dimension infinie.

Exemple 1.8. Sur $\mathcal{C}([0, 1]; \mathbf{K})$, l'espace vectoriel des fonctions à valeurs dans \mathbf{K} qui sont continues sur $[0, 1]$, on peut définir les normes suivantes

$$\|v\|_1 := \int_0^1 |v(x)| dx, \quad \|v\|_2 := \left(\int_0^1 |v(x)|^2 dx \right)^{1/2} \quad \text{et} \quad \|v\|_\infty := \sup_{x \in [0, 1]} |v(x)|.$$

La norme $\|\cdot\|_\infty$ est appelée la *norme de la convergence uniforme*.

La démonstration du fait que les applications ci-dessus sont bien des normes sur l'espace $\mathcal{C}([0, 1]; \mathbf{K})$ ne pose aucune difficulté. On remarquera que, pour démontrer que $\|\cdot\|_1$ et $\|\cdot\|_2$ sont des normes, on utilise le résultat suivant : si l'intégrale d'une fonction positive continue est égale à 0 alors, cette fonction est nulle. On remarquera aussi que, pour démontrer que $\|\cdot\|_2$ vérifie l'inégalité triangulaire, on a besoin de l'inégalité de Cauchy-Schwarz

$$\int_0^1 |f(x)g(x)| dx \leq \left(\int_0^1 |f(x)|^2 dx \right)^{1/2} \left(\int_0^1 |g(x)|^2 dx \right)^{1/2},$$

qui est valable pour toutes $f, g \in \mathcal{C}([0, 1]; \mathbf{K})$. Nous reviendrons sur la démonstration de cette inégalité dans le chapitre consacré aux espaces de Lebesgue et dans celui consacré aux espaces de Hilbert.

Les espaces de suites sont aussi des exemples très utiles qui permettent de se familiariser avec le maniement des d'espaces de dimension infinie.

Exemple 1.9. On note $\ell^1(\mathbf{N}; \mathbf{K})$ l'ensemble des suites $\mathbf{x} := (x_n)_{n \geq 0}$ qui sont indexées par \mathbf{N} , à valeurs dans \mathbf{K} et pour lesquelles la série de terme général $|x_n|$ converge. On vérifie que $\ell^1(\mathbf{N}; \mathbf{K})$ est un espace vectoriel et que l'application définie par

$$\|\mathbf{x}\|_1 := \sum_{n \in \mathbf{N}} |x_n|,$$

est une norme sur cet espace.

On note $\ell^2(\mathbf{N}; \mathbf{K})$ l'ensemble des suites $\mathbf{x} := (x_n)_{n \geq 0}$ qui sont indexées par \mathbf{N} , à valeurs dans \mathbf{K} et pour lesquelles la série de terme général $|x_n|^2$ converge. On vérifie que $\ell^2(\mathbf{N}; \mathbf{K})$ est un espace vectoriel et que l'application définie par

$$\|\mathbf{x}\|_2 := \left(\sum_{n \in \mathbf{N}} |x_n|^2 \right)^{1/2},$$

est une norme sur $\ell^2(\mathbf{N}; \mathbf{K})$.

Enfin, l'application

$$\|\mathbf{x}\|_\infty := \sup_{n \in \mathbf{N}} |x_n|,$$

est une norme sur $\ell^\infty(\mathbf{N}; \mathbf{K})$, l'espace vectoriel des suites bornées qui sont indexées par \mathbf{N} et à valeurs dans \mathbf{K} .

Dans un même espace, l'allure des boules (ouvertes ou fermées) change considérablement en fonction de la distance considérée. Pour s'en convaincre, on invite le lecteur à dessiner la boule unité de \mathbf{R}^2 pour les distances associées aux normes suivantes

$$\|(x_1, x_2)\|_1 := |x_1| + |x_2|, \quad \|(x_1, x_2)\|_2 := \sqrt{x_1^2 + x_2^2},$$

et

$$\|(x_1, x_2)\|_\infty := \max(|x_1|, |x_2|).$$

Exemple 1.10. Pour tout $p \geq 1$, on peut définir sur \mathbf{K}^N la norme

$$\|x\|_p := \left(\sum_{i=1}^N |x_i|^p \right)^{1/p},$$

où $x := (x_1, \dots, x_N)$. Ici, le point délicat consiste à vérifier l'inégalité triangulaire qui, pour cette norme, porte le nom d'inégalité de Minkowski. Quand $p > 1$, cette inégalité est une conséquence de l'*inégalité de Hölder*

$$\sum_{i=1}^N |x_i y_i| \leq \left(\sum_{i=1}^N |x_i|^p \right)^{1/p} \left(\sum_{i=1}^N |y_i|^q \right)^{1/q},$$

où $q = \frac{p}{p-1}$. Nous renvoyons le lecteur au chapitre consacré aux espaces de Lebesgue pour une démonstration générale de cette inégalité (voir la démonstration du Théorème 6.2 page 112 et du Théorème 6.3 page 113).

On pourra étudier la forme de la boule unité de \mathbf{R}^2 pour les distances associées à la norme $\|\cdot\|_p$ quand $p \geq 1$ varie et s'intéresser à ce qui se passe quand p tend vers $+\infty$.

Exemple 1.11. Reprenons l'exemple précédent quand $N \geq 2$. Pour tout $p > 0$, on peut définir

$$\|x\|_p := \left(\sum_{i=1}^N |x_i|^p \right)^{1/p},$$

où, comme d'habitude, $x := (x_1, \dots, x_N)$. Considérons

$$x := (1, 0, 0, \dots, 0) \in \mathbf{K}^N \quad \text{et} \quad y := (0, 1, 0, \dots, 0) \in \mathbf{K}^N,$$

Pour ces deux vecteurs, on trouve

$$\|x + y\|_p = 2^{1/p} \quad \text{et} \quad \|x\|_p + \|y\|_p = 2.$$

On constate que l'inégalité $\|x + y\|_p \leq \|x\|_p + \|y\|_p$ n'est pas satisfaite si $p \in]0, 1[$. Conclusion, l'application $\|\cdot\|_p$, qui est bien définie sur \mathbf{K}^N , n'est pas une norme sur cet espace quand $p \in]0, 1[$.

Devant la multitude des normes définies sur un même espace vectoriel, nous sommes conduits à donner la :

Définition 1.3. Soit E un espace vectoriel. On dit que deux normes \mathcal{N}_1 et \mathcal{N}_2 définies sur E sont équivalentes s'il existe deux constantes $C_1, C_2 > 0$ telles que

$$\forall x \in E, \quad \mathcal{N}_1(x) \leq C_1 \mathcal{N}_2(x) \quad \text{et} \quad \mathcal{N}_2(x) \leq C_2 \mathcal{N}_1(x).$$

Géométriquement, deux normes définies sur un même espace vectoriel sont équivalentes si la boule unité ouverte centrée à l'origine pour la norme \mathcal{N}_1 contient une boule ouverte non vide centrée à l'origine pour la norme \mathcal{N}_2 et si la boule unité ouverte centrée à l'origine pour la norme \mathcal{N}_2 contient une boule ouverte non vide centrée à l'origine pour la norme \mathcal{N}_1 .

Comme nous le verrons un peu plus tard, sur un espace vectoriel de dimension finie, toutes les normes sont équivalentes. Malheureusement, ce n'est pas le cas sur les espaces de dimension infinie. Ce phénomène, est illustré dans l'exemple suivant :

Exemple 1.12. Reprenons les notations de l'exemple 1.8 page 12. On vérifie que

$$\|v\|_1 \leq \|v\|_2 \leq \|v\|_\infty,$$

pour toute $v \in \mathcal{C}([0, 1]; \mathbf{K})$ (la première inégalité est une conséquence de l'inégalité de Cauchy-Schwarz rappelée dans l'exemple 1.8 page 12). En revanche, on ne peut pas trouver de constante $C > 0$ telle que $\|v\|_2 \leq C \|v\|_1$ pour toute $v \in \mathcal{C}([0, 1]; \mathbf{K})$. Pour le voir, on construit pour tout $j \in \mathbf{N} - \{0\}$ la fonction

$$v_j(x) := (1 - jx)^+ := \max(1 - jx, 0).$$

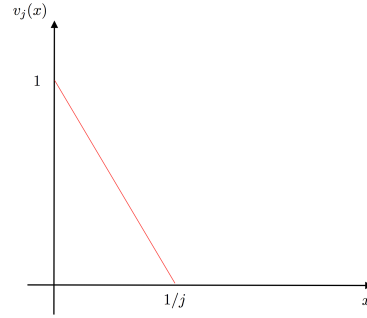


FIGURE 1.1. Graphe de la fonction v_j .

On vérifie que

$$\|v_j\|_1 = \frac{1}{2j}, \quad \|v_j\|_2 = \frac{1}{\sqrt{3j}} \quad \text{et} \quad \|v_j\|_\infty = 1.$$

S'il existait une constante $C > 0$ telle que $\|v\|_2 \leq C \|v\|_1$ pour toute $v \in \mathcal{C}([0, 1]; \mathbf{K})$, on pourrait écrire

$$\|v_j\|_2 \leq C \|v_j\|_1,$$

pour tout $j \geq 1$. Ce qui est manifestement faux quand j est assez grand.

En utilisant cette suite de fonctions, on montre qu'il n'existe pas de constante $C > 0$ telle que $\|v\|_\infty \leq C \|v\|_2$ pour toute $v \in \mathcal{C}([0, 1]; \mathbf{K})$ et on montre également qu'il n'existe pas de constante $C > 0$ telle que $\|v\|_\infty \leq C \|v\|_1$ pour toute $v \in \mathcal{C}([0, 1]; \mathbf{K})$. Conclusion, parmi ces trois normes, il n'y en a pas deux qui sont équivalentes.

On retiendra que, suivant le choix de la norme (ou plus généralement de la distance), deux éléments d'un même espace (ou plus généralement d'un même ensemble) peuvent paraître plus ou moins éloignés.

3. Topologie des espaces métriques

Nous donnons dans cette section les définitions usuelles utilisées en topologie.

Définition 1.4. Soit (X, d) un espace métrique. Un sous-ensemble U de X est un ouvert de (X, d) si, pour tout $x \in U$, il existe un rayon $r > 0$ tel que $B(x, r)$, la boule ouverte centrée en x et de rayon $r > 0$, est incluse dans U . On appelle topologie associée à la métrique d et l'on note \mathcal{T}_d , l'ensemble des ouverts de (X, d) .

On vérifie que \emptyset, X sont toujours des ouverts de (X, d) .

Exemple 1.13. Soit $x \in X$ et $r > 0$. Alors, $B(x, r)$, la boule ouverte de centre x et de rayon $r > 0$, est un ouvert de (X, d) .

Voici deux propriétés importantes vérifiées par les ouverts d'un espace métrique.

Proposition 1.1. Une intersection finie d'ouverts de (X, d) est un ouvert de (X, d) et une réunion quelconque d'ouverts de (X, d) est un ouvert de (X, d) .

On appelle *voisinage* d'un point $x \in X$ (ou d'un ensemble $Y \subset X$) tout ensemble V qui contient un ouvert de (X, d) qui lui même contient le point x (ou l'ensemble Y). On remarquera qu'un ouvert est un voisinage de chacun de ses points (le lecteur intéressé pourra le démontrer à titre d'exercice).

Définition 1.5. Un sous-ensemble F de X est un fermé de (X, d) si son complémentaire $X - F$ est un ouvert.

Par passage aux complémentaires, on vérifie que \emptyset, X sont toujours des fermés de (X, d) .

Exemple 1.14. Soit $x \in X$ et $r > 0$. Alors, $B_f(x, r)$, la boule *fermée* de centre $x \in X$ et de rayon $r > 0$, est un fermé de (X, d) .

Par passage aux complémentaires, on a également :

Proposition 1.2. Une réunion finie de fermés de (X, d) est un fermé de (X, d) et qu'une intersection quelconque de fermés de (X, d) est un fermé de (X, d) .

Donnons quelques exemples qui permettront au lecteur de se familiariser avec ces notions.

Exemple 1.15. On considère $X = \mathbf{R}$, muni de la distance usuelle $d(x, y) = |y - x|$. Si $a < b$, on vérifie que :

- (i) les ensembles $]a, b[,]b, +\infty[$ et $] - \infty, a[$ sont des ouverts de (\mathbf{R}, d) ;
- (ii) les ensembles $[a, b], [b, +\infty[,] - \infty, a]$ et $\{a\}$ sont des fermés de (\mathbf{R}, d) ;
- (iii) les ensembles $[a, b[$ et $]a, b]$ ne sont ni des ouverts, ni des fermés de (\mathbf{R}, d) .

Exemple 1.16. On considère $X = \mathbf{R}$, muni de la distance usuelle. Clairement,

$$[0, 1] = \bigcap_{j \geq 1}] - \frac{1}{j}, 1 + \frac{1}{j}[,$$

ce qui montre qu'une intersection quelconque d'ouverts n'est pas forcément un ouvert. De même

$$]0, 1[= \bigcup_{j \geq 2} [\frac{1}{j}, 1 - \frac{1}{j}],$$

ce qui montre qu'une réunion quelconque de fermés n'est pas forcément un fermé.

Exemple 1.17. On considère toujours $X = \mathbf{R}$, muni de la distance usuelle. Les seuls sous-ensembles de \mathbf{R} qui sont à la fois ouverts et fermés dans \mathbf{R} sont \emptyset et \mathbf{R} (on verra plus tard que cette propriété traduit le fait que \mathbf{R} est *connexe*). Pour démontrer cette propriété, on considère un sous-ensemble non vide $X \subset \mathbf{R}$ qui est à la fois ouvert et fermé. Choisissons $x \in X$. On note

$$Y := \{y \geq x : [x, y] \subset X\}.$$

On vérifie facilement que Y est un intervalle de \mathbf{R} . Montrons que $Y = [x, +\infty[$. Supposons le contraire et notons

$$y_* := \sup_{y \in Y} y.$$

Supposons que $y_* \in X$. L'ensemble X étant ouvert, on peut trouver $r > 0$ tel que $]y_* - 2r, y_* + 2r[\subset X$. Par définition de y_* , on a $Y \cap [x, y_*[\neq \emptyset$ et par conséquent on conclut que $[x, y_* + r] \subset X$ ce qui contredit la définition de y_* . Supposons maintenant que $y_* \in \mathbf{R} - X$. L'ensemble X étant fermé, son complémentaire est ouvert, donc il existe $r > 0$ tel que $]y_* - r, y_* + r[\subset \mathbf{R} - X$ et par conséquent $Y \subset [x, y_* - r]$, ce qui contredit la définition de y_* .

On montre de même que $] - \infty, x] \subset X$ ce qui termine la démonstration.

Le lecteur intéressé pourra montrer, en utilisant des arguments semblables à ceux utilisés ci-dessus, que dans l'espace \mathbf{R}^N muni par exemple de la distance associée à la norme euclidienne, les seuls sous-ensembles qui sont à la fois ouverts et fermés sont \emptyset et \mathbf{R}^N .

Donnons maintenant quelques exemples un peu moins intuitifs.

Exemple 1.18. On considère sur \mathbf{Z} la distance $d(x, y) := |x - y|$ induite par la distance usuelle sur \mathbf{R} . Pour tout $x \in \mathbf{Z}$, l'ensemble $\{x\}$ est à la fois un ouvert et un fermé dans cet espace métrique. Plus généralement tout sous-ensemble de \mathbf{Z} est à la fois ouvert et fermé pour la topologie associée à cette distance.

L'exemple suivant montre que la notion d'ouvert dépend fortement de l'espace métrique que l'on considère et qu'il faut faire attention aux conclusions hâtives qui pourraient être déduites des notations.

Exemple 1.19. Comme dans l'exemple 1.4 page 10, on considère l'ensemble $[0, 1[$ muni de la distance usuelle $d(x, y) = |y - x|$. Soient $a, b \in]0, 1[$ tels que $a < b$, on vérifie que :

- (i) les ensembles $[0, a[$, $]a, b[$ et $[0, 1[$ sont des ouverts de $([0, 1[, d)$;
- (ii) les ensembles $[0, a]$, $[a, b]$, $[0, 1[$ et $\{a\}$ sont des fermés de $([0, 1[, d)$;
- (iii) l'ensemble $[a, b]$ n'est ni un ouvert, ni un fermé de $([0, 1[, d)$.

On vérifie que les seuls sous-ensembles de $([0, 1[, d)$ qui sont à la fois ouverts et fermés sont \emptyset et $[0, 1[$.

Topologie induite. Soit $Y \subset X$ un sous-ensemble d'un espace métrique (X, d) . Vérifions que les ouverts de (Y, d) (c'est-à-dire les ouverts de Y pour la distance induite) sont les *traces* $U \cap Y$ des ouverts U de (X, d) . Autrement dit, $V \subset Y$ est un ouvert de (Y, d) si et seulement s'il existe un ouvert $U \subset X$ de (X, d) tel que

$$V = U \cap Y.$$

La topologie correspondante sur Y s'appelle la *topologie induite* (sous-entendu, par la topologie de (X, d)). Passons maintenant à la démonstration de cette propriété. Pour tout $x \in X$, on note $B_X(x, r)$, la boule ouverte de X , de centre x et de rayon $r > 0$ et, pour tout $y \in Y$, on

note $B_Y(y, r)$, la boule ouverte de Y , de centre y et de rayon $r > 0$. Soit $V \subset Y$ un ouvert de (Y, d) . Pour tout $y \in V$, il existe $r_y > 0$ tel que $B_Y(y, r_y) \subset V$. Autrement dit

$$B_Y(y, r_y) := \{z \in Y : d(y, z) < r_y\} \subset V.$$

Alors

$$U := \bigcup_{y \in V} B_X(y, r_y),$$

est un ouvert de X et $V = U \cap Y$. Inversement, si U est un ouvert de X et si $x \in U \cap Y$, alors il existe $r > 0$ tel que $B_X(x, r) \subset U$. Dans ce cas,

$$B_Y(x, r) := \{z \in Y : d(x, z) < r\} \subset Y \cap U,$$

ce qui montre que $U \cap Y$ est un ouvert de Y .

Remarquons qu'en général, si $V \subset Y$ est un ouvert de (Y, d) alors V n'est pas forcément un ouvert de (X, d) . En reprenant l'exemple 1.19, on voit que $[0, 1/2[$ est un ouvert de $([0, 1], d)$ mais ce n'est pas un ouvert de (\mathbf{R}, d) . En revanche, dans le cas particulier où $Y \subset X$ est lui-même un ouvert de (X, d) , on vérifie que $V \subset Y$ est un ouvert de (Y, d) si et seulement si V est un ouvert de (X, d) . Nous laissons au lecteur le soin d'énoncer et de prouver une propriété analogue pour les fermés.

Topologies associées à des normes équivalentes. Les distances associées à deux normes équivalentes, définies sur un même espace vectoriel normé E , définissent la même topologie sur tout sous-ensemble non vide $X \subset E$. Autrement dit, si d_j est la distance associée à la norme \mathcal{N}_j définie sur un espace vectoriel E et si les normes \mathcal{N}_1 et \mathcal{N}_2 sont équivalentes, alors, pour tout ensemble non vide $X \subset E$, tout ouvert (fermé) de (X, d_1) est un ouvert (fermé) de (X, d_2) et réciproquement.

Plus généralement, on dit que deux distances d_1 et d_2 , définies sur un ensemble X , sont Lipschitz-équivalentes s'il existe deux constantes $C_1, C_2 > 0$ telle que

$$d_1(x, y) \leq C_1 d_2(x, y) \quad \text{et} \quad d_2(x, y) \leq C_2 d_1(x, y),$$

pour tous $x, y \in X$. Deux distances Lipschitz-équivalentes sur X définissent sur X la même topologie.

Terminons par la définition de distances sur le produit de deux espaces métriques (cette définition s'étend immédiatement au cas d'un produit d'un nombre fini d'espaces métriques).

Exemple 1.20. Si (X, d) et (X', d') sont deux espaces métriques, on peut munir l'espace produit $X \times X'$ de la *distance somme* qui est définie par

$$d_s((x_1, x'_1), (x_2, x'_2)) := d(x_1, x_2) + d'(x'_1, x'_2),$$

ou bien de la *distance produit* qui est définie par

$$d_p((x_1, x'_1), (x_2, x'_2)) := \max(d(x_1, x_2), d'(x'_1, x'_2)).$$

On vérifie que l'on a là deux distances qui permettent de munir $X \times X'$ d'une structure d'espace métrique. On vérifie en outre que

$$d_p \leq d_s \leq 2d_p,$$

ce qui montre que ces deux distances sont Lipschitz-équivalentes, en particulier, les topologies \mathcal{T}_{d_s} et \mathcal{T}_{d_p} associées aux distances d_s et d_p sont les mêmes (i.e. qu'elles ont les mêmes ouverts, les mêmes fermés). On montre également que $O \times O'$, produit d'un ouvert O de (X, d) et d'un ouvert O' de (X', d') , est un ouvert de $(X \times X', d_p)$.

4. Intérieur, adhérence et densité

Deux autres notions de topologie seront fréquemment utilisées dans ce cours.

Définition 1.6. Soit (X, d) un espace métrique et $Y \subset X$. L'intérieur de Y , qui sera noté $\overset{\circ}{Y}$, est le plus grand ouvert contenu dans Y , c'est aussi la réunion de tous les ouverts inclus dans Y

$$\overset{\circ}{Y} := \bigcup_{\substack{U \text{ ouvert} \\ U \subset Y}} U.$$

L'adhérence de $Y \subset X$, qui sera notée \overline{Y} , est le plus petit fermé contenant Y , c'est aussi l'intersection de tous les fermés qui contiennent Y

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

Exemple 1.21. Dans \mathbf{R} , muni de la distance usuelle, l'adhérence de $]0, 1]$ est $[0, 1]$ et l'intérieur de $[0, 1]$ est $]0, 1[$.

Exemple 1.22. Dans $[0, 1[$, muni de la distance usuelle, l'adhérence de $]0, 1[$ est $[0, 1[$ et l'intérieur de $[0, 1[$ est $]0, 1[$.

Exemple 1.23. Sur $X = \mathbf{Z}$ muni de la distance usuelle $d(x, y) = |x - y|$, la boule ouverte de rayon 1 centrée en 0 est réduite à $\{0\}$

$$B(0, 1) = \{0\}.$$

Étant donné que $\{0\}$ est un fermé de (\mathbf{Z}, d) , l'adhérence de $B(0, 1)$ est aussi réduite à $\{0\}$

$$\overline{B(0, 1)} = \{0\},$$

alors que, pour cette même distance, la boule fermée de centre 0 et de rayon 1 est l'ensemble $\{-1, 0, 1\}$

$$B_f(0, 1) = \{-1, 0, 1\}.$$

On vérifie aussi que l'intérieur de $B_f(0, 1)$, la boule fermée de centre 0 et de rayon 1, est égal à $\{-1, 0, 1\}$

$$\overbrace{B_f(0, 1)}^{\circ} = \{-1, 0, 1\}.$$

Remarque 1.1. Attention, on prendra garde de ne pas confondre $B_f(x, r)$, qui est la boule fermée de centre x et de rayon $r > 0$, et $\overline{B(x, r)}$, qui est l'adhérence de la boule ouverte de centre x et de rayon $r > 0$. Par définition,

$$\overline{B(x, r)} \subset B_f(x, r),$$

car $B(x, r) \subset B_f(x, r)$, mais on n'a pas forcément égalité comme le montre l'exemple ci-dessus. On retiendra que, dans un espace métrique, l'adhérence de la boule ouverte de rayon r n'est pas nécessairement égale à la boule fermée de rayon r .

Remarque 1.2. En revanche, dans un espace vectoriel normé (E, \mathcal{N}) , on a bien $\overline{B(x, r)} = B_f(x, r)$. Pour démontrer cette propriété, on utilise la remarque ci-dessus qui permet d'affirmer que $\overline{B(x, r)} \subset B_f(x, r)$. Maintenant, $E - \overline{B(x, r)}$ est un ouvert et par conséquent pour tout $y \in E - \overline{B(x, r)}$, il existe $\rho > 0$ tel que $B(y, \rho) \subset E - \overline{B(x, r)}$, autrement dit, $B(x, r) \cap B(y, \rho) = \emptyset$. On note

$$z := y + \frac{\rho}{2} \frac{x - y}{\mathcal{N}(x - y)}.$$

Clairement $z \in B(y, \rho)$ et par conséquent, $z \notin B(x, r)$. En particulier

$$r \leq \mathcal{N}(z - x) = \left(1 - \frac{\rho}{2} \frac{1}{\mathcal{N}(x - y)}\right) \mathcal{N}(y - x).$$

Finalement $\mathcal{N}(y - x) > r$, ce qui montre que $y \notin B_f(x, r)$. Conclusion, $E - \overline{B(x, r)} \subset E - B_f(x, r)$ ou autrement dit $B_f(x, r) \subset \overline{B(x, r)}$.

Donnons quelques exemples supplémentaires, pour nous permettre de nous familiariser avec ces notions d'intérieur et d'adhérence.

Exemple 1.24. On considère \mathbf{R} muni de la distance usuelle. On vérifie que $\overline{\mathbf{Q}} = \mathbf{R}$ et que $\overset{\circ}{\mathbf{Q}} = \emptyset$. De même, $\overline{\mathbf{R} - \mathbf{Q}} = \mathbf{R}$ et $\overset{\circ}{\mathbf{R} - \mathbf{Q}} = \emptyset$.

Exemple 1.25. Nous reprenons les notations de l'exemple 1.9 page 12. On note $\ell_c(\mathbf{N}; \mathbf{K})$ le sous-espace de $\ell^\infty(\mathbf{N}; \mathbf{K})$ constitué des suites identiquement nulles à partir d'un certain rang. Montrons que

$$\overset{\circ}{\ell_c(\mathbf{N}; \mathbf{K})} = \emptyset,$$

dans $(\ell^\infty(\mathbf{N}, \mathbf{K}), \| \cdot \|_\infty)$. Pour démontrer ce résultat, on raisonne par l'absurde et l'on suppose

qu'il existe un élément \mathbf{x} dans $\overset{\circ}{\ell_c(\mathbf{N}; \mathbf{K})}$, qui est par définition un ouvert $\ell^\infty(\mathbf{N}, \mathbf{K})$. En particulier, il existe un réel $r > 0$ tel que la boule ouverte de $\ell^\infty(\mathbf{N}; \mathbf{K})$ centrée en \mathbf{x} et de rayon r est incluse dans $\overset{\circ}{\ell_c(\mathbf{N}; \mathbf{K})}$. Notons $\mathbf{1} \in \ell^\infty(\mathbf{N}, \mathbf{K})$, la suite constante égale à 1. D'une part

$$\mathbf{x} + s \mathbf{1} \in B(\mathbf{x}, r) \subset \overset{\circ}{\ell_c(\mathbf{N}; \mathbf{K})} \subset \ell_c(\mathbf{N}; \mathbf{K}),$$

pour tout $s \in]0, r[$, et d'autre part $(\mathbf{x} + s' \mathbf{1}) - (\mathbf{x} + s \mathbf{1}) = (s' - s) \mathbf{1}$ n'est pas une suite nulle à partir d'un certain rang si $s \neq s'$, ce qui constitue une contradiction.

Support d'une fonction. Le *support d'une fonction* définie sur un espace métrique, à valeurs réelles ou complexes, est, par définition, l'adhérence des points où cette fonction ne s'annule pas. On remarquera que, par définition, le support d'une fonction est un fermé.

Définition 1.7. On dit qu'une suite $(x_n)_{n \geq 0}$ d'un espace métrique (X, d) est convergente et qu'elle a pour limite x , si $\lim_{n \rightarrow +\infty} d(x_n, x) = 0$.

Les espaces métriques sont des espaces topologiques *séparés* (on parle aussi d'espace de Hausdorff) ce qui se traduit par le fait qu'étant donnés deux points distincts x_1 et x_2 d'un espace métrique (X, d) , il existe deux ouverts U_1 et U_2 disjoints, tels que $x_i \in U_i$ (on peut par exemple prendre les deux boules ouvertes centrées en x_1 et x_2 , de rayon $r > 0$ choisi de telle sorte que $2r < d(x_1, x_2)$). Une conséquence du caractère séparé des espaces métriques est que

la limite d'une suite convergente est unique : si une suite $(x_n)_{n \geq 0}$ converge vers x et aussi vers \bar{x} alors, grâce à l'inégalité triangulaire, on obtient

$$d(x, \bar{x}) \leq d(x, x_n) + d(x_n, \bar{x}),$$

pour tout $n \in \mathbf{N}$; en passant à la limite quand n tend vers l'infini, on conclut que $d(x, \bar{x}) = 0$ et finalement (propriété de séparation de la distance) que $x = \bar{x}$.

Dans un espace métrique la propriété d'être fermé peut être vérifiée en étudiant les limites de suites. C'est l'objet de la proposition suivante.

Proposition 1.3. *Un sous-ensemble $F \subset X$ est fermé dans (X, d) si et seulement si la limite de toute suite d'éléments de F qui converge dans X , appartient à F . Autrement dit, F est un fermé si et seulement si, toute suite $(x_n)_{n \geq 0}$ d'éléments de F qui converge dans X , converge vers un élément de F .*

DÉMONSTRATION. Supposons que F est un fermé et soit $(x_n)_{n \geq 0}$ une suite d'éléments de F qui converge vers $x \in X$. Montrons que $x \in F$. Pour ce faire, raisonnons par l'absurde et supposons que $x \notin F$, c'est à dire que $x \in X - F$. L'ensemble $X - F$ est un ouvert, par conséquent, il existe $r > 0$ tel que $B(x, r) \subset X - F$. La suite $(x_n)_{n \geq 0}$ convergeant vers x , on a, pour tout n assez grand (par exemple pour $n \geq m$) $d(x, x_n) < r$, donc $x_n \in B(x, r) \subset X - F$, pour tout $n \geq m$, ce qui contredit le fait que $x_n \in F$. Conclusion, $x \in F$.

Inversement, supposons que la limite de toute suite d'éléments de F qui converge dans X , appartient à F et montrons que F est un fermé. Une fois de plus raisonnons par l'absurde et supposons que F n'est pas un fermé. Alors $X - F$ n'est pas un ouvert ce qui se traduit par le fait qu'il existe $x \in X - F$ tel que, pour tout $r > 0$, $B(x, r)$ n'est pas incluse dans $X - F$. Autrement dit, pour tout $r > 0$, $B(x, r) \cap F \neq \emptyset$. Pour tout $n \geq 1$, choisissons $x_n \in F \cap B(x, 1/n)$. La suite $(x_n)_{n \geq 1}$ est une suite d'éléments de F qui converge vers x et $x \notin F$. Ce qui constitue une contradiction. \square

On peut par exemple utiliser cette caractérisation des fermés pour montrer que \mathbf{Q} n'est pas un fermé de \mathbf{R} (muni de la distance usuelle). En effet, on peut construire une suite de nombres rationnels (par exemple les approximations décimales par défaut de $\sqrt{2}$) qui converge vers $\sqrt{2}$ et $\sqrt{2} \notin \mathbf{Q}$, ce qui montre que \mathbf{Q} n'est pas un fermé de \mathbf{R} .

De la même façon, on dispose de caractérisations particulièrement pratiques de l'adhérence d'un sous-ensemble d'un espace métrique.

Proposition 1.4. *L'adhérence dans (X, d) d'un ensemble $Y \subset X$ est égale à :*

- (i) *l'ensemble des limites de suites d'éléments de Y ;*
- (ii) *l'ensemble des $x \in X$ tels que, $Y \cap B(x, \varepsilon) \neq \emptyset$, pour tout $\varepsilon > 0$.*

DÉMONSTRATION. On note \bar{Y} l'adhérence de Y et \tilde{Y} l'ensemble des limites de suites d'éléments de Y . Il est clair que $Y \subset \tilde{Y}$ (considérer pour tout $y \in Y$, la suite constante dont tous les éléments sont égaux à y , suite qui converge manifestement vers y). Soit $\tilde{y} \in \tilde{Y}$ et F un fermé qui contient Y . Par définition de \tilde{Y} , il existe une suite $(y_n)_{n \geq 0}$ d'éléments de Y qui converge vers \tilde{y} . Dans ce cas, le résultat de la Proposition précédente nous assure que $\tilde{y} \in F$. On conclut que $\tilde{Y} \subset F$ et par conséquent que $\tilde{Y} \subset \bar{Y}$. Nous avons donc démontré les inclusions

$$Y \subset \tilde{Y} \subset \bar{Y}.$$

Montrons que \tilde{Y} est un fermé, ce qui montrera que $\bar{Y} \subset \tilde{Y}$ et qui terminera la démonstration du fait que $\tilde{Y} = \bar{Y}$. Pour ce faire, donnons nous $(\tilde{y}_n)_{n \geq 1}$, une suite d'éléments de \tilde{Y} qui

converge vers $\tilde{y} \in X$. Pour tout $n \geq 1$, $\tilde{y}_n \in \tilde{Y}$ est la limite d'une suite d'éléments de Y , en particulier, il existe $y_n \in Y$ tel que $d(y_n, \tilde{y}_n) \leq 1/n$. On vérifie que la suite $(y_n)_{n \geq 1}$, qui est une suite d'éléments de Y , converge vers \tilde{y} , ce qui, par définition de \tilde{Y} , prouve que $\tilde{y} \in \tilde{Y}$. Le résultat de la Proposition précédente nous assure alors que \tilde{Y} est un fermé.

On note \hat{Y} l'ensemble des $x \in X$ tels que, $B(x, \varepsilon) \cap Y \neq \emptyset$, pour tout $\varepsilon > 0$. Soit $x \in \hat{Y}$. Pour tout $n \geq 1$, choisissons $x_n \in Y \cap B(x, 1/n)$. La suite $(x_n)_{n \geq 1}$ est une suite d'éléments de Y qui converge vers x . D'après le point (i), $x \in \bar{Y}$. Donc $\hat{Y} \subset \bar{Y}$. Inversement, grâce à (i), on sait que tout élément $x \in \bar{Y}$ est la limite d'une suite d'éléments de Y . En particulier, pour tout $\varepsilon > 0$, $B(x, \varepsilon)$ contient des éléments de cette suite, donc des éléments de Y . Conclusion $x \in \hat{Y}$. Ce qui termine la démonstration du point (ii). \square

Ce résultat permet de traiter rapidement l'exemple qui suit.

Exemple 1.26. Nous reprenons les notations de l'exemple 1.25 page 19. On note $\ell_0(\mathbf{N}; \mathbf{K})$ le sous-espace de $\ell^\infty(\mathbf{N}; \mathbf{K})$ constitué des suites qui tendent vers 0. Montrons que

$$\overline{\ell_c(\mathbf{N}; \mathbf{K})} = \ell_0(\mathbf{N}; \mathbf{K}),$$

dans $(\ell^\infty(\mathbf{N}; \mathbf{K}), \|\cdot\|_\infty)$.

Soit $\mathbf{x} := (x_n)_{n \geq 0} \in \ell_0(\mathbf{N}; \mathbf{K})$. Pour tout $i \geq 0$, définissons $\mathbf{x}^i := (x_n^i)_{n \geq 0} \in \ell_c(\mathbf{N}; \mathbf{K})$ la suite dont les éléments sont donnés par $x_n^i = x_n$ si $n \leq i$ et $x_n^i = 0$ si $n > i$. Le fait que la suite \mathbf{x} tend vers 0 se traduit par le fait que

$$\lim_{i \rightarrow \infty} \|\mathbf{x} - \mathbf{x}^i\|_\infty = 0.$$

Le point (i) de la Proposition 1.4 page précédente permet de conclure que $\mathbf{x} \in \overline{\ell_c(\mathbf{N}; \mathbf{K})}$, ce qui montre que $\ell_0(\mathbf{N}; \mathbf{K}) \subset \overline{\ell_c(\mathbf{N}; \mathbf{K})}$.

Soit $\mathbf{x} := (x_n)_{n \geq 0} \in \overline{\ell_c(\mathbf{N}; \mathbf{K})}$. Grâce au point (i) de la Proposition 1.4 page ci-contre, on sait qu'il existe une suite d'éléments de $\ell_c(\mathbf{N}; \mathbf{K})$ qui converge vers \mathbf{x} . Notons $(\mathbf{x}^i)_{i \geq 0}$ cette suite (attention, chaque $\mathbf{x}^i := (x_n^i)_{n \geq 0}$ est un élément de $\ell_c(\mathbf{N}; \mathbf{K})$). Fixons $\varepsilon > 0$ et choisissons i assez grand pour que

$$\|\mathbf{x} - \mathbf{x}^i\|_\infty \leq \varepsilon.$$

Par définition, la suite \mathbf{x}^i est nulle à partir d'un certain rang. Donc, il existe $m \in \mathbf{N}$ tel que $x_n^i = 0$ pour tout $n \geq m$. En particulier $|x_n| \leq \varepsilon$ pour tout $n \geq m$. Ce qui montre que $\mathbf{x} \in \ell_0(\mathbf{N}; \mathbf{K})$ et par conséquent que $\overline{\ell_c(\mathbf{N}; \mathbf{K})} \subset \ell_0(\mathbf{N}; \mathbf{K})$.

La notion de *densité* d'un sous-ensemble dans un espace métrique est une notion très importante que nous retrouverons à plusieurs reprises dans ce cours.

Définition 1.8. On dit qu'un sous-ensemble $Y \subset X$ d'un espace métrique (X, d) est dense dans X si $\bar{Y} = X$. Autrement dit, Y est dense dans (X, d) si Y rencontre tous les ouverts non vides de (X, d) .

Grâce au résultat de la Proposition 1.4 page précédente, on dispose de caractérisations assez simple des sous-ensembles denses d'un espace métrique :

Proposition 1.5. Un sous-ensemble $Y \subset X$ est dense dans (X, d) si et seulement si

$$\forall x \in X, \quad \forall \varepsilon > 0, \quad B(x, \varepsilon) \cap Y \neq \emptyset.$$

Autrement dit, $Y \subset X$ est dense dans (X, d) si et seulement si Y rencontre toute boule ouverte de (X, d) .

On a également la :

Proposition 1.6. *Un sous-ensemble $Y \subset X$ est dense dans (X, d) si et seulement si, pour tout élément $x \in X$, il existe une suite $(y_n)_{n \geq 0}$ d'éléments de Y telle que*

$$\lim_{n \rightarrow +\infty} d(x, y_n) = 0.$$

Autrement dit, $Y \subset X$ est dense dans (X, d) si et seulement si tout élément $x \in X$ est limite d'une suite d'éléments de Y .

Nous aurons l'occasion de rencontrer dans ce cours de nombreux exemples de sous-ensembles denses dans des espaces métriques variés. Pour l'instant contentons nous de l'exemple le plus simple.

Exemple 1.27. L'ensemble des rationnels \mathbf{Q} et l'ensemble des irrationnels $\mathbf{R} - \mathbf{Q}$ sont denses dans \mathbf{R} (muni de la distance usuelle).

5. Valeurs d'adhérence

On rappelle qu'une suite $(y_n)_{n \geq 0}$ est une *suite extraite* (ou une *sous-suite*) de la suite $(x_n)_{n \geq 0}$, s'il existe une application $\varphi : \mathbf{N} \rightarrow \mathbf{N}$ qui est strictement croissante et telle que $y_n = x_{\varphi(n)}$, pour tout $n \geq 0$.

Définition 1.9. *Soit (X, d) un espace métrique et soit $(x_n)_{n \geq 0}$ une suite d'éléments de X . Un point $x \in X$ est une valeur d'adhérence de la suite $(x_n)_{n \geq 0}$ si x est limite d'une suite extraite de la suite $(x_n)_{n \geq 0}$.*

Exemple 1.28. Dans \mathbf{R} muni de la distance usuelle, la suite $((-1)^n)_{n \geq 0}$ admet exactement deux valeurs d'adhérence -1 et $+1$. La suite $(n)_{n \geq 0}$ n'admet pas de valeur d'adhérence dans \mathbf{R} (muni de la distance usuelle).

On démontre immédiatement une caractérisation très commode des valeurs d'adhérence d'une suite : si $(x_n)_{n \geq 0}$ est une suite d'éléments de X , le point $x \in X$ est une valeur d'adhérence de la suite $(x_n)_{n \geq 0}$ si et seulement si, pour tout $\varepsilon > 0$

$$\text{card} \{k \in \mathbf{N} : x_k \in B(x, \varepsilon)\} = +\infty,$$

où $\text{card } A$ désigne le cardinal de l'ensemble A .

Le lien entre valeurs d'adhérence d'une suite et adhérence d'un ensemble est précisé dans la proposition suivante qui est une conséquence immédiate du point (i) de la Proposition 1.4 page 20.

Proposition 1.7. *L'adhérence \overline{Y} d'un sous-ensemble $Y \subset X$ est égale à l'ensemble des valeurs d'adhérence dans X des suites d'éléments de Y .*

Pour se familiariser avec les notions de valeur d'adhérence et d'adhérence d'un ensemble, le lecteur pourra vérifier que l'ensemble des valeurs d'adhérence d'une suite est un fermé. Plus précisément, on a la :

Proposition 1.8. *L'ensemble des valeurs d'adhérence de la suite $(x_n)_{n \geq 0}$ est donné par*

$$\bigcap_{n \geq 0} \overline{\{x_k : k \geq n\}}.$$

DÉMONSTRATION. Soit $x \in X$ une valeur d'adhérence de la suite $(x_n)_{n \geq 0}$ et $\varepsilon > 0$. D'après la caractérisation des valeurs d'adhérence telle qu'elle est donnée ci-dessus, nous savons que

$$\text{card} \{k \in \mathbf{N} : x_k \in B(x, \varepsilon)\} = +\infty,$$

ce qui nous assure que pour tout entier n et tout $\varepsilon > 0$, la boule $B(x, \varepsilon)$ rencontre $\{x_k : k \geq n\}$. D'après le point (ii) de la Proposition 1.4 page 20, on conclut que $x \in \overline{\{x_k : k \geq n\}}$.

Inversement, soit

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

et $\varepsilon > 0$. Étant donné que $B(x, \varepsilon)$ rencontre chacun des ensembles $X_n := \{x_k : k \geq n\}$, on construit par récurrence (grâce à la décroissance des X_n), une fonction strictement croissante φ telle que $x_{\varphi(n)} \in \{x_k : k \geq n\} \cap B(x, \varepsilon)$ ce qui nous garantit que le cardinal de $\{k \in \mathbf{N} : x_k \in B(x, \varepsilon)\}$ est bien infini. \square

6. Continuité des applications

Venons en maintenant à la continuité des applications définies entre deux espaces métriques. La définition précise de la continuité a été introduite par A. L. Cauchy¹ dans son cours d'analyse de l'École Royale Polytechnique (1821) : « *La fonction $f(x)$ sera, entre les deux limites assignées à la variable x , fonction continue de cette variable, si, pour chaque valeur de x intermédiaire entre ces limites, la valeur numérique de la différence $f(x + \alpha) - f(x)$ décroît indéfiniment avec celle de α* ». En langage moderne, cette définition peut être traduite par la :

Définition 1.10. Soient (X, d) et (Y, d') deux espaces métriques. On dit qu'une application $f : X \rightarrow Y$ est continue en $x_0 \in X$ si, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que

$$d(x_0, x) < \delta \quad \Rightarrow \quad d'(f(x_0), f(x)) < \varepsilon.$$

On dit que f est continue sur X si f est continue en tout point de X .²

On a également la :

Définition 1.11. Soient (X, d) et (Y, d') deux espaces métriques. On dit qu'une application $f : X \rightarrow Y$ est uniformément continue sur X si, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que, pour tous $x, x' \in X$

$$d(x, x') < \delta \quad \Rightarrow \quad d'(f(x), f(x')) < \varepsilon.$$

On prendra soin de ne pas confondre continuité et continuité uniforme. Clairement, une fonction uniformément continue sur X est continue sur X , mais la réciproque n'est pas forcément vraie (nous verrons prochainement - Corollaire 3.2 page 39 - une classe d'espaces métriques pour lesquelles toute application continue est uniformément continue). Par exemple, la fonction $x \mapsto x^2$ est continue sur \mathbf{R} mais n'est pas uniformément continue sur \mathbf{R} . En revanche, la fonction $x \mapsto \sqrt{x}$ est continue et uniformément continue sur $[0, +\infty[$ en vertu de l'inégalité

$$|\sqrt{x'} - \sqrt{x}| \leq \sqrt{|x' - x|},$$

1. A. L. Cauchy X 1804, professeur à l'École Polytechnique de 1815 à 1830.

2. Dans la définition de la continuité en un point, on trouve également la formulation $d(x_0, x) \leq \delta \Rightarrow d'(f(x_0), f(x)) \leq \varepsilon$, et le lecteur vérifiera facilement que les deux formulations sont en fait équivalentes.

pour tous $x, x' > 0$, qui permet de démontrer facilement la continuité uniforme de cette fonction.

On dispose d'une caractérisation pratique des applications continues :

Proposition 1.9. *On vérifie que l'on a équivalence entre les propriétés suivantes :*

- (i) $f : X \rightarrow Y$ est continue sur X ;
- (ii) l'image réciproque par f de tout ouvert de (Y, d') est un ouvert de (X, d) ;
- (iii) l'image réciproque par f de tout fermé de (Y, d') est un fermé de (X, d) .

DÉMONSTRATION. Démontrons l'équivalence entre (i) et (ii). Supposons dans un premier temps que f est continue sur X et donnons nous un ouvert U de Y . Soit $x \in f^{-1}(U)$. Par définition, $y = f(x) \in U$ qui est ouvert. Il existe donc $\varepsilon > 0$ tel que $B_Y(y, \varepsilon) \subset U$. La continuité de f en x nous assure qu'il existe $\delta > 0$ tel que $f(B_X(x, \delta)) \subset B_Y(y, \varepsilon) \subset U$, ce qui montre que $B_X(x, \delta) \subset f^{-1}(U)$. Par conséquent, $f^{-1}(U)$ est un ouvert de X . Inversement, supposons que l'image réciproque de tout ouvert de Y est un ouvert de X . Pour tout $x \in X$ et pour tout $\varepsilon > 0$, l'image réciproque de $B_Y(f(x), \varepsilon)$ par f est un ouvert de X qui contient x , donc il existe $\delta > 0$ tel que $B_X(x, \delta) \subset f^{-1}(B_Y(f(x), \varepsilon))$. Autrement dit, l'image de $B_X(x, \delta)$ par f est incluse dans $B_Y(f(x), \varepsilon)$, ce qui démontre la continuité de f au point x .

L'équivalence entre (ii) et (iii) est une conséquence du fait que l'image réciproque du complémentaire dans Y d'un ensemble A est égale au complémentaire dans X de l'image réciproque de A , autrement dit

$$f^{-1}(Y - A) = X - f^{-1}(A).$$

Démontrons par exemple que (ii) \Rightarrow (iii). Si F un fermé de Y , on peut écrire $F = Y - U$ où U est un ouvert de Y . Dans ce cas, $f^{-1}(F) = f^{-1}(Y - U) = X - f^{-1}(U)$ est un fermé de X comme complémentaire de l'ouvert $f^{-1}(U)$. \square

On vérifie que la composée d'applications continues est aussi une application continue. Plus précisément, si (X, d) , (Y, d') et (Z, d'') sont des espaces métriques, si $f : X \rightarrow Y$ est continue au point $x \in X$ et si $g : Y \rightarrow Z$ est continue au point $f(x) \in Y$, alors $g \circ f$ est continue au point x .

De même si $f : X \rightarrow E$ et $g : X \rightarrow E$ sont deux applications continues sur un espace métrique (X, d) (continues en $x \in X$) à valeurs dans un espace vectoriel normé $(E, \| \cdot \|)$ et si $\alpha, \beta : X \rightarrow \mathbf{K}$ sont deux fonctions continues sur X (continues en $x \in X$), alors $\alpha f + \beta g$ est continue sur X (continue en $x \in X$).

On dit qu'une application $f : X \rightarrow Y$ définie entre deux espaces métriques, est *lipschitzienne* de rapport $k > 0$ (ou encore *k-lipschitzienne*), si

$$d'(f(x), f(y)) \leq k d(x, y),$$

pour tous $x, y \in X$. Une application lipschitzienne est bien entendu continue.

Exemple 1.29. Soit (X, d) un espace métrique et $x_0 \in X$. Remarquons que l'inégalité triangulaire nous assure que

$$d(x, x_0) \leq d(x, y) + d(y, x_0),$$

pour tous $x, y \in X$. En particulier,

$$d(x, x_0) - d(y, x_0) \leq d(x, y).$$

En échangeant le rôle de x et de y , on conclut que

$$|d(x, x_0) - d(y, x_0)| \leq d(x, y),$$

ce qui montre que l'application $f_{x_0} : X \rightarrow \mathbf{R}$ définie par $f_{x_0}(x) := d(x, x_0)$ est continue et qu'elle est même 1-lipshitzienne.

Dans le cas où la distance d est associée à une norme $\| \cdot \|$ définie sur un espace vectoriel E , l'inégalité ci-dessus est une conséquence de l'inégalité bien connue

$$|\|y\| - \|x\|| \leq \|y - x\|,$$

pour tous $x, y \in E$.

Exercice 1.1. Soit (X, d) un espace métrique et $Y \subset X$.

1) Vérifier que l'application $f_Y : X \rightarrow \mathbf{R}$ définie par

$$f_Y(x) := \inf_{y \in Y} d(x, y),$$

est continue et qu'elle est en fait 1-lipshitzienne i.e. $|f_Y(x) - f_Y(y)| \leq d(x, y)$ pour tous $x, y \in X$.

2) Montrer que $x \in \bar{Y}$ (adhérence de Y) si et seulement si $f_Y(x) = 0$.

3) Montrer que les fermés de (X, d) sont les ensembles de zéros des fonctions continues sur X à valeurs réelles.

4) Soient $A, B \subset X$. On suppose que $\bar{A} \cap B = A \cap \bar{B} = \emptyset$. Montrer qu'il existe deux ouverts disjoints U et V tels que $A \subset U$ et $B \subset V$.

Le résultat qui suit nous donne une caractérisation de la continuité en terme de convergence de suites.

Proposition 1.10. Soient (X, d) et (Y, d') deux espaces métriques et $x \in X$. Une application $f : X \rightarrow Y$ est continue (continue au point x) si et seulement si pour toute suite $(x_n)_{n \geq 0}$ qui converge dans X (converge vers x dans X), on a

$$\lim_{n \rightarrow +\infty} f(x_n) = f\left(\lim_{n \rightarrow +\infty} x_n\right).$$

Dans le cas des applications linéaires définies entre espaces vectoriels normés, la définition de la continuité prend une forme particulièrement simple. En effet, on a la :

Proposition 1.11. Soient $(E, \| \cdot \|_E)$ et $(F, \| \cdot \|_F)$ deux espaces vectoriels normés et $L : E \rightarrow F$ une application linéaire. Alors, les propositions suivantes sont équivalentes :

(i) L est continue sur E ;

(ii) L est continue en 0 ;

(iii) il existe une constante $C > 0$ telle que

$$\|L(x)\|_F \leq C \|x\|_E,$$

pour tout $x \in E$.

DÉMONSTRATION. Clairement, (i) implique (ii). Supposons que L est continue en 0 et, dans la définition de la continuité de L en 0, choisissons $\varepsilon = 1$. Il existe donc une constante $\delta > 0$ telle que

$$\|x - 0\|_E \leq \delta \quad \Rightarrow \quad \|L(x) - L(0)\|_F \leq \varepsilon = 1.$$

Étant donné que $L(0) = 0$, on conclut que

$$\|x\|_E \leq \delta \Rightarrow \|L(x)\|_F \leq 1.$$

Maintenant, pour tout $x \in E - \{0\}$, on note

$$y := \delta \frac{x}{\|x\|_E}.$$

Par construction $\|y\|_E \leq \delta$, donc $\|L(y)\|_F \leq 1$. En utilisant la linéarité de f et l'homogénéité de la norme, on conclut que

$$\|L(x)\|_F \leq \frac{1}{\delta} \|x\|_E,$$

pour tout $x \in E$. Ce qui démontre (iii).

Enfin, si (iii) est vérifiée, on obtient par linéarité de L , l'inégalité

$$\|L(y) - L(x)\|_F = \|L(y - x)\|_F \leq C \|y - x\|_E,$$

pour tous $x, y \in E$. Autrement dit, l'application L est C -lipschitzienne sur E , donc continue sur E . \square

Définition 1.12. Si $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ sont des espaces vectoriels normés, on notera $L(E, F)$ l'espace vectoriel des applications linéaires qui sont définies sur E et à valeurs dans F . On notera $\mathcal{L}(E, F)$ le sous-espace vectoriel de $L(E, F)$ constitué des applications linéaires continues de E dans F .

Enfin, dans le cas où $F = \mathbf{K}$, on notera E' l'espace $\mathcal{L}(E, \mathbf{K})$ des formes linéaires continues sur E , i.e. les applications linéaires continues de E dans \mathbf{K} . L'espace E' est appelé dual topologique de $(E, \|\cdot\|_E)$.³

Pour toute $L \in \mathcal{L}(E, F)$, on peut définir

$$\|L\|_{\mathcal{L}(E, F)} := \sup_{x \in E - \{0\}} \frac{\|L(x)\|_F}{\|x\|_E}.$$

On laisse le soin au lecteur de vérifier que l'on définit ainsi une norme sur $\mathcal{L}(E, F)$ et que

$$\|L\|_{\mathcal{L}(E, F)} = \sup_{\substack{x \in E \\ \|x\|_E = 1}} \|L(x)\|_F = \sup_{\substack{x \in E \\ \|x\|_E \leq 1}} \|L(x)\|_F.$$

Par définition, on a

$$\|L(x)\|_F \leq \|L\|_{\mathcal{L}(E, F)} \|x\|_E,$$

pour tout $x \in E$. On vérifie aussi que si $L \in \mathcal{L}(E, F)$ et si $\tilde{L} \in \mathcal{L}(F, G)$, où E, F et G sont des espaces vectoriels normés, alors $\tilde{L} \circ L \in \mathcal{L}(E, G)$ et

$$\|\tilde{L} \circ L\|_{\mathcal{L}(E, G)} \leq \|L\|_{\mathcal{L}(E, F)} \|\tilde{L}\|_{\mathcal{L}(F, G)}.$$

En particulier, si $L \in \mathcal{L}(E, E)$, on a

$$\|L^n\|_{\mathcal{L}(E, E)} \leq \|L\|_{\mathcal{L}(E, E)}^n,$$

pour tout $n \in \mathbf{N}$. On comparera ces définitions et ces résultats avec ceux de l'exemple 1.7 page 12.

Donnons un premier exemple d'application linéaire continue, qui est définie sur un espace de dimension infinie.

3. On l'appelle *dual topologique* pour le distinguer du *dual algébrique* l'espace des formes linéaires (pas forcément continues). Il arrive lorsque le contexte est clair, que l'on omette l'adjectif « topologique ».

Exemple 1.30. On munit $\mathcal{C}([0, 1]; \mathbf{K})$ de la norme

$$\|f\|_\infty := \sup_{t \in [0, 1]} |f(t)|.$$

L'application linéaire $I : \mathcal{C}([0, 1]; \mathbf{K}) \rightarrow \mathcal{C}([0, 1]; \mathbf{K})$ définie par

$$I(f)(x) := \int_0^x f(t) dt,$$

pour tout $x \in [0, 1]$, est une application continue. On remarque immédiatement que $I(f) \in \mathcal{C}([0, 1]; \mathbf{K})$ si $f \in \mathcal{C}([0, 1]; \mathbf{K})$. De plus, pour tout $x \in [0, 1]$, on a

$$|I(f)|(x) \leq \int_0^x |f|(t) dt \leq \|f\|_\infty.$$

Donc

$$\|I(f)\|_\infty \leq \|f\|_\infty,$$

ce qui montre que l'application I est continue. On prendra soin de ne pas confondre le fait que $I(f)$ est une *fonction continue* sur $[0, 1]$ et le fait que l'application $f \mapsto I(f)$ est une *application continue* de $\mathcal{C}([0, 1]; \mathbf{K})$ dans lui-même.

On prendra garde que les applications linéaires définies entre espaces vectoriels normés de dimension infinie ne sont pas forcément continues, comme l'illustre les exemples suivants (nous reviendrons sur ce point ultérieurement).

Exemple 1.31. Dans un exemple précédent, nous avons déjà défini $\ell_c(\mathbf{N}; \mathbf{K})$ comme étant l'espace des suites indexées par \mathbf{N} , à valeurs dans \mathbf{K} , qui sont nulles à partir d'un certain rang. Cet espace vectoriel est muni de la norme

$$\|(x_n)_{n \geq 0}\|_\infty := \sup_{n \in \mathbf{N}} |x_n|.$$

Soit $(a_n)_{n \geq 0}$ une suite d'éléments de \mathbf{K} et $A : \ell_c(\mathbf{N}; \mathbf{K}) \rightarrow \ell_c(\mathbf{N}; \mathbf{K})$ l'application linéaire définie par

$$A((x_n)_{n \geq 0}) := (a_n x_n)_{n \geq 0}.$$

L'application A est clairement une application linéaire de $\ell_c(\mathbf{N}; \mathbf{K})$ dans lui-même mais n'est pas forcément continue. En fait, on vérifie que A est continue si et seulement si la suite $(a_n)_{n \geq 0}$ est bornée et que dans ce cas

$$\|A\|_{\mathcal{L}(\ell_c(\mathbf{N}; \mathbf{K}), \ell_c(\mathbf{N}; \mathbf{K}))} = \sup_{n \in \mathbf{N}} |a_n| = \|(a_n)_{n \geq 0}\|_\infty.$$

Le deuxième exemple reprend des éléments et notations de l'exemple 1.12 page 14.

Exemple 1.32. On munit $\mathcal{C}([0, 1]; \mathbf{K})$ de la norme

$$\|v\|_1 := \int_0^1 |v(t)| dt.$$

L'application

$$A : \mathcal{C}([0, 1]; \mathbf{K}) \rightarrow \mathbf{K},$$

définie par $A(v) = v(0)$ est clairement linéaire mais elle n'est pas continue de $(\mathcal{C}([0, 1]; \mathbf{K}), \|\cdot\|_1)$ dans $(\mathbf{K}, |\cdot|)$. Pour le voir, comme dans l'exemple 1.12 page 14, on construit pour tout $j \in \mathbf{N} - \{0\}$ la fonction

$$v_j(x) := (1 - jx)^+.$$

On a

$$\|v_j\|_1 = \frac{1}{2j} \quad \text{et} \quad A(v_j) = 1.$$

Ce qui montre que l'on ne peut pas avoir une inégalité de la forme $|A(v)| \leq C \|v\|_1$, valable pour toute $v \in \mathcal{C}([0, 1]; \mathbf{K})$.

En revanche, si l'on munit l'espace $\mathcal{C}([0, 1]; \mathbf{K})$ de la norme

$$\|v\|_\infty := \sup_{[0,1]} |v|,$$

on vérifie que $|A(v)| \leq \|v\|_\infty$, ce qui montre que, cette fois-ci, l'application A est continue de $(\mathcal{C}([0, 1]; \mathbf{K}), \|\cdot\|_\infty)$ dans $(\mathbf{K}, |\cdot|)$.

On retiendra que, pour les applications linéaires définies sur des espaces vectoriels de dimension infinie, la continuité dépend de la norme utilisée.

7. Connexité

Intuitivement, un espace métrique connexe est un espace qui est *fait d'un seul tenant*.

Définition 1.13. *Un espace métrique (X, d) est connexe s'il n'existe pas de sous-ensemble de X autre que \emptyset et X qui soit à la fois ouvert et fermé. Un sous-ensemble $Y \subset X$ d'un espace métrique (X, d) est connexe si (Y, d) , muni de la distance induite est un espace métrique connexe.*

Exemple 1.33. Nous avons déjà vu ci-dessus que \mathbf{R} , muni de la distance usuelle, est un espace métrique connexe. En revanche, $X := \mathbf{R} - \{0\}$, muni de la distance usuelle, n'est pas connexe car $] -\infty, 0[$ est ouvert ouvert et également fermé dans X en tant que complémentaire dans $\mathbf{R} - \{0\}$ de l'ouvert $]0, +\infty[$ de $\mathbf{R} - \{0\}$.

La connexité de X s'exprime aussi en disant qu'il n'existe pas de fonction $f : X \rightarrow \mathbf{R}$, continue, non constante et à valeurs dans $\{0, 1\}$: en effet si l'on pouvait écrire $X = U \cup V$ avec U, V ouverts disjoints non vides, la fonction caractéristique de U définie par 1 sur U et 0 sur le complémentaire de U , serait une fonction continue (l'image réciproque d'un ouvert de \mathbf{R} est soit l'ensemble vide, soit égale à U ou à V , soit égale à X tout entier, donc dans tous les cas c'est un ouvert) et non constante. Inversement, si $f : X \rightarrow \mathbf{R}$ est une fonction continue non constante qui prend ses valeurs dans $\{0, 1\}$, les ensembles $U = f^{-1}(] -\infty, 1/2[)$ et $V = f^{-1}(]1/2, +\infty[)$ sont des ouverts non vides, disjoints de (X, d) , dont la réunion est égale à X .

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

Proposition 1.12. *Soient (X, d) et (Y, d') deux espaces métriques et f une application continue de X dans Y . Si X est connexe, alors $f(X)$ est connexe.*

DÉMONSTRATION. Si $f(X) = A \cup B$ où A et B sont des ouverts disjoints, alors $f^{-1}(A)$ et $f^{-1}(B)$ sont ouverts disjoints et leur réunion est égale à X . Par connexité de X , l'un des deux ensembles $f^{-1}(A)$ ou $f^{-1}(B)$ est vide, par exemple $f^{-1}(A)$ est vide. Comme f est surjective sur son image, A est contenu dans $f(f^{-1}(A))$ et l'on conclut que A est vide. \square

La notion de connexité est une propriété relativement délicate à manier, même si la définition peut paraître simple. Une propriété plus intuitive est la *connexité par arcs* que nous allons maintenant définir.

Définition 1.14. *Un espace métrique (X, d) est connexe par arcs si deux points quelconques de X peuvent être reliés par un arc continu. C'est-à-dire que, pour tous $x, y \in X$, il existe $\gamma \in \mathcal{C}([0, 1]; X)$ telle que $\gamma(0) = x$ et $\gamma(1) = y$.*

La connexité par arcs entraîne de manière évidente la connexité. En effet, supposons que (X, d) est connexe par arc et que f est une fonction continue définie sur X , qui prend ses valeurs dans $\{0, 1\}$. Pour tous $x, y \in X$, il existe un chemin continu γ qui relie x à y , le théorème des valeurs intermédiaires impose à la fonction $f \circ \gamma$ de prendre la même valeur en 0 et 1, donc $f(x) = f(y)$ et l'on montre ainsi que f est une fonction constante. Donc, (X, d) est un espace métrique connexe. Insistons sur le fait que tous les espaces métriques connexe, qui sont utilisés dans ce cours, sont en fait des espaces métriques connexes par arcs.

Exemple 1.34. Le cercle, la sphère et le tore sont des sous-ensembles connexes par arc de l'espace euclidien. Un espace vectoriel normé, un sous-ensemble convexe d'un espace vectoriel normé sont des ensembles connexes par arc. Un sous-ensemble X d'un espace vectoriel normé est dit *étoilé*, s'il existe un point $x \in X$ tel que tout segment entre x et un autre point de X est contenu dans X . Un ensemble étoilé est connexe par arc.

On vérifie qu'une réunion de sous-ensembles connexes (connexes par arc) d'un même espace métrique, qui ont un point commun, est un ensemble connexe (connexe par arc). Nous terminons ce chapitre par la :

Définition 1.15. *Soit x un point d'un espace métrique (X, d) . La réunion de tous les sous-ensembles connexes de X qui contiennent le point x est un ensemble connexe que l'on appelle la composante connexe du point x dans X .*

CHAPITRE 2

Quelques mots sur la dénombrabilité

1. Définitions et propriétés

Commençons par rappeler la définition d'un ensemble dénombrable.

Définition 2.1. *On dit qu'un ensemble X est dénombrable¹ s'il est fini ou s'il est en bijection avec \mathbf{N} .*

Donnons quelques exemples classiques :

Exemple 2.1. Les ensembles $\mathbf{N} - \{0\}$, $2\mathbf{N}$, \mathbf{Z} sont dénombrables. En effet, l'application $\phi_0(n) := n + 1$ réalise une bijection de \mathbf{N} sur $\mathbf{N} - \{0\}$, l'application $\phi_1(n) := 2n$ est une bijection de \mathbf{N} sur $2\mathbf{N}$ et l'application ϕ_2 définie par

$$\begin{cases} \phi_2(2n) &:= -n \\ \phi_2(2n+1) &:= n+1, \end{cases}$$

est une bijection de \mathbf{N} sur \mathbf{Z} .

Proposition 2.1. *Un sous-ensemble $X \subset \mathbf{N}$ est dénombrable.*

DÉMONSTRATION. Si X est fini, c'est terminé. Supposons que X est infini. L'idée est de numéroter les éléments de X par ordre croissant. On définit par récurrence une application $\phi : \mathbf{N} \rightarrow X$ de la manière suivante :

$$\phi(0) := \inf\{x \in X\} \quad \text{et} \quad \phi(n+1) := \inf\{x \in X : x > \phi(n)\},$$

pour tout $n \geq 1$. On vérifie que ϕ est une bijection de \mathbf{N} sur X . □

La propriété suivante permet de trouver des démonstrations simples du fait qu'un ensemble est dénombrable :

Proposition 2.2. *Soit X un ensemble. S'il existe une application $f : X \rightarrow \mathbf{N}$ qui est injective alors X est dénombrable. S'il existe une application $f : \mathbf{N} \rightarrow X$ qui est surjective alors X est dénombrable.*

DÉMONSTRATION. Supposons que X est infini, autrement c'est terminé. Commençons par le cas où il existe une application $f : X \rightarrow \mathbf{N}$ qui est injective. Dans ce cas, X est en bijection avec $f(X)$, qui est un sous-ensemble infini de \mathbf{N} . D'après le résultat précédent, il existe une bijection h entre $f(X)$ et \mathbf{N} . Alors, $h \circ f$ réalise une bijection de X sur \mathbf{N} .

Supposons maintenant qu'il existe une application $f : \mathbf{N} \rightarrow X$ qui est surjective. Pour tout $x \in X$, définissons $g(x) := \inf\{y \in \mathbf{N} : f(y) = x\}$. On vérifie que $g : X \rightarrow \mathbf{N}$ est injective et l'on conclut, en utilisant la première partie de la démonstration, que X est dénombrable. □

1. Attention, certain auteurs définissent un ensemble dénombrable comme un ensemble en bijection avec \mathbf{N} .

La première partie de ce résultat permet en particulier d'affirmer qu'un sous-ensemble d'un ensemble dénombrable est dénombrable. On pourra également vérifier tout l'intérêt du résultat précédent en essayant de démontrer que \mathbf{N}^2 est dénombrable. En effet, l'application

$$\psi(n, m) := 2^n 3^m,$$

est une application injective (Théorème fondamental de l'arithmétique) de \mathbf{N}^2 dans \mathbf{N} . Le résultat précédent nous permet de conclure immédiatement que \mathbf{N}^2 est dénombrable. Autrement, le lecteur pourra essayer de démontrer que la *fonction de couplage de Cantor*

$$\phi(n, m) := \frac{(n+m)(n+m+1)}{2} + n,$$

réalise effectivement une bijection de \mathbf{N}^2 sur \mathbf{N} .

La démonstration du fait que \mathbf{N}^2 est dénombrable se généralise pour démontrer la :

Proposition 2.3. *Pour tout $N \geq 1$, l'ensemble \mathbf{N}^N est dénombrable.*

DÉMONSTRATION. On se donne des nombres premiers distincts p_1, \dots, p_N et l'on considère l'application

$$\psi(n_1, \dots, n_N) := p_1^{n_1} \dots p_N^{n_N},$$

qui est injective de \mathbf{N}^N dans \mathbf{N} . Le résultat précédent permet de conclure que \mathbf{N}^N est dénombrable. \square

Le résultat précédent permet également de démontrer la :

Proposition 2.4. *Les deux propositions suivantes sont vraies :*

- (i) *un produit fini d'ensembles dénombrables est dénombrable ;*
- (ii) *une réunion dénombrable d'ensembles dénombrables est dénombrable.*

DÉMONSTRATION. Commençons par démontrer qu'un produit fini d'ensembles dénombrables est dénombrable. On se donne X_1, \dots, X_N des ensembles dénombrables et, pour tout $i = 1, \dots, N$, une application injective $f_i : X_i \rightarrow \mathbf{N}$. Alors, l'application

$$\psi((x_1, \dots, x_N)) = (f_1(x_1), \dots, f_N(x_N)),$$

est injective de $X_1 \times \dots \times X_N$ dans \mathbf{N}^N . Étant donné que \mathbf{N}^N est dénombrable, on construit facilement une application injective de $X_1 \times \dots \times X_N$ dans \mathbf{N} en composant l'application ψ construite ci-dessus avec une application bijective de \mathbf{N}^N dans \mathbf{N} , ce qui prouve que $X_1 \times \dots \times X_N$ est dénombrable.

Passons maintenant à la démonstration du fait qu'une réunion dénombrable d'ensembles dénombrables est dénombrable. Soit I un ensemble dénombrable et $(X_i)_{i \in I}$ une famille d'ensembles dénombrables indexée par I . Pour chaque $i \in I$, il existe une application injective $f_i : X_i \rightarrow \mathbf{N}$. On note

$$Z := \bigcup_{i \in I} \{(i, f_i(x)) : x \in X_i\} \subset I \times \mathbf{N}.$$

D'après la première partie de la preuve, l'ensemble $I \times \mathbf{N}$ est dénombrable, donc Z est dénombrable comme un sous-ensemble d'un ensemble dénombrable. Définissons enfin l'application $\phi : Z \rightarrow \bigcup_{i \in I} X_i$ par $\phi(z) := f_i^{-1}(x)$ si $z = (i, x)$. Cette application est clairement surjective et Z est dénombrable, donc $\bigcup_{i \in I} X_i$ est dénombrable. \square

Exemple 2.2. L'ensemble des rationnels \mathbf{Q} est dénombrable. Pour le vérifier, il suffit de considérer l'application $\phi : \mathbf{Z} \times (\mathbf{N} - \{0\}) \rightarrow \mathbf{Q}$ définie par

$$\phi(p, q) := \frac{p}{q},$$

qui est surjective et de la composer avec une bijection entre \mathbf{N} et $\mathbf{Z} \times (\mathbf{N} - \{0\})$ (qui sont deux ensembles dénombrables infinis).

2. Topologie des réels et dénombrabilité

La notion de dénombrabilité permet de montrer que les ouverts de \mathbf{R} sont finalement relativement simples à comprendre.

Proposition 2.5. *Un ouvert U de \mathbf{R} est une réunion dénombrable d'intervalles ouverts deux à deux disjoints.*

DÉMONSTRATION. Pour tout $x \in U$, on note I_x la réunion des intervalles ouverts qui sont inclus dans U et qui contiennent x : c'est le plus grand intervalle ouvert qui est inclus dans U et qui contient x . On note

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

La remarque clef est que deux éléments 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 inclus à la fois dans I_x et I_y (maximalité) de sorte que $I_x = I_y = I$. On peut donc écrire

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

Montrons que \mathcal{F} est dénombrable. Pour tout $I \in \mathcal{F}$, choisissons un rationnel $r_I \in \mathbf{Q} \cap I$ (tout intervalle non-vidé contient un rationnel, par densité des rationnels dans \mathbf{R}). On définit ainsi une application $f(I) = r_I$ de \mathcal{F} dans \mathbf{Q} . Clairement cette application est injective car, si $f(I) = f(J)$, les intervalles I et J ont un point commun et donc sont égaux. Donc \mathcal{F} , comme \mathbf{Q} , est dénombrable. \square

La non dénombrabilité de \mathbf{R} est bien connue. Nous donnons ci-dessous une très jolie preuve due à Cantor, le fondateur de l'arithmétique des cardinaux infinis².

Théorème 2.1 (Argument diagonal de Cantor). *L'ensemble $[0, 1[$ (et donc \mathbf{R}) n'est pas dénombrable.*

DÉMONSTRATION. On raisonne par l'absurde. Supposons que $[0, 1[$ soit dénombrable et que l'on a une bijection $\varphi : \mathbf{N} - \{0\} \rightarrow [0, 1[$. On écrit alors le développement (propre) en base 10 de chaque élément de $[0, 1[$

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

Pour tout $n \geq 1$, choisissons $a_n := 0$ si $a_{n,n} = 1$ et $a_n := 1$ si $a_{n,n} \neq 1$. On définit

$$x := \sum_{k \geq 1} a_k 10^{-k} \in [0, 1[.$$

2. Le lecteur pourra consulter l'article de P. Dehornoy *Cantor et les Infinités* paru en 2009 dans la Gazette de la SMF et disponible à l'adresse suivante http://smf4.emath.fr/Publications/Gazette/2009/121/smf_gazette_121_28-46.pdf

Pour tout $n \in \mathbf{N} - \{0\}$, par construction, $x \neq \varphi(n)$ car $a_n \neq a_{n,n}$, ce qui contredit le fait que φ est surjective. \square

CHAPITRE 3

Compacité et complétude

Il nous reste à définir et comprendre deux notions particulièrement importantes en topologie : la compacité et la complétude. Ces deux notions seront au cœur des développements que nous effectuerons dans les chapitres suivants, notamment dans le chapitre sur les espaces de Lebesgue et le chapitre sur les espaces de Hilbert.

1. La compacité

1.1. Définition. Soit (X, d) un espace métrique. Étant donnée une famille d'ouverts qui recouvrent X , peut-on extraire de ce recouvrement un sous-recouvrement fini ? Autrement dit, si $(A_i)_{i \in I}$ est une famille d'ouverts de (X, d) et si

$$X = \bigcup_{i \in I} A_i,$$

existe-t-il un sous-ensemble *fini* $J \subset I$ tel que

$$X = \bigcup_{i \in J} A_i ?$$

Si X est un ensemble fini (muni par exemple de la distance discrète définie par $d(x, x) = 0$ et $d(x, y) = 1$ si $x \neq y$), c'est évidemment le cas. Si X n'est pas de cardinal fini, ce n'est plus le cas en général, même si X est borné. Par exemple, si l'on considère l'ensemble $]0, 1[$ muni de la topologie associée à la distance usuelle, on vérifie que

$$\bigcup_{n \geq 3}]\frac{1}{n}, \frac{n-1}{n}[,$$

constitue un recouvrement de $]0, 1[$ par les ouverts $] \frac{1}{n}, \frac{n-1}{n} [$, pour $n \geq 3$, dont on ne peut pas extraire de sous-recouvrement fini.

Proposition 3.1. *Soit Y un sous-ensemble d'un espace métrique (X, d) , alors (Y, d) est un espace compact si et seulement si de tout recouvrement de Y par des ouverts de (X, d) , on peut extraire un sous-recouvrement fini.*

DÉMONSTRATION. On remarque que tout ouvert V de (Y, d) s'écrit sous la forme $V = Y \cap U$ où U est un ouvert de (X, d) . □

Remarque 3.1. Si (X, d) est un espace métrique et si $Y \subset X$ est tel que (Y, d) est un espace compact, on dira simplement que « Y est un compact de X » (ou de (X, d) , étant sous-entendu que Y est (naturellement) muni de la distance d).

Les espaces métriques jouissant de cette propriété sont par définition les espaces *compacts* :

Définition 3.1. *Un espace métrique (X, d) est un espace compact si l'une des deux propriétés équivalentes suivantes est vérifiée :*

- (i) de tout recouvrement de X par des ouverts, on peut extraire un sous-recouvrement fini ;
- (ii) toute famille de fermés de X d'intersection vide admet une sous-famille finie d'intersection vide.

Pour vérifier l'équivalence entre les propositions (i) et (ii), il suffit de passer aux ensembles complémentaires.

Proposition 3.2. *Soit (X, d) un espace métrique. Les propriétés suivantes sont vérifiées :*

- (a) si (X, d) est compact et $Y \subset X$ est un fermé de (X, d) , alors (Y, d) est un compact ;
- (b) si $Y \subset X$ et (Y, d) compact, alors Y est un fermé de (X, d) ;
- (c) une intersection décroissante de compacts non vides est non vide ;
- (d) l'image d'un compact par une application continue est un compact.

DÉMONSTRATION. La proposition (a) découle immédiatement de (ii) dans la définition ci-dessus. La proposition (d) découle de (i). Enfin, la proposition (c) découle de (ii) et de (b). Reste à démontrer que, dans un espace métrique (X, d) , un compact Y est fermé.

Cette propriété résulte du fait que la topologie sur X est *séparée* : pour tout $x, x' \in Y$, $x \neq x'$, il existe des ouverts disjoints $U_{x,x'}, U_{x',x}$ contenant x et x' respectivement (il suffit de considérer les boules ouvertes $B(x, \varepsilon)$ et $B(x', \varepsilon)$ avec $2\varepsilon < d(x, x')$). Fixons $x \in X - Y$ et choisissons pour tout $y \in Y$ des ouverts disjoints $U_{x,y}, U_{y,x}$ contenant x, y respectivement. Extrayons du recouvrement de Y

$$Y \subset \bigcup_{y \in Y} U_{y,x},$$

un sous-recouvrement fini

$$Y \subset V := \bigcup_{i=1}^n U_{y_i, x}.$$

Par construction,

$$U := \bigcap_{i=1}^n U_{x, y_i},$$

ne rencontre pas V et, *a fortiori*, ne rencontre pas Y . Or, U est ouvert dans X (en tant qu'intersection finie d'ouverts), et est donc un voisinage ouvert de x dans $X - Y$. Conclusion, $X - Y$ est ouvert et Y est fermé dans (X, d) . \square

1.2. Caractérisation séquentielle des compacts. Le résultat qui suit est un des résultats importants en topologie. Sa démonstration utilise de manière cruciale que nous travaillons dans un espace métrique :

Théorème 3.1 (Théorème de Bolzano-Weierstrass). *Un espace métrique (X, d) est compact si et seulement si, de toute suite d'éléments de X , on peut extraire une sous-suite qui converge.*

DÉMONSTRATION. Supposons X compact et soit $(x_n)_{n \geq 0}$ une suite d'éléments de X . Il suffit de montrer que l'intersection décroissante des fermés

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

est non vide, car tout élément dans cette intersection est une valeur d'adhérence de la suite $(x_n)_{n \geq 0}$ et est donc limite d'une suite extraite de la suite $(x_n)_{n \geq 0}$ (voir la Proposition 1.8 page 22). On raisonne par l'absurde. Supposons que ce ne soit pas le cas i.e. que cette intersection est vide. La compacité de X entraînerait l'existence de $\bar{n} \geq 0$ tel que

$$\bigcap_{n \leq \bar{n}} \overline{\{x_k : k \geq n\}} = \overline{\{x_k : k \geq \bar{n}\}} = \emptyset,$$

ce qui conduit à une contradiction.

Pour démontrer la réciproque, on utilise le lemme clef :

Lemme 3.1 (Nombre de Lebesgue). *On suppose que (X, d) est un espace métrique dans lequel la propriété suivante est vérifiée : de toute suite d'éléments de X on peut extraire une sous-suite qui converge. Soit $(U_i)_{i \in I}$ un recouvrement de X par des ouverts. Alors, il existe $\delta > 0$ tel que, pour tout $x \in X$, il existe $i \in I$, tel que $B(x, \delta) \subset U_i$.*

DÉMONSTRATION. On raisonne par l'absurde. Supposons que la propriété ne soit pas vérifiée. On construit alors une suite $(x_n)_{n \geq 1}$ d'éléments de X , telle que pour tous n et i , $B(x_n, \frac{1}{n}) \not\subset U_i$. Choisissons une valeur d'adhérence x de la suite $(x_n)_{n \geq 1}$, de sorte que l'on peut écrire

$$x = \lim_{n \rightarrow +\infty} x_{\varphi(n)},$$

où φ est strictement croissante. Comme la réunion des U_i forme un recouvrement de X par des ouverts, on peut choisir j tel que $x \in U_j$. Comme U_j est ouvert, on peut choisir $r > 0$ tel que $B(x, r) \subset U_j$.

Étant donné que

$$\lim_{n \rightarrow +\infty} x_{\varphi(n)} = x \quad \text{et} \quad \lim_{n \rightarrow +\infty} \frac{1}{\varphi(n)} = 0,$$

on a

$$B\left(x_{\varphi(n)}, \frac{1}{\varphi(n)}\right) \subset B(x, r) \subset U_j,$$

pour tout n suffisamment grand, et ceci contredit la définition de x_n . \square

Terminons la démonstration du Théorème de Bolzano-Weierstrass. On suppose que toute union finie des $(U_i)_{i \in I}$ est différente de Y . Soit δ un nombre de Lebesgue qui est donné par le lemme ci-dessus. On construit alors, par récurrence, une suite $(x_n)_{n \geq 1}$ d'éléments de Y et une suite $(U^{(n)})_{n \geq 0}$ d'éléments de $\{U_i : i \in I\}$ de la manière suivante : tout d'abord, $U^{(0)} = \emptyset$, ensuite, si x_1, \dots, x_n et $U^{(0)}, \dots, U^{(n)}$ ont déjà été construits, on choisit $x_{n+1} \in X$ tel que

$$x_{n+1} \in X - \left(U^{(1)} \cup \dots \cup U^{(n)}\right),$$

et l'on choisit $U^{(n+1)} \in \{U_i : i \in I\}$ tel que

$$B(x_{n+1}, \delta) \subset U^{(n+1)}.$$

Par construction, pour tous $m < n$, on a $x_n \notin B(x_m, \delta) \subset U^{(m)}$, donc, on en déduit que $d(x_m, x_n) \geq \delta$ pour $m \neq n$. La suite $(x_n)_{n \geq 1}$ ne peut donc pas avoir de suite extraite qui converge, ce qui constitue une contradiction. \square

Remarque 3.2. Dans le cas où X est fini, ce résultat est une conséquence du *principe des tiroirs* qui dit qu'une suite, qui ne prend qu'un nombre fini de valeurs, prend une infinité de fois la même valeur !

Corollaire 3.1. *Le produit $X \times Y$ de deux espaces métriques compacts (X, d) et (Y, d') , muni de la distance somme ou de la distance produit (voir l'exemple 1.20 page 17), est un espace métrique compact.*

DÉMONSTRATION. Soit $((x_n, y_n))_{n \geq 0}$ une suite d'éléments de $X \times Y$. On utilise dans un premier temps la compacité de (X, d) pour extraire de la suite $(x_n)_{n \geq 0}$, une sous-suite $(x_{\varphi(n)})_{n \geq 0}$ qui converge vers x . Ensuite, on utilise la compacité de (Y, d) pour extraire de la suite $(y_{\varphi(n)})_{n \geq 0}$, une sous-suite $(y_{\varphi(\psi(n))})_{n \geq 0}$ qui converge vers y . Enfin, on observe que la suite extraite $(x_{\varphi(\psi(n))}, y_{\varphi(\psi(n))})_{n \geq 0}$ converge vers (x, y) dans $X \times Y$. En particulier, (x, y) est une valeur d'adhérence de la suite $((x_n, y_n))_{n \geq 0}$. \square

Bien entendu, ce résultat se généralise au produit d'un nombre fini de compacts.

Lemme 3.2. *L'intervalle $[0, 1]$ est un compact de \mathbf{R} (muni de la topologie usuelle associée à la distance $d(x, y) = |y - x|$).*

DÉMONSTRATION. Donnons de ce résultat, une première démonstration qui utilise la caractérisation des compacts par les recouvrements. Soit $(U_i)_{i \in I}$ une famille d'ouverts qui constitue un recouvrement de $[0, 1]$. On note

$$W := \{m \in [0, 1] : [0, m] \text{ admet un recouvrement fini par des } U_i\}.$$

On vérifie que $W \neq \emptyset$ (car $0 \in [0, 1]$ donc il existe $i \in I$ tel que $0 \in U_i$, ce qui prouve que $0 \in W$). Ensuite, on vérifie que, par construction, W est un sous-intervalle de $[0, 1]$, il est donc de la forme $[0, c[$ ou $[0, c]$. Supposons que $c < 1$. Dans ce cas, on remarque qu'il existe $i \in I$ tel que $c \in U_i$. L'ensemble U_i étant ouvert, il existe $\varepsilon > 0$ tel que $[c - \varepsilon, c + \varepsilon] \cap [0, 1] \subset U_i$ et $c + \varepsilon \leq 1$. On vérifie alors que $[0, c + \varepsilon]$ peut être recouvert par un nombre fini des U_i . Donc $[0, c + \varepsilon] \subset W$, ce qui contredit la définition de c . Conclusion, $c = 1$ et en utilisant un raisonnement analogue, on montre que nécessairement $W = [0, 1]$.

On peut donner une autre démonstration de ce résultat en utilisant cette fois-ci le Théorème de Bolzano-Weierstrass. Soit $(x_n)_{n \geq 0}$ une suite d'éléments de $[0, 1]$. On définit deux suites $(a_n)_{n \geq 0}$ et $(b_n)_{n \geq 0}$ d'éléments de $[0, 1]$ de la manière suivante : on pose $a_0 = 0$ et $b_0 = 1$ et, si a_n et b_n ont déjà été construits, on prend $a_{n+1} = a_n$ et $b_{n+1} = \frac{a_n + b_n}{2}$ si $\{i \geq 0 : x_i \in [a_n, \frac{a_n + b_n}{2}]\}$ est infini ; sinon on prend $a_{n+1} = \frac{a_n + b_n}{2}$ et $b_{n+1} = b_n$. On vérifie que la suite $(a_n)_{n \geq 0}$ est croissante et que la suite $(b_n)_{n \geq 0}$ est décroissante. Enfin, $|b_n - a_n| = 2^{-n}$ et $\{i \geq 0 : x_i \in [a_n, b_n]\}$ est infini. Les deux suites $(a_n)_{n \geq 0}$ et $(b_n)_{n \geq 0}$ étant adjacentes, elles convergent vers une limite commune que nous noterons x . Pour tout $\varepsilon > 0$, il existe $n \geq 0$ tel que $x - \varepsilon \leq a_n \leq b_n \leq x + \varepsilon$. En particulier, $\{i \geq 0 : x_i \in]x - \varepsilon, x + \varepsilon[\}$ est infini. Donc x est une valeur d'adhérence de la suite $(x_n)_{n \geq 0}$. Ce qui montre que $[0, 1]$ est un compact. \square

Soient $a < b$. En utilisant la fonction $f(t) = (b - a)t + a$ et le fait que l'image d'un compact par une fonction continue est un compact (voir le point (d) de la Proposition 3.2 page 36), on déduit du résultat précédent que tout intervalle fermé $[a, b]$ est un compact de \mathbf{R} (muni de la topologie usuelle).

On suppose, dans le résultat suivant, que l'espace \mathbf{R}^N est muni de la topologie associée à la norme

$$\|(x_1, \dots, x_N)\|_\infty = \max_{i=1, \dots, N} |x_i|.$$

Comme nous le verrons un peu plus tard, pour définir la topologie de \mathbf{R}^N , le choix de la norme n'a pas d'importance.

Théorème 3.2 (Théorème de Borel-Lebesgue). *Un sous-ensemble de \mathbf{R}^N , muni de la norme somme $\| \cdot \|_\infty$, est compact si et seulement s'il est fermé et borné.*

DÉMONSTRATION. Nous avons déjà vu qu'un compact est un fermé. En utilisant le Théorème de Bolzano-Weierstrass on vérifie qu'un compact X de \mathbf{R}^N est nécessairement borné. Autrement, on pourrait construire une suite $(x_n)_{n \geq 0}$ d'éléments de X telle que $\|x_n\|_\infty \geq n$ et l'on ne peut certainement pas extraire d'une telle suite, une sous-suite qui converge.

Supposons maintenant que X est fermé et borné. Rappelons qu'un fermé inclus dans un compact est un compact. L'ensemble X étant borné, il existe $a > 0$ tel que $X \subset [-a, a]^N$. Il suffit donc de prouver que $[-a, a]^N$ est un compact de \mathbf{R}^N mais c'est une simple conséquence du fait que l'intervalle $[-a, a]$ est un compact de \mathbf{R} et qu'un produit fini de compacts est un compact. \square

Un résultat particulièrement important pour les applications est le suivant :

Théorème 3.3. *Une fonction continue, à valeurs réelles et définie sur un espace métrique compact est bornée et atteint ses bornes.*

DÉMONSTRATION. Afin de démontrer la deuxième propriété, utilisons le fait que l'image d'un compact par une application continue est un compact (voir la Proposition 3.2 page 36), l'image d'un compact X par une fonction continue $f : X \rightarrow \mathbf{R}$ est un compact de \mathbf{R} , dont on vient de voir que c'était un sous-ensemble fermé et borné de \mathbf{R} . En particulier, la fonction f est bornée et le fait qu'elle atteigne ses bornes est une simple traduction de la fermeture de son image qui implique que $\inf_X f$ et $\sup_X f$ appartiennent à $f(X)$. \square

On rappelle qu'une fonction f définie sur (X, d) à valeurs dans (X', d') est *uniformément continue* si, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que, pour tous $x, y \in X$

$$d(x, y) < \delta \quad \Rightarrow \quad d'(f(x), f(y)) < \varepsilon.$$

Nous avons le :

Corollaire 3.2 (Théorème de Heine). *Soit $f : X \rightarrow Y$ une application continue d'un espace métrique compact (X, d) dans un espace métrique (Y, d') (quelconque), alors f est uniformément continue.*

DÉMONSTRATION. Soit $\varepsilon > 0$ et définissons, pour tout $\delta > 0$, l'ensemble

$$K_\delta := \{(x, y) \in X \times X : d(x, y) \leq \delta \quad \text{et} \quad d'(f(x), f(y)) \geq \varepsilon\}.$$

Il suffit de montrer qu'il existe $\delta > 0$ tel que $K_\delta = \emptyset$. Par continuité de la fonction distance et de la fonction f , K_δ est fermé dans $X \times X$, en tant qu'image réciproque d'un fermé par une application continue, c'est donc un compact (car $X \times X$ est un compact d'après le Corollaire 3.1 page ci-contre). Donc, la famille des K_δ est une famille décroissante de compacts d'intersection vide. En vertu du point (c) de la Proposition 3.2 page 36, il existe $\delta_0 > 0$ tel que $K_{\delta_0} = \emptyset$. \square

2. Espaces vectoriels de dimension finie

Commençons par nous affranchir du choix des normes.

Théorème 3.4. *Toutes les normes sur \mathbf{R}^N sont équivalentes.*

DÉMONSTRATION. Soit N une norme sur \mathbf{R}^N . On note (e_1, \dots, e_N) la base canonique de \mathbf{R}^N et l'on définit $a = \max_{i=1, \dots, N} \mathcal{N}(e_i)$. Clairement $a > 0$.

Si $x = (x_1, \dots, x_N) \in \mathbf{R}^N$, on note $\|x\|_\infty := \max_{i=1, \dots, N} |x_i|$. L'inégalité triangulaire donne le sens facile, à savoir

$$(3.1) \quad \mathcal{N}(x) = \mathcal{N}\left(\sum_{i=1}^N x_i e_i\right) \leq \sum_{i=1}^N |x_i| \mathcal{N}(e_i) \leq \left(\sum_{i=1}^N \mathcal{N}(e_i)\right) \|x\|_\infty.$$

L'inégalité inverse est plus difficile à établir. Tout d'abord, l'inégalité (3.1) et le fait que

$$|\mathcal{N}(y) - \mathcal{N}(x)| \leq \mathcal{N}(y - x),$$

nous assurent que l'application

$$\mathcal{N} : (\mathbf{R}^N, \|\cdot\|_\infty) \rightarrow (\mathbf{R}, |\cdot|),$$

est continue. Mais dans ce cas, le Théorème 3.2 nous assure que $S := \{x \in \mathbf{R}^N : \|x\|_\infty = 1\}$, la sphère unité de $(\mathbf{R}^N, \|\cdot\|_\infty)$ est fermée et bornée et est donc compacte. Le Théorème 3.3 page précédente nous assure que \mathcal{N} atteint ses bornes sur S et qu'en particulier cette fonction est minorée par $\mathcal{N}(x_0) > 0$, pour un certain $x_0 \in S$. On a donc, pour tout $x \in S$,

$$\mathcal{N}(x) \geq \mathcal{N}(x_0).$$

Par homogénéité des normes, on en déduit que $\mathcal{N}(x) \geq \mathcal{N}(x_0) \|x\|_\infty$. \square

Ce résultat se généralise à tout espace vectoriel de dimension finie, une fois que un tel espace est identifié à un espace euclidien. En effet, si E est un \mathbf{R} -espace vectoriel de dimension finie N qui est muni de deux normes \mathcal{N}_1 et \mathcal{N}_2 et si (e_1, \dots, e_N) est une base de E , on peut définir sur \mathbf{R}^N les applications

$$\tilde{\mathcal{N}}_j(x_1, \dots, x_N) := \mathcal{N}_j\left(\sum_{i=1}^N x_i e_i\right),$$

dont on vérifie que ce sont bien des normes sur \mathbf{R}^N . Le résultat précédent nous assure que ces deux normes $\tilde{\mathcal{N}}_1$ et $\tilde{\mathcal{N}}_2$ sont équivalentes et ceci montre immédiatement que les normes \mathcal{N}_1 et \mathcal{N}_2 sont elles aussi équivalentes.

En revanche, dans les espaces de dimension infinie, deux normes ne sont pas forcément équivalentes, comme le montre l'exemple 1.12 page 14.

Remarque 3.3. Le Théorème 3.4 page précédente nous assure que, sur un espace vectoriel de dimension finie, le choix de la norme n'a pas d'influence sur la topologie, qui est appelée « topologie de la norme ».

Comme corollaire, nous obtenons le :

Corollaire 3.3 (Théorème de Borel-Lebesgue). *Pour toute norme sur \mathbf{R}^N , les sous-ensembles compacts de \mathbf{R}^N sont les fermés, bornés.*

Comme le résultat précédent, ce résultat se généralise à tout espace vectoriel normé de dimension finie : dans un espace vectoriel normé de dimension finie, les sous-ensembles compacts sont les fermés, bornés (remarquer que l'on n'a pas besoin de préciser la norme utilisée).

Attention : le résultat précédent tombe en défaut dans le cas des espaces de dimension infinie. Plus précisément, dans un espace vectoriel normé de dimension *infinie*, la boule unité fermée n'est jamais compacte. En fait, le Théorème de Riesz (Théorème 3.8 page 47) nous assure que la boule unité fermée d'un espace vectoriel normé est compacte si et seulement si

l'espace est de dimension finie. Nous verrons une démonstration de ce Théorème dans un cadre général et nous verrons également une autre démonstration (plus simple) du même résultat dans le cadre des espaces de Hilbert (voir la démonstration du Théorème 10.8 page 213).

Exemple 3.1. Pour tout $i \in \mathbf{N}$, on note \mathbf{e}_i la suite dont tous les éléments sont nuls, sauf le i -ième qui vaut 1. Clairement, la suite $(\mathbf{e}_i)_{i \geq 0}$ est une suite de la boule unité fermée de $(\ell^\infty(\mathbf{N}; \mathbf{K}), \|\cdot\|_\infty)$. Néanmoins, on ne peut pas extraire de cette suite une sous-suite qui converge car

$$\|\mathbf{e}_j - \mathbf{e}_i\|_\infty = 1,$$

si $i \neq j$.

Une autre conséquence importante de l'équivalence des normes en dimension finie est le fait que, si E est un espace vectoriel normé de dimension finie et F est un espace vectoriel normé, toute application linéaire de E dans F est continue.

Proposition 3.3. *Si E est un espace vectoriel normé de dimension finie et F est un espace vectoriel normé, alors $L(E, F)$, l'espace des applications linéaires de E dans F coïncide avec $\mathcal{L}(E, F)$ l'espace des applications linéaires continues de E dans F .*

DÉMONSTRATION. L'espace E étant de dimension finie, toutes les normes sur E sont équivalentes. Par conséquent, si (e_1, \dots, e_N) est une base de E , on peut considérer sur E la norme

$$\left\| \sum_{i=1}^N x_i e_i \right\|_E := \sup_{i=1, \dots, N} |x_i|.$$

Dans ce cas, en utilisant la linéarité de L et l'inégalité triangulaire, on trouve

$$\|L(x)\|_F = \left\| L \left(\sum_{i=1}^N x_i e_i \right) \right\|_F \leq \sum_{i=1}^N |x_i| \|L(e_i)\|_F \leq \left(\sum_{i=1}^N \|L(e_i)\|_F \right) \|x\|_E.$$

Ce qui, en vertu de la Proposition 1.11 page 25, montre que L est continue. \square

Dans le cas où les applications sont définies sur un espace de dimension infinie, la linéarité n'entraîne pas systématiquement la continuité comme le montrent l'exemple 1.31 page 27 et l'exemple 1.32 page 28.

3. Quelques applications

3.1. Théorème de Dini. Le Théorème de Dini nous sera utile dans la construction de l'intégrale de Lebesgue.

Proposition 3.4 (Théorème de Dini). *Soit (X, d) un espace métrique et $(f_n)_{n \geq 0}$ une suite décroissante de fonctions continues à support compact qui converge simplement vers 0, i.e. $\forall x \in X, \lim_{n \rightarrow +\infty} f_n(x) = 0$. Alors, la suite $(f_n)_{n \geq 0}$ converge uniformément vers 0 i.e. $\lim_{n \rightarrow +\infty} (\sup_X f_n) = 0$.*

Autrement dit, une suite monotone de fonctions $(f_n)_{n \geq 0}$ qui converge *simplement* vers une fonction f sur un compact, converge *uniformément* vers f .

DÉMONSTRATION. Remarquons que les hypothèses de décroissance de la suite font que, nécessairement, $f_n \geq 0$ et donc, pour montrer la convergence uniforme, il suffit de montrer que la suite $(\sup_X f_n)_{n \geq 0}$ converge vers 0. Soit $\varepsilon > 0$. Définissons

$$F_n := \{x \in X : f_n(x) \geq \varepsilon\}.$$

La suite de sous-ensembles $(F_n)_{n \geq 0}$ décroît puisque la suite de fonctions $(f_n)_{n \geq 0}$ est une suite décroissante et F_n est compact puisque c'est un fermé dans un compact (le support de f_0 par exemple). Par hypothèse $\lim_{n \rightarrow +\infty} f_n(x) = 0$ pour tout $x \in X$, l'intersection des F_n est donc vide. Or, une intersection décroissante de compacts non vides est non vide (voir le point (c) de la Proposition 3.2 page 36) de sorte que nécessairement, il existe $\bar{n} \in \mathbf{N}$ tel que $F_{\bar{n}} = \emptyset$. La décroissance des F_n assure que, pour tout $n \geq \bar{n}$, $F_n = \emptyset$ et donc $0 \leq \sup_X f_n \leq \varepsilon$. Ce qui termine la démonstration. \square

3.2. Théorème d'Urysohn. Le Théorème d'Urysohn trouve de nombreuses applications en analyse mais aussi en géométrie (existence de partitions de l'unité subordonnées à un recouvrement par des ouverts).

Théorème 3.5 (Théorème d'Urysohn). *Soit U un ouvert d'un espace métrique (X, d) et K un compact inclus dans U . Alors, il existe une fonction continue $f \in \mathcal{C}(X; \mathbf{R})$ à valeurs dans $[0, 1]$, qui vaut 1 sur K et 0 sur $X - U$.*

DÉMONSTRATION. Pour tout $A \subset X$, on note

$$\delta(x, A) := \inf_{y \in A} d(x, y).$$

Puisque U est un ouvert, $X - U$ est un fermé. Ceci implique que pour tout $x \in U$, $\delta(x, X - U)$, qui représente la distance de x au fermé $X - U$, est strictement positive (voir l'exercice 1.1 page 25). Par compacité de K et continuité de la fonction $x \mapsto \delta(x, X - U)$ (c.f. *loc. cit.* et le Théorème 3.3 page 39), la fonction $x \mapsto \delta(x, X - U)$ atteint ses bornes sur K . Étant donné que $K \subset U$, on conclut qu'il existe $\varepsilon > 0$ tel que, pour tout $x \in K$ et pour tout $y \in X - U$, on a $d(x, y) \geq \varepsilon$.

La fonction définie sur X par

$$f(x) := \max \left(0, 1 - \frac{\delta(x, K)}{\varepsilon} \right),$$

est continue, vaut 1 sur K et 0 sur $X - U$. \square

Remarque 3.4. Dans le cas où $X = \mathbf{R}^N$, on peut raffiner ce résultat et montrer que la fonction f peut être choisie parmi les fonctions de classe \mathcal{C}^∞ sur \mathbf{R}^N . Ce résultat utilise des outils qui seront introduits dans les chapitres sur l'intégration, notamment la convolution de fonctions.

Si U est un ouvert d'un espace métrique (X, d) , on note $\mathcal{C}_c(U; \mathbf{R})$ l'espace vectoriel des fonctions continues à valeurs réelles qui sont définies sur U et dont le support est un compact de U . Soit $\Lambda : \mathcal{C}_c(U; \mathbf{R}) \rightarrow \mathbf{R}$ une *forme linéaire positive*, i.e. une application linéaire de $\mathcal{C}_c(U; \mathbf{R})$ à valeurs dans \mathbf{R} qui vérifie

$$f \geq 0 \quad \Rightarrow \quad \Lambda(f) \geq 0.$$

Si tel est le cas, on a évidemment

$$f \geq g \quad \Rightarrow \quad \Lambda(f) \geq \Lambda(g).$$

Donnons trois exemples importants.

Exemple 3.2. Si $X = \mathbf{R}$ et $U =]a, b[$, l'application

$$\Lambda(f) := \int_a^b f(t) dt,$$

est une forme linéaire positive sur $\mathcal{C}_c(U; \mathbf{R})$.

Exemple 3.3. Si $X = \mathbf{R}$ et si $x_0 \in \mathbf{R}$, l'application

$$\Lambda_{x_0}(f) := f(x_0),$$

est une forme linéaire positive sur $\mathcal{C}_c(U; \mathbf{R})$.

Le dernier exemple est un peu plus compliqué.

Exemple 3.4. Si $X = \mathbf{N}$ muni de la distance usuelle, l'espace $\mathcal{C}_c(\mathbf{N}; \mathbf{R})$ s'identifie à l'espace des suites réelles qui sont indexées par \mathbf{N} et qui sont nulles à partir d'un certain rang. L'application

$$\Lambda(f) := \sum_{n \in \mathbf{N}} f(n),$$

est une forme linéaire positive sur $\mathcal{C}_c(\mathbf{N}; \mathbf{R})$.

Dans le cours sur l'intégrale de Lebesgue, nous aurons besoin de la :

Proposition 3.5. *Soit U un ouvert de \mathbf{R}^N et $(f_n)_{n \geq 0}$ une suite décroissante de fonctions continues à support compact dans U , qui converge simplement vers 0. Alors*

$$\lim_{n \rightarrow +\infty} \Lambda(f_n) = 0.$$

DÉMONSTRATION. Soit $K \subset U$ un compact en dehors duquel f_0 est nulle. Étant donnée la décroissance de la suite $(f_n)_{n \geq 0}$, chaque f_n est nulle en dehors de K . Comme dans la démonstration du Théorème d'Urysohn, on montre qu'il existe $\varepsilon > 0$ tel que, pour tout $x \in K$ et pour tout $y \in \mathbf{R}^N - U$, $d(x, y) \geq 2\varepsilon$.

On note $U_\varepsilon := \{x \in \mathbf{R}^N : d(x, K) < \varepsilon\}$, c'est un ouvert borné. On note $K_\varepsilon := \{x \in \mathbf{R}^N : d(x, K) \leq \varepsilon\}$ qui est un fermé, borné, donc compact. On remarque que $K \subset U_\varepsilon \subset K_\varepsilon \subset U$. La fonction définie sur \mathbf{R}^N par

$$g(x) := \max \left(0, 1 - \frac{d(x, K)}{\varepsilon} \right),$$

est continue, vaut 1 sur K et 0 sur $\mathbf{R}^N - K_\varepsilon$. De plus, par construction

$$f_n = g f_n,$$

car $g = 1$ sur le support de la fonction f_n et donc

$$\Lambda(f_n) = \Lambda(g f_n) \leq \left(\sup_{\mathbf{R}^N} |f_n| \right) \Lambda(g),$$

par positivité de Λ . On applique enfin le Théorème de Dini (Théorème 3.4 page 41), pour conclure que $\sup_{\mathbf{R}^N} |f_n|$ tend vers 0, ce qui termine la démonstration. \square

4. Espaces métriques complets et espaces de Banach

La complétude est une notion de toute première importance en analyse. Commençons par la :

Définition 3.2. *Une suite $(x_n)_{n \geq 0}$ d'un espace métrique (X, d) est une suite de Cauchy si*

$$\forall \varepsilon > 0, \quad \exists n_0 \in \mathbf{N}, \quad \text{tel que} \quad (\forall n, m \geq n_0, \quad d(x_n, x_m) < \varepsilon).$$

Remarquons qu'une suite qui converge est automatiquement une suite de Cauchy et qu'une suite de Cauchy est toujours bornée. En revanche, une suite de Cauchy n'est, en général, pas convergente. Par exemple, si l'on considère $X = \mathbf{Q}$ muni de la distance usuelle, la densité de \mathbf{Q} dans \mathbf{R} nous montre qu'il existe une suite de rationnels $(r_n)_{n \geq 0}$ qui converge dans \mathbf{R} vers $\sqrt{2}$. En particulier, la suite $(r_n)_{n \geq 0}$ est une suite de Cauchy de \mathbf{Q} (muni de la distance usuelle sur \mathbf{R}), mais cette suite n'a pas de limite dans \mathbf{Q} car $\sqrt{2} \notin \mathbf{Q}$! Donnons un autre exemple pour convaincre le lecteur. Considérons, dans $X =]0, 1[$ muni de la distance $d(x, y) = |x - y|$, la suite $(x_n)_{n \geq 1}$ définie par $x_n := \frac{1}{n}$. On vérifie que c'est une suite de Cauchy, mais qu'elle ne converge pas dans $]0, 1[$.

Lemme 3.3. *Dans un espace métrique (X, d) , une suite de Cauchy $(x_n)_{n \geq 0}$ qui possède une valeur d'adhérence est convergente.*

DÉMONSTRATION. Soit x une valeur d'adhérence de la suite $(x_n)_{n \geq 0}$ et $\varepsilon > 0$. Choisissons $n_0 \in \mathbf{N}$ tel que pour tous $n, m \geq n_0$, $d(x_n, x_m) < \varepsilon/2$. Ensuite, choisissons $n_1 \geq n_0$, tel que $d(x_{n_1}, x) < \varepsilon/2$. Alors, pour tout $n \geq n_1$, on a $d(x_n, x) < \varepsilon$. Ce qui termine la démonstration. \square

Définition 3.3 (Espace complet). *Un espace métrique (X, d) est complet si et seulement si toute suite de Cauchy converge dans X .*

Exemple 3.5. Une conséquence immédiate du Lemme 3.3 est qu'un espace métrique compact est un espace métrique complet. En effet, si $(x_n)_{n \geq 0}$ est une suite de Cauchy dans un espace métrique compact, elle admet une valeur d'adhérence et, en vertu du Lemme ci-dessus, elle converge.

On vérifie sans trop de difficulté qu'une partie F d'un espace métrique complet X est complète si et seulement si elle est fermée (il suffit pour le voir d'appliquer la caractérisation des fermés donnée par la Proposition 1.3 page 20). Cette propriété permet de donner l'exemple fondamental d'espace métrique complet.

Proposition 3.6. *L'espace \mathbf{R}^N , muni de distance associée à la norme*

$$\|(x_1, \dots, x_N)\|_\infty := \max_{i=1, \dots, N} |x_i|,$$

est un espace métrique complet.

DÉMONSTRATION. Soit $(x_n)_{n \geq 0}$ une suite de Cauchy. Cette suite est bornée, elle est donc incluse dans un pavé $[-a, a]^N$ de \mathbf{R}^N . Un tel pavé est compact, donc, on peut extraire de la suite $(x_n)_{n \geq 0}$ une sous-suite qui converge. En particulier, la suite $(x_n)_{n \geq 0}$ admet une valeur d'adhérence et, en vertu du Lemme 3.3, elle converge. \square

Il convient de faire attention au fait que la complétude est une propriété qui dépend de la distance et qui ne dépend pas seulement de la topologie sur l'ensemble X . Prenons par exemple, $X = \mathbf{R}$ muni de la distance

$$d'(x, y) := |e^{-x} - e^{-y}|.$$

On vérifie qu'un sous-ensemble de \mathbf{R} est un ouvert pour la topologie associée à la distance d' si et seulement si c'est un ouvert pour la topologie associée à la distance usuelle $d(x, y) := |x - y|$. Néanmoins, (\mathbf{R}, d') n'est pas un espace métrique complet. En effet la suite $(n)_{n \geq 0}$ est une suite de Cauchy dans (\mathbf{R}, d') , pourtant elle ne converge pas dans \mathbf{R} pour cette distance. Nous laissons au lecteur le soin de rédiger les détails des propositions avancées ci-dessus.

En revanche, si d_1 et d_2 sont deux distances Lipschitz-équivalentes sur X , on vérifie qu'une suite est une suite de Cauchy pour la distance d_1 si et seulement si c'est une suite de Cauchy pour la distance d_2 et on vérifie également que (X, d_1) est un espace métrique complet si et seulement si (X, d_2) est un espace métrique complet.

Nous avons vu (Théorème 3.4 page 39) qu'en dimension finie toutes les normes sont équivalentes, en particulier, en utilisant la Proposition ci-dessus, on montre que :

Théorème 3.6. *Un espace vectoriel normé de dimension finie, muni de la distance associée à la norme, est un espace métrique complet.*

Définition 3.4. *Un espace de Banach est un espace vectoriel normé complet.*

Les espaces de Banach jouent un rôle absolument fondamental dans plusieurs branches des mathématiques et des mathématiques appliquées. Ils jouent un rôle particulièrement important dans la résolution des problèmes variationnels et dans la résolution des équations aux dérivées partielles.

Exemple 3.6. Nous avons déjà rencontré $\ell^\infty(\mathbf{N}; \mathbf{K})$, l'espace des suites qui sont indexées par \mathbf{N} , à valeurs dans \mathbf{K} et qui sont bornées, muni de la norme

$$\|(x_n)_{n \geq 0}\|_\infty := \sup_{n \geq 0} |x_n|.$$

Vérifions maintenant que $(\ell^\infty(\mathbf{N}; \mathbf{K}), \|\cdot\|_\infty)$ est un espace de Banach, c'est-à-dire que c'est un espace vectoriel complet pour la distance associée à la norme définie ci-dessus.

Soit $(\mathbf{x}^m)_{m \geq 0}$ est une suite de Cauchy d'éléments de $(\ell^\infty(\mathbf{N}; \mathbf{K}), \|\cdot\|_\infty)$. On note

$$\mathbf{x}^m := (x_n^m)_{n \in \mathbf{N}},$$

où $x_n^m \in \mathbf{K}$. Par définition, pour tout $\varepsilon > 0$, il existe $m_0 \geq 0$ tel que, pour tous $m, m' \geq m_0$,

$$\|\mathbf{x}^m - \mathbf{x}^{m'}\|_\infty = \sup_{n \geq 0} |x_n^m - x_n^{m'}| < \varepsilon.$$

En particulier, ceci entraîne que, pour chaque $n \geq 0$, la suite $(x_n^m)_{m \geq 0}$ est une suite de Cauchy dans $(\mathbf{K}, |\cdot|)$, qui est un espace métrique complet. Donc, cette suite converge vers une limite que l'on note $z_n \in \mathbf{K}$, i.e. $\lim_{m \rightarrow +\infty} x_n^m = z_n$.

On note $\mathbf{z} := (z_n)_{n \geq 0}$. Vérifions d'une part que $\mathbf{z} \in \ell^\infty(\mathbf{N}; \mathbf{K})$ et d'autre part que l'on a $\lim_{m \rightarrow +\infty} \mathbf{x}^m = \mathbf{z}$. Exploitions à nouveau le fait que, pour tout $\varepsilon > 0$, il existe m_0 tel que, pour tous $m, m' \geq m_0$ et tout $n \geq 0$, on a $|x_n^m - x_n^{m'}| < \varepsilon$. Par passage à la limite quand m' tend vers l'infini, on conclut que $|x_n^m - z_n| \leq \varepsilon$. En particulier, pour tout $m \geq m_0$, on a $\|\mathbf{x}^m - \mathbf{z}\|_\infty \leq \varepsilon$, ce qui d'une part implique que la suite \mathbf{z} est bornée (prendre par exemple $\varepsilon = 1$) et d'autre part, traduit le fait que la suite $(\mathbf{x}^m)_{m \geq 0}$ converge vers \mathbf{z} pour la norme $\|\cdot\|_\infty$.

Exemple 3.7. Nous verrons dans le chapitre sur les espaces de Hilbert (voir l'exemple 10.2 page 190), que l'espace $\ell^2(\mathbf{N}; \mathbf{K})$ des suites $(x_n)_{n \geq 0}$ à valeurs dans \mathbf{K} telles que $\sum_{n \geq 0} |x_n|^2 < \infty$

muni de la norme définie par

$$\|(x_n)_{n \geq 0}\|_2 := \left(\sum_{n \geq 0} |x_n|^2 \right)^{1/2},$$

est un espace de Banach (en fait, c'est un espace de Hilbert). Nous verrons que c'est l'exemple standard d'espace de Hilbert de dimension infinie (voir le Corollaire 10.4 page 210).

Terminons cette première introduction aux espaces métriques complets par le :

Lemme 3.4. *Le produit de deux espaces métriques complets $(X_i, d_i), i = 1, 2$ muni de la distance somme ou de la distance produit (voir l'exemple 1.20 page 17) est un espace métrique complet.*

DÉMONSTRATION. On note d_p la métrique produit sur $X_1 \times X_2$. Dans un premier temps, on remarque que si $((x_{1,n}, x_{2,n}))_{n \geq 0}$ est une suite de Cauchy dans $(X_1 \times X_2, d_p)$, alors les suites $(x_{i,n})_{n \geq 0}$ sont des suites de Cauchy dans (X_i, d_i) . Elles convergent donc vers une limite notée x_i . Enfin, on montre que la suite $((x_{1,n}, x_{2,n}))_{n \geq 0}$ converge vers (x_1, x_2) dans (X_1, X_2) . Le lemme suit. \square

Le Théorème des fermés emboîtés. Soit (X, d) un espace métrique complet. Définissons le diamètre d'un sous-ensemble $Y \subset X$ par la formule

$$\text{diam}(Y) := \sup_{y, y' \in Y} d(y, y') \in \bar{\mathbf{R}}^+ := [0, +\infty) \cup \{+\infty\}.$$

Théorème 3.7 (Théorème des fermés emboîtés). *Soit $(F_n)_{n \geq 0}$ une suite décroissante de fermés non vides d'un espace métrique complet (X, d) , dont la suite des diamètres tend vers 0. Alors, $\bigcap_{n \geq 0} F_n$, l'intersection des F_n , est réduite à un point.*

DÉMONSTRATION. Pour tout entier $n \geq 0$, choisissons $x_n \in F_n$. La suite $(F_n)_{n \geq 0}$ étant décroissante, pour tous $p \leq q$, on a $x_p, x_q \in F_p$ donc $d(x_p, x_q) \leq \text{diam}(F_p)$. Étant donné que $\lim_{p \rightarrow +\infty} \text{diam}(F_p) = 0$, la suite $(x_n)_{n \geq 0}$ est une suite de Cauchy de (X, d) , qui est supposé complet, et donc elle converge vers une limite que nous noterons x . Fixons alors $p \geq 0$. On a $x = \lim_{n \rightarrow +\infty} x_{n+p}$ de sorte que x est aussi limite d'une suite d'éléments de F_p . Étant donné que F_p est fermé, on en déduit que $x \in F_p$, donc $x \in \bigcap_{n \geq 0} F_n$.

Montrons maintenant l'unicité du point qui appartient à l'intersection des F_n . Si $x, x' \in \bigcap_{n \geq 0} F_n$ alors $\text{diam}(F_n) \geq d(x, x')$ et donc $d(x, x') = 0$, ce qui implique que $x = x'$. \square

5. Théorème de Riesz

Nous donnons ici un énoncé et une démonstration du Théorème de Riesz, ce qui nous permettra de nous familiariser avec certaines notions introduites dans ce chapitre. Une autre démonstration, dans un cadre plus simple, sera donnée ultérieurement dans le chapitre sur les espaces de Hilbert. Commençons par le :

Lemme 3.5 (Lemme de Riesz). *Soit $(E, \| \cdot \|)$ un espace vectoriel normé et F un sous-espace fermé de E . On suppose que $F \neq E$. Alors, pour tout $\varepsilon \in]0, 1[$, il existe $x \in E$ tel que $\|x\| = 1$ et*

$$\min_{y \in F} \|x - y\| \geq 1 - \varepsilon.$$

DÉMONSTRATION. Soit $y \in E - F$. Par hypothèse F est fermé et $y \notin F$ donc

$$\alpha := \inf_{z \in F} \|y - z\| > 0.$$

Choisissons $z \in F$ tel que $\|y - z\| \leq \frac{\alpha}{1-\varepsilon}$. On note

$$x := \frac{y - z}{\|y - z\|}.$$

Par construction, $\|x\| = 1$ et, pour tout $z' \in F$, on peut écrire

$$\|x - z'\| = \frac{1}{\|y - z\|} (\|y - z - \|y - z\| z'\|)$$

Étant donné que $z + \|y - z\| z' \in F$, on conclut que

$$\|x - z'\| \geq \frac{\alpha}{\|y - z\|} \geq 1 - \varepsilon.$$

Ce qui termine la démonstration. \square

Nous aurons également besoin du :

Lemme 3.6. *Dans un espace vectoriel normé, un sous-espace vectoriel de dimension finie est fermé.*

DÉMONSTRATION. On commence par démontrer que, dans un espace vectoriel normé, un sous-espace vectoriel de dimension finie F , muni de la norme induite, est un espace de Banach.

Pour démontrer cette assertion, on considère une base (e_1, \dots, e_N) de F et, toutes les normes étant équivalentes sur F , on peut choisir la norme

$$\|x\| := \sup_{i=1, \dots, N} |x_i|, \quad \text{si} \quad x = \sum_{i=1}^N x_i e_i.$$

On voit immédiatement que si $(x^j)_{j \geq 0}$ est une suite de Cauchy de F , alors $((x_1^j, \dots, x_N^j))_{j \geq 0}$, la suite des coordonnées des x^j dans la base (e_1, \dots, e_N) , est une suite de Cauchy de \mathbf{K}^N , donc elle converge dans cet espace vers $(x_1, \dots, x_N) \in \mathbf{K}^N$. On vérifie ensuite que la suite $(x^j)_{j \geq 0}$

converge vers $y := \sum_{i=1}^N x_i e_i$ car

$$\lim_{j \rightarrow +\infty} \|x^j - x\| = \lim_{j \rightarrow +\infty} \left(\sup_{i=1, \dots, N} |x_i^j - x_i| \right) = 0.$$

Conclusion, F , muni de la norme induite, est un espace de Banach, donc c'est un fermé. \square

Nous pouvons maintenant énoncer le :

Théorème 3.8 (Théorème de Riesz). *Un espace vectoriel normé $(E, \|\cdot\|)$ est de dimension finie si et seulement si $B_f(0, 1)$, la boule unité fermée de E , est compacte.*

DÉMONSTRATION. Nous avons déjà vu que, dans un espace vectoriel normé de dimension finie, les compacts sont les fermés bornés. Donc, $B_f(0, 1)$ est compacte si E est de dimension finie.

Inversement, on suppose que $B_f(0, 1)$ est compacte. Raisonnons par l'absurde et supposons que E n'est pas de dimension finie. Choisissons alors un vecteur $e_0 \in E$ tel que $\|e_0\| = 1$. On note $F_0 := \text{Vect}\{e_0\}$. D'après le Lemme de Riesz, on peut trouver un vecteur $e_1 \notin F_0$ tel que $\|e_1\| = 1$ et $\inf_{x \in F_0} \|x - e_1\| \geq 1/2$. On note $F_1 := \text{Vect}\{e_0, e_1\}$. En utilisant le Lemme de Riesz, on construit ainsi par récurrence une suite de vecteurs $(e_n)_{n \geq 0}$ de norme 1 et une suite de sous-espaces de dimension finie $F_n := \text{Vect}\{e_0, \dots, e_n\}$ tels que $e_n \notin F_{n-1}$ et $\inf_{x \in F_{n-1}} \|x - e_n\| \geq 1/2$. En particulier, $\|e_m - e_n\| \geq 1/2$ si $m < n$.

Clairement, la suite $(e_n)_{n \geq 0}$ est une suite d'éléments de $B_f(0, 1)$ dont aucune sous-suite ne converge, ce qui contredit la compacité de $B_f(0, 1)$. \square

6. Espaces d'applications continues

Dans cette section, nous démontrons que deux espaces d'applications à valeurs dans un espace complet sont aussi, pourvus qu'ils soient munis des normes adéquates, des espaces complets.

Théorème 3.9. *Soient (X, d) un espace métrique compact et (X', d') est un espace métrique complet. L'espace $\mathcal{C}(X; X')$ des fonctions continues définies sur X à valeurs dans X' , muni de la distance d_∞ définie par*

$$d_\infty(f, g) := \sup_{x \in X} d'(f(x), g(x)),$$

est un espace métrique complet.

DÉMONSTRATION. Soit $(f_n)_{n \geq 0}$ une suite de Cauchy dans $(\mathcal{C}(X; X'), d_\infty)$. Pour tout $\varepsilon > 0$, il existe $\bar{n} \in \mathbf{N}$ tel que, pour tous $n, m \geq \bar{n}$, on a

$$d_\infty(f_n, f_m) < \varepsilon.$$

Donc, si $x \in X$, on a pour tous $n, m \geq \bar{n}$

$$(3.2) \quad d(f_n(x), f_m(x)) < \varepsilon.$$

Ceci montre que la suite $(f_n(x))_{n \geq 0}$ est une suite de Cauchy dans (X', d') qui est un espace métrique complet. Cette suite converge donc vers une limite que l'on note $f(x)$.

Montrons que l'application $f : X \rightarrow X'$ ainsi définie est bien continue. Fixons $x \in X$ et $\varepsilon > 0$. On sait qu'il existe $\bar{m} \in \mathbf{N}$ tel que pour tous $n, m \geq \bar{m}$, on a

$$d_\infty(f_n, f_m) < \varepsilon/3.$$

En particulier, pour tout $y \in X$ et pour tout $m \geq \bar{m}$

$$d'(f_{\bar{m}}(y), f_m(y)) < \varepsilon/3.$$

Faisons tendre m vers l'infini pour conclure que

$$d'(f_{\bar{m}}(y), f(y)) \leq \varepsilon/3.$$

Maintenant, utilisons la continuité de l'application $f_{\bar{m}}$. Il existe $\delta > 0$ tel que, si $d(x, y) < \delta$ alors $d'(f_{\bar{m}}(x), f_{\bar{m}}(y)) < \varepsilon/3$. L'inégalité triangulaire nous assure que

$$d'(f(y), f(x)) \leq d'(f(y), f_{\bar{m}}(y)) + d'(f_{\bar{m}}(y), f_{\bar{m}}(x)) + d'(f_{\bar{m}}(x), f(x)) < \varepsilon,$$

pourvu que $d(x, y) < \delta$. Ce qui démontre la continuité de f .

Enfin, montrons que $(f_n)_{n \geq 0}$ converge vers f . Pour ceci reprenons (3.2) et faisons tendre m vers l'infini pour conclure que, pour tout $x \in X$ et pour tout $n \geq \bar{n}$,

$$d(f_n(x), f(x)) \leq \varepsilon.$$

Autrement dit, $d_\infty(f_n, f) \leq \varepsilon$ pour tout $n \geq \bar{n}$. Ce qui montre la convergence souhaitée. \square

Dans le cas particulier où (X, d) un espace métrique compact et où $(F, \|\cdot\|_F)$ est un espace de Banach, le résultat précédent nous assure que $\mathcal{C}(X; F)$, muni de la norme de la convergence uniforme

$$\|f\|_\infty := \sup_{x \in X} \|f(x)\|_F,$$

est un espace de Banach.

Exemple 3.8. L'espace $\mathcal{C}([0, 1]; \mathbf{K})$ muni de la norme de la convergence uniforme

$$\|f\|_\infty := \sup_{t \in [0, 1]} |f(t)|,$$

est un espace de Banach.

Rappelons que l'on peut définir la norme

$$\|L\|_{\mathcal{L}(E, F)} := \sup_{x \in E - \{0\}} \frac{\|L(x)\|_F}{\|x\|_E},$$

pour toute application linéaire continue $L \in \mathcal{L}(E, F)$. Le résultat qui suit nous assure que, muni de cette norme, l'espace $\mathcal{L}(E, F)$ est un espace de Banach dès lors que $(F, \|\cdot\|_F)$ est un espace de Banach.

Théorème 3.10. *Supposons que $(F, \|\cdot\|_F)$ est un espace de Banach. Alors, l'espace $\mathcal{L}(E, F)$ muni de la norme $\|\cdot\|_{\mathcal{L}(E, F)}$ définie ci-dessus est un espace de Banach.*

DÉMONSTRATION. La démonstration du fait que $\mathcal{L}(E, F)$ est un espace de Banach reprend les arguments utilisés dans la démonstration du résultat précédent. Soit $(L_n)_{n \geq 0}$ une suite d'applications linéaires continues sur E et à valeurs dans F , qui soit une suite de Cauchy dans $(\mathcal{L}(E, F), \|\cdot\|_{\mathcal{L}(E, F)})$.

Pour tout $\varepsilon > 0$, il existe $\bar{n} \in \mathbf{N}$ tel que, pour tous $n, m \geq \bar{n}$

$$\|L_n - L_m\|_{\mathcal{L}(E, F)} < \varepsilon.$$

Autrement dit, pour tout $x \in E$,

$$(3.3) \quad \|L_n(x) - L_m(x)\|_F \leq \|L_n - L_m\|_{\mathcal{L}(E, F)} \|x\|_E < \varepsilon \|x\|_E.$$

En particulier, pour tout x de E , la suite $(L_n(x))_{n \geq 0}$ est une suite de Cauchy de F . Comme F est un espace de Banach, cette suite converge vers une limite notée $L(x) \in F$.

Vérifions que $x \mapsto L(x)$ est une application linéaire et continue de E dans F , et ensuite que la suite $(L_n)_{n \geq 0}$ converge vers L . La démonstration de la linéarité ne pose aucun problème (il suffit pour cela d'utiliser l'unicité de la limite des suites dans un espace vectoriel normé), seule la continuité est plus délicate à démontrer. Nous voulons montrer qu'il existe une constante $C > 0$ telle que $\|L(x)\|_F \leq C \|x\|_E$, pour tout $x \in E$. Par passage à la limite dans (3.3) quand n tend vers $+\infty$, on trouve que, pour tout $x \in E$

$$\begin{aligned} \|L(x)\|_F &\leq \|L(x) - L_{\bar{n}}(x)\|_F + \|L_{\bar{n}}(x)\|_F \\ &\leq \lim_{n \rightarrow +\infty} \|L_n(x) - L_{\bar{n}}(x)\|_F + \|L_{\bar{n}}(x)\|_F \\ &\leq (\varepsilon + \|L_{\bar{n}}\|_{\mathcal{L}(E, F)}) \|x\|_E, \end{aligned}$$

ce qui prouve la continuité de L .

Enfin, par passage à la limite dans (3.3) quand n tend vers $+\infty$, on obtient aussi

$$\|L(x) - L_m(x)\|_F = \lim_{n \rightarrow +\infty} \|L_n(x) - L_m(x)\|_F \leq \varepsilon \|x\|_E,$$

pour tout $m \geq \bar{n}$ et pour tout $x \in E$. Autrement dit

$$\|L - L_m\|_{\mathcal{L}(E, F)} \leq \varepsilon,$$

pour tout $m \geq \bar{n}$, ce qui montre que la suite $(L_n)_{n \geq 0}$ converge vers L dans $\mathcal{L}(E, F)$. \square

Une application classique du résultat précédent est donnée par le :

Théorème 3.11. *Soit $(E, \|\cdot\|_E)$ un espace de Banach et $L \in \mathcal{L}(E, E)$ une application linéaire continue. On suppose que*

$$\|L\|_{\mathcal{L}(E, E)} < 1.$$

Alors, $I_E - L$ est inversible (i.e. c'est une application bijective et son inverse est une application linéaire continue) et

$$(I_E - L)^{-1} = \sum_{n \geq 0} L^n.$$

Ici I_E est l'application identité de E et L^n est définie par récurrence par $L^{n+1} = L \circ L^n$ et $L^0 = I_E$.

DÉMONSTRATION. Pour tout $n \in \mathbf{N}$, on note

$$V_n := \sum_{i=0}^n L^i.$$

Clairement, pour tout $m > n$, on peut écrire

$$\begin{aligned} \|V_m - V_n\|_{\mathcal{L}(E, E)} &= \left\| \sum_{i=n+1}^m L^i \right\|_{\mathcal{L}(E, E)} \\ &\leq \sum_{i=n+1}^m \|L^i\|_{\mathcal{L}(E, E)} \\ &\leq \sum_{i=n+1}^m \|L\|_{\mathcal{L}(E, E)}^i \\ &\leq \frac{1}{1 - \|L\|_{\mathcal{L}(E, E)}} \|L\|_{\mathcal{L}(E, E)}^{n+1}. \end{aligned}$$

Le membre de droite tendant vers 0 quand n tend vers l'infini, on en déduit que la suite $(V_n)_{n \geq 0}$ est une suite de Cauchy dans $(\mathcal{L}(E, E), \|\cdot\|_{\mathcal{L}(E, E)})$, qui est un espace de Banach. Donc, cette suite converge vers une limite notée V . En utilisant les mêmes arguments, on montre que

$$\|V(x)\|_E \leq \frac{1}{1 - \|L\|_{\mathcal{L}(E, E)}} \|x\|_E,$$

pour tout $x \in E$, ce qui permet de conclure que l'application V , qui est clairement linéaire, est continue.

Maintenant, pour tout $n \in \mathbf{N}$

$$(I_E - L) \circ V_n = V_n \circ (I_E - L) = I_E - L^{n+1}.$$

En particulier, nous obtenons l'estimation

$$\|(I_E - L) \circ V_n - I_E\|_{\mathcal{L}(E, E)} \leq \|L\|_{\mathcal{L}(E, E)}^{n+1},$$

qui, par passage à la limite quand n tend vers l'infini, nous permet de conclure que

$$V \circ (I_E - L) = (I_E - L) \circ V = I_E.$$

ce qui montre que V est bien l'inverse de $I_E - L$. □

Plus généralement, si

$$f(x) := \sum_{n \geq 0} a_n x^n,$$

est une série entière de rayon de convergence $R > 0$ et si l'on se donne une application $L \in \mathcal{L}(E, E)$ telle que $\|L\|_{\mathcal{L}(E, E)} < R$, on vérifie comme ci-dessus que

$$\sum_{n \geq 0} a_n L^n,$$

définit une application linéaire continue de E dans lui même.

7. Topologie de la droite achevée, limite inf et limite sup

Rappelons que la droite numérique achevée est définie par $\bar{\mathbf{R}} = \mathbf{R} \cup \{\pm\infty\}$. On peut munir $\bar{\mathbf{R}}$ d'une distance qui nous assure qu'une suite réelle $(x_n)_{n \geq 0}$ tend vers $\pm\infty$ au sens de la distance de $\bar{\mathbf{R}}$ si et seulement si $\lim_{n \rightarrow +\infty} x_n = \pm\infty$ au sens usuel.

Une telle distance est par exemple donnée par

$$d(x, y) := |f(x) - f(y)|,$$

où

$$f(x) := \frac{x}{1 + |x|},$$

et où, par convention,

$$f(\pm\infty) = \pm 1.$$

Essayons de comprendre quelle est la topologie de $(\bar{\mathbf{R}}, d)$. Pour cela, remarquons que, pour tout intervalle $[a, b] \subset \mathbf{R}$, le théorème des accroissements finis nous assure que l'on peut trouver deux constantes $C, \bar{C} > 0$ telles que

$$\forall x, y \in [a, b], \quad C|x - y| \leq \left| \frac{x}{1 + |x|} - \frac{y}{1 + |y|} \right| \leq \bar{C}|x - y|.$$

Cette inégalité permet de démontrer aisément que, sur \mathbf{R} , la topologie associée à d est la même que la topologie associée à la distance usuelle. Autrement dit, si $U \subset \mathbf{R}$, alors U est un ouvert (un fermé) de $(\bar{\mathbf{R}}, d)$ si et seulement si U est un ouvert (un fermé) de $(\mathbf{R}, |\cdot|)$. Maintenant, pour tout $a \in \mathbf{R}$, on vérifie que $[-\infty, a[$ et $]a, +\infty]$ sont des ouverts de $(\bar{\mathbf{R}}, d)$, de même, $[-\infty, a]$ et $[a, +\infty]$ sont des fermés de $(\bar{\mathbf{R}}, d)$. Nous invitons le lecteur à déterminer $B(+\infty, r)$, la boule ouvert de centre $+\infty$ et de rayon $r > 0$, dans $(\bar{\mathbf{R}}, d)$.

Nous avons le :

Lemme 3.7. *L'espace métrique $(\bar{\mathbf{R}}, d)$ est un espace métrique compact et, par conséquent, $(\bar{\mathbf{R}}, d)$ est également un espace métrique complet.*

DÉMONSTRATION. Remarquons que, par construction de d , la fonction

$$h : x \in [-1, 1] \mapsto \frac{x}{1 - |x|} \in \bar{\mathbf{R}},$$

prolongée par continuité par $h(\pm 1) = \pm\infty$, est continue de $[-1, 1]$ muni de la distance usuelle dans $(\bar{\mathbf{R}}, d)$. On en déduit que $(\bar{\mathbf{R}}, d)$ est un compact comme image du compact $([-1, 1], |\cdot|)$ par l'application continue h . \square

Ce résultat nous assure que toute suite de $\bar{\mathbf{R}}$ admet au moins une valeur d'adhérence dans $\bar{\mathbf{R}}$ (pour la distance d). L'ensemble des valeurs d'adhérences d'une suite $(x_n)_{n \geq 0}$ de $\bar{\mathbf{R}}$ étant fermé, on conclut qu'une suite de \mathbf{R} admet une plus grande valeur d'adhérence notée $\bar{\lim}_{n \rightarrow +\infty} x_n$, c'est la *limite supérieure de la suite* $(x_n)_{n \geq 0}$ (respectivement une plus petite valeur d'adhérence notée $\underline{\lim}_{n \rightarrow +\infty} x_n$, c'est la *limite inférieure de la suite* $(x_n)_{n \geq 0}$).

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

Proposition 3.7. *Soit $(x_n)_{n \geq 0}$ une suite de $\bar{\mathbf{R}}$. Alors :*

(i) *l'inégalité*

$$\liminf_{n \rightarrow +\infty} x_n \leq \overline{\lim}_{n \rightarrow +\infty} x_n,$$

est une égalité si et seulement si la suite $(x_n)_{n \geq 0}$ converge dans $\bar{\mathbf{R}}$. Dans ce cas

$$\lim_{n \rightarrow +\infty} x_n = \liminf_{n \rightarrow +\infty} x_n = \overline{\lim}_{n \rightarrow +\infty} x_n;$$

(ii) *la suite réelle $(x_n)_{n \geq 0}$ converge dans \mathbf{R} vers $x \in \mathbf{R}$ si et seulement si*

$$\liminf_{n \rightarrow +\infty} x_n = \overline{\lim}_{n \rightarrow +\infty} x_n = x;$$

(iii) *les suites $(\inf_{k \geq n} x_k)_{n \geq 0}$, $(\sup_{k \geq n} x_k)_{n \geq 0}$ sont monotones de limites respectives*

$$\liminf_{n \rightarrow +\infty} x_n \quad \text{et} \quad \overline{\lim}_{n \rightarrow +\infty} x_n.$$

Le point (iii) de cette Proposition permet de définir la limite supérieure (inférieure) d'une suite de \mathbf{R} par les formules

$$\liminf_{n \rightarrow +\infty} x_n = \lim_{n \rightarrow \infty} \left(\inf_{k \geq n} x_k \right) \quad \text{et} \quad \overline{\lim}_{n \rightarrow +\infty} x_n = \lim_{n \rightarrow \infty} \left(\sup_{k \geq n} x_k \right).$$

CHAPITRE 4

Théorèmes fondamentaux de l'analyse fonctionnelle

1. Théorème de Baire

Le Théorème de Baire est un des théorèmes importants en analyse fonctionnelle. Son utilisation permet d'obtenir des résultats d'existence assez surprenants.

Théorème 4.1 (Théorème de Baire). *Soit (X, d) un espace métrique complet et $(F_n)_{n \geq 0}$ une suite de fermés de X d'intérieurs vides i.e. $\overset{\circ}{F}_n = \emptyset$. Alors $\bigcup_{n \geq 0} F_n$ est d'intérieur vide.*

Par passage aux complémentaires, on obtient immédiatement le théorème équivalent :

Théorème 4.2 (Théorème de Baire). *Soit (X, d) un espace métrique complet et $(U_n)_{n \geq 0}$ une suite d'ouverts de X qui sont denses dans X i.e. $\overline{U}_n = X$. Alors $\bigcap_{n \geq 0} U_n$ est dense dans X .*

Remarquons que, dans le premier énoncé $\bigcup_{n \geq 0} F_n$ n'est pas forcément un fermé et que, dans le deuxième énoncé, $\bigcap_{n \geq 0} U_n$ n'est pas forcément un ouvert.

Il y a bien entendu équivalence entre ces deux énoncés étant donné que

$$\overline{X - Y} = X - \overset{\circ}{Y} \quad \text{et} \quad X - \overline{Y} = \overset{\circ}{X - Y}.$$

Donc,

$$\begin{aligned} \overline{\bigcap_{n \geq 0} U_n} = X & \Leftrightarrow X - \overline{\bigcap_{n \geq 0} U_n} = \emptyset \\ & \Leftrightarrow X - \overset{\circ}{\bigcap_{n \geq 0} U_n} = \emptyset \\ & \Leftrightarrow \overset{\circ}{\bigcup_{n \geq 0} (X - U_n)} = \emptyset. \end{aligned}$$

DÉMONSTRATION. Soit $(U_n)_{n \geq 0}$ une suite d'ouverts denses dans X et U un ouvert non vide de X . L'ouvert U_0 est dense dans X et U est un ouvert non vide, donc, il existe $(x_0, r_0) \in X \times]0, +\infty[$ tel que

$$B_f(x_0, r_0) \subset B(x_0, 2r_0) \subset U \cap U_0.$$

On reprend maintenant le raisonnement ci-dessus. L'ouvert U_1 est un ouvert dense et $B(x_0, r_0) \cap U_1$ est un ouvert non vide de X , il existe donc $(x_1, r_1) \in X \times]0, +\infty[$ tel que

$$B_f(x_1, r_1) \subset B(x_1, 2r_1) \subset B(x_0, r_0) \cap U_1 \subset U \cap U_0 \cap U_1.$$

De plus, quitte à diminuer r_1 si c'est nécessaire, on peut toujours supposer que $r_1 \leq r_0/2$. On construit ainsi par récurrence une suite $((x_n, r_n))_{n \geq 0}$ d'éléments de $X \times]0, +\infty[$ telle que

$$B_f(x_n, r_n) \subset B(x_n, 2r_n) \subset B(x_{n-1}, r_{n-1}) \cap U_n \subset U \cap (U_0 \cap \dots \cap U_n),$$

et $r_n \leq r_{n-1}/2$. On vérifie que

$$\forall m \geq n \quad d(x_n, x_m) \leq r_n \leq \frac{r_0}{2^n}.$$

Ce qui prouve que la suite $(x_n)_{n \geq 0}$ est une suite de Cauchy dans X , qui est un espace complet. Cette suite converge donc vers une limite que l'on note x_∞ .

Remarquons que, pour tous $m \geq n$, on a $x_m \in B_f(x_n, r_n)$, donc $x_\infty \in B_f(x_n, r_n)$. On en déduit que, pour tout $n \geq 0$,

$$x_\infty \in U \cap \left(\bigcap_{i=0}^n U_i \right).$$

Donc

$$U \cap \left(\bigcap_{n \geq 0} U_n \right) \neq \emptyset,$$

ce qui termine la démonstration. □

Donnons un exemple d'application de ce résultat.

Exemple 4.1. Soit $(U_n)_{n \geq 0}$ une suite d'ouverts denses dans $(\mathbf{R}, | \cdot |)$. Montrons que

$$U := \bigcap_{n \geq 0} U_n,$$

n'est pas un ensemble dénombrable. Raisonnons par l'absurde et supposons que U est dénombrable infini (le cas où U est fini se traite de manière identique). Dans ce cas, il existe une application bijective ϕ de \mathbf{N} dans U . On vérifie que $V_n := U_n - \{\phi(n)\}$ est encore un ouvert dense de \mathbf{R} et l'on applique le Théorème de Baire qui nous assure que $\bigcap_{n \geq 0} V_n$ est dense dans \mathbf{R} . Mais

$$\bigcap_{n \geq 0} V_n = \left(\bigcap_{n \geq 0} U_n \right) - U = \emptyset,$$

ce qui constitue la contradiction recherchée.

Comme nous le verrons dans les exercices, le Théorème de Baire permet de démontrer l'existence d'une infinité de solutions à un problème, sans pour autant donner une méthode pour exhiber ne serait-ce qu'une solution !

2. Théorème de Banach-Steinhaus

Le Théorème de Banach-Steinhaus est une conséquence du Théorème de Baire, c'est un résultat qui nous sera utile dans l'étude des espaces de Hilbert.

Théorème 4.3 (Théorème de Banach-Steinhaus). *Soit $(E, \| \cdot \|_E)$ un espace de Banach et $(F, \| \cdot \|_F)$ un espace vectoriel normé. Soit $(T_i)_{i \in I}$ une famille quelconque d'applications linéaires, continues de E dans F . On suppose que*

$$\forall x \in E, \quad \exists M_x > 0, \quad \text{tel que} \quad (\forall i \in I, \quad \|T_i(x)\|_F \leq M_x).$$

Alors, il existe une constante $M > 0$ telle que, pour tout $x \in E$ et pour tout $i \in I$,

$$\|T_i(x)\|_F \leq M \|x\|_E.$$

DÉMONSTRATION. Pour tout $n \geq 0$, on note

$$A_n := \{x \in E : \forall i \in I, \|T_i(x)\|_F \leq n\}.$$

Par hypothèse, les applications T_i sont continues donc A_n est un fermé de E , comme intersection de fermés (images réciproques de fermés par des applications continues). De plus, pour tout $x \in E$, $x \in A_n$ dès que $n \geq M_x$. En particulier

$$\bigcup_{n \geq 0} A_n = E.$$

Montrons qu'il existe $n \in \mathbf{N}$ pour lequel l'intérieur de A_n n'est pas vide, i.e. $\overset{\circ}{A}_n \neq \emptyset$. On raisonne par l'absurde. Supposons que, pour tout $n \in \mathbf{N}$, $\overset{\circ}{A}_n = \emptyset$. Dans ce cas, le Théorème de Baire, nous assure que $E = \emptyset$, ce qui constitue une contradiction.

Il existe donc $n \in \mathbf{N}$ tel que $\overset{\circ}{A}_n \neq \emptyset$. En particulier, il existe $(x_0, r_0) \in E \times]0, \infty[$ tel que $B(x_0, r_0) \subset A_n$. Alors, pour tout $x \in B(0, 1)$, on a

$$\|T_i(x_0 + rx)\|_F \leq n.$$

pour tout $i \in I$. En utilisant la linéarité de T_i , on conclut que

$$\|T_i(x)\|_F \leq \frac{1}{r}(n + \|T_i(x_0)\|_F) \leq \frac{1}{r}(n + M_{x_0}).$$

Ce qui termine la démonstration. \square

Remarquons que l'inégalité obtenue dans ce résultat s'exprime aussi sous la forme

$$\|T_i\|_{\mathcal{L}(E, E)} \leq M,$$

pour tout $i \in I$.

3. Théorème de point fixe de Banach

Soit T une application d'un espace métrique (X, d) dans un espace métrique (X', d') .

Définition 4.1. On dit que T est contractante s'il existe $k \in [0, 1[$ tel que, pour tous $x, y \in X$,

$$d'(T(x), T(y)) \leq k d(x, y),$$

(autrement dit si T est k -lipschitzienne de rapport $k \in [0, 1[$).

Une application contractante est, par définition, lipschitzienne. C'est donc une application continue!

Le Théorème de point fixe de Banach est à l'origine de nombreux résultats d'existence de solutions en mathématiques et en mathématiques appliquées.

Théorème 4.4 (Théorème de point fixe de Banach). Soit (X, d) un espace métrique complet non-vide et $T : X \rightarrow X$ une application contractante. Alors, T possède un unique point fixe dans X (i.e. il existe un unique $x \in X$ tel que $T(x) = x$).

DÉMONSTRATION. Soit $x_0 \in X$. Montrons que la suite $(x_n)_{n \geq 0}$ définie par récurrence par $x_{n+1} = T(x_n)$ pour tout $n \geq 0$, est une suite de Cauchy dans (X, d) . Ceci impliquera que cette suite est convergente puis, que sa limite est le point fixe cherché.

Pour tous $n \geq m$, on peut estimer

$$d(x_n, x_m) = d(T^m(x_{n-m}), T^m(x_0)) \leq k^m d(x_{n-m}, x_0).$$

En particulier, l'inégalité triangulaire implique que

$$d(x_n, x_0) \leq \sum_{j=0}^{n-1} d(x_{j+1}, x_j) \leq \sum_{j=0}^{n-1} k^j d(x_1, x_0) \leq \frac{1-k^n}{1-k} d(x_1, x_0) \leq \frac{1}{1-k} d(x_1, x_0).$$

Donc, pour tous $n \geq m$,

$$d(x_n, x_m) \leq \frac{k^{n-m}}{1-k} d(x_1, x_0),$$

ce qui prouve que la suite $(x_n)_{n \geq 0}$ est bien une suite de Cauchy. Étant donné que X est complet, cette suite converge. On note $x_\infty \in X$ sa limite. Par continuité de l'application T on a

$$T(x_\infty) = T\left(\lim_{n \rightarrow +\infty} x_n\right) = \lim_{n \rightarrow +\infty} T(x_n) = \lim_{n \rightarrow +\infty} x_{n+1} = x_\infty.$$

Donc, x_∞ est un point fixe de T .

L'unicité du point fixe résulte immédiatement de la propriété de contraction vérifiée par T : si x, y sont deux points fixes de T on a $d(x, y) = d(T(x), T(y)) \leq k d(x, y)$ et, puisque $k \in [0, 1[$, on a nécessairement $d(x, y) = 0$, i.e. $x = y$. \square

Le Théorème de point fixe de Banach est un résultat important qui trouve de nombreuses applications, notamment en théorie des systèmes dynamiques, en analyse non linéaire, ... Comme nous le verrons ultérieurement, il peut être utilisé pour démontrer l'existence de solutions pour des problèmes d'évolution.

Ce théorème fait partie des « théorèmes de point fixe » qui ont été développés et utilisés en mathématiques et qui ont eux aussi trouvé de nombreuses applications. Mentionnons par exemple le *Théorème de point fixe de Brouwer*

Théorème 4.5 (Théorème de Brouwer). *Soit K un convexe compact non vide d'un espace vectoriel normé de dimension finie et f une application continue de K dans K . Alors f admet un point fixe dans K .*

Nous invitons le lecteur à vérifier, qu'en dimension $N = 1$, ce résultat est une conséquence du théorème des valeurs intermédiaires. Ce résultat se généralise également en dimension infinie :

Théorème 4.6 (Théorème de Schauder). *Soit K un convexe compact non vide d'un espace de Banach et f une application continue de K dans K . Alors f admet un point fixe dans K .*

Ce dernier théorème est notamment utilisé pour démontrer l'existence de solutions pour certaines équations aux dérivées partielles non linéaires. On remarque que, dans les théorèmes de Brouwer et Schauder, l'unicité du point fixe n'est plus garantie.

4. Critères de densité dans $\mathcal{C}(X; \mathbf{K})$

Dans ce paragraphe, nous donnons une illustration des notions introduites et des résultats démontrés. On suppose que (X, d) est un espace métrique compact et l'on s'intéresse à l'espace $\mathcal{C}(X; \mathbf{K})$ des fonctions continues sur X à valeurs dans \mathbf{K} , où $\mathbf{K} = \mathbf{R}$ ou \mathbf{C} . On munit $\mathcal{C}(X; \mathbf{K})$ de la distance associée à la norme

$$\|f\|_\infty := \sup_{x \in X} |f(x)|.$$

Nous avons la :

Définition 4.2. On dit qu'un sous-ensemble $H \subset \mathcal{C}(X; \mathbf{K})$ est séparant si, pour tous $x \neq y \in X$, il existe $f \in H$ telle que $f(x) \neq f(y)$.

Voici un critère qui permet d'assurer la densité d'un sous-ensemble de $\mathcal{C}(X; \mathbf{R})$, espace des fonctions continues sur X à valeurs réelles.

Théorème 4.7 (Théorème de Stone-Weierstrass). *On suppose que (X, d) est un espace compact. Soit H une sous-algèbre¹ de $\mathcal{C}(X; \mathbf{R})$ qui contient les fonctions constantes, est séparante. Alors H est dense dans $\mathcal{C}(X; \mathbf{R})$, muni de la norme de la convergence uniforme.*

DÉMONSTRATION. Nous aurons besoin du résultat suivant :

Lemme 4.1. *Il existe une suite de polynômes $P_n \in \mathbf{R}[X]$ tels que la suite de fonctions $(P_n)_{n \geq 0}$ converge uniformément vers $x \mapsto |x|$ sur $[-1, 1]$.*

DÉMONSTRATION. On note $P_0 \equiv 0$ et l'on définit pour tout $n \geq 0$

$$P_{n+1}(x) = P_n(x) + \frac{1}{2} (x^2 - P_n^2(x)).$$

Une récurrence permet de voir que

$$0 \leq P_n(x) \leq P_{n+1}(x) \leq |x|,$$

pour tout $x \in [-1, 1]$. En particulier, pour tout $x \in [-1, 1]$, la suite $(P_n(x))_{n \geq 0}$ est une suite croissante et majorée de \mathbf{R} , donc elle converge vers une limite notée $P(x)$. Bien entendu, par passage à la limite quand n tend vers l'infini, on trouve que $0 \leq P(x) \leq |x|$ et que $P(x) = P(x) + \frac{1}{2} (x^2 - P(x)^2)$, donc $P(x) = |x|$. Ceci nous permet de conclure que la suite de fonctions $(P_n)_{n \geq 0}$ converge simplement vers la fonction $x \mapsto |x|$ sur $[-1, 1]$. Le Théorème de Dini nous permet ensuite d'affirmer que la suite de fonctions $(P_n)_{n \geq 0}$ converge uniformément vers $x \mapsto |x|$ sur $[-1, 1]$. \square

On remarque que, H étant une sous-algèbre, \overline{H} , l'adhérence de H dans $\mathcal{C}(X; \mathbf{R})$, est également une sous-algèbre (le vérifier). Si $f \in \overline{H}$, on note

$$f_n := \|f\|_\infty P_n \left(\frac{f}{\|f\|_\infty} \right),$$

où P_n est la suite de polynômes définie dans le lemme ci-dessus. Par construction $f_n \in \overline{H}$ (car \overline{H} est une sous-algèbre). Pour tout $\varepsilon > 0$, il existe $n \in \mathbf{N}$ tel que $|P_n(x) - |x|| \leq \varepsilon$ pour tout $x \in [-1, 1]$. De plus, pour tout $x \in X$, on vérifie que

$$|f_n(x) - |f(x)|| = \|f\|_\infty \left| \frac{|f(x)|}{\|f\|_\infty} - P_n \left(\frac{f(x)}{\|f\|_\infty} \right) \right|.$$

En particulier, $\|f - f_n\|_\infty \leq \varepsilon$. Conclusion, $|f|$ est limite uniforme d'une suite d'éléments de \overline{H} , donc $|f| \in \overline{H}$.

Quelques remarques élémentaires :

$$\min(f, g) = \frac{1}{2} (f + g - |f - g|) \quad \text{et} \quad \max(f, g) = \frac{1}{2} (f + g + |f - g|).$$

En particulier, pour toutes $f, g \in \mathcal{C}(X; \mathbf{R})$, on a

$$\min(f, g) \quad \text{et} \quad \max(f, g) \in \mathcal{C}(X; \mathbf{R}).$$

1. On dit que $H \subset \mathcal{C}(X; \mathbf{K})$ est une sous-algèbre de $\mathcal{C}(X; \mathbf{K})$ si un produit de deux éléments de H est un élément de H et si une combinaison linéaire à coefficients dans \mathbf{K} , de deux éléments de H est un élément de H .

De même

$$f^+ := \max(f, 0) \in \mathcal{C}(X; \mathbf{R}) \quad \text{et} \quad f^- := -\min(f, 0) \in \mathcal{C}(X; \mathbf{R}).$$

Pour terminer remarquons que $f = f^+ - f^-$ et que $|f| = f^+ + f^-$. On en déduit que, pour toutes $f, g \in \overline{H}$, $\min(f, g)$ et $\max(f, g)$ appartiennent à \overline{H} .

Montrons maintenant que, sous les hypothèses du théorème, pour tous $x \neq x' \in X$ et pour tous $\alpha, \alpha' \in \mathbf{R}$, il existe $f \in H$ telle que $f(x) = \alpha$ et $f(x') = \alpha'$. Pour ceci, remarquons que, H étant séparante, il existe $h \in H$ telle que $h(x) \neq h(x')$. Il suffit alors de définir

$$\tilde{h} = \frac{1}{h(x') - h(x)} ((\alpha' - \alpha)h + \alpha h(x') - \alpha' h(x)),$$

qui est une fonction qui appartient à H et qui répond à la question.

Nous pouvons maintenant terminer la démonstration du théorème. Soit $f \in \mathcal{C}(X; \mathbf{R})$ et $x \in X$ fixés. D'après ce que nous venons de voir, pour tout $y \in X$, on peut trouver une fonction $f_y \in H$ telle que $f_y(x) = f(x)$ et $f_y(y) = f(y)$. On note

$$O_y := \{x' \in X : f_y(x') > f(x') - \varepsilon\},$$

où $\varepsilon > 0$ est fixé. Clairement, O_y est un ouvert non vide de X qui contient x . De plus, $X = \bigcup_{y \neq x} O_y$. En utilisant la compacité de X , on peut extraire de ce recouvrement, un sous-recouvrement fini

$$X = \bigcup_{j=1}^n O_{y_j},$$

où chaque $y_j \neq x$. Afin d'alléger les notations, on notera $f_j := f_{y_j}$. Par définition $f_j(x) = f(x)$ et $f_j(y_j) = f(y_j)$. De plus, pour tout $x' \in X$, il existe $j \in \{1, \dots, n\}$ tel que $f_j(x') > f(x') - \varepsilon$. On note

$$g_x := \max_{j=1, \dots, n} f_j.$$

Clairement $g_x \in \overline{H}$ et, par construction, $g_x(x) = f(x)$ et, pour tout $x' \in X$ on vérifie que $g_x(x') > f(x') - \varepsilon$. On note

$$U_x := \{x' \in X : g_x(x') < f(x') + \varepsilon\}.$$

Comme précédemment U_x est un ouvert non vide et $X = \bigcup_{x \in X} U_x$. On peut donc recouvrir X à l'aide d'un nombre fini de U_x ,

$$X = \bigcup_{i=1}^m U_{x_i}.$$

On note $g_i := g_{x_i}$. Par définition $g_i(x_i) = f(x_i)$ et, pour tout $x' \in X$, il existe $i \in \{1, \dots, m\}$ tel que $g_i(x') < f(x') + \varepsilon$. Enfin, on définit

$$g := \min_{i=1, \dots, m} g_i,$$

qui, par construction, appartient à \overline{H} . De plus $f(x') - \varepsilon < g(x') < f(x') + \varepsilon$ pour tout $x' \in X$. Autrement dit, il existe $g \in \overline{H}$ tel que $\|f - g\|_\infty \leq \varepsilon$, ce qui termine la démonstration. \square

Exemple 4.2. On suppose que (X, d) est un espace compact. Définissons H comme étant le sous-espace de $\mathcal{C}(X; \mathbf{R})$ des *fonctions lipschitziennes* sur (X, d) . C'est-à-dire l'ensemble des fonctions $f \in \mathcal{C}(X; \mathbf{R})$ pour lesquelles, il existe une constante $k > 0$ telle que

$$|f(x) - f(y)| \leq k d(x, y).$$

L'ensemble H est clairement une sous-algèbre de $\mathcal{C}(X; \mathbf{R})$. Montrons que H est séparante. Soient $x \neq y \in X$, considérons la fonction f définie par $f(x') = d(x, x')$. Clairement $f(x) \neq f(y)$. De plus f est lipschitzienne car on a toujours l'inégalité

$$|d(x, x') - d(x, y')| \leq d(x', y').$$

Bien entendu, H contient les fonctions constantes. Le Théorème de Stone-Weierstrass nous assure que H est dense dans $\mathcal{C}(X; \mathbf{R})$, muni de la norme de la convergence uniforme.

Exemple 4.3. Soient $a < b$. Définissons H comme étant l'ensemble des restrictions à $[a, b]$ des fonctions polynômes à coefficients dans \mathbf{R} et $X = [a, b] \subset \mathbf{R}$. Les hypothèses du Théorème de Stone-Weierstrass sont vérifiées donc H est dense dans $\mathcal{C}([a, b]; \mathbf{R})$. Autrement dit : Toute fonction continue sur $[a, b]$ est limite uniforme d'une suite de polynômes.

Plus généralement, si $(\lambda_n)_{n \geq 1}$ est une suite croissante de réels positifs, distincts, le Théorème de Müntz-Szász (1914-1916) permet d'affirmer que les combinaisons linéaires des fonctions $x \mapsto 1$ et $x \mapsto x^{\lambda_n}$, pour tout $n \geq 1$ sont denses dans $\mathcal{C}([a, b]; \mathbf{R})$ (on suppose bien entendu que $b > a > 0$) si et seulement si

$$\sum_{n \geq 0} \frac{1}{\lambda_n} = +\infty.$$

Exemple 4.4. Soient (X, d) et (X', d') deux espaces métriques compacts. Définissons sur $X \times Y$ la topologie induite par la distance produit

$$d_p((x_1, y_1), (x_2, y_2)) := \max(d(x_1, x_2), d'(y_1, y_2)).$$

On note

$$\mathcal{C}(X; \mathbf{R}) \otimes \mathcal{C}(Y; \mathbf{R}) := \left\{ (x, y) \mapsto \sum_{\text{finie}} f_i(x) g_i(y) : f_i \in \mathcal{C}(X; \mathbf{R}), \quad g_i \in \mathcal{C}(Y; \mathbf{R}) \right\}.$$

On vérifie que $\mathcal{C}(X; \mathbf{R}) \otimes \mathcal{C}(Y; \mathbf{R})$ est une sous-algèbre de $\mathcal{C}(X \times Y; \mathbf{R})$ qui contient les constantes. Vérifions que $\mathcal{C}(X; \mathbf{R}) \otimes \mathcal{C}(Y; \mathbf{R})$ est séparante. Soient $(x_1, y_1) \neq (x_2, y_2) \in X \times Y$. Choisissons la fonction f définie sur $X \times Y$ par

$$f(x, y) = d(x, x_1) + d'(y, y_1).$$

Alors $f(x_1, y_1) = 0$ et $f(x_2, y_2) \neq 0$, ce qui montre que $\mathcal{C}(X; \mathbf{R}) \otimes \mathcal{C}(Y; \mathbf{R})$ est séparante. Conclusion, $\mathcal{C}(X; \mathbf{R}) \otimes \mathcal{C}(Y; \mathbf{R})$ est dense dans $\mathcal{C}(X \times Y; \mathbf{R})$.

En particulier, toute fonction continue sur $[0, 1] \times [0, 1]$ est limite uniforme sur $[0, 1] \times [0, 1]$ de sommes finies de fonctions de la forme $(x, y) \mapsto f(x)g(y)$, où f et g sont continues sur $[0, 1]$.

Une question naturelle est de savoir si les résultats précédents s'étendent aux fonctions à valeurs dans \mathbf{C} . Avant de répondre à cette question, commençons par étudier l'exemple où $X := \{z \in \mathbf{C} : |z| \leq 1\}$ est le disque unité fermé de \mathbf{C} , et H est l'ensemble des fonctions polynômes à coefficients dans \mathbf{C} . On vérifie que H est bien une sous-algèbre de $\mathcal{C}(X; \mathbf{C})$, qui est séparante et qui contient les fonctions constantes. Pourtant, H n'est pas dense dans $\mathcal{C}(X; \mathbf{C})$! En effet, pour tout $h \in H$, on vérifie que

$$\int_0^{2\pi} e^{i\theta} h(e^{i\theta}) d\theta = 0,$$

en particulier

$$\int_0^{2\pi} e^{i\theta} (e^{-i\theta} - h(e^{i\theta})) d\theta = 2\pi,$$

pour tout $h \in H$. On ne pourra donc pas approcher la fonction $z \mapsto \bar{z}$ uniformément sur X par des éléments de H .

Afin de généraliser le Théorème de Stone-Weierstrass au cas des fonctions à valeurs complexes, on introduit la :

Définition 4.3. *On dit que qu'un sous-ensemble $H \subset \mathcal{C}(X; \mathbf{C})$ est auto-conjugué si, $\forall f \in H$, la fonction \bar{f} , fonction complexe conjuguée de f , appartient à H .*

Voici maintenant la version complexe du Théorème de Stone-Weierstrass :

Théorème 4.8 (Théorème de Stone-Weierstrass). *On suppose que (X, d) est un espace compact. Soit H une sous-algèbre de $\mathcal{C}(X; \mathbf{C})$ qui contient les fonctions constantes, est séparante et auto-conjuguée. Alors H est dense dans $\mathcal{C}(X; \mathbf{C})$, muni de la norme de la convergence uniforme.*

DÉMONSTRATION. Définissons

$$H_{\mathbf{R}} = \{f \in H : \forall x \in X, f(x) \in \mathbf{R}\},$$

qui est une sous-algèbre de $\mathcal{C}(X; \mathbf{R})$ qui contient les fonctions constantes. Vérifions qu'elle est séparante. Soient $x \neq y$, on sait qu'il existe $f \in H$ telle que $f(x) \neq f(y)$. Quitte à échanger x et y on peut supposer que $f(x) \neq 0$. Définissons alors la fonction $g \in H$ par

$$g(z) = \frac{f(z) - f(y)}{f(x) - f(y)}.$$

pour $z \in X$, puis,

$$h = \frac{1}{2}(g + \bar{g}) \in H_{\mathbf{R}}$$

On vérifie que $h(x) = 1 \neq h(y) = 0$. Conclusion, $H_{\mathbf{R}}$ est séparante et, en vertu du Théorème 4.7 page 57, $H_{\mathbf{R}}$ est dense dans $\mathcal{C}(X; \mathbf{R})$. On montre de même que

$$H_I = \{f \in H : \forall x \in X, f(x) \in i\mathbf{R}\},$$

est dense dans $i\mathcal{C}(X; \mathbf{R})$. Enfin, $H = H_{\mathbf{R}} \oplus H_I$ est dense dans $\mathcal{C}(X; \mathbf{C})$. \square

On note $\mathcal{C}(S^1; \mathbf{C})$, l'espace des fonctions continues, 2π -périodiques, à valeurs dans \mathbf{C} . Remarquons que l'on peut identifier cet espace à l'espace des fonctions continues qui sont définies sur S^1 , le cercle unité de \mathbf{C} . En particulier, nous avons le :

Théorème 4.9. *Le \mathbf{C} -espace vectoriel engendré par les fonctions $x \mapsto e^{inx}$, pour $n \in \mathbf{Z}$ est dense dans $\mathcal{C}(S^1; \mathbf{C})$, muni de la norme de la convergence uniforme.*

5. Théorème d'Ascoli

Le Théorème de Riesz (Théorème 3.8 page 47) nous enseigne que, dans un espace vectoriel normé de dimension infinie, la boule unité fermée n'est jamais compacte.

Exemple 4.5. On voit facilement que la boule unité fermée de $(\mathcal{C}([0, 1]; \mathbf{R}), \| \cdot \|_{\infty})$ n'est pas compacte en considérant la suite de fonctions $(f_n)_{n \geq 0}$ définie par

$$f_n(x) = x^n.$$

En effet, la seule limite possible est la fonction qui vaut 0 sur $[0, 1[$ et 1 quand $x = 1$, mais cette fonction n'est pas continue !

Le Théorème d'Ascoli nous donne un critère simple qui permet de garantir qu'un sous-ensemble de $(\mathcal{C}([0, 1]; \mathbf{R}), \|\cdot\|_\infty)$ est compact. Nous rencontrerons dans la suite du cours d'autres résultats qui visent à palier le défaut de compacité des sous-ensembles fermés bornés des espaces vectoriels de dimension infinie.

Définition 4.4. Soit (X, d) un espace métrique compact et $\mathcal{F} \subset \mathcal{C}(X; \mathbf{K})$. On dit que la famille \mathcal{F} est équicontinue sur X si, pour tout $x \in X$ et pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que

$$\forall f \in \mathcal{F}, \quad \forall y \in X, \quad (d(x, y) < \delta \Rightarrow |f(y) - f(x)| < \varepsilon).$$

Donnons quelques exemples pour illustrer cette définition.

Exemple 4.6. On considère $X = [0, 1]$ muni de la distance usuelle et

$$\mathcal{F} := \left\{ f \in \mathcal{C}^1([0, 1]; \mathbf{R}) : \sup_{t \in [0, 1]} |f'(t)| \leq 1 \right\}.$$

Si $f \in \mathcal{F}$, on peut majorer

$$|f(y) - f(x)| \leq \left(\sup_{t \in [0, 1]} |f'(t)| \right) |y - x| \leq |y - x|.$$

On montre alors facilement que \mathcal{F} est équicontinue sur $[0, 1]$ (prendre $\delta = \varepsilon$).

Plus généralement, on a :

Exemple 4.7. Soit (X, d) un espace métrique compact. Fixons $k > 0$ et considérons

$$\mathcal{F} := \{f \in \mathcal{C}(X; \mathbf{K}) : |f(x) - f(y)| \leq k d(x, y)\}$$

l'ensemble des fonctions k -Lipschitziennes définies sur X à valeurs dans \mathbf{K} . Cette famille est équicontinue sur X .

Exemple 4.8. On considère $X = [0, 1]$ muni de la distance usuelle et

$$\mathcal{F} := \left\{ f \in \mathcal{C}^1([0, 1]; \mathbf{R}) : \int_0^1 |f'(t)|^2 dt \leq 1 \right\}.$$

Remarquons que, si $f \in \mathcal{C}^1([0, 1]; \mathbf{R})$ et si $x < y$, on peut écrire

$$f(y) - f(x) = \int_x^y f'(t) dt.$$

En utilisant l'inégalité de Cauchy-Schwarz, on obtient l'inégalité

$$|f(y) - f(x)| \leq \left(\int_x^y |f'(t)|^2 dt \right)^{\frac{1}{2}} |y - x|^{\frac{1}{2}} \leq |y - x|^{\frac{1}{2}},$$

pourvu que $f \in \mathcal{F}$. Grace à cette inégalité, on vérifie immédiatement que la famille \mathcal{F} est équicontinue sur $[0, 1]$ (prendre $\delta = \varepsilon^2$).

Exemple 4.9. On considère dans $\mathcal{C}([0, 1]; \mathbf{R})$ la famille

$$\mathcal{F} := \{x \mapsto e^{-nx} : n \geq 0\}.$$

On vérifie que la famille \mathcal{F} n'est pas équicontinue sur $[0, 1]$. Pour le voir il suffit de remarquer que pour tout $x > 0$, on peut toujours trouver un entier $n \in \mathbf{N}$ tel que

$$|1 - e^{-nx}| \geq \frac{1}{2}.$$

Donc, si l'on choisit $\varepsilon = 1/2$, pour tout $\delta > 0$, on peut trouver $x \in [0, \delta]$ et $n \in \mathbf{N}$ tels que

$$|x| < \delta \quad \text{et} \quad |1 - e^{-nx}| \geq 1/2.$$

Exemple 4.10. On considère toujours l'espace $\mathcal{C}([0, 1]; \mathbf{R})$, mais cette fois définissons

$$\mathcal{F} := \{x \mapsto \sin(nx) : n \geq 1\}.$$

On vérifie que la famille \mathcal{F} n'est pas équicontinue sur $[0, 1]$.

Venons en maintenant à l'énoncé et la démonstration du Théorème d'Ascoli que, pour simplifier, nous avons restreint aux fonctions à valeurs dans \mathbf{K} .

Théorème 4.10 (Théorème d'Ascoli). *On suppose que (X, d) est un espace métrique compact, que $\mathcal{F} \subset \mathcal{C}(X; \mathbf{K})$ est une famille équicontinue sur X et que pour tout $x \in X$, l'ensemble $\{f(x) : f \in \mathcal{F}\}$ est borné. Alors, de toute suite d'éléments de \mathcal{F} on peut extraire une sous-suite qui converge dans $(\mathcal{C}(X; \mathbf{K}), \|\cdot\|_\infty)$.*

DÉMONSTRATION. Exploitions pour commencer l'équicontinuité de \mathcal{F} . Pour tout $x \in X$ et pour tout $\varepsilon > 0$ il existe $\delta_x > 0$ (qui dépend de x) tel que

$$\forall y \in X, \quad \forall f \in \mathcal{F}, \quad (d(x, y) < \delta_x \Rightarrow |f(x) - f(y)| < \varepsilon).$$

Par hypothèse, l'espace (X, d) est supposé compact et

$$X = \bigcup_{x \in X} B(x, \delta_x/2).$$

On peut donc extraire de ce recouvrement de X par des ouverts, un sous-recouvrement fini. Il existe donc $x_1, \dots, x_n \in X$ tels que

$$(4.1) \quad X = \bigcup_{i=1}^n B(x_i, \delta_{x_i}/2).$$

On note

$$\delta := \min_{i=1, \dots, n} \delta_{x_i}/4.$$

On remarque, pour tout $x, y \in X$ si $d(x, y) < \delta$ alors il existe $i \in \{1, \dots, n\}$ tel que $x \in B(x_i, \delta_{x_i}/2)$ en vertu de (4.1). De plus, $y \in B(x_i, \delta_{x_i})$ car

$$d(y, x_i) \leq d(y, x) + d(x, x_i) \leq \frac{\delta_{x_i}}{2} + \delta < \delta_{x_i}.$$

On peut donc conclure que $|f(x) - f(y)| \leq \varepsilon$. Pour résumer, on a montré que, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que

$$(4.2) \quad \forall x, y \in X, \quad \forall f \in \mathcal{F}, \quad (d(x, y) < \delta \Rightarrow |f(x) - f(y)| < \varepsilon).$$

C'est ce que l'on appelle l'*uniforme équicontinuité* de \mathcal{F} sur X .

Ce résultat préliminaire étant acquis, on se donne maintenant $(f_n)_{n \geq 0}$, une suite d'éléments de \mathcal{F} , et l'on cherche à extraire de cette suite, une sous-suite qui converge dans $(\mathcal{C}(X; \mathbf{K}), \|\cdot\|_\infty)$. La démonstration consiste à utiliser une récurrence pour extraire une suite de sous-suites extraites de la suite $(f_n)_{n \geq 0}$, puis à utiliser une extraction de suite diagonale.

Expliquons comment fonctionne la récurrence. Pour initier la récurrence, i.e. $k = 0$, on utilise (4.2) avec $\varepsilon = \frac{1}{3}$ pour démontrer l'existence de $\delta_0 > 0$ tel que

$$\forall x, y \in X, \quad \forall f \in \mathcal{F}, \quad \left(d(x, y) < \delta_0 \quad \Rightarrow \quad |f(x) - f(y)| < \frac{1}{3} \right).$$

L'espace X étant compact, on peut recouvrir X à l'aide d'un nombre fini de boules ouvertes de rayon δ_0 . Il existe donc $x_{0,1}, \dots, x_{0,n_0} \in X$ tels que

$$X = \bigcup_{i=1}^{n_0} B(x_{0,i}, \delta_0).$$

Pour tout $i \in \{1, \dots, n_0\}$, remarquons que la suite $(f_n(x_{0,i}))_{n \geq 0}$ est, par hypothèse, bornée dans \mathbf{K} . On conclut donc qu'il existe une sous-suite de la suite $(f_n)_{n \geq 0}$, que l'on note $(f_{0,n})_{n \geq 0}$, telle que la suite $(f_{0,n}(x_{0,i}))_{n \geq 0}$ converge. Quitte à éliminer les premiers termes de la suite extraite, on peut même demander que

$$\forall n, m \in \mathbf{N}, \quad \forall i \in \{0, \dots, n_0\}, \quad |f_{0,n}(x_{0,i}) - f_{0,m}(x_{0,i})| < \frac{1}{3}.$$

Finalement, pour tout $x \in X$, il existe $i \in \{1, \dots, n_0\}$ tel que $x \in B(x_i^0, \delta_0)$, donc nous pouvons écrire

$$\begin{aligned} |f_{0,n}(x) - f_{0,m}(x)| &\leq |f_{0,n}(x) - f_{0,n}(x_{0,i})| + |f_{0,n}(x_{0,i}) - f_{0,m}(x_{0,i})| + |f_{0,m}(x_{0,i}) - f_{0,m}(x)| \\ &< 1. \end{aligned}$$

Expliquons maintenant l'étape générale de la récurrence. On suppose avoir construit la suite $(f_{k-1,n})_{n \geq 0}$ (extraite de la suite $(f_{k-2,n})_{n \geq 0}$ si $k \geq 2$) et pour laquelle

$$\sup_{x \in X} |f_{k-1,n}(x) - f_{k-1,m}(x)| < \frac{1}{2^{k-1}},$$

pour tous $n, m \geq 0$. Choisissons $\varepsilon = \frac{1}{2^k 3}$ et utilisons (4.2) pour conclure qu'il existe $\delta_k > 0$ tel que

$$\forall x, y \in X, \quad \forall f \in \mathcal{F}, \quad \left(d(x, y) < \delta_k \quad \Rightarrow \quad |f(x) - f(y)| < \frac{1}{2^k 3} \right).$$

L'espace X étant compact, on peut recouvrir X à l'aide d'un nombre fini de boules ouvertes de rayon δ_k . Il existe donc $x_{k,1}, \dots, x_{k,n_k} \in X$ tels que

$$X = \bigcup_{i=1}^{n_k} B(x_{k,i}, \delta_k).$$

Pour tout $i \in \{1, \dots, n_k\}$, remarquons que la suite $(f_{k-1,n}(x_{k,i}))_{n \geq 0}$ est par hypothèse bornée dans \mathbf{K} . On conclut donc qu'il existe $(f_{k,n})_{n \geq 0}$, une suite extraite de la suite $(f_{k-1,n})_{n \geq 0}$, telle que $(f_{k,n}(x_{k,i}))_{n \geq 0}$ converge. Quitte à éliminer les premiers termes de la suite extraite, on peut même demander que

$$\forall n, m \in \mathbf{N}, \quad \forall i \in \{0, \dots, n_k\}, \quad |f_{k,n}(x_{k,i}) - f_{k,m}(x_{k,i})| < \frac{1}{2^k 3}.$$

Comme dans l'étape initiale, pour tout $x \in X$, il existe $i \in \{1, \dots, n_k\}$ tel que $x \in B(x_{k,i}, \delta_k)$, donc nous pouvons écrire

$$|f_{k,n}(x) - f_{k,m}(x)| < \frac{1}{2^k}.$$

Considérons enfin la *suite diagonale* $(f_{n,n})_{n \geq 0}$. Par construction, on a

$$\forall m \geq n, \quad \sup_{x \in X} |f_{n,n}(x) - f_{m,m}(x)| < \frac{1}{2^n}.$$

Autrement dit, la suite $(f_{n,n})_{n \geq 0}$, qui est extraite de la suite $(f_n)_{n \geq 0}$, est une suite de Cauchy dans $(\mathcal{C}(X; \mathbf{K}), \|\cdot\|_\infty)$ qui est un espace de Banach, donc cette suite converge. Ceci termine la démonstration. \square

Donnons une application immédiate de ce résultat.

Exemple 4.11. Soit $(f_n)_{n \geq 0}$ une suite de $\mathcal{C}([0, 1]; \mathbf{K})$. On suppose qu'il existe une constante $k > 0$ telle que

$$|f_n(y) - f_n(x)| \leq k|y - x|,$$

pour tout $n \geq 0$ et pour tous $x, y \in [0, 1]$ et l'on suppose qu'il existe une constante $C > 0$ telle que

$$\forall n \geq 0, \quad \sup_{x \in [0, 1]} |f_n(x)| \leq C.$$

Alors, on peut extraire de la suite $(f_n)_{n \geq 0}$ une sous-suite qui converge dans $(\mathcal{C}([0, 1]; \mathbf{K}), \|\cdot\|_\infty)$, c'est-à-dire, une sous-suite qui converge uniformément sur $[0, 1]$.

6. Introduction à la résolution des équations différentielles ordinaires

Afin d'illustrer les résultats obtenus dans ce chapitre et les chapitres précédents, étudions le problème de la résolution des équations différentielles ordinaires non linéaires qui seront étudiées en plus de détail dans le cours de MAT 431 *Systèmes Dynamiques*. Donnons pour commencer deux exemples classiques d'équations différentielles ordinaires non linéaires qui ont fait et font toujours l'objet de recherches intenses.

Exemple 4.12. (Le problème à n corps). Il s'agit de comprendre l'évolution au cours du temps d'un système de n particules massives (de masses respectives m_1, \dots, m_n) qui interagissent sous l'effet des forces gravitationnelles. Les positions des particules seront notées x_1, \dots, x_n où chaque $x_i \in \mathbf{R}^3$ est une fonction du temps et leurs vitesses seront notées v_1, \dots, v_n .

Les équations qui régissent le mouvement, s'écrivent

$$\begin{cases} \frac{dx_i}{dt} = v_i \\ \frac{dv_i}{dt} = -G \sum_{k \neq i} m_k \frac{x_i - x_k}{|x_i - x_k|^3}, \end{cases}$$

pour $i = 1, \dots, n$, où G est la constante de gravitation. Il s'agit d'un système d'équations différentielles d'ordre 2. On suppose que l'on connaît la position initiale $x_1(0), \dots, x_n(0)$ ainsi que la vitesse initiale $v_1(0), \dots, v_n(0)$ de chaque particule (on suppose également qu'au temps initial les n particules sont distinctes).

Le problème consiste dans un premier temps à démontrer l'existence d'une solution pour des temps petits (c'est cette question que nous allons étudier) et ensuite à en étudier le

comportement pour des temps longs. Une autre question importante est celle de l'existence et de la stabilité des solutions périodiques pour ce système d'équations²

Exemple 4.13. Un autre exemple important pour ses applications en géométrie et en physique (notamment en théorie de la relativité générale) est celui de la détermination des géodésiques sur des variétés Riemanniennes ou Lorentziennes. En dimension N , le problème consiste à trouver des fonctions $s \mapsto x^\alpha(s)$ scalaires, pour $\alpha = 1, \dots, N$, solutions du système d'équations différentielles non linéaires d'ordre 2

$$\frac{d^2 x^\alpha}{ds^2} + \Gamma_{\beta\gamma}^\alpha \frac{dx^\beta}{ds} \frac{dx^\gamma}{ds} = 0,$$

pour $\alpha = 1, \dots, N$ (la sommation sur toutes les valeurs que peuvent prendre les indices β et γ est sous-entendue).

Les fonctions $\Gamma_{\beta\gamma}^\alpha$, pour $\alpha, \beta, \gamma \in \{1, \dots, N\}$, sont appelées *symboles de Christoffel*. Ce sont des fonctions des variables x_1, \dots, x_N , qui sont calculées à partir du tenseur métrique sur la variété, ce sont donc des données du problème. Nous renvoyons aux cours de Relativité Générale PHY 568 et MAT 568 pour plus de précisions.

Le problème consiste à démontrer l'existence et l'unicité d'une solution du système décrit ci-dessus pour s proche de 0, pour des positions initiales $x^1(0), \dots, x^N(0)$ et des vitesses initiales $\frac{dx^1}{ds}(0), \dots, \frac{dx^N}{ds}(0)$ fixées. Le système ci-dessus est un système d'équations différentielles d'ordre 2, mais on se ramène à un système d'équations différentielles d'ordre 1 en regardant les équations vérifiées par le vecteur d'état constitué des positions x^α et des vitesses $\frac{dx^\alpha}{ds}$.

Fixons $N \geq 1$. Sur \mathbf{R}^N , il sera commode de choisir la norme

$$\|x\| := \max_{i=1, \dots, N} |x_i|,$$

si $x = (x_1, \dots, x_N) \in \mathbf{R}^N$. Toutes les normes étant équivalentes sur \mathbf{R}^N , ce choix n'a pas beaucoup d'importance.

On se propose de résoudre l'équation différentielle non linéaire

$$(4.3) \quad \frac{dx}{dt}(t) = F(t, x(t)),$$

avec condition initiale $x(0) = x_0 \in \mathbf{R}^N$. Dans cette équation,

$$F : [-t_*, t_*] \times B_f(x_0, r_*) \rightarrow \mathbf{R}^N,$$

est une application continue, définie sur $[-t_*, t_*] \times B_f(x_0, r_*) \subset \mathbf{R} \times \mathbf{R}^N$ où $t_*, r_* > 0$. Remarquons que, la fonction F étant continue sur le compact $[-t_*, t_*] \times B_f(x_0, r_*)$, elle est bornée sur cet ensemble, on peut donc affirmer qu'il existe une constante $M > 0$ telle que

$$\|F(t, x)\| \leq M,$$

pour tout $(t, x) \in [-t_*, t_*] \times B_f(x_0, r_*)$.

2. Cette dernière question a fait couler beaucoup d'encre et a mobilisé de nombreux chercheurs dont font partie Joseph-Louis Lagrange et Henri Poincaré (X1873). En introduisant à cette occasion une multitude d'idées et de concepts nouveaux, Henri Poincaré a jeté les bases de la théorie moderne des systèmes dynamiques. La recherche de solutions périodiques fait toujours l'objet de travaux et, à ce sujet, le lecteur intéressé pourra consulter l'article d'Alain Chenciner et Richard Montgomery, *A remarkable periodic solution of the three-body problem in the case of equal masses*, <http://fr.arxiv.org/pdf/math/0011268v1>.

On fait maintenant l'hypothèse supplémentaire que F est lipschitzienne par rapport à la deuxième variable, c'est-à-dire qu'il existe un réel $k > 0$ tel que

$$\|F(t, y) - F(t, x)\| \leq k \|x - y\|,$$

pour tous $(t, x), (t, y) \in [-t_*, t_*] \times B_f(x_0, r_*)$.

Pour préciser le problème, on recherche une fonction $t \mapsto x(t)$ qui est définie et de classe \mathcal{C}^1 sur un intervalle $]t_0, t_0[$, où $t_0 \in]0, t_*]$, à valeurs dans \mathbf{R}^N , qui est solution de (4.3) et qui vérifie $x(0) = x_0$. Après intégration, on se ramène à la recherche d'une solution $t \mapsto x(t)$ de l'équation intégrale

$$(4.4) \quad x(t) = x_0 + \int_0^t F(s, x(s)) ds.$$

Choisissons $r_0 \in]0, r_*]$ et $t_0 \in]0, t_*]$ de telle sorte que

$$k t_0 < 1, \quad \text{et} \quad M t_0 \leq r_0.$$

On note $\mathcal{B}_f(0, r_0)$, la boule fermée centrée en 0 et de rayon r_0 dans l'espace $\mathcal{C}([-t_0, t_0]; \mathbf{R}^N)$, muni de la norme

$$\|x\|_\infty := \sup_{t \in [-t_0, t_0]} \|x(t)\|,$$

et l'on définit l'application Φ sur $\mathcal{B}_f(0, r_0)$ par

$$\Phi(y)(t) := \int_0^t F(s, x_0 + y(s)) ds,$$

pour $t \in [-t_0, t_0]$. Grâce aux hypothèses faites sur l'application F , on peut estimer

$$\|\Phi(y)\|_\infty \leq \sup_{t \in [-t_0, t_0]} \int_0^t \|F(s, x_0 + y(s))\| ds \leq t_0 M,$$

et l'on a également

$$\begin{aligned} \|\Phi(y) - \Phi(y')\|_\infty &\leq \sup_{t \in [-t_0, t_0]} \int_0^t \|F(s, x_0 + y(s)) - F(s, x_0 + y'(s))\| ds \\ &\leq k \sup_{t \in [-t_0, t_0]} \int_0^t \|y(s) - y'(s)\| ds \\ &\leq k t_0 \|y - y'\|_\infty, \end{aligned}$$

pour $y, y' \in \mathcal{B}_f(0, r_0)$. En particulier, le choix de t_0 permet de conclure que Φ est une application contractante de $\mathcal{B}_f(0, r_0)$ dans elle même.

L'espace $(\mathcal{C}([-t_0, t_0]; \mathbf{R}^N), \|\cdot\|_\infty)$ est un espace de Banach, donc $\mathcal{B}_f(0, r_0)$, muni de la distance induite par $\|\cdot\|_\infty$, est un espace métrique complet. Le Théorème de point fixe de Banach nous permet de conclure que l'application Φ admet un unique point fixe dans $\mathcal{B}_f(0, r_0)$. On obtient ainsi l'existence et l'unicité de $y \in \mathcal{B}_f(0, r_0)$ telle que

$$y(t) = \int_0^t F(s, x_0 + y(s)) ds,$$

pour $t \in [-t_0, t_0]$. Dans ce cas, $x(t) = x_0 + y(t)$ est une solution de (4.4) dont on vérifie qu'elle est bien de classe \mathcal{C}^1 sur $[-t_0, t_0]$ (on laisse le soin au lecteur de vérifier cette dernière assertion). On démontre ainsi le Théorème de Cauchy-Lipschitz, sur la résolution des équations différentielles ordinaires.

Remarque 4.1. Le lecteur pourra vérifier que, sous les hypothèses que nous avons faites, le temps maximal d'existence de la solution de (4.3) est minoré par $\min(t_*, r_*/M)$.

Supposons maintenant que l'application F est seulement continue sur $[-t_*, t_*] \times B_f(0, r_*)$ (on ne fait plus l'hypothèse que F est lipschitzienne par rapport à la deuxième variable), ce qui assure en particulier qu'il existe une constante $M > 0$ telle que

$$\|F(t, x)\| \leq M,$$

pour tout $(t, x) \in [-t_*, t_*] \times B_f(0, r_*)$.

Définissons $\mathcal{E} \subset \mathcal{C}([-t_0, t_0]; \mathbf{R}^N)$ comme étant l'ensemble des fonctions continues qui sont définies sur $[-t_0, t_0]$, à valeurs dans \mathbf{R}^N , qui sont M -Lipschitziennes et qui valent 0 quand $t = 0$. L'ensemble \mathcal{E} est clairement convexe et l'on vérifie que c'est un fermé de $(\mathcal{C}([-t_0, t_0]; \mathbf{R}^N), \|\cdot\|_\infty)$. Enfin, on montre que \mathcal{E} est un compact de $(\mathcal{C}([-t_0, t_0]; \mathbf{R}^N), \|\cdot\|_\infty)$ en appliquant le Théorème d'Ascoli.

On laisse le soin au lecteur de vérifier que, pourvu que $t_0 \in]0, t_*]$ soit choisi assez petit, l'application Φ est une application continue de \mathcal{E} dans lui-même et, cette fois-ci, le Théorème de point fixe de Schauder nous assure l'existence d'une solution de (4.4), donc de (4.3), dans le cas où F n'est que continue.

Remarque 4.2. On remarquera que, sous cette hypothèse plus faible (i.e. F seulement continue), on n'a pas nécessairement unicité de la solution. Par exemple, choisissons $\alpha \in]0, 1[$ et définissons $\beta := \frac{1}{1-\alpha}$. On considère alors l'équation différentielle

$$\frac{dx}{dt} = \beta^{1-\alpha} |x|^\alpha,$$

où la fonction x est à valeurs réelles, avec pour donnée initiale $x(0) = 0$. Cette équation admet pour solution la fonction $t \mapsto x(t) \equiv 0$ qui est définie pour tout $t \in \mathbf{R}$. Mais, on vérifie que la fonction $t \mapsto x(t) := |t|^{\beta-1} t$ est, elle aussi, une solution de l'équation ci-dessus pour tout $t \in \mathbf{R}$.

Dans le cas où la fonction F est seulement continue (mais n'est pas supposée lipschitzienne par rapport à la deuxième variable), le lecteur pourra trouver une autre démonstration du résultat d'existence d'une solution de (4.3) en utilisant dans un premier temps le Théorème de Stone-Weierstrass qui permet d'approcher la fonction F par une suite $(F_n)_{n \geq 0}$ de fonctions lipschitziennes. On peut alors appliquer le Théorème de Cauchy-Lipschitz pour démontrer l'existence (et l'unicité) d'une solution x_n de

$$x_n(t) = x_0 + \int_0^t F_n(s, x_n(s)) ds,$$

qui est définie sur un intervalle $[-t_n, t_n]$. Ensuite, on montre que le temps maximal d'existence de x_n est en fait minoré par une constante $t_0 > 0$ qui ne dépend pas de n . Enfin, on utilise une fois de plus le Théorème d'Ascoli pour extraire de la suite $(x_n)_{n \geq 0}$ une sous suite qui converge uniformément vers une fonction $t \mapsto x(t)$ solution de (4.4) sur un intervalle $[-t_0, t_0]$.

Construction de l'intégrale de Lebesgue

1. Motivation

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 que l'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 difficiles à mettre en œuvre 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 sur l'intégrale de Lebesgue est donc double :

- (i) définir l'intégrale pour une classe de fonctions plus générale que celle des fonctions continues par morceaux sur un intervalle de \mathbf{R} ;
- (ii) disposer de théorèmes d'intégration terme à terme pour des suites ou les séries de fonctions, pour lesquels il n'est pas nécessaire de vérifier que la fonction limite est continue par morceaux.

Pour le point (ii), on aimerait disposer, pour les suites de fonctions, d'un énoncé du type suivant : *Soit $(U_n)_{n \geq 0}$ suite croissante de fonctions « intégrables » sur un ouvert non vide $\Omega \subset \mathbf{R}^N$ alors*

$$\Omega \ni x \mapsto \lim_{n \rightarrow +\infty} U_n(x),$$

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

$$\int_{\Omega} \lim_{n \rightarrow +\infty} U_n(x) dx = \lim_{n \rightarrow +\infty} \int_{\Omega} U_n(x) dx,$$

l'égalité ayant lieu dans $] -\infty, +\infty]$.

Voici un énoncé équivalent pour les séries de fonctions : *Soit $(u_n)_{n \geq 0}$ suite de fonctions positives « intégrables » sur un ouvert non vide $\Omega \subset \mathbf{R}^N$; alors*

$$\Omega \ni x \mapsto \sum_{n \geq 0} u_n,$$

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

$$\int_{\Omega} \left(\sum_{n \geq 0} u_n \right) dx = \sum_{n \geq 0} \int_{\Omega} u_n(x) dx,$$

l'égalité ayant lieu dans $[0, +\infty]$.

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

$$U_n := \sum_{k=0}^n u_k.$$

Les points (i) et (ii) sont intimement liés. Pour disposer d'un énoncé d'intégration 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 5.1 (Fonction indicatrice de \mathbf{Q}). Rappelons que \mathbf{Q} est dénombrable infini. Il existe donc une bijection

$$\mathbf{N} \ni n \mapsto r_n \in \mathbf{Q},$$

qui énumère tous les rationnels. 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 et continue par morceaux et

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

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

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

Rappelons que, pour toute fonction $f : [0, 1] \rightarrow \mathbf{R}$, les sommes de Riemann de la fonction f sont définies par

$$\mathcal{R}_n(f) := \frac{1}{n} \sum_{k=1}^n f\left(\frac{k}{n}\right),$$

pour tout $n \geq 1$.

Observons cependant que la somme de Riemann

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

On ne peut donc aboutir à une notion d'intégrale vérifiant la propriété (ii) ci-dessus en définissant la classe des fonctions intégrables sur $[0, 1]$ comme l'ensemble des fonctions $f : [0, 1] \rightarrow \mathbf{R}$ telles que la suite des sommes de Riemann

$$\mathcal{R}_n(f) := \frac{1}{n} \sum_{k=1}^n f\left(\frac{k}{n}\right),$$

converge quand $n \rightarrow +\infty$ et l'intégrale d'une telle fonction f comme la limite de la suite $\mathcal{R}_n(f)$.

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.

Dans toute la suite de ce chapitre, Ω désignera un ouvert non vide de \mathbf{R}^N , où N est un entier strictement positif et, afin d'alléger les notations, on notera simplement $\mathcal{C}(\Omega)$ à la place de $\mathcal{C}(\Omega; \mathbf{R})$.

2. 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 prolongement¹. 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 :

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

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

Dans toute la suite, on notera $\mathcal{C}_c(\Omega)$ 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*

$$\mathcal{R}_j(f) := \frac{1}{2^{jN}} \sum_{\substack{k \in \mathbf{Z}^N \\ 2^{-j}k \in \Omega}} f\left(\frac{k}{2^j}\right).$$

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

Lemme 5.1. *Soit $f \in \mathcal{C}_c(\mathbf{R}^N)$. Alors la suite de réels indexée par $j \in \mathbf{N}$*

$$\mathcal{R}_j(f) := \frac{1}{2^{jN}} \sum_{k \in \mathbf{Z}^N} f\left(\frac{k}{2^j}\right),$$

converge pour $j \rightarrow +\infty$.

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.

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

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

1. Il existe plusieurs manières de construire l'intégrale de Lebesgue; celle qui est adoptée ici remonte à F. Riesz, *Sur l'intégrale de Lebesgue*, Acta Mathematica **42** (1920), 1–15, et P. J. Daniell, *A general form of integral*, Annals of Mathematics **19** (1918), 279–294. Cette méthode 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 [11]. Nous avons suivi ici la présentation concise de M. Willem, en la complétant par de nombreux exemples.

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

$$I_j^-(f) = \frac{1}{2^{jN}} \sum_{k \in \mathbf{Z}^N} \min_{x \in C_{j,k}} f(x) \quad \text{et} \quad I_j^+(f) = \frac{1}{2^{jN}} \sum_{k \in \mathbf{Z}^N} \max_{x \in C_{j,k}} f(x).$$

Bien entendu

$$I_j^-(F) \leq \mathcal{R}_j(f) \leq I_j^+(f),$$

pour tout $j \in \mathbf{N}$. Montrons maintenant que les suites $(I_j^-(f))_{j \geq 1}$ et $(I_j^+(f))_{j \geq 1}$ sont adjacentes.

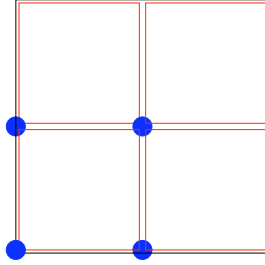


FIGURE 5.1. Le carré noir étant $C_{j,k} \subset \mathbf{R}^2$, les quatre points bleus sont les points de la forme $2^{-j-1}l$ pour $l \in S_k$, et les quatre carrés rouges sont les $C_{j+1,l}$ lorsque l décrit S_k .

Dans un premier temps, montrons que la suite $(I_j^-(f))_{j \geq 1}$ est croissante et $(I_j^+(f))_{j \geq 1}$ est décroissante. C'est là que l'on utilise de manière essentielle le fait que nous avons utilisé une décomposition dyadique de l'espace : les sommets des cubes de la forme $C_{j,k}$, pour $k \in \mathbf{Z}^N$, sont des sommets des cubes de la forme $C_{j+1,l}$, pour $l \in \mathbf{Z}^N$.

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

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

(voir la Figure 5.1). 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

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

De même, on montre que $I_j^+(f) \geq I_{j+1}^+(f)$. Enfin, on montre que $I_j^+(f) - I_j^-(f)$ tend vers 0 quand j tend vers l'infini. Étant donné que la fonction f est à support compact, il existe

$L \in \mathbf{N} - \{0\}$ tel que $f(x) = 0$ si $x \notin [-L, L]^N$. Donc ²

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

où $\#A$ désigne le cardinal de l'ensemble A . La fonction f étant continue et à support compact sur \mathbf{R}^N , elle est uniformément continue (d'après le Corollaire 3.2 page 39), on en déduit que

$$\lim_{j \rightarrow +\infty} \sup_{|x-y| \leq 2^{-j} \sqrt{N}} |f(x) - f(y)| = 0.$$

Conclusion, les suites $(I_j^-(f))_{j \geq 0}$ et $(I_j^+(f))_{j \geq 0}$ sont donc adjacentes, et convergent lorsque $j \rightarrow +\infty$ vers une limite commune. Étant donné que l'on a, pour tout $j \in \mathbf{N}$,

$$I_j^-(f) \leq \mathcal{R}_j(f) \leq I_j^+(f),$$

on conclut que la suite $(\mathcal{R}_j(f))_{j \geq 0}$ converge. \square

Étant donnée $f \in \mathcal{C}_c(\Omega)$, on déduit du Lemme 5.1 page 71 appliqué à la fonction $\tilde{f} \in \mathcal{C}_c(\mathbf{R}^N)$ définie par

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

que la suite $\mathcal{R}_j(f)$ est convergente lorsque $j \rightarrow +\infty$.

Définition 5.1 (Intégrale des fonctions de $\mathcal{C}_c(\Omega)$). *Pour tout $f \in \mathcal{C}_c(\Omega)$, on pose*

$$\int_{\Omega} f(x) dx := \lim_{j \rightarrow +\infty} \mathcal{R}_j(f).$$

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

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

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

Cette notion d'intégrale, que nous venons de définir sur $\mathcal{C}_c(\Omega)$, coïncide donc avec la notion d'intégrale multiple étudiée dans les classes préparatoires.

La seule propriété de l'intégrale sur $\mathcal{C}_c(\Omega)$ que nous allons utiliser pour définir l'intégrale de Lebesgue est la suivante : l'application

$$\mathcal{C}_c(\Omega) \ni f \longmapsto \int_{\Omega} f(x)dx \in \mathbf{R},$$

est une forme linéaire sur le \mathbf{R} -espace vectoriel $\mathcal{C}_c(\Omega)$, qui est **positive**, au sens où :

$$\int_{\Omega} f(x)dx \geq 0,$$

pour tout fonction $f \in \mathcal{C}_c(\Omega)$ telle que $f(x) \geq 0$ pour tout $x \in \Omega$. Voici un premier résultat d'intégration terme à terme de suites convergeant simplement, qui, compte-tenu de la positivité de l'intégrale, est un cas particulier de la Proposition 3.5 page 43.

Lemme 5.2. *Soit $(f_n)_{n \geq 0}$ suite décroissante de fonctions de $\mathcal{C}_c(\Omega)$, qui converge simplement vers 0 sur Ω . Alors*

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

3. Définition de l'intégrale de Lebesgue

À partir de l'intégrale usuelle définie ci-dessus, nous allons construire l'intégrale de Lebesgue en deux étapes :

- (a) on *impose* l'interversion limite \leftrightarrow intégrale pour les suites croissantes de fonctions de $\mathcal{C}_c(\Omega)$ dont la suite des intégrales est majorée ;
- (b) on *prolonge par linéarité* cette première extension.

Cette construction fournit un \mathbf{R} -espace vectoriel contenant strictement $\mathcal{C}_c(\Omega)$, et une *forme linéaire* qui est positive et définie sur cet espace vectoriel, prolongeant l'intégrale usuelle sur $\mathcal{C}_c(\Omega)$, et vérifiant l'énoncé (ii) de l'introduction (voir la section 1 page 69).

3.1. Étape 1 : Définition de la classe de fonctions $\mathcal{L}^+(\Omega)$. On commence par la :

Définition 5.2 (Suites de Levi). *Une suite $(f_n)_{n \geq 0}$ de fonctions de $\mathcal{C}_c(\Omega)$ est dite de Levi si c'est une suite croissante sur Ω (i.e. $f_{n+1}(x) \geq f_n(x)$, pour tout $x \in \Omega$) et si elle vérifie*

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

Attention : on prendra garde de ne pas confondre une suite croissante de fonctions et une suite de fonction croissantes (qui par ailleurs n'a aucun sens pour les fonctions de plus d'une variable) !

Pour que l'intégrale que l'on cherche à construire vérifie l'énoncé (ii) de l'introduction (voir la section 1 page 69), cet énoncé doit déjà être vrai pour les suites de Levi. Ceci motive la définition suivante :

Définition 5.3 (Classe $\mathcal{L}^+(\Omega)$). *On définit $\mathcal{L}^+(\Omega)$ comme l'ensemble des fonctions $f : \Omega \rightarrow \mathbf{R} \cup \{+\infty\}$ pour lesquelles il existe une suite de Levi qui converge simplement vers f sur Ω .*

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

$$\left(\int_{\Omega} f_n(x) dx \right)_{n \geq 0},$$

est croissante et majorée, donc converge vers sa borne supérieure, et l'on peut donc prolonger la définition de l'intégrale aux fonctions qui appartiennent à la classe $\mathcal{L}^+(\Omega)$ en posant

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

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

Lemme 5.3. *Soient $(f_n)_{n \geq 0}$ et $(g_n)_{n \geq 0}$, deux suites de Levi définies sur Ω . On suppose que $\lim_{n \rightarrow +\infty} f_n(x) = \lim_{n \rightarrow +\infty} g_n(x)$ pour tout $x \in \Omega$, alors*

$$\lim_{n \rightarrow +\infty} \int_{\Omega} f_n(x) dx = \lim_{n \rightarrow +\infty} \int_{\Omega} g_n(x) dx.$$

Avant de donner la démonstration de ce lemme, rappelons les notions de parties positives et négatives d'un réel, qui ont déjà été utilisées dans un chapitre précédent et que nous utiliserons systématiquement par la suite.

Parties positives et négatives d'un réel. Pour tout $a \in \mathbf{R}$, on note

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

on a

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

Évidemment

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

Étant 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. On note $f(x) := \lim_{n \rightarrow +\infty} f_n(x)$ et $g(x) := \lim_{n \rightarrow +\infty} g_n(x)$, pour tout $x \in \Omega$. Pour $m \in \mathbf{N}$ fixé et pour tout $n \in \mathbf{N}$, on pose

$$h_n := (f_m - g_n)^+.$$

Évidemment $h_n \in \mathcal{C}_c(\Omega)$, car f_m et $g_n \in \mathcal{C}_c(\Omega)$, et $f_m - g_n \leq h_n$.

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

$$\lim_{n \rightarrow +\infty} h_n(x) = (f_m(x) - g(x))^+ \leq (f(x) - g(x))^+ = 0.$$

On déduit alors du Lemme 5.2 page ci-contre que

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

Mais puisque $f_m \leq g_n + h_n$, on conclut que

$$\int_{\Omega} f_m(x) dx \leq \lim_{n \rightarrow +\infty} \int_{\Omega} (g_n(x) + h_n(x)) dx = \lim_{n \rightarrow +\infty} \int_{\Omega} g_n(x) dx.$$

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

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

On conclut en échangeant les rôles des suites $(f_n)_{n \geq 0}$ et $(g_n)_{n \geq 0}$. \square

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

Lemme 5.4. *Soit $f : \Omega \rightarrow \mathbf{R} \cup \{+\infty\}$. Il y a équivalence entre les propriétés suivantes :*

- (a) *pour tout $\lambda \in \mathbf{R}$, l'ensemble $f^{-1}(] \lambda, +\infty])$ est un ouvert de Ω ;*
- (b) *pour tout $x \in \Omega$ et pour toute suite $(x_n)_{n \geq 0}$ de points de Ω qui converge vers x , on a*

$$f(x) \leq \liminf_{n \rightarrow +\infty} f(x_n).$$

DÉMONSTRATION. Montrons que (a) implique (b). En effet, soit $(x_n)_{n \geq 0}$ une suite de points de Ω qui converge vers $x \in \Omega$. Il existe donc une suite extraite $(x_{\varphi(n)})_{n \geq 0}$ telle que

$$\lambda := \lim_{n \rightarrow +\infty} f(x_{\varphi(n)}) = \liminf_{n \rightarrow +\infty} f(x_n).$$

Pour tout $\varepsilon > 0$, il existe $n_\varepsilon \geq 0$ tel que, pour tout $n \geq n_\varepsilon$

$$x_{\varphi(n)} \in f^{-1}(] -\infty, \lambda + \varepsilon]).$$

Or d'après la propriété (a), l'ensemble $f^{-1}(] -\infty, \lambda + \varepsilon])$ est un fermé de Ω (en tant que complémentaire de $f^{-1}(] \lambda + \varepsilon, +\infty])$ qui lui est un ouvert) ; comme de plus $\lim_{n \rightarrow +\infty} x_{\varphi(n)} = x$, on conclut que $x \in f^{-1}(] -\infty, \lambda + \varepsilon])$, c'est-à-dire que $f(x) \leq \lambda + \varepsilon$. Cette inégalité étant valable pour tout $\varepsilon > 0$, on conclut que $f(x) \leq \lambda$, qui est l'inégalité de la propriété (b).

Réciproquement, supposons que la propriété (b) est vérifiée, et soit $\lambda \in \mathbf{R}$. On cherche à montrer que $f^{-1}(] -\infty, \lambda])$ est fermé dans Ω . Soit $(x_n)_{n \geq 0}$ une suite de points de $f^{-1}(] -\infty, \lambda])$ qui converge vers $x \in \Omega$; il suffit de vérifier que l'on a $x \in f^{-1}(] -\infty, \lambda])$, c'est-à-dire que $f(x) \leq \lambda$. Or, par hypothèse, $f(x_n) \leq \lambda$ pour tout $n \geq 0$, on déduit de la propriété (b) que

$$f(x) \leq \liminf_{n \rightarrow +\infty} f(x_n) \leq \lambda,$$

ce qui établit la propriété (a). \square

Définition 5.4 (Fonction s.c.i.). *Une fonction $f : \Omega \rightarrow \mathbf{R} \cup \{+\infty\}$ est dite semi-continue inférieurement (s.c.i.) sur Ω si elle vérifie une des deux propriétés équivalentes du lemme ci-dessus.*

Exemple 5.2. La fonction f définie sur \mathbf{R} par $f(x) := 0$ pour tout $x \leq 0$ et $f(x) := 1$ pour tout $x > 0$ est s.c.i. sur \mathbf{R} . En revanche, $-f$ n'est pas s.c.i. sur \mathbf{R} .

Venons en aux propriétés des fonctions de $\mathcal{L}^+(\Omega)$.

Proposition 5.1. *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 de Levi 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. \square

Voici quelques exemples et contre-exemples de fonctions appartenant à la classe $\mathcal{L}^+(\Omega)$. Mais avant tout, rappelons la notation pour les fonctions indicatrices : étant donné un sous-ensemble $A \subset \mathbf{R}^N$, la fonction indicatrice de A dans \mathbf{R}^N est définie par

$$\mathbf{1}_A : x \mapsto \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{si } x \notin A. \end{cases}$$

Exemple 5.3. Soient $a < b$. On vérifie que $\mathbf{1}_{]a,b[} \in \mathcal{L}^+(\mathbf{R})$ et que

$$\int_{\mathbf{R}} \mathbf{1}_{]a,b[}(x) dx = b - a.$$

Par exemple, si $a = 0$ et $b = 1$, on peut définir pour tout $n \geq 1$, la fonction

$$g_n(x) := \begin{cases} 0 & \text{si } x \leq 0 \\ 2n x & \text{si } x \in [0, \frac{1}{2n}] \\ 1 & \text{si } x \in [\frac{1}{2n}, 1 - \frac{1}{2n}] \\ 2n(1 - x) & \text{si } x \in [1 - \frac{1}{2n}, 1] \\ 0 & \text{si } x \geq 1, \end{cases}$$

dont on vérifie que c'est une suite de Levi qui converge simplement vers $\mathbf{1}_{]0,1[}$.

Exemple 5.4. Pour tout $\alpha > 0$, la fonction

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

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

Exemple 5.5. La fonction $\mathbf{1}_{\mathbf{Q} \cap [0,1]}$ n'appartient pas à $\mathcal{L}^+(\mathbf{R})$ (vérifier qu'elle n'est pas s.c.i.).

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 $\mathbf{R} \cup \{+\infty\}$ En ce qui concerne les règles de calcul dans $[0, +\infty]$, les conventions sont les suivantes :

- (i) pour tout $z \in \mathbf{R} \cup \{+\infty\}$, $z + \infty = +\infty$;
- (ii) pour tout $\alpha > 0$, $\alpha \times (+\infty) = +\infty$;
- (iii) $0 \times (+\infty) = 0$.

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

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

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

Exemple 5.6. Si $a < b$ et si $f \in \mathcal{L}^+(\mathbf{R})$ est une fonction positive. On vérifie que $\mathbf{1}_{]a,b[} f \in \mathcal{L}^+(\mathbf{R})$.

En effet, $f \in \mathcal{L}^+(\mathbf{R})$ donc il existe une suite de Levi $f_n \in \mathcal{C}_c(\mathbf{R})$ qui converge simplement vers f . Comme nous avons supposé que $f \geq 0$, la suite $(f_n^+)_{n \geq 0}$ est aussi une suite de Levi qui converge simplement vers f . Si $(g_n)_{n \geq 0}$ est une suite de Levi qui converge simplement vers $\mathbf{1}_{]a,b[}$ (voir l'exemple 5.3 page précédente) alors, $(f_n g_n)_{n \geq 0}$ est une suite de Levi qui converge simplement vers $\mathbf{1}_{]a,b[} f$.

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

Proposition 5.2. On a :

(a) $\mathcal{C}_c(\Omega) \subset \mathcal{L}^+(\Omega)$;

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

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

(c) pour toutes $f, g \in \mathcal{L}^+(\Omega)$, on a $\max(f, g)$ et $\min(f, g) \in \mathcal{L}^+(\Omega)$ et si $f \leq g$ sur Ω , alors

$$\int_{\Omega} f(x) dx \leq \int_{\Omega} g(x) dx.$$

Tous ces énoncés sont triviaux, et leurs démonstrations sont laissées au lecteur.

Par construction de la classe $\mathcal{L}^+(\Omega)$, l'énoncé (ii) d'interversion intégrale \leftrightarrow limite de l'introduction (voir la section 1 page 69) est vérifié pour les suites de Levi. Mais ce n'est pas tout : ce même énoncé vaut encore pour les suites croissantes de fonctions de $\mathcal{L}^+(\Omega)$.

Proposition 5.3. Soit $(f_n)_{n \geq 0}$ suite croissante de fonctions de $\mathcal{L}^+(\Omega)$ telle que

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

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

$$f(x) := \lim_{n \rightarrow +\infty} f_n(x),$$

pour tout $x \in \Omega$. Alors $f \in \mathcal{L}^+(\Omega)$ et

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

La démonstration de cette proposition est laissée au lecteur, car nous allons établir bientôt un résultat un peu plus général (voir la Proposition 5.4 page 81 ci-dessous).

Terminons cette section par quelques remarques importantes. Dans la mesure où l'on a construit la classe de fonctions $\mathcal{L}^+(\Omega)$ et où l'on a défini sur celle-ci une notion d'intégrale qui prolonge l'intégrale usuelle définie initialement sur $\mathcal{C}_c(\Omega)$, dans la mesure où cette intégrale vérifie l'énoncé (ii) d'interversion intégrale \leftrightarrow limite de la section 1 page 69, on pourrait croire que l'on est arrivé à la généralisation voulue de l'intégrale usuelle. Pourtant, il n'en est rien ! D'une part, on a vu dans l'exemple 5.4 page précédente que la fonction $\mathbf{1}_{\mathbf{Q} \cap [0,1]}$ n'appartenant pas à la classe $\mathcal{L}^+(\Omega)$: le prolongement de l'intégrale à $\mathcal{L}^+(\Omega)$ ne suffit donc pas pour répondre complètement aux problèmes soulevés au début de ce chapitre dans la section 1 page 69. 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 $\lim_{x \rightarrow 0^+} -f(x) = -\infty$ de sorte que

$-f$ ne peut être minorée par une fonction appartenant à $\mathcal{C}_c([0, 1])$. 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 $\mathcal{C}_c(\Omega)$. Or, pour des raisons évidentes, il est indispensable que la nouvelle notion d'intégrale que l'on cherche à construire soit une opération linéaire, définie sur un espace vectoriel !

3.2. Ensembles 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 5.5. Soit $\mathcal{Z} \subset \Omega$. On dira que l'ensemble \mathcal{Z} est négligeable, s'il existe $f \in \mathcal{L}^+(\Omega)$ tel que $f(x) = +\infty$ pour tout $x \in \mathcal{Z}$.

Voici un premier exemple d'ensemble négligeable, nous en donnerons d'autres plus loin.

Exemple 5.7. Tout singleton de Ω est négligeable. Vérifions par exemple que $\{0\}$ est négligeable dans \mathbf{R} , la démonstration s'adapte facilement pour démontrer que $\{0\}$ est négligeable dans \mathbf{R}^N (exercice). Considérons la fonction définie par

$$f(x) := \sum_{n \geq 1} n(1 - n^3|x|)^+,$$

On vérifie que $f(x) = +\infty$ si et seulement si $x = 0$.

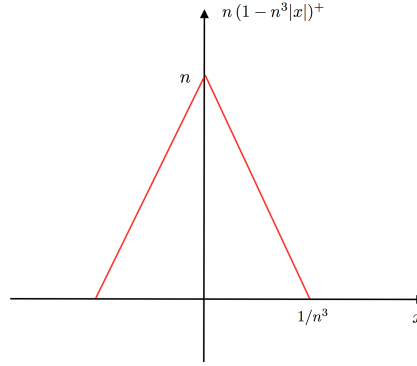


FIGURE 5.2. Graphe de la fonction $x \mapsto n(1 - n^3|x|)^+$.

Comme f est somme d'une série de fonctions positives appartenant à $\mathcal{C}_c(\mathbf{R})$ et que

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

(voir la Figure 5.2), on a

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

d'après la Proposition 5.3 page précédente. Donc, $\{0\}$ est négligeable.

Venons-en à la définitions de propriétés qui sont « vraies presque partout ».

Définition 5.6. Une propriété $\mathcal{P}(x)$, qui dépend du point $x \in \Omega$, est dite vraie presque partout sur Ω (ce que l'on abrège en \mathcal{P} p.p. sur Ω) ou encore vraie pour presque tout $x \in \Omega$ si

$$\{x \in \Omega : \mathcal{P}(x) \text{ est fausse}\},$$

est négligeable.

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

De même, la propriété $\mathcal{P}(x)$ peut être $\ll \lim_{n \rightarrow +\infty} f_n(x) = f(x) \gg$, 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 $\ll \lim_{n \rightarrow +\infty} f_n = f \text{ p.p. sur } \Omega \gg$.

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

D'abord, on peut généraliser le Lemme 5.3 page 75 en remplaçant l'hypothèse de la convergence simple sur Ω (en tout point de Ω) par la convergence p.p. sur Ω .

Lemme 5.5. Soient $(f_n)_{n \geq 0}$ et $(g_n)_{n \geq 0}$, deux suites de Levi définies sur Ω . On suppose que $\lim_{n \rightarrow +\infty} f_n(x) \leq \lim_{n \rightarrow +\infty} g_n(x)$ pour presque tout $x \in \Omega$, alors

$$\lim_{n \rightarrow +\infty} \int_{\Omega} f_n(x) dx \leq \lim_{n \rightarrow +\infty} \int_{\Omega} g_n(x) dx.$$

Autrement dit, si $f, g \in \mathcal{L}^+(\Omega)$, et si $f \leq g$ p.p. sur Ω , alors

$$\int_{\Omega} f(x) dx \leq \int_{\Omega} g(x) dx.$$

Remarquons qu'en échangeant les rôles des suites $(f_n)_{n \geq 0}$ et $(g_n)_{n \geq 0}$, on voit immédiatement que le Lemme 5.3 page 75 est un corollaire de ce résultat.

DÉMONSTRATION. Il s'agit d'une variante plus technique de la preuve du Lemme 5.3 page 75.

Pour tout $x \in \Omega$, notons $f(x) := \lim_{n \rightarrow +\infty} f_n(x)$ et $g(x) := \lim_{n \rightarrow +\infty} g_n(x)$. On fixe $m \in \mathbf{N}$. Pour tout $n \in \mathbf{N}$, définissons la suite

$$h_n := (f_m - g_n)^+.$$

On note \mathcal{Z} , l'ensemble des points de Ω tels que $\lim_{n \rightarrow +\infty} f_n(x) > \lim_{n \rightarrow +\infty} g_n(x)$. Par hypothèse \mathcal{Z} est négligeable. Il existe donc une fonction $k \in \mathcal{L}^+(\Omega)$ telle que $k(x) = +\infty$ si et seulement si $x \in \mathcal{Z}$. Quitte à changer k en k^+ , on peut supposer que $k \geq 0$ sur Ω . Sans perdre de généralité, on peut aussi supposer que

$$\int_{\Omega} k(x) dx = 1.$$

Soit $(k_n)_{n \geq 0}$ suite de Levi convergeant simplement vers k ; quitte à changer k_n en k_n^+ , on peut supposer que $k_n \geq 0$ sur Ω pour tout $n \geq 0$.

Soit $\varepsilon > 0$ fixé. Définissons alors, pour tout $n \in \mathbf{N}$, la fonction

$$l_n := (h_n - \varepsilon k_n)^+.$$

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 en déduit que la suite $(l_n)_{n \geq 0}$ est décroissante.

En utilisant le fait que $k_n(x) \rightarrow +\infty$ pour tous les points $x \in \Omega$ tels que $h_n(x)$ ne tend pas vers 0 lorsque $n \rightarrow +\infty$, on vérifie que la suite $(l_n)_{n \geq 0}$ converge simplement vers 0 sur Ω .

D'autre part, $l_n \in \mathcal{C}_c(\Omega)$ pour tout $n \in \mathbf{N}$. D'après le Lemme 5.2 page 74 (basé sur le Théorème de Dini),

$$\lim_{n \rightarrow +\infty} \int_{\Omega} l_n(x) dx = 0.$$

D'autre part,

$$f_m \leq h_n + g_n \leq l_n + \varepsilon k_n + g_n,$$

sur Ω , de sorte que

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

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

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

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 $m \rightarrow +\infty$. \square

Dans le même ordre d'idées, on peut aussi généraliser la Proposition 5.3 page 78 en y remplaçant l'hypothèse de *convergence simple (partout)* sur Ω par la *convergence p.p.* sur Ω .

Proposition 5.4 (Convergence monotone dans $\mathcal{L}^+(\Omega)$). *Soit $(f_k)_{k \geq 0}$ suite croissante de fonctions de $\mathcal{L}^+(\Omega)$ telle que*

$$\sup_{k \geq 0} \int_{\Omega} f_k(x) dx < +\infty.$$

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

$$f(x) := \lim_{k \rightarrow +\infty} f_k(x),$$

pour presque tout $x \in \Omega$ et

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

DÉMONSTRATION. Soit $\mathcal{Z} \subset \Omega$ un sous-ensemble négligeable tel que la suite $(f_n(x))_{n \geq 0}$ soit croissante pour tout $x \in \Omega - \mathcal{Z}$. D'autre part, pour tout $k \in \mathbf{N}$, soit $(f_{k,n})_{n \geq 0}$, une suite de Levi qui converge simplement vers f_k sur Ω .

Étape 1 : construction de la fonction f . Posons

$$F_n := \max(f_{0,n}, \dots, f_{n,n}),$$

pour tout $n \in \mathbf{N}$. Évidemment, $F_n \in \mathcal{C}_c(\Omega)$; d'autre part la suite $(F_n(x))_{n \geq 0}$ est croissante pour tout $x \in \Omega$ et

$$F_n(x) \leq \max(f_0(x), \dots, f_n(x)) = f_n(x),$$

pour tout $x \in \Omega - \mathcal{Z}$. D'après le Lemme 5.5 page précédente,

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

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

$$f(x) := \lim_{n \rightarrow +\infty} F_n(x),$$

pour tout $x \in \Omega$. Par construction $f \in \mathcal{L}^+(\Omega)$.

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

$$F_n(x) \leq f_n(x),$$

pour tout $x \in \Omega - \mathcal{Z}$; d'autre part, par construction, pour tout $n \geq k \geq 0$, on a

$$f_{k,n}(x) \leq F_n(x),$$

pour tout $x \in \Omega$. En passant à la limite dans ces deux inégalités pour $n \rightarrow +\infty$, on trouve que

$$f_k(x) \leq f(x) \leq \lim_{n \rightarrow +\infty} f_n(x),$$

pour tout $x \in \Omega - \mathcal{Z}$; puis en passant à la limite lorsque $k \rightarrow +\infty$ dans l'inégalité ci-dessus, on conclut que $\lim_{k \rightarrow +\infty} f_k(x) = f(x)$ pour tout $x \in \Omega - \mathcal{Z}$.

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

$$f_{k,n}(x) \leq F_n(x) \leq f_n(x),$$

pour tout $x \in \Omega - \mathcal{Z}$; de sorte que, d'après le Lemme 5.5 page 80

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

Passons à la limite lorsque $n \rightarrow +\infty$: comme les suites $(f_{k,n})_{n \geq 0}$ et $(F_n)_{n \geq 0}$ sont des suites de Levi sur Ω et que la suite $\left(\int_{\Omega} f_n(x) dx \right)_{n \geq 0}$ est croissante et bornée supérieurement, on trouve que

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

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

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

Proposition 5.5. *Toute réunion dénombrable d'ensembles négligeables est négligeable.*

DÉMONSTRATION. Soit $(\mathcal{Z}_k)_{k \geq 0}$, une suite de sous-ensembles négligeables de Ω . On note $\mathcal{Z} := \bigcup_{k \geq 0} \mathcal{Z}_k$. Pour chaque $k \in \mathbf{N}$, il existe $f_k \in \mathcal{L}^+(\Omega)$ telle que $f_k(x) = +\infty$ pour tout $x \in \mathcal{Z}_k$. Évidemment $f_k^+ \in \mathcal{L}^+(\Omega)$ pour tout $k \in \mathbf{N}$, d'après le point (c) dans la Proposition 5.2 page 78. Quitte à multiplier f_k^+ par une constante positive bien choisie, on peut supposer que

$$\int_{\Omega} f_k^+(x) dx \leq 2^{-k},$$

pour tout $k \geq 0$. Posons, pour tout $x \in \Omega$ et tout $n \in \mathbf{N}$,

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

Évidemment $F_n \in \mathcal{L}^+(\Omega)$ 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 $F(x) := \sum_{k \geq 0} f_k(x)^+$ définit une fonction $F \in \mathcal{L}^+(\Omega)$. De plus, si $x \in \mathcal{Z}$, alors $F(x) = +\infty$. Donc, nous pouvons conclure que \mathcal{Z} est négligeable. \square

Remarquons que dans ce résultat l'hypothèse de dénombrabilité est essentielle : une réunion non dénombrable 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 5.8. Toute partie dénombrable \mathcal{D} de \mathbf{R}^N est négligeable. En effet, $\mathcal{D} = \bigcup_{x \in \mathcal{D}} \{x\}$ est réunion dénombrable de singletons, qui sont négligeables (Exemple 5.7 page 79). D'après la Proposition 5.5 page précédente, \mathcal{D} est négligeable.

Exemple 5.9. L'ensemble des rationnels \mathbf{Q} est négligeable dans \mathbf{R} et, plus généralement, \mathbf{Q}^N est négligeable dans \mathbf{R}^N pour tout $N \geq 1$.

Exemple 5.10. Un ouvert non vide $\Omega \subset \mathbf{R}^N$ n'est pas négligeable dans \mathbf{R}^N .

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

$$k|f(x)| \leq \lim_{n \rightarrow +\infty} g_n(x) = +\infty,$$

pour tout $x \in \Omega$, on déduirait du Lemme 5.5 page 80 que

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

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

$$\int_{\Omega} |f(x)| dx = 0,$$

d'où $f(x) = 0$ pour tout $x \in \Omega$. Comme f est arbitraire dans $\mathcal{C}_c(\Omega)$, ceci entraînerait que $\mathcal{C}_c(\Omega) = \{0\}$, ce qui est impossible puisque $\Omega \neq \emptyset$.

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

Proposition 5.6. *Tout ouvert non vide de \mathbf{R}^N , $N \geq 1$, est non dénombrable.*

DÉMONSTRATION. Soit Ω ouvert non vide de \mathbf{R}^N . Si Ω était un ensemble dénombrable, il serait négligeable (voir l'exemple 5.8). Or c'est impossible d'après l'exemple 5.10. \square

On espère que la simplicité de cette démonstration, 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.

N.B. Comme nous l'avons signalé dans la section 2 page 71, la construction de l'intégrale de Lebesgue n'utilise en réalité que la positivité de la forme linéaire associée à l'intégrale usuelle sur $\mathcal{C}_c(\Omega)$. On pourrait donc effectuer la même construction en partant d'une autre forme linéaire positive sur $\mathcal{C}_c(\Omega)$ différente de l'intégrale usuelle : on aboutirait ainsi à un objet analogue à l'intégrale de Lebesgue, qui cette fois est appelé *intégrale par rapport à une mesure de Radon*. Il s'agit d'objets sur lesquels nous reviendrons plus loin dans le Chapitre 7. Lorsque l'on effectue ce type de construction, on aboutit également à une nouvelle notion d'ensemble négligeable qui est associée à l'intégrale par rapport à cette mesure de Radon. On prendra

garde que la propriété d'être une fonction intégrable ou d'être un sous-ensemble négligeable dépend de la forme linéaire positive choisie.

En réalité, dans tout le cours, à l'exception du Chapitre 7, il sera toujours question en pratique de l'intégrale de Lebesgue construite à partir de l'intégrale usuelle sur $\mathcal{C}_c(\Omega)$, de sorte que le lecteur peut sans inconvénient oublier cette distinction en première lecture et parler d'ensembles négligeables sans plus de précision.

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. Étape 2 : fonctions intégrables. 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 5.7. Une fonction f définie p.p. sur Ω et à valeurs dans \mathbf{R} est intégrable au sens de Lebesgue (ou sommable) s'il existe $g, h \in \mathcal{L}^+(\Omega)$ telles que $f = g - h$ p.p. sur Ω . 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 intégrables sur Ω à valeurs réelles.

Exemple 5.11. La fonction $\mathbf{1}_{\mathbf{Q}}$ appartient à $\mathcal{L}^1(\mathbf{R})$ puisque $\mathbf{1}_{\mathbf{Q}} = 0$ p.p. sur \mathbf{R} .

Vérifions que cette définition de l'intégrale d'une fonction 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

$$f = g_1 - h_1 = g_2 - h_2 \quad \text{p.p. sur } \Omega,$$

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

$$g_1 + h_2 = g_2 + h_1 \quad \text{p.p. sur } \Omega,$$

et comme $g_1 + h_2$ et $g_2 + h_1$ appartiennent à $\mathcal{L}^+(\Omega)$ d'après le point (b) de la Proposition 5.2 page 78, on déduit du Lemme 5.5 page 80 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.

Donnons maintenant plus de précisions sur la décomposition des fonctions intégrables en différence de fonctions de la classe \mathcal{L}^+ .

Lemme 5.6. Soit $f \in \mathcal{L}^1(\Omega)$. Pour tout $\varepsilon > 0$, il existe $u, v \in \mathcal{L}^+(\Omega)$ telles que

$$f = u - v \quad \text{p.p. sur } \Omega,$$

avec $v \geq 0$ sur Ω et

$$\int_{\Omega} v(x) dx \leq \varepsilon.$$

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

$$\lim_{n \rightarrow +\infty} \int_{\Omega} h_n(x) dx = \int_{\Omega} h(x) dx.$$

Il existe donc $m > 0$ tel que, pour tout $n \geq m$,

$$\int_{\Omega} (h(x) - h_n(x)) dx \leq \varepsilon.$$

Posons alors $u = g - h_m$ et $v = h - h_m$; évidemment $u \in \mathcal{L}^+(\Omega)$ et $v \geq 0$ sur Ω tandis que

$$\int_{\Omega} v(x) dx \leq \varepsilon.$$

Il reste à montrer que $v \in \mathcal{L}^+(\Omega)$. Pour cela, remarquons que la suite $(v_n)_{n \geq 0}$ définie par $v_n = h_n - h_m$ est une suite de Levi sur Ω puisqu'elle est croissante 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 $v_n \rightarrow v$ simplement sur Ω , on en déduit que $v \in \mathcal{L}^+(\Omega)$. Enfin $f = u - v$ p.p. sur Ω . \square

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 intégrables 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.

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

Soient f_1, \dots, f_N des fonctions définies p.p. sur Ω et soit Φ une fonction définie sur \mathbf{R}^N à valeurs dans \mathbf{R} . La fonction $x \mapsto \Phi(f_1(x), \dots, f_N(x))$ est définie p.p. sur Ω . En effet, par hypothèse, pour tout $i = 1, \dots, N$, il existe \mathcal{Z}_i négligeable dans Ω tel que la fonction f_i soit définie sur $\Omega - \mathcal{Z}_i$, de sorte que la fonction $x \mapsto \Phi(f_1(x), \dots, f_N(x))$ est définie sur $\Omega - \bigcup_{i=1}^N \mathcal{Z}_i$, c'est-à-dire p.p. sur Ω , puisque $\bigcup_{i=1}^N \mathcal{Z}_i$ est négligeable en tant que réunion finie d'ensembles négligeables. Ce raisonnement se généralise au cas où l'on considère une fonction définie à partir d'une famille dénombrable de fonction définies p.p. sur Ω .

Par exemple, on peut prendre $\Phi(f_1, f_2) = f_1 f_2$ ou $\Phi(f_1, f_2) = f_1 + f_2$. Ainsi les opérations élémentaires mettant en jeu une famille dénombrable de fonctions définies p.p. sur Ω fournissent une fonction définie p.p. sur Ω .

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

Proposition 5.7. *L'intégrale de Lebesgue vérifie les énoncés suivants :*

(a) *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*

$$\mathcal{L}^1(\Omega) \ni f \quad \mapsto \quad \int_{\Omega} f(x) dx \in \mathbf{R},$$

une forme \mathbf{R} -linéaire ;

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

(c) si $f, g \in \mathcal{L}^1(\Omega)$ et $f \leq g$ p.p. sur Ω , alors

$$\int_{\Omega} f(x) dx \leq \int_{\Omega} g(x) dx ;$$

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

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

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

DÉMONSTRATION. Les énoncés (a) et (b) sont triviaux, ainsi que l'énoncé (c), qui découle de la définition des fonctions intégrables et du Lemme 5.5 page 80.

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

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

Comme g et h appartiennent à $\mathcal{L}^+(\Omega)$, les fonctions $\max(g, h)$ et $\min(g, h)$ appartiennent à $\mathcal{L}^+(\Omega)$ (voir le point (c) dans la Proposition 5.2 page 78) ; et donc $|f| \in \mathcal{L}^1(\Omega)$. Puis $f \leq |f|$ p.p. sur Ω , de sorte que, d'après l'énoncé (c)

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

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 (e) découle du (d) en observant que

$$\max(f, g) = \frac{1}{2}(f + g + |f - g|), \quad \text{et} \quad \min(f, g) = \frac{1}{2}(f + g - |f - g|).$$

Ce qui termine la démonstration. \square

Remarque 5.1. On prendra bien garde au fait suivant : si une fonction f est intégrable, son module (ou sa valeur absolue, si f est à valeurs réelles) est aussi une fonction intégrable.

En particulier, dans les formules du type

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

le membre de gauche n'est pas une intégrale au sens de Lebesgue, et la fonction $x \mapsto \frac{\sin x}{x}$ n'est pas intégrable sur \mathbf{R}_+ , puisque

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

On rappelle que la formule

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

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

$$F(y) := \int_0^y \frac{\sin x}{x} dx$$

vérifie $\lim_{y \rightarrow +\infty} F(y) = \frac{\pi}{2}$. La limite de F en $+\infty$, est notée

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

ce qui introduit une confusion avec les notations

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

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

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

3.4. Fonctions intégrables et ensembles négligeables. Terminons cette présentation de l'intégrale de Lebesgue avec quelques remarques sur la relation qui existe entre les ensembles négligeables et les fonctions intégrables.

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

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

il suffit de décomposer ϕ sous la forme $\phi = g - g$ p.p. sur Ω .

Exemple 5.12. Comme \mathbf{Q} est négligeable dans \mathbf{R} , la fonction $\mathbf{1}_{\mathbf{Q}} = 0$ p.p. sur \mathbf{R} , de sorte que la définition de l'intégrale de Lebesgue montre que

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

Ceci 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.

L'énoncé ci-dessus admet une sorte de réciproque extrêmement importante pour la suite de la théorie de l'intégrale de Lebesgue.

Théorème 5.1. Si $f \in \mathcal{L}^1(\Omega)$, et si

$$\int_{\Omega} |f(x)| dx = 0,$$

alors $f = 0$ p.p. sur Ω .

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

$$\int_a^b |f(x)| dx = 0,$$

alors $f = 0$ sur $[a, b]$.

DÉMONSTRATION. Comme $f \in \mathcal{L}^1(\Omega)$, sa valeur absolue $|f| \in \mathcal{L}^1(\Omega)$. D'après le Lemme 5.6 page 84, pour tout $\varepsilon > 0$, il existe $u, v \in \mathcal{L}^+(\Omega)$ telles que $|f| = u - v$ p.p. sur Ω , avec $v \geq 0$ sur Ω et

$$\int_{\Omega} v(x) dx \leq \varepsilon.$$

Donc u vérifie $u = |f| + v \geq |f| \geq 0$ p.p. sur Ω et

$$\int_{\Omega} u(x) dx = \int_{\Omega} v(x) dx \leq \varepsilon.$$

Effectuons la décomposition ci-dessus pour $\varepsilon = 2^{-2n}$ 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 $u_n \in \mathcal{L}^+(\Omega)$, $u_n \geq |f| \geq 0$ p.p. sur Ω et

$$\int_{\Omega} u_n(x) dx \leq 2^{-2n}.$$

Par convergence monotone

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

puisque U est limite simple d'une suite croissante de fonctions (les sommes partielles de la série) et puisque

$$\sum_{n \geq 0} \int_{\Omega} 2^n u_n(x) dx \leq \sum_{n \geq 0} 2^{-n} = 2 < +\infty.$$

Or, pour tout $n \geq 0$, on a $u_n \geq |f|$ p.p. sur Ω , de sorte que

$$U(x) \geq |f(x)| \sum_{n \geq 0} 2^n = +\infty,$$

dès que $|f(x)| > 0$. D'où la conclusion. \square

Par construction, une fonction intégrable 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 5.8. *Soit $f \in \mathcal{L}^1(\Omega)$. Alors $|f| < +\infty$ p.p. sur Ω .*

DÉMONSTRATION. Comme $f \in \mathcal{L}^1(\Omega)$, nous avons vu que $|f| \in \mathcal{L}^1(\Omega)$. Il existe donc $u, v \in \mathcal{L}^+(\Omega)$ telles que $|f| = u - v$ p.p. sur Ω . Soit $(v_n)_{n \geq 0}$ une suite de Levi convergeant simplement vers v sur Ω . Évidemment, $v \geq v_0$ sur Ω , de sorte que $u = |f| + v \geq |f| + v_0$ p.p. sur Ω , c'est-à-dire qu'il existe $\mathcal{Z} \subset \Omega$ négligeable tel que $u \geq |f| + v_0$ sur $\Omega - \mathcal{Z}$. Étant donné que la fonction $v_0 \in \mathcal{C}_c(\Omega)$, elle est bornée sur Ω , de sorte que

$$\{x \in \Omega - \mathcal{Z} : |f(x)| = +\infty\} \subset \{x \in \Omega : u(x) = +\infty\},$$

qui est négligeable, par définition d'un ensemble négligeable. \square

4. Théorème de la convergence dominée

L'intégrale de Lebesgue que nous venons de construire conduit à des résultats d'inter-version intégrale \leftrightarrow limite pour des suites de fonctions convergeant simplement (ou même presque partout) qui sont d'une facilité d'utilisation tout à fait remarquable. En particulier, ces résultats n'exigent aucune hypothèse sur la fonction limite, qui dans la plupart des cas est mal connue. Le résultat le plus utile est le *Théorème de la convergence dominée de Lebesgue*, qui s'énonce comme suit.

Théorème 5.2 (Théorème de la convergence dominée de Lebesgue). *Soit $(f_n)_{n \geq 0}$ une suite de fonctions de $\mathcal{L}^1(\Omega)$. Supposons que $f_n \rightarrow f$ p.p. sur Ω et qu'il existe $F \in \mathcal{L}^1(\Omega)$ telle que $|f_n| \leq F$ p.p. sur Ω . Alors, $f \in \mathcal{L}^1(\Omega)$ et*

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

La démonstration du Théorème de la 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 par un énoncé de la convergence monotone dans le cadre des fonctions intégrables.

Proposition 5.9 (Théorème de la convergence monotone de Beppo Levi). *Soit $(f_n)_{n \geq 0}$ suite croissante p.p. sur Ω de fonctions de $\mathcal{L}^1(\Omega)$, telle que 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 $f_n \rightarrow f$ p.p. sur Ω et

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

DÉMONSTRATION. Pour tout $n \in \mathbb{N}$, posons $u_n := f_{n+1} - f_n$. Par hypothèse $u_n \geq 0$ p.p. sur Ω et $u_n \in \mathcal{L}^1(\Omega)$. On utilise le Lemme 5.6 page 84 pour décomposer u_n sous la forme $u_n = v_n - w_n$ p.p. sur Ω avec $v_n, w_n \in \mathcal{L}^+(\Omega)$ où $w_n \geq 0$ sur Ω et

$$\int_{\Omega} w_n(x) dx \leq 2^{-2n}.$$

On a donc $v_n = u_n + w_n \geq 0$ p.p. sur Ω ; d'autre part

$$\sum_{k=0}^n \int_{\Omega} w_k(x) dx \leq 2,$$

et par conséquent

$$\begin{aligned} \sum_{k=0}^n \int_{\Omega} v_k(x) dx &= \int_{\Omega} (f_n(x) - f_0(x)) dx + \sum_{k=0}^n \int_{\Omega} w_k(x) dx \\ &\leq \sup_{k \geq 0} \int_{\Omega} f_k(x) dx - \int_{\Omega} f_0(x) dx + 2. \end{aligned}$$

D'après la Proposition 5.4 page 81, il existe $V, W \in \mathcal{L}^+(\Omega)$ telles que

$$\lim_{n \rightarrow +\infty} \sum_{k=0}^n v_k = V \quad \text{et} \quad \lim_{n \rightarrow +\infty} \sum_{k=0}^n w_k = W,$$

p.p. sur Ω et

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

Définissons $f := f_0 + V - W$ p.p. sur Ω . Étant donné que $f_0 \in \mathcal{L}^1(\Omega)$ et $V, W \in \mathcal{L}^+(\Omega)$, il s'ensuit que $f \in \mathcal{L}^1(\Omega)$ comme différence de deux fonctions de $\mathcal{L}^+(\Omega)$.

D'autre part

$$\lim_{n \rightarrow +\infty} f_n = \lim_{n \rightarrow +\infty} \left(f_0 + \sum_{k=0}^{n-1} v_k - \sum_{k=0}^{n-1} w_k \right) = f_0 + V - W = f,$$

p.p. sur Ω . Enfin, on conclut de la même manière que

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

ce qu'il fallait démontrer. \square

Si l'on abandonne l'hypothèse de monotonie sur la suite de fonctions considérées, l'inter-version intégrale \leftrightarrow limite n'est plus légitime en toute généralité.

Exemple 5.13. On considère la suite de fonctions $f_n(x) = 2nx(1-x^2)^{n-1}$ définies sur $]0, 1[$. On vérifie que $\lim_{n \rightarrow +\infty} f_n(x) = 0$, pour tout $x \in]0, 1[$. De plus

$$0 = \int_0^1 \lim_{n \rightarrow +\infty} f_n(x) dx < \int_0^1 f_n(x) dx = 1.$$

Toutefois, on dispose d'une inégalité, fort utile dans la pratique.

Lemme 5.7 (Lemme de Fatou). *Soit $(f_n)_{n \geq 0}$ une suite de fonctions de $\mathcal{L}^1(\Omega)$ telles que $f_n \geq 0$ p.p. sur Ω . On suppose que la suite des intégrales des fonctions f_n vérifie*

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

Alors, $\underline{\lim}_{n \rightarrow +\infty} f_n \in \mathcal{L}^1(\Omega)$ et

$$\int_{\Omega} \underline{\lim}_{n \rightarrow +\infty} f_n(x) dx \leq \underline{\lim}_{n \rightarrow +\infty} \int_{\Omega} f_n(x) dx.$$

DÉMONSTRATION. L'idée de la démonstration de ce résultat est de se ramener à l'application du Théorème de la convergence monotone de Beppo Levi. L'observation clef est la suivante : pour toute suite de réels $(x_n)_{n \geq 0}$,

$$\underline{\lim}_{n \rightarrow +\infty} x_n = \lim_{k \rightarrow +\infty} \left(\lim_{m \rightarrow +\infty} \min(x_k, \dots, x_m) \right)$$

et les suites $(\lim_{m \rightarrow +\infty} \min(x_k, \dots, x_m))_{k \geq 0}$ et $(\min(x_k, \dots, x_m))_{m \geq k}$ sont monotones (voir le point (iii) dans la Proposition 3.7 page 52).

Pour $m \geq k \geq 0$, on pose $f_{k,m} := \min(f_k, \dots, f_m)$, qui est une fonction définie p.p. sur Ω . Évidemment, $f_{k,m} \in \mathcal{L}^1(\Omega)$, d'après le point (5) dans la Proposition 5.7 page 85. Fixons un entier k et faisons tendre m vers l'infini. On voit alors que, pour tout $k \geq 0$ fixé, on a :

- (i) la suite $f_{k,m} \rightarrow \phi_k := \inf_{n \geq k} f_n$ p.p. sur Ω quand $m \rightarrow +\infty$;
- (ii) la suite $(f_k(x) - f_{k,m}(x))_{m \geq k}$ est croissante p.p. sur Ω ;
- (iii) pour tout $m \geq k$,

$$\int_{\Omega} (f_k - f_{k,m})(x) dx \leq \int_{\Omega} f_k(x) dx < +\infty.$$

D'après la Proposition 5.9 page 89 appliquée à la suite $(f_k - f_{k,m})_{m \geq k}$, la fonction $\phi_k \in \mathcal{L}^1(\Omega)$ et

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

Maintenant, nous avons :

- (i) la suite $\phi_k \rightarrow \underline{\lim}_{n \rightarrow +\infty} f_n$ p.p. sur Ω ;
- (ii) la suite $(\phi_k)_{k \geq 0}$ est croissante p.p. sur Ω ;
- (iii) pour tout $k \geq 0$,

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

On appliquant à nouveau la Proposition 5.9 page 89, cette fois à la suite $(\phi_k)_{k \geq 0}$, on trouve alors que la fonction

$$\underline{\lim}_{n \rightarrow +\infty} f_n \in \mathcal{L}^1(\Omega),$$

et que

$$\lim_{k \rightarrow +\infty} \int_{\Omega} \phi_k(x) dx = \int_{\Omega} \underline{\lim}_{n \rightarrow +\infty} f_n(x) dx.$$

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

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

on conclut que

$$\int_{\Omega} \underline{\lim}_{n \rightarrow +\infty} f_n(x) dx \leq \underline{\lim}_{n \rightarrow +\infty} \int_{\Omega} f_n(x) dx.$$

Ce qui termine la démonstration. □

Le Théorème de la convergence dominée de Lebesgue est une conséquence du Lemme de Fatou.

DÉMONSTRATION DU THÉORÈME DE LA CONVERGENCE DOMINÉE. Pour tout $n \geq 0$, on note $g_n := F - f_n$. Par hypothèse, $g_n \in \mathcal{L}^1(\Omega)$ comme différence de deux fonctions intégrables, et $g_n \geq 0$ 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 $\underline{\lim}_{n \rightarrow +\infty} g_n \in \mathcal{L}^1(\Omega)$. Or $F - f = \underline{\lim}_{n \rightarrow +\infty} g_n$ p.p. sur Ω , donc $f \in \mathcal{L}^1(\Omega)$.

Définissons maintenant $h_n := 2F - |f_n - f|$. Étant donné que F, f_n et $f \in \mathcal{L}^1(\Omega)$, la fonction $h_n \in \mathcal{L}^1(\Omega)$ (voir le point (4) dans la Proposition 5.7 page 85) et $h_n \geq 0$ 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

$$\begin{aligned} 2 \int_{\Omega} F(x) dx &= \int_{\Omega} \varliminf_{n \rightarrow +\infty} h_n(x) dx \\ &\leq \varliminf_{n \rightarrow +\infty} \int_{\Omega} h_n(x) dx \\ &= 2 \int_{\Omega} F(x) dx - \overline{\lim}_{n \rightarrow +\infty} \int_{\Omega} |f_n(x) - f(x)| dx, \end{aligned}$$

d'où

$$0 = \overline{\lim}_{n \rightarrow +\infty} \int_{\Omega} |f_n(x) - f(x)| dx.$$

Finalement, il suffit d'utiliser l'inégalité

$$\left| \int_{\Omega} f_n(x) dx - \int_{\Omega} f(x) dx \right| \leq \int_{\Omega} |f_n(x) - f(x)| dx,$$

pour achever la démonstration. \square

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

Remarque 5.2. L'hypothèse de domination est absolument essentielle dans le Théorème de la convergence dominée. En l'absence de cette hypothèse, les deux phénomènes suivants peuvent se produire. Pour décrire ces phénomènes, considérons $\Omega = \mathbf{R}$ et choisissons une fonction positive $\phi \in \mathcal{C}_c(\mathbf{R})$, telle que

$$\int_{\mathbf{R}} \phi(x) dx = 1.$$

On peut par exemple prendre $\phi(x) := (1 - |x|)^+$.

Phénomène de concentration : Pour tout $n \geq 0$, définissons la fonction $f_n(x) := n \phi(nx)$.

Évidemment $f_n \in \mathcal{C}_c(\mathbf{R})$ et $f_n(x) = 0$, pour tout $n \geq 1$ si $|x| \geq a/n$ où $a > 0$ est choisi tel que le support de la fonction ϕ est inclus dans $[-a, a]$. En particulier, $\lim_{n \rightarrow +\infty} f_n(x) = 0$ si $x \neq 0$. Donc $f_n \rightarrow 0$ p.p. sur \mathbf{R} . Pourtant, un changement de variable dans l'intégrale de f_n montre que

$$\int_{\mathbf{R}} f_n(x) dx = 1,$$

pour tout $n \geq 1$. Donc

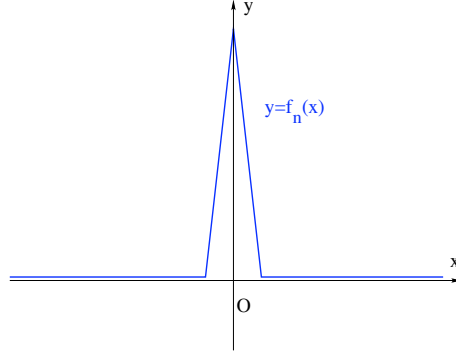
$$1 = \lim_{n \rightarrow +\infty} \int_{\mathbf{R}} f_n(x) dx > \int_{\mathbf{R}} \lim_{n \rightarrow +\infty} f_n(x) dx = 0.$$

Dans cet exemple, les fonctions positives f_n peuvent être vues comme des *densités de masse*, et 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.

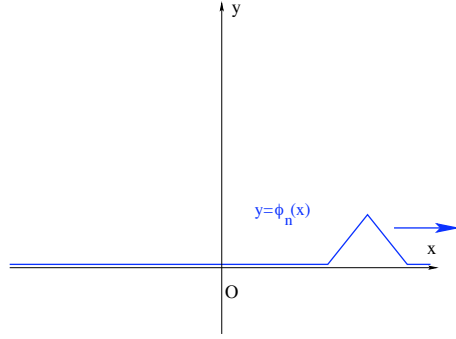
Phénomène d'évanescence : Pour tout $n \in \mathbf{N}$, on définit la fonction $g_n(x) := \phi(x - n)$.

De nouveau, la suite g_n converge simplement vers 0 sur \mathbf{R} . Pourtant

$$\lim_{n \rightarrow +\infty} \int_{\mathbf{R}} g_n(x) dx = \int_{\mathbf{R}} \phi(y) dy > \int_{\mathbf{R}} \lim_{n \rightarrow +\infty} g_n(x) dx = 0.$$

FIGURE 5.3. Phénomène de concentration : graphe de f_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 que l'on ne peut la détecter par convergence simple, puisque la convergence simple ne montre ce qui se passe que localement.

FIGURE 5.4. Phénomène d'évanescence : graphe de g_n .

5. Caractérisation géométrique des ensembles 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 de Levi. 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 négligeables, qui elle est de nature plus géométrique.

Attention : tout ce qui est dit dans cette section à propos des ensembles négligeables ne vaut que pour la construction partant de la seule forme linéaire positive sur $\mathcal{C}_c(\Omega)$ définie par l'intégrale usuelle, c'est-à-dire pour des ensembles négligeables, et pas pour les généralisations de cette construction qui seront envisagées ci-dessous dans le Chapitre 7, correspondant à une forme linéaire positive *quelconque* sur $\mathcal{C}_c(\Omega)$.

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

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

l'(hyper)cube ouvert de centre $a \in \mathbf{R}^N$ et de côté $2r > 0$. On notera $|C(a, r)|$ le *volume* de ce cube. Bien entendu,

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

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

Théorème 5.3. *Un sous-ensemble \mathcal{Z} de \mathbf{R}^N est négligeable si et seulement si, pour tout $\varepsilon > 0$, il existe une famille **dénombrable** $(C_i)_{i \in I}$ de cubes de \mathbf{R}^N telle que*

$$\mathcal{Z} \subset \bigcup_{i \in I} C_i, \quad \text{et} \quad \sum_{i \in I} |C_i| \leq \varepsilon.$$

DÉMONSTRATION. Nous donnons la démonstration dans le seul cas où la dimension de l'espace $N = 1$, laissant le soin au lecteur de traiter le cas des dimensions $N \geq 2$.

Supposons que $\mathcal{Z} \subset \mathbf{R}$ est négligeable. Soit $f \in \mathcal{L}^+(\mathbf{R})$ telle que $f(x) = +\infty$ pour tout $x \in \mathcal{Z}$. On peut supposer que $f \geq 0$ sur \mathbf{R} . Donnons nous un entier $n \geq 1$. Alors

$$\mathcal{Z} \subset f^{-1}(\{+\infty\}) \subset f^{-1}(]n, +\infty]).$$

Étant donné que $f \in \mathcal{L}^+(\mathbf{R})$, elle est s.c.i. sur \mathbf{R} (voir la Proposition 5.1 page 76), de sorte que $f^{-1}(]n, +\infty])$ est un ouvert de \mathbf{R} .

D'après la Proposition 2.5 page 33, l'ensemble $f^{-1}(]n, +\infty])$ est la réunion d'une famille dénombrable $(I_k^n)_{k \in J_n}$ d'intervalles ouverts disjoints de \mathbf{R} , de sorte que

$$\mathcal{Z} \subset f^{-1}(]n, +\infty]) = \bigcup_{k \in J_n} I_k^n.$$

Étant donné que $f(x) > n$ pour tout $x \in \Omega(n)$, pour tout sous-ensemble fini $K \subset J_n$, on a

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

où $|I_k^n|$ est la longueur de l'intervalle I_k^n . On remarquera que la fonction $\mathbf{1}_{I_k^n}(f - n) \in \mathcal{L}^+(\mathbf{R})$ (voir l'exemple 5.6 page 78) et que $\mathbf{1}_{I_k^n} \in \mathcal{L}^+(\mathbf{R})$ (voir l'exemple 5.3 page 77). Au total, comme $f \geq 0$ sur \mathbf{R} , on a

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

pour tout $K \subset J_n$, sous-ensemble fini d'indices, on en déduit que

$$\sum_{k \in J_n} |I_k^n| \leq \frac{1}{n} \int_{\mathbf{R}} f(x) dx.$$

Étant donné $\varepsilon > 0$, en choisissant $n > 0$ assez grand pour que

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

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

Réciproquement, soit $\mathcal{Z} \subset \mathbf{R}$, et supposons que, pour tout $n \geq 0$, il existe une suite $(I_k^n)_{k \in \mathbf{N}}$ d'intervalles ouverts de \mathbf{R} tels que

$$\mathcal{Z} \subset \bigcup_{k \geq 0} I_k^n, \quad \text{et} \quad \sum_{k \geq 0} |I_k^n| < 2^{-2n}.$$

On vérifie alors sans difficulté que la fonction

$$f := \sum_{n \geq 0} 2^n \left(\sum_{k \geq 0} \mathbf{1}_{I_k^n} \right) \in \mathcal{L}^+(\mathbf{R}),$$

avec

$$\int_{\mathbf{R}} f(x) dx = \sum_{n \geq 0} 2^n \left(\sum_{k \geq 0} |I_k^n| \right) \leq \sum_{n \geq 0} 2^{-n} = 2,$$

et que

$$f(x) = \sum_{n \geq 1} 2^n = +\infty$$

pour tout $x \in \mathcal{Z}$. □

Exemple 5.14. On retrouve très simplement grâce au Théorème 5.3 page ci-contre le fait qu'un ouvert non vide de \mathbf{R}^N n'est jamais négligeable — résultat déjà établi plus haut, voir l'exemple 5.10 page 83. Plus généralement, une partie de \mathbf{R}^N d'intérieur non vide n'est jamais négligeable.

En effet, d'après le Théorème 5.3 page ci-contre, 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 5.15. 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 . Étant donné $n > 1$, on découpe I en n segments de longueur $1/n$, notés I_1, \dots, I_n . Le segment I est donc recouvert par la réunion des disques D_1, \dots, D_n de diamètres $1/n$. Pour tout $j = 1, \dots, n$, soit C_j le plus petit carré dont les côtés sont parallèles aux axes et qui contient D_j . La longueur du côté d'un tel carré est égale à $1/n$. Donc, on a $I \subset C_1 \cup \dots \cup C_n$ et

$$|C_1| + \dots + |C_n| = n \cdot \left(\frac{1}{n} \right)^2 = \frac{1}{n}.$$

Ainsi, pour tout $\varepsilon > 0$, le segment I peut être recouvert par $n = [1/\varepsilon] + 1$ carrés dont la somme des aires reste inférieure à ε . Une droite affine de \mathbf{R}^2 étant la réunion dénombrable de segments de longueur 1, c'est un sous-ensemble négligeable de \mathbf{R}^2 .

Les autres énoncés sont laissés au lecteur à titre d'exercice.

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 négligeables.

Exemple 5.16 (Ensemble triadique de Cantor). Pour tous $a < b \in \mathbf{R}$, on note T l'opération suivante

$$T[a, b] := [a, a + \frac{b-a}{3}] \cup [b - \frac{b-a}{3}, b].$$

Géométriquement, on coupe le segment $[a, b]$ en 3 et l'on enlève l'intervalle ouvert du milieu.

Par exemple $T[0, 1] = [0, 1/3] \cup [2/3, 1]$.

On note $K_0 := [0, 1]$, $K_1 := T K_0 = [0, 1/3] \cup [2/3, 1]$, puis, par récurrence,

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

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

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

Par exemple $K_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1]$. L'ensemble triadique de Cantor est défini par

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

Voir sur la Figure 5.5 les premières étapes de la construction de l'ensemble triadique de Cantor.



FIGURE 5.5. Les 6 premières étapes de la construction de l'ensemble triadique de Cantor.

Vérifier que K est un compact de \mathbf{R} qui est *totalelement discontinu* (c'est-à-dire que la composante connexe de tout point x de K est réduite au singleton $\{x\}$), et qui est un *ensemble parfait* (c'est à dire que K n'a pas de point isolé).

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

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

Exemple 5.17 (Triangle de Sierpiński). Le triangle de Sierpiński est l'analogue de l'ensemble triadique de Cantor dans le plan \mathbf{R}^2 . On a représenté sur la Figure 5.6 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'*ensembles fractals*, c'est-à-dire de *dimension non entière*.³ On vérifie que l'ensemble de 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 strictement inférieure à celle de

3. Étant donné $0 < \varepsilon \ll 1$, on vérifie qu'il faut environ $O(1/\varepsilon^N)$ hypercubes de côté de longueur ε pour recouvrir l'hypercube $[0, 1]^N \subset \mathbf{R}^N$. La dimension N apparaît donc comme l'exposant de $1/\varepsilon$ dans le nombre d'hypercubes de côté ε qui sont nécessaires pour recouvrir $[0, 1]^N$. En général, étant donné un compact K de \mathbf{R}^N , on note $\mathcal{Z}(K, \varepsilon)$ la borne inférieure du nombre d'hypercubes de côté $\varepsilon \ll 1$ qui sont nécessaires pour



FIGURE 5.6. Le triangle de Sierpiński. À chaque étape, on enlève, dans chaque triangle équilatéral noir, le triangle blanc dont les sommets sont les milieux des côtés du triangle noir.

l'espace ambiant. On remarquera que tout ensemble fractal de dimension strictement inférieure à la dimension de l'espace ambiant est négligeable.

Voici un dernier exemple d'ensemble négligeable pathologique.

Exemple 5.18 (Ensemble de Besicovich). Le problème de Kakeya (1917) consiste à trouver un domaine du plan \mathbf{R}^2 d'aire la plus petite possible dans lequel on peut retourner une aiguille (en fait un segment) de longueur 1. Bien entendu, dans un disque de diamètre 1 on peut retourner une aiguille de longueur 1. Néanmoins, un moment de réflexion permet de se convaincre qu'il existe des ensembles dont l'aire est plus petite que celle du disque de diamètre 1 dans lesquels on peut retourner une aiguille de longueur 1, par exemple : un triangle équilatéral de côté $2/\sqrt{3}$. On pourrait être tenté de croire qu'il existe une *forme optimale* pour ce problème. Contrairement à l'intuition, il n'en est rien et Besicovich a montré en 1928 qu'il existait en fait des ensembles négligeables dans lesquels on pouvait retourner une aiguille de longueur 1.

Plus généralement, un ensemble de Besicovich dans \mathbf{R}^2 est un ensemble qui contient au moins un segment de longueur 1 dans chaque direction. Évidemment, un domaine du plan dans lequel on peut retourner une aiguille de longueur 1 est un ensemble de Besicovich. Un problème intéressant est de construire des ensembles de Besicovich aussi petits que possible et de déterminer leur dimension fractale. Il s'agit d'un problème difficile qui fait apparaître des liens surprenants entre ce problème de nature géométrique, l'analyse de Fourier et l'arithmétique.⁴

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 que l'on le rencontre au voisinage de tout point. Examinons plus en détail les rapports entre ces deux notions.

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

Exemple 5.19. Par exemple \mathbf{Q} est négligeable et dense dans \mathbf{R} ; de même, \mathbf{Q}^N est négligeable et dense dans \mathbf{R}^N .

D'autre part, on a vu plus haut (voir l'exemple 5.14 page 95) 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 :

recouvrir K . On définit alors $\dim_{Cube}(K)$, la *dimension de Minkowski* de K , comme la limite (si elle existe)

$$\dim_{Cube}(K) := \lim_{\varepsilon \rightarrow 0} \frac{\ln \mathcal{Z}(K, \varepsilon)}{\ln(1/\varepsilon)}.$$

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 $\mathcal{Z}(K, 3^{-n}) \leq 2^n$. Donc $\dim_{Cube}(K) \leq \frac{\ln 2}{\ln 3}$.

4. Voir l'article de Tao, Terence. *From Rotating Needles to Stability of Waves : Emerging Connections between Combinatorics, Analysis and PDE*. Notices of the AMS 48 (3) : 297-303.

Proposition 5.10. *Soit Ω ouvert de \mathbf{R}^N . Alors le complémentaire d'une partie négligeable de Ω est dense dans Ω .*

Terminons ce paragraphe par un résultat qui explicite les liens entre l'intégration au sens de Riemann et l'intégration au sens de Lebesgue et dont la démonstration est proposée à titre d'exercice (voir l'exercice 5.1).

Théorème 5.4. *Une fonction bornée définie sur un intervalle $[a, b]$ est intégrable au sens de Riemann si et seulement si l'ensemble des points de discontinuité de f est négligeable.*

Exercice 5.1. [Intégrabilité au sens de Riemann]. Soient $a < b \in \mathbf{R}$ et $f : [a, b] \rightarrow \mathbf{R}$ une fonction bornée. Une subdivision de $[a, b]$ est une suite finie σ de la forme $\sigma := \{a = x_0 < x_1 < \dots < x_n = b\}$. On note $|\sigma| = \max_{1 \leq j \leq n} (x_j - x_{j-1})$ le pas de la subdivision σ et l'on note $S(a, b)$ l'ensemble des subdivisions de $[a, b]$. On note

$$\begin{cases} I^+(\sigma) : &= \sum_{j=1}^n (x_j - x_{j-1}) \sup_{x_{j-1} \leq y < x_j} f(y), \\ I^-(\sigma) : &= \sum_{j=1}^n (x_j - x_{j-1}) \inf_{x_{j-1} \leq y < x_j} f(y). \end{cases}$$

La fonction f est dite intégrable au sens de Riemann si et seulement si

$$\sup_{\sigma \in S(a, b)} I^-(\sigma) = \inf_{\sigma \in S(a, b)} I^+(\sigma).$$

Si tel est le cas, on pose

$$\int_a^b f(x) dx := \sup_{\sigma \in S(a, b)} I^-(\sigma) = \inf_{\sigma \in S(a, b)} I^+(\sigma).$$

- 1) Vérifier que $\sup_{\sigma \in S(a, b)} I^-(\sigma) \leq \inf_{\sigma \in S(a, b)} I^+(\sigma)$.
- 2) Pour tout intervalle $I \subset [a, b]$, on note $\omega(I) = \sup_{x, y \in I} |f(y) - f(x)|$ l'oscillation de f sur l'intervalle I . Montrer que f est intégrable au sens de Riemann sur $[a, b]$ si et seulement si

$$\lim_{\substack{\sigma = (x_0, \dots, x_n) \\ |\sigma| \rightarrow 0}} \sum_{j=1}^n (x_j - x_{j-1}) \omega([x_{j-1}, x_j]) = 0.$$

- 3) Pour tout $x_0 \in [a, b]$, on note

$$v(x_0) := \inf_{h > 0} \omega([x_0 - h, x_0 + h] \cap [a, b]).$$

Montrer que f est continue en $x_0 \in [a, b]$ si et seulement si $v(x_0) = 0$.

- 4) Soit $(\sigma^n)_{n \geq 1}$ une suite croissante de subdivisions de $[a, b]$, i.e.

$$\sigma^n = \{a = x_0^n < x_1^n < \dots < x_{n-1}^n < x_n^n = b\},$$

et $\sigma^n \subset \sigma^{n+1}$. On suppose que $\lim_{n \rightarrow +\infty} |\sigma^n| = 0$. On note

$$\phi_n := \sum_{k=0}^{n-1} \left(\inf_{x \in [x_k^n, x_{k+1}^n[} f(x) \right) \mathbf{1}_{[x_k^n, x_{k+1}^n[},$$

et

$$\psi_n := \sum_{k=0}^{n-1} \left(\sup_{x \in [x_k^n, x_{k+1}^n[} f(x) \right) \mathbf{1}_{[x_k^n, x_{k+1}^n[}.$$

Montrer que les suites $(\phi_n)_{n \geq 1}$ et $(\psi_n)_{n \geq 1}$ sont monotones.

5) On note $X := \{x_k^n : n \geq 1, k = 0, \dots, n\}$. Montrer que X est négligeable. Montrer que

$$v(x) = \lim_{n \rightarrow +\infty} \psi_n(x) - \lim_{n \rightarrow +\infty} \phi_n(x),$$

p.p. sur $]a, b[$.

6) Montrer que l'ensemble des points de discontinuité de f dans $[a, b]$ est négligeable si et seulement si f est intégrable au sens de Riemann.

7) On considère la fonction $\phi : \mathbf{R} \rightarrow \mathbf{R}$ définie comme suit :

$$\begin{cases} \phi(0) = 0, \\ \phi(x) = 0 & \text{si } x \in \mathbf{R} - \mathbf{Q}, \\ \phi\left(\frac{p}{q}\right) = \frac{1}{q} & \text{pour tous } (p, q) \in \mathbf{Z} \times (\mathbf{N} - \{0\}) \text{ premiers entre eux.} \end{cases}$$

Vérifier que ϕ est intégrable au sens de Riemann sur $[0, 1]$ et calculer son intégrale.

6. Fonctions mesurables

Dans la théorie de l'intégration que nous venons de construire, les fonctions 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 la 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} . Ceci va nous conduire à la notion de *fonction mesurable*.

De plus, vérifier qu'une fonction donnée est intégrable, en revenant à la définition de $\mathcal{L}^1(\Omega)$ peut s'avérer fastidieux. Il est donc nécessaire d'avoir un critère simple permettant de savoir si une fonction est intégrable. Nous donnons un critère s'appuyant sur la notion de fonction mesurable.

Définition 5.8. Une fonction $f : \Omega \rightarrow [-\infty, +\infty]$ est mesurable s'il existe une suite $(f_n)_{n \geq 0}$ de fonctions continues à support compact qui sont définies sur Ω et qui converge vers f p.p. sur Ω .

Voici les premières propriétés des fonctions mesurables.

Proposition 5.11. La classe des fonctions mesurables sur Ω vérifie les propriétés suivantes :

- (a) toute fonction continue sur Ω est mesurable ;
- (b) toute fonction continue par morceaux sur un intervalle $I \subset \mathbf{R}$ est mesurable ;
- (c) toute fonction appartenant à $\mathcal{L}^+(\Omega)$ est mesurable ;
- (d) si f_1, \dots, f_N , définies sur Ω , sont mesurables et si $\Phi : \mathbf{R}^N \rightarrow \mathbf{R}$ est continue, alors la fonction

$$\Phi(f_1, \dots, f_N)(x) := \Phi(f_1(x), \dots, f_N(x)),$$

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 (d), on pourra dans un premier temps commencer par le démontrer dans

le cas où $\Phi(0, \dots, 0) = 0$, puis vérifier que si ϕ est une fonction mesurable sur Ω , il en va de même pour $\phi + C$, pour tout $C \in \mathbf{R}$.

Voici quelques applications importantes du résultat précédent, notamment de l'énoncé (d) de la proposition précédente.

Corollaire 5.1. *On a les énoncés suivants :*

- (a) *si $f, g : \Omega \rightarrow \mathbf{R}$ sont mesurables et si $\lambda, \mu \in \mathbf{R}$, alors les fonctions $\lambda f + \mu g$ et fg sont mesurables sur Ω ;*
- (b) *toute fonction appartenant à $\mathcal{L}^1(\Omega)$ est mesurable sur Ω ;*
- (c) *si $f, g : \Omega \rightarrow \mathbf{R}$ sont mesurables, alors les fonctions $\max(f, g)$ et $\min(f, g)$ sont mesurables. En particulier, f^+ , f^- et $|f|$ sont mesurable si f est mesurable.*

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

Théorème 5.5. *Soit f , fonction mesurable sur Ω . Supposons qu'il existe $g \in \mathcal{L}^1(\Omega)$ telle que $|f| \leq g$ p.p. sur Ω . Alors $f \in \mathcal{L}^1(\Omega)$.*

DÉMONSTRATION. Comme f est mesurable sur Ω , il existe une suite $(f_n)_{n \geq 0}$ de fonctions appartenant à $\mathcal{C}_c(\Omega)$ telle que $f_n \rightarrow f$ p.p. sur Ω . Pour tout $n \geq 0$, définissons

$$h_n := \max(\min(f_n, g), -g),$$

p.p. sur Ω . En utilisant la propriété (5) de la Proposition 5.7 page 85 on conclut que $h_n \in \mathcal{L}^1(\Omega)$. De plus, $h_n \rightarrow f$ p.p. sur Ω et $|h_n| \leq g$ p.p. sur Ω . Le Théorème de la convergence dominée nous assure que $f \in \mathcal{L}^1(\Omega)$. \square

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

Théorème 5.6. *Soit $(f_n)_{n \geq 0}$ une suite de fonctions mesurables, définies sur Ω et f une fonction définie sur Ω , telles que $f_n \rightarrow f$ p.p. sur Ω . Alors, la fonction f est mesurable.*

DÉMONSTRATION. Nous allons démontrer ce résultat en plusieurs étapes. Dans un premier temps, on démontre le résultat en supposant que la suite $(f_n)_{n \geq 0}$ est croissante et est constituée d'éléments appartenant à $\mathcal{L}^+(\Omega)$, ensuite on suppose que les éléments de cette suite croissante appartiennent à $\mathcal{L}^1(\Omega)$. Enfin, on démontre le résultat sans l'hypothèse de croissance de la suite.

Étape 1 : Soit $(f_n)_{n \geq 0}$ suite de fonctions de $\mathcal{L}^+(\Omega)$ qui est croissante p.p. sur Ω et converge p.p. sur Ω vers une fonction f . Montrons que f est une fonction mesurable sur Ω .

Pour chaque $k \geq 1$, il existe $(f_{k,n})_{n \geq 0}$, une suite de Levi qui converge simplement vers f_k sur Ω . Posons

$$g_n := \max(f_{0,n}, \dots, f_{n,n}).$$

La suite $(g_n(x))_{n \geq 0}$ est croissante sur Ω et converge simplement vers une limite notée g , à valeurs dans $\mathbf{R} \cup \{+\infty\}$. Évidemment g est mesurable sur Ω puisque $g_n \in \mathcal{C}_c(\Omega)$. Comme $g_n \leq \max(f_0, \dots, f_n) = f_n$ p.p. sur Ω , on en déduit que $g \leq f$ p.p. sur Ω . En particulier $g < +\infty$ p.p. sur Ω .

D'autre part, pour $k \leq n$, on a $f_{k,n} \leq g_n$ sur Ω , en passant à la limite pour $n \rightarrow +\infty$, on trouve que $f_k \leq g$ sur Ω . Puis, en passant à la limite en $k \rightarrow +\infty$, on trouve que $f \leq g$ p.p. sur Ω . Au total, $f = g$ p.p. sur Ω , et comme g est mesurable sur Ω , il en est de même pour f .

Étape 2 : Supposons maintenant que la suite $(f_n)_{n \geq 0}$ est une suite de fonctions de $\mathcal{L}^1(\Omega)$ qui est croissante p.p. sur Ω et qui converge p.p. sur Ω vers f . Montrons que f est une fonction mesurable sur Ω .

D'après le Lemme 5.6 page 84, pour tout $n \geq 1$, il existe $g_n, h_n \in \mathcal{L}^+(\Omega)$ telles que

$$0 \leq f_n - f_{n-1} = g_n - h_n,$$

p.p. sur Ω , tandis que $h_n \geq 0$ sur Ω et

$$\int_{\Omega} h_n(x) dx < 2^{-n}.$$

Remarquons que, pour tout $n \geq 1$, on a $g_n = f_n - f_{n-1} + h_n \geq 0$, p.p. sur Ω , de sorte que les suites

$$H_n := \sum_{k=1}^n h_n \quad \text{et} \quad G_n = \sum_{k=1}^n g_n,$$

sont croissantes p.p. sur Ω (en fait la suite $(H_n)_{n \geq 1}$ est même croissante sur Ω) et appartiennent à $\mathcal{L}^+(\Omega)$. On note H la limite simple de la suite $(H_n)_{n \geq 1}$ sur Ω et G la limite simple de la suite $(G_n)_{n \geq 1}$ p.p. sur Ω . D'après l'étape 1, G et H sont mesurable sur Ω et, au total, $f = G - H + f_0$ est mesurable comme somme de fonctions mesurables sur Ω .

Étape 3 : Supposons que la suite $(f_n)_{n \geq 0}$ est une suite de fonctions mesurables sur Ω qui sont positives p.p. sur Ω et qui converge simplement vers f p.p. sur Ω . Montrons que f est mesurable sur Ω .

Pour tout $n \geq 0$, il existe une suite $(f_{n,k})_{k \geq 1}$ de fonctions de $\mathcal{C}_c(\Omega)$ qui sont positives et telles que

$$\lim_{k \rightarrow +\infty} f_{n,k} = f_n,$$

p.p. sur Ω . De plus, on peut toujours supposer que

$$F_{n,k} := \int_{\Omega} f_{n,k}(x) dx > 0,$$

pour tout $k, n \geq 0$ (exercice). D'après la Proposition 5.4 page 81

$$g := \sum_{n \geq 0} \left(\sum_{k \geq n} \frac{1}{2^k F_{n,k}} f_{n,k} \right),$$

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

Soit alors $m \in \mathbf{N} - \{0\}$. D'après la Proposition 5.5 page ci-contre, la fonction $h_{m,n} := \min(f_n, m g)$ définie p.p. sur Ω appartient à $\mathcal{L}^1(\Omega)$. D'autre part, pour tout $m \geq 1$ fixé, $\lim_{n \rightarrow +\infty} h_{m,n} = h_m := \min(f, m g)$ p.p. sur Ω et $0 \leq h_{m,n} \leq m g$ p.p. sur Ω . Donc, d'après le Théorème de la convergence dominée de Lebesgue, pour tout $m \geq 1$, la fonction $h_m \in \mathcal{L}^1(\Omega)$.

De plus la suite $(h_m)_{m \geq 0}$ est croissante p.p. sur Ω et converge vers f p.p. sur Ω . D'après le résultat de l'étape 2, la fonction f est mesurable sur Ω .

Pour le cas général, on décompose chaque terme de la suite $(f_n)_{n \geq 1}$ sous la forme $f_n = (f_n)^+ - (f_n)^-$, et l'on applique le résultat de l'étape 3, ainsi que le Corollaire 5.1 page précédente pour conclure. \square

Le Corollaire 5.1 page ci-contre 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 5.2. *Soit $f : \Omega \rightarrow \mathbf{R}$ une fonction mesurable telle que $f(x) \neq 0$ p.p. sur Ω . Alors la fonction $x \mapsto 1/f(x)$, qui est définie p.p. sur Ω , est mesurable sur Ω .*

DÉMONSTRATION. Il suffit d'observer que, lorsque $n \rightarrow +\infty$,

$$h_n := \frac{n f}{1 + n|f|^2} \rightarrow \frac{1}{f},$$

en tout point de Ω où f ne s'annule pas. Or, pour tout $n \geq 1$, la fonction h_n est mesurable sur Ω d'après le point (4) de la Proposition 5.11 page 99), appliquée à la fonction continue

$$\Phi(z) = \frac{nz}{1 + n|z|^2}.$$

On conclut alors grâce au Théorème 5.6 page 100. □

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.

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 choix⁵

5. Cet axiome peut se formuler ainsi : pour tout ensemble $I \neq \emptyset$, et toute famille $(X_i)_{i \in I}$ d'ensembles indexée par I

$$X_i \neq \emptyset \text{ pour tout } i \in I \quad \Rightarrow \quad \prod_{i \in I} X_i \neq \emptyset.$$

Autrement dit, il existe $(x_i)_{i \in I}$ avec $x_i \in X_i$ 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 $X_1 \neq \emptyset$, on peut choisir $x_1 \in X_1$, puis comme $X_2 \neq \emptyset$, 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 que l'on ait pu terminer la construction en temps fini, et donc que l'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 $f : I \rightarrow X$ 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 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 que l'on songe à s'en passer — par exemple, c'est cet axiome qui garantit que tout espace vectoriel, même de dimension infinie, admet une base. Mais, en contrepartie, 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 $\mathbf{R}^{\mathbf{N}}$. On verra dans le chapitre suivant une conséquence encore plus dérangeante de ce même axiome.

7. Intégration et fonctions mesurables

Jusqu'ici, on a construit l'intégrale sur l'espace \mathcal{L}^1 des fonctions intégrables, avec un théorème d'interversion intégrale \leftrightarrow limite (le théorème de la convergence dominée) qui correspond bien à ce que l'on cherchait, voir notamment les points (i) et (ii) de l'introduction, section 1 page 69.

Cette construction laisse toutefois plusieurs questions en suspens. Par exemple, *étant donnée une fonction f définie p.p. sur Ω , comment vérifie-t-on si $f \in \mathcal{L}^1(\Omega)$?* Nous allons donner quelques éléments de réponse à cette question pratique, en nous basant en particulier sur la notion de fonction mesurable que nous venons justement d'introduire.

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 est nécessaire qu'elle soit mesurable sur Ω (voir le Corollaire 5.1 page 100). 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 que la fonction f est mesurable sur Ω , et cherchons quelle(s) condition(s) supplémentaire(s) f doit vérifier pour être intégrable 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 (voir le Théorème 5.5 page 100.)

Toutefois, cet énoncé n'est pas entièrement satisfaisant pour l'esprit, dans la mesure où, pour vérifier que la fonction f est intégrable sur Ω , il faut déjà savoir que la fonction g , à laquelle on compare $|f|$, l'est elle-même. En réalité, on peut modifier très légèrement l'énoncé de ce résultat pour arriver à l'énoncé suivant qui lui est d'ailleurs équivalent.

Théorème 5.7 (Théorème de comparaison). *Soit $f : \Omega \rightarrow [-\infty, +\infty]$ une fonction mesurable sur Ω . Alors $f \in \mathcal{L}^1(\Omega)$ si et seulement s'il existe $F \in \mathcal{L}^+(\Omega)$ telle que $|f| \leq F$ p.p. sur Ω .*

DÉMONSTRATION. Comme $\mathcal{L}^+(\Omega) \subset \mathcal{L}^1(\Omega)$, la condition de comparaison est évidemment suffisante d'après le Théorème 5.5 page 100.

Réciproquement, si $f \in \mathcal{L}^1(\Omega)$, on sait, d'après le point (4) de la Proposition 5.7 page 85), que $|f| \in \mathcal{L}^1(\Omega)$. Il existe donc deux fonctions $g, h \in \mathcal{L}^+(\Omega)$ telles que

$$|f| = g - h = (g^+ + h^-) - (g^- + h^+),$$

p.p. sur Ω . Posons $F := g^+ + h^-$. D'après (2) et (3) de la Proposition 5.2 page 78, on a $F \in \mathcal{L}^+(\Omega)$ et ce qui précède montre que $|f| \leq F$ p.p. sur Ω . \square

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 $|f| \leq F$ p.p. sur Ω , dans la mesure où il suffit de chercher si l'on peut construire $(F_n)_{n \geq 0}$ une suite de Levi sur Ω telle que

$$\lim_{n \rightarrow +\infty} F_n \geq |f|.$$

Dans le même ordre d'idées, pour vérifier la condition de domination dans le Théorème de la convergence dominée 5.2 page 89, on cherchera en pratique une fonction $F \in \mathcal{L}^+(\Omega)$.

Remarque 5.3. Comparons l'énoncé du Théorème 5.5 page 100 (ou du Théorème 5.7 page précédente) avec celui du Théorème de la convergence dominée de Lebesgue 5.2 page 89.

Dans le Théorème 5.5 page 100 ou le Théorème 5.7 page précédente, comme f est mesurable sur Ω , il existe une suite $(f_n)_{n \geq 0}$ de fonctions de $\mathcal{C}_c(\Omega)$ telle que $f_n(x) \rightarrow f(x)$ p.p. en $x \in \Omega$. En particulier $f_n \in \mathcal{L}^1(\Omega)$ pour tout $n \in \mathbf{N}$ car $\mathcal{C}_c(\Omega) \subset \mathcal{L}^1(\Omega)$. Mais comme, dans le Théorème 5.5 page 100 ou le Théorème 5.7 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 la convergence dominée, on a seulement $f \in \mathcal{L}^1(\Omega)$ sans pouvoir conclure que

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

Ce n'est pas gênant, dans la mesure où le Théorème 5.5 page 100 ou le Théorème 5.7 page précédente servent uniquement à vérifier que f est intégrable, et pas à calculer son intégrale!

On peut même aller plus loin dans le cas des fonctions mesurables à valeurs *positives*, c'est-à-dire dans $[0, +\infty]$.

Définition 5.9 (Intégration des fonctions mesurables positives). *Soit $f : \Omega \rightarrow [0, +\infty]$ une fonction mesurable définie p.p. sur Ω . Alors, par définition :*

- (i) soit $f \in \mathcal{L}^1(\Omega)$;
- (ii) soit $f \notin \mathcal{L}^1(\Omega)$, et auquel cas on pose alors

$$\int_{\Omega} f(x) dx := +\infty.$$

Autrement dit, d'après le Théorème 5.7 page précédente, soit il existe une fonction $F \in \mathcal{L}^+(\Omega)$ telle que $f \leq F$ p.p. sur Ω , soit $\int_{\Omega} f(x) dx := +\infty$.

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 (ii) de l'introduction (voir la section 1 page 69.)

Théorème 5.8 (Théorème de la convergence monotone de B. Levi). *Soit $(f_n)_{n \geq 0}$ une suite de fonctions mesurables sur Ω , à valeurs dans $[0, +\infty]$. Alors*

$$\int_{\Omega} \left(\sum_{n=0}^{\infty} f_n(x) \right) dx = \sum_{n=0}^{\infty} \int_{\Omega} f_n(x) dx \in [0, +\infty].$$

DÉMONSTRATION. Définissons

$$F_n := \sum_{k=0}^n f_k, \quad \text{et} \quad F := \sum_{k=0}^{\infty} f_k.$$

Évidemment, la série ci-dessus converge dans $[0, +\infty]$, et définit une fonction F mesurable sur Ω .

Si

$$\sum_{n=0}^{\infty} \int_{\Omega} f_n(x) dx < +\infty,$$

on déduit de la Proposition 5.9 page 89 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 (avec la définition ci-dessus)

$$\int_{\Omega} F(x) dx = +\infty = \sum_{n=0}^{\infty} \int_{\Omega} f_n(x) dx.$$

En effet, si l'on avait

$$\int_{\Omega} F(x) dx < +\infty,$$

alors, d'après la Définition 5.9 page ci-contre, $F \in \mathcal{L}^1(\Omega)$ et l'on aurait $F_n \rightarrow F$ p.p. sur Ω et $0 \leq F_n \leq F$ p.p. sur Ω , pour tout $n \geq 0$. Grâce au Théorème de la convergence dominée de Lebesgue (Théorème 5.2 page 89), on en déduirait que

$$\int_{\Omega} F_n(x) dx = \sum_{k=0}^n \int_{\Omega} f_k(x) dx \rightarrow \int_{\Omega} F(x) dx,$$

lorsque $n \rightarrow +\infty$, 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. Intégration des fonctions à valeurs complexes

On aura éventuellement besoin d'intégrer les fonctions à valeurs complexes (ou encore à valeurs dans un espace vectoriel de dimension finie).

Définition 5.10. On dit qu'une fonction f définie p.p. sur Ω à valeurs dans \mathbf{C} est intégrable si les fonctions à valeurs réelles $\Re f$ et $\Im f$ sont intégrables ; auquel cas, on pose

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

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

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

Proposition 5.12 (Intégration des fonctions à valeurs complexes). *L'intégrale de Lebesgue des fonctions à valeurs complexes vérifie les énoncés suivants :*

- (a) *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*

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

une forme \mathbf{C} -linéaire ;

- (b) *$\mathcal{C}_c(\Omega; \mathbf{C}) \subset \mathcal{L}^1(\Omega; \mathbf{C})$ et l'intégrale de Lebesgue coïncide avec l'intégrale usuelle sur $\mathcal{C}_c(\Omega; \mathbf{C})$;*

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

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

DÉMONSTRATION. Les énoncés (a) et (b) sont triviaux. Quant à l'énoncé (c), le fait que $|f| \in \mathcal{L}^1(\Omega)$ est une conséquence du Théorème 5.7 page 103. En effet, $|f| = \sqrt{|\Re f|^2 + |\Im f|^2}$ est une fonction mesurable d'après le point (d) de la Proposition 5.11 page 99, et $|f|$ est majorée par $|f| \leq |\Re f| + |\Im f|$. Or $|\Re f| + |\Im f| \in \mathcal{L}^1(\Omega)$ donc $|f| \in \mathcal{L}^1(\Omega)$.

Maintenant, choisissons $\xi \in \mathbf{C}$ tel que $|\xi| = 1$ et

$$\xi \int_{\Omega} f(x) dx = \left| \int_{\Omega} f(x) dx \right|.$$

Alors

$$\left| \int_{\Omega} f(x) dx \right| = \xi \int_{\Omega} f(x) dx = \int_{\Omega} \Re(\xi f(x)) dx + i \int_{\Omega} \Im(\xi f(x)) dx = \int_{\Omega} \Re(\xi f(x)) dx.$$

Puis, d'après les énoncés (c) et (d) de la Proposition 5.7 page 85,

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

d'où le résultat. \square

9. Intégrales paramétriques

Nous présentons ici des applications du Théorème de la convergence dominée de Lebesgue à l'étude des intégrales paramétriques.

Soient I un intervalle ouvert de \mathbf{R} et Ω un ouvert de \mathbf{R}^N , tous deux supposés non vides. Soit $f : I \times \Omega \rightarrow \mathbf{C}$ une fonction telle que, pour tout $t \in I$, la fonction $f(t, \cdot) \in \mathcal{L}^1(\Omega; \mathbf{C})$. Définissons, pour tout $t \in I$, la fonction F sur I par

$$F(t) := \int_{\Omega} f(t, x) dx.$$

Les deux résultats qui suivent donnent des conditions suffisantes pour que la fonction F soit continue ou bien dérivable. Leurs démonstrations reposent sur le Théorème de la convergence dominée de Lebesgue.

Théorème 5.9 (Continuité des intégrales paramétriques). *Soient I un intervalle ouvert de \mathbf{R} et Ω un ouvert de \mathbf{R}^N , supposés non vides. Soit $f : I \times \Omega \rightarrow \mathbf{C}$ telle que, pour tout $t \in I$, la fonction $f(t, \cdot) \in \mathcal{L}^1(\Omega; \mathbf{C})$ et, pour presque tout $x \in \Omega$, la fonction $t \mapsto f(t, x)$ est continue en $t_0 \in I$. On suppose de plus qu'il existe une fonction $\Phi \in \mathcal{L}^1(\Omega)$ telle que, pour presque tout $x \in \Omega$ et pour tout $t \in I$, on ait*

$$|f(t, x)| \leq \Phi(x).$$

Alors, la fonction F définie sur I par

$$F(t) := \int_{\Omega} f(t, x) dx,$$

est continue en t_0 et

$$\lim_{t \rightarrow t_0} \int_{\Omega} f(t, x) dx = \int_{\Omega} \lim_{t \rightarrow t_0} f(t, x) dx.$$

DÉMONSTRATION. La démonstration du résultat est une application directe du Théorème de la convergence dominée de Lebesgue. \square

Présentons maintenant le Théorème de dérivation sous le signe somme.

Théorème 5.10 (Dérivation sous le signe somme). *Soient I un intervalle ouvert de \mathbf{R} et Ω un ouvert de \mathbf{R}^N , tous deux supposés non vides. Soit $f : I \times \Omega \rightarrow \mathbf{C}$ telle que, pour tout $t \in I$, la fonction $f(t, \cdot) \in \mathcal{L}^1(\Omega; \mathbf{C})$ et, pour presque tout $x \in \Omega$, la fonction $t \mapsto f(t, x)$ est dérivable sur I . On suppose de plus qu'il existe $\Phi \in \mathcal{L}^1(\Omega)$ telle que, pour presque tout $x \in \Omega$ et pour tout $t \in I$, on ait*

$$\left| \frac{\partial f}{\partial t}(t, x) \right| \leq \Phi(x).$$

Alors, la fonction F définie sur I par

$$F(t) := \int_{\Omega} f(t, x) dx,$$

est dérivable sur I et sa dérivée est donnée par

$$F'(t) = \int_{\Omega} \frac{\partial f}{\partial t}(t, x) dx.$$

De plus, si pour presque tout $x \in \Omega$, la fonction $f(\cdot, x) \in \mathcal{C}^1(I; \mathbf{C})$ alors la fonction $F \in \mathcal{C}^1(I; \mathbf{C})$.

DÉMONSTRATION. Soit $(t_n)_{n \geq 0}$ suite de I qui converge vers $t \in I$. Étudions la suite

$$\frac{F(t_n) - F(t)}{t_n - t} = \int_{\Omega} \frac{f(t_n, x) - f(t, x)}{t_n - t} dx.$$

Par hypothèse, il existe $\mathcal{Z} \subset \Omega$ négligeable tel que

$$\lim_{n \rightarrow +\infty} \frac{f(t_n, x) - f(t, x)}{t_n - t} = \frac{\partial f}{\partial t}(t, x),$$

et, grâce au Théorème des accroissements finis,

$$\left| \frac{f(t_n, x) - f(t, x)}{t_n - t} \right| \leq \Phi(x),$$

pour tout $x \in \Omega - \mathcal{Z}$. En appliquant le Théorème de la convergence dominée de Lebesgue, on conclut que

$$\lim_{n \rightarrow +\infty} \frac{F(t_n) - F(t)}{t_n - t} = \lim_{n \rightarrow +\infty} \int_{\Omega} \frac{f(t_n, x) - f(t, x)}{t_n - t} dx = \int_{\Omega} \frac{\partial f}{\partial t}(t, x) dx.$$

On en déduit que F est dérivable sur I et que

$$F'(t) = \int_{\Omega} \frac{\partial f}{\partial t}(t, x) dx,$$

pour tout $t \in I$.

Supposons maintenant qu'il existe un sous-ensemble négligeable $\mathcal{Z}' \subset \Omega$ tel que $f(\cdot, x) \in \mathcal{C}^1(I; \mathbf{C})$ pour tout $x \in \Omega - \mathcal{Z}'$. Soit $t \in I$ et $(t_n)_{n \geq 0}$ une suite qui converge vers t . Alors, pour tout $x \in \Omega - \mathcal{Z}'$, on a

$$\lim_{n \rightarrow +\infty} \frac{\partial f}{\partial t}(t_n, x) = \frac{\partial f}{\partial t}(t, x),$$

et

$$\left| \frac{\partial f}{\partial t}(t_n, x) \right| \leq \Phi(x).$$

En appliquant une fois de plus le Théorème de la convergence dominée de Lebesgue, on conclut que

$$\lim_{n \rightarrow +\infty} F'(t_n) = \lim_{n \rightarrow +\infty} \int_{\Omega} \frac{\partial f}{\partial t}(t_n, x) dx = \int_{\Omega} \frac{\partial f}{\partial t}(t, x) dx = F'(t),$$

ce qui montre que F' est continue en t . Il s'ensuit que $F \in \mathcal{C}^1(I)$. □

CHAPITRE 6

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 intégrables 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 (Théorème de Fubini) ;
- (b) la formule du changement de variables dans les intégrales multiples.

Ce chapitre est une première occasion de mettre en œuvre les notions présentées 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.

Dans tout ce chapitre, Ω désigne un ouvert non vide de \mathbf{R}^N , avec $N \geq 1$.

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 intégrable.

1.1. Inégalité de Jensen. La plus fondamentale de toutes les inégalités intégrales basées sur la convexité est l'inégalité de Jensen. Commençons par rappeler quelques propriétés essentielles des fonctions convexes sur \mathbf{R} .

Lemme 6.1 (Fonctions convexes sur \mathbf{R}). *Soit I un intervalle ouvert de \mathbf{R} et soit $F : I \rightarrow \mathbf{R}$ une fonction 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) si $x < y$ appartiennent à I , alors $F'_d(x) \leq F'_g(y)$;

(e) pour tout $x_0 \in I$ et tout $a \in [F'_g(x_0), F'_d(x_0)]$, on a

$$F(x) \geq F(x_0) + a(x - x_0),$$

pour tout $x \in I$.

Lorsque F est dérivable en $x_0 \in I$, c'est-à-dire lorsque $F'_g(x_0) = F'_d(x_0)$, la propriété (d) 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.

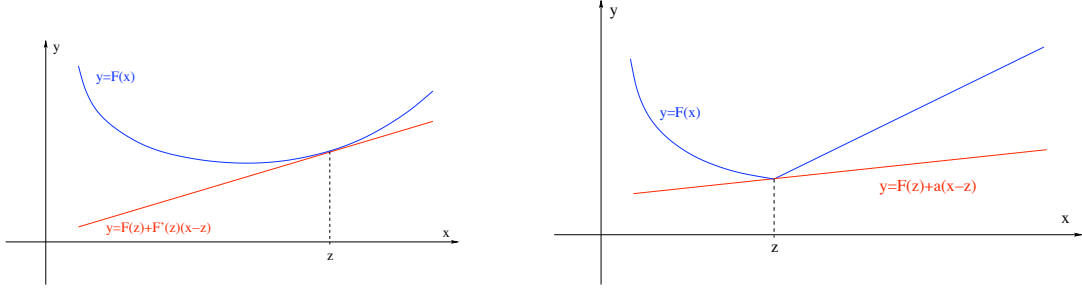


FIGURE 6.1. À gauche, le graphe de la fonction convexe F reste au-dessus de sa tangente au point d'abscisse z ; à droite, le graphe de la fonction convexe F reste au dessus de la droite de pente $a \in [F'_g(z), F'_d(z)]$ passant par le point d'abscisse $(z, F(z))$.

DÉMONSTRATION. Partons de l'inégalité de convexité

$$F((1 - \theta)x + \theta z) \leq (1 - \theta)F(x) + \theta F(z),$$

vérifiée pour tous $x < z \in I$ et tout $\theta \in [0, 1]$.

Si $x < y < z$ sont trois points de I , en prenant $\theta = \frac{y-x}{z-x}$, on trouve que

$$\frac{F(y) - F(x)}{y - x} \leq \frac{F(z) - F(x)}{z - x} \leq \frac{F(z) - F(y)}{z - y},$$

Ces inégalités permettent de montrer que la fonction $a \mapsto \frac{F(a) - F(x)}{a - x}$ est croissante sur $I - \{x\}$. Elle admet donc une limite à droite en x (qui est $F'_d(x)$ par définition) et une limite à gauche en x (qui est $F'_g(x)$ par définition). On trouve également que $F'_g(x) \leq F'_d(x)$, ce qui établit le (b) et la deuxième moitié de (c).

Réécrivons cette dernière inégalité, cette fois avec quatre points successifs $x < y < z < t$, en raisonnant successivement avec les triplets de points $x < y < z$, puis $y < z < t$. On trouve que

$$F'_d(x) \leq \frac{F(y) - F(x)}{y - x} \leq \frac{F(z) - F(y)}{z - y} \leq \frac{F(t) - F(z)}{t - z} \leq F'_g(t),$$

ce qui établit le (d). En particulier, F est lipschitzienne sur $[x, t]$ puisque

$$|F(z) - F(y)| \leq \max(|F'_d(x)|, |F'_g(t)|)|z - y|,$$

et donc, comme $x < t$ 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 $z \searrow y$, puis pour $y \nearrow z$ dans l'avant-dernière chaîne d'inégalités, on trouve que

$$F'_d(x) \leq \frac{F(y) - F(x)}{y - x} \leq F'_d(y), \quad \text{et} \quad F'_g(z) \leq \frac{F(t) - F(z)}{t - z} \leq F'_g(t),$$

d'où l'on tire le fait que F'_d et F'_g sont toutes les deux croissantes sur I , ce qui termine la démonstration de (c).

On en tire aussi les deux inégalités

$$F(y) \geq F(x) + F'_d(x)(y - x) \quad \text{et} \quad F(z) \geq F(t) + F'_g(t)(z - t),$$

que l'on peut réécrire sous la forme

$$F(y) \geq F(x) + F'_d(x)(y - x) \quad \text{si } x < y \in I,$$

$$F(y) \geq F(x) + F'_g(x)(y - x) \quad \text{si } x > y \in I,$$

qui se combinent pour donner le point (e). □

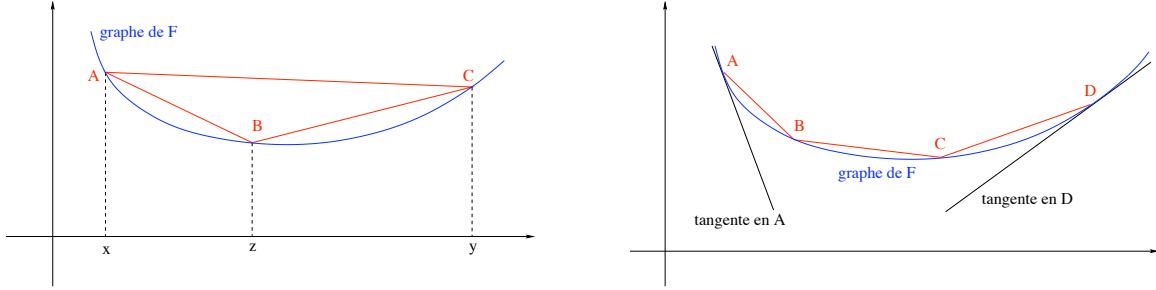


FIGURE 6.2. À gauche : les pentes des segments $[A, B]$, $[A, C]$ et $[B, C]$ sont ordonnées de façon croissante ; à droite : les pentes de la tangente au point A , des segments $[A, B]$, $[B, C]$, $[C, D]$ et de la tangente en D sont ordonnées de manière croissante.

Théorème 6.1 (Inégalité de Jensen). *Soient f, g deux fonctions mesurables définies p.p. sur Ω et à valeurs réelles, et soit $\Phi : \mathbf{R} \rightarrow \mathbf{R}$ une fonction convexe. Supposons que $g \geq 0$ p.p. sur Ω et que*

$$\int_{\Omega} g(x) dx = 1.$$

Si $f g$ et $\Phi(f) g \in \mathcal{L}^1(\Omega)$, alors

$$\Phi \left(\int_{\Omega} f(x) g(x) dx \right) \leq \int_{\Omega} \Phi(f(x)) g(x) dx.$$

DÉMONSTRATION. Posons

$$m = \int_{\Omega} f(x) g(x) dx,$$

et appliquons la propriété (e) du Lemme précédent au cas où $I = \mathbf{R}$, $F = \Phi$ et $x_0 = m$. On trouve alors que

$$\Phi(f(x)) - \Phi(m) - \Phi'_d(m)(f(x) - m) \geq 0,$$

pour presque tout $x \in \Omega$ de sorte que, en multipliant chaque membre de cette inégalité par $g(x)$ qui est ≥ 0 pour presque tout $x \in \Omega$, on trouve que

$$\Phi(f(x))g(x) - \Phi(m)g(x) - \Phi'_d(m)(f(x)g(x) - mg(x)) \geq 0,$$

pour presque tout $x \in \Omega$. Par hypothèse, le membre de gauche de cette inégalité est une fonction intégrable de x , qui est définie p.p. sur Ω (comme combinaison linéaire des fonctions intégrables $\Phi(f)g$, g et fg) et à valeurs positives. Son intégrale est donc un nombre positif et l'on a

$$\begin{aligned} 0 &\leq \int_{\Omega} (\Phi(f(x))g(x) - \Phi(m)g(x) - \Phi'_d(m)(f(x)g(x) - mg(x))) dx \\ &= \int_{\Omega} \Phi(f(x))g(x)dx - \Phi(m) \int_{\Omega} g(x)dx - \Phi'_d(m) \left(m - m \int_{\Omega} g(x)dx \right) \\ &= \int_{\Omega} \Phi(f(x))g(x)dx - \Phi(m), \end{aligned}$$

puisque

$$\int_{\Omega} g(x)dx = 1.$$

Ce qui donne l'inégalité cherchée. \square

Remarque 6.1. Dans le cas particulier où $f(x) = x$, on obtient l'inégalité

$$\Phi \left(\int_{\Omega} x g(x) dx \right) \leq \int_{\Omega} \Phi(x) g(x) dx,$$

que l'on retrouve dans le cours de probabilités, sous la forme

$$\Phi(\mathbb{E}(X)) \leq \mathbb{E}(\Phi(X)).$$

où X est une variable aléatoire et $g(x) dx$ une mesure de probabilité.

1.2. Inégalité de Hölder. L'inégalité de Hölder est une généralisation de l'inégalité de Cauchy-Schwarz (sur laquelle nous reviendrons dans le chapitre sur les espaces de Hilbert). On s'en sert le plus souvent pour montrer qu'un produit fg de fonctions mesurables est intégrable, pourvu que des puissances bien choisies de f et de g soient intégrables.

Théorème 6.2 (Inégalité de Hölder). *Soient $f, g : \Omega \rightarrow \mathbf{R}$ deux fonctions mesurables, et soient $p, q > 1$ tels que $\frac{1}{p} + \frac{1}{q} = 1$ (on dit que les exposants p et q sont conjugués). Alors*

$$\int_{\Omega} |f(x)g(x)|dx \leq \left(\int_{\Omega} |f(x)|^p dx \right)^{1/p} \left(\int_{\Omega} |g(x)|^q dx \right)^{1/q}.$$

Avant de démontrer ce résultat, établissons un résultat intermédiaire.

Lemme 6.2 (Une inégalité auxiliaire). *Soit $p > 1$ et $q = \frac{p}{p-1}$, de sorte que $\frac{1}{p} + \frac{1}{q} = 1$. Alors, pour tous $X, Y > 0$, on a*

$$X^{1/p} Y^{1/q} \leq \frac{X}{p} + \frac{Y}{q}.$$

DÉMONSTRATION. On utilise la concavité de la fonction logarithme (ou bien la convexité de la fonction $-\ln$) pour conclure que $\frac{1}{p} \ln X + \frac{1}{q} \ln Y \leq \ln \left(\frac{X}{p} + \frac{Y}{q} \right)$. \square

Grâce à ce résultat, pour pouvons maintenant passer à la démonstration de l'inégalité de Hölder.

DÉMONSTRATION DE L'INÉGALITÉ DE HÖLDER. On note

$$I := \left(\int_{\Omega} |f(x)|^p dx \right)^{1/p} \quad \text{et} \quad J := \left(\int_{\Omega} |g(x)|^q dx \right)^{1/q}.$$

Supposons pour commencer que $0 < I, J < +\infty$ et écrivons alors l'inégalité obtenue dans le Lemme précédent pour $X = |f(x)|/I$ et $Y = |g(x)|/J$. On aboutit à

$$\frac{|f||g|}{IJ} \leq \frac{1}{p} \frac{|f|^p}{I^p} + \frac{1}{q} \frac{|g|^q}{J^q},$$

p.p. sur Ω . Chaque membre de cette inégalité est une fonction mesurable positive de la fonction x ; d'autre part, le membre de droite est intégrable comme combinaison linéaire des fonctions $|f|^p$ et $|g|^q$ qui, pour l'instant, sont supposées intégrable. D'après le théorème de comparaison (voir le Théorème 5.7 page 103), on conclut que $|f g|$ est également intégrable sur Ω et que

$$\frac{1}{IJ} \int_{\Omega} |f(x)||g(x)| dx \leq \frac{1}{p} \frac{1}{I^p} \int_{\Omega} |f(x)|^p dx + \frac{1}{q} \frac{1}{J^q} \int_{\Omega} |g(x)|^q dx.$$

Or, dans cette inégalité, le membre de droite est égal à

$$\frac{1}{p} \frac{1}{I^p} I^p + \frac{1}{q} \frac{1}{J^q} J^q = \frac{1}{p} + \frac{1}{q} = 1.$$

Nous avons donc

$$\int_{\Omega} |f(x)||g(x)| dx \leq IJ,$$

ce qui est l'inégalité cherchée.

Dans le cas où $I > 0$ et $J > 0$ et où 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. \square

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 très important en analyse. Nous la retrouverons au Chapitre 8 page 155, où elle nous servira à définir certains espaces de fonctions généralisant l'espace $\mathcal{L}^1(\Omega)$ des fonctions intégrables sur Ω .

Théorème 6.3 (Inégalité de Minkowski). *Soient $f, g : \Omega \rightarrow \mathbf{R}$ deux fonctions mesurables, et $p \in [1, +\infty[$. Alors*

$$\left(\int_{\Omega} |f(x) + g(x)|^p dx \right)^{1/p} \leq \left(\int_{\Omega} |f(x)|^p dx \right)^{1/p} + \left(\int_{\Omega} |g(x)|^p dx \right)^{1/p}$$

DÉMONSTRATION. Le cas $p = 1$ résulte immédiatement de l'inégalité triangulaire $|f + g| \leq |f| + |g|$, p.p. sur Ω .

Supposons donc que $p > 1$ et écrivons que

$$\int_{\Omega} |f(x) + g(x)|^p dx \leq \int_{\Omega} |f(x)| |f(x) + g(x)|^{p-1} dx + \int_{\Omega} |g(x)| |f(x) + g(x)|^{p-1} dx = A + B.$$

D'après l'inégalité de Hölder

$$A \leq \left(\int_{\Omega} |f(x)|^p dx \right)^{1/p} \left(\int_{\Omega} |f(x) + g(x)|^p dx \right)^{(p-1)/p},$$

puisque $q = \frac{p}{p-1}$ et $\frac{1}{p} + \frac{1}{q} = 1$.

De même

$$B \leq \left(\int_{\Omega} |g(x)|^p dx \right)^{1/p} \left(\int_{\Omega} |f(x) + g(x)|^p dx \right)^{(p-1)/p},$$

de sorte que

$$\begin{aligned} \int_{\Omega} |f(x) + g(x)|^p dx &\leq \left(\int_{\Omega} |f(x) + g(x)|^p dx \right)^{(p-1)/p} \\ &\quad \times \left(\left(\int_{\Omega} |f(x)|^p dx \right)^{1/p} + \left(\int_{\Omega} |g(x)|^p dx \right)^{1/p} \right). \end{aligned}$$

De deux choses l'une, ou bien

$$\int_{\Omega} |f(x) + g(x)|^p dx = 0,$$

et dans ce cas, 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

$$\left(\int_{\Omega} |f(x) + g(x)|^p dx \right)^{(p-1)/p} > 0.$$

Dans tous les cas l'inégalité est démontrée. \square

2. Intégrales multiples

Commençons par rappeler le résultat classique sur l'interversion de l'ordre des intégrations successives dans une intégrale double. Soit $f \in \mathcal{C}(K \times K)$, où K est un segment de \mathbf{R} . Alors :

(a) les fonctions

$$x \mapsto \int_K f(x, y) dy, \quad \text{et} \quad y \mapsto \int_K f(x, y) dx,$$

sont continues sur K ;

(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) 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 5 page 69, c'est-à-dire que

$$\iint_{K \times K} f(x, y) dx dy = \int_K \left(\int_K f(x, y) dy \right) dx = \int_K \left(\int_K f(x, y) dx \right) dy.$$

Plus généralement, on dispose du résultat suivant.

Théorème 6.4. Soient $\Omega_1 \subset \mathbf{R}^{N_1}$ et $\Omega_2 \subset \mathbf{R}^{N_2}$ deux ouverts non vides, et $f \in \mathcal{C}_c(\Omega_1 \times \Omega_2)$. Alors, la fonction

$$x_1 \mapsto \int_{\Omega_2} f(x_1, x_2) dx_2,$$

est bien définie sur Ω_1 et appartient à $\mathcal{C}_c(\Omega_1)$. Enfin,

$$\iint_{\Omega_1 \times \Omega_2} f(x_1, x_2) dx_1 dx_2 = \int_{\Omega_1} \left(\int_{\Omega_2} f(x_1, x_2) dx_2 \right) dx_1 = \int_{\Omega_2} \left(\int_{\Omega_1} f(x_1, x_2) dx_1 \right) dx_2.$$

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 6.5 (Théorème de Fubini). Soient $\Omega_1 \subset \mathbf{R}^{N_1}$ et $\Omega_2 \subset \mathbf{R}^{N_2}$ deux ouverts non vides et $f \in \mathcal{L}^1(\Omega_1 \times \Omega_2)$. Alors :

- (a) pour presque tout $x_2 \in \Omega_2$, la fonction $x_1 \mapsto f(x_1, x_2)$ appartient à $\mathcal{L}^1(\Omega_1)$;
- (b) la fonction définie p.p. sur Ω_2 par $x_2 \mapsto \int_{\Omega_1} f(x_1, x_2) dx_1$ appartient à $\mathcal{L}^1(\Omega_2)$;
- (c) on a l'égalité

$$\iint_{\Omega_1 \times \Omega_2} f(x_1, x_2) dx_1 dx_2 = \int_{\Omega_1} \left(\int_{\Omega_2} f(x_1, x_2) dx_2 \right) dx_1 = \int_{\Omega_2} \left(\int_{\Omega_1} f(x_1, x_2) dx_1 \right) dx_2.$$

Nous allons démontrer le Théorème de Fubini pour les fonctions intégrables 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 de Levi.

Lemme 6.3. Soit $f \in \mathcal{L}^+(\Omega_1 \times \Omega_2)$. Alors :

- (a) pour presque tout $x_2 \in \Omega_2$, la fonction $x_1 \mapsto f(x_1, x_2)$ appartient à $\mathcal{L}^+(\Omega_1)$;
- (b) la fonction définie p.p. sur Ω_2 par $x_2 \mapsto \int_{\Omega_1} f(x_1, x_2) dx_1$ est p.p. égale à une fonction qui appartient à $\mathcal{L}^+(\Omega_2)$;
- (c) on a

$$\iint_{\Omega_1 \times \Omega_2} f(x_1, x_2) dx_1 dx_2 = \int_{\Omega_1} \left(\int_{\Omega_2} f(x_1, x_2) dx_2 \right) dx_1 = \int_{\Omega_2} \left(\int_{\Omega_1} f(x_1, x_2) dx_1 \right) dx_2.$$

DÉMONSTRATION. Soit $(f_n)_{n \geq 0}$ suite de Levi convergeant simplement vers f sur $\Omega_1 \times \Omega_2$.

Étape 1 : Pour tout $n \geq 0$ et pour tout $x_2 \in \Omega_2$, posons

$$F_n(x_2) := \int_{\Omega_1} f_n(x_1, x_2) dx_1.$$

Comme $f_n \in \mathcal{C}_c(\Omega_1 \times \Omega_2)$, la fonction $F_n \in \mathcal{C}_c(\Omega_2)$, d'après le résultat sur les fonctions continues à support compact rappelé avant l'énoncé du Théorème de Fubini. De plus, la suite

$(f_n)_{n \geq 0}$ est croissante sur $\Omega_1 \times \Omega_2$, ce qui implique que la suite $(F_n)_{n \geq 0}$ est croissante sur Ω_2 . Enfin

$$\int_{\Omega_2} F_n(x_2) dx_2 = \iint_{\Omega_1 \times \Omega_2} f_n(x_1, x_2) dx_1 dx_2 \leq \iint_{\Omega_1 \times \Omega_2} f(x_1, x_2) dx_1 dx_2 < +\infty,$$

de sorte que la suite $(F_n)_{n \geq 0}$ est une suite de Levi sur Ω_2 . On note F la limite simple sur Ω_2 de la suite $(F_n)_{n \geq 0}$. Par définition $F \in \mathcal{L}^+(\Omega_2)$ et

$$\int_{\Omega_2} F(x_2) dx_2 = \lim_{n \rightarrow +\infty} \int_{\Omega_2} F_n(x_2) dx_2.$$

Étape 2 : Soit $\mathcal{Z} := \{x_2 \in \Omega_2 : F(x_2) = +\infty\}$. Étant donné que $F \in \mathcal{L}^+(\Omega_2)$, l'ensemble \mathcal{Z} est négligeable. Remarquons que, pour tout $x_2 \in \Omega_2 - \mathcal{Z}$, la suite de fonctions $(f_n(\cdot, x_2))_{n \geq 0}$ est une suite de Levi sur Ω_1 . En effet, cette suite est clairement croissante et de plus, pour tout $n \geq 0$

$$\int_{\Omega_2} f_n(x_1, x_2) dx_1 = F_n(x_2) \leq F(x_2) < \infty,$$

par définition de \mathcal{Z} . Enfin, $(f_n)_{n \geq 0}$ converge simplement vers f sur $\Omega_1 \times \Omega_2$, on en déduit que $f(\cdot, x_2) \in \mathcal{L}^+(\Omega_1)$ pour tout $x_2 \in \Omega_2 - \mathcal{Z}$, c'est-à-dire pour presque tout $x_2 \in \Omega_2$, ce qui établit le point (a).

Étape 3 : Par convergence monotone

$$\int_{\Omega_1} f(x_1, x_2) dx_1 = \lim_{n \rightarrow +\infty} \int_{\Omega_1} f_n(x_1, x_2) dx_1 = \lim_{n \rightarrow +\infty} F_n(x_2) = F(x_2),$$

pour tout $x_2 \in \Omega_2 - \mathcal{Z}$. Or, on a vu plus haut que $F \in \mathcal{L}^+(\Omega_2)$, de sorte que la fonction

$$x_2 \mapsto \int_{\Omega_1} f(x_1, x_2) dx_1,$$

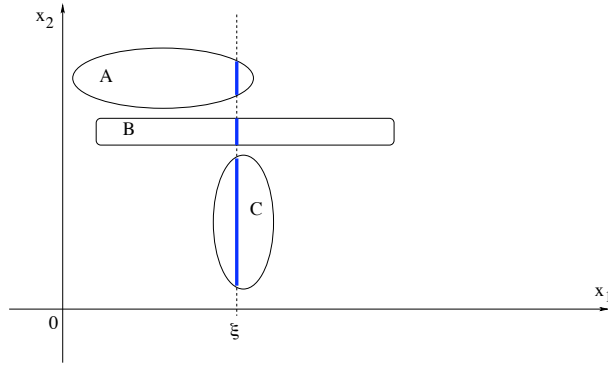
définie sur $\Omega_2 - \mathcal{Z}$, est p.p. égale à une fonction qui appartient à $\mathcal{L}^+(\Omega_2)$, ce qui établit le point (b).

Étape 4 : Enfin

$$\begin{aligned} \iint_{\Omega_1 \times \Omega_2} f(x_1, x_2) dx_1 dx_2 &= \lim_{n \rightarrow +\infty} \iint_{\Omega_1 \times \Omega_2} f_n(x_1, x_2) dx_1 dx_2 \\ &= \lim_{n \rightarrow +\infty} \int_{\Omega_2} \left(\int_{\Omega_1} f_n(x_1, x_2) dx_1 \right) dx_2 \\ &= \lim_{n \rightarrow +\infty} \int_{\Omega_2} F_n(x_2) dx_2 \\ &= \int_{\Omega_2} F(x_2) dx_2 = \int_{\Omega_2} \left(\int_{\Omega_1} f(x_1, x_2) dx_1 \right) dx_2. \end{aligned}$$

En effet, la première égalité découle du fait que $(f_n)_{n \geq 0}$ est une suite de Levi 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 troisième de la définition de F_n , la quatrième de l'étape 1 ci-dessus et la cinquième de l'égalité obtenue dans l'étape 3 ci-dessus. \square

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 $\Omega_1 \times \Omega_2$.

FIGURE 6.3. La fibre de $A \cup B \cup C$ au-dessus de ξ (en trait bleu).

Lemme 6.4 (Fibres des ensembles négligeables). *Soit $\mathcal{Z} \subset \Omega_1 \times \Omega_2$ un ensemble négligeable. Pour $x_1 \in \Omega_1$, on note*

$$\mathcal{Z}_{x_1} := \{x_2 \in \Omega_2 : (x_1, x_2) \in \mathcal{Z}\} \subset \Omega_2$$

la fibre de \mathcal{Z} au-dessus de x_1 (voir la Figure 6.3). Alors, pour presque tout $x_1 \in \Omega_1$, \mathcal{Z}_{x_1} est négligeable dans Ω_2 .

DÉMONSTRATION. Soit $f \in \mathcal{L}^+(\Omega_1 \times \Omega_2)$ telle que $f(x_1, x_2) = +\infty$ pour tout $(x_1, x_2) \in \mathcal{Z}$.

D'après le point (a) du Lemme 6.3 page 115 (en échangeant les rôles des variables x_1 et x_2), il existe $\mathcal{Z}_1 \subset \Omega_1$ négligeable tel que, pour $x_1 \in \Omega_1 - \mathcal{Z}_1$, la fonction $f(x_1, \cdot) \in \mathcal{L}^+(\Omega_2)$. Soit $x_1 \in \Omega - \mathcal{Z}_1$, par définition,

$$\mathcal{Z}_{x_1} \subset \{x_2 \in \Omega_2 : f(x_1, x_2) = +\infty\},$$

et $f(x_1, \cdot) \in \mathcal{L}^+(\Omega_2)$, donc l'ensemble \mathcal{Z}_{x_1} est négligeable dans Ω_2 . Conclusion, \mathcal{Z}_{x_1} est négligeable pour presque tout $x_1 \in \Omega_1$. \square

On prendra bien garde au fait qu'en général, les fibres \mathcal{Z}_{x_1} de l'ensemble négligeable \mathcal{Z} ne sont négligeables dans Ω_2 que pour presque tout $x_1 \in \Omega_1$, et pas pour tout $x_1 \in \Omega_1$. Le lecteur est invité à trouver un contre-exemple à titre d'exercice.

DÉMONSTRATION DU THÉORÈME DE FUBINI. Étant donné que $f \in \mathcal{L}^1(\Omega_1 \times \Omega_2)$, 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 un ensemble $\mathcal{Z} \subset \Omega_1 \times \Omega_2$ négligeable tel que $f = g - h$ sur $\Omega_1 \times \Omega_2 - \mathcal{Z}$.

Appliquons le point (a) du Lemme 6.3 page 115 aux fonctions g et h . Il existe $\mathcal{Z}'_2 \subset \Omega_2$ négligeable tel que les fonctions $g(\cdot, x_2)$ et $h(\cdot, x_2)$ appartiennent à $\mathcal{L}^+(\Omega_1)$ pour tout $x_2 \in \Omega_2 - \mathcal{Z}'_2$; d'après le Lemme 6.4, il existe $\mathcal{Z}_2 \subset \Omega_2$ négligeable tel que la fibre \mathcal{Z}_{x_2} soit négligeable dans Ω_1 pour tout $x_2 \in \Omega_2 - \mathcal{Z}_2$. Alors, pour tout $x_2 \in \Omega_2 - (\mathcal{Z}'_2 \cup \mathcal{Z}_2)$, on a

$$f(\cdot, x_2) = g(\cdot, x_2) - h(\cdot, x_2) \in \mathcal{L}^1(\Omega_1),$$

ce qui établit le point (a).

D'après le point (b) du Lemme 6.3 page 115, la fonction définie sur $\Omega_2 - (\mathcal{Z}'_2 \cup \mathcal{Z}_2)$ par

$$x_2 \mapsto \int_{\Omega_1} g(x_1, x_2) dx_1,$$

et la fonction définie sur $\Omega_2 - (\mathcal{Z}'_2 \cup \mathcal{Z}_2)$ par

$$x_2 \mapsto \int_{\Omega_1} h(x_1, x_2) dx_1,$$

sont presque partout égale à des fonctions qui appartiennent à $\mathcal{L}^+(\Omega_2)$, de sorte que la fonction définie sur $\Omega_2 - (\mathcal{Z}'_2 \cup \mathcal{Z}_2)$ par

$$x_2 \mapsto \int_{\Omega_1} g(x_1, x_2) dx_1 - \int_{\Omega_1} h(x_1, x_2) dx_1 = \int_{\Omega_1} f(x_1, x_2) dx_1,$$

appartient à $\mathcal{L}^1(\Omega_2)$, ce qui établit le point (b).

Enfin, d'après le point (c) du Lemme 6.3 page 115, on trouve que

$$\begin{aligned} \iint_{\Omega_1 \times \Omega_2} f(x_1, x_2) dx_1 dx_2 &= \iint_{\Omega_1 \times \Omega_2} g(x_1, x_2) dx_1 dx_2 - \iint_{\Omega_1 \times \Omega_2} h(x_1, x_2) dx_1 dx_2 \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} g(x_1, x_2) dx_1 \right) dx_2 - \int_{\Omega_2} \left(\int_{\Omega_1} h(x_1, x_2) dx_1 \right) dx_2 \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} f(x_1, x_2) dx_1 \right) dx_2, \end{aligned}$$

l'égalité restante étant obtenue en échangeant les variables x_1 et x_2 . \square

Comme pour le Théorème de la 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 6.6 (Théorème de Tonelli). *Soient $\Omega_1 \subset \mathbf{R}^{N_1}$ et $\Omega_2 \subset \mathbf{R}^{N_2}$ deux ouverts non vides et $f : \Omega_1 \times \Omega_2 \rightarrow [0, +\infty]$ une fonction mesurable. Alors :*

(a) *pour presque tout $x_2 \in \Omega_2$, la fonction $x_1 \mapsto f(x_1, x_2) \in [0, +\infty]$ définie p.p. sur Ω_1 est mesurable ;*

(b) *la fonction $x_2 \mapsto \int_{\Omega_1} f(x_1, x_2) dx_1 \in [0, +\infty]$, définie p.p. sur Ω_2 , est mesurable ;*

(c) *on a l'égalité dans $[0, +\infty]$*

$$\iint_{\Omega_1 \times \Omega_2} f(x_1, x_2) dx_1 dx_2 = \int_{\Omega_1} \left(\int_{\Omega_2} f(x_1, x_2) dx_2 \right) dx_1 = \int_{\Omega_2} \left(\int_{\Omega_1} f(x_1, x_2) dx_1 \right) dx_2.$$

DÉMONSTRATION. Pour tout $n \geq 1$, on note

$$U_n := (B(0, n) \times B(0, n)) \cap (\Omega_1 \times \Omega_2),$$

et l'on définit

$$F_n(x_1, x_2) := n \mathbf{1}_{U_n}(x_1, x_2),$$

qui appartient à $\mathcal{L}^1(\Omega_1 \times \Omega_2)$.

Pour tout $n \geq 1$, définissons alors $f_n := \min(f, F_n)$. Par construction, f_n est une fonction mesurable définie p.p. sur $\Omega_1 \times \Omega_2$ et à valeurs dans \mathbf{R}_+ . Notons $\mathcal{Z}' \subset \Omega_1 \times \Omega_2$ l'ensemble négligeable en dehors duquel la fonction mesurable f est définie. Alors $f_n \leq F_n$ sur $\Omega_1 \times \Omega_2 - \mathcal{Z}'$, de sorte que, d'après le théorème de comparaison (voir le Théorème 5.7 page 103), $f_n \in \mathcal{L}^1(\Omega_1 \times \Omega_2)$.

Le Théorème de Fubini appliqué à la fonction f_n nous permet d'affirmer qu'il existe un ensemble $\mathcal{Z}_{2,n} \subset \Omega_2$ négligeable tel que $f_n(\cdot, x_2) \in \mathcal{L}^1(\Omega_1)$ pour tout $x_2 \in \Omega_2 - \mathcal{Z}_{2,n}$, et tel que

la fonction définie sur $\Omega_2 - \mathcal{Z}_{2,n}$ par

$$x_2 \mapsto \int_{\Omega_1} f_n(x_1, x_2) dx_1,$$

appartienne à $\mathcal{L}^1(\Omega_1)$.

Posons

$$\mathcal{Z}_2 := \bigcup_{n \geq 0} \mathcal{Z}_{2,n}.$$

L'ensemble $\mathcal{Z}_2 \subset \Omega_2$ est négligeable comme réunion dénombrable d'ensembles négligeables. D'après le Lemme 6.4 page 117, il existe un ensemble $\mathcal{Z}'_2 \subset \Omega_2$ négligeable tel que, pour tout $x_2 \in \Omega_2 - \mathcal{Z}'_2$, la fibre \mathcal{Z}'_{x_2} , de l'ensemble négligeable \mathcal{Z}' en dehors duquel la fonction f est définie, soit négligeable dans Ω_1 .

Alors, pour tout $x_2 \in \Omega_2 - (\mathcal{Z}_2 \cup \mathcal{Z}'_2)$, 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 $x_2 \in \Omega_2 - (\mathcal{Z}_2 \cup \mathcal{Z}'_2)$ (comme limite p.p. d'une suite de fonctions mesurables sur Ω_1), c'est-à-dire pour presque tout $x_2 \in \Omega_2$. Ceci établit le point (a) du Théorème de Tonelli.

De plus, toujours par construction, la suite $(f_n)_{n \geq 0}$ est croissante sur $\Omega_1 \times \Omega_2 - \mathcal{Z}'$ de sorte que la suite $(f_n(\cdot, x_2))_{n \geq 0}$ est croissante p.p. sur Ω_1 , dès que $x_2 \in \Omega_2 - \mathcal{Z}'_2$. Alors, par convergence monotone

$$\lim_{n \rightarrow +\infty} \int_{\Omega_1} f_n(x_1, x_2) dx_1 = \int_{\Omega_1} f(x_1, x_2) dx_1,$$

pour tout $x_2 \in \Omega_2 - (\mathcal{Z}_2 \cup \mathcal{Z}'_2)$. Ainsi, la fonction définie sur $\Omega_2 - (\mathcal{Z}_2 \cup \mathcal{Z}'_2)$ par

$$x_2 \mapsto \int_{\Omega_1} f(x_1, x_2) dx_1,$$

est mesurable sur Ω_2 comme limite p.p. d'une suite de fonctions intégrables 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 la suite $(f_n)_{n \geq 0}$ est p.p. croissante sur $\Omega_1 \times \Omega_2$ et que

$$\left(\int_{\Omega_1} f_n(x_1, x_2) dx_1 \right)_{n \geq 0},$$

est croissante p.p. sur Ω_2 et converge p.p. vers $\int_{\Omega_1} f(x_1, x_2) dx_1$. Donc, en vertu du Théorème de la convergence monotone, on obtient l'égalité suivante dans $[0, +\infty]$:

$$\begin{aligned} \iint_{\Omega_1 \times \Omega_2} f(x_1, x_2) dx_1 dx_2 &= \lim_{n \rightarrow +\infty} \iint_{\Omega_1 \times \Omega_2} f_n(x_1, x_2) dx_1 dx_2 \\ &= \lim_{n \rightarrow +\infty} \int_{\Omega_2} \left(\int_{\Omega_1} f_n(x_1, x_2) dx_1 \right) dx_2 \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} f(x_1, x_2) dx_1 \right) dx_2. \end{aligned}$$

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 intégrables. 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 intégrable sur $\Omega_1 \times \Omega_2$?*

Étant donnée une fonction $f : \Omega_1 \times \Omega_2 \rightarrow \mathbf{C}$ 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. Changements 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 non vides de \mathbf{R}^N et Φ un \mathcal{C}^1 -difféomorphisme de Ω_1 sur Ω_2 . Rappelons que ceci signifie que Φ est une bijection de classe \mathcal{C}^1 sur Ω_1 dont la réciproque Φ^{-1} est de classe \mathcal{C}^1 sur Ω_2 .

En pratique, on cherchera à éviter de vérifier que Φ^{-1} est de classe \mathcal{C}^1 sur Ω_2 , car il n'est pas toujours possible de calculer explicitement Φ^{-1} . On utilise alors la caractérisation suivante : l'application

$$\Phi : \Omega_1 \rightarrow \Omega_2,$$

est un \mathcal{C}^1 -difféomorphisme si et seulement si Φ est une bijection de classe \mathcal{C}^1 sur Ω_1 et $\det J_\Phi(x) \neq 0$ pour tout $x \in \Omega_1$ où $J_\Phi(x)$ désigne la matrice jacobienne de Φ au point $x \in \Omega_1$.

Rappelons la formule du changement de variables dans le cadre de l'intégrale usuelle des fonctions continues. Pour tout $f \in \mathcal{C}_c(\Omega_2)$, on a $f \circ \Phi \in \mathcal{C}_c(\Omega_1)$ 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 intégrables.

Théorème 6.7 (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 l'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 6.5. *Soit $f \in \mathcal{L}^+(\Omega_2)$. Alors, la fonction $f \circ \Phi | \det J_\Phi | \in \mathcal{L}^+(\Omega_1)$ et l'on a la formule*

$$\int_{\Omega_2} f(y) dy = \int_{\Omega_1} f(\Phi(x)) | \det J_\Phi(x) | dx.$$

DÉMONSTRATION. Soit $(f_n)_{n \geq 0}$ une suite de Levi convergeant simplement vers f sur Ω_2 .

Alors, pour tout $n \geq 0$, la fonction $f_n \circ \Phi \in \mathcal{C}_c(\Omega_1)$ puisque $f_n \in \mathcal{C}_c(\Omega_2)$. De plus la suite $(f_n)_{n \geq 0}$ étant croissante sur Ω_2 , on en déduit que la suite $(f_n \circ \Phi | \det J_\Phi |)_{n \geq 0}$ est croissante sur Ω_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 $(f_n \circ \Phi | \det J_\Phi |)_{n \geq 0}$ est une suite de Levi sur Ω_1 , et elle 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

$$\begin{aligned} \int_{\Omega_2} f(y) dy &= \lim_{n \rightarrow +\infty} \int_{\Omega_2} f_n(y) dy \\ &= \lim_{n \rightarrow +\infty} \int_{\Omega_1} f_n(\Phi(x)) | \det J_\Phi(x) | dx \\ &= \int_{\Omega_1} f(\Phi(x)) | \det J_\Phi(x) | dx. \end{aligned}$$

Ce qui termine la démonstration. \square

Une conséquence immédiate du lemme ci-dessus est le fait que les \mathcal{C}^1 difféomorphismes transforment les ensembles négligeables en ensembles négligeables.

Lemme 6.6. *Si $\mathcal{Z} \subset \Omega_2$ est négligeable, alors $\Phi^{-1}(\mathcal{Z}) \subset \Omega_1$ est négligeable.*

DÉMONSTRATION. Si $\mathcal{Z} \subset \Omega_2$ est négligeable, il existe $f \in \mathcal{L}^+(\Omega_2)$ telle que $f(y) = +\infty$ pour tout $y \in \mathcal{Z}$. Donc la fonction $f \circ \Phi$ vérifie $f \circ \Phi(x) = +\infty$ pour tout $x \in \Phi^{-1}(\mathcal{Z})$. 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 $x \in \Phi^{-1}(\mathcal{Z})$. Or, la fonction $f \circ \Phi | \det J_\Phi | \in \mathcal{L}^+(\Omega_1)$ d'après le Lemme 6.5, donc $\Phi^{-1}(\mathcal{Z})$ est négligeable. \square

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 $\mathcal{Z} \subset \Omega_2$ un ensemble négligeable et $g, h \in \mathcal{L}^+(\Omega_2)$ tels que $f = g - h$ sur $\Omega_2 - \mathcal{Z}$. D'après le Lemme 6.5, les fonctions $g \circ \Phi | \det J_\Phi |$ et $h \circ \Phi | \det J_\Phi |$ appartiennent à $\mathcal{L}^+(\Omega_1)$ et

$$f \circ \Phi | \det J_\Phi | = g \circ \Phi | \det J_\Phi | - h \circ \Phi | \det J_\Phi |,$$

sur $\Omega_1 - \Phi^{-1}(\mathcal{Z})$, c'est-à-dire p.p. sur Ω_1 , puisque $\Phi^{-1}(\mathcal{Z})$ est négligeable d'après le Lemme 6.6. Par conséquent, $f \circ \Phi | \det J_\Phi | \in \mathcal{L}^1(\Omega_1)$.

Enfin, en utilisant à nouveau le Lemme 6.5 page précédente,

$$\begin{aligned}
 \int_{\Omega_1} f \circ \Phi(x) |\det J_\Phi(x)| dx &= \int_{\Omega_1} g \circ \Phi(x) |\det J_\Phi(x)| dx - \int_{\Omega_1} h \circ \Phi(x) |\det J_\Phi(x)| dx \\
 &= \int_{\Omega_2} g(y) dy - \int_{\Omega_2} h(y) dy \\
 &= \int_{\Omega_2} f(y) dy.
 \end{aligned}$$

□

Terminons cette présentation de la formule du changement de variables dans l'intégrale de Lebesgue par quelques exemples que l'on rencontre fréquemment.

Exemple 1 : changements de variables affines

Soit $A \in GL_N(\mathbf{R})^1$ et $b \in \mathbf{R}^N$. On pose $\Phi(x) = Ax + b$. Alors Φ est de classe \mathcal{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 \mathcal{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 l'on a

$$\int_{\mathbf{R}^N} f(Ax + b) dx = \frac{1}{|\det A|} \int_{\mathbf{R}^N} f(y) dy.$$

Le cas des isométries affines : si de plus $AA^T = A^T A = I$, alors Φ est une isométrie affine et $|\det A| = 1$. Donc

$$\int_{\mathbf{R}^N} f(Ax + b) dx = \int_{\mathbf{R}^N} f(y) dy.$$

Le 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

$$\int_{\mathbf{R}^N} f(\lambda x + b) dx = \frac{1}{|\lambda|^N} \int_{\mathbf{R}^N} f(y) dy.$$

Exemple 6.1. En effectuant un changement de variables (coordonnées polaires), on calcule

$$\int_{\mathbf{R}^2} e^{-(x^2+y^2)} dx dy = 2\pi \int_0^\infty e^{-r^2} r dr = \pi.$$

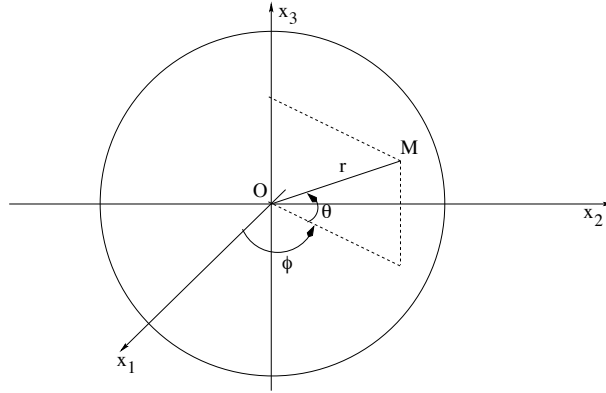
D'autre part, en utilisant le Théorème de Fubini, on trouve

$$\int_{\mathbf{R}^2} e^{-(x^2+y^2)} dx dy = \left(\int_{\mathbf{R}} e^{-x^2} dx \right) \left(\int_{\mathbf{R}} e^{-y^2} dy \right),$$

ce qui permet finalement de conclure que

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

1. On note $GL_N(\mathbf{R})$ l'ensemble des matrices carrées inversibles à N lignes et N colonnes et à coefficients réels.

FIGURE 6.4. Les coordonnées sphériques (r, ϕ, θ) du point $M \in \mathbf{R}^3$.**Exemple 2 : coordonnées sphériques**

Soit

$$\Phi : \mathbf{R}_+^* \times]-\pi, \pi[\times]-\frac{\pi}{2}, \frac{\pi}{2}[\rightarrow \mathbf{R}^3 - P_-,$$

définie par

$$(6.1) \quad \Phi(r, \phi, \theta) := (r \cos \phi \cos \theta, r \sin \phi \cos \theta, r \sin \theta),$$

où $P_- := \{x = (x_1, x_2, x_3) \in \mathbf{R}^3 : x_1 \leq 0, x_2 = 0\}$ est un demi-plan affine fermé dans \mathbf{R}^3 . On vérifie que $|\det J_\Phi(r, \phi, \theta)| = r^2 \cos \theta$. Donc, pour $g \in \mathcal{L}^1(\mathbf{R}^3 - P_-)$,

$$\int_{\mathbf{R}^3 - P_-} g(x) dx = \int_0^{+\infty} \int_{-\pi}^{\pi} \int_{-\pi/2}^{\pi/2} g(\Phi(r, \phi, \theta)) r^2 \cos \theta dr d\theta d\phi.$$

En fait, comme P_- est négligeable dans \mathbf{R}^3 , on peut prolonger g n'importe comment sur P_- sans pour autant changer la valeur de l'intégrale. Ainsi, la formule du passage en coordonnées sphériques vaut pour tout $f \in \mathcal{L}^1(\mathbf{R}^3)$, et s'écrit

$$\int_{\mathbf{R}^3} f(x) dx = \int_0^{+\infty} \int_{-\pi}^{\pi} \int_{-\pi/2}^{\pi/2} f(\Phi(r, \phi, \theta)) r^2 \cos \theta dr d\theta d\phi.$$

Exemple 6.2. Calculons le volume de la boule unité de \mathbf{R}^3

$$V = \int_{\mathbf{R}^3} \mathbf{1}_{\{|x| < 1\}} dx = \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi/2}^{\pi/2} r^2 \cos \theta dr d\phi d\theta = \frac{4\pi}{3}.$$

CHAPITRE 7

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 $f \geq 0$,

$$\int_a^b f(x)dx = \text{Aire de } \{(x, y) \in \mathbf{R}^2 : a \leq x \leq b, \quad 0 \leq y \leq f(x)\}.$$

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 siècle¹

On a vu que l'intégrale de Lebesgue permet d'intégrer des fonctions extrêmement 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 définir l'intégrale comme l'aire d'une portion de plan, il vaut mieux utiliser l'intégrale que l'on a déjà construite pour définir et éventuellement calculer l'aire des surfaces ou le volume 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 mesurer l'aire ou le volume. Cette question est loin d'être évidente, comme le montre le paradoxe ci-dessous.

Le paradoxe de Banach-Tarski (1924)²

Soient A et B deux boules ouvertes non vides de \mathbf{R}^N , avec $N \geq 3$. Il existe alors $n \in \mathbf{N} - \{0\}$, et deux partitions finies des boules A et B :

$$\begin{aligned} A &= A_1 \cup \dots \cup A_n, & A_i \cap A_j &= \emptyset \text{ si } i \neq j, \\ B &= B_1 \cup \dots \cup B_n, & B_i \cap B_j &= \emptyset \text{ si } i \neq j, \end{aligned}$$

1. En particulier, B. Pascal avait expliqué dans son *Traité de la roulette* comment calculer l'aire de 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*, 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 vieilles (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>.

ainsi que des isométries affines T_i de l'espace euclidien \mathbf{R}^N telles que $T_i(A_i) = B_i$, pour $1 \leq i \leq n$.

Cet énoncé montre par exemple que l'on peut découper une orange en un nombre fini de morceaux, puis réassembler ces morceaux après déplacement pour obtenir une boule dont le rayon est égal au rayon de Jupiter !

Ceci est évidemment contre-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

$$\text{vol}(A) = \sum_{i=1}^n \text{vol}(A_i) \quad \text{et} \quad \text{vol}(B) = \sum_{i=1}^n \text{vol}(B_i).$$

D'autre part, puisque les applications T_i sont des isométries affines, elles préservent 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

$$\text{vol}(B_i) = \text{vol}(T_i(A_i)) = \text{vol}(A_i),$$

pour tout $i = 1, \dots, n$ et donc, grâce aux deux égalités ci-dessus, on conclurait que

$$\text{vol}(A) = \text{vol}(B).$$

Or le Théorème de Banach-Tarski affirme qu'il existe des partitions finies de A et B , ainsi que des isométries affines T_i pour $i = 1, \dots, n$ sans aucune autre hypothèse que le fait que A et B soient non vides.

Faisons quelques remarques avant d'aller plus loin : la démonstration de Banach-Tarski utilise l'axiome du choix, dont nous avons déjà parlé au chapitre 5 page 69. 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 :

- (i) le volume d'une réunion finie d'ensembles disjoints est la somme de leurs volumes ;
- (ii) 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 que (i) ; 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 correctement le volume des sous-ensembles A_i et B_i qui forment la partition de A et B . 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 7.1 (Ensembles mesurables). *Un sous-ensemble $A \subset \mathbf{R}^N$ est mesurable si sa fonction indicatrice $\mathbf{1}_A$ est mesurable sur \mathbf{R}^N .*

Voici les premières propriétés des ensembles mesurables.

Proposition 7.1. *Soit $N \geq 1$. Alors :*

- (a) \emptyset et \mathbf{R}^N sont des ensembles mesurables ;
- (b) si A et $B \subset \mathbf{R}^N$ sont des ensembles mesurables, $A - B$ est un ensemble mesurable ;
- (c) si $(A_n)_{n \in \mathbf{N}}$ est une famille **dénombrable** d'ensembles mesurables de \mathbf{R}^N , alors leur réunion et leur intersection

$$\bigcup_{n \in \mathbf{N}} A_n \quad \text{et} \quad \bigcap_{n \in \mathbf{N}} A_n,$$

sont des ensembles mesurables (l'énoncé contient évidemment le cas des réunions finies).

DÉMONSTRATION. Le point (a) est trivial car $\mathbf{1}_{\emptyset} = 0$ et $\mathbf{1}_{\mathbf{R}^N} = 1$ sont des fonctions mesurables.

Pour le point (b), observons que

$$\mathbf{1}_{A-B} = \mathbf{1}_A (1 - \mathbf{1}_B),$$

qui est une fonction mesurable si $\mathbf{1}_A$ et $\mathbf{1}_B$ sont des fonctions mesurables.

Démontrons le point (c). Pour cela, posons

$$B_n := \bigcap_{k=0}^n A_k.$$

Alors

$$\mathbf{1}_{B_n} = \prod_{k=0}^n \mathbf{1}_{A_k};$$

et ceci entraîne que la suite $(\mathbf{1}_{B_n})_{n \geq 0}$ est décroissante. De plus, on vérifie facilement (exercice) que

$$\inf_{n \geq 0} \mathbf{1}_{B_n} = \mathbf{1}_{\bigcap_{k \geq 0} A_k}.$$

Au total, la suite $(\mathbf{1}_{B_n})_{n \geq 0}$ converge simplement vers $\mathbf{1}_{\bigcap_{k \geq 0} A_k}$ sur \mathbf{R}^N , ce qui montre que

$$\mathbf{1}_{\bigcap_{k \geq 0} A_k},$$

est une fonction mesurable comme limite simple d'une suite de fonctions mesurables.

L'énoncé sur les réunions dénombrables se déduit facilement de l'énoncé sur les intersections dénombrables par passage au complémentaire, c'est-à-dire en écrivant que

$$\bigcup_{n \geq 0} A_n = \mathbf{R}^N - \bigcap_{n \geq 0} (\mathbf{R}^N - A_n),$$

et en appliquant le point (b). □

La proposition ci-dessus motive la définition suivante :

Définition 7.2. *Étant donné un ensemble Ω quelconque, on appelle tribu sur Ω , tout sous-ensemble \mathcal{F} de l'ensemble $\mathcal{P}(\Omega)$ des parties de Ω vérifiant :*

- (a) $\emptyset \in \mathcal{F}$;

- (b) $A \in \mathcal{F} \Rightarrow \Omega - A \in \mathcal{F}$;
 (c) si $(A_n)_{n \in \mathbf{N}}$ est une suite d'éléments de \mathcal{F} , alors

$$\bigcup_{n \in \mathbf{N}} A_n \in \mathcal{F}.$$

Remarquons que, par passage au complémentaire, une intersection dénombrable d'éléments d'une tribu est un élément de cette tribu. 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.

Proposition 7.2. *Quelques 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 point (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 (voir le point (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

$$F = \bigcup_{n \geq 0} (F \cap B_f(0, n)),$$

et $F \cap B_f(0, n)$ 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.

Soit $K \subset \mathbf{R}^N$ un compact, et, pour tout $n \in \mathbf{N} - \{0\}$, définissons la fonction f_n par

$$f_n(x) := (1 - n \operatorname{dist}(x, K))^+,$$

où $\operatorname{dist}(x, K) := \inf_{y \in K} |y - x|$. Rappelons que la fonction $x \mapsto \operatorname{dist}(x, K)$ définie sur \mathbf{R}^N , est continue car, pour $x, y \in \mathbf{R}^N$,

$$|\operatorname{dist}(x, K) - \operatorname{dist}(y, K)| \leq \operatorname{dist}(x, y),$$

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 $\operatorname{dist}(x, K) > 1$, on a $f_n(x) = 0$. Par conséquent, $f_n \in \mathcal{C}_c(\mathbf{R}^N)$ pour tout $n \geq 1$. Enfin, comme K est fermé, $\operatorname{dist}(x, K) = 0$ 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 $\mathcal{C}_c(\mathbf{R}^N)$, et donc que K est une partie mesurable de \mathbf{R}^N .

Pour le point (d), observons que, si $\mathcal{Z} \subset \mathbf{R}^N$ est négligeable, alors $\mathbf{1}_{\mathcal{Z}} = 0$ 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 point (c). Si P' est un pavé quelconque, il est de la forme $P' = \overline{P'} - B$ où B est une partie de la frontière de P' . Or la frontière de P' est négligeable (comme réunion finie de parties d'hyperplans affines de \mathbf{R}^N). Donc B est également négligeable, et donc mesurable d'après le point (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 $N = 1$. \square

Remarque 7.1. La Proposition 7.1 page 127 montre que l'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). Néanmoins, 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.

Intégration sur un ouvert de \mathbf{R}^N . 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 .

Étant donné un ouvert non vide $\Omega \subset \mathbf{R}^N$ et une fonction $f \in \mathcal{L}^1(\Omega)$, définissons le prolongement de f à \mathbf{R}^N par la formule

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

Clairement, F appartient à $\mathcal{L}^1(\mathbf{R}^N)$ 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 $f \in \mathcal{C}_c(\Omega)$, puisqu'alors $F \in \mathcal{C}_c(\mathbf{R}^N)$, et qu'il s'agit d'une égalité classique entre intégrales au sens usuel. Puis on montre cette égalité pour des fonctions de $\mathcal{L}^1(\Omega)$, 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 7.3 (Intégration sur un ensemble mesurable). *Soient $A \subset \mathbf{R}^N$ un ensemble mesurable et f une fonction définie p.p. sur A . On dira que $f \in \mathcal{L}^1(A)$ si son prolongement*

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

appartient à $\mathcal{L}^1(\mathbf{R}^N)$, auquel cas nous définissons

$$\int_A f(x) dx := \int_{\mathbf{R}^N} F(x) dx.$$

La relation de Chasles³ bien connue pour 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 : étant donnés $A, B \subset \mathbf{R}^N$ deux ensembles mesurables tels que $A \cap B$ est négligeable, et $f \in \mathcal{L}^1(A \cup B)$, on a

$$\int_{A \cup B} f(x) dx = \int_A f(x) dx + \int_B f(x) dx.$$

Remarquons que 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.

3. Michel Chasles X 1812.

Proposition 7.3 (Fonctions/Ensembles mesurables). *Soit Ω un ouvert de \mathbf{R}^N non vide et $f : \Omega \rightarrow \mathbf{R}$ une fonction définie p.p. sur Ω . 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

$$f^{-1}([\lambda, +\infty[) = \bigcap_{n \geq 1} f^{-1}(]\lambda - \frac{1}{n}, +\infty[),$$

et

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

ce qui montre que les deux premières conditions sont équivalentes.

Tout intervalle I de \mathbf{R} est de la forme $I = A - B$, où A et B sont de la forme $[\lambda, +\infty[$ ou $]\lambda, +\infty[$. La troisième condition est donc évidemment équivalente aux deux premières puisque $f^{-1}(A - B) = f^{-1}(A) - f^{-1}(B)$.

Vérifions par exemple la première équivalence. Si f est mesurable, définissons, pour tout $n \geq 1$ et pour presque tout $x \in \mathbf{R}^N$,

$$h_n(x) := n \left(\min \left(f(x), \lambda + \frac{1}{n} \right) - \min(f(x), \lambda) \right),$$

(voir sur la Figure 7.1 la signification géométrique de cette formule). Pour tout $n \geq 1$, la fonction h_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 . Étant donné que $\lim_{n \rightarrow +\infty} h_n(x) = \mathbf{1}_{] \lambda, +\infty[}(f(x))$ pour tout $x \in \Omega$ tel que $f(x)$ est défini, c'est-à-dire p.p. sur Ω , on en déduit que la fonction $\mathbf{1}_{] \lambda, +\infty[} \circ f$ est mesurable comme limite p.p. d'une suite de fonctions mesurables. Or

$$\mathbf{1}_{] \lambda, +\infty[} \circ f = \mathbf{1}_{f^{-1}(] \lambda, +\infty[)},$$

p.p. et, comme cette fonction est mesurable, on en déduit que l'ensemble $f^{-1}(] \lambda, +\infty[)$ est mesurable.

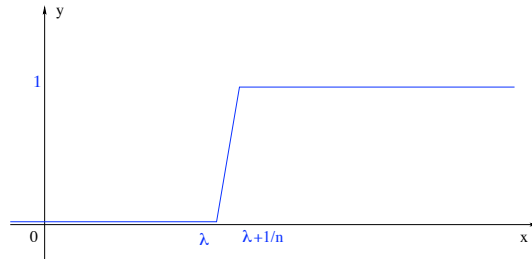


FIGURE 7.1. Graphe de la fonction $x \mapsto n \left(\min \left(x, \lambda + \frac{1}{n} \right) - \min(x, \lambda) \right)$.

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 $n \geq 0$, on pose

$$f_n := \sum_{k=0}^{4^n} \frac{k}{2^n} \mathbf{1}_{f^{-1}([2^{-n}k, 2^{-n}(k+1)])}.$$

Évidemment, pour tout $n \geq 0$, la fonction f_n est mesurable comme combinaison linéaire 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, $f_n \rightarrow f$ simplement sur Ω , de sorte que f est mesurable comme limite simple de la suite de fonctions mesurables $(f_n)_{n \geq 0}$.

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 $f > -\infty$ p.p. sur Ω par hypothèse, $f_n \rightarrow f$ p.p. sur Ω , de sorte que f est mesurable comme limite p.p. d'une suite de fonctions mesurables. \square

3. Mesure de Lebesgue

Définition 7.4 (Mesure de Lebesgue). *Pour tout $A \subset \mathbf{R}^N$ mesurable, la mesure de Lebesgue de A est définie par*

$$|A| := \int_{\mathbf{R}^N} \mathbf{1}_A(x) dx.$$

Cette définition éclaire la section précédente. En effet, la partie A de \mathbf{R}^N est mesurable si 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 7.4. *La mesure de Lebesgue vérifie les propriétés suivantes :*

(a) *la mesure de Lebesgue d'un pavé de \mathbf{R}^N est égale à son volume*

$$|(a_1, b_1) \times \dots \times (a_N, b_N)| = \prod_{k=1}^N (b_k - a_k)^+$$

(ici (a, b) est un intervalle de la forme $[a, b]$, $[a, b[$, $]a, b[$ ou $]a, b[)$ — en particulier $|\emptyset| = 0$;

(b) *pour tout sous-ensemble mesurable $A \subset \mathbf{R}^N$ et pour toute isométrie affine T de \mathbf{R}^N , on a*

$$|T(A)| = |A|;$$

(c) *si $(A_n)_{n \geq 0}$ est une suite dénombrable de sous-ensembles mesurables de \mathbf{R}^N , alors*

$$A_k \cap A_l = \emptyset \text{ pour tous } k \neq l \quad \Rightarrow \quad \left| \bigcup_{n \geq 0} A_n \right| = \sum_{n \geq 0} |A_n|;$$

- (d) si $A, B \subset \mathbf{R}^N$ sont des sous-ensembles mesurables et si $A \subset B$, alors $|A| \leq |B|$;
 (e) si $(B_n)_{n \geq 0}$ est une suite dénombrable de sous-ensembles mesurables de \mathbf{R}^N , alors

$$\left| \bigcup_{n \geq 0} B_n \right| \leq \sum_{n \geq 0} |B_n|.$$

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 (c'est la 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 volumes. Cette généralisation au cas de réunions dénombrables est particulièrement subtile — voir en particulier la Proposition 5.6 page 83 et sa démonstration.

DÉMONSTRATION. Dans le cas $N = 1$, la fonction $\mathbf{1}_{(a,b)}$ est continue par morceaux (ici (a, b) est un intervalle fermé, ouvert ou semi-ouvert), le calcul intégral usuel nous apprend que

$$\int_{\mathbf{R}} \mathbf{1}_{(a,b)}(x) dx = \begin{cases} \int_a^b dx = b - a & \text{si } b \geq a, \\ 0 & \text{si } a > b. \end{cases}$$

Le cas $N > 1$ découle du cas $N = 1$ en appliquant le Théorème de Tonelli. Ceci démontre le point (a).

Le point (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, si

$$A_k \cap A_l = \emptyset \text{ pour tous } k \neq l \quad \Rightarrow \quad \mathbf{1}_{\bigcup_{n \geq 0} A_n} = \sum_{n \geq 0} \mathbf{1}_{A_n}.$$

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

$$\sum_{n \geq 0} \mathbf{1}_{A_n}(x),$$

et d'ailleurs x appartient au plus à l'un des ensembles A_n , la formule ci-dessus est donc triviale.

La formule du point (c) découle alors de la linéarité de l'intégrale dans le cas d'une réunion finie, et du Théorème de la convergence monotone dans le cas d'une réunion infinie.

Le point (d) est une conséquence du point (c). En effet, $B = A \cup (B - A)$ est une réunion disjointe, de sorte que

$$|B| = |A| + |B - A| \geq |A|.$$

Pour ce qui est du point (e), supposons que la suite $(B_n)_{n \geq 0}$ 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 $A_0 = B_0$, $A_1 = B_1 - B_0$, et plus généralement

$$A_n = B_n - (B_0 \cup \dots \cup B_{n-1}), \quad n \geq 1.$$

Alors

$$\bigcup_{n \geq 0} A_n = \bigcup_{n \geq 0} B_n;$$

de plus

$$A_n \cap A_m \subset A_n \cap B_m = \emptyset,$$

si $n > m$. Donc, d'après les points (c) et (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 . \square

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 7.1 (La mesure de Lebesgue des ouverts de \mathbf{R}). Commençons par une remarque : 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 $\Omega \subset \mathbf{R}$ non vide 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 non vide Ω quelconque de \mathbf{R} : on écrit la décomposition ci-dessus de Ω , à savoir

$$\Omega = \bigcup_{n \geq 0}]a_n, b_n[\quad \text{où } a_n < b_n \text{ et }]a_m, b_m[\cap]a_n, b_n[= \emptyset \text{ si } m \neq n,$$

et par additivité dénombrable de la mesure de Lebesgue

$$|\Omega| = \sum_{n \geq 0} (b_n - a_n).$$

Nous aurons besoin du résultat suivant qui est connu sous le nom de *régularité de la mesure de Lebesgue*. Ce résultat montre que les ensembles mesurables de \mathbf{R}^N peut être approchés par des fermés et des ouverts. Plus précisément, nous avons la :

Proposition 7.5. *Soit $A \subset \mathbf{R}^N$ un ensemble mesurable, de mesure finie. Alors, pour tout $\varepsilon > 0$, il existe un ouvert O et un fermé F tels que $F \subset A \subset O$ et $|O - F| \leq \varepsilon$.*

DÉMONSTRATION. Choisissons $t > 1$ tel que $(t - 1)|A| < \varepsilon/3$. Par définition, la fonction $\mathbf{1}_A$ est une fonction intégrable et, d'après le Lemme 5.6 page 84, il existe deux fonctions $u, v \in \mathcal{L}^+(\mathbf{R}^N)$ telles que $t \mathbf{1}_A = u - v$ et $v \geq 0$ sur $\mathbf{R}^N - \mathcal{Z}$, où \mathcal{Z} est un ensemble négligeable. De plus

$$\int_{\mathbf{R}^N} v \, dx \leq \varepsilon/3.$$

L'ensemble \mathcal{Z} étant négligeable, il existe une fonction positive, $w \in \mathcal{L}^+(\mathbf{R}^N)$ telle que $\mathcal{Z} \subset \{x \in \mathbf{R}^N : w(x) = +\infty\}$ et l'on peut toujours supposer que

$$\int_{\mathbf{R}^N} w \, dx \leq 1.$$

Donc

$$t \mathbf{1}_A \leq t \mathbf{1}_A + v + (\varepsilon/3) w = u + (\varepsilon/3) w.$$

On note $O := \{x : (u + (\varepsilon/3)w)(x) > 1\}$. Étant donné que $u + (\varepsilon/3)w \geq t \mathbf{1}_A$, on a $A \subset O$. De plus O est un ouvert car la fonction $u + (\varepsilon/3)w$ qui appartient à $\mathcal{L}^+(\mathbf{R}^N)$ est s.c.i. (voir le Lemme 5.4 page 76). Enfin

$$\begin{aligned} |A| &= \int_{\mathbf{R}^N} \mathbf{1}_A dx \leq \int_{\mathbf{R}^N} \mathbf{1}_O dx = |O| \leq \int_{\mathbf{R}^N} (u + (\varepsilon/3)w) dx \\ &\leq \int_{\mathbf{R}^N} (t \mathbf{1}_A + v) dx + \varepsilon/3 \leq t |A| + 2\varepsilon/3 < |A| + \varepsilon. \end{aligned}$$

Pour conclure, on a construit un ouvert O tel que $A \subset O$ et $|A| \leq |O| \leq |A| + \varepsilon$, donc $|O - A| \leq \varepsilon$.

Pour démontrer l'existence d'un fermé $F \subset A$ tel que $|A - F| \leq \varepsilon$, on remarque que

$$|A| = \lim_{n \rightarrow +\infty} |A_n|,$$

où $A_n := A \cap [-n, n]^N$.

En particulier, $\varepsilon > 0$ étant fixé, il existe un entier $n \geq 1$ tel que $|A - A_n| \leq \varepsilon/2$. D'après la première partie, il existe un ouvert O tel que

$$\mathbf{R}^N - A_n \subset O \quad \text{et} \quad |O - (\mathbf{R}^N - A_n)| \leq \varepsilon/2.$$

On note $F := \mathbf{R}^N - O$. Bien entendu, F est un fermé en tant que complémentaire d'un ouvert, et $F \subset A_n \subset A$. Enfin, on vérifie que $A_n - F = O - (\mathbf{R}^N - A_n)$ et donc, $|A - F| \leq |A - A_n| + |O - (\mathbf{R}^N - A_n)| \leq \varepsilon$, ce qui termine la démonstration. \square

4. Fonction de répartition complémentaire

On a vu dans la section 2 page 126 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 7.5 (Fonction de répartition complémentaire). *Soient Ω un ouvert de \mathbf{R}^N supposé non vide, et $f : \Omega \rightarrow \mathbf{R}_+$ une fonction mesurable. La fonction de répartition complémentaire de f est la fonction $\rho_f : \mathbf{R}_+ \rightarrow [0, +\infty]$ définie par*

$$\rho_f(\lambda) := |f^{-1}(]\lambda, +\infty[)| = |\{x \in \Omega : f(x) > \lambda\}|.$$

Évidemment, 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 $\lambda \geq 0$. De plus, par construction, la fonction ρ_f est décroissante au sens large, puisque

$$\lambda' \geq \lambda \quad \Rightarrow \quad f^{-1}(]\lambda', +\infty[) \subset f^{-1}(]\lambda, +\infty[) \quad \Rightarrow \quad \rho_f(\lambda') \leq \rho_f(\lambda).$$

Voici une première relation entre la fonction de répartition complémentaire d'une fonction et son intégrale de Lebesgue.

Proposition 7.6 (Inégalité de Markov). *Soient Ω ouvert non vide de \mathbf{R}^N et $f \in \mathcal{L}^1(\Omega)$. Alors, pour tout $\lambda > 0$, on a*

$$\rho_{|f|}(\lambda) = |\{x \in \Omega : |f(x)| \geq \lambda\}| \leq \frac{1}{\lambda} \int_{\Omega} |f(x)| dx.$$

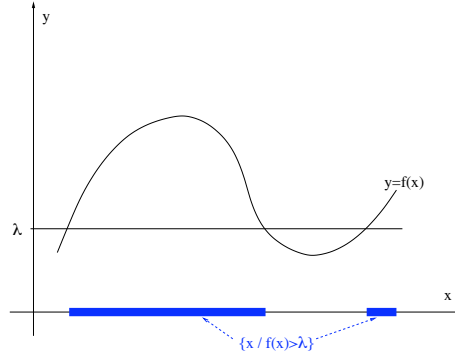


FIGURE 7.2. Fonction de répartition complémentaire.

DÉMONSTRATION. En effet

$$\begin{aligned}
 \int_{\Omega} |f(x)| dx &\geq \int_{\Omega} |f(x)| \mathbf{1}_{[\lambda, +\infty]}(f(x)) dx \\
 &\geq \lambda \int_{\Omega} \mathbf{1}_{[\lambda, +\infty]}(f(x)) dx \\
 &\geq \lambda |f^{-1}([\lambda, +\infty])| \\
 &\geq \lambda \rho_{|f|}(\lambda),
 \end{aligned}$$

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 pour presque tout $x \in \Omega$. \square

L'inégalité de Markov a de nombreuses conséquences importantes. En particulier, elle montre que les ensembles négligeables sont les ensembles de mesure nulle.

Théorème 7.1 (Mesure et ensembles négligeables). *Soient $A \subset \mathbf{R}^N$ et f une fonction mesurable positive définie sur Ω ouvert non vide de \mathbf{R}^N . Alors :*

- (i) A est négligeable $\Leftrightarrow |A| = 0$;
- (ii) si $\int_{\Omega} f(x) dx = 0$, alors $|\{x \in \Omega : f(x) \neq 0\}| = 0$;
- (iii) si $\int_{\Omega} f(x) dx < +\infty$, alors $|\{x \in \Omega : f(x) = +\infty\}| = 0$.

DÉMONSTRATION. Établissons le point (iii). Pour tout $n \in \mathbf{N} - \{0\}$, 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 point (ii), remarquons que

$$\{x \in \Omega : f(x) \neq 0\} = \bigcup_{n \geq 1} \{x \in \Omega : f(x) > 1/n\} ;$$

or d'après l'inégalité de Markov

$$|\{x \in \Omega : f(x) > 1/n\}| \leq n \int_{\Omega} f(x) dx = 0,$$

de sorte que

$$|\{x \in \Omega : f(x) \neq 0\}| \leq \sum_{n \geq 1} |\{x \in \Omega : f(x) > 1/n\}| = 0,$$

d'après le point (e) de la Proposition 7.4 page 131.

Terminons par la démonstration du point (i). Si A est 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 point (iii), on a

$$|A| \leq |\{x \in \mathbf{R}^N : |f(x)| = +\infty\}| = 0.$$

Inversement, si $|A| = 0$, la fonction positive $\mathbf{1}_A$ est intégrable et vérifie

$$\int_{\mathbf{R}^N} \mathbf{1}_A(x) dx = |A| = 0.$$

D'après le Théorème 5.1 page 87, on conclut que $\mathbf{1}_A = 0$ p.p. sur \mathbf{R}^N , ce qui équivaut précisément à dire que A est négligeable. \square

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 7.2 (Principe de Cavalieri). *Soient Ω un ouvert non vide de \mathbf{R}^N et $f : \Omega \rightarrow [0, +\infty]$ une fonction mesurable. Soit ρ_f la fonction de répartition complémentaire de f . Alors, pour tout $\Phi : \mathbf{R}_+ \rightarrow \mathbf{R}_+$ croissante de classe \mathcal{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, puisque l'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 Riemann⁴ (voir le Théorème 5.4 page 98).

Autrement dit, toute intégrale de Lebesgue peut se ramener à une intégrale de Riemann... une fois que l'on a défini la mesure de Lebesgue. Nous reviendrons plus loin sur cette remarque, qui est lourde de conséquences.

4. Si $\phi : \mathbf{R}_+ \rightarrow \mathbf{R}_+$ 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.

Le cas particulier ci-dessus du principe de Cavalieri suggère une autre remarque importante. L'intégrale usuelle, dont la définition est due à Cauchy, puis Riemann, consiste à poser, pour $f \in \mathcal{C}([a, b])$,

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

L'idée clef 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 petite⁵ lorsque $n \rightarrow +\infty$, de sorte que l'on peut assimiler f à une constante sur chacun de ces intervalles lorsque leur longueur tend vers 0.

Autrement dit, le cœur 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 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 $\varepsilon > 0$ 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 clef consiste à écrire

$$\Phi(f(x)) = \int_0^{f(x)} \Phi'(\lambda) d\lambda,$$

pour presque tout $x \in \Omega$. Comme f est mesurable et Φ de classe \mathcal{C}^1 croissante sur \mathbf{R}_+ , la fonction

$$\Omega \times \mathbf{R}_+ \ni (x, \lambda) \mapsto \mathbf{1}_{[0, f(x)[}(\lambda) \Phi'(\lambda) \in \mathbf{R}_+,$$

est mesurable. D'après le Théorème de Tonelli,

$$\begin{aligned} \int_{\Omega} \Phi(f(x)) dx &= \int_{\Omega} \left(\int_0^{f(x)} \Phi'(\lambda) d\lambda \right) dx \\ &= \int_{\Omega} \left(\int_{\mathbf{R}_+} \mathbf{1}_{[0, f(x)[}(\lambda) \Phi'(\lambda) d\lambda \right) dx \\ &= \iint_{\Omega \times \mathbf{R}_+} \mathbf{1}_{[0, f(x)[}(\lambda) \Phi'(\lambda) dx d\lambda. \end{aligned}$$

On remarque que

$$\mathbf{1}_{[0, f(x)[}(\lambda) = \mathbf{1}_{\lambda, +\infty[}(f(x)),$$

5. En effet, f étant continue sur le compact $[a, b]$, elle y est uniformément continue, d'après le Corollaire 3.2 page 39.

pour tout $\lambda \in \mathbf{R}_+$ et presque tout $x \in \Omega$, de sorte que, toujours d'après le Théorème de Tonelli

$$\begin{aligned}
 \int_{\Omega} \Phi(f(x)) dx &= \iint_{\Omega \times \mathbf{R}_+} \mathbf{1}_{[0, f(x)]}(\lambda) \Phi'(\lambda) dx d\lambda \\
 &= \iint_{\Omega \times \mathbf{R}_+} \mathbf{1}_{] \lambda, +\infty[}(f(x)) \Phi'(\lambda) dx d\lambda \\
 &= \int_{\mathbf{R}_+} \Phi'(\lambda) \left(\int_{\Omega} \mathbf{1}_{] \lambda, +\infty[}(f(x)) dx \right) d\lambda \\
 &= \int_{\mathbf{R}_+} \Phi'(\lambda) |\{x \in \Omega : f(x) > \lambda\}| d\lambda \\
 &= \int_{\mathbf{R}_+} \Phi'(\lambda) \rho_f(\lambda) d\lambda.
 \end{aligned}$$

Ce qui termine la démonstration. □

5. Mesures de Radon

Nous donnons dans cette section de nouveaux exemples de mesures qui sont construites en utilisant la méthode que nous avons utilisée pour construire la mesure de Lebesgue.

Au lieu de partir de l'intégrale usuelle sur $\mathcal{C}_c(\Omega)$, on se donne cette fois-ci une forme linéaire positive⁶,

$$\mu : \mathcal{C}_c(\Omega) \rightarrow \mathbf{R},$$

au sens où

$$f \in \mathcal{C}_c(\Omega) \quad \text{et} \quad f \geq 0 \quad \Rightarrow \quad \langle \mu, f \rangle \geq 0.$$

Exemple 7.2. Soit $w \in \mathcal{C}(\Omega)$ une fonction positive, on peut définir une forme linéaire positive sur $\mathcal{C}_c(\Omega)$ par la formule

$$\langle \mu, f \rangle := \int_{\Omega} f(t) w(t) dt,$$

pour tout $f \in \mathcal{C}_c(\Omega)$.

Exemple 7.3 (Masse de Dirac en un point). Soit $x_0 \in \Omega$. On peut définir une forme linéaire positive sur $\mathcal{C}_c(\Omega)$ par la formule

$$\langle \delta_{x_0}, f \rangle := f(x_0),$$

pour tout $f \in \mathcal{C}_c(\Omega)$.

Le procédé de *prolongement par monotonie* qui nous a permis de construire l'intégrale de Lebesgue s'applique de même et aboutit à l'espace vectoriel $\mathcal{L}^1(\Omega, \mu)$ des fonctions μ -intégrables et à un prolongement à $\mathcal{L}^1(\Omega, \mu)$ de la forme linéaire μ , noté

$$\mathcal{L}^1(\Omega, \mu) \ni f \quad \mapsto \quad \int_{\Omega} f(x) d\mu(x) \in \mathbf{R},$$

par analogie avec l'intégrale de Lebesgue — c'est-à-dire que $\mathcal{C}_c(\Omega) \subset \mathcal{L}^1(\Omega, \mu)$ et

$$f \in \mathcal{C}_c(\Omega) \quad \Rightarrow \quad \int_{\Omega} f(x) d\mu(x) = \langle \mu, f \rangle.$$

6. On utilisera la notation $\langle \mu, f \rangle$ pour désigner l'image de la fonction $f \in \mathcal{C}_c(\Omega)$ par la forme linéaire μ .

De même que dans le cas de l'intégrale de Lebesgue, le prolongement de la forme linéaire ainsi obtenu vérifie les Théorèmes de la convergence monotone et le Théorème de la convergence 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 (i.e. l'intégrale de Lebesgue). Ils doivent être définis au moyen des suites fondamentales pour la forme linéaire μ , suivant la Définition 5.5 page 79, et surtout pas en utilisant la caractérisation géométrique du Théorème 5.3 page 94, qui ne vaut que pour les ensembles négligeables. Plus précisément, tous les résultats de la section 3.2 page 79, sauf les exemples 5.7 page 79, 5.8 page 83 et 5.10 page 83 restent valables pour les ensembles μ -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 $\mathcal{C}_c(\Omega)$).

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

$$A \text{ est } \mu\text{-négligeable} \quad \Leftrightarrow \quad \mu(A) = 0.$$

Remarquons que la clef de toute cette construction est la Proposition 3.5 page 43, qui est une conséquence du Théorème de Dini.

La terminologie est la suivante : on appelle *mesure de Radon sur Ω* une forme linéaire positive μ définie sur $\mathcal{C}_c(\Omega)$ comme ci-dessus ; son prolongement à la classe des fonctions μ -sommables est appelé *intégrale par rapport à la mesure μ* . Remarquons que 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 $\mathcal{C}_c(\Omega)$ 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 elle est le prolongement de l'intégrale usuelle à la classe $\mathcal{L}^1(\Omega)$ des fonctions intégrables 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 :

Exemple 7.4 (Masse de Dirac). À $x_0 \in \Omega$, nous avons associé une forme linéaire positive δ_{x_0} définie sur $\mathcal{C}_c(\Omega)$ 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 les sous-ensembles de Ω qui ne contiennent pas x_0 . Ces ensembles ne sont évidemment pas Lebesgue-négligeables en général. Cet exemple montre donc clairement que la notion d'ensemble négligeable dépend très fortement de la mesure de Radon considérée.

Exemple 7.5 (Mesure de comptage). À tout sous-ensemble dénombrable sans point d'accumulation $A \subset \Omega$, on associe la forme linéaire positive définie sur $\mathcal{C}_c(\Omega)$ par

$$\mu_A := \sum_{a \in A} \delta_a.$$

Autrement dit

$$\langle \mu_A, f \rangle = \sum_{a \in A} f(a) =: \int_{\Omega} f(x) d\mu_A(x).$$

Un exemple important est celui où $A = \mathbf{N}$ (ou $A = \mathbf{Z}$). En effet, on constate alors que la théorie des séries numériques coïncide avec la théorie de l'intégration par rapport à la mesure de comptage associée au sous-ensemble \mathbf{N} (ou \mathbf{Z}) de \mathbf{R} . Par exemple,

$$\langle \mu_{\mathbf{N}}, f \rangle = \sum_{n \in \mathbf{N}} f(n).$$

Une fonction $f \in \mathcal{L}^1(\mathbf{R}, \mu_{\mathbf{N}})$ si et seulement si

$$\sum_{n \in \mathbf{N}} |f(n)| < +\infty.$$

Les ensembles $\mu_{\mathbf{N}}$ négligeables sont les ensembles qui ne contiennent aucun élément de \mathbf{N} et l'on peut alors identifier $f \in \mathcal{L}^1(\mathbf{R}, \mu_{\mathbf{N}})$ avec la suite réelle $(a_n)_{n \geq 0}$ telle que $\sum_{n \in \mathbf{N}} |a_n| < +\infty$,

en posant $a_n := f(n)$.

La plupart des résultats que nous avons démontrés s'appliquent. Par exemple, nous avons l'inégalité de Hölder

$$\left| \sum_{n \in \mathbf{N}} a_n b_n \right| \leq \left(\sum_{n \in \mathbf{N}} |a_n|^p \right)^{1/p} \left(\sum_{n \in \mathbf{N}} |b_n|^q \right)^{1/q},$$

pour tous $p, q > 1$ tels que $\frac{1}{p} + \frac{1}{q} = 1$. Nous avons aussi l'inégalité de Minkowski

$$\left(\sum_{n \in \mathbf{N}} |a_n + b_n|^p \right)^{1/p} \leq \left(\sum_{n \in \mathbf{N}} |a_n|^p \right)^{1/p} + \left(\sum_{n \in \mathbf{N}} |b_n|^p \right)^{1/p},$$

pour tout $p \geq 1$.

Le lecteur curieux pourra essayer de comprendre comment interpréter le Théorème de Fubini pour la mesure $\mu_{\mathbf{N}}$ définie ci-dessus.

Le point de vue des mesures de Radon est un cas particulier d'une théorie plus vaste : la *théorie des distributions*, qui sera étudiée dans la seconde partie du cours MAT431 de deuxième année.

6. Théorie abstraite de la mesure

Nous avons vu dans la Proposition 7.1 page 127 que les parties mesurables d'un ouvert Ω forment une tribu sur Ω — c'est-à-dire un sous ensemble de $\mathcal{P}(\Omega)$ qui contient \emptyset et qui est stable par passage au complémentaire et par réunion dénombrable. Plus généralement (voir la Définition 7.2 page 128), définissons :

Définition 7.6 (Tribu ou σ -algèbre). *Soit X un ensemble et $\mathcal{A} \subset \mathcal{P}(X)$. On dit que \mathcal{A} est une tribu (ou σ -algèbre) sur X si les propriétés suivantes sont vérifiées :*

- (i) $\emptyset \in \mathcal{A}$;
- (ii) \mathcal{A} est stable par passage au complémentaire, i.e. si $A \in \mathcal{A}$ alors $X - A \in \mathcal{A}$;
- (iii) \mathcal{A} est stable par réunion dénombrable. i.e. si $A_n \in \mathcal{A}$ pour tout $n \in \mathbf{N}$, alors $\bigcup_{n \geq 0} A_n \in \mathcal{A}$.

On dit dans ce cas que (X, \mathcal{A}) est un espace mesurable et les éléments de \mathcal{A} sont les sous-ensembles mesurables de X .

Exemple 7.6. Soit E un ensemble contenant au moins deux éléments et $A \subset X$ un sous-ensemble non vide tel que $A \neq X$. Alors $\{\emptyset, X\}$, $\{\emptyset, A, X - A, X\}$ et $\mathcal{P}(X)$ sont des tribus sur X .

Par passage aux complémentaires, on remarque que, si \mathcal{A} est une tribu, alors $X \in \mathcal{A}$, car X est le complémentaire de \emptyset dans X , de plus \mathcal{A} est stable par intersection dénombrable, i.e. si $A_n \in \mathcal{A}$ pour tout $n \in \mathbf{N}$, alors $\bigcap_{n \geq 0} A_n \in \mathcal{A}$. Une intersection quelconque de tribus est une tribu. Enfin, si $A, B \in \mathcal{A}$ alors $A - B \in \mathcal{A}$ (car $A - B = A \cap (X - B)$).

Définition 7.7 (Tribu engendrée par un sous-ensemble de $\mathcal{P}(X)$). *Soient X un ensemble et $\mathcal{S} \subset \mathcal{P}(X)$. On appelle tribu engendrée par \mathcal{S} la plus petite tribu contenant \mathcal{S} , c'est-à-dire la tribu $\mathcal{A}(\mathcal{S})$ qui est définie comme l'intersection de toutes les tribus sur X contenant \mathcal{S} (cette intersection est non vide car $\mathcal{P}(X)$ est une tribu contenant \mathcal{S}).*

Exemple 7.7. La tribu des boréliens sur \mathbf{R}^N est la tribu engendrée par les ouverts de \mathbf{R}^N (muni de la topologie usuelle).

Le lecteur démontra à titre d'exercice que :

Exemple 7.8. On note \mathcal{O}_1 l'ensemble des intervalles ouverts de \mathbf{R} , \mathcal{O}_2 l'ensemble des intervalles ouverts bornés de \mathbf{R} et $\mathcal{O}_3 := \{]a, \infty[: a \in \mathbf{R}\}$. Les tribus engendrées par \mathcal{O}_1 , \mathcal{O}_2 ou \mathcal{O}_3 sont identiques et égales à la tribu des boréliens sur \mathbf{R} (pour le vérifier, on utilisera le fait que tout ouvert est réunion dénombrable d'intervalles ouverts).

Remarque 7.2. Remarquons que, dans le cas où $X = \mathbf{R}^N$, nous avons défini les ensemble mesurables (pour la mesure de Lebesgue) de \mathbf{R}^N , dont nous avons démontré qu'ils formaient une tribu (voir la Proposition 7.1 page 127). Cette tribu contient la tribu des boréliens (voir la Proposition 7.2 page 128 qui nous assure que la tribu des ensembles mesurables contient les ouverts, donc elle contient la tribu engendrée par les ouverts de \mathbf{R}^N). En fait, on peut montrer que la tribu des ensembles mesurables (pour la mesure de Lebesgue) contient strictement la tribu des boréliens.

Maintenant que nous avons défini la notion d'espace mesurable, nous pouvons donner la :

Définition 7.8. Soit (X, \mathcal{A}) un espace mesurable. On dit qu'une application $\mu : \mathcal{A} \rightarrow [0, +\infty]$ est une mesure (positive) sur (X, \mathcal{A}) si les propriétés suivantes sont vérifiées :

- (i) $\mu(\emptyset) = 0$;
- (ii) μ est σ -additive, i.e. pour toute suite $(A_n)_{n \geq 0}$ d'éléments deux à deux disjoints de \mathcal{A} , on a

$$\mu \left(\bigcup_{n \geq 0} A_n \right) = \sum_{n \geq 0} \mu(A_n).$$

On dit alors que (X, \mathcal{A}, μ) est un espace mesuré. On dira que la mesure μ est finie si $\mu(X) < +\infty$ et l'on dit qu'une mesure μ est une mesure de probabilité si $\mu(X) = 1$.

Exemple 7.9 (Mesure de Dirac). Soit (X, \mathcal{A}) un espace mesurable et $a \in X$. L'application δ_a définie par

$$\delta_a(A) := \begin{cases} 1 & \text{si } a \in A \\ 0 & \text{si } a \notin A, \end{cases}$$

pour tout $A \in \mathcal{A}$, est une mesure sur \mathcal{A} . Nous avons déjà vu cet exemple dans la section dédiée aux mesures de Radon.

Exemple 7.10. Nous avons construit dans les chapitres précédents les ensembles mesurables dont nous avons démontré qu'ils formaient une tribu et la mesure de Lebesgue pour laquelle nous avons démontré qu'elle vérifie bien les axiomes ci-dessus (voir la Proposition 7.4 page 131). Bien entendu, toute la difficulté est justement de démontrer l'existence d'une mesure (en l'occurrence la mesure de Lebesgue) et d'une tribu (en l'occurrence la tribu des ensembles mesurables) qui vérifient bien les axiomes ci-dessus.

Remarque 7.3. On montre qu'il n'existe pas de mesure m définie sur $\mathcal{P}(\mathbf{R})$ telle que la mesure de $[a, b]$ soit égale à $b - a$ et telle que $m(A) = m(A + t)$ pour tous $A \in \mathcal{P}(\mathbf{R})$ et $t \in \mathbf{R}$, où $A + t := \{a + t : a \in A\}$.

Les propriétés suivantes sont de simples conséquence des propriétés de σ -additivité des mesures.

Proposition 7.7. Soit (X, \mathcal{A}, μ) un espace mesuré. Alors, les propriétés suivantes sont vérifiées :

- (i) soient $A, B \in \mathcal{A}$ tels que $A \subset B$, alors $\mu(A) \leq \mu(B)$;
- (ii) soient $A_n \in \mathcal{A}$, pour tout $n \in \mathbf{N}$. Alors

$$\mu \left(\bigcup_{n \geq 0} A_n \right) \leq \sum_{n \geq 0} \mu(A_n);$$

- (iii) Soit $(A_n)_{n \geq 0}$ une suite croissante d'éléments de \mathcal{A} , i.e. $A_n \subset A_{n+1}$, pour tout $n \in \mathbf{N}$. Alors

$$\lim_{n \rightarrow +\infty} \mu(A_n) = \mu \left(\bigcup_{n \geq 0} A_n \right);$$

(iv) Soit $(A_n)_{n \geq 0}$ une suite décroissante d'éléments de \mathcal{A} , i.e. $A_{n+1} \subset A_n$, pour tout $n \in \mathbf{N}$. On suppose que $\mu(A_0) < \infty$. Alors

$$\lim_{n \rightarrow +\infty} \mu(A_n) = \mu\left(\bigcap_{n \geq 0} A_n\right).$$

DÉMONSTRATION. Pour démontrer la première propriété, on se donne $A, B \in \mathcal{A}$ tels que $A \subset B$. On écrit $B = A \cup (B - A)$. La σ -additivité de μ implique que $\mu(B) = \mu(A) + \mu(B - A)$. Étant donné que $\mu(B - A) \geq 0$, on conclut que $\mu(A) \leq \mu(B)$.

Soit $(A_n)_{n \in \mathbf{N}}$ une suite d'éléments de \mathcal{A} . On définit par récurrence la suite $(B_n)_{n \geq 0}$ de la manière suivante : $B_0 = A_0$ et $B_n := A_n - \left(\bigcup_{i=0}^{n-1} A_i\right)$. On vérifie que $B_n \subset A_n$ et que $\bigcup_{n \geq 0} A_n = \bigcup_{n \geq 0} B_n$. De plus, les B_n sont deux à deux disjoints. La σ -dditivité de m implique que

$$\mu\left(\bigcup_{n \geq 0} A_n\right) = \mu\left(\bigcup_{n \geq 0} B_n\right) = \sum_{n \geq 0} \mu(B_n) \leq \sum_{n \geq 0} \mu(A_n).$$

Ce qui démontre le point (ii).

Pour démontrer le point (iii), on commence par définir la suite $(B_n)_{n \geq 0}$ par $B_0 = A_0$ et $B_n := A_n - A_{n-1}$. Les B_n sont deux à deux disjoints et $\bigcup_{n \geq 0} A_n = \bigcup_{n \geq 0} B_n$. Donc

$$\begin{aligned} \lim_{n \rightarrow +\infty} \mu(A_n) &= m\left(\bigcup_{n \geq 0} A_n\right) = \mu\left(\bigcup_{n \geq 0} B_n\right) \\ &= \sum_{n \geq 0} \mu(B_n) = \lim_{\bar{n} \rightarrow \infty} \sum_{n=0}^{\bar{n}} \mu(B_n) \\ &= \lim_{\bar{n} \rightarrow +\infty} \mu\left(\bigcup_{n=0}^{\bar{n}} B_n\right) = \lim_{\bar{n} \rightarrow +\infty} \mu(A_n). \end{aligned}$$

Pour démontrer le point (iv), on note $B_n := A_0 - A_n$. Alors, la suite $(B_n)_{n \geq 0}$ est une suite croissante et

$$A_0 - \bigcap_{n \geq 0} A_n = \bigcup_{n \geq 0} B_n,$$

donc, d'après le point (iii), on peut écrire

$$\mu\left(A_0 - \bigcap_{n \geq 0} A_n\right) = \lim_{n \rightarrow +\infty} \mu(A_0 - A_n).$$

Ou encore,

$$\mu(A_0) - \mu\left(\bigcap_{n \geq 0} A_n\right) = \lim_{n \rightarrow +\infty} (\mu(A_0) - m(A_n)).$$

Étant donné que $\mu(A_0)$ est fini, on conclut que

$$\mu\left(\bigcap_{n \geq 0} A_n\right) = \lim_{n \rightarrow +\infty} \mu(A_n).$$

Ce qui termine la démonstration. □

Une fois que nous avons un espace mesuré, nous pouvons donner la :

Définition 7.9. Soit (X, \mathcal{A}, μ) un espace mesuré. On dit qu'un ensemble $A \in \mathcal{P}(X)$ est négligeable s'il existe $B \in \mathcal{A}$ tel que $A \subset B$ et $\mu(B) = 0$. Une propriété $\mathcal{P}(x)$ dépendant de $x \in X$ est dite vraie presque partout sur X (ou vraie pour presque tout $x \in X$) si l'ensemble des $x \in X$ pour lesquels $\mathcal{P}(x)$ est fausse est négligeable dans X .

Soit (X, \mathcal{A}) un espace mesurable. Nous dirons qu'une fonction $f : X \rightarrow \mathbf{R}$ est \mathcal{A} -mesurable si, pour tout intervalle $I \subset \mathbf{R}$, $f^{-1}(I) \in \mathcal{A}$. On vérifie qu'une combinaison linéaire ou le produit de deux fonctions mesurables sur X à valeurs dans \mathbf{R} sont mesurables et que, si f_1, \dots, f_N sont mesurables sur X à valeurs dans \mathbf{R} , et si Φ est une application continue de \mathbf{R}^N dans \mathbf{R} , la fonction composée $x \mapsto \Phi(f_1(x), \dots, f_N(x))$ est mesurable sur X .

Ensuite, si (X, \mathcal{A}, m) est un espace mesuré, on peut définir l'intégrale d'une fonction mesurable positive f sur X par rapport à la mesure m grâce à la formule de Cavalieri :

$$\int_X f(x) dm(x) := \int_{\mathbf{R}^+} m(\{x \in X : f(x) > \lambda\}) d\lambda.$$

Comme nous l'avons déjà évoqué plus haut, tout l'intérêt de cette définition réside dans le fait 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 et le Théorème 5.4 page 98).

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 [1] pour plus de détails sur cette présentation. On trouvera dans les références [8] et [9] une présentation plus classique mais très détaillée de l'intégration abstraite, ainsi que dans l'appendice de [5].

Le formalisme développé ci-dessus est au coeur de la théorie des probabilités, même si le vocabulaire utilisé en théorie des probabilités est légèrement différent du vocabulaire utilisé ci-dessus. 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. À 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, ...)

Comme nous l'avons mentionné, le vocabulaire utilisé en théorie des probabilité est différent de celui utilisé en analyse. Si (X, \mathcal{A}, m) est un espace mesuré, l'ensemble X est, en théorie des probabilités, un *espace d'états*, les éléments de X sont des *résultats d'expériences* ou des *éventualités*, les éléments de la tribu \mathcal{A} sont des *événements aléatoires*. Lorsque la mesure

m vérifie $m(X) = 1$, on parle alors de mesure de *probabilité* et dans ce cas (X, \mathcal{A}, m) est un *espace probabilisé*. Enfin, une application mesurable de X dans \mathbf{R} sera une *variable aléatoire*.

Les différentes notions de convergence de suites de fonctions (ou de variables aléatoires) portent aussi des noms différents. Par exemple, une suite de fonctions mesurables $(f_n)_{n \geq 0}$ qui converge simplement vers une fonction f , p.p. sur X est une *suite de variables aléatoires qui converge presque sûrement vers une variable aléatoire f* .

7. Théorème de récurrence de Poincaré

Donnons une application intéressante de la notion de mesure à l'étude des systèmes dynamiques. Face à l'impossibilité de résoudre explicitement les équations différentielles non linéaires gouvernant les trajectoires des planètes, H. Poincaré a développé des outils permettant de décrire qualitativement les solutions de ces équations.⁷ Dans son mémoire *Sur le problème des trois corps et les équations de la dynamique*, il écrit : *Je n'ai pu résoudre rigoureusement et complètement le problème de la stabilité du système solaire, en entendant ce mot dans un sens strictement mathématique. L'emploi des invariants intégraux m'a cependant permis d'atteindre certains résultats partiels, s'appliquant surtout au problème dit restreint où les deux corps principaux circulent dans des orbites sans excentricité, pendant que le corps troublé a une masse négligeable. Dans ce cas, si on laisse de côté certaines trajectoires exceptionnelles, dont la réalisation est infiniment peu probable, on peut démontrer que le système repassera une infinité de fois aussi près que l'on voudra de sa position initiale. C'est ce que j'ai appelé la « stabilité à la Poisson ».*

Soit (X, \mathcal{A}, μ) un espace mesuré de mesure finie (i.e. $\mu(X) < +\infty$) et soit

$$S : X \rightarrow X,$$

une application mesurable, c'est-à-dire que, pour tout ensemble mesurable $A \in \mathcal{A}$, $S^{-1}(A) \in \mathcal{A}$. On suppose que S préserve la mesure μ , c'est-à-dire que

$$\mu(S^{-1}(A)) = \mu(A),$$

pour tout $A \in \mathcal{A}$.

Exemple 7.11 (La transformation du boulanger). On considère $X := [0, 1] \times [0, 1]$ muni de la mesure de Lebesgue et de la tribu des ensembles mesurables. L'application $S : [0, 1]^2 \rightarrow [0, 1]^2$ définie par

$$S(x, y) := \begin{cases} (2x, (1+y)/2) & x \in [0, 1/2] \\ (2(1-x), (1-y)/2) & x \in [1/2, 1], \end{cases}$$

est mesurable et préserve la mesure.

Théorème 7.3. *Soit $A \in \mathcal{A}$ un ensemble mesurable de X . Alors, l'ensemble des points $x \in A$ pour lesquels il existe $n \in \mathbf{N} - \{0\}$ tel que $S^k(x) \notin A$, pour tout $k \geq n$, est de mesure nulle.*

Autrement dit, pour presque tout $x \in A$, l'orbite de x , $(S^n(x))_{n \geq 0}$ passe par A une infinité de fois. Citons une fois de plus H. Poincaré : *On peut dire que les [trajectoires non récurrentes] sont l'exception et que les [trajectoires récurrentes] sont la règle au même titre que les nombres rationnels sont l'exception et les nombres incommensurables sont la règle. Je démontre en effet*

7. Le lecteur intéressé trouvera une analyse très intéressante du théorème de récurrence de Poincaré dans l'article d'Etienne Ghys *Variations autour du théorème de récurrence de Poincaré* <http://www.umpa.ens-lyon.fr/~ghys/articles/recurrence.pdf>, article dont nous avons repris quelques citations

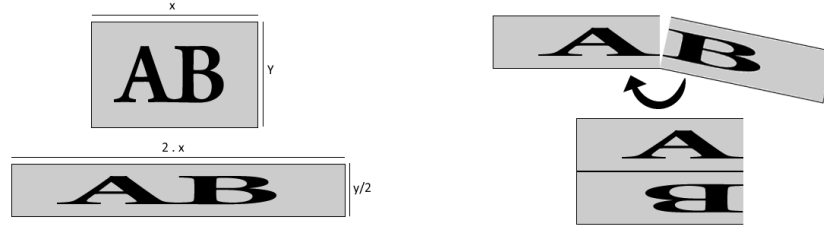


FIGURE 7.3. Transformation du boulanger.

que la probabilité pour que les conditions initiales du mouvement soient celles qui correspondent à une solution instable [non récurrente], que cette probabilité, dis-je, est nulle. Ce mot n'a par lui même aucun sens : j'en donne dans mon Mémoire une définition précise.

DÉMONSTRATION. On note

$$E_n := \{x \in A : \forall k \geq n, S^k(x) \notin A\},$$

L'objectif est de montrer que

$$E := \bigcup_{n \geq 0} E_n,$$

qui correspond à l'ensemble des points de A dont l'orbite ne passe qu'un nombre fini de fois pas A , est de mesure nulle.

On remarque que, pour tous $k_2 > k_1$

$$S^{-k_2 n}(E_n) \cap S^{-k_1 n}(E_n) = \emptyset.$$

En effet, si $x \in S^{-k_2 n}(E_n) \cap S^{-k_1 n}(E_n)$ alors

$$S^{k_2 n}(x) = S^{(k_2 - k_1)n}(S^{k_1 n}(x)) \in A,$$

et $S^{k_1 n}(x) \in E_n$ ce qui constitue une contradiction.

On en déduit que les ensembles mesurables $S^{-kn}(E_n)$, pour $k \in \mathbf{N}$, sont deux à deux disjoints, en particulier

$$\sum_{k \geq 0} \mu(S^{-kn}(E_n)) = \mu\left(\bigcup_{k \geq 0} S^{-kn}(E_n)\right) \leq \mu(X) < +\infty,$$

L'application S préservant la mesure des ensembles, on a $\mu(E_n) = \mu(S^{-kn}(E_n))$. Donc, nécessairement $\mu(E_n) = 0$ et l'on conclut que E est de mesure nulle comme réunion dénombrable d'ensembles de mesure nulle. \square

D'autres développements et applications seront donnés dans le cours de *Systèmes dynamiques* MAT 551.

8. Dérivabilité et intégration

Maintenant que nous avons généralisé la notion d'intégrale à une classe de fonction assez générale, on s'intéresse aux liens entre dérivabilité et intégration. Plus précisément, on souhaite aborder les deux problèmes suivants :

(i) Soit f une fonction intégrable sur l'intervalle $[a, b]$. Pour tout $x \in [a, b]$, on note

$$F(x) := \int_{[a, x]} f(t) dt.$$

Si f est continue sur $[a, b]$, alors F est dérivable et $F' = f$ en tout point de $[a, b]$. Plus généralement, si f est seulement intégrable, peut-on toujours affirmer que F est dérivable (ou au moins dérivable p.p.) et si tel est le cas, peut-on dire que $F' = f$ sur $[a, b]$ (ou au moins que $F' = f$ p.p. sur $[a, b]$) ?

(ii) Sous quelle condition, portant sur la fonction F définie sur $[a, b]$, peut-on affirmer que F' , la dérivée de F , existe (ou au moins existe p.p.) et sous quelle condition portant sur F l'égalité

$$F(b) - F(a) = \int_{[a, b]} F'(t) dt,$$

est-elle vérifiée ? Bien entendu, le résultat est vrai si l'on suppose que F est de classe \mathcal{C}^1 sur $[a, b]$.

Énonçons maintenant deux résultats qui sont des applications directes des Théorèmes de convergence démontrés dans le Chapitre 5.

Lemme 7.1. *On suppose que $F \in \mathcal{C}([a, b])$ est une fonction dérivable en tout point de $[a, b]$ et qu'il existe une constante $M > 0$ telle que $|F'(x)| \leq M$ pour tout $x \in [a, b]$. Alors $F' \in \mathcal{L}^1([a, b])$ et*

$$\int_{[a, b]} F'(t) dt = F(b) - F(a).$$

DÉMONSTRATION. Définissons, pour tout $n \in \mathbf{N}$ assez grand

$$G_n(x) := n(F(x + 1/n) - F(x)).$$

On a clairement

$$(7.1) \quad \int_{[a, b-1/n]} G_n(x) dx = n \left(\int_{[b-1/n, b]} F(x) dx - \int_{[a, a+1/n]} F(x) dx \right),$$

pour tout $n \in \mathbf{N}$ assez grand. Il suffit maintenant d'appliquer le Théorème de la convergence dominée de Lebesgue pour démontrer que la fonction F' est intégrable et, par passage à la limite dans (7.1), que

$$\int_{[a, b]} F'(t) dt = F(b) - F(a).$$

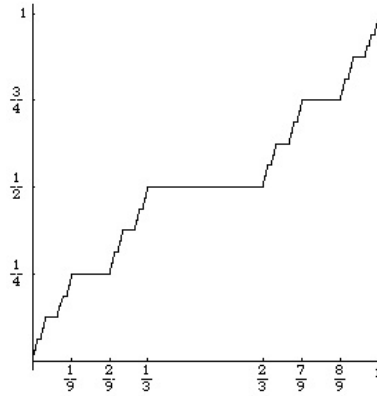
Ce qui termine la démonstration. □

Donnons un autre résultat qui est une simple conséquence du Lemme de Fatou.

Lemme 7.2. *On suppose que F est une fonction croissante sur $[a, b]$ et dérivable presque partout sur $[a, b]$. Alors, F' est intégrable et*

$$\int_{[a, b]} F'(t) dt \leq F(b) - F(a).$$

Ce résultat est en quelque sorte optimal comme le montre l'exemple suivant.

FIGURE 7.4. Le graphe de la fonction F : l'escalier du Diable.

Exemple 7.12 (Fonction de Cantor). On considère la suite de fonctions $(F_n)_{n \geq 0}$ définies par récurrence de la manière suivante : $F_0(x) = x$ pour tout $x \in [0, 1]$, puis, pour tout $n \geq 0$

$$F_{n+1}(x) := \frac{1}{2} \begin{cases} F_n(3x) & \text{si } x \in [0, 1/3] \\ 1 & \text{si } x \in [1/3, 2/3] \\ 1 + F_n(3 - 2x) & \text{si } x \in [2/3, 1] \end{cases}.$$

On laisse le soin au lecteur de vérifier que chaque F_n est une fonction croissante et que la suite $(F_n)_{n \geq 0}$ converge uniformément vers une fonction F qui est croissante et constante sur le complément de l'ensemble de Cantor (voir l'exemple 5.16 page 96 pour la définition de l'ensemble de Cantor). La fonction F étant constante sur le complément de l'ensemble de Cantor (voir la Figure 7.4), elle est donc dérivable p.p. sur $[0, 1]$ et $F' = 0$ p.p. sur $[0, 1]$. On peut donc affirmer que

$$0 = \int_{[0,1]} F'(x) dx < F(1) - F(0) = 1,$$

ce qui montre que, dans le Lemme 7.2, l'inégalité peut être stricte.

On peut aussi définir la fonction de Cantor F comme étant l'unique point fixe de l'application $\Phi : \mathcal{C}([0, 1]) \rightarrow \mathcal{C}([0, 1])$ définie par

$$\Phi(f)(x) := \frac{1}{2} \begin{cases} f(3x) & \text{si } x \in [0, 1/3] \\ 1 & \text{si } x \in [1/3, 2/3] \\ 1 + f(3 - 2x) & \text{si } x \in [2/3, 1]. \end{cases}$$

Le lecteur pourra démontrer l'existence d'un tel point fixe à titre d'exercice.

Pour répondre aux deux questions posées, nous introduisons la notion de *fonction absolument continue*.

Définition 7.10. On dit qu'une fonction F définie sur $[a, b]$ est absolument continue si, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que

$$\sum_{i=0}^n |b_i - a_i| \leq \delta \quad \Rightarrow \quad \sum_{i=0}^n |F(b_i) - F(a_i)| \leq \varepsilon$$

pour tout $n \geq 1$ et pour tous les intervalles $]a_i, b_i[$ qui sont deux à deux disjoints.

Remarquons qu'une fonction absolument continue est continue et même qu'elle est uniformément continue. Donnons quelques exemples qui nous permettront de nous familiariser avec cette définition. Clairement, une fonction k -lipschitzienne sur $[a, b]$ est absolument continue sur cet intervalle. En revanche, on vérifie que la fonction de Cantor définie dans l'exemple ci-dessus n'est pas absolument continue.

Exemple 7.13. Soit $f \in \mathcal{L}^1([a, b])$, alors, la fonction

$$F(x) := \int_{[a, x]} f(t) dt,$$

est absolument continue. C'est une conséquence de l'exercice ci-dessous.

Exercice 7.1. Soient $\Omega \subset \mathbf{R}^N$ un ouvert non vide et $f \in \mathcal{L}^1(\Omega; \mathbf{C})$. Montrer que pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que, pour tout $A \subset \Omega$ mesurable

$$|A| < \alpha \quad \Rightarrow \quad \int_A |f(x)| dx < \varepsilon.$$

On a le résultat suivant qui répond entièrement à la deuxième question posée ci-dessus.

Théorème 7.4. Si f est intégrable sur $[a, b]$, alors la fonction

$$F(x) := \int_{[a, x]} f(t) dt$$

est absolument continue, dérivable p.p. sur $[a, b]$ et $F' = f$ p.p. sur $[a, b]$.

Le reste de cette section est consacré à la démonstration de ce résultat. Nous aurons besoin d'un résultat technique concernant les recouvrements.

Lemme 7.3 (Argument de recouvrement de Vitali). Soit (I_1, \dots, I_n) une famille finie d'intervalles ouverts de \mathbf{R} , qui ne sont pas nécessairement disjoints. Alors, on peut extraire de cette famille une sous-famille $(I_{i_1}, \dots, I_{i_k})$ d'intervalles deux à deux disjoints tels que

$$\left| \bigcup_{i=1}^n I_i \right| \leq 3 \sum_{j=1}^k |I_{i_j}|.$$

DÉMONSTRATION. L'idée est très simple mais la rédaction est délicate. On note I_{i_1} l'intervalle de la famille (I_1, \dots, I_n) dont la mesure est la plus grande. On note \hat{I}_{i_1} l'intervalle qui a le même centre que I_{i_1} , mais dont le diamètre est le triple de celui de I_{i_1} . On remarque que tous les intervalles de la famille (I_1, \dots, I_n) qui rencontrent I_{i_1} sont inclus dans \hat{I}_{i_1} .

On élimine de la famille (I_1, \dots, I_n) tous les intervalles qui rencontrent I_{i_1} et l'on prend, dans l'ensemble des intervalles qui nous restent, l'intervalle I_{i_2} dont la mesure est maximale. On note \hat{I}_{i_2} l'intervalle ouvert qui a le même centre que I_{i_2} mais dont le diamètre est le triple de celui de I_{i_2} . On remarque que tous les intervalles de la famille (I_1, \dots, I_n) qui rencontrent $I_{i_1} \cup I_{i_2}$ sont inclus dans $\hat{I}_{i_1} \cup \hat{I}_{i_2}$. Ensuite, on élimine de la famille (I_1, \dots, I_n) tous les intervalles qui rencontrent $I_{i_1} \cup I_{i_2}$ et l'on poursuit le processus de construction jusqu'à épuisement de la famille initiale. On obtient ainsi une sous-famille $(I_{i_1}, \dots, I_{i_k})$ d'intervalles ouverts qui sont deux à deux disjoints et tels que les intervalles de la famille initiales sont inclus dans la réunion

des \hat{I}_{i_j} , intervalles centrés au même point de les B_{i_j} mais dont le diamètre a été multiplié par 3.

En particulier

$$\left| \bigcup_{i=1}^n I_i \right| \leq \left| \bigcup_{j=1}^k \hat{I}_{i_j} \right| \leq \sum_{j=1}^k |\hat{I}_{i_j}| \leq 3 \sum_{j=1}^k |I_{i_j}|.$$

Ce qui termine la démonstration. \square

Soit f une fonction continue sur \mathbf{R} . On vérifie immédiatement que

$$\lim_{\substack{b-a \rightarrow 0 \\ x \in]a,b[}} \frac{1}{b-a} \int_{]a,b[} f(y) dy = f(x),$$

où la limite est prise sur l'ensemble des intervalles ouverts, qui contiennent le point x , et la convergence $b-a \rightarrow 0$ exprime simplement le fait que le diamètre de ces intervalles tend vers 0. Notre objectif est de démontrer que, dans le cas où la fonction f est intégrable, cette propriété est vraie pour presque tout x .

Si f est une fonction intégrable sur \mathbf{R} , on définit la *fonction maximale de Hardy-Littlewood* par

$$f^*(x) := \sup_{\substack{]a,b[\subset \mathbf{R} \\ x \in]a,b[}} \frac{1}{b-a} \int_{]a,b[} |f(y)| dy,$$

où le sup est pris sur l'ensemble des intervalles ouverts de \mathbf{R} qui contiennent le point x .

Proposition 7.8. *La fonction f^* est mesurable et, pour tout $\lambda > 0$,*

$$|\{x \in \mathbf{R} : f^*(x) > \lambda\}| \leq \frac{3}{\lambda} \int_{\mathbf{R}} |f| dx.$$

Enfin, $f^*(x) < +\infty$ pour presque tout $x \in \mathbf{R}$.

On pourra comparer l'inégalité obtenue dans cette Proposition avec l'inégalité obtenue grâce à la propriété de Markov

$$|\{x \in \mathbf{R} : g(x) > \lambda\}| \leq \frac{1}{\lambda} \int_{\mathbf{R}} |g| dx,$$

qui est valable dès que $g \in \mathcal{L}^1(\mathbf{R})$. Remarquons que, dans l'énoncé ci-dessus, nous n'affirmons pas que f^* est une fonction intégrable et, en fait, il peut s'avérer que ce ne soit pas le cas.

DÉMONSTRATION. Montrons que, pour tout $\lambda \in \mathbf{R}$,

$$E_\lambda := \{x \in \mathbf{R} : f^*(x) > \lambda\},$$

est un ouvert de \mathbf{R} . En effet, si $f^*(x) > \lambda$, il existe un intervalle ouvert I contenant x tel que

$$\frac{1}{|I|} \int_I |f(y)| dy > \lambda.$$

Mais dans ce cas $f^*(y) > \lambda$ pour tous les $y \in I$, ce qui montre que E_λ est un ouvert. Le résultat de la Proposition 7.3 page 130 nous assure alors que la fonction f^* est mesurable.

Soit $K \subset E_\lambda$ un compact. Pour tout $x \in K$, il existe un intervalle ouvert, notée I_x tel que

$$\frac{1}{|I_x|} \int_{I_x} |f(y)| dy > \lambda$$

Clairement $(I_x)_{x \in K}$ constitue un recouvrement de K par des ouverts. L'ensemble K étant compact, on peut extraire de ce recouvrement, un sous-recouvrement fini I_{x_1}, \dots, I_{x_n} . On applique alors le Lemme 7.3 page 149 pour estimer

$$|K| \leq \left| \bigcup_{i=1}^n I_{x_i} \right| \leq 3 \sum_{j=1}^k |I_{x_{i_j}}|.$$

Or, si $x \in K \subset E_\lambda$, on a

$$|I_x| \leq \frac{1}{\lambda} \int_{I_x} |f(y)| dy.$$

On peut donc majorer

$$|K| \leq \frac{3}{\lambda} \sum_{j=1}^k \int_{I_{x_{i_j}}} |f(y)| dy.$$

Étant donné que les intervalles $I_{x_{i_j}}$ sont deux à deux disjoints, on conclut que

$$|K| \leq \frac{3}{\lambda} \int_{\mathbf{R}} |f| dx.$$

Nous avons donc démontré que la mesure de tout compact K inclus dans E_λ est majorée par une constante qui ne dépend que de λ et de f . On sait que E_λ est un ouvert. On peut donc l'écrire comme une réunion dénombrable d'ouverts disjoints

$$E_\lambda = \bigcup_{j \in J} U_j,$$

où J est dénombrable et les U_j sont des intervalles. De plus, la mesure de E_λ est égale à la somme des mesures de ces intervalles U_j (voir l'exemple 7.1 page 133). Il est alors facile de démontrer que, pour tout $\varepsilon > 0$ il existe un compact $K_\varepsilon \subset E$ tel que $|E - K_\varepsilon| \leq \varepsilon$ (voir aussi la Proposition 7.5 page 133). En utilisant cette propriété, on conclut que

$$|E_\lambda| \leq \frac{3}{\lambda} \int_{\mathbf{R}} |f| dx.$$

Ce qui termine la démonstration. Le fait que $f^* < +\infty$ p.p. sur \mathbf{R} est une conséquence de cette inégalité en faisant tendre λ vers $+\infty$. \square

Le résultat précédent nous permet de démontrer le :

Théorème 7.5 (Théorème de dérivation de Lebesgue). *Soit f une fonction intégrable sur \mathbf{R} , alors*

$$(7.2) \quad \lim_{\substack{|I| \rightarrow 0 \\ x \in I}} \frac{1}{|I|} \int_I f(y) dy = f(x),$$

pour presque tout $x \in \mathbf{R}$, où la limite est entendue sur les intervalles ouverts qui contiennent x et dont le diamètre tend vers 0.

DÉMONSTRATION. Pour tout $\lambda > 0$, on définit

$$E_\lambda := \left\{ x \in \mathbf{R} : \overline{\lim}_{\substack{|I| \rightarrow 0 \\ x \in I}} \left| \frac{1}{|I|} \int_I f(y) dy - f(x) \right| > 2\lambda \right\}.$$

Montrons que E_λ est de mesure nulle. La fonction f est intégrable. Donc, par construction de $\mathcal{L}^1(\mathbf{R})$, pour tout $\varepsilon > 0$, il existe une fonction continue à support compact g telle que

$$\int_{\mathbf{R}} |f - g| dx < \varepsilon.$$

Nous renvoyons à la démonstration du Théorème 8.3 page 162 pour les détails.

On peut alors écrire

$$\frac{1}{|I|} \int_I f(y) dy - f(x) = \frac{1}{|I|} \int_I (f(y) - g(y)) dy + \frac{1}{|I|} \int_I g(y) dy - f(x).$$

La fonction g étant continue, on a

$$\lim_{\substack{|I| \rightarrow 0 \\ I \ni x}} \frac{1}{|I|} \int_I g(y) dy = g(x),$$

pour tout $x \in \mathbf{R}$. En particulier nous avons la majoration

$$\lim_{|I| \rightarrow 0: x \in I} \left| \frac{1}{|I|} \int_I f(y) dy - f(x) \right| \leq (f - g)^*(x) + |g(x) - f(x)|.$$

Enfin, on note

$$F_\lambda := \{x \in \mathbf{R} : (f - g)^*(x) > \lambda\} \quad \text{et} \quad G_\lambda := \{x \in \mathbf{R} : |f(x) - g(x)| > \lambda\}.$$

La propriété de Markov entraîne que

$$|G_\lambda| \leq \frac{1}{\lambda} \int_{\mathbf{R}} |f - g| dx \leq \frac{\varepsilon}{\lambda}.$$

Alors que, grâce au résultat de la Proposition 7.8, nous avons

$$|F_\lambda| \leq \frac{C}{\lambda} \int_{\mathbf{R}} |f - g| dx \leq 3 \frac{\varepsilon}{\lambda}.$$

Conclusion, pour tout $\varepsilon > 0$, nous avons démontré que $|E_\lambda| \leq 4 \frac{\varepsilon}{\lambda}$. Cette inégalité étant vraie pour tout $\varepsilon > 0$, on conclut que $|E_\lambda| = 0$. Enfin, on remarque que $\bigcup_{n \geq 1} E_{1/n}$ est de mesure nulle, comme réunion dénombrable d'ensembles de mesure nulle, et donc que (7.2) est vraie pour tous les points dans le complémentaire de cet ensemble, i.e. pour presque tout point de \mathbf{R} . \square

Le Théorème 7.4 page 149 est maintenant un simple corollaire du Théorème de dérivation de Lebesgue. Le résultat suivant permet de répondre complètement à la deuxième question posée dans l'introduction de cette section :

Théorème 7.6. *Soit F une fonction absolument continue sur $[a, b]$. Alors, F est dérivable presque partout et F' est une fonction intégrable. De plus*

$$F(x) - F(a) = \int_{[a, x]} F'(t) dt,$$

pour tout $x \in [a, b]$.

La démonstration de ce résultat est abordable avec les outils développés dans ce cours mais elle est assez technique. Elle est basée sur le fait qu'une fonction absolument continue sur un intervalle $[a, b]$ est en fait la différence de deux fonctions continues croissantes et sur le

fait que les fonctions croissantes sont dérivables presque partout. Plus précisément, si F est une fonction absolument continue, on définit la *variation totale de F* entre a et $x \in [a, b]$ par

$$V_F(x) := \sup_{a=t_0 < \dots < t_n=x} \sum_{i=0}^n |F(t_{i+1}) - F(t_i)|,$$

où le sup est entendu sur toutes les subdivisions de l'intervalle $[a, b]$ (donc pour tout $n \geq 1$). On laisse au lecteur le soin de vérifier que V_F est une fonction (continue) croissante et que $F - V_F$ est une fonction (continue) décroissante. Ce qui montre que F est la différence entre deux fonctions croissantes (continues). Ensuite, on utilise le résultat suivant :

Théorème 7.7. *Soit G une fonction continue croissante sur $[a, b]$, alors G est dérivable presque partout sur $[a, b]$.*

Comme nous l'avons vu ci-dessus, le Lemme de Fatou permet d'affirmer que G' , la dérivée de G qui est définie presque partout, est intégrable sur $[a, b]$. On conclut que la fonction F est dérivable presque partout et que F' est une fonction intégrable. On peut alors définir la fonction

$$H(x) := F(x) - F(a) - \int_{[a,x]} F'(y) dy.$$

Le Théorème 7.4 nous assure que H est une fonction qui est dérivable presque partout et que $H'(x) = 0$ pour presque tout $x \in [a, b]$. L'exemple de la fonction de Cantor montre qu'en général, le fait que $H' = 0$ p.p. sur $[a, b]$ ne permet pas de conclure que H est une fonction constante mais, dans le cas où la fonction H est absolument continue, ce résultat est vrai. Le lecteur intéressé par les démonstrations pourra consulter [10].

CHAPITRE 8

Espaces de Lebesgue

1. Motivation

Dans tout ce chapitre, Ω désigne un ouvert non vide de \mathbf{R}^N . En partant de l'espace $\mathcal{C}_c(\Omega)$, on a construit dans le Chapitre 5 page 69 l'intégrale de Lebesgue et l'espace vectoriel $\mathcal{L}^1(\Omega)$ des fonctions intégrables sur l'ouvert Ω 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 intégrables 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 intégrables, les méthodes de la topologie, c'est-à-dire de considérer les fonctions intégrables 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 $\mathcal{C}_c(\Omega)$, la *convergence en moyenne* des suites ou séries de fonctions est définie par la norme

$$\mathcal{N}_1(f) := \int_{\Omega} |f(x)| dx.$$

Malheureusement, l'espace vectoriel $\mathcal{C}_c(\Omega)$ muni de la norme \mathcal{N}_1 n'est pas complet.

Exemple 8.1. Pour tout $n \geq 1$, on définit $f_n : \mathbf{R} \rightarrow \mathbf{R}$ comme la fonction impaire telle que

$$f_n(x) := \min(1, nx),$$

pour tout $x \geq 0$. La fonction f_n est continue et, pour $n \geq m \geq 1$, on a

$$\mathcal{N}_1(f_n - f_m) = \int_{\mathbf{R}} |f_n(x) - f_m(x)| dx = 2 \int_0^{+\infty} |f_n(x) - f_m(x)| dx = \frac{1}{m} - \frac{1}{n}.$$

La suite $(f_n)_{n \geq 1}$ est donc une suite de Cauchy pour la norme \mathcal{N}_1 , puisque, pour tous $m, n \geq n_0$

$$\mathcal{N}_1(f_n - f_m) \leq \frac{1}{n_0},$$

tend vers 0 lorsque n_0 tend vers $+\infty$.

De plus, on remarque que la suite $(f_n)_{n \geq 1}$ converge, pour la norme \mathcal{N}_1 , vers la fonction f définie par

$$f(x) := \begin{cases} +1 & \text{si } x > 0, \\ -1 & \text{si } x < 0, \end{cases}$$

puisque

$$\int_{-1}^1 |f_n(x) - f(x)| dx = 2 \int_0^{1/n} |1 - nx| dx = \frac{1}{n}.$$

Pour conclure, la suite $(f_n - f_1)_{n \geq 1}$ est une suite d'éléments de $\mathcal{C}_c(\mathbf{R})$, c'est une suite de Cauchy pour la norme \mathcal{N}_1 , mais elle ne converge pas vers la fonction $f - f_1$ qui n'est pas une fonction continue, et donc qui n'appartient pas à $\mathcal{C}_c(\mathbf{R})$.

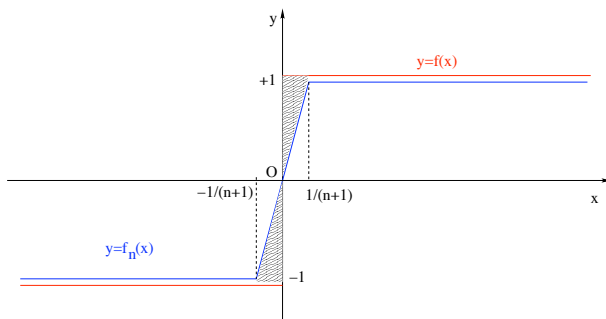


FIGURE 8.1. Graphes de f_n et de f ; la distance de f_n à f est l'aire de la surface hachurée.

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 supplémentaire pour développer une théorie de l'intégration plus puissante que celle de l'intégrale usuelle de Riemann sur $\mathcal{C}_c(\Omega)$.

Le but de ce chapitre est donc de construire un espace complet à partir de $\mathcal{L}^1(\Omega)$ et de la norme \mathcal{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.)

En quelque sorte, 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.

Afin de simplifier les notations, nous nous restreignons dans ce chapitre au cas des fonctions à valeurs réelles, l'extension des résultats aux fonctions à valeur complexes est laissée en exercice.

2. L'espace de Lebesgue L^1

L'idée naturelle est de prolonger la norme \mathcal{N}_1 de la convergence en moyenne à l'espace $\mathcal{L}^1(\Omega)$. Considérons donc l'application $\mathcal{N}_1 : \mathcal{L}^1(\Omega) \mapsto \mathbf{R}_+$ qui est définie par

$$\mathcal{N}_1(f) := \int_{\Omega} |f(x)| dx.$$

Cette application est une *semi-norme*, c'est-à-dire qu'elle vérifie :

- (i) Pour tout $f \in \mathcal{L}^1(\Omega)$ et pour tout $\lambda \in \mathbf{R}$, $\mathcal{N}_1(\lambda f) = |\lambda| \mathcal{N}_1(f)$;
- (ii) Pour tous $f, g \in \mathcal{L}^1(\Omega)$, $\mathcal{N}_1(f + g) \leq \mathcal{N}_1(f) + \mathcal{N}_1(g)$.

Mais \mathcal{N}_1 n'est pas une norme sur $\mathcal{L}^1(\Omega)$. En effet, d'après le Théorème 5.1 page 87, si $f \in \mathcal{L}^1(\Omega)$ et si $\mathcal{N}_1(f) = 0$ alors $f = 0$ p.p. sur Ω . Donc $\mathcal{N}_1(f) = 0$ n'implique pas que $f = 0$ partout sur Ω .

Cette difficulté n'existe pas sur $\mathcal{C}_c(\Omega)$, car $\mathcal{N}_1(f) = 0$ implique que $f = 0$ p.p. sur Ω , 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 5.1 page 87 et la Proposition 5.10 page 98. Dire que $f = 0$ p.p. sur Ω , c'est dire que $\{x \in \Omega : f(x) \neq 0\}$ est ensemble négligeable. La fonction f étant supposée continue, cet ensemble est un ouvert (comme image réciproque par une

application continue de $\mathbf{R} - \{0\}$ qui est un ouvert de \mathbf{R}). Le complémentaire de cet ensemble, $\{x \in \Omega : f(x) = 0\}$, est donc fermé et dense dans Ω (voir la Proposition 5.10 page 98). On conclut que cet ensemble est Ω tout entier, de sorte que $f = 0$ sur Ω .

L'idée clef pour pallier cette difficulté consiste à identifier deux fonctions intégrables égales p.p. sur Ω . C'est tout à fait naturel dans le contexte de l'intégration, puisque deux fonctions intégrables sur Ω et qui sont égales p.p. sur Ω , ont la même intégrale.

Voici comment on effectue cette identification dans le langage mathématique : On définit sur $\mathcal{L}^1(\Omega)$ la *relation d'équivalence* \simeq par

$$f \simeq g \quad \text{si et seulement si} \quad (f - g = 0 \quad \text{p.p. sur } \Omega).$$

On invite le lecteur à vérifier que l'on a bien là une relation d'équivalence.

Observons que les opérations élémentaires sur $\mathcal{L}^1(\Omega)$ sont compatibles avec cette relation d'équivalence :

(1) Pour tous $\alpha, \beta \in \mathbf{R}$ et $f_1, f_2, g_1, g_2 \in \mathcal{L}^1(\Omega)$

$$f_1 \simeq g_1 \quad \text{et} \quad f_2 \simeq g_2 \quad \Rightarrow \quad \alpha f_1 + \beta f_2 \simeq \alpha g_1 + \beta g_2.$$

De plus l'intégrale de Lebesgue et la semi-norme \mathcal{N}_1 sont également compatibles avec cette relation d'équivalence :

(2) Pour tout $f, g \in \mathcal{L}^1(\Omega)$

$$f \simeq g \quad \Rightarrow \quad \int_{\Omega} f(x) dx = \int_{\Omega} g(x) dx \quad \text{et} \quad \mathcal{N}_1(f) = \mathcal{N}_1(g).$$

Définition 8.1 (Espace de Lebesgue L^1). *Pour tout $f \in \mathcal{L}^1(\Omega)$, on note*

$$[f] := \{\phi \in \mathcal{L}^1(\Omega) : \phi \simeq f\},$$

la classe d'équivalence de f . L'espace de Lebesgue $L^1(\Omega)$ est défini par

$$L^1(\Omega) := \{[f] : f \in \mathcal{L}^1(\Omega)\},$$

c'est l'espace quotient de $\mathcal{L}^1(\Omega)$ par le sous-espace l'espace vectoriel des fonctions nulles p.p. sur Ω .

L'espace de Lebesgue $L^1(\Omega)$ est un \mathbf{R} -espace vectoriel pour l'addition des fonctions (ou plutôt pour l'addition des classes d'équivalences de fonctions) et la multiplication par les scalaires réels (resp. complexes), qui sont définies comme suit :

Pour tous $\alpha \in \mathbf{R}$ et $f, g \in \mathcal{L}^1(\Omega)$, on pose

$$[f] + [g] := [f + g] \quad \text{et} \quad \alpha[f] := [\alpha f].$$

On vérifie grâce à la propriété (1) ci-dessus que ces définitions sont indépendantes des éléments choisis dans les classes $[f]$ et $[g]$. On peut ainsi munir $L^1(\Omega)$ d'une structure de \mathbf{R} espace vectoriel.

L'intégrale de Lebesgue définit une forme \mathbf{R} -linéaire sur $L^1(\Omega)$, comme suit :

$$L^1(\Omega) \ni [f] \quad \mapsto \quad \int_{\Omega} [f](x) dx := \int_{\Omega} f(x) dx \in \mathbf{R}.$$

La propriété (2) ci-dessus montre que cette définition est indépendante de l'élément choisi dans la classe $[f]$.

La semi-norme \mathcal{N}_1 définie sur $\mathcal{L}^1(\Omega)$, définit une semi-norme sur $L^1(\Omega)$, par la formule

$$\mathcal{N}_1([f]) := \mathcal{N}_1(f) = \int_{\Omega} |f(x)| dx.$$

Il s'agit bien d'une semi-norme, car pour tout $[f], [g] \in L^1(\Omega)$ et $\alpha \in \mathbf{R}$, on a

$$\mathcal{N}_1(\alpha[f]) = |\alpha|\mathcal{N}_1([f]) \quad \text{et} \quad \mathcal{N}_1([f] + [g]) \leq \mathcal{N}_1([f]) + \mathcal{N}_1([g]).$$

Mais l'avantage de cette construction est que la semi-norme \mathcal{N}_1 ainsi définie sur $L^1(\Omega)$ est une norme, car

$$\mathcal{N}_1([f]) = 0 \quad \Leftrightarrow \quad \int_{\Omega} |f(x)| dx = 0 \quad \Leftrightarrow \quad f = 0 \text{ p.p. sur } \Omega \quad \Leftrightarrow \quad f \simeq 0 \quad \Leftrightarrow \quad [f] = 0.$$

Convention de langage et notations : Dorénavant, on identifiera toujours l'élément $[f]$ de $L^1(\Omega)$ avec n'importe laquelle des fonctions de $[f]$ — par exemple f elle-même. De la même façon que l'on ne pense pas à un élément de $\mathbf{Z}/n\mathbf{Z}$ comme à un sous-ensemble de \mathbf{Z} , aucun mathématicien ne pense à un élément de $L^1(\Omega)$ comme à une classe d'équivalence de fonctions. On pense donc à un élément de $L^1(\Omega)$ 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

$$f = g \quad \text{p.p. sur } \Omega.$$

Autrement dit, la définition rigoureuse de $L^1(\Omega)$ 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 $L^1(\Omega)$ comme à des fonctions définies p.p. sur Ω , en remplaçant la notion habituelle d'égalité de fonctions définies partout sur Ω , par la notion d'égalité p.p. sur Ω de fonctions définies 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.

Notation : il est d'usage de noter la norme \mathcal{N}_1 ainsi définie sur l'espace de Lebesgue $L^1(\Omega)$ par

$$\| \cdot \|_1, \quad \| \cdot \|_{L^1} \quad \text{ou encore} \quad \| \cdot \|_{L^1(\Omega)},$$

selon le degré d'ambiguïté du contexte.

Ce que l'on a gagné avec cette construction, c'est évidemment que $L^1(\Omega)$ 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 $x_0 \in \Omega$ 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 $p(x)$ au 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 un domaine régulier $A \subset \mathbf{R}^3$ — 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, 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 — nous renvoyons, pour plus de détails, au 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 par rapport aux théories antérieures de l'intégration tient au résultat suivant.

Théorème 8.1 (Complétude de L^1 , Théorème de Fischer-Riesz). *L'espace $L^1(\Omega)$ 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'en se limitant à la notion usuelle d'intégrale définie sur les fonctions continues, c'est un peu comme si l'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.

Évidemment, 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 XIX^{ème} 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 8.2 (Réciproque du Théorème de la convergence dominée). *Soit $(f_n)_{n \geq 0}$ suite de Cauchy de $\mathcal{L}^1(\Omega)$. Il existe alors une suite extraite $(f_{\varphi(n)})_{n \geq 0}$ vérifiant les hypothèses du Théorème de la convergence dominée. C'est-à-dire qu'il existe $f, F \in \mathcal{L}^1(\Omega)$ tels que :*

- (i) $\lim_{n \rightarrow +\infty} f_{\varphi(n)} = f$ p.p. sur Ω ;
- (ii) pour tout $n \geq 0$, $|f_{\varphi(n)}| \leq F$ p.p. sur Ω .

De plus

$$\lim_{n \rightarrow +\infty} \int_{\Omega} |f_n - f| dx = 0.$$

Outre le fait que ce dernier théorème entraîne la complétude de l'espace de Lebesgue $L^1(\Omega)$, dont nous avons déjà souligné l'importance, il constitue une quasi-réciproque du Théorème de

la convergence dominée de Lebesgue. 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 la convergence dominée de Lebesgue.

DÉMONSTRATION. Comme $(f_n)_{n \geq 0}$ est une suite de Cauchy dans $\mathcal{L}^1(\Omega)$, pour tout $\varepsilon > 0$, il existe $N(\varepsilon) \geq 0$ tel que, pour tous $m, n \geq N(\varepsilon)$, on a

$$\int_{\Omega} |f_m - f_n| dx \leq \varepsilon.$$

Choisissons $\varepsilon = 1$, et posons $\varphi(0) := N(1)$. Alors, pour tout $m \geq \varphi(0)$, on a

$$\int_{\Omega} |f_m - f_{\varphi(0)}| dx \leq 1.$$

Puis, prenons $\varepsilon = 1/2$, on pose $\varphi(1) := \max(\varphi(0) + 1, N(1/2))$, de sorte que, pour tout $m \geq \varphi(1)$, on a

$$\int_{\Omega} |f_m - f_{\varphi(1)}| dx \leq \frac{1}{2}.$$

Par récurrence, si $\varphi(n)$ est construit, on considère $\varepsilon = 1/2^{n+1}$ et l'on pose $\varphi(n+1) := \max(\varphi(n) + 1, N(1/2^{n+1}))$, de telle sorte que, pour tout $m \geq \varphi(n+1)$, on a

$$\int_{\Omega} |f_m - f_{\varphi(n+1)}| dx \leq \frac{1}{2^{n+1}}.$$

Par construction

$$\int_{\Omega} |f_{\varphi(n+1)} - f_{\varphi(n)}| dx \leq \frac{1}{2^n}.$$

Définissons, pour presque tout $x \in \Omega$, la fonction $\Phi : \Omega \mapsto [0, +\infty]$, par

$$\Phi = \sum_{n \geq 0} |f_{\varphi(n+1)} - f_{\varphi(n)}|.$$

La fonction Φ est mesurable (comme somme d'une série de terme général mesurable, d'après le Théorème 5.6 page 100), à valeurs dans $[0, +\infty]$; d'autre part, le Théorème de la convergence monotone de Beppi Levi nous assure que

$$\int_{\Omega} \Phi(x) dx = \sum_{n \geq 0} \int_{\Omega} |f_{\varphi(n+1)}(x) - f_{\varphi(n)}(x)| dx \leq \sum_{n \geq 0} \frac{1}{2^n} = 2,$$

de sorte que $\Phi \in \mathcal{L}^1(\Omega)$. D'après la Proposition 5.8 page 88, il existe donc un sous-ensemble négligeable $\mathcal{Z} \subset \Omega$ tel que $\Phi(x) < +\infty$ pour tout $x \in \Omega - \mathcal{Z}$.

Considérons alors la série de fonctions

$$\sum_{n \geq 0} (f_{\varphi(n+1)} - f_{\varphi(n)}).$$

D'après ce qui précède, pour tout $x \in \Omega - \mathcal{Z}$, la série ci-dessus est absolument convergente dans \mathbf{R} (qui est complet) et donc, elle converge. Par conséquent, comme

$$f_{\varphi(n)} = f_{\varphi(0)} + \sum_{k=0}^{n-1} (f_{\varphi(k+1)} - f_{\varphi(k)}),$$

on conclut que la suite $(f_{\varphi(n)})_{n \geq 0}$ converge vers une fonction f en tout point de $\Omega - \mathcal{Z}$, donc p.p. sur Ω . Enfin

$$|f_{\varphi(n)}| \leq |f_{\varphi(0)}| + \Phi$$

p.p. sur Ω et $F := |f_{\varphi(0)}| + \Phi \in \mathcal{L}^1(\Omega)$. La suite extraite $(f_{\varphi(n)})_{n \geq 0}$ vérifie donc toutes les hypothèses du Théorème de la convergence dominée de Lebesgue.

En particulier

$$\lim_{n \rightarrow +\infty} |f_{\varphi(n)} - f| = 0,$$

p.p. sur Ω et

$$|f_{\varphi(n)} - f| = \sum_{k \geq n} |f_{\varphi(k+1)} - f_{\varphi(k)}| \leq \Phi,$$

p.p. sur Ω . Le Théorème de la convergence dominée de Lebesgue nous assure que

$$\lim_{n \rightarrow +\infty} \int_{\Omega} |f_{\varphi(n)}(x) - f(x)| dx = 0.$$

Conclusion, en partant d'une suite de Cauchy $(f_n)_{n \geq 0}$ dans $\mathcal{L}^1(\Omega)$, on a construit une suite extraite $(f_{\varphi(n)})_{n \geq 0}$ qui vérifie de plus les hypothèses du Théorème de la convergence dominée de Lebesgue et telle que

$$\lim_{n \rightarrow +\infty} \int_{\Omega} |f_{\varphi(n)} - f| dx = 0.$$

En utilisant le fait que $(f_n)_{n \geq 0}$ est une suite de Cauchy, on vérifie que

$$\lim_{n \rightarrow +\infty} \int_{\Omega} |f_n - f| dx = 0.$$

Ce qui termine la démonstration. □

Le théorème ci-dessus implique en particulier que, pour toute suite $(f_n)_{n \geq 0}$ d'éléments de $\mathcal{L}^1(\Omega)$ pour laquelle il existe $f \in \mathcal{L}^1(\Omega)$ telle que

$$\lim_{n \rightarrow +\infty} \int_{\Omega} |f_n - f| dx = 0,$$

il existe une sous-suite qui converge vers f p.p. sur Ω . À nouveau ce résultat est à peu près optimal, dans la mesure où une suite convergente dans $\mathcal{L}^1(\Omega)$ n'est pas nécessairement convergente p.p. sur Ω .

Exemple 8.2. Pour tout $n \geq 1$ et pour tout $k = 0, \dots, n-1$ on note

$$f_{n,k}(x) = \mathbf{1}_{]k/n, (k+1)/n[}.$$

On considère la suite

$$g_0 := f_{1,0}, \quad g_1 := f_{2,0}, \quad g_2 := f_{2,1}, \quad g_3 := f_{3,0}, \quad g_4 := f_{3,1}, \quad g_5 := f_{3,2}, \dots$$

et plus généralement

$$g_{\frac{n(n-1)}{2} + k} = f_{n,k},$$

pour tout $n \geq 0$ et pour tout $k = 0, \dots, n-1$. On vérifie que $g_m \in \mathcal{L}^1(]0, 1[)$, que

$$\lim_{m \rightarrow +\infty} \int_{]0, 1[} g_m(x) dx = 0.$$

Mais la suite $(g_m)_{m \geq 0}$ ne tend pas vers 0 p.p. sur $]0, 1[$.

3.2. Densité de \mathcal{C}_c dans L^1 . À partir de $\mathcal{C}_c(\Omega)$, 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 si $L^1(\Omega)$ est le plus petit espace possible ayant ces propriétés. Une réponse à cette question est fournie par le :

Théorème 8.3 (Densité de $\mathcal{C}_c(\Omega)$ dans $L^1(\Omega)$). *L'espace $\mathcal{C}_c(\Omega)$ s'identifie à un sous-espace dense de $L^1(\Omega)$. Autrement dit, pour tout $\varepsilon > 0$ et toute fonction f intégrable sur Ω , il existe $f_\varepsilon \in \mathcal{C}_c(\Omega)$ telle que*

$$\|f - f_\varepsilon\|_{L^1(\Omega)} \leq \varepsilon.$$

DÉMONSTRATION. Montrons dans un premier temps que $\mathcal{C}_c(\Omega)$ s'identifie à un sous-espace de $L^1(\Omega)$. Pour cela, considérons l'application

$$\mathcal{C}_c(\Omega) \ni \phi \mapsto [\phi] \in L^1(\Omega).$$

Cette application est évidemment linéaire ; montrons qu'elle est injective. En effet, dire qu'une fonction $\phi \in \mathcal{C}_c(\Omega)$ 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 la Proposition 5.10 page 98), donc $\phi \equiv 0$. Par conséquent, $\mathcal{C}_c(\Omega)$ est isomorphe à son image par l'application linéaire injective ci-dessus, qui est un sous-espace vectoriel de $L^1(\Omega)$ et auquel on identifiera désormais $\mathcal{C}_c(\Omega)$.

Démontrons maintenant la densité de $\mathcal{C}_c(\Omega)$ dans $L^1(\Omega)$. Soit $[f] \in L^1(\Omega)$ et $f \in \mathcal{L}^1(\Omega)$ une fonction intégrable de cette classe. Par définition de $\mathcal{L}^1(\Omega)$, il existe deux fonctions $g, h \in \mathcal{L}^+(\Omega)$ telles que $f = g - h$ p.p. sur Ω . Soient $(g_n)_{n \geq 0}$ et $(h_n)_{n \geq 0}$, deux suites croissantes de fonctions définies sur Ω qui convergent simplement, respectivement vers g et h et telles que

$$\lim_{n \rightarrow +\infty} \int_{\Omega} |g(x) - g_n(x)| dx = 0, \quad \text{et} \quad \lim_{n \rightarrow +\infty} \int_{\Omega} |h(x) - h_n(x)| dx = 0.$$

Posons alors, $f_n := g_n - h_n$. Évidemment $f_n \in \mathcal{C}_c(\Omega)$ puisque $g_n, h_n \in \mathcal{C}_c(\Omega)$, de plus

$$\lim_{n \rightarrow +\infty} \int_{\Omega} |f(x) - f_n(x)| dx \leq \lim_{n \rightarrow +\infty} \left(\int_{\Omega} |g(x) - g_n(x)| dx + \int_{\Omega} |h(x) - h_n(x)| dx \right) = 0.$$

Étant donné $\varepsilon > 0$, il existe donc $n_\varepsilon > 0$ tel que

$$\int_{\Omega} |f(x) - f_{n_\varepsilon}(x)| dx \leq \varepsilon.$$

Autrement dit

$$\|[f] - [f_{n_\varepsilon}]\|_{L^1(\Omega)} = \int_{\Omega} |f(x) - f_{n_\varepsilon}(x)| dx \leq \varepsilon.$$

□

La densité de $\mathcal{C}_c(\Omega)$ dans $L^1(\Omega)$, bien qu'étant une conséquence directe de la définition même de la notion de fonction intégrable, a de nombreuses conséquences très importantes. En voici une que nous utiliserons à plusieurs reprises dans le chapitre sur l'analyse de Fourier.

Corollaire 8.1 (Continuité L^1 des translations). *Pour toute fonction f intégrable sur \mathbf{R}^N*

$$\lim_{|y| \rightarrow 0} \int_{\mathbf{R}^N} |f(x-y) - f(x)| dx = 0.$$

En général, une fonction intégrable 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 5 page 69, qui est intégrable 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 $L^1(\mathbf{R})$, qui est la trace résiduelle de la continuité des fonctions à partir desquelles on a construit l'espace des fonctions intégrables, ou l'espace de Lebesgue $L^1(\mathbf{R})$.

DÉMONSTRATION. Soit $\varepsilon > 0$. Par densité de $\mathcal{C}_c(\mathbf{R}^N)$ dans $L^1(\mathbf{R}^N)$, il existe $f_\varepsilon \in \mathcal{C}_c(\mathbf{R})$ telle que

$$\|f - f_\varepsilon\|_{L^1(\mathbf{R})} \leq \frac{\varepsilon}{3}.$$

Soit $r > 0$ tel que le support de la fonction f_ε soit inclus dans $B_f(0, r)$. Alors

$$\begin{aligned} \int_{\mathbf{R}^N} |f(x-y) - f(x)| dx &\leq \int_{\mathbf{R}^N} |f(x-y) - f_\varepsilon(x-y)| dx \\ &+ \int_{\mathbf{R}^N} |f_\varepsilon(x-y) - f_\varepsilon(x)| dx \\ &+ \int_{\mathbf{R}^N} |f_\varepsilon(x) - f(x)| dx. \end{aligned}$$

En effectuant le changement de variables $x \mapsto x+z$, on vérifie que la première et troisième intégrale sont égales et, en raison du choix de f_ε , on trouve

$$\int_{\mathbf{R}^N} |f(x-y) - f(x)| dx \leq \frac{2}{3} \varepsilon + \int_{\mathbf{R}^N} |f_\varepsilon(x-y) - f_\varepsilon(x)| dx.$$

Enfin, la fonction f_ε est continue et à support compact, elle est donc uniformément continue (Théorème de Heine 3.2 page 39). En particulier, il existe $\delta > 0$ (que l'on peut choisir ≤ 1) tel que si $|y| \leq \delta$ alors

$$|f_\varepsilon(x-y) - f_\varepsilon(x)| \leq \frac{1}{3\Omega_{r+1}} \varepsilon,$$

pour tout $x \in \mathbf{R}^N$. Danc cette inégalité Ω_{r+1} désigne le volume de la boule de rayon $r+1$ dans \mathbf{R}^N .

Au total, étant donné $\varepsilon > 0$, il existe donc $\delta > 0$ tel que, si $|y| \leq \delta$ alors

$$\int_{\mathbf{R}^N} |f(x-y) - f(x)| dx \leq \varepsilon.$$

Comme $\varepsilon > 0$ peut être choisi arbitrairement petit, on conclut que

$$\lim_{|y| \rightarrow +\infty} \int_{\mathbf{R}^N} |f(x-y) - f(x)| dx = 0.$$

□

4. Généralisations

L'espace de Lebesgue $L^1(\Omega)$ est un cas particulier d'une famille d'espaces d'une importance fondamentale en analyse.

L'inégalité de Minkowski étudiée au Chapitre 6 page 109 montre que la quantité

$$\mathcal{N}_p(f) := \left(\int_{\Omega} |f(x)|^p dx \right)^{1/p},$$

définit pour tout $p \in [1, +\infty[$ une norme sur $\mathcal{C}_c(\Omega)$.

Définition 8.2 (Espace \mathcal{L}^p). Soit $p \in]1, +\infty[$. On désigne par $\mathcal{L}^p(\Omega)$ l'ensemble des fonctions mesurables à valeurs réelles, qui sont définies p.p. sur Ω , et qui vérifient la condition

$$\int_{\Omega} |f(x)|^p dx < +\infty.$$

On vérifie sans peine, grâce à l'inégalité de Minkowski, que $\mathcal{L}^p(\Omega)$ est un \mathbf{R} -espace vectoriel pour l'addition et la multiplication par les scalaires des fonctions définies p.p. sur Ω , et que \mathcal{N}_p est une semi-norme sur $\mathcal{L}^p(\Omega)$, mais pas une norme, pour la même raison que \mathcal{N}_1 n'est pas une norme sur $\mathcal{L}^1(\Omega)$.

On notera en particulier que, pour $p = 2$, la norme \mathcal{N}_2 définit la convergence en moyenne quadratique des suites de fonctions dans $\mathcal{C}_c(\Omega)$; dans ce cas, l'espace $\mathcal{C}_c(\Omega)$ muni de la norme \mathcal{N}_2 est un exemple d'espace préhilbertien, qui n'est pas complet. (On laisse au lecteur le soin de construire une suite de Cauchy de $\mathcal{C}([-1, 1])$ pour la norme \mathcal{N}_2 qui ne converge pas en moyenne quadratique vers une fonction appartenant à $\mathcal{C}_c([-1, 1])$).

Comme dans le cas $p = 1$, on construit à partir de $\mathcal{L}^p(\Omega)$, un espace pour lequel \mathcal{N}_p est une norme, en identifiant les fonctions de $\mathcal{L}^p(\Omega)$ égales p.p. sur Ω .

Définition 8.3 (Espace de Lebesgue L^p). Soit $p \in]1, +\infty[$. Pour $f \in \mathcal{L}^p(\Omega)$, on note $[f]$ l'ensemble de toutes les fonctions mesurables définies p.p. sur Ω , qui sont égales à f p.p. sur Ω , et l'on pose

$$L^p(\Omega) := \{[f] : f \in \mathcal{L}^p(\Omega)\}.$$

L'addition des éléments de $L^p(\Omega)$ et la multiplication par les scalaires est définie comme dans le cas $p = 1$: pour tous $[f], [g] \in L^p(\Omega)$ et tout $\alpha \in \mathbf{R}$, on pose

$$[f] + [g] := [f + g], \quad \alpha[f] := [\alpha f].$$

On définit ainsi une structure d'espace vectoriel sur $L^p(\Omega)$. Enfin la semi-norme \mathcal{N}_p induit sur $L^p(\Omega)$ l'application

$$\|[f]\|_{L^p} := \left(\int_{\Omega} |f(x)|^p dx \right)^{1/p},$$

qui est une norme.

Comme dans le cas $p = 1$, le fait que l'application $L^p(\Omega) \ni [f] \mapsto \|[f]\|_{L^p}$ est une norme sur $L^p(\Omega)$ 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 $L^p(\Omega)$ vérifie les propriétés résumées dans le théorème suivant.

Théorème 8.4. Soit $p \in]1, +\infty[$. Alors :

- (1) l'espace $L^p(\Omega)$ muni de la norme $\|\cdot\|_{L^p}$ est complet (Théorème de Fischer-Riesz) ;
- (2) l'espace $\mathcal{C}_c(\Omega)$ s'identifie à un sous-espace vectoriel dense dans $L^p(\Omega)$;

(3) pour tout $f \in \mathcal{L}^p(\mathbf{R}^N)$ on a

$$\lim_{|y| \rightarrow 0} \int_{\mathbf{R}^N} |f(x-y) - f(x)|^p dx = 0.$$

Autrement dit

$$\lim_{|y| \rightarrow 0} \|f(\cdot - y) - f\|_{L^p(\mathbf{R})} = 0,$$

pour tout $f \in L^p(\mathbf{R})$.

DÉMONSTRATION DU POINT (1). Soit $(f_n)_{n \geq 0}$ suite de Cauchy dans $\mathcal{L}^p(\Omega)$. Construisons, comme dans la démonstration du Théorème 8.1, une suite extraite telle que

$$\left(\int_{\Omega} |f_{\varphi(n+1)} - f_{\varphi(n)}|^p(x) dx \right)^{1/p} \leq \frac{1}{2^n}.$$

Définissons la fonction mesurable Φ comme dans la démonstration du Théorème 8.1. D'après l'inégalité de Minkowski, pour tout $n \geq 1$,

$$\left(\int_{\Omega} \left(\sum_{k=0}^{n-1} |f_{\varphi(k+1)} - f_{\varphi(k)}|(x) \right)^p dx \right)^{1/p} \leq \sum_{k=0}^{n-1} \left(\int_{\Omega} |f_{\varphi(k+1)} - f_{\varphi(k)}|^p(x) dx \right)^{1/p} \leq 2,$$

donc

$$\int_{\Omega} \left(\sum_{k=0}^{n-1} |f_{\varphi(k+1)} - f_{\varphi(k)}|(x) \right)^p dx \leq 2^p.$$

D'après le Théorème de la convergence monotone de Beppo Levi 5.9 page 89, en passant à la limite quand n tend vers $+\infty$, on trouve que

$$\begin{aligned} \int_{\Omega} |\Phi(x)|^p dx &= \int_{\Omega} \lim_{n \rightarrow +\infty} \left(\sum_{k=0}^{n-1} |f_{\varphi(k+1)} - f_{\varphi(k)}|(x) \right)^p dx \\ &\leq \lim_{n \rightarrow +\infty} \int_{\Omega} \left(\sum_{k=0}^{n-1} |f_{\varphi(k+1)} - f_{\varphi(k)}|(x) \right)^p dx \\ &\leq 2^p. \end{aligned}$$

Comme Φ^p est intégrable sur Ω , il existe un sous-ensemble négligeable $\mathcal{Z} \subset \Omega$ tel que $\Phi(x) < +\infty$ pour tout $x \in \Omega - \mathcal{Z}$.

Définissons alors la fonction mesurable ϕ comme dans la démonstration du Théorème 8.1. La série

$$\phi := \sum_{k \geq 0} (f_{\varphi(k+1)} - f_{\varphi(k)}),$$

est absolument convergente, donc convergente dans \mathbf{R} pour tout $x \in \Omega - \mathcal{Z}$, c'est-à-dire p.p. sur Ω . Elle définit donc une fonction mesurable comme limite p.p. sur Ω d'une suite de fonctions mesurables.

De plus, l'inégalité triangulaire entraîne que $|\phi| \leq \Phi$ sur $\Omega - \mathcal{Z}$, de sorte que

$$\int_{\Omega} |\phi(x)|^p dx \leq \int_{\Omega} \Phi(x)^p dx < +\infty.$$

Donc $\phi \in \mathcal{L}^p(\Omega)$.

Posons $f := f_{\varphi(0)} + \phi$. Évidemment $f \in \mathcal{L}^p(\Omega)$ comme somme de deux fonctions de $\mathcal{L}^p(\Omega)$. Ensuite, pour tous $n < m$

$$\begin{aligned} \left(\int_{\Omega} |f_{\varphi(m)} - f_{\varphi(n)}|^p(x) dx \right)^{1/p} &\leq \sum_{k=n}^{m-1} \left(\int_{\Omega} |f_{\varphi(k+1)} - f_{\varphi(k)}|^p(x) dx \right)^{1/p} \\ &\leq \sum_{k=n}^{m-1} \frac{1}{2^k} \\ &\leq \frac{1}{2^{n-1}}, \end{aligned}$$

de sorte que, d'après le Lemme de Fatou

$$\begin{aligned} \int_{\Omega} |f - f_{\varphi(n)}|^p(x) dx &= \int_{\Omega} \lim_{m \rightarrow +\infty} |f_{\varphi(m)} - f_{\varphi(n)}|^p(x) dx \\ &\leq \lim_{m \rightarrow +\infty} \int_{\Omega} |f_{\varphi(m)} - f_{\varphi(n)}|^p(x) dx \\ &\leq \frac{1}{2^{n-1}}. \end{aligned}$$

Conclusion, en revenant aux classes d'équivalence dans $L^p(\Omega)$, on a montré que la suite extraite $(f_{\varphi(n)})_{n \geq 0}$ 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 .

Démontrons maintenant le point (2). C'est-à-dire que $\mathcal{C}_c(\Omega)$ est un sous-espace vectoriel dense dans $\mathcal{L}^p(\Omega)$. En décomposant $f = f^+ - f^-$ on se ramène au cas où $f \geq 0$ p.p. sur Ω . Comme dans la démonstration du Théorème 8.3 page 162, on construit une suite $(f_n)_{n \geq 0}$ de fonctions positives dans $\mathcal{C}_c(\Omega)$ telles que $\lim_{n \rightarrow +\infty} f_n^p = f^p$ p.p. sur Ω et telles que

$$\lim_{n \rightarrow +\infty} \int_{\Omega} f_n^p dx = \int_{\Omega} f(x) dx.$$

Enfin, il suffit d'appliquer le Lemme de Brézis-Lieb, (voir l'exercice 8.1 page ci-contre), pour conclure que

$$\lim_{n \rightarrow +\infty} \int_{\Omega} f_n^p(x) dx = \int_{\Omega} f^p(x) dx.$$

Ce qui termine la démonstration du point (2). \square

Le cas où $p = 2$ et où les fonctions considérées sont à valeurs complexes, 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 hermitien¹ défini par

$$\langle \phi, \psi \rangle_{L^2} := \int_{\Omega} \phi(x) \overline{\psi(x)} dx,$$

pour tous $\phi, \psi \in L^2(\Omega; \mathbf{C})$. L'espace $L^2(\Omega; \mathbf{C})$ est un *espace préhilbertien* et de plus il est 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.

1. Charles Hermite X 1843, Professeur à l'École polytechnique en 1869.

Concluons ce chapitre par une dernière remarque. Il est possible de définir une notion d'espace \mathcal{L}^∞ — 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 $\mathcal{L}^\infty(\Omega)$ comme suit. Pour toute fonction f mesurable définie p.p. sur Ω , on note

$$\mathcal{N}_\infty(f) := \inf \{M > 0 : \{x \in \Omega : |f(x)| > M\} \text{ est négligeable}\},$$

le *supremum essentiel* de la fonction f , avec la convention que $\inf \emptyset = +\infty$. On définit $\mathcal{L}^\infty(\Omega)$ comme l'ensemble des fonctions mesurables définies p.p. sur Ω telles que $\mathcal{N}_\infty(f) < +\infty$ et $L^\infty(\Omega)$ comme l'ensemble des classes de fonctions de $\mathcal{L}^\infty(\Omega)$ égales p.p. sur Ω . Pour tout élément $[f]$ de $L^\infty(\Omega)$, on définit

$$\|[f]\|_{L^\infty} := \mathcal{N}_\infty(f),$$

laquelle est évidemment indépendante du représentant f de $[f]$.

On montre que l'espace $L^\infty(\Omega)$ ainsi obtenu muni de la norme $\|\cdot\|_{L^\infty}$ est complet (exercice), et que $\mathcal{C}_c(\Omega)$ s'identifie à un sous-espace vectoriel de cet espace $L^\infty(\Omega)$, qui cette fois *n'est pas dense* dans $L^\infty(\Omega)$. La norme $\|\cdot\|_{L^\infty}$ restreinte à $\mathcal{C}_c(\Omega)$ 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 $\mathcal{L}^\infty(\Omega)$ des fonctions qui ne sont pas p.p. égales à des fonctions continues (exercice).

Enfin, la propriété (3) du théorème ci-dessus est fausse pour l'espace $L^\infty(\mathbf{R}^N)$.

Exemple 8.3. La fonction indicatrice $\mathbf{1}_{[0,1]}$ appartient à $L^\infty(\mathbf{R})$ mais on vérifie que

$$\|\mathbf{1}_{[0,1]}(\cdot - y) - \mathbf{1}_{[0,1]}\|_{L^\infty(\mathbf{R})} = 1,$$

pour tout $y \in]-1, 1[- \{0\}$. En particulier

$$\lim_{y \rightarrow 0} \|\mathbf{1}_{[0,1]}(\cdot - y) - \mathbf{1}_{[0,1]}\|_{L^\infty(\mathbf{R})} \neq 0.$$

L'étude plus détaillée de l'espace L^∞ pourra être poursuivie dans les petites classes.

Exercice 8.1 (Lemme de Brézis-Lieb). Soit $p \geq 1$ et $(f_n)_{n \in \mathbf{N}}$ une suite de fonctions mesurables définies sur un ouvert non vide $\Omega \subset \mathbf{R}^N$. On suppose que $|f_n|^p \in \mathcal{L}^1(\Omega)$ et qu'il existe une fonction f définie sur Ω telle que $|f|^p \in \mathcal{L}^1(\Omega)$ et $\lim_{n \rightarrow +\infty} f_n = f$ p.p. sur Ω . On suppose de plus que

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

On se propose de montrer que

$$\lim_{n \rightarrow +\infty} \left(\int_{\Omega} |f_n(x)|^p dx - \int_{\Omega} |f(x) - f_n(x)|^p dx \right) = \int_{\Omega} |f(x)|^p dx$$

- 1) Montrer que $0 \leq |f_n - f| + |f| - |f_n| \leq 2|f|$ p.p. sur Ω .
- 2) Établir le résultat lorsque $p = 1$.
- 3) On suppose maintenant que $p > 1$. Montrer que, pour tout $\varepsilon > 0$, il existe $C_\varepsilon > 0$ tel que

$$||s + t|^p - |s|^p| \leq \varepsilon |s|^p + C_\varepsilon |t|^p,$$

pour tous $s, t \in \mathbf{R}$.

- 4) On note

$$W(s, t) := (||s + t|^p - |s|^p - |t|^p| - \varepsilon |s|^p)^+,$$

et

$$F_n = W(f_n - f, f).$$

Calculer

$$\lim_{n \rightarrow +\infty} \int_{\Omega} F_n(x) dx.$$

5) Conclure.

CHAPITRE 9

La transformation de Fourier sur L^1

1. Motivation

L'analyse de Fourier¹ 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 les idées fondamentales de l'analyse de Fourier dans l'étude des fonctions périodiques à travers la *théorie des séries de Fourier*, pour des fonctions qui sont supposées assez régulières (par exemple de classe \mathcal{C}^1 par morceaux). Mais les grands principes de l'analyse de Fourier s'appliquent aussi à l'étude de fonctions plus générales, qui ne sont pas forcément continues ou périodiques.

Dans ce chapitre, nous allons tout particulièrement nous intéresser à l'action de la transformation de Fourier sur les fonctions intégrables qui sont définies sur l'espace euclidien \mathbf{R}^N , en insistant sur l'apport de la théorie de l'intégration de Lebesgue à l'analyse de Fourier.

Il existe de nombreuses motivations pour étudier la transformation de Fourier ; certaines sont internes aux mathématiques, d'autres ont pour origine certains modèles mathématiques de la physique².

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 $A dS$, 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 différentes sources secondaires.

En supposant l'onde plane incidente de phase nulle sur la pupille, l'onde émise au point M de l'écran, de coordonnées (x, y) , par l'élément de surface dS de la pupille P centré au point de coordonnées (X, Y) vaut

$$\frac{e^{i\frac{2\pi}{\lambda}\sqrt{r^2+(x-X)^2+(y-Y)^2}}}{\sqrt{r^2+(x-X)^2+(y-Y)^2}} A dS,$$

1. Professeur à l'École polytechnique en 1797.

2. 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 le cours de F. Golse, *Distributions, analyse de Fourier et équations aux dérivées partielles*, MAT431.

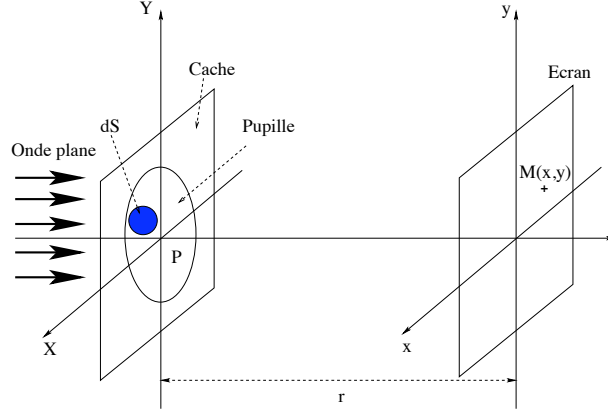


FIGURE 9.1. La diffraction de Fraunhofer.

à une constante multiplicative près. Dans cette formule, 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 sur la pupille de coordonnées (X, Y) et le point M de l'écran de coordonnées (x, y) .

La somme de ces contributions intégrée sur la surface de la pupille P vaut donc

$$A \iint_P \frac{e^{i\frac{2\pi}{\lambda} \sqrt{r^2 + (x-X)^2 + (y-Y)^2}}}{\sqrt{r^2 + (x-X)^2 + (y-Y)^2}} dX dY.$$

On note $\tilde{r} := \sqrt{r^2 + x^2 + y^2}$ et l'on fait alors l'approximation suivante

$$\frac{e^{i\frac{2\pi}{\lambda} \sqrt{r^2 + (x-X)^2 + (y-Y)^2}}}{\sqrt{r^2 + (x-X)^2 + (y-Y)^2}} \simeq \frac{1}{\tilde{r}} e^{i\frac{2\pi}{\lambda} \tilde{r} - i\frac{2\pi}{\lambda \tilde{r}} (xX + yY)},$$

sous l'hypothèse que $X^2 + Y^2 \ll \lambda \tilde{r}$, qui est valable pourvu que le diamètre de la pupille soit petit devant r . On remarque que \tilde{r} ne dépend pas de X ni de Y .

Au point M de coordonnées (x, y) sur l'écran, la somme des contributions des ondes élémentaires rayonnées par chaque élément de surface sur la pupille vaut donc, en première approximation

$$\frac{A}{\tilde{r}} e^{i\frac{2\pi}{\lambda} \tilde{r}} \iint_P e^{-i\frac{2\pi}{\lambda \tilde{r}} (xX + yY)} dX dY,$$

et 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

$$\left| \iint_{\mathbf{R}^2} \mathbf{1}_P(X, Y) e^{-i\frac{2\pi}{\lambda \tilde{r}} (xX + yY)} dX dY \right|.$$

L'intégrale

$$\iint_{\mathbf{R}^2} \mathbf{1}_P(X, Y) e^{-i\frac{2\pi}{\lambda \tilde{r}} (xX + yY)} dX dY,$$

est un exemple d'intégrale de Fourier, plus précisément, c'est la transformée de Fourier de $\mathbf{1}_P$, la fonction indicatrice de la pupille.

Voici quelques exemples de figures de diffraction.

Exemple 1 : diffraction par une pupille carrée.

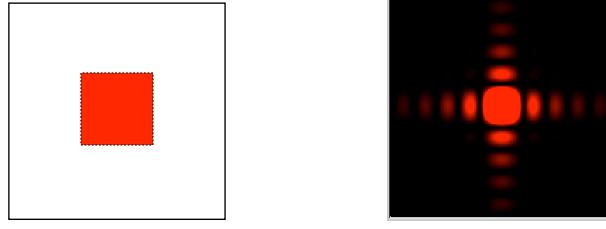


FIGURE 9.2. À gauche, pupille carrée ; à droite, l'image diffractée sur l'écran.

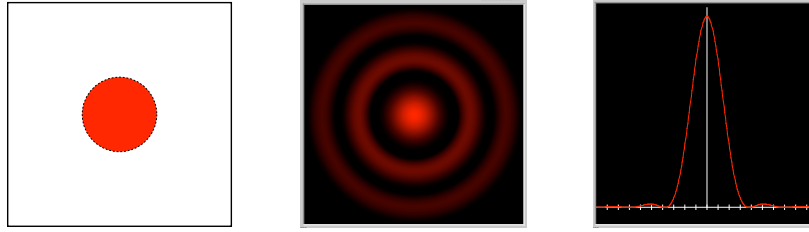
Exemple 2 : diffraction par une pupille circulaire

FIGURE 9.3. À 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 connaissance 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 [6], Malliavin-Airault [8] et Lieb-Loss [7].

2. Propriétés de la transformation de Fourier

Définition 9.1 (Transformation de Fourier sur L^1). À tout $f \in L^1(\mathbf{R}^N; \mathbf{C})$ la transformation de Fourier \mathcal{F} associe la fonction $\hat{f} = \mathcal{F}(f)$ qui est définie sur \mathbf{R}^N par la formule

$$\hat{f}(\xi) := \int_{\mathbf{R}^N} e^{-i\xi \cdot x} f(x) dx,$$

pour tout $\xi \in \mathbf{R}^N$.

Il existe dans la littérature, différentes manières de normaliser la transformation de Fourier. Dans ce cours, nous avons adoptée la normalisation qui est utilisée par les spécialistes des équations aux dérivées partielles³.

Remarquons une différence importante entre f et sa transformée $\mathcal{F}(f)$: l'élément $f \in L^1(\mathbf{R}^N; \mathbf{C})$ est une classe d'équivalence de fonctions que l'on 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 $f \in L^1(\mathbf{R}^N; \mathbf{C})$,

3. Le lecteur notera la différence avec la normalisation utilisée dans le cours PHYS311.

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 intégrable puisque

$$\int_{\mathbf{R}^N} |e^{-i\xi \cdot x} f(x)| dx = \int_{\mathbf{R}^N} |f(x)| dx = \|f\|_{L^1} < +\infty.$$

Plus précisément, nous avons le :

Théorème 9.1 (Théorème de Riemann-Lebesgue). *Soit $f \in L^1(\mathbf{R}^N; \mathbf{C})$ et $\hat{f} = \mathcal{F}(f)$ la transformée de Fourier de f . Alors la fonction \hat{f} est continue sur \mathbf{R}^N ,*

$$\lim_{|\xi| \rightarrow +\infty} \hat{f}(\xi) = 0,$$

et

$$|\hat{f}(\xi)| \leq \|f\|_{L^1},$$

pour tout $\xi \in \mathbf{R}^N$.

DÉMONSTRATION. Pour tout $\xi \in \mathbf{R}^N$, on vient de voir que

$$|\hat{f}(\xi)| = \left| \int_{\mathbf{R}^N} e^{-i\xi \cdot x} f(x) dx \right| \leq \int_{\mathbf{R}^N} |e^{-i\xi \cdot x} f(x)| dx = \int_{\mathbf{R}^N} |f(x)| dx = \|f\|_{L^1}.$$

Montrons que la fonction \hat{f} est continue sur \mathbf{R}^N . Soit $(\xi_n)_{n \geq 0}$ une suite de \mathbf{R}^N qui converge vers $\xi \in \mathbf{R}^N$. Alors, la suite de fonctions $x \mapsto e^{-i\xi_n \cdot x} f(x)$ converge simplement vers $x \mapsto e^{-i\xi \cdot x} f(x)$ p.p. sur \mathbf{R}^N et

$$|e^{-i\xi_n \cdot x} f(x)| \leq |f(x)|,$$

p.p. sur \mathbf{R}^N . Étant donné que $f \in L^1(\mathbf{R}^N; \mathbf{C})$, on en déduit que $|f| \in L^1(\mathbf{R}^N)$ et le Théorème de la convergence dominée de Lebesgue nous assure que

$$\lim_{n \rightarrow +\infty} \hat{f}(\xi_n) := \lim_{n \rightarrow +\infty} \int_{\mathbf{R}^N} e^{-i\xi_n \cdot x} f(x) dx = \int_{\mathbf{R}^N} e^{-i\xi \cdot x} f(x) dx =: \hat{f}(\xi).$$

Ce qui démontre que \hat{f} est continue sur \mathbf{R}^N .

Démontrons enfin que la fonction \hat{f} tend vers 0 à l'infini. Pour $\xi \neq 0$, on a

$$\hat{f}(\xi) = \int_{\mathbf{R}^N} e^{-i\xi \cdot x} f(x) dx = - \int_{\mathbf{R}} e^{-i\xi \cdot \left(x + \frac{\pi}{|\xi|^2} \xi\right)} f(x) dx = - \int_{\mathbf{R}^N} e^{-i\xi \cdot y} f\left(y - \frac{\pi}{|\xi|^2} \xi\right) dy,$$

grâce au changement de variables $y = x + \frac{\pi}{|\xi|^2} \xi$. Ainsi

$$\hat{f}(\xi) = \frac{1}{2} \int_{\mathbf{R}^N} e^{-i\xi \cdot x} \left(f(x) - f\left(x - \frac{\pi}{|\xi|^2} \xi\right) \right) dx,$$

de sorte que

$$|\hat{f}(\xi)| \leq \frac{1}{2} \int_{\mathbf{R}^N} \left| f(x) - f\left(x - \frac{\pi}{|\xi|^2} \xi\right) \right| dx = \frac{1}{2} \left\| f - f\left(\cdot - \frac{\pi}{|\xi|^2} \xi\right) \right\|_{L^1},$$

tend vers 0 lorsque $|\xi| \rightarrow +\infty$, d'après le Corollaire 8.1 page 163. \square

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 précisément, si u est une fonction continue sur \mathbf{R} , périodique de période 1 et si $u \in \mathcal{C}^m(\mathbf{R}; \mathbf{C})$ alors les coefficients de Fourier de u

$$\hat{u}(k) := \int_0^1 e^{i2\pi kx} u(x) dx,$$

vérifient $|\widehat{u}(k)| \leq C(1 + |k|)^{-m}$, où $C > 0$ ne dépend pas de k . Il en va de même pour la transformation de Fourier, comme le montre l'énoncé suivant.

Théorème 9.2 (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 $\widehat{f} \in \mathcal{C}^1(\mathbf{R}^N; \mathbf{C})$ et l'on a*

$$\frac{\partial \widehat{f}}{\partial \xi_\ell}(\xi) = -i \int_{\mathbf{R}^N} x_\ell e^{-i\xi \cdot x} f(x) dx = -i \widehat{x_\ell f}(\xi),$$

pour tout $\ell = 1, \dots, N$;

(b) *Si $f \in \mathcal{C}^1(\mathbf{R}^N; \mathbf{C})$ et si $\frac{\partial f}{\partial x_k} \in L^1(\mathbf{R}^N; \mathbf{C})$, pour tout $k = 1, \dots, N$, alors*

$$\frac{\partial \widehat{f}}{\partial x_k}(\xi) = \int_{\mathbf{R}^N} e^{-i\xi \cdot x} \frac{\partial f}{\partial x_k}(x) dx = i \xi_k \widehat{f}(\xi),$$

pour tout $k = 1, \dots, N$.

Cet énoncé est l'analogie, dans le cadre de la transformation de Fourier, de ce que nous avons rappelé à propos des coefficients de Fourier des fonctions périodiques.

Plus généralement, on dispose des deux propriétés suivantes dont la démonstration est laissée en exercice :

- (a) plus une fonction décroît vite à l'infini, plus sa transformée de Fourier est régulière. Plus précisément, pour tout $k \geq 1$, si $f \in \mathcal{L}^1(\mathbf{R}^N; \mathbf{C})$ et si $|x|^k f \in \mathcal{L}^1(\mathbf{R}^N; \mathbf{C})$ alors $\widehat{f} \in \mathcal{C}^k(\mathbf{R}^N; \mathbf{C})$. (Appliquer successivement le théorème ci-dessus à \widehat{f} et à ses dérivées partielles).
- (b) plus une fonction est régulière, plus sa transformée de Fourier décroît vite à l'infini. Plus précisément, pour tout $k \geq 1$, si $f \in \mathcal{C}^k(\mathbf{R}^N; \mathbf{C})$ et si les dérivées partielles de f jusqu'à l'ordre k appartiennent à $L^1(\mathbf{R}^N; \mathbf{C})$ alors $\lim_{|\xi| \rightarrow +\infty} |\xi|^k |\widehat{f}(\xi)| = 0$. (Appliquer successivement le théorème ci-dessus à f et à ses dérivées partielles).

Ces remarques montrent que la transformation de Fourier \mathcal{F} échange régularité et décroissance à l'infini.

Passons à la démonstration du Théorème 9.2. Elle repose sur le Théorème de dérivation sous le signe somme, que nous avons présenté dans un chapitre précédent.

DÉMONSTRATION DU THÉORÈME 9.2. Soit $\mathcal{Z} \subset \Omega$ un sous-ensemble négligeable tel que $|f(x)| < +\infty$ pour tout $x \in \mathbf{R}^N - \mathcal{Z}$. Définissons la fonction $\phi : \mathbf{R}^N \times \mathbf{R}^N \rightarrow \mathbf{C}$ par la formule

$$\phi(\xi, x) := e^{-i\xi \cdot x} f(x).$$

On vérifie que $\phi(\cdot, x) \in \mathcal{C}^1(\mathbf{R}^N; \mathbf{C})$ pour tout $x \in \mathbf{R}^N - \mathcal{Z}$ et que $\phi(\xi, \cdot) \in L^1(\mathbf{R}^N; \mathbf{C})$ pour tout $\xi \in \mathbf{R}^N$. D'autre part,

$$\left| \frac{\partial \phi}{\partial \xi_k}(\xi, x) \right| = |x_k| |f(x)| \leq |x| |f(x)|,$$

pour tout $x \in \mathbf{R}^N - \mathcal{Z}$ et pour tout $\xi \in \mathbf{R}^N$. Le Théorème de dérivation sous le signe somme (voir le Théorème 5.10) nous permet de conclure que $\widehat{f} \in \mathcal{C}^1(\mathbf{R}^N; \mathbf{C})$ et que

$$\frac{\partial \widehat{f}}{\partial \xi_k}(\xi) = -i \int_{\mathbf{R}^N} x_k e^{-i\xi \cdot x} f(x) dx,$$

pour tout $x \in \mathbf{R}^N$, ce qui établit le point (a).

Pour simplifier les notations, nous allons traiter le seul cas de la dérivation par rapport à la première variable dans la démonstration du point (b). Comme f est de classe \mathcal{C}^1 sur \mathbf{R}^N , une intégration par parties montre que

$$\int_a^b e^{-i\xi \cdot x} \frac{\partial f}{\partial x_1}(x) dx_1 = \left[e^{-i\xi \cdot x} f(x) \right]_{x_1=a}^{x_1=b} + i\xi_1 \int_a^b e^{-i\xi \cdot x} f(x) dx_1.$$

Notons $x' = (x_2, \dots, x_N)$; comme f et $\frac{\partial f}{\partial x_1}$ sont intégrables sur \mathbf{R}^N , on peut appliquer le Théorème de Fubini pour obtenir les égalités

$$\begin{aligned} \int_{\mathbf{R}^{N-1}} \left(\int_a^b e^{-i\xi \cdot x} \frac{\partial f}{\partial x_1}(x) dx_1 \right) dx' &= \int_{\mathbf{R}^N} e^{-i\xi \cdot x} \frac{\partial f}{\partial x_1}(x) \mathbf{1}_{[a,b]}(x_1) dx, \\ \int_{\mathbf{R}^{N-1}} \left(\int_a^b e^{-i\xi \cdot x} f(x) dx_1 \right) dx' &= \int_{\mathbf{R}^N} e^{-i\xi \cdot x} f(x) \mathbf{1}_{[a,b]}(x_1) dx. \end{aligned}$$

Étant donné que la fonction $\mathbf{1}_{[a,b]}$ converge simplement vers 1 lorsque $a \rightarrow -\infty$ et $b \rightarrow +\infty$, et que f et $\frac{\partial f}{\partial x_1}$ sont intégrables sur \mathbf{R}^N , on déduit du Théorème de la convergence dominée de Lebesgue et des deux égalités ci-dessus que

$$\begin{aligned} \lim_{a \rightarrow -\infty} \lim_{b \rightarrow +\infty} \int_{\mathbf{R}^{N-1}} \left(\int_a^b e^{-i\xi \cdot x} \frac{\partial f}{\partial x_1}(x) dx_1 \right) dx' &= \int_{\mathbf{R}^N} e^{-i\xi \cdot x} \frac{\partial f}{\partial x_1}(x) dx, \\ \lim_{a \rightarrow -\infty} \lim_{b \rightarrow +\infty} \int_{\mathbf{R}^{N-1}} \left(\int_a^b e^{-i\xi \cdot x} f(x) dx_1 \right) dx' &= \int_{\mathbf{R}^N} e^{-i\xi \cdot x} f(x) dx. \end{aligned}$$

D'autre part, comme f est intégrable sur \mathbf{R}^N , le Théorème de Fubini nous assure que la fonction

$$x_1 \mapsto \int_{\mathbf{R}^{N-1}} e^{-i\xi \cdot x} f(x_1, x') dx',$$

est intégrable sur \mathbf{R} . Il existe donc deux suites $(a_n)_{n \geq 0}$ et $(b_n)_{n \geq 0}$ convergeant respectivement vers $-\infty$ et $+\infty$, telles que

$$\lim_{n \rightarrow +\infty} \int_{\mathbf{R}^{N-1}} e^{-i\xi \cdot x} f(a_n, x') dx' = 0, \quad \text{et} \quad \lim_{n \rightarrow +\infty} \int_{\mathbf{R}^{N-1}} e^{-i\xi \cdot x} f(b_n, x') dx' = 0.$$

En injectant ces deux suites $(a_n)_{n \geq 0}$ et $(b_n)_{n \geq 0}$ dans l'identité

$$\int_{a_n}^{b_n} e^{-i\xi \cdot x} \frac{\partial f}{\partial x_1}(x) dx_1 = \left[e^{-i\xi \cdot x} f(x) \right]_{x_1=a_n}^{x_1=b_n} + i\xi_1 \int_{a_n}^{b_n} e^{-i\xi \cdot x} f(x) dx_1,$$

que l'on intègre sur \mathbf{R}^{N-1} par rapport à x' , et en passant à la limite pour $n \rightarrow +\infty$, on aboutit à l'égalité

$$\int_{\mathbf{R}^N} e^{-i\xi \cdot x} \frac{\partial f}{\partial x_1}(x) dx = i\xi_1 \int_{\mathbf{R}^N} e^{-i\xi \cdot x} f(x) dx,$$

ce qui termine la démonstration du point (b). \square

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 9.2 (Densité gaussienne). *Pour tout $a > 0$, on appelle densité gaussienne sur \mathbf{R}^N de matrice de covariance $a I_N$ la fonction*

$$G_a(x) := \frac{1}{(2\pi a)^{N/2}} e^{-|x|^2/2a},$$

définie pour tout $x \in \mathbf{R}^N$.

Dans cet énoncé, I_N désigne la matrice identité de dimension $N \times N$.

En particulier, pour tout $a > 0$, la fonction G_a est continue sur \mathbf{R}^N , $G_a(x) > 0$ pour tout $x \in \mathbf{R}^N$ et (voir l'exercice 6.1 page 122)

$$\int_{\mathbf{R}^N} G_a(x) dx = 1.$$

On dit aussi que G_a est une *densité de probabilité* sur \mathbf{R}^N .

Proposition 9.1 (Transformée de Fourier des gaussiennes). *Pour tout $a > 0$, on a*

$$\widehat{G}_a(\xi) = e^{-a|\xi|^2/2},$$

pour tout $\xi \in \mathbf{R}^N$. 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 de la dimension $N = 1$.

Pour tout $a > 0$, la fonction $x \mapsto G_a(x)$ appartient à $\mathcal{C}^\infty(\mathbf{R})$ et vérifie

$$a G'_a + x G_a = 0,$$

pour tout $x \in \mathbf{R}$. Comme les fonctions G_a , $x G_a$ et G'_a appartiennent à $\mathcal{L}^1(\mathbf{R})$, on déduit du Théorème 9.2 page 173 que la transformée de Fourier $\widehat{G}_a \in \mathcal{C}^1(\mathbf{R}; \mathbf{C})$, et que

$$\widehat{G}'_a = i \xi \widehat{G}_a, \quad \text{et} \quad \widehat{G}_a' = -i x \widehat{G}_a.$$

On déduit de ces deux identités, et de l'égalité ci-dessus vérifiée par G_a et G'_a , que

$$a \xi \widehat{G}_a + \widehat{G}_a' = 0.$$

Cette égalité est une équation différentielle satisfaite par $\xi \mapsto \widehat{G}_a(\xi)$, équation différentielle dont la solution générale est de la forme

$$\widehat{G}_a(\xi) = C e^{-a\xi^2/2},$$

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

$$\begin{aligned} \int_{\mathbf{R}^N} e^{-i\xi \cdot x} \frac{1}{(2\pi a)^{N/2}} e^{-|x|^2/(2a)} dx &= \int_{\mathbf{R}^N} \prod_{k=1}^N e^{-i\xi_k x_k} \frac{1}{\sqrt{2\pi a}} e^{-x_k^2/(2a)} dx \\ &= \prod_{k=1}^N \int_{\mathbf{R}} e^{-i\xi_k x_k} \frac{1}{\sqrt{2\pi a}} e^{-x_k^2/(2a)} dx_k \\ &= \prod_{k=1}^N e^{-a\xi_k^2/2} = e^{-a|\xi|^2/2}. \end{aligned}$$

□

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 9.3 (Théorème d'inversion de Fourier dans L^1). *Soit $f \in L^1(\mathbf{R}^N; \mathbf{C})$ telle que $\widehat{f} \in L^1(\mathbf{R}^N; \mathbf{C})$. Alors, pour presque tout $x \in \mathbf{R}^N$*

$$f(x) = \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{+ix \cdot \xi} \widehat{f}(\xi) d\xi.$$

Une remarque sur l'énoncé de la formule d'inversion. On observera que le membre de droite, à savoir la fonction

$$x \mapsto \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{+ix \cdot \xi} \widehat{f}(\xi) d\xi,$$

est définie et continue sur tout \mathbf{R}^N et même tend vers 0 lorsque $|x| \rightarrow +\infty$ d'après le Théorème de Riemann-Lebesgue, puisque cette fonction est, à une constante multiplicative près, la transformée de Fourier de $\xi \mapsto \widehat{f}(-\xi)$ et que $\widehat{f} \in L^1(\mathbf{R}^N; \mathbf{C})$ par hypothèse. Mais l'égalité qu'exprime le Théorème d'inversion n'a de sens que pour presque tout $x \in \mathbf{R}^N$, puisque f est un élément de $L^1(\mathbf{R}^N)$, c'est-à-dire que f est identifiée à une fonction intégrable qui est définie seulement p.p. sur \mathbf{R}^N .

Ainsi, le théorème d'inversion ci-dessus ne vaut que pour des fonctions f intégrables, qui sont 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

$$f(x) = \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{+ix \cdot \xi} \left(\int_{\mathbf{R}^N} e^{-i\xi \cdot y} f(y) dy \right) d\xi.$$

Il semble naturel d'essayer d'échanger l'ordre des intégrations par rapport aux variables y et ξ

$$f(x) = \int_{\mathbf{R}^N} f(y) \left(\frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{+i(x-y) \cdot \xi} d\xi \right) dy,$$

ce qui conduit à la formule

$$f(x) = \int_{\mathbf{R}^N} K(x-y) f(y) dy,$$

avec

$$K(z) := \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{-iz \cdot \xi} d\xi.$$

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 l'intégration de Lebesgue, car la fonction $\xi \mapsto e^{-iz \cdot \xi}$ n'est intégrable sur \mathbf{R}^N pour aucune valeur de $z \in \mathbf{R}^N$, puisque

$$\int_{\mathbf{R}^N} |e^{-iz \cdot \xi}| d\xi = \int_{\mathbf{R}^N} d\xi = +\infty,$$

pour tout $z \in \mathbf{R}^N$.

D'autre part l'interversion des intégrations par rapport aux variables ξ et y ci-dessus ne peut pas être justifiée grâce au Théorème de Fubini puisque la fonction

$$(y, \xi) \mapsto e^{i\xi \cdot (x-y)} f(y),$$

n'est pas intégrable sur $\mathbf{R}^N \times \mathbf{R}^N$. En effet,

$$\iint_{\mathbf{R}^N \times \mathbf{R}^N} |e^{i\xi \cdot (x-y)} f(y)| d\xi dy = \iint_{\mathbf{R}^N \times \mathbf{R}^N} |f(y)| d\xi dy = \int_{\mathbf{R}^N} d\xi \int_{\mathbf{R}^N} |f(y)| dy = +\infty,$$

sauf si $f(y) = 0$ p.p. sur \mathbf{R}^N .

Ces deux obstacles disparaissent en même temps si l'on rend l'intégrande dans la définition de K intégrable en le multipliant par un poids gaussien, c'est-à-dire si l'on remplace K par la fonction définie par

$$K_\varepsilon(z) := \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{-iz \cdot \xi} e^{-\frac{1}{2}\varepsilon^2 |\xi|^2} d\xi,$$

pour tout $\varepsilon > 0$ et tout $z \in \mathbf{R}^N$. D'une part, on remarque que, grâce à la Proposition 9.1 page 175, on sait calculer explicitement K_ε pour tout $\varepsilon > 0$ et l'on trouve

$$K_\varepsilon = \frac{1}{(2\pi\varepsilon^2)^{N/2}} \widehat{G_{1/\varepsilon^2}} = G_{\varepsilon^2}.$$

D'autre part, pour tout $\varepsilon > 0$, l'interversion d'intégrales ci-dessus devient licite, et l'on a

$$\begin{aligned} \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{+ix \cdot \xi - \frac{1}{2}\varepsilon^2 |\xi|^2} \widehat{f}(\xi) d\xi &= \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{+ix \cdot \xi - \frac{1}{2}\varepsilon^2 |\xi|^2} \left(\int_{\mathbf{R}^N} e^{-i\xi \cdot y} f(y) dy \right) d\xi \\ &= \int_{\mathbf{R}^N} f(y) \left(\frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{i(x-y) \cdot \xi} e^{-\frac{1}{2}\varepsilon^2 |\xi|^2} d\xi \right) dy \\ &= \int_{\mathbf{R}^N} f(y) K_\varepsilon(x-y) dy. \end{aligned}$$

La démonstration du théorème va consister à montrer que le membre de gauche converge vers

$$\frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{+ix \cdot \xi} \widehat{f}(\xi) d\xi,$$

pour tout $x \in \mathbf{R}^N$, et que, à extraction d'une sous-suite près, le membre de droite converge pour presque tout $x \in \mathbf{R}^N$ vers $f(x)$ lorsque $\varepsilon \rightarrow 0$.

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 correcte⁴.

4. En fait, ce schéma de démonstration, et notamment l'expression

$$K(z) = \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{i\xi \cdot z} d\xi,$$

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 *Distributions, analyse de Fourier, EDP*. Dans cette théorie, on montre que $K = \delta_0$, à savoir la masse de Dirac en 0, introduite dans le Chapitre 7,

DÉMONSTRATION. Étape 1 : D'après la Proposition 9.1 page 175, pour tout $\varepsilon > 0$,

$$\begin{aligned} K_\varepsilon(z) &= \frac{1}{(2\pi\varepsilon^2)^{N/2}} \widehat{G}_{1/\varepsilon^2}(z) \\ &= \frac{1}{(2\pi\varepsilon^2)^{N/2}} \int_{\mathbf{R}^N} e^{-i\xi \cdot z} \frac{e^{-\frac{1}{2}\varepsilon^2|\xi|^2}}{(2\pi/\varepsilon^2)^{N/2}} d\xi \\ &= \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{-i\xi \cdot z - \frac{1}{2}\varepsilon^2|\xi|^2} d\xi \\ &= \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{i\xi \cdot z - \frac{1}{2}\varepsilon^2|\xi|^2} d\xi, \end{aligned}$$

pour tout $z \in \mathbf{R}^N$. On en déduit l'égalité

$$\begin{aligned} \int_{\mathbf{R}^N} f(y) K_\varepsilon(x-y) dy &= \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} f(y) \left(\int_{\mathbf{R}^N} e^{i\xi \cdot (x-y) - \frac{1}{2}\varepsilon^2|\xi|^2} d\xi \right) dy \\ &= \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{i\xi \cdot x - \frac{1}{2}\varepsilon^2|\xi|^2} \left(\int_{\mathbf{R}^N} e^{-i\xi \cdot y} f(y) dy \right) d\xi \\ &= \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{i\xi \cdot x - \frac{1}{2}\varepsilon^2|\xi|^2} \widehat{f}(\xi) d\xi. \end{aligned}$$

En effet, l'interversion des intégrales en y et ξ (la seconde égalité ci-dessus) est justifiée grâce au Théorème de Fubini, car la fonction

$$(y, \xi) \mapsto f(y) e^{i\xi \cdot (x-y) - \frac{1}{2}\varepsilon^2|\xi|^2},$$

appartient à $L^1(\mathbf{R}^N \times \mathbf{R}^N; \mathbf{C})$ puisque

$$\iint_{\mathbf{R}^N \times \mathbf{R}^N} \left| f(y) e^{i\xi \cdot (x-y) - \frac{1}{2}\varepsilon^2|\xi|^2} \right| dy d\xi = \int_{\mathbf{R}^N} |f(y)| dy \int_{\mathbf{R}^N} e^{-\frac{1}{2}\varepsilon^2|\xi|^2} d\xi < +\infty.$$

Étape 2 : Passons à la limite en $\varepsilon \rightarrow 0^+$ dans le membre de droite. Pour tout $x \in \mathbf{R}^N$ et pour presque tout $\xi \in \mathbf{R}^N$,

$$\lim_{\varepsilon \rightarrow 0} e^{i\xi \cdot x - \frac{1}{2}\varepsilon^2|\xi|^2} \widehat{f}(\xi) = e^{i\xi \cdot x} \widehat{f}(\xi),$$

et

$$\left| e^{i\xi \cdot x - \frac{1}{2}\varepsilon^2|\xi|^2} \widehat{f}(\xi) \right| \leq |\widehat{f}(\xi)|.$$

Grâce au Théorème de la convergence dominée de Lebesgue (puisque $\widehat{f} \in L^1(\mathbf{R}^N; \mathbf{C})$), on conclut que

$$\lim_{\varepsilon \rightarrow 0} \int_{\mathbf{R}} e^{i\xi \cdot x - \frac{1}{2}\varepsilon^2|\xi|^2} \widehat{f}(\xi) \frac{d\xi}{(2\pi)^N} = \frac{1}{(2\pi)^N} \int_{\mathbf{R}} e^{i\xi \cdot x} \widehat{f}(\xi) d\xi,$$

pour tout $x \in \mathbf{R}^N$.

Étape 3 : Étudions le membre de gauche

$$\int_{\mathbf{R}^N} f(y) G_{\varepsilon^2}(x-y) dy = \frac{1}{(2\pi\varepsilon^2)^{N/2}} \int_{\mathbf{R}^N} f(x-\varepsilon z) e^{-|z|^2/2} dz,$$

et la formule d'inversion de Fourier se réduit alors à l'identité

$$f(x) = \int_{\mathbf{R}^N} f(y) K(x-y) dy = \int_{\mathbf{R}^N} f(x-y) K(y) dy = \langle \delta_0, f(x-\cdot) \rangle,$$

qui découle de la définition même de la masse de Dirac.

grâce au changement de variables $y = x - \varepsilon z$. Comme

$$\frac{1}{(2\pi)^{N/2}} \int_{\mathbf{R}^N} e^{-|z|^2/2} dz = 1,$$

on voit que

$$\begin{aligned} \int_{\mathbf{R}^N} \left| \int_{\mathbf{R}^N} f(y) K_\varepsilon(x-y) dy - f(x) \right| dx \\ = \frac{1}{(2\pi)^{N/2}} \int_{\mathbf{R}^N} \left| \int_{\mathbf{R}^N} (f(x-\varepsilon z) - f(x)) e^{-|z|^2/2} dz \right| dx \\ \leq \frac{1}{(2\pi)^{N/2}} \int_{\mathbf{R}^N} e^{-|z|^2/2} \left(\int_{\mathbf{R}^N} |f(x-\varepsilon z) - f(x)| dx \right) dz. \end{aligned}$$

D'après le Corollaire 8.1 page 163, on a

$$\lim_{\varepsilon \rightarrow 0} \int_{\mathbf{R}^N} |f(x - \varepsilon_n z) - f(x)| dx = 0,$$

pour tout $z \in \mathbf{R}$ et

$$e^{-|z|^2/2} \int_{\mathbf{R}^N} |f(x - \varepsilon_n z) - f(x)| dx \leq 2\|f\|_{L^1(\mathbf{R}^N)} e^{-|z|^2/2},$$

pour tout $\varepsilon > 0$. De sorte qu'en appliquant le Théorème de la convergence dominée de Lebesgue, on conclut que

$$\lim_{\varepsilon \rightarrow 0} \int_{\mathbf{R}^N} \left| \int_{\mathbf{R}^N} f(y) K_\varepsilon(x-y) dy - f(x) \right| dx = 0.$$

Étape 4 : Pour tout $x \in \mathbf{R}^N$ et pour tout $\varepsilon > 0$, on a l'égalité

$$\int_{\mathbf{R}^N} f(y) G_{\varepsilon^2}(x-y) dy = \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{i\xi \cdot x - \frac{1}{2}\varepsilon^2|\xi|^2} \widehat{f}(\xi) d\xi.$$

D'après l'étape 2, pour tout $x \in \mathbf{R}^N$, on a

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{i\xi \cdot x - \frac{1}{2}\varepsilon^2|\xi|^2} \widehat{f}(\xi) d\xi = \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{i\xi \cdot x} \widehat{f}(\xi) d\xi.$$

D'après l'étape 3

$$x \mapsto \int_{\mathbf{R}^N} f(y) K_\varepsilon(x-y) dy,$$

converge vers f dans $L^1(\mathbf{R}^N; \mathbf{C})$ quand ε tend vers 0, donc, quitte à extraire une sous-suite $\varepsilon_k \rightarrow 0$, converge vers f p.p. sur \mathbf{R}^N d'après le Théorème 8.2 page 159.

Conclusion, nous avons l'égalité

$$f(x) = \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{i\xi \cdot x} \widehat{f}(\xi) d\xi,$$

pour presque tout $x \in \mathbf{R}^N$. □

4. 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 . La fonction g n'a aucune raison d'être dans $L^1(\mathbf{R}; \mathbf{C})$, toutefois, on vérifie que c'est une fonction de $L^1(S^1; \mathbf{C})$, grâce au Théorème de la convergence monotone qui nous permet d'intervertir somme et intégrale dans le calcul suivant

$$\begin{aligned} \int_0^{2\pi} \left| \sum_{n \in \mathbf{Z}} f(t + 2\pi n) \right| dt &\leq \int_0^{2\pi} \sum_{n \in \mathbf{Z}} |f(t + 2\pi n)| dt \\ &= \sum_{n \in \mathbf{Z}} \int_0^{2\pi} |f(t + 2\pi n)| dt \\ &= \sum_{n \in \mathbf{Z}} \int_{2\pi n}^{2(n+1)\pi} |f(t)| dt \\ &= \int_{\mathbf{R}} |f(t)| dt. \end{aligned}$$

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. En utilisant le Théorème de la convergence dominée, on a

$$\begin{aligned} c_k(g) &= \frac{1}{2\pi} \int_0^{2\pi} \sum_{n \in \mathbf{Z}} f(t + 2\pi n) e^{-ikt} dt \\ &= \frac{1}{2\pi} \sum_{n \in \mathbf{Z}} \int_0^{2\pi} f(t + 2\pi n) e^{-ikt} dt \\ &= \frac{1}{2\pi} \sum_{n \in \mathbf{Z}} \int_{2\pi n}^{2\pi(n+1)} f(t) e^{-ikt} e^{i2\pi kn} dt \\ &= \frac{1}{2\pi} \int_{\mathbf{R}} f(t) e^{-ikt} dt \\ &= \frac{1}{2\pi} \hat{f}(k). \end{aligned}$$

Étant donné que la formule $g(t) = \sum_{k \in \mathbf{Z}} c_k(g) e^{ikt}$ est valable par exemple si g est dans $L^2(S^1; \mathbf{C})$, la convergence de la série étant entendue dans ce même espace (et non pas ponctuelle), on obtient, au moins formellement, l'égalité

$$(9.1) \quad \sum_{n \in \mathbf{Z}} f(t + 2\pi n) = \frac{1}{2\pi} \sum_{k \in \mathbf{Z}} \hat{f}(k) e^{ikt}.$$

Dans le résultat ci-dessous, nous précisons les hypothèses qui permettent d'affirmer que la formule ci-dessus a un sens.

Proposition 9.2. *Soit $f : \mathbf{R} \mapsto \mathbf{C}$ une fonction pour laquelle il existe une constante $C > 0$ telle que*

$$|f(t)| \leq \frac{C}{1+t^2},$$

pour tout $t \in \mathbf{R}$. Alors, on a

$$\sum_{n \in \mathbf{Z}} f(t + 2\pi n) = \frac{1}{2\pi} \sum_{k \in \mathbf{Z}} \hat{f}(k) e^{ikt},$$

l'égalité et la convergence des séries étant entendues dans $L^2(S^1; \mathbf{C})$.

DÉMONSTRATION. Il suffit de vérifier que $g \in L^2(S^1; \mathbf{C})$. On remarque que

$$\begin{aligned} \int_0^{2\pi} |g(t)|^2 dt &= \int_0^{2\pi} \sum_{m, n \in \mathbf{Z}} f(t + 2\pi m) \overline{f(t + 2\pi n)} dt \\ &\leq \sum_{n, m \in \mathbf{Z}} \int_{2\pi m}^{2\pi(m+1)} |f(t)| |f(t + 2\pi(m - n))| dt \\ &= \sum_{q \in \mathbf{Z}} \int_{\mathbf{R}} |f(t)| |f(t + 2\pi q)| dt. \end{aligned}$$

Il suffit donc de démontrer que la série

$$\sum_{q \in \mathbf{Z}} \int_{\mathbf{R}} |f(t)| |f(t + 2\pi q)| dt,$$

converge. Or, par hypothèse, la fonction f tend vers 0 à l'infini, comme $t \mapsto t^2$. Nous pouvons donc écrire

$$\int_{|t| \geq t_0} |f(t)|^2 dt \leq C^2 \int_{|t| \geq t_0} \frac{dt}{(1 + |t|^2)^2} \leq \frac{C'}{t_0^3},$$

pour tout $t_0 \geq 1$. De plus, pour tout $q \geq 0$ (le cas où $q \leq 0$ se traite de la même manière), on peut écrire

$$\begin{aligned} \int_{\mathbf{R}} |f(t)| |f(t + 2\pi q)| dt &= \int_{t \leq -\pi q} |f(t)| |f(t + 2\pi q)| dt + \int_{t \geq -\pi q} |f(t)| |f(t + 2\pi q)| dt \\ &= \int_{t \leq -\pi q} |f(t)| |f(t + 2\pi q)| dt + \int_{t \geq \pi q} |f(t - 2\pi q)| |f(t)| dt \\ &\leq 2 \left(\int_{\mathbf{R}} |f(t)|^2 dt \right)^{1/2} \left(\int_{|t| \geq \pi q} |f(t)|^2 dt \right)^{1/2}. \end{aligned}$$

Utilisons l'inégalité ci-dessus avec $t_0 = \pi q$ pour conclure que

$$\int_{\mathbf{R}} |f(t)| |f(t + 2\pi q)| dt \leq \frac{C''}{q^{3/2}},$$

pour une constante $C'' > 0$ (qui dépend de f). En particulier, la série de terme général

$$\sum_{q \in \mathbf{Z}} \int_{\mathbf{R}} |f(t)| |f(t + 2\pi q)| dt,$$

converge. On conclut que g est dans $L^2(S^1; \mathbf{C})$. Elle est donc égale dans $L^2(S^1; \mathbf{C})$ à la somme de sa série de Fourier, ce qui démontre la proposition. \square

Nous voudrions maintenant pouvoir utiliser cette égalité pour $t = 0$. 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 \mathcal{C}^2 et si g est dans L^2 .

Corollaire 9.1 (Formule de Poisson⁵). *Si f est de classe \mathcal{C}^2 et s'il existe une constante $C > 0$ telle que $\|f(t)\| \leq \frac{C}{1+t^2}$ pour tout $t \in \mathbf{R}$, alors*

$$\sum_{n \in \mathbf{Z}} f(2\pi n) = \frac{1}{2\pi} \sum_{k \in \mathbf{Z}} \hat{f}(k).$$

DÉMONSTRATION. Comme f est de classe \mathcal{C}^2 , $\xi^2 \hat{f}(\xi)$ tend vers 0 d'après le Théorème de Riemann-Lebesgue (Théorème 9.1 page 172). 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 qui est donc continue.

La proposition précédente affirme alors que l'on a égalité dans $L^2(S^1; \mathbf{C})$ de ces deux fonctions continues. Il en résulte que les fonctions sont égales en tout point. \square

Pour clore ce chapitre, expliquons brièvement comment le résultat précédent est utilisé dans les problèmes d'échantillonnage des signaux continus (i.e. problème de la transformation des signaux analogiques en signaux numériques). Le Théorème d'échantillonnage de Shannon-Nyquist affirme que « la fréquence d'échantillonnage d'un signal continu doit être égale ou supérieure au double de la fréquence maximale contenue dans ce signal afin de pouvoir reconstituer ce signal à partir des données discrètes ».

Soit f une fonction de classe \mathcal{C}^2 et dont le support est contenu dans $] -\pi, \pi[$. Elle vérifie bien entendu les hypothèses permettant d'écrire la formule de Poisson, et, pour $t \in [-\pi, \pi]$, on a

$$g(t) = \frac{1}{2\pi} \sum_{k \in \mathbf{Z}} \hat{f}(k) e^{ikt}.$$

Mais, la fonction f étant à support dans $] -\pi, \pi[$, on a aussi $f(t) = g(t)$ pour tout $t \in [-\pi, \pi]$. Autrement dit,

$$f(t) = \frac{1}{2\pi} \sum_{k \in \mathbf{Z}} \hat{f}(k) e^{ikt},$$

pour tout $t \in [-\pi, \pi]$. Conclusion, il suffit de connaître la valeur de la fonction \hat{f} aux points $k \in \mathbf{Z}$ pour reconstituer la fonction f . Ceci n'a évidemment rien de surprenant, car on peut toujours étendre f en une fonction 2π -périodique et la formule ci-dessus n'est rien d'autre que la formule d'inversion de Fourier!

Inversement, en remplaçant f par \hat{f} dans (9.1) 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 une fonction à support dans $] -\pi, \pi[$, on conclut que

$$\hat{f}(t) = \sum_{n \in \mathbf{Z}} \hat{f}(t + 2\pi n) = \sum_{k \in \mathbf{Z}} f(k) e^{-ikt},$$

pour tout $t \in] -\pi, \pi[$ et la connaissance de la valeur de f aux points $k \in \mathbf{Z}$ permet cette fois-ci de reconstituer la fonction \hat{f} , ce qui permet ensuite de reconstituer la fonction f . En

5. Siméon Denis Poisson X 1798, Professeur à l'École polytechnique en 1806.

d'autres termes si l'on a un signal $t \mapsto f(t)$ dont on sait que les fréquences sont contenues dans l'intervalle $]-\pi, \pi[$, l'échantillonnage $f(k)$, pour $k \in \mathbf{Z}$, permet de reconstituer le complètement ce signal. Un simple changement de variable montre que, pour tout $\alpha > 0$ fixé, si l'on a un signal $t \mapsto f(t)$ dont on sait que les fréquences sont contenues dans l'intervalle $]-\pi/\alpha, \pi/\alpha[$, l'échantillonnage $f(\alpha k)$, pour $k \in \mathbf{Z}$, permet de reconstituer le complètement ce signal.

C'est ce résultat qui est utilisé pour déterminer la fréquence d'échantillonnage des signaux sonores (de l'ordre de 44 kHz) qui est utilisée pour les enregistrements numériques, l'oreille humaine n'étant sensible qu'aux signaux dont la fréquence est inférieure à 22 kHz.

5. Bilan et remarques finales

On note $\mathcal{C}_0(\mathbf{R}^N; \mathbf{C})$ l'espace vectoriel normé des fonctions continues sur \mathbf{R}^N qui tendent vers 0 à l'infini, muni de la norme de la convergence uniforme

$$\|\phi\|_\infty = \sup_{x \in \mathbf{R}^N} |\phi(x)|.$$

Commençons par un énoncé qui synthétise les propriétés de la transformation de Fourier que nous venons d'établir.

Théorème 9.4. *La transformation de Fourier*

$$\mathcal{F} : L^1(\mathbf{R}^N; \mathbf{C}) \rightarrow \mathcal{C}_0(\mathbf{R}^N; \mathbf{C}),$$

est une application linéaire continue injective. De plus $\|\mathcal{F}\|_{\mathcal{L}(L^1, L^\infty)} \leq 1$, c'est-à-dire que

$$\|\mathcal{F}(f)\|_{L^\infty(\mathbf{R}^N)} \leq \|f\|_{L^1(\mathbf{R}^N)}.$$

Le fait que \mathcal{F} est une application linéaire continue qui vérifie $\|\mathcal{F}\|_{\mathcal{L}(L^1, L^\infty)} \leq 1$ est une conséquence immédiate du Théorème de Riemann-Lebesgue. Le fait que \mathcal{F} est injective découle du Théorème d'inversion de Fourier. 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 $\text{Ker } \mathcal{F} = \{0\}$, 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 $\mathcal{C}_0(\mathbf{R}^N; \mathbf{C})$. On établit ce point par un argument de nature topologique sur les espaces de Banach.

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 l'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 intégrable en $\xi \in \mathbf{R}$, car $\mathbf{1}_P$ n'est jamais p.p. égale à une fonction continue tendant vers 0 à l'infini sauf si elle est p.p. nulle, ce qui correspond au cas trivial où P est négligeable (c'est-à-dire en pratique, quand il n'y a pas d'ouverture).

Le Théorème d'inversion de Fourier dans L^1 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 se prolonge en une bijection linéaire isométrique de l'espace $L^2(\mathbf{R}^N; \mathbf{C})$ des fonctions de carré intégrable 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 $\widehat{f} \in L^1(\mathbf{R}; \mathbf{C})$, on définit \widehat{f} la transformation de Fourier de f par

$$\widehat{f}(\xi) = \int_{\mathbf{R}} e^{-i\xi x} f(x) dx,$$

et l'on reconstruit f à partir de \widehat{f} par la formule d'inversion

$$f(x) = \frac{1}{2\pi} \int_{\mathbf{R}} e^{ix\xi} \widehat{f}(\xi) d\xi.$$

- (b) Pour toute fonction $F \in \mathcal{C}^1(\mathbf{R}; \mathbf{C})$ 2π -périodique, on définit la suite $(\widehat{F}(k))_{k \in \mathbf{Z}}$ de ses coefficients de Fourier

$$\widehat{F}(k) := \int_{-\pi}^{\pi} e^{-ikx} F(x) dx,$$

et l'on reconstruit F à partir de la suite $(\widehat{F}(k))_{k \in \mathbf{Z}}$ par la formule

$$F(x) = \frac{1}{2\pi} \sum_{k \in \mathbf{Z}} \widehat{F}(k) e^{ikx}.$$

Cependant, on ne peut pas dire que la transformation de Fourier soit plus générale que les séries de Fourier car toute fonction périodique intégrable sur \mathbf{R} est nulle p.p. sur \mathbf{R} .

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 émergé 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 années 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 où $N = 1$ (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 $G_a \geq 0$ sur \mathbf{R} et

$$\int_{\mathbf{R}} G_a(x) dx = 1,$$

que la *moyenne* de G_a vaut 0, c'est-à-dire que

$$\int_{\mathbf{R}} x G_a(x) dx = 0,$$

et que la *variance* de G_a vaut a , c'est-à-dire que

$$\int_{\mathbf{R}} x^2 G_a(x) dx = a.$$

On invite le lecteur à consulter le cours de S. Méléard *Aléatoire : Introduction à la théorie et au calcul des probabilités* sur les notions de moyenne et variance d'une densité de probabilité.

Ces formules montrent que plus $a > 0$ est petit, plus la densité de probabilité G_a est concentrée sur sa valeur moyenne qui, dans ce cas, est égale à 0. La formule donnant \widehat{G}_a montre alors que plus $a > 0$ est petit, moins la transformée de Fourier \widehat{G}_a est concentrée sur sa valeur moyenne (également nulle.) Autrement dit, plus G_a est localisée et moins \widehat{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. À une fonction $\phi \equiv \phi(x)$ de la variable x de position — typiquement, la fonction d'onde d'une particule — la transformation de Fourier associe la fonction $\widehat{\phi} \equiv \widehat{\phi}(\xi)$ 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 \widehat{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 que

$$\frac{1}{2\pi} \int_{\mathbf{R}} x^2 |G_a(x)|^2 dx \int_{\mathbf{R}} \xi^2 |\widehat{G}_a(\xi)|^2 d\xi = \frac{1}{4} \left(\int_{\mathbf{R}} |G_a(x)|^2 dx \right)^2,$$

alors qu'une formulation mathématique du principe d'incertitude de Heisenberg est l'inégalité

$$\frac{1}{2\pi} \int_{\mathbf{R}} x^2 |\phi(x)|^2 dx \int_{\mathbf{R}} \xi^2 |\widehat{\phi}(\xi)|^2 d\xi \geq \frac{1}{4} \left(\int_{\mathbf{R}} |\phi(x)|^2 dx \right)^2,$$

pour toute fonction $\phi \in \mathcal{C}^1(\mathbf{R}; \mathbf{C})$ telle que

$$\int_{\mathbf{R}} (|x^2 \phi(x)|^2 + |\phi'(x)|^2) dx < \infty.$$

Nous reviendrons sur ce point dans le chapitre suivant qui est consacré 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.

Espaces de Hilbert

La théorie des *espaces de Hilbert* est la théorie la plus simple concernant des espaces vectoriels de dimension infinie. Elle généralise la théorie des *espaces hermitiens* (ou euclidiens) de dimension finie. Le cas plus général des *espaces de Banach*, c'est-à-dire des espaces vectoriels normés complets pour lesquels la norme n'est pas nécessairement induite par un produit scalaire, en est une généralisation plus vaste que nous n'étudierons pas en détail ici, bien que certains résultats de cette théorie sont donnés dans le premier chapitre de ce cours. La théorie des espaces de Hilbert permet déjà de résoudre de nombreux problèmes concrets, notamment certains problèmes variationnels. De plus, c'est le cadre naturel pour la mécanique quantique.

1. Définitions et premières propriétés

Dans ce chapitre tous les espaces vectoriels seront supposés définis sur \mathbf{C} , néanmoins la théorie s'étend aux \mathbf{R} -espaces vectoriels. Dans la suite, on dira qu'une application L entre deux \mathbf{C} -espaces vectoriels E et F est *anti-linéaire* si elle vérifie

$$L(\lambda x + \mu y) = \bar{\lambda} L(x) + \bar{\mu} L(y),$$

quels que soient $\lambda, \mu \in \mathbf{C}$ et $x, y \in E$.

Définition 10.1. On dit qu'une application $\langle \cdot, \cdot \rangle : E \times E \rightarrow \mathbf{C}$ est une :

- (i) une forme sesquilinéaire si $x \mapsto \langle x, y \rangle$ est linéaire et si $y \mapsto \langle x, y \rangle$ est anti-linéaire ;¹
- (ii) une forme hermitienne si c'est une forme sesquilinéaire et si de plus $\langle x, y \rangle = \overline{\langle y, x \rangle}$, pour tous $x, y \in E$. En particulier, $\langle x, x \rangle$ est réel pour tout $x \in E$;
- (iii) un produit hermitien si c'est une forme hermitienne définie positive, c'est-à-dire, si $\langle x, x \rangle > 0$ pour $x \neq 0$.

Remarque 10.1. Dans le cadre des \mathbf{R} -espaces vectoriels, les formes sesquilinéaires (resp. hermitiennes) sont simplement les formes bilinéaires (resp. bilinéaires symétriques) et les produits hermitiens des produits scalaires.

Proposition 10.1. (*Inégalité de Cauchy-Schwarz*) Soit $\langle \cdot, \cdot \rangle$ un produit hermitien sur E . Alors, pour tous $x, y \in E$

$$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle},$$

avec égalité si et seulement si la famille $\{x, y\}$ est liée.

DÉMONSTRATION. Supposons que $x \neq 0$ et $y \neq 0$ autrement le résultat est immédiat. Pour tout $t > 0$, on calcule

$$\langle x - ty, x - ty \rangle = \langle x, x \rangle - 2t \Re \langle x, y \rangle + t^2 \langle y, y \rangle \geq 0,$$

1. Il existe d'autres conventions, notamment en physique, pour lesquelles les formes sesquilinéaires sont anti-linéaires par rapport à la première variable et linéaires par rapport à la deuxième variable.

De telle sorte que, pour tout $t > 0$

$$2\Re\langle x, y \rangle \leq \frac{1}{t} \langle x, x \rangle + t \langle y, y \rangle.$$

On optimise cette inégalité en choisissant t de telle sorte que le membre de droite soit minimal. Autrement dit

$$t = \sqrt{\frac{\langle x, x \rangle}{\langle y, y \rangle}}.$$

On conclut que

$$\Re\langle x, y \rangle \leq \sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle}.$$

On obtient l'inégalité voulue en remplaçant x par $e^{i\theta}x$ pour $\theta \in \mathbf{R}$ et en utilisant le fait que

$$|z| = \sup_{\theta \in \mathbf{R}} \Re(e^{i\theta} z),$$

pour tout nombre complexe $z \in \mathbf{C}$.

L'étude des cas d'égalité est laissée en exercice au lecteur. \square

Si $\langle \cdot, \cdot \rangle$ est un produit hermitien sur E , on peut définir

$$\|x\| := \sqrt{\langle x, x \rangle},$$

dont on vérifie que c'est une norme sur E . Le seul point délicat est la démonstration de l'inégalité triangulaire qui est une conséquence de la formule de Cauchy-Schwarz. En effet, pour tous $x, y \in E$, en développant $\langle x + y, x + y \rangle$ on trouve

$$\|x + y\|^2 = \|x\|^2 + 2\Re\langle x, y \rangle + \|y\|^2 \leq \|x\|^2 + 2\|x\|^{1/2}\|y\|^{1/2} + \|y\|^2 = (\|x\| + \|y\|)^2.$$

Définition 10.2. On appelle espace préhilbertien un espace vectoriel E muni d'un produit hermitien $\langle \cdot, \cdot \rangle$ et de la norme associée $\|\cdot\|$. C'est un espace métrique pour la distance $d(x, y) := \|x - y\|$.

Exemple 10.1. Sur \mathbf{C}^N , la forme sesquilinéaire

$$\langle x, y \rangle := \sum_{j=1}^N x_j \bar{y}_j,$$

définit un produit hermitien. Sur un espace vectoriel de dimension finie, tout produit hermitien se ramène à cette forme dans une base adaptée (voir l'exercice ci-dessous).

Exercice 10.1. Soit $\langle \cdot, \cdot \rangle$ un produit hermitien sur H , un \mathbf{C} -espace vectoriel de dimension finie $N \geq 1$. On veut montrer qu'il existe (e_1, \dots, e_N) , une base de H , dans laquelle

$$\langle x, y \rangle = \sum_{j=1}^N x_j \bar{y}_j, \quad \text{si } x := \sum_{i=1}^N x_i e_i, \quad \text{et } y := \sum_{i=1}^N y_i e_i.$$

1) Montrer que le résultat est vrai lorsque $\dim H = 1$.

2) Montrer que, si H est de dimension $N \geq 1$, il existe e_1 tel que $\langle e_1, e_1 \rangle = 1$. Montrer que

$$(\mathbf{C}e_1)^\perp := \{x \in H : \langle x, e_1 \rangle = 0\},$$

est un espace de dimension $N - 1$ et que la restriction de $\langle \cdot, \cdot \rangle$ à $(\mathbf{C}e_1)^\perp$ est un produit hermitien. Conclure.

1.1. Propriétés du produit hermitien. On suppose que E est un espace préhilbertien.

Proposition 10.2. *Les propriétés suivantes sont valables pour tous $x, y, a, b \in E$:*

(i) *Théorème de Pythagore*

$$\Re \langle x, y \rangle = 0 \iff \|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

(ii) *Égalité du parallélogramme*

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

(iii) *Égalité de la médiane*

$$\|x - a\|^2 + \|x - b\|^2 = 2 \left\| x - \frac{a + b}{2} \right\|^2 + 2 \left\| \frac{a - b}{2} \right\|^2.$$

(iv) *Formule de polarisation*

$$\langle x, y \rangle = \frac{\|x + y\|^2 - \|x - y\|^2}{4} + i \left(\frac{\|x + iy\|^2 - \|x - iy\|^2}{4} \right).$$

DÉMONSTRATION. Simple calcul à partir de la formule

$$\|x + y\|^2 = \|x\|^2 + 2\Re \langle x, y \rangle + \|y\|^2.$$

Les détails sont laissés en exercice au lecteur. \square

On prendra garde que, dans un \mathbf{C} -espace préhilbertien, l'égalité

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2,$$

n'entraîne pas forcément que $\langle x, y \rangle = 0$, mais simplement $\Re \langle x, y \rangle = 0$. On dira que x et y sont *orthogonaux* si $\langle x, y \rangle = 0$ et l'on dira que x et y sont *perpendiculaires* si $\Re \langle x, y \rangle = 0$. En particulier, x et ix sont perpendiculaires puisque $\langle x, ix \rangle = -i \langle x, x \rangle$ et donc $\Re \langle x, ix \rangle = 0$, pour tout $x \in E$. Remarquons que, dans le cas des espaces préhilbertiens réels, ces deux notions sont confondues.

1.2. Espaces de Hilbert. Soit H un \mathbf{C} -espace vectoriel muni d'un produit hermitien $\langle \cdot, \cdot \rangle$ (i.e. un espace préhilbertien). On note $\| \cdot \|$ la norme associée à $\langle \cdot, \cdot \rangle$.

Définition 10.3. *On dit que H (sous-entendu, muni de la norme $\| \cdot \|$ associée au produit hermitien $\langle \cdot, \cdot \rangle$) est un espace de Hilbert si $(H, \| \cdot \|)$ est un espace de Banach.*

Étant donné que tout sous-ensemble fermé d'un espace métrique complet est un espace métrique complet (muni de la distance induite), 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, et plus généralement dans les espaces de dimension infinie, il existe des sous-espaces qui ne sont pas fermés (par exemple des sous-espaces denses qui ne sont pas égaux à l'espace entier). Nous en verrons un exemple un peu plus tard (voir les exemples 10.10 page 193).

Remarque 10.2. Si E est un \mathbf{C} -espace vectoriel normé, on peut se demander si sa norme provient d'une norme hilbertienne. La réponse est donnée par la *formule de polarisation* : la norme $\| \cdot \|$ est une norme hermitienne si et seulement si

$$(10.1) \quad \langle x, y \rangle := \frac{\|x + y\|^2 - \|x - y\|^2}{4} + i \left(\frac{\|x + iy\|^2 - \|x - iy\|^2}{4} \right),$$

est une forme hermitienne. Si tel est le cas, on a bien $\langle x, x \rangle = \|x\|^2$.

Venons-en maintenant aux exemples classiques d'espaces de Hilbert.

Exemple 10.2. 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.

Exemple 10.3. L'espace $\ell^2(\mathbf{N}; \mathbf{C})$ des suites complexes $\mathbf{x} := (x_n)_{n \geq 0}$ telles que

$$\sum_{n \geq 0} |x_n|^2 < +\infty,$$

muni du produit hermitien défini par

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\ell^2} := \sum_{n=0}^{+\infty} x_n \overline{y_n},$$

est un espace de Hilbert. On notera $\|\cdot\|_{\ell^2}$ la norme associée à $\langle \cdot, \cdot \rangle_{\ell^2}$.

Il est clair que $\langle \cdot, \cdot \rangle_{\ell^2}$ est un produit hermitien sur $\ell^2(\mathbf{N}; \mathbf{C})$. Montrons que cet espace, muni de la norme $\|\cdot\|_{\ell^2}$, est complet. Si $(\mathbf{x}^{(k)})_{k \geq 0}$ est une suite de Cauchy de $\ell^2(\mathbf{N}; \mathbf{C})$, on peut écrire

$$\mathbf{x}^{(k)} = (x_n^{(k)})_{n \geq 0}.$$

On vérifie que, pour tout $n \in \mathbf{N}$, la suite $(x_n^{(k)})_{k \geq 0}$ est une suite de Cauchy dans \mathbf{C} (car $|x_n^{(k)} - x_n^{(l)}| \leq \|\mathbf{x}^{(k)} - \mathbf{x}^{(l)}\|_{\ell^2}$) et donc, elle converge vers $z_n \in \mathbf{C}$. On prétend maintenant que la suite $\mathbf{z} := (z_n)_{n \geq 0}$ appartient à $\ell^2(\mathbf{N}; \mathbf{C})$ et est la limite des $(\mathbf{x}^{(k)})_{k \geq 0}$. En effet, pour tout n , on peut écrire

$$\sum_{j=0}^n |x_j^{(k)} - z_j|^2 = \lim_{m \rightarrow +\infty} \sum_{j=0}^n |x_j^{(k)} - x_j^{(m)}|^2 \leq \sup_{m \geq k} \|\mathbf{x}^{(k)} - \mathbf{x}^{(m)}\|_{\ell^2}^2.$$

Étant donné que le terme de droite est indépendant de n , on obtient en faisant tendre n vers l'infini

$$\|\mathbf{x}^{(k)} - \mathbf{z}\|_{\ell^2} \leq \sup_{m \geq k} \|\mathbf{x}^{(k)} - \mathbf{x}^{(m)}\|_{\ell^2},$$

qui tend vers 0 quand k tend vers $+\infty$, car la suite $(\mathbf{x}^{(k)})_{k \geq 0}$ est une suite de Cauchy. On en déduit d'une part que $\mathbf{z} \in \ell^2(\mathbf{N}; \mathbf{C})$ et d'autre part que $(\mathbf{x}^{(k)})_{k \geq 0}$ tend vers \mathbf{z} dans cet espace.

Exemple 10.4 (Espace de Sobolev discret). Pour tout $k \in \mathbf{N}$, on définit l'espace $H^k(\mathbf{N}; \mathbf{C})$ comme étant le \mathbf{C} -espace vectoriel des suites complexes $\mathbf{x} := (x_n)_{n \geq 0}$ telles que

$$\sum_{n \in \mathbf{N}} (1 + n^2)^k |x_n|^2 < +\infty,$$

muni du produit hermitien

$$\langle \mathbf{x}, \mathbf{y} \rangle_{H^k} := \sum_{n \in \mathbf{N}} (1 + n^2)^k x_n \overline{y_n}.$$

On vérifie que $H^k(\mathbf{N}; \mathbf{C})$ est ainsi muni d'une structure d'espace de Hilbert (la démonstration est analogue à celle de l'exemple précédent).

Exemple 10.5. L'espace $L^2(\mathbf{R}; \mathbf{C})$ des fonctions de carré intégrable sur \mathbf{R} , muni du produit hermitien

$$\langle f, g \rangle_{L^2(\mathbf{R}; \mathbf{C})} := \int_{\mathbf{R}} f(t) \overline{g(t)} dt,$$

est un espace de Hilbert (voir le Théorème 8.4 page 164).

Exemple 10.6. L'espace $L^2(S^1; \mathbf{C})$ des fonctions à valeurs complexes qui sont 2π -périodiques, c'est-à-dire telles que $f(t + 2\pi) = f(t)$ pour presque tout $t \in \mathbf{R}$ et telles que

$$\int_{[0, 2\pi]} |f(t)|^2 dt < +\infty,$$

muni du produit hermitien

$$\langle f, g \rangle_{L^2} := \frac{1}{2\pi} \int_{[0, 2\pi]} f(t) \overline{g(t)} dt,$$

est un espace de Hilbert. Cet espace est canoniquement isomorphe à l'espace $L^2([0, 2\pi]; \mathbf{C})$, car toute fonction définie sur $[0, 2\pi]$ coïncide p.p. avec une unique fonction 2π -périodique. On utilisera dans la suite indifféremment les espaces $L^2(S^1; \mathbf{C})$ et l'espace $L^2([0, 2\pi]; \mathbf{C})$ sachant que ces deux espaces s'identifient de manière naturelle.

Si $(E, \langle \cdot, \cdot \rangle)$ est un espace préhilbertien, on peut le compléter² en un espace de Hilbert dont E est un sous-espace dense.

Proposition 10.3. *Tout espace préhilbertien $(E, \langle \cdot, \cdot \rangle)$ admet un complété H dans lequel il est dense. Inversement, si E est un sous-espace vectoriel dense d'un espace de Hilbert $(H, \langle \cdot, \cdot \rangle)$, le complété de E est isomorphe à H .*

DÉMONSTRATION. On considère le \mathbf{C} -espace vectoriel \tilde{E} des suites de Cauchy de E , muni de la relation d'équivalence

$$(x_n)_{n \geq 0} \simeq (y_n)_{n \geq 0} \iff \lim_{n \rightarrow +\infty} \|x_n - y_n\| = 0.$$

Si $(x_n)_{n \geq 0} \in \tilde{E}$, on définit la classe d'équivalence de $(x_n)_{n \geq 0}$ par

$$[(x_n)_{n \geq 0}] := \{(y_n)_{n \geq 0} : (y_n)_{n \geq 0} \simeq (x_n)_{n \geq 0}\}.$$

On appelle H l'ensemble des classes d'équivalence³ des éléments de \tilde{E} . On vérifie que H est un \mathbf{C} -espace vectoriel et que l'on peut le munir du produit hermitien

$$([(x_n)_{n \geq 0}], [(y_n)_{n \geq 0}])_H := \lim_{n \rightarrow +\infty} \langle x_n, y_n \rangle.$$

Étant donné que

$$|\langle x_n, y_n \rangle - \langle x_m, y_m \rangle| \leq \|x_n - x_m\| \|y_n\| + \|x_m\| \|y_n - y_m\|,$$

on montre que $\langle x_n, y_n \rangle$ est une suite de Cauchy, et que sa limite ne dépend pas des représentants des classes d'équivalence $[(x_n)_{n \geq 0}]$ et $[(y_n)_{n \geq 0}]$.

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 .

On identifie ensuite l'espace E avec un sous-espace de H en associant à $x \in E$, la classe d'équivalence $[(x)_{n \geq 0}]$ de la suite de \tilde{E} qui est constante égale à x . On vérifie sans difficulté que

$$E \ni x \mapsto [(x)_{n \geq 0}] \in H,$$

est une isométrie, c'est-à-dire que

$$\|x\| = \|[(x)_{n \geq 0}]\|_H.$$

La densité de E dans H se démontre comme suit. Soit $\varepsilon > 0$ un réel, et $(x_n)_{n \geq 0}$ une suite de \tilde{E} . On choisit N assez grand pour que $\|x_n - x_m\| \leq \varepsilon$ pour $m, n \geq N$. Alors la classe d'équivalence $[(x_N)_{n \geq 0}]$ de la suite constante égale à x_N est à distance au plus ε de $[(x_n)_{n \geq 0}]$, ce qui démontre la densité de E dans H . On vérifie (exercice) que H est un espace métrique complet. \square

Exemple 10.7. On considère l'espace $L^2(\mathbf{R}, e^{-\frac{x^2}{2}} dx)$ qui est l'espace des fonctions mesurables, qui sont définies sur \mathbf{R} à valeurs dans \mathbf{C} et dont le carré du module est intégrable pour la mesure $e^{-\frac{x^2}{2}} dx$, i.e. le complété de l'espace préhilbertien des fonctions continues à support compact qui vérifient

$$\int_{\mathbf{R}} |f(x)|^2 e^{-\frac{x^2}{2}} dx < +\infty,$$

et qui est muni du produit hermitien

$$\langle f, g \rangle := \int_{\mathbf{R}} f(x) \overline{g(x)} e^{-\frac{x^2}{2}} dx.$$

On peut bien entendu remplacer la densité $e^{-\frac{x^2}{2}}$ par n'importe quelle fonction $x \mapsto w(x)$ continue, strictement positive sur \mathbf{R} . On note $L^2(\mathbf{R}, w(x) dx)$ cet espace, et l'on l'appelle *espace L^2 à poids w* .

Exemple 10.8. Soient H_1, H_2 deux espaces de Hilbert munis des produits hermitiens $\langle \cdot, \cdot \rangle_1$ et $\langle \cdot, \cdot \rangle_2$. Alors $H_1 \oplus H_2 = \{(x_1, x_2) : x_1 \in H_1, x_2 \in H_2\}$ 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* . On vérifiera à titre d'exercice que $H_1 \oplus H_2$ est complet.

Définition 10.4. Soient H_1, H_2 deux espaces de Hilbert et $L : H_1 \rightarrow H_2$ une application linéaire. On dit que L est un isomorphisme d'espaces préhilbertiens (ou un isomorphisme isométrique) si L est une application bijective et si, pour tout $x \in H_1$ on a

$$\|L(x)\|_2 = \|x\|_1.$$

Remarquons que la formule de polarisation montre que si L est une isométrie, elle préserve le produit hermitien : pour tous $x, y \in H_1$, on a

$$\langle L(x), L(y) \rangle_2 = \langle x, y \rangle_1.$$

Remarquons aussi, qu'en dimension infinie, la bijectivité de L ne résulte pas de la conservation de la norme (qui pourtant garantit l'injectivité).

Exemple 10.9. L'application $u : \ell^2(\mathbf{N}) \rightarrow \ell^2(\mathbf{N})$ définie par

$$u((x_0, x_1, x_2, \dots, x_n, \dots)) = (0, x_0, x_1, x_2, \dots, x_{n-1}, \dots),$$

préserve la norme mais ce n'est pas un isomorphisme. C'est aussi un exemple d'application qui est injective mais qui n'est pas surjective.

De même, il existe des applications surjectives qui ne sont pas injectives, comme par exemple l'application

$$v((x_0, x_1, x_2, \dots, x_n, \dots)) = (x_1, x_2, x_3, \dots, x_{n+1}, \dots),$$

définie de $\ell^2(\mathbf{N})$ dans lui même.

2. Sous-espaces fermés, applications continues

Sauf mention contraire, H désignera dans la suite un espace de Hilbert. On notera $\langle \cdot, \cdot \rangle$ le produit hermitien et $\| \cdot \|$ la norme associée. Soit F un sous-espace vectoriel d'un espace de H . Contrairement au cas de la dimension finie, en dimension infinie F n'est pas nécessairement fermé. Donnons quelques exemples de ce phénomène qui est spécifique aux espaces de dimension infinie.

Exemple 10.10. Soit $\ell_c(\mathbf{N})$ le sous-espace des suites de $\ell^2(\mathbf{N})$ qui sont nulles à partir d'un certain rang. Ce sous-espace est dense dans $\ell^2(\mathbf{N})$, car pour tout $\mathbf{x} = (x_n)_{n \geq 0} \in \ell^2(\mathbf{N})$ et pour tout $\varepsilon > 0$, il existe $m \in \mathbf{N}$ tel que

$$\sum_{n \geq m} |x_n|^2 < \varepsilon.$$

Il en résulte que si la suite $\bar{\mathbf{x}} := (\bar{x}_n)_{n \geq 0}$ est définie par $\bar{x}_n = x_n$ pour $n < m$ et $\bar{x}_n = 0$ pour $n \geq m$, on aura $\|\mathbf{x} - \bar{\mathbf{x}}\|_{\ell^2} \leq \varepsilon$. Bien entendu, il existe des suites de $\ell^2(\mathbf{N})$ qui ont une infinité de termes non nuls, et le sous-espace $\ell_c^2(\mathbf{N})$ est donc distinct de $\ell^2(\mathbf{N})$.

Nous verrons une deuxième démonstration de ce résultat de densité, qui elle sera basée sur le Théorème de Riesz (voir le Théorème 10.16 page 202).

Exemple 10.11. On sait que l'espace vectoriel $\mathcal{C}_c(\mathbf{R}; \mathbf{C})$ des fonctions continues à support compact est dense dans $L^2(\mathbf{R}; \mathbf{C})$. Étant donné qu'il existe des fonctions de $L^2(\mathbf{R}; \mathbf{C})$ qui ne sont pas continues, le sous-espace $\mathcal{C}_c(\mathbf{R}; \mathbf{C})$ n'est pas fermé dans $L^2(\mathbf{R}; \mathbf{C})$.

Remarque 10.3. On démontre aisément que, si F est un sous-espace vectoriel d'un espace de Hilbert H , alors \overline{F} , l'adhérence de F dans H , est un sous-espace vectoriel fermé de H .

Définition 10.5. Si F est un sous-espace vectoriel de H , On note

$$F^\perp := \{x \in H : \forall y \in F, \langle x, y \rangle = 0\},$$

l'orthogonal de F .

Proposition 10.4. Soit F un sous-espace vectoriel d'un espace de Hilbert H , alors :

- (i) le sous-espace F^\perp est fermé;
- (ii) si G est un sous-espace vectoriel inclus dans F , alors $F^\perp \subset G^\perp$;
- (iii) on a l'égalité $F^\perp = \overline{F}^\perp$.

DÉMONSTRATION. Démontrons le point (i). Soit $(x_n)_{n \geq 0}$ une suite de H convergeant vers x et $y \in F$. On suppose que $\langle x_n, y \rangle = 0$ pour tout $n \geq 0$. Alors, en utilisant la sesquilinearité de $\langle \cdot, \cdot \rangle$ et l'inégalité de Cauchy-Schwarz, on obtient

$$|\langle x, y \rangle| = |\langle x, y \rangle - \langle x_n, y \rangle| = |\langle x - x_n, y \rangle| \leq \|x_n - x\| \|y\|,$$

qui tend vers 0 quand n tend vers $+\infty$. On en déduit que $\langle x, y \rangle = 0$. Conclusion, $x \in F^\perp$ et F^\perp est donc fermé.

Démontrons maintenant le point (iii). Tout d'abord, $F \subset \overline{F}$ donc $\overline{F}^\perp \subset F^\perp$. Soit $y \in F^\perp$ et $x \in \overline{F}$. Il existe une suite $(x_n)_{n \geq 0}$ d'éléments de F qui converge vers x . Comme ci-dessus, on vérifie que $\langle x, y \rangle = \lim_{n \rightarrow +\infty} \langle x_n, y \rangle = 0$. Donc $y \in \overline{F}^\perp$, d'où $F^\perp \subset \overline{F}^\perp$. Ce qui termine la démonstration. \square

Dans les espaces de dimension infinie, de même qu'il existe des sous-espaces qui ne sont pas fermés, il existe des applications linéaires qui ne sont pas continues. Donnons quelques exemples d'applications linéaires continues définies entre espaces de Hilbert de dimensions infinies.

Exemple 10.12. Si g est une fonction continue bornée, l'application M_g définie de $L^2(\mathbf{R}; \mathbf{C})$ dans lui-même, par

$$M_g(f) := g f,$$

est continue. En effet,

$$\|M_g(f)\|_{L^2} \leq \left(\sup_{t \in \mathbf{R}} |g(t)| \right) \|f\|_{L^2}.$$

On laisse au lecteur le soin de vérifier qu'il suffit que g soit mesurable et appartienne à $L^\infty(\mathbf{R}; \mathbf{C})$ pour que M_g soit une application continue. Si tel est le cas, on montrera que la norme de M_g est égale à $\|g\|_{L^\infty}$.

Exemple 10.13. L'application $I : L^2([0, 1]; \mathbf{C}) \rightarrow L^2([0, 1]; \mathbf{C})$ définie par

$$I(f)(x) := \int_{[0, x]} f(y) dy,$$

pour tout $x \in [0, 1]$, est continue. En effet, on a

$$|I(f)(x)| \leq \int_{[0, 1]} |f(x)| dx \leq \left(\int_{[0, 1]} |f(x)|^2 dx \right)^{1/2} = \|f\|_{L^2},$$

donc

$$\|I(f)\|_{L^2} = \left(\int_{[0, 1]} |I(f)(x)|^2 dx \right)^{1/2} \leq \left(\int_{[0, 1]} \|f\|_{L^2}^2 dx \right)^{1/2} \leq \|f\|_{L^2},$$

l'application I est donc continue.

Remarquons que l'inégalité de Cauchy-Schwarz nous assure que, pour tout $x < y$

$$|I(f)(y) - I(f)(x)| = \int_{[x, y]} |f(x)| dx \leq \left(\int_{[x, y]} |f(x)|^2 dx \right)^{1/2} |y - x|^{1/2} = \|f\|_{L^2} |y - x|^{1/2},$$

donc $I(f) \in \mathcal{C}([0, 1]; \mathbf{C})$. On prendra soin de ne pas confondre le fait que $I(f)$ est une *fonction continue* sur $[0, 1]$ et le fait que l'application $f \mapsto I(f)$ est une *application continue* de $L^2([0, 1]; \mathbf{C})$ dans lui-même.

Exemple 10.14. Soit $A : \ell^2(\mathbf{N}) \rightarrow \ell^2(\mathbf{N})$ l'application définie par

$$A((x_n)_{n \geq 0}) := \left(\frac{x_n}{n+1} \right)_{n \geq 0}.$$

On vérifie que

$$\|A(\mathbf{x})\|_{\ell^2} \leq \|\mathbf{x}\|_{\ell^2},$$

pour tout $\mathbf{x} \in \ell^2(\mathbf{N})$. Donc, A est une application linéaire continue. Clairement, l'image de l'application A contient $\ell_c(\mathbf{N})$, l'espace des suites nulles à partir d'un certain rang, donc l'image de A est dense dans $\ell^2(\mathbf{N})$. En revanche, l'image de A n'est pas fermée : par exemple, la suite $((n+1)^{-3/2})_{n \geq 0}$ est clairement un élément de $\ell^2(\mathbf{N})$ mais elle n'appartient pas à l'image de A .

Concentrons nous maintenant sur les formes linéaires continues définies sur un espace de Hilbert H (i.e. les applications linéaires, continues qui sont définies sur H et à valeurs dans \mathbf{R} ou \mathbf{C}). Rappelons que l'espace vectoriel H' des formes linéaires *continues* définies sur H est appelé *dual topologique* de H (voir la Définition 1.12 page 26).

L'espace H' est muni de la norme définie par

$$\|u\|_{H'} := \sup_{\|x\|=1} |u(x)|.$$

Nous pouvons énoncer la :

Proposition 10.5. *L'espace H' est un espace de Hilbert, pour la norme ci-dessus.*

Que, muni de la norme ci-dessus, l'espace H' soit un espace de Banach, c'est clair puisque c'est justement le résultat de la Proposition 3.10. Que ce soit un espace de Hilbert, c'est-à-dire le fait que la norme définie ci-dessus est induite par un produit hermitien, c'est moins clair et cela résultera du Théorème de représentation de Riesz (voir le Théorème 10.3 page 198).

Remarque 10.4. Le résultat ci-dessus suggère une méthode bien plus simple que celle présentée plus haut pour construire le complété d'un espace préhilbertien. En effet, soit E un espace préhilbertien et E' le dual topologique de E , c'est-à-dire l'ensemble des formes linéaires continues sur E . On le munit de la norme

$$\|u\|_{E'} := \sup_{\|x\|_E=1} |u(x)|,$$

dont on ignore pour le moment si elle est induite par un produit hermitien. L'espace E s'identifie à un sous-espace de E' grâce à l'identification de $y \in E$ à la forme linéaire continue $\Lambda_y \in E'$ définie pour tout $x \in E$ par

$$\Lambda_y(x) := \langle x, y \rangle.$$

Il est facile de vérifier que cette application est bien injective. De plus, on sait que E^* est un espace de Banach. On peut alors prendre pour H l'adhérence de E dans E' . C'est un espace hilbertien, car il est complet (c'est un fermé dans un espace de Banach), et sa norme est définie par un produit hermitien, car d'après la Remarque 10.2 page 189, il suffit de vérifier que l'expression donnée par la formule de polarisation (10.1) vérifie des identités algébriques. Or ces identités étant vérifiées sur un espace dense (l'image de E dans E'), elles le sont aussi sur H tout entier (principe de prolongement des identités par continuité). On pourra ensuite vérifier *a posteriori* que $H = E'$.

3. Théorème de la projection sur un convexe fermé

Commençons par le Théorème de la projection sur un ensemble convexe fermé :

Théorème 10.1 (Théorème de la projection sur un convexe fermé). *Soit H un espace de Hilbert et F un sous-ensemble convexe fermé de H . Pour tout $x \in H$, il existe un unique $y \in F$ tel que*

$$\|x - y\| = d(x, F) := \inf_{z \in F} \|x - z\|.$$

De plus, si $x \in F$ alors $y = x$ et si $x \notin F$, y est caractérisé par la condition

$$\Re \langle x - y, z - y \rangle \leq 0,$$

pour tout $z \in F$.

DÉMONSTRATION. Soit $(y_n)_{n \geq 0}$ une suite d'éléments de F telle que

$$(10.2) \quad \lim_{n \rightarrow +\infty} \|x - y_n\| = d(x, F).$$

Montrons que $(y_n)_{n \geq 0}$ est une suite de Cauchy dans F . Pour ceci, remarquons que F étant supposé convexe

$$\frac{y_m + y_n}{2} \in F,$$

en particulier,

$$\left\| x - \frac{y_m + y_n}{2} \right\| \geq d(x, F).$$

Pour tout $\varepsilon > 0$, il existe un entier $n_\varepsilon \geq 0$ tel que pour tout $n \geq n_\varepsilon$ on a $\|x - y_n\| \leq d(x, F) + \varepsilon$.

Utilisons maintenant l'égalité du parallélogramme

$$\|y_n - y_m\|^2 + \|2x - y_n - y_m\|^2 = 2(\|y_n - x\|^2 + \|y_m - x\|^2),$$

et l'inégalité ci-dessus, pour conclure que

$$\|y_n - y_m\|^2 \leq 4((d(x, F) + \varepsilon)^2 - d(x, F)^2) = 4\varepsilon(2d(x, F) + \varepsilon),$$

pour tous $n, m \geq n_\varepsilon$. On en déduit que $(y_n)_{n \geq 0}$ est une suite de Cauchy de H , qui est complet, donc elle converge vers un élément $y \in H$. Comme F est supposé, fermé, on conclut que $y \in F$. En passant à limite dans (10.2) quand n tend vers $+\infty$, on obtient

$$\|y - x\| = d(x, F).$$

Montrons maintenant l'unicité de y . Si y_1, y_2 sont deux points réalisant le minimum de la distance de x à F , l'égalité du parallélogramme nous assure que

$$\|y_1 - y_2\|^2 = 2(\|y_1 - x\|^2 + \|y_2 - x\|^2) - 4 \left\| x - \frac{y_1 + y_2}{2} \right\|^2 \leq 4d(x, F)^2 - 4d(x, F)^2 = 0,$$

donc $y_1 = y_2$.

Montrons enfin que y est caractérisé par le fait que $\Re \langle x - y, z - y \rangle \leq 0$ pour tout $z \in F$. Pour cela, on utilise le fait que l'ensemble F étant convexe, le segment $[y, z] \subset F$. On peut alors calculer

$$d(x, F)^2 \leq \|x - z_t\|^2 = \|x - y\|^2 + 2\Re \langle x - y, y - z_t \rangle + \|y - z_t\|^2$$

où $z_t = (1 - t)y + tz$, pour tout $t \in [0, 1]$. Comme $\|x - y\|^2 = d(x, F)^2$, on conclut que

$$2\Re \langle x - y, y - z_t \rangle + \|y - z_t\|^2 \geq 0$$

Donc $2t \Re \langle x - y, z - y \rangle \leq t^2 \|y - z\|^2$, pour tout $t \in [0, 1]$ en divisant par $t > 0$ et en faisant tendre t vers 0, on en déduit que $\Re \langle x - y, z - y \rangle \leq 0$. Ce qui termine la démonstration. \square

Un cas particulièrement important d'ensemble convexe est celui des sous-espaces vectoriels. On obtient alors comme corollaire du résultat ci-dessus :

Théorème 10.2 (Théorème de la projection). *Soit H un espace de Hilbert et F un sous-espace vectoriel fermé de H . Il existe une unique application linéaire $P_F : E \rightarrow F$ telle que, pour tout $x \in H$,*

$$\|x - P_F(x)\| = d(x, F) := \inf_{z \in F} \|x - z\|.$$

De plus $P_F(x)$ est l'unique élément de F vérifiant cette égalité et $x - P_F(x)$ est orthogonal à tout vecteur de F . Enfin P_F est 1-lipschitzienne (donc continue). Autrement dit

$$\|P_F(x) - P_F(y)\| \leq \|x - y\|,$$

pour tous $x, y \in E$.

DÉMONSTRATION. L'existence de P_F est garantie grâce au résultat ci-dessus. Montrons que $x - P_F(x) \in F^\perp$. On sait déjà que, pour tout $z \in F$, on a

$$\Re \langle x - y, z - y \rangle \leq 0,$$

où $y := P_F(x)$. Mais, F étant un sous-espace vectoriel, on en déduit que $\Re \langle x - y, w \rangle \leq 0$ pour tout $w \in F$. En remplaçant w par $-w$ puis par iw , on conclut que $\langle x - y, w \rangle = 0$ pour tout $w \in F$.

La linéarité est une conséquence de la caractérisation que nous verrons de démontrer et de l'unicité de la projection. Reste enfin à vérifier que P_F est 1-lipschitzienne. Or cela résulte immédiatement du Théorème de Pythagore

$$\|x\|^2 = \|x - P_F(x)\|^2 + \|P_F(x)\|^2 \geq \|P_F(x)\|^2.$$

Ce qui termine la démonstration. \square

L'exemple le plus élémentaire de sous-espace fermé dans un espace de Hilbert est celui d'un sous-espace de dimension finie. Supposons que F est engendré par les vecteurs e_1, \dots, e_n qui forment une famille orthonormée, c'est-à-dire que $\langle e_i, e_j \rangle = \delta_i^j$, la projection P_F peut alors s'écrire

$$P_F(x) = \sum_{j=1}^n \langle x, e_j \rangle e_j.$$

Pour démontrer ce résultat, il suffit 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

$$\langle x - P_F(x), e_j \rangle = \langle x, e_j \rangle - \langle x, e_j \rangle = 0.$$

Une autre conséquence du Théorème de la projection est le :

Corollaire 10.1. *Si F est un sous-espace fermé de H , on a la décomposition $H = F \oplus F^\perp$. Par conséquent, si F est un sous-espace fermé, $(F^\perp)^\perp = F$. Plus généralement, si F est un sous-espace de H - qui n'est pas nécessairement fermé - on a $(F^\perp)^\perp = \overline{F}$.*

DÉMONSTRATION. La première partie de l'énoncé n'est qu'une reformulation du Théorème de la projection sur un sous-espace fermé qui permet de décomposer

$$x = P_F(x) + (x - P_F(x)),$$

avec $x - P_F(x) \in F^\perp$ et le fait que $F \cap F^\perp = \{0\}$. La seconde partie de l'énoncé résulte de ce que, par définition, $F \subset (F^\perp)^\perp$. Si l'inclusion était stricte, il existerait, en vertu de la décomposition déjà démontrée, un élément $x \neq 0$ qui appartiendrait à $(F^\perp)^\perp$ et à F^\perp . Mais nécessairement $\langle x, x \rangle = 0$ donc $x = 0$, ce qui constitue une contradiction.

Enfin, dans le cas général, on a vu dans la Proposition 10.4 page 193 que $F^\perp = \overline{F}^\perp$, donc $(F^\perp)^\perp = (\overline{F}^\perp)^\perp = \overline{F}$. \square

Il résulte du corollaire précédent qu'un sous-espace $F \subset H$ est fermé si et seulement s'il existe un espace vectoriel normé G et une application linéaire continue $L : H \rightarrow G$ telle que $F = \text{Ker } L$. Le fait que la condition soit suffisante est évidente, la nécessité s'obtient en prenant $G = H$ et $L := \text{Id}_H - P_F$.

4. Théorèmes de représentation de Riesz

Soit H un espace de Hilbert. Pour tout $a \in H$, on note Λ_a la forme linéaire définie par

$$\Lambda_a(x) := \langle x, a \rangle,$$

pour tout $x \in H$. On vérifie que $\Lambda_a \in H'$ et que

$$\|\Lambda_a\|_{H'} = \|a\|.$$

En effet, l'inégalité de Cauchy-Schwarz nous assure que

$$|\Lambda_a(x)| = |\langle x, a \rangle| \leq \|a\| \|x\|,$$

donc $\|\Lambda_a\|_{H'} := \sup_{\|x\|=1} |\Lambda_a(x)| \leq \|a\|$. Enfin, $\Lambda_a(a) = \|a\|^2$ ce qui montre que $\|\Lambda_a\|_{H'} \geq \|a\|$. D'où l'égalité.

Tirons les conséquences des résultats précédents, en commençant par le Théorème de représentation de Riesz.

Théorème 10.3 (Théorème de représentation de Riesz). *Soit H un espace de Hilbert et $u \in H'$ une forme linéaire continue sur H . Alors, il existe un unique $a \in H$ tel que, pour tout $x \in H$,*

$$u(x) = \Lambda_a(x).$$

De plus, l'application $a \mapsto \Lambda_a$ qui est définie de H sur H' , est un isomorphisme anti-linéaire isométrique.

DÉMONSTRATION. On se donne $u \in H'$. Démontrons l'existence de $a \in H$ tel que $u = \Lambda_a$. Si $u = 0$, il suffit de prendre $a = 0$. On suppose donc désormais que $u \neq 0$. Comme u est une forme linéaire continue sur H , l'espace $F = \text{Ker } u$ est un sous-espace fermé de H et le Corollaire 10.1 nous permet de décomposer

$$H = F \oplus F^\perp.$$

On remarque que F^\perp est de dimension un, car si $a \in F^\perp$ est un vecteur non nul et si $x \in F^\perp$, on peut écrire

$$u\left(x - \frac{u(x)}{u(a)} a\right) = 0,$$

donc

$$x - \frac{u(x)}{u(a)} a \in F \cap F^\perp = \{0\}.$$

Conclusion $x = \frac{u(x)}{u(a)} a$, ce qui montre que F^\perp est engendré par a et est donc de dimension un.

Maintenant, choisissons $a \in F^\perp$ tel que $u(a) = \langle a, a \rangle = \|a\|^2$. On prétend que $u(x) = \langle x, a \rangle$, pour tout $x \in H$. En effet, si $x \in F$ les deux termes sont nuls, si $x = \lambda a$ pour un certain $\lambda \in \mathbf{C}$, les deux termes valent $\lambda u(a)$. Par linéarité, les deux termes sont égaux sur $F \oplus F^\perp = H$.

Nous savons maintenant que l'application $a \mapsto \Lambda_a$ est surjective de H dans H' . Elle est évidemment injective, car si $a \neq 0$, Λ_a est non nulle (par exemple $\Lambda_a(a) = \|a\|^2 \neq 0$). Enfin, nous avons déjà vérifié que cette application est une isométrie, i.e. $\|\Lambda_a\|_{H'} = \|a\|$. \square

Remarque 10.5. Ce résultat généralise donc le résultat classique qui, en dimension finie, affirme que, si u est une forme linéaire sur \mathbf{R}^N , il existe $y \in \mathbf{R}^N$ tel que $u(x) = x \cdot y$, où \cdot désigne le produit scalaire euclidien.

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 $\langle \cdot, \cdot \rangle_{H'}$ sur H' est défini en posant, pour tous $u, v \in H'$

$$\langle u, v \rangle_{H'} := \langle b, a \rangle,$$

où $a, b \in H$ sont définis par $u = \Lambda_a$ et $v = \Lambda_b$.

Remarque 10.6. Notons aussi que la notation des physiciens en **bra** et **kets**, introduite par P. Dirac, correspond exactement à cette identification de H et H' . En physique, et contrairement à notre convention, le produit hermitien $(x, y) \mapsto \langle x, y \rangle$ est linéaire par rapport à y et anti-linéaire par rapport à x . Le bra $|y\rangle$ correspond au vecteur ordinaire y et $\langle x|$ correspond à l'élément de H' défini par $x \mapsto \langle x, y \rangle$. Si T est une application linéaire de H dans H , $\langle x|T|y\rangle$ correspond donc à $\langle x, Ty \rangle$.

L'extension suivante du théorème de Riesz aux formes sesquilinéaires sera utile pour l'étude des équations de Sturm-Liouville.⁴

Définition 10.6 (Critère de continuité pour les formes sesquilinéaires). *On dit qu'une forme sesquilinéaire $\Phi : H \times H \rightarrow \mathbf{C}$ est continue sur H s'il existe une constante C telle que*

$$|\Phi(x, y)| \leq C \|x\| \|y\|,$$

pour tous $x, y \in H$.

Remarque 10.7. Cette définition est en réalité un théorème : la continuité au sens usuel de l'application de $\Phi : H \times H \rightarrow \mathbf{C}$ est équivalente à la continuité au sens de la définition précédente (preuve laissée en exercice).

On a alors :

Théorème 10.4 (Théorème de représentation de Riesz (version sesquilinéaire)). *Soit Φ une forme sesquilinéaire continue sur un espace de Hilbert H . Alors, il existe une unique application linéaire continue $A : H \rightarrow H$ telle que*

$$\Phi(x, y) = \langle x, A(y) \rangle,$$

pour tous $x, y \in H$.

DÉMONSTRATION. D'après l'hypothèse de continuité, pour chaque $y \in H$ fixé la forme linéaire $x \mapsto \Phi(x, y)$ est continue. Le Théorème de Riesz nous assure qu'il existe $a_y \in H$ unique tel que $\Phi(x, y) = \langle x, a_y \rangle$. Vérifions que l'application $A : H \rightarrow H$ définie par $A(y) := a_y$ est linéaire et continue.

4. Charles Sturm, Professeur à l'École polytechnique en 1840, Joseph Liouville, X 1825, Professeur à l'École polytechnique en 1838.

La linéarité résulte immédiatement de l'unicité dans le Théorème de représentation de Riesz et de la sesquilinearité de Φ . Pour la continuité, nous avons par hypothèse

$$\|A(y)\|^2 = \langle A(y), A(y) \rangle = \Phi(A(y), y) \leq C \|y\| \|A(y)\|.$$

On en déduit que

$$\|A(y)\| \leq C \|y\|,$$

et la continuité de A en découle. \square

Une conséquence du Théorème de représentation de Riesz est l'existence d'un *adjoint* pour les applications linéaires continues d'un espace de Hilbert dans lui même. Commençons par la :

Définition 10.7. Soit $A : H \rightarrow H$ une application linéaire continue sur un espace de Hilbert. On appelle adjoint de A et l'on note A^* , un endomorphisme continu $A^* : H \rightarrow H$ vérifiant

$$\langle A(x), y \rangle = \langle x, A^*(y) \rangle,$$

pour tous $x, y \in H$.

Remarquons que si A possède un adjoint, il résulte de l'injectivité de l'application $a \mapsto \Lambda_a$ que cet adjoint est unique. En ce qui concerne l'existence de l'adjoint d'une application linéaire continue, nous avons le :

Corollaire 10.2. Soit A une application linéaire continue d'un espace de Hilbert H dans lui même. Alors, l'adjoint de A est bien défini et c'est encore une application linéaire continue.

DÉMONSTRATION. La forme sesquilinéaire $\Phi(x, y) := \langle A(x), y \rangle$ est continue, vu que

$$|\langle A(x), y \rangle| \leq \|A(x)\| \|y\| \leq \|A\|_{\mathcal{L}(H, H)} \|x\| \|y\|.$$

D'après le Théorème de représentation de Riesz, il existe donc une unique application linéaire continue A^* telle que $\langle A(x), y \rangle = \langle x, A^*(y) \rangle$. \square

Définition 10.8. Soit $A : H \rightarrow H$ une application linéaire continue définie sur un espace de Hilbert complexe H . On dit que A est un opérateur hermitien (ou auto-adjoint ou encore, dans le cadre des espaces de Hilbert réels, opérateur symétrique) si $A = A^*$. Autrement dit, si

$$\langle A(x), y \rangle = \langle x, A(y) \rangle,$$

pour tous $x, y \in H$.

Exemple 10.15. L'opérateur de projection orthogonale sur un sous-espace vectoriel fermé $F \subset H$ est un opérateur hermitien.

Nous donnons maintenant une autre application du Théorème de représentation de Riesz. Nous aurons besoin de la :

Définition 10.9. On dit qu'une forme sesquilinéaire Φ définie sur un espace de Hilbert H est coercive, s'il existe une constante $c > 0$ telle que $\Phi(x, x) \geq c \|x\|^2$.

Proposition 10.6. On suppose que Φ est une forme hermitienne continue et coercive, alors l'application A définie dans le Théorème 10.4 page précédente est inversible et d'inverse continu. De plus $A = A^*$ (on dit alors que A est une application hermitienne).

DÉMONSTRATION. Utilisons la coercivité de Φ . Par construction de A on a

$$\langle x, A(x) \rangle = \Phi(x, x) \geq c \|x\|^2,$$

par conséquent, si $A(x) = 0$ alors $x = 0$, ce qui montre que A est injective.

Montrons que A est surjective. Pour cela remarquons que $\mathcal{N} := \sqrt{\Phi}$ définit une norme sur H qui est équivalente à la norme $\| \cdot \|$. En effet, la continuité de Φ nous assure l'existence d'une constante $C > 0$ telle que

$$\Phi(x, x) \leq C \|x\|^2,$$

et la coercivité de Φ nous assure l'existence d'une constante $c > 0$ telle que

$$\Phi(x, x) \geq c \|x\|^2,$$

ce qui démontre que les deux normes sont équivalentes. En particulier, l'espace H muni du produit hermitien Φ et de la norme \mathcal{N} est un espace de Hilbert (c'est pour démontrer la complétude de cet espace que nous utilisons le fait que les deux normes sont équivalentes, donc que toute suite de Cauchy pour une norme est une suite de Cauchy pour l'autre).

Soit $z \in H$. Définissons la forme linéaire continue $\Lambda \in H'$ par

$$\Lambda(x) := \langle x, z \rangle.$$

D'après le Théorème de représentation de Riesz (appliqué à la forme linéaire Λ dans l'espace de Hilbert H muni du produit hermitien Φ), il existe un unique $y \in H$ tel que

$$\Phi(x, y) = \langle x, z \rangle,$$

pour tout $x \in H$. Revenons à la définition de A pour conclure que

$$\langle x, A(y) \rangle = \langle x, z \rangle,$$

pour tout $x \in H$. Donc $A(y) = z$, ce qui démontre que A est surjective.

Pour démontrer que A^{-1} est continue, il suffit de remarquer que la coercivité et la continuité de Φ nous permettent d'écrire que

$$c \|x\|^2 \leq \langle x, A(x) \rangle \leq C \|A(x)\| \|x\|,$$

donc $c \|x\| \leq C \|A(x)\|$, ce qui nous permet de conclure que $c \|A^{-1}(y)\| \leq \|y\|$ en prenant $x = A^{-1}(y)$. Ceci démontre la continuité de A^{-1} .

Enfin, par définition

$$\Phi(x, y) = \langle x, y \rangle = \langle A^*(x), y \rangle,$$

pour tous $x, y \in H$. D'autre part, Φ étant supposée hermitienne, on peut écrire

$$\Phi(x, y) = \overline{\Phi(y, x)} = \overline{\langle y, A(x) \rangle} = \langle A(x), y \rangle.$$

Donc $\langle A^*(x), y \rangle = \langle A(x), y \rangle$, pour tous $x, y \in H$ et l'on conclut que $A = A^*$. \square

Terminons ce paragraphe par l'énoncé du Théorème de Hahn-Banach, ce qui nous donnera un critère très simple de densité d'un sous-espace vectoriel. Ce critère nous sera bien utile dans la section suivante.

Théorème 10.5 (Théorème de Hahn-Banach dans un espace de Hilbert). *Soit F un sous-espace d'un espace de Hilbert H . On suppose qu'il existe $x \notin \overline{F}$. Alors, il existe une forme linéaire continue $\Lambda \in H'$ telle que $\Lambda \equiv 0$ sur F et $\Lambda(x) = 1$.*

DÉMONSTRATION. On note $G := F^\perp = \overline{F}^\perp$, qui est un sous-espace fermé de H . Grâce au Théorème de la projection (voir le Théorème 10.2 page 197), on a la décomposition

$$x = P_G(x) + (x - P_G(x)),$$

où $x - P_G(x) \in G^\perp = \overline{F}$ (voir le Corollaire 10.1 page 197). Étant donné que $P_G(x) \neq 0$ (sinon on aurait $x \in \overline{F}$), on peut définir

$$y := \frac{P_G(x)}{\|P_G(x)\|^2} \in F^\perp = \overline{F}^\perp,$$

de sorte que $\Lambda_y(x) = \langle x, y \rangle = 1$ et $\Lambda_y(z) = \langle z, y \rangle = 0$ pour $z \in \overline{F}$. Il suffit de prendre $\Lambda := \Lambda_y$. \square

Le résultat suivant permet de démontrer facilement des résultats de densité, et il sera particulièrement utile pour montrer qu'une suite de vecteurs forme une base hilbertienne.

Corollaire 10.3 (Critère de densité). *Soit F un sous-espace vectoriel de H . Alors, F est dense dans H si et seulement si il n'existe pas d'élément $a \neq 0$ de H tel que $\langle x, a \rangle = 0$ pour tout $x \in F$, autrement dit, si et seulement si $F^\perp = \{0\}$.*

Exemple 10.16. Soit $\ell_c(\mathbf{N})$ l'ensemble des suites de $\ell^2(\mathbf{N})$ qui sont nulles à partir d'un certain rang, et \mathbf{e}_n la suite dont tous les termes sont nuls sauf le n -ième qui est égal à 1. Si $\mathbf{a} := (a_n)_{n \geq 0}$ est orthogonal à $\ell_c(\mathbf{N})$, on a $\langle \mathbf{a}, \mathbf{e}_n \rangle_{\ell^2} = a_n = 0$, donc $\mathbf{a} = 0$. On conclut que $\ell_c(\mathbf{N})$ est dense dans $\ell^2(\mathbf{N})$.

5. Bases hilbertiennes

Soit H un espace de Hilbert de dimension infinie et $(e_n)_{n \geq 0}$ une famille dénombrable de vecteurs de H .

Définition 10.10. On dit que $(e_n)_{n \geq 0}$ est une base hilbertienne de H si :

- (i) pour tous $n \neq m$, $\langle e_n, e_m \rangle = 0$ et $\|e_n\| = 1$ pour tout $n \in \mathbf{N}$;
- (ii) l'espace vectoriel $\text{Vect} \{e_n : n \in \mathbf{N}\}$ des combinaisons linéaires finies des vecteurs e_n , pour $n \in \mathbf{N}$, est dense dans H

Attention, une base hilbertienne n'est pas nécessairement une base algébrique car pour une base algébrique, tout élément de l'espace est combinaison linéaire finie d'éléments de la base alors qu'en dimension infinie, tout élément de l'espace n'est pas forcément combinaison linéaire des éléments d'une base hilbertienne de l'espace. Nous précisons justement la notion de base hilbertienne dans le résultat qui suit.

Théorème 10.6. Soit $(e_n)_{n \geq 0}$ une base hilbertienne d'un espace de Hilbert H . Alors, tout élément $x \in H$ s'écrit de manière unique comme la somme d'une série convergente dans H

$$x = \sum_{n \geq 0} x_n e_n,$$

où $x_n = \langle x, e_n \rangle \in \mathbf{C}$. De plus, on a l'égalité de Parseval

$$(10.3) \quad \|x\|^2 = \sum_{n \geq 0} |x_n|^2 = \sum_{n \geq 0} |\langle x, e_n \rangle|^2.$$

Réciproquement, si $\sum_{n \geq 0} |x_n|^2 < +\infty$, alors la série $\sum_{n \geq 0} x_n e_n$ converge dans H .

DÉMONSTRATION. Soit F le sous-espace de H constitué des éléments qui s'écrivent comme combinaison linéaire finie des e_n ,

$$F := \text{Vect}\{e_n : n \in \mathbf{N}\}.$$

La décomposition et l'égalité de Parseval sont vraies pour les éléments de F , puisqu'elles le sont dans $F_n := \text{Vect}\{e_0, \dots, e_n\}$ et que tout élément de F est contenu dans un tel sous-espace.

Montrons maintenant l'existence de la décomposition. Soit P_n la projection orthogonale sur le sous-espace de dimension finie F_n . Clairement, $\lim_{n \rightarrow +\infty} P_n(x) = x$, puisque $P_n(x)$ est le point de F_n le plus proche de x et que $F = \bigcup_{n \in \mathbf{N}} F_n$ est dense dans H . On définit alors $x_n \in \mathbf{C}$ par $x_n := \langle x, e_n \rangle$, de telle sorte que

$$P_n(x) = \sum_{k=0}^n x_k e_k.$$

Étant donné que les e_k sont deux à deux orthogonaux, on a

$$\|P_n(x)\|^2 = \sum_{k=0}^n |x_k|^2,$$

grâce au Théorème de Pythagore. Étant donné que la suite $(P_n(x))_{n \geq 0}$ converge vers x et que la norme d'une limite est la limite de la norme (remarquer que l'application $x \mapsto \|x\|$ est continue), on obtient la formule de Parseval

$$\|x\|^2 = \sum_{k \geq 0} |x_k|^2.$$

L'unicité de la décomposition est évidente.

Enfin, pour montrer qu'une série $\sum_{n \geq 0} x_n e_n$ converge dans H si $\sum_{n \geq 0} |x_n|^2 < +\infty$, il suffit de remarquer que

$$\left\| \sum_{k=m}^n x_k e_k \right\|^2 = \sum_{k=m}^n |x_k|^2,$$

d'après le Théorème de Pythagore. En particulier, la suite $\left(\sum_{k=1}^n x_k e_k \right)_{n \geq 0}$ est une suite de Cauchy dans H , donc elle converge. □

Remarque 10.8. Remarquons que la série

$$\sum_{n \geq 0} x_n e_n,$$

n'est pas en général une série *absolument convergente*, en effet, $\sum_{n \geq 0} |x_n|^2$ converge, mais il n'y

a pas de raison que $\sum_{n \geq 0} |x_n|$ converge.

5.1. Orthonormalisation de Schmidt, familles totales, inégalité de Bessel. En dimension finie, le procédé d'*orthonormalisation de Schmidt* permet, à partir d'une famille de vecteurs engendrant l'espace, de construire une base orthonormée. Ce procédé s'applique aussi dans un espace de Hilbert. Plus précisément soit $(f_n)_{n \geq 0}$ une famille de vecteurs que l'on peut supposer, quitte à en oublier éventuellement certains, libre. On note $F_n := \text{Vect}\{f_0, \dots, f_n\}$ l'espace vectoriel engendré par les vecteurs f_0, \dots, f_n , et l'on suppose avoir déjà construit (e_0, \dots, e_n) , une base orthonormée de F_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 $(e_n)_{n \geq 0}$ est donc telle que toute somme finie des f_n s'écrit comme somme finie des e_n . Si le sous-espace vectoriel engendré par les f_n est dense, on aura la densité des combinaisons linéaires finies des e_n , et donc on conclut que $(e_n)_{n \geq 0}$ forme une base hilbertienne.

Soit $(e_n)_{n \geq 0}$ une suite dénombrable de vecteurs de H . On dit que c'est une *famille totale* si l'espace engendré par les e_n est dense dans H . Un système orthonormé forme une base hilbertienne si et seulement s'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 10.1. *Soit $(e_n)_{n \geq 0}$ une famille orthonormale de vecteurs de H . Alors*

$$\sum_{n=0}^{+\infty} \langle x, e_n \rangle e_n,$$

est la projection orthogonale de x sur \overline{F} , l'adhérence du sous-espace vectoriel F engendré par les e_n .

DÉMONSTRATION. On note F_n le sous-espace vectoriel de H engendré par les vecteurs e_0, \dots, e_n . Nous avons vu que $P_{F_n}(x)$, la projection orthogonale de x sur F_n , est donnée par

$$P_{F_n}(x) = \sum_{k=0}^n \langle x, e_k \rangle e_k,$$

et de plus

$$\|x\|^2 = \|P_{F_n}(x)\|^2 + \|P_{F_n}(x) - x\|^2 \geq \|P_{F_n}(x)\|^2,$$

d'où

$$\|P_{F_n}(x)\|^2 = \sum_{k=0}^n |\langle x, e_k \rangle|^2 \leq \|x\|^2.$$

On déduit de cette inégalité que la série $\sum_{k \geq 0} |\langle x, e_k \rangle|^2$ converge et par conséquent que la série

$$\sum_{k \geq 0} \langle x, e_k \rangle e_k,$$

est convergente dans H et qu'elle définit un élément $y \in \overline{F}$. Enfin, $\langle x - y, e_n \rangle = 0$ pour tout $n \in \mathbf{N}$. Donc, par unicité de la décomposition, $y = P_{\overline{F}}(x)$. \square

Remarquons que le Théorème de Pythagore nous permet d'écrire

$$\|x\|^2 = \|x - P_{\overline{F}}(x)\|^2 + \|P_{\overline{F}}(x)\|^2 = \|x - P_{\overline{F}}(x)\|^2 + \sum_{n \geq 0} |\langle x, e_n \rangle|^2.$$

On en déduit d'une part l'inégalité de Bessel

$$\sum_{n \geq 0} |\langle x, e_n \rangle|^2 \leq \|x\|^2,$$

et d'autre part que l'égalité de Parseval équivaut à $x = P_{\overline{F}}(x)$ quel que soit $x \in H$ et donc que F est dense dans H ou encore que la famille $(e_n)_{n \geq 0}$ forme une base hilbertienne.

Exemple 10.17 (Séries de Fourier dans $L^2(S^1; \mathbf{C})$). Considérons l'espace $H := L^2(S^1; \mathbf{C})$ que l'on identifie à l'espace des fonctions 2π -périodiques dont la restriction à $[0, 2\pi]$ est dans $L^2([0, 2\pi]; \mathbf{C})$. On munit cet espace du produit hermitien

$$\langle f, g \rangle_{L^2} := \frac{1}{2\pi} \int_{S^1} f(t) \overline{g(t)} dt.$$

On vérifie que la famille $(e_n)_{n \in \mathbf{Z}}$ où $e_n(t) := e^{int}$, est une famille orthonormée de H , en effet $\langle e_n, e_m \rangle_{L^2} = \delta_{m,n}$, pour tous $m, n \in \mathbf{Z}$. Dans ce cas,

$$\langle f, e_n \rangle_{L^2} = \frac{1}{2\pi} \int_{S^1} f(t) e^{-int} dt =: c_n(f),$$

est le n -ième coefficient de Fourier de la fonction f . On montre que la fermeture de l'espace vectoriel engendré par les combinaisons finies des e_n est dense dans $L^2(S^1; \mathbf{C})$. Pour ce faire, on peut par exemple utiliser le Théorème de Stone-Weierstrass qui nous assure que les combinaisons linéaires des e_n sont denses dans $\mathcal{C}([0, 2\pi]; \mathbf{C})$ (muni de la norme de la convergence uniforme), espace qui est lui-même dense dans $L^2([0, 2\pi]; \mathbf{C})$ (muni de la norme L^2) – voir le Théorème 8.4 page 164. Il en résulte que la série de Fourier de f converge pour la norme L^2 vers f , et que l'égalité de Parseval est dans ce cas donnée par

$$\sum_{n \in \mathbf{Z}} |c_n(f)|^2 = \frac{1}{2\pi} \int_{S^1} |f(t)|^2 dt.$$

Inversement si $(a_n)_{n \in \mathbf{Z}}$ est une suite de nombres complexes telle que

$$\sum_{n \in \mathbf{Z}} |a_n|^2 < +\infty,$$

la série de Fourier $\sum_{n \in \mathbf{Z}} a_n e^{int}$ converge dans $L^2(S^1; \mathbf{C})$.

Exemple 10.18. Soit $w \in \mathcal{C}([a, b])$ une fonction continue, à valeurs réelles, qui est strictement positive sur l'intervalle borné $[a, b]$ et $H = L^2([a, b]; \mathbf{C})$ l'espace des fonctions L^2 à valeurs complexes. On définit sur H le produit hermitien

$$\langle f, g \rangle_{L_w^2} := \int_a^b f(t) \overline{g(t)} w(t) dt.$$

Le Théorème de Stone-Weierstrass montre que les fonctions polynômes à coefficients complexes forment un sous-espace dense de $\mathcal{C}([a, b]; \mathbf{C})$ pour la norme de la convergence uniforme qui est plus forte que la norme L_w^2 . Le sous-espace $\mathcal{C}([a, b]; \mathbf{C})$ étant dense dans $L^2([a, b]; \mathbf{C})$, on conclut que les fonctions polynômes forment un sous-espace dense de $L^2([a, b]; \mathbf{C})$. Le procédé d'orthonormalisation de Schmidt de la base (non orthogonale) $(1, t, t^2, \dots, t^n, \dots)$ fournit alors des familles de polynômes remarquables qui portent le nom des mathématiciens qui les ont définies.

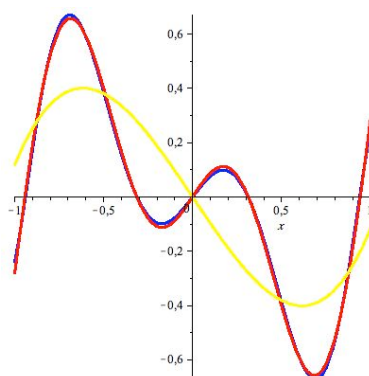


FIGURE 10.1. Décomposition de la fonction $x \mapsto \cos(5x)\exp(-x)$ (en rouge) en polynômes de Legendre (en jaune 5 termes, en bleu 9 termes).

Exemple 10.19. Dans l'exemple ci-dessus, si $w \equiv 1$ et $a = -1$, $b = 1$, on obtient ainsi les *polynômes de Legendre*. Si $w(x) = (1 - x^2)^{1/2}$ et $a = -1$, $b = 1$, on obtient les *polynômes de Chebychev*. Si l'on prend $a = -\infty$, $b = +\infty$ et $w(x) := e^{-x^2}$, la densité des fonctions polynômes est encore vraie dans $L^2(\mathbf{R}, e^{-x^2} dx)$. Le procédé d'orthonormalisation de Schmidt appliqué à la famille $(1, x, x^2, \dots, x^n, \dots)$ fournit alors les *polynômes d'Hermite*. Si $a = 0$, $b = +\infty$ et $w(x) = e^{-x}$, on obtient les *polynômes de Laguerre*. On renvoie à la Proposition ci-dessous pour une démonstration du fait que les fonctions polynômes sont denses dans ces espaces.

Comme dans le cas des séries de Fourier, toute fonction de $L^2([-1, 1], \mathbf{C})$ admet 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.

Le fait que les fonctions polynômes sont denses dans $L^2(\mathbf{R}, e^{-x^2} dx)$ et dans $L^2([0, +\infty[, e^{-x} dx)$ est une conséquence du Corollaire 10.3.

Proposition 10.7. *Les fonctions polynômes sont denses dans $L^2(\mathbf{R}, e^{-x^2} dx)$. Elles sont aussi denses dans $L^2([0, +\infty[, e^{-x} dx)$.*

DÉMONSTRATION. Montrons que l'espace des fonctions polynômes forme un sous-espace dense de $L^2(\mathbf{R}, e^{-x^2} dx)$. Pour ce faire, en vertu du Corollaire 10.3, il suffit de démontrer qu'une fonction $f \in L^2(\mathbf{R}, e^{-x^2} dx)$ qui est orthogonale, pour le produit hermitien de $L^2(\mathbf{R}, e^{-x^2} dx)$, à toutes les fonctions polynômes est en fait la fonction nulle.

Soit f une telle fonction. Définissons pour tout $\xi \in \mathbf{R}$,

$$g(\xi) := \int_{\mathbf{R}} f(x) e^{-i\xi x} e^{-x^2} dx.$$

On peut écrire

$$g(\xi) := \sum_{n=0}^{\infty} \frac{(-i\xi)^n}{n!} \int_{\mathbf{R}} f(x) x^n e^{-x^2} dx,$$

et, par hypothèse $g(\xi) = 0$ pour tout $\xi \in \mathbf{R}$. Autrement dit, la transformée de Fourier de $x \mapsto f(x) e^{-x^2}$ est nulle. On conclut que $f = 0$, ce qui termine la démonstration. Nous laissons le soin au lecteur d'adapter cette démonstration pour démontrer la deuxième assertion. \square

5.2. Les ondelettes de Haar. La *Théorie des ondelettes* s'est fortement développée à partir du milieu des années 1980, suite aux travaux de Jean Morlet⁵ et Alex Grossmann et surtout suite aux travaux d'Yves Meyer, Stéphane Mallat⁶, Ingrid Daubechie, Stéphane Jaffard⁷, ... Les bases d'ondelettes, dont les premiers exemples datent du début du XX-ième siècle, sont des bases hilbertiennes qui sont particulièrement bien adaptées à l'analyse des images et au traitement du signal. La théorie des ondelettes trouve maintenant des applications dans de nombreux domaines, citons leur utilisation dans les algorithmes de compression d'images comme par exemple dans le standard JPEG 2000.

Dans ce cours, nous nous contentons de définir les ondelettes de Haar (1909), nous renvoyons le lecteur à l'ouvrage d'I. Daubechies⁸ ou au cours de S. Mallat⁹ pour plus de détails sur les bases d'ondelettes et leur utilisation.

On note

$$\varphi := \mathbf{1}_{[0,1/2[} - \mathbf{1}_{[1/2,1[}.$$

Pour tout $n \geq 0$ et pour tout $k = 0, \dots, 2^n - 1$, on note

$$\varphi_{n,k}(x) := 2^{n/2} \varphi(2^n x - k).$$

Par définition, $\varphi_{0,0} = \varphi$ et pose $\mathbf{1} := \mathbf{1}_{[0,1]}$. Nous allons montrer que la fonction $\mathbf{1}$ et les fonctions $\varphi_{n,k}$ pour $n \in \mathbf{N}$ et pour $0 \leq k < 2^n$, forment une base hilbertienne de $L^2([0, 1]; \mathbf{C})$.

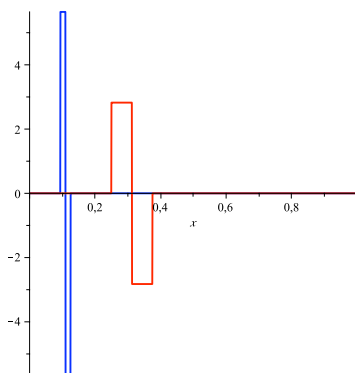


FIGURE 10.2. Les ondelettes de Haar $\varphi_{5,3}$ (en bleu) et $\varphi_{3,2}$ (en rouge).

Tout d'abord, il est facile de voir que $\langle \varphi_{n,k}, \varphi_{m,j} \rangle_{L^2} = 0$ si $(n, k) \neq (m, j)$ et vaut 1 sinon. Maintenant, d'après le Corollaire 10.3 page 202, pour démontrer que la famille est totale, il suffit de montrer que si $f \in L^2([0, 1]; \mathbf{C})$ est orthogonale à $\mathbf{1}$ et à toutes les fonctions $\varphi_{n,k}$, elle est nulle p.p. sur $[0, 1]$. Pour ce faire, on montre par récurrence sur $n \geq 0$ que

$$\int_{\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right]} f(x) dx = 0,$$

5. Jean Morlet X 1954.

6. Stéphane Mallat X 1981.

7. Stéphane Jaffard X 1981.

8. I. Daubechies, Ten lectures on Wavelets, CBMS regional conference series in Applied Mathematics, SIAM, Philadelphia, 1992.

9. S. Mallat, Une exploration des signaux en ondelettes, éditions de l'École Polytechnique, 2000.

pour $0 \leq k < 2^n$. Par exemple, $\langle f, \mathbf{1} \rangle_{L^2} = 0$, donc

$$\int_{[0,1]} f(t) dt = 0.$$

Ensuite, comme $\langle f, \varphi_{0,0} \rangle_{L^2} = 0$, on en déduit que

$$\int_{[0,1/2]} f(t) dt - \int_{[1/2,1]} f(t) dt = 0.$$

Et, en utilisant la première identité, on conclut que

$$\int_{[0,1/2]} f(t) dt = \int_{[1/2,1]} f(t) dt = 0.$$

De même, en utilisant le fait que $\langle f, \varphi_{1,0} \rangle_{L^2} = \langle f, \varphi_{1,1} \rangle_{L^2} = 0$ pour montrer que

$$\int_{[0,1/4]} f(t) dt - \int_{[1/4,1/2]} f(t) dt = \int_{[1/2,3/4]} f(t) dt - \int_{[3/4,1]} f(t) dt = 0,$$

ce qui, grâce aux identités ci-dessus, permet de montrer que

$$\int_{[0,1/4]} f(t) dt = \int_{[1/4,1/2]} f(t) dt = \int_{[1/2,3/4]} f(t) dt = \int_{[3/4,1]} f(t) dt = 0.$$

On poursuit en utilisant une récurrence sur $n \geq 0$ (exercice).

Nous pouvons maintenant conclure, car si $f \in L^2([0,1]; \mathbf{C})$, nous avons vu que la fonction

$$F(x) := \int_{[0,x]} f(t) dt,$$

est continue (voir l'exemple 10.13 page 194). Mais cette fonction prend la même valeur en tous les points de la forme $\frac{k}{2^m}$ pour $0 \leq k < 2^m$ et tout $m \in \mathbf{N}$. Comme ces points forment un sous-ensemble dense de $[0,1]$ on conclut que F est une fonction constante sur $[0,1]$ donc identiquement nulle (car elle est nulle en $x = 0$).

En particulier, f est orthogonale à toutes les fonctions indicatrices des sous-intervalles de $[0,1]$. Comme les combinaisons linéaires de ces fonctions sont denses dans $L^2([0,1]; \mathbf{C})$ (voir l'exercice 10.2 page suivante), on conclut que $f = 0$ p.p. sur $[0,1]$.

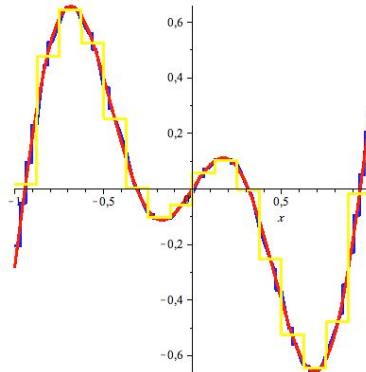


FIGURE 10.3. Décomposition en ondelettes de Haar de la fonction $x \cos(5x)$ sur $[0,1]$ avec 16 et 64 coefficients.

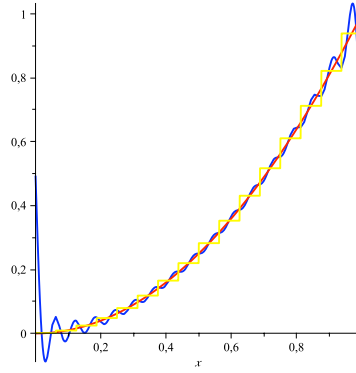


FIGURE 10.4. Décomposition en Fourier (en bleu) et en ondelettes de Haar (en jaune) de x^2 (en rouge) sur $[0, 1]$ avec 16 coefficients.

Exercice 10.2. Soit $p \geq 1$. Montrer que l'espace des fonctions en escalier s'identifie à un sous espace dense de $L^p(\mathbf{R})$.

6. Espaces de Hilbert séparables

Définition 10.11. On dit qu'un espace de Hilbert H est séparable, s'il existe un sous-ensemble de H qui est à la fois dénombrable et dense.

Théorème 10.7. Tout espace de Hilbert séparable possède une base hilbertienne.

DÉMONSTRATION. C'est essentiellement le procédé d'orthonormalisation de Schmidt appliqué à une suite de vecteurs formant un sous-ensemble dense. Plus précisément, si $(x_n)_{n \geq 0}$ est une suite dense dans H , on note $F_n := \text{Vect}\{x_0, \dots, x_n\}$ le sous-espace vectoriel engendré par x_0, \dots, x_n . Les F_n forment une suite croissante (au sens large) de sous-espaces vectoriels de H 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 qui soit distinct de F_n . On aura alors $\dim(F_n) = n$.

Choisissons un vecteur unitaire e_0 dans F_0 et, pour tout $n \geq 1$, choisissons un vecteur unitaire $e_n \in F_{n-1}^\perp \cap F_n$. Alors $(e_n)_{n \geq 0}$ est une base hilbertienne, puisque c'est une famille orthonormale, et que les combinaisons linéaires finies des e_n engendrent $\bigcup_{n \geq 0} F_n$ qui, par hypothèse, est dense dans H . \square

Remarque 10.9. On remarque que, si H possède une base hilbertienne, alors H est séparable. En effet, si $(e_n)_{n \geq 0}$ est une base hilbertienne de H , le \mathbf{Q} -espace vectoriel engendré par les vecteurs e_n

$$\left\{ \sum_{j=0}^n a_j e_j : n \geq 0, \quad a_j \in \mathbf{Q}, \quad j = 0, \dots, n \right\},$$

est dénombrable, comme réunion dénombrable d'ensembles dénombrables, et dense dans H .

Enfin, remarquons que dans un espace de Hilbert séparable, une famille orthonormée de vecteurs $(f_n)_{n \geq 0}$, 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 10.1 page 197 on a une décomposition de l'espace H en $H = F \oplus F^\perp$. En prenant

une base hilbertienne $(g_m)_{m \geq 0}$ de F^\perp et en la juxtaposant à $(f_n)_{n \geq 0}$ on obtient une base hilbertienne de H .

La plupart des espaces de Hilbert usuels sont séparables. Par exemple, $\ell^2(\mathbf{N}; \mathbf{C})$, $L^2(\mathbf{R}; \mathbf{C})$, $L^2([a, b]; \mathbf{C})$, ... sont tous des espaces de Hilbert séparables. C'est clair pour $\ell^2(\mathbf{N}; \mathbf{C})$ qui possède la base hilbertienne donnée par $(e_n)_{n \geq 1}$ où e_n est la suite identiquement nulle à l'exception du n -ième terme, égal à 1. De même pour $L^2([0, 2\pi]; \mathbf{C})$ qui possède une base hilbertienne donnée par les fonctions $e_n(t) = e^{int}$ pour $n \in \mathbf{Z}$. La situation est un peu plus délicate pour $L^2(\mathbf{R}; \mathbf{C})$ (voir l'exercice 10.3 page suivante). Tous les espaces de Hilbert qui seront considérés dans ce cours sont séparables.

Corollaire 10.4. *Deux espaces de Hilbert séparables, de dimension infinie, sont isomorphes.*

DÉMONSTRATION. On va montrer que H est isomorphe à $\ell^2(\mathbf{N})$. Soit $(e_n)_{n \geq 0}$ une base hilbertienne de H . À un élément $\mathbf{x} = (x_n)_{n \geq 0}$ de $\ell^2(\mathbf{N})$ on associe

$$u(\mathbf{x}) := \sum_{n \geq 0} x_n e_n,$$

qui est un élément de H . L'application $u : \ell^2(\mathbf{N}) \rightarrow H$ est clairement linéaire et elle vérifie $\|u(\mathbf{x})\| = \|\mathbf{x}\|$ d'après l'égalité de Parseval. Elle est donc continue et injective. Montrons qu'elle est surjective, d'inverse continu.

En reprenant la définition d'une base hilbertienne, on voit que $\mathbf{y} \in H$ s'écrit de manière unique comme

$$y = \sum_{n=0}^{\infty} x_n e_n,$$

où $\mathbf{x} := (x_n)_{n \geq 0}$ est dans $\ell^2(\mathbf{N})$. Ceci prouve que $y = u(\mathbf{x})$, d'où la surjectivité de u . L'égalité des normes de \mathbf{x} et y entraîne alors la continuité de l'application réciproque. \square

Remarquons 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, $(u(f_n))_{n \geq 0}$ est une base hilbertienne de H .

Proposition 10.8. *Un sous-espace fermé d'un espace de Hilbert séparable est séparable.*

DÉMONSTRATION. Soit $X = \{x_n : n \geq 0\}$ un sous-ensemble dénombrable et dense de H . Vérifions que $P_F(X)$, qui est clairement dénombrable, est bien dense dans F . En effet, si $y \in F$, pour tout $\varepsilon > 0$, il existe $n \in \mathbf{N}$ tel que $\|y - x_n\| \leq \varepsilon$. Mais, l'application P_F étant 1-lipschitzienne et égale à l'identité sur F , on peut écrire

$$\|P_F(x_n) - y\| = \|P_F(x_n) - P_F(y)\| \leq \|x_n - y\| \leq \varepsilon,$$

ce qui montre la densité de $P_F(X)$ dans F . \square

Remarque 10.10. Soit $(e_i)_{i \in I}$ une famille orthonormée de vecteurs d'un espace de Hilbert séparable H . Alors I est nécessairement dénombrable. En effet, pour chaque vecteur $x \in H$, l'inégalité de Bessel nous assure que

$$\sum_{i \in J} |\langle x, e_i \rangle|^2 \leq \|x\|^2,$$

pour tout sous-ensemble $J \subset I$. Alors, pour tout $n \geq 0$, l'ensemble $\{i \in I : |\langle x, e_i \rangle| \geq 1/n\}$ est fini et l'on en déduit que $\{i \in I : \langle x, e_i \rangle \neq 0\} = \bigcup_{n \in \mathbf{N} - \{0\}} \{i \in I : |\langle x, e_i \rangle| \geq 1/n\}$ est

dénombrable. Donc, pour chaque $x \in H$, seul un nombre dénombrable des $\langle x, e_i \rangle$ est non nul. Mais alors, si X est un sous-ensemble dénombrable dense de H , l'ensemble $\{i \in I : \exists x \in X, \langle x, e_i \rangle \neq 0\}$ est dénombrable comme réunion dénombrable d'ensembles dénombrables. De plus

$$\{i \in I : \exists x \in X, \langle x, e_i \rangle \neq 0\} \subset \{i \in I : \exists x \in H, \langle x, e_i \rangle \neq 0\} \subset I.$$

Par densité de X dans H , on conclut qu'il y a égalité entre ces deux ensembles, donc l'ensemble $\{i \in I : \exists x \in H, \langle x, e_i \rangle \neq 0\}$ est dénombrable. Enfin, comme $\langle e_i, e_i \rangle = 1 \neq 0$, on en déduit que

$$I = \{i \in I : \exists x \in H, \langle x, e_i \rangle \neq 0\},$$

et finalement que I est lui-même dénombrable.

Exercice 10.3. Montrer que $L^2(\mathbf{R}; \mathbf{C})$ est séparable¹⁰. 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.

7. Opérateurs compacts

Dans l'étude des espaces de Hilbert de dimension quelconque, il y a une classe d'opérateurs particulièrement importante. Il s'agit des opérateurs compacts.

Définition 10.12 (Opérateurs compacts). *On dit qu'un opérateur $A : H \rightarrow H$ est compact si l'image par A de la boule unité fermée $B_f(0, 1)$ de H est contenu dans un sous-ensemble compact de H . Autrement dit, un opérateur A est compact si, pour toute suite bornée $(x_n)_{n \geq 0}$ d'éléments de H , la suite $(A(x_n))_{n \geq 0}$ admet une sous-suite qui converge.*

On remarque qu'un opérateur compact est automatiquement borné (exercice). Donnons maintenant quelques exemples d'opérateurs compacts.

Exemple 10.20. Un opérateur borné, de rang fini est un opérateur compact.

Exemple 10.21. D'après le Théorème de Riesz 10.8 page 213, l'opérateur de projection orthogonale sur un sous-espace vectoriel fermé $F \subset H$ est un opérateur compact si et seulement si F est de dimension finie, vu que $P_F(B_f(0, 1)) = B_f(0, 1) \cap F$.

Exemple 10.22. Soient B est une application linéaire continue de H dans H et A une application linéaire compacte, alors $B \circ A$ et $A \circ B$ sont des opérateurs compacts. Une combinaison linéaire d'opérateurs compacts est un opérateur compact.

¹⁰. On pourrait bien entendu utiliser l'existence d'une base hilbertienne donnée par les ondelettes de Haar, c.f. 5.2 page 207

Exemple 10.23. L'application $f \mapsto I(f)$ définie sur $L^2([0, 1]; \mathbf{C})$ dans lui-même par

$$I(f)(x) := \int_{[0, x]} f(t) dt,$$

est compacte.

Nous avons déjà vu (voir l'exemple 10.13 page 194) que

$$|I(f)(y) - I(f)(x)| \leq \|f\|_{L^2} |y - x|^{1/2}.$$

Par conséquent, l'image par I d'une suite bornée de $L^2([0, 1]; \mathbf{C})$ est une suite de fonctions continues, qui est bornée dans $\mathcal{C}([0, 1]; \mathbf{C})$ et équicontinue. Le Théorème d'Ascoli nous permet d'extraire de cette suite, une sous suite qui converge dans $\mathcal{C}([0, 1]; \mathbf{C})$ pour la norme de la convergence uniforme, donc qui converge dans $L^2([0, 1]; \mathbf{C})$ pour la norme L^2 .

La résultat qui suit permet de démontrer facilement que certains opérateurs sont compacts.

Proposition 10.9. *Une limite (pour la norme des applications linéaires continues) d'une suite d'opérateurs compacts est un opérateur compact. En particulier, la limite d'une suite d'opérateurs de rang fini est compact.*

DÉMONSTRATION. Soit $(A_n)_{n \geq 0}$ une suite d'opérateurs compacts qui converge pour la norme des applications linéaires continues, vers A et soit $(x_n)_{n \geq 0}$ une suite d'éléments de $B_f(0, 1)$. On va extraire de la suite $(x_n)_{n \geq 0}$ une sous-suite $(x_{\varphi(n)})_{n \geq 0}$ telle que $(A(x_{\varphi(n)}))_{n \geq 0}$ est une suite de Cauchy. Comme nous sommes dans un espace de Hilbert, cette suite converge et le résultat sera démontré.

Étant donné que A_0 est compact, on peut extraire de la suite $(x_n)_{n \geq 0}$, une sous-suite $(x_n^{(0)})_{n \geq 0}$ telle que la suite $(A_0(x_n^{(0)}))_{n \geq 0}$ converge. Ensuite, A_1 étant compact, on peut extraire de la suite $(x_n^{(0)})_{n \geq 0}$, une sous-suite $(x_n^{(1)})_{n \geq 0}$ telle que la suite $(A_1(x_n^{(1)}))_{n \geq 0}$ converge. On poursuit ainsi par récurrence et, en utilisant le fait que A_k est compact, on extrait de la suite $(x_n^{(k-1)})_{n \geq 0}$, une sous-suite $(x_n^{(k)})_{n \geq 0}$ telle que la suite $(A_k(x_n^{(k)}))_{n \geq 0}$ converge.

Définissons enfin la suite extraite diagonale $(x_{\varphi(n)})_{n \geq 0}$ où $x_{\varphi(n)} = x_n^{(n)}$. On prétend que $(A(x_{\varphi(n)}))_{n \geq 0}$ est une suite de Cauchy. En effet, on peut écrire pour tous $k, m, n \in \mathbf{N}$

$$\begin{aligned} \|A(x_{\varphi(m)}) - A(x_{\varphi(n)})\| &\leq \|A(x_{\varphi(m)}) - A_k(x_{\varphi(m)})\| + \|A_k(x_{\varphi(m)}) - A_k(x_{\varphi(n)})\| \\ &\quad + \|A_k(x_{\varphi(n)}) - A(x_{\varphi(n)})\|. \end{aligned}$$

Pour tout $\varepsilon > 0$, choisissons $k \geq 0$ tel que $\|A - A_k\| \leq \varepsilon/3$. Ensuite, étant donné que la suite $(A_k(x_{\varphi(n)}))_{n \geq 0}$ converge, il existe $n_* \geq 0$ tel que pour tous $m, n \geq n_*$, $\|A_k(x_{\varphi(m)}) - A_k(x_{\varphi(n)})\| \leq \varepsilon/3$. Conclusion, pour tous $m, n \geq n_*$, on a $\|A(x_{\varphi(m)}) - A(x_{\varphi(n)})\| \leq \varepsilon$. Ce qui termine la démonstration. \square

Inversement, on montre que tout opérateur compact défini sur un espace de Hilbert est limite d'une suite d'opérateurs de rang fini. C'est l'objet de l'exercice ci-dessous.

Exercice 10.4. Soit C un sous-ensemble fermé et borné d'un espace de Hilbert H .

1) On suppose que C est compact. Montrer que, pour tout $\varepsilon > 0$, il existe un sous-espace de dimension finie $F \subset H$, tel que

$$\sup_{x \in C} \|x - P_F(x)\| = \sup_{x \in C} d(x, F) \leq \varepsilon.$$

2) Inversement, on suppose que, pour tout $\varepsilon > 0$, il existe un sous-espace de dimension finie F tel que pour tout $x \in C$, $d(x, F) \leq \varepsilon$. Montrer que C est compact.

- 3) Soit $A : H \rightarrow H$ un opérateur borné. Montrer que, A est compact si et seulement si A est limite d'une suite d'opérateurs bornés de rang fini.
- 4) Soit $A : H \rightarrow H$ un opérateur borné. On suppose que A est de rang fini, montrer que A^* , l'adjoint de A est de rang fini.
- 5) Montrer que, si $A : H \rightarrow H$ est un opérateur compact, alors A^* , l'adjoint de A , est aussi un opérateur compact.

8. Compacité et convergence faible

La compacité est une propriété particulièrement importante en analyse fonctionnelle. Dans les espaces vectoriels normés de dimension finie, nous avons vu que les fermés bornés sont compacts. Malheureusement, dans les espaces vectoriels de dimension infinie, qui sont omniprésents dans les applications, cette propriété n'est plus vraie.

Par exemple, nous disposons du Théorème de Riesz (voir le Théorème 3.8 page 47) qui nous assure que la boule unité d'un espace vectoriel normé de dimension infinie n'est jamais compacte. Dans le cadre des espaces de Hilbert, la démonstration de ce résultat est considérablement simplifiée par l'existence d'un produit hermitien.

Théorème 10.8 (Théorème de Riesz). *Soit H un espace de préhilbertien de dimension infinie, alors $B_f(0, 1)$, la boule unité fermée de H , n'est pas compacte.*

DÉMONSTRATION. Étant donné que H est de dimension infinie, on peut construire par récurrence $(e_n)_{n \geq 0}$, une famille orthonormale de vecteurs de H . Bien entendu chaque $e_n \in B_f(0, 1)$, de plus

$$\|e_n - e_m\|^2 = \|e_n\|^2 + \|e_m\|^2 = 2,$$

si $n \neq m$. On ne peut donc pas extraire, de la suite $(e_n)_{n \geq 0}$, une sous-suite qui converge. Conclusion, la boule unité fermée de H n'est pas compacte. \square

Il est très souvent souhaitable de récupérer une forme plus faible de compacité pour la boule unité d'un espace vectoriel normé de dimension infinie. Mais, pour cela, il faut *changer de topologie*. Nous introduisons alors la définition suivante :

Définition 10.13. *Dans un espace de Hilbert H . On dit que la suite $(x_n)_{n \geq 0}$ converge faiblement vers x dans H si*

$$\forall y \in H, \quad \lim_{n \rightarrow +\infty} \langle x_n, y \rangle = \langle x, y \rangle.$$

On notera $x_n \rightharpoonup x$ dans H .

Comme son nom l'indique, la convergence faible est une condition plus faible que la convergence ordinaire, souvent appelée *convergence forte* lorsque l'on veut éviter toute ambiguïté. En effet, l'inégalité de Cauchy-Schwarz nous indique que

$$|\langle x_n, y \rangle - \langle x, y \rangle| = |\langle x_n - x, y \rangle| \leq \|y\| \|x_n - x\|,$$

pour tout $y \in H$. Donc, si $\lim_{n \rightarrow +\infty} \|x - x_n\| = 0$, on en déduit que $\lim_{n \rightarrow +\infty} \langle x_n, y \rangle = \langle x, y \rangle$.

Exemple 10.24. Soit $(e_n)_{n \geq 0}$ une base hilbertienne d'un espace de Hilbert H . Alors, $e_n \rightharpoonup 0$ dans H . En effet, si $y \in H$, on sait, grâce à l'égalité de Parseval que la série des $|\langle x, e_n \rangle|^2$ converge, donc

$$\lim_{n \rightarrow +\infty} \langle x, e_n \rangle = 0.$$

Plus délicat, les ensembles fermés pour la convergence forte ne le sont pas nécessairement pour la convergence faible. Par exemple, dans un espace de Hilbert séparable, 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 $(e_n)_{n \geq 0}$ des vecteurs d'une base hilbertienne est une suite de vecteurs de norme 1 qui n'a pas de limite forte, mais dont la limite faible est 0. Une autre difficulté vient de ce que la convergence faible n'est pas une convergence associée à une distance, mais seulement à une topologie. Néanmoins, si une suite possède une limite faible, celle-ci est unique (exercice). Nous disposons de la :

Proposition 10.10. *Soit $(x_n)_{n \geq 0}$ une suite d'un espace de Hilbert H . On suppose que $x_n \rightharpoonup x$. Alors, on a l'inégalité*

$$(10.4) \quad \|x\| \leq \liminf_{n \rightarrow +\infty} \|x_n\|.$$

DÉMONSTRATION. Remarquons que

$$|\langle x_n, x \rangle| \leq \|x\| \|x_n\|,$$

grâce à l'inégalité de Cauchy-Schwarz. Par définition de la convergence faible, si $x_n \rightharpoonup x$ dans H , le membre de gauche tend vers $\|x\|^2$ quand n tend vers $+\infty$. On en déduit que

$$\|x\|^2 \leq \|x\| \liminf_{n \rightarrow +\infty} \|x_n\|,$$

ce qui termine la démonstration. \square

L'exemple de la suite $(e_n)_{n \geq 0}$ constituée des vecteurs d'une base hilbertienne d'un espace de Hilbert H , dont on a vu qu'elle converge faiblement vers 0, montre que l'inégalité (10.9) du théorème n'est pas, en général, une égalité.

Nous pouvons maintenant énoncer le :

Théorème 10.9 (Compacité faible de la boule unité fermée). *Soit H un espace de Hilbert. Alors, toute suite $(x_n)_{n \geq 0}$ d'éléments de H qui est bornée, possède une sous-suite qui converge faiblement dans H .*

DÉMONSTRATION. On utilise dans la démonstration le procédé d'extraction diagonale que nous avons déjà rencontré dans la preuve du Théorème 2.1 page 33. Soit $(x_n)_{n \geq 0}$ une suite bornée de H . Remarquons que, la suite $(x_n)_{n \geq 0}$ étant bornée, pour tout $k \in \mathbf{N}$, la suite $(\langle x_n, x_k \rangle)_{n \geq 0}$ est bornée, donc elle admet une sous-suite qui converge.

On commence par extraire, de la suite $(x_n)_{n \geq 0}$, une sous-suite $(x_{\varphi_0(n)})_{n \geq 0}$, telle que $(\langle x_0, x_{\varphi_0(n)} \rangle)_{n \geq 0}$ converge dans \mathbf{C} (ici φ_0 est une application strictement croissante de \mathbf{N} dans \mathbf{N}). Ensuite, on extrait de la sous-suite $(x_{\varphi_0(n)})_{n \geq 0}$, une nouvelle sous-suite $(x_{\varphi_1(n)})_{n \geq 0}$, telle que $(\langle x_1, x_{\varphi_1(n)} \rangle)_{n \geq 0}$ converge dans \mathbf{C} (ici φ_1 est une application strictement croissante de \mathbf{N} dans \mathbf{N}), ... on extrait de la sous-suite $(x_{\varphi_k(n)})_{n \geq 0}$, une nouvelle sous-suite $(x_{\varphi_{k+1}(n)})_{n \geq 0}$ telle que $(\langle x_{k+1}, x_{\varphi_{k+1}(n)} \rangle)_{n \geq 0}$ converge dans \mathbf{C} , ...

Par construction, pour tout $k \in \mathbf{N}$, la suite $(\langle x_k, x_{\varphi_n(n)} \rangle)_{n \geq 0}$ converge dans \mathbf{C} . Pour simplifier les notations, on note

$$\tilde{x}_n := x_{\varphi_n(n)},$$

les éléments de la suite extraite diagonale. Clairement, la suite $(\tilde{x}_n)_{n \geq 0}$ est une suite extraite de la suite $(x_n)_{n \geq 0}$. On note

$$E := \text{Vect} \{x_n : n \in \mathbf{N}\},$$

l'espace vectoriel engendré par les éléments de la suite $(x_n)_{n \geq 0}$ (i.e. les combinaisons linéaires finies des x_n). Par linéarité du produit hermitien par rapport à la première variable, on peut affirmer que, pour tout $y \in E$, la suite $(\langle y, \tilde{x}_n \rangle)_{n \geq 0}$ converge dans \mathbf{C} vers une limite (qui sera notée $A(y)$). On vérifie facilement que pour tout $y \in \overline{E}$, la suite $(\langle y, \tilde{x}_n \rangle)_{n \geq 0}$ converge aussi dans \mathbf{C} , vers une limite qui sera notée $A(y)$. Enfin, on vérifie que U est une application linéaire de \overline{E} dans \mathbf{C} .

La suite $(x_n)_{n \geq 0}$ étant bornée, il existe une constante $M > 0$ telle que $|\langle y, \tilde{x}_n \rangle| \leq M \|y\|$, donc

$$|A(y)| \leq M \|y\|.$$

Conclusion, nous avons démontré qu'il existe une forme linéaire continue A définie sur \overline{E} telle que

$$\lim_{n \rightarrow +\infty} \langle y, \tilde{x}_n \rangle = A(y),$$

pour tout $y \in \overline{E}$. L'espace \overline{E} , muni du produit hermitien $\langle \cdot, \cdot \rangle$, est un espace de Hilbert et le Théorème de représentation de Riesz nous permet d'affirmer qu'il existe (un unique) $x \in \overline{E}$ tel que $A(y) = \langle y, x \rangle$ pour tout $y \in \overline{E}$.

Enfin, le Corollaire 10.1 page 197 nous permet d'écrire $H = \overline{E} \oplus \overline{E}^\perp$ et, en utilisant cette décomposition, on conclut que

$$\lim_{n \rightarrow +\infty} \langle y, \tilde{x}_n \rangle = \langle y, x \rangle,$$

pour tout $y \in H$. Ce qui montre que $\tilde{x}_n \rightharpoonup x$ dans H . \square

Exemple 10.25. Soit $N \geq 1$. On peut munir $L^2([0, 1]; \mathbf{R}^N)$ du produit hermitien

$$\langle u, v \rangle := \int_{[0, 1]} u(t) \cdot v(t) dt,$$

où \cdot désigne le produit scalaire euclidien dans \mathbf{R}^N . Cet espace est un espace de Hilbert et le Théorème 10.9 appliqué à chaque composante d'une fonction de $L^2([0, 1]; \mathbf{R}^N)$ permet d'énoncer le résultat suivant : si $(u_n)_{n \geq 0}$ est une suite bornée de $L^2([0, 1]; \mathbf{R}^N)$, on peut extraire de cette suite une sous-suite qui converge faiblement, vers $u \in L^2([0, 1]; \mathbf{R}^N)$ i.e.

$$\lim_{n \rightarrow +\infty} \langle u_n, v \rangle = \langle u, v \rangle,$$

pour tout $v \in L^2([0, 1]; \mathbf{R}^N)$. De plus

$$\|u\|_{L^2} \leq \varliminf_{n \rightarrow +\infty} \|u_n\|_{L^2},$$

grâce à la Proposition 10.10 page ci-contre.

Exemple 10.26. Une suite $(x_n)_{n \geq 0}$ qui converge faiblement (vers x) est nécessairement bornée. Pour démontrer ce résultat, on utilise le Théorème de Banach-Steinhaus 4.3 page 54 que l'on applique à la suite d'applications linéaires continues définies par

$$T_n(z) := \langle z, x_n \rangle.$$

Pour tout $z \in H$, la suite $(T_n(z))_{n \geq 0}$ converge (car $x_n \rightharpoonup x$ dans H), donc elle est bornée. Il existe donc $M_z > 0$ tel que

$$|T_n(z)| \leq M_z.$$

Le Théorème de Banach-Steinhaus nous assure qu'il existe $M > 0$ tel que

$$|T_n(z)| \leq M \|z\|,$$

pour tout $z \in H$. En appliquant cette inégalité à $z = x_n$, on en déduit que

$$\|x_n\|^2 \leq M \|x_n\|.$$

D'où le résultat.

Remarque 10.11. Les applications linéaires continues préservent la convergence faible. Plus précisément, si $x_n \rightharpoonup x$ dans H et si $A : H \rightarrow H$ est une application linéaire continue, alors $A(x_n) \rightharpoonup A(x)$. En effet, pour tout $y \in H$, on a

$$\langle A(x_n), y \rangle = \langle x_n, A^*(y) \rangle,$$

et la suite $(\langle x_n, A^*(y) \rangle)_{n \geq 0}$ converge par hypothèse vers $\langle x, A^*(y) \rangle = \langle A(x), y \rangle$. Par conséquent, pour tout $y \in H$, la suite $(\langle A(x_n), y \rangle)_{n \geq 0}$ converge vers $\langle A(x), y \rangle$ ce qui traduit bien le fait que $A(x_n) \rightharpoonup A(x)$.

Revenons maintenant à l'étude de la compacité pour la topologie usuelle d'un espace de Hilbert. Bien entendu, 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 10.11 (Compacité du cube de Hilbert). *On note $(e_n)_{n \geq 0}$ la base canonique de $\ell^2(\mathbf{N})$. Le sous-ensemble de $\ell^2(\mathbf{N})$ donné par*

$$K := \left\{ \sum_{n \geq 0} x_n e_n : |x_n| \leq \frac{1}{n+1} \right\},$$

est compact.

DÉMONSTRATION. Remarquons tout d'abord que, la série $\sum_{n \geq 0} \frac{1}{(n+1)^2}$ étant convergente, la condition $|x_n| \leq \frac{1}{n+1}$ garantit que $\sum_{n \geq 0} x_n e_n$ définit bien un élément de $\ell^2(\mathbf{N})$.

Soit $(\mathbf{x}_k)_{k \geq 0}$ une suite d'éléments du cube de Hilbert. Cette suite est bornée car

$$\|\mathbf{x}_k\|_{\ell^2}^2 \leq \sum_{n \geq 0} \frac{1}{(n+1)^2},$$

et d'après le Théorème 10.9 page 214, cette suite admet une sous-suite qui converge faiblement vers $\mathbf{x} \in \ell^2(\mathbf{N})$. Quitte à modifier la suite initiale, on peut directement supposer que $\mathbf{x}_k \rightharpoonup \mathbf{x}$ dans $\ell^2(\mathbf{N})$. En appliquant la définition de la limite faible aux éléments $(e_n)_{n \geq 0}$ de la base canonique Hilbertienne de $\ell^2(\mathbf{N})$, on voit immédiatement que $\mathbf{x} \in K$. En effet, pour tout $k \geq 0$ et pour tout $n \geq 0$, $\langle \mathbf{x}_k, e_n \rangle_{\ell^2} \leq \frac{1}{n+1}$ et, par passage à la limite quand k tend vers $+\infty$, on en déduit que $\langle \mathbf{x}, e_n \rangle_{\ell^2} \leq \frac{1}{n+1}$.

Montrons maintenant que la suite $(x_k)_{k \geq 0}$ converge fortement vers x . Pour tout $\varepsilon > 0$, choisissons $n_0 \in \mathbf{N}$ assez grand pour que

$$\sum_{n \geq n_0} \frac{1}{(n+1)^2} \leq \frac{\varepsilon^2}{4}.$$

Ensuite, choisissons $k_0 \in \mathbf{N}$ assez grand pour que

$$|\langle \mathbf{x}_k - \mathbf{x}, e_n \rangle_{\ell^2}|^2 \leq \frac{\varepsilon^2}{2n_0},$$

pour tout $n = 0, \dots, n_0 - 1$. On a alors pour $k \geq N$. Alors, pour tout $k \geq k_0$, on peut évaluer

$$\|\mathbf{x}_k - \mathbf{x}\|_{\ell^2}^2 \leq \sum_{n=0}^{n_0-1} |\langle \mathbf{x}_k - \mathbf{x}, e_n \rangle_{\ell^2}|^2 + \sum_{n \geq n_0} |\langle \mathbf{x}_k - \mathbf{x}, e_n \rangle_{\ell^2}|^2 \leq \frac{\varepsilon^2}{2} + 2 \sum_{n \geq n_0} \frac{1}{(n+1)^2} \leq \varepsilon^2,$$

ce qui permet de terminer la démonstration. \square

9. Opérateurs à noyau

Une classe d'application joue un rôle particulièrement important en analyse fonctionnelle. Il s'agit de celle des *opérateurs à noyau*. Nous illustrons les résultats de ce chapitre sur cette classe d'opérateurs que nous commençons par définir.

Définition 10.14. Soit $K \in L^2([0, 1]^2; \mathbf{C})$. On définit l'opérateur à noyau

$$A_K : L^2([0, 1]; \mathbf{C}) \rightarrow L^2([0, 1]; \mathbf{C}),$$

par

$$A_K(f)(x) := \int_{[0,1]} K(x, y) f(y) dy.$$

Exemple 10.27. Remarquons que l'exemple 10.13 est un exemple particulier d'opérateur à noyau si l'on prend comme noyau

$$K(x, y) := \mathbf{1}_{[0, +\infty[}(x - y),$$

qui appartient bien à $L^2([0, 1]^2; \mathbf{C})$.

Remarque 10.12. Plus généralement, si Ω est un ouvert de \mathbf{R}^N et si $K \in L^2(\Omega \times \Omega; \mathbf{C})$. On peut définir l'application

$$A_K : L^2(\Omega; \mathbf{C}) \rightarrow L^2(\Omega; \mathbf{C}),$$

par

$$A_K(f)(x) := \int_{\Omega} K(x, y) f(y) dy,$$

dont on vérifie qu'elle est bien définie et continue. De plus, on dispose de l'inégalité

$$\|A_K(f)\|_{L^2(\Omega)} \leq \|K\|_{L^2(\Omega \times \Omega)} \|f\|_{L^2(\Omega)}.$$

Passons maintenant en revue quelques propriétés de ces opérateurs.

Lemme 10.2. L'application A_K est bien définie et continue. De plus

$$\|A_K(f)\|_{L^2([0,1])} \leq \|K\|_{L^2([0,1]^2)} \|f\|_{L^2([0,1])}.$$

DÉMONSTRATION. Le Théorème de Fubini (Théorème 6.5 page 115) nous permet d'affirmer que $y \mapsto |K(x, y)|^2$ est intégrable pour presque tout $x \in [0, 1]$ et que

$$a(x) := \int_{[0,1]} |K(x, y)|^2 dy,$$

est bien définie pour presque tout $x \in [0, 1]$. On sait également que a est intégrable sur $[0, 1]$ et que

$$\int_{[0,1]} a(x) dx = \|K\|_{L^2([0,1]^2)}^2.$$

Soit $f \in L^2([0, 1]; \mathbf{C})$. Pour presque tout $x \in [0, 1]$, la fonction $y \mapsto K(x, y) f(y)$ est intégrable et

$$\left| \int_{[0,1]} K(x, y) f(y) dy \right|^2 \leq \left(\int_{[0,1]} |K(x, y)|^2 dy \right) \left(\int_{[0,1]} |f(y)|^2 dy \right) \leq a(x) \|f\|_{L^2([0,1])}^2,$$

ce qui montre que la fonction

$$x \mapsto \left| \int_{[0,1]} K(x, y) f(y) dy \right|^2,$$

qui est bien entendu une fonction mesurable, est intégrable. De plus

$$\int_{[0,1]} \left| \int_{[0,1]} K(x, y) f(y) dy \right|^2 dx \leq \|f\|_{L^2([0,1])}^2 \iint_{[0,1]^2} |K(x, y)|^2 dx dy.$$

Ce qui démontre l'inégalité mentionné dans le Lemme. \square

En particulier, nous obtenons le :

Lemme 10.3. *L'adjoint de l'opérateur A_K est l'opérateur à noyau défini par*

$$A_K^*(f)(x) := \int_{[0,1]} \overline{K(y, x)} f(y) dy.$$

En particulier A_K est hermitien (ou auto-adjoint) si et seulement si $K(x, y) = \overline{K(y, x)}$ pour presque tout $(x, y) \in [0, 1]^2$.

DÉMONSTRATION. Il s'agit d'une simple application du Théorème de Fubini.

$$\begin{aligned} \langle f, A_K(g) \rangle_{L^2} &= \int_{[0,1]} f(s) \overline{A_K(g)(s)} ds \\ &= \int_{[0,1]} f(s) \left(\int_{[0,1]} \overline{K(s, t)} \overline{g(t)} dt \right) ds \\ &= \int_{[0,1]} \left(\int_{[0,1]} f(s) \overline{K(s, t)} ds \right) \overline{g(t)} dt \\ &= \int_{[0,1]} A_{\tilde{K}}(f)(s) \overline{g(t)} dt \\ &= \langle A_{\tilde{K}}(f), g \rangle_{L^2}, \end{aligned}$$

où $\tilde{K}(x, y) := \overline{K(y, x)}$. Nous laissons le soin au lecteur de justifier proprement chaque étape. \square

Lemme 10.4. *L'application A_K , définie de $L^2([0, 1]; \mathbf{C})$ dans lui-même, est compacte.*

DÉMONSTRATION. Nous proposons deux démonstrations de ce résultat. La première est basée sur le Théorème de Stone-Weierstrass.

On note F le sous-espace de $L^2([0, 1]^2; \mathbf{C})$ constitué des combinaisons linéaires de fonctions de la forme $(x, y) \mapsto f(x) g(y)$ où f et g sont continues sur $[0, 1]$, à valeurs dans \mathbf{C} . Le Théorème de Stone-Weierstrass (voir l'exemple 4.4 page 59) nous permet d'affirmer que F est dense dans $L^2([0, 1]^2; \mathbf{C})$. Le résultat de la Proposition 10.9 permet de conclure que l'opérateur A_K est compact.

Nous donnons maintenant de ce résultat, une autre démonstration basée sur la notion de compacité faible. Soit $(f_n)_{n \geq 0}$ une suite bornée de $L^2([0, 1]; \mathbf{C})$. Quitte à extraire une sous-suite, on peut supposer que cette suite converge faiblement vers une fonction f dans $L^2([0, 1]; \mathbf{C})$. Le Théorème de Fubini nous permet de conclure que

$$\lim_{n \rightarrow +\infty} \int_{[0,1]} K(x, y) f_n(y) dy = \int_{[0,1]} K(x, y) f(y) dy.$$

pour presque tout $x \in [0, 1]$. On en déduit que $(A_K(f_n))_{n \geq 0}$ est une suite qui converge presque partout. Enfin, on utilise l'inégalité

$$|A_K(f_n - f)|(x) \leq \|f_n - f\|_{L^2} \left(\int_{[0,1]} |K(x, y)|^2 dy \right)^{1/2},$$

et le Théorème de convergence dominée pour conclure que $A_K(f_n)$ converge dans $L^2([0, 1]; \mathbf{C})$ vers $A_K(f)$. \square

Spectre des opérateurs et problèmes variationnels

Nous donnons, dans ce chapitre, un certain nombre d'applications de la théorie des espaces de Hilbert que nous venons juste de développer. Rappelons que tous les espaces de Hilbert que nous considérons dans ce cours sont séparables.

1. Prolongement d'opérateurs

On considère deux espaces de Hilbert $(H_1, \langle \cdot, \cdot \rangle_1)$ et $(H_2, \langle \cdot, \cdot \rangle_2)$. Lorsqu'aucune confusion n'est possible sur l'espace ambiant, on note simplement $\langle \cdot, \cdot \rangle$ le produit hermitien et $\| \cdot \|$ la norme associée. On souhaite prolonger une application qui est *a priori* définie sur un sous-espace dense de H_1 et à valeurs dans H_2 , en une application linéaire continue définie sur H_1 à valeurs dans H_2 .

Théorème 11.1. *Soit E un sous-espace dense de H_1 et $A : E \rightarrow H_2$ une application linéaire. On suppose que*

$$\mathcal{N}_E(A) := \sup_{\substack{x \in E \\ \|x\|=1}} \|A(x)\| < +\infty.$$

Alors, A se prolonge en une application linéaire \tilde{A} , c'est-à-dire que $A = \tilde{A}$ sur E , qui est continue de H_1 dans H_2 et telle que $\|\tilde{A}\| = \mathcal{N}_E(A)$.

DÉMONSTRATION. Si $x \in H_1$ est limite d'une suite $(x_n)_{n \geq 0}$ d'éléments de E , on pose

$$\tilde{A}(x) := \lim_{n \rightarrow +\infty} A(x_n).$$

Étant donné que $\|A(x_n) - A(x_m)\| \leq \mathcal{N}_E(A)\|x_n - x_m\|$ et étant donné que la suite $(x_n)_{n \geq 0}$ est une suite de Cauchy, on en déduit que la suite $(A(x_n))_{n \geq 0}$ est elle aussi une suite de Cauchy de H_2 , et donc qu'elle converge dans cet espace vers une limite notée $\tilde{A}(x)$. On vérifie que cette limite ne dépend que de x et ne dépend pas du choix de la suite $(x_n)_{n \geq 0}$ qui converge vers x . En particulier, $\tilde{A}(x) = A(x)$ si $x \in E$ (car x est limite de la suite identiquement égale à x).

D'autre part,

$$\|\tilde{A}(x)\| = \lim_{n \rightarrow +\infty} \|A(x_n)\| \leq \mathcal{N}_E(A) \lim_{n \rightarrow +\infty} \|x_n\| = \mathcal{N}_E(A)\|x\|,$$

ce qui montre que $\|\tilde{A}\| \leq \mathcal{N}_E(A)$. Enfin, on a trivialement l'inégalité $\mathcal{N}_E(A) \leq \|\tilde{A}\|$ ce qui termine la démonstration. \square

2. Transformation de Fourier dans L^2

Le théorème précédent va nous permettre d'étendre la transformation de Fourier, que nous avons définie sur l'espace $L^1(\mathbf{R}^N; \mathbf{C})$, à l'espace des fonctions $L^2(\mathbf{R}^N; \mathbf{C})$. On rappelle que la

transformation de Fourier \mathcal{F} est définie par

$$\mathcal{F}(f)(\xi) := \hat{f}(\xi) = \int_{\mathbf{R}^N} e^{-ix \cdot \xi} f(x) dx,$$

pour toute fonction $f \in L^1(\mathbf{R}^N; \mathbf{C})$. Nous définissons maintenant la *transformée de Fourier normalisée* de f par

$$\Phi(f)(\xi) := \frac{1}{(2\pi)^{N/2}} \int_{\mathbf{R}^N} e^{-ix \cdot \xi} f(x) dx,$$

et la *transformée inverse normalisée* de \hat{f} par

$$\Psi(\hat{f})(x) := \frac{1}{(2\pi)^{N/2}} \int_{\mathbf{R}^N} e^{ix \cdot \xi} \hat{f}(\xi) d\xi.$$

Avant d'énoncer le résultat principal de cette section, démontrons le :

Lemme 11.1. *Pour tout $k \geq 1$, $\mathcal{C}_c^k(\mathbf{R}^N; \mathbf{C})$, l'espace vectoriel des fonctions de classe \mathcal{C}^k à support compact, est dense dans $\mathcal{C}_c(\mathbf{R}^N; \mathbf{C})$ (muni topologie de la convergence uniforme associée à la norme).*

DÉMONSTRATION. Soit $g : \mathbf{R}^N \rightarrow [0, +\infty[$ une fonction \mathcal{C}^∞ à support compact dans $B(0, 1)$. On suppose que

$$\int_{\mathbf{R}^N} g(x) dx = 1.$$

On pourra par exemple considérer un multiple de la fonction qui est définie sur la boule unité par

$$x \mapsto e^{-\frac{|x|^2}{1-|x|^2}},$$

que l'on prolonge par 0 en dehors de la boule unité fermée de \mathbf{R}^N .

Pour tout $n \geq 1$, on note $g_n(x) := n^N g(nx)$. On vérifie que l'intégrale de g_n sur \mathbf{R}^N est encore égale à 1 et que g_n est à support compact dans la boule de rayon $1/n$ centrée à l'origine. Si $f \in \mathcal{C}_c(\mathbf{R}^N; \mathbf{C})$, définissons

$$f_n(x) := f \star g_n(x) := \int_{\mathbf{R}^N} g_n(x - y) f(y) dy.$$

On vérifie que f_n est à support compact et, grâce au Théorème de dérivation sous le signe somme (voir le Théorème 5.10 page 107), qu'elle est de classe \mathcal{C}^∞ sur \mathbf{R}^N . Enfin, la convergence uniforme de la suite $(f_n)_{n \geq 1}$ vers f est assurée par le fait que la fonction f est continue et à support compact, donc elle est uniformément continue (voir le Théorème de Heine, Théorème 3.2 page 39). Pour tout $\varepsilon > 0$, il existe donc $\delta > 0$ tel que, pour tous $x, y \in \mathbf{R}^N$, si $|y - x| \leq \delta$ alors $|f(y) - f(x)| \leq \varepsilon$. En particulier, on a

$$\begin{aligned} |f(x) - f_n(x)| &= \left| \int_{\mathbf{R}^N} g_n(y - x) (f(x) - f(y)) dy \right| \\ &\leq \sup_{z \in B(x, 1/n)} |f(z) - f(x)| \int_{\mathbf{R}^N} g_n(y - x) dy \\ &\leq \varepsilon, \end{aligned}$$

pour tout $n \geq 1/\varepsilon$. Ce qui démontre le résultat. \square

Nous pouvons maintenant démontrer le :

Théorème 11.2. *Avec les notations ci-dessus, les applications Φ et Ψ , qui ne sont a priori définies que sur $L^2(\mathbf{R}^N; \mathbf{C}) \cap L^1(\mathbf{R}^N; \mathbf{C})$, se prolongent en des isométries bijectives de $L^2(\mathbf{R}^N; \mathbf{C})$ dans lui-même et leurs extensions vérifient $\Phi \circ \Psi = \Psi \circ \Phi = \text{Id}_{L^2}$.*

DÉMONSTRATION. Le Théorème 9.3 page 176 nous assure que

$$\Psi \circ \Phi(f) = f,$$

si f et Φf appartiennent à $L^1(\mathbf{R}^N; \mathbf{C})$. Ceci est en particulier vérifié si $f \in \mathcal{C}_c^{N+1}(\mathbf{R}^N; \mathbf{C})$, car dans ce cas $(1 + |\xi|)^{N+1}\Phi(f)$ est bornée et $\Phi(f) \in L^1(\mathbf{R}^N; \mathbf{C})$ d'après le Théorème de Riemann-Lebesgue (Théorème 9.1 page 172). Le Lemme 11.1 nous assure que $\mathcal{C}_c^{N+1}(\mathbf{R}^N; \mathbf{C})$ est dense dans $\mathcal{C}_c(\mathbf{R}^N; \mathbf{C})$ (muni de la norme de la convergence uniforme) qui lui-même est dense dans l'espace $L^2(\mathbf{R}^N; \mathbf{C})$ (muni de la norme L^2), le théorème ci-dessus permet donc de prolonger Φ à $L^2(\mathbf{R}^N; \mathbf{C})$.

Par ailleurs, on vérifie que Ψ est l'adjoint de Φ pour le produit hermitien de $L^2(\mathbf{R}^N; \mathbf{C})$, c'est-à-dire que

$$\langle \Phi(f), g \rangle_{L^2} = \langle f, \Psi(g) \rangle_{L^2}.$$

En effet, si $f, g \in \mathcal{C}_c^{N+1}(\mathbf{R}^N; \mathbf{C})$, alors comme $(x, \xi) \mapsto f(x)g(\xi)$ est dans $L^1(\mathbf{R}^{2N}; \mathbf{C})$, le Théorème de Fubini permet d'intervertir les intégrales dans le calcul suivant

$$\begin{aligned} \langle \Phi(f), g \rangle_{L^2} &= \frac{1}{(2\pi)^{N/2}} \int_{\mathbf{R}^N} \left(\int_{\mathbf{R}^N} e^{-ix\xi} f(x) dx \right) \bar{g}(\xi) d\xi \\ &= \frac{1}{(2\pi)^{N/2}} \int_{\mathbf{R}^N} \int_{\mathbf{R}^N} f(x) \bar{g}(\xi) e^{-ix\xi} dx d\xi \\ &= \frac{1}{(2\pi)^{N/2}} \int_{\mathbf{R}^N} f(x) \left(\int_{\mathbf{R}^N} g(\xi) e^{ix\xi} d\xi \right) dx \\ &= \langle f, \Psi(g) \rangle_{L^2}. \end{aligned}$$

On en déduit que

$$\langle \Phi(f), \Phi(g) \rangle_{L^2} = \langle f, \Psi \circ \Phi(g) \rangle_{L^2} = \langle f, g \rangle_{L^2},$$

et de manière analogue

$$\langle \Psi(f), \Psi(g) \rangle_{L^2} = \langle f, g \rangle_{L^2}.$$

En particulier, les extensions de Φ et Ψ à l'espace $L^2(\mathbf{R}; \mathbf{C})$ sont des isométries. Enfin, $\Phi \circ \Psi = \Psi \circ \Phi = \text{Id}$, car ces égalités sont vérifiées sur $\mathcal{C}_c^{N+1}(\mathbf{R}^N; \mathbf{C})$ qui est un sous-espace dense de $L^2(\mathbf{R}^N; \mathbf{C})$ et une identité vraie sur un sous-espace dense est, par continuité, vraie sur tout l'espace. \square

Ce résultat permet de résoudre le problème de la diffraction de Fraunhofer tel qu'il est mentionné dans le Chapitre 9. En effet, si l'on connaît $\Phi(\mathbf{1}_P)$, où $\mathbf{1}_P$ 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 $\Psi(\Phi(\mathbf{1}_P)) = \mathbf{1}_P$. Donc, en appliquant Ψ (qui est, à un coefficient multiplicatif près, la transformation de Fourier) à $\widehat{\mathbf{1}_P}$, on retrouve $\mathbf{1}_P$.

3. Inégalité de Heisenberg

Nous avons déjà mentionné l'inégalité de Heisenberg à la fin du chapitre sur la transformation de Fourier. Cette inégalité fonctionnelle trouve de nombreuses applications de la physique quantique ou à la théorie du signal. Elle affirme qu'un signal ne peut être localisé à la fois en temps et en fréquence, ou encore qu'une particule ne peut avoir simultanément une position et une quantité de mouvement bien déterminées (nous renvoyons au cours de mécanique

quantique [2] pour plus de précisions sur les conséquences de cette inégalité), et elle explicite cette restriction de manière quantitative.

Proposition 11.1. *Si f est une fonction de $\mathcal{C}_c^1(\mathbf{R}; \mathbf{C})$. On a*

$$\left(\int_{\mathbf{R}} x^2 |f(x)|^2 dx \right)^{1/2} \left(\int_{\mathbf{R}} \left| \frac{df}{dx}(x) \right|^2 dx \right)^{1/2} \geq \frac{1}{2} \int_{\mathbf{R}} |f(x)|^2 dx.$$

Cette inégalité est équivalente à l'inégalité suivante, vérifiée par toute fonction de $L^2(\mathbf{R}; \mathbf{C})$

$$\left(\int_{\mathbf{R}} x^2 |f(x)|^2 dx \right)^{1/2} \left(\int_{\mathbf{R}} \xi^2 |\hat{f}(\xi)|^2 d\xi \right)^{1/2} \geq \sqrt{\frac{\pi}{2}} \int_{\mathbf{R}} |f(x)|^2 dx.$$

Bien entendu, dans l'expression ci-dessus, les termes de gauche ne sont finis que si les fonctions $x \mapsto x f(x)$ et $\xi \mapsto \xi \hat{f}(\xi)$ sont dans $L^2(\mathbf{R}; \mathbf{C})$. Remarquons que $\xi \mapsto \xi \hat{f}(\xi)$ est dans $L^2(\mathbf{R}; \mathbf{C})$ si et seulement si sa transformée de Fourier y est, c'est-à-dire si et seulement si $x \mapsto \frac{df}{dx}(x)$ est dans $L^2(\mathbf{R}; \mathbf{C})$. On peut vérifier que l'inégalité de Heisenberg est satisfaite dès que les termes de gauche sont définis.

DÉMONSTRATION DE L'INÉGALITÉ DE HEISENBERG. Nous ne démontrons ici l'inégalité de Heisenberg, que le cas où la fonction f est dans $\mathcal{C}_c^1(\mathbf{R}; \mathbf{C})$, l'extension au cas où $x \mapsto x f(x)$ et $\xi \mapsto \xi \hat{f}(\xi)$ sont dans $L^2(\mathbf{R}; \mathbf{C})$ sont laissée en exercice au lecteur. On notera f' la dérivée de f .

Grâce à une intégration par parties, on peut écrire

$$\int_{\mathbf{R}} |f(x)|^2 dx = - \int_{\mathbf{R}} x \frac{d}{dx} |f(x)|^2 dx = - \int_{\mathbf{R}} (f'(x) \bar{f}(x) + f(x) \overline{f'(x)}) x dx.$$

On utilise alors l'inégalité de Cauchy-Schwarz pour conclure que

$$\int_{\mathbf{R}} |f(x)|^2 dx \leq 2 \left(\int_{\mathbf{R}} x^2 |f(x)|^2 dx \right)^{1/2} \left(\int_{\mathbf{R}} |f'(x)|^2 dx \right)^{1/2}.$$

Si de plus $f \in \mathcal{C}_c^2(\mathbf{R}; \mathbf{C})$ alors $\hat{f}'(\xi) = i \xi \hat{f}(\xi) \in L^2(\mathbf{R}; \mathbf{C})$ et la seconde inégalité est alors équivalente à la première. \square

On peut étudier les cas d'égalité dans l'inégalité de Heisenberg. On constate facilement que, pour tout $t > 0$, les fonctions

$$f(x) := e^{-\frac{x^2}{2t}},$$

réalisent l'égalité dans l'inégalité de Heisenberg. On laisse au lecteur le soin de démontrer que les multiples de ces fonctions sont en fait toutes les fonctions qui réalisent les cas d'égalité.

L'interprétation en termes de concentration est la suivante : si l'on souhaite mesurer la concentration de f au voisinage d'un point x_0 et celle des fréquences de f au voisinage de ξ_0 , on regarde les quantités

$$\left(\int_{\mathbf{R}} (x - x_0)^2 |f(x)|^2 dx \right) \quad \text{et} \quad \left(\int_{\mathbf{R}} (\xi - \xi_0)^2 |\hat{f}(\xi)|^2 d\xi \right),$$

qui satisfont une inégalité semblable à celle énoncée dans le cas où $x_0 = 0$ et $\xi_0 = 0$, à savoir

$$(11.1) \quad \left(\int_{\mathbf{R}} (x - x_0)^2 |f(x)|^2 dx \right)^{1/2} \left(\int_{\mathbf{R}} (\xi - \xi_0)^2 |\hat{f}(\xi)|^2 d\xi \right)^{1/2} \geq \sqrt{\frac{\pi}{2}} \left(\int_{\mathbf{R}} |f(x)|^2 dx \right)^{1/2}.$$

Les cas d'égalités correspondent maintenant aux fonctions

$$f_{x_0, \xi_0}(x) := e^{-\frac{(x-x_0)^2}{2t}} e^{i\xi_0 x},$$

que les physiciens appellent *états cohérents*. Ces fonctions représentent une particule localisée dans l'espace au voisinage de x_0 et en fréquence au voisinage de ξ_0 , autant que faire se peut, c'est-à-dire dans la limite autorisée par l'inégalité de Heisenberg.

La décomposition de Fourier permet d'écrire toute fonction (raisonnable) comme somme des $x \mapsto e^{i\xi x}$ qui sont des états parfaitement localisés en fréquence (égale à ξ), mais qui ne sont pas du tout localisés en espace. Il est souvent utile d'avoir une décomposition qui tient compte de la localisation spatiale des fonctions. Une méthode couramment utilisée avant l'avènement des ondelettes—dont nous avons déjà parlé dans le chapitre précédent—consiste en une transformation de Fourier « avec fenêtre » dont le principe de base est le suivant : on restreint la fonction $x \mapsto f(x)$ à $[2\pi n, 2\pi(n+1)]$ et l'on décompose cette fonction en série de Fourier (on peut par exemple la prolonger par périodicité). L'un des problèmes est l'ajout artificiel de hautes fréquences dues à la troncature. Les ondelettes permettent, elles, de conserver une information qui tient compte à la fois de la localisation spatiale et fréquentielle.

Pour terminer cette étude, nous faisons le lien avec l'inégalité de Heisenberg telle qu'elle peut être présentée dans un cours de Mécanique Quantique [2]. On vérifie qu'en tant que fonction des variables x_0 et ξ_0 , le terme de gauche de (11.1) est minimal lorsque

$$x_0 = \frac{\int_{\mathbf{R}} x |f(x)|^2 dx}{\int_{\mathbf{R}} |f(x)|^2 dx} \quad \text{et} \quad \xi_0 = \frac{\int_{\mathbf{R}} \xi |\hat{f}(\xi)|^2 d\xi}{\int_{\mathbf{R}} |\hat{f}(\xi)|^2 d\xi}.$$

En mécanique quantique, la fonction f représente une fonction d'onde et elle est normalisée de telle sorte que $|f(x)|^2 dx$ soit une mesure de probabilité. Autrement dit, on suppose que f est normalisée de telle sorte que

$$\int_{\mathbf{R}} |f(x)|^2 dx = 1.$$

On interprète alors x_0 comme la valeur de l'observable « position »

$$\langle x \rangle := \langle f | A | f \rangle = \int_{\mathbf{R}} x |f(x)|^2 dx,$$

où l'opérateur position A est défini par $A(f)(x) := x f(x)$ et l'on interprète ξ_0 comme la valeur de l'observable « quantité de mouvement »

$$\langle p \rangle := \langle f | B | f \rangle = \int_{\mathbf{R}} \overline{f(x)} i \frac{df}{dx}(x) dx = \frac{1}{2\pi} \int_{\mathbf{R}} \xi |\hat{f}(\xi)|^2 d\xi.$$

où l'opérateur quantité de mouvement B est défini par $B(f)(x) := i \frac{df}{dx}(x)$ (attention au fait que la convention des physiciens pour le produit hermitien est différente de la convention choisie dans ce cours ce qui explique la première égalité). Si

$$\langle x^2 \rangle := \int_{\mathbf{R}} x^2 |f(x)|^2 dx \quad \text{et} \quad \langle p^2 \rangle := \frac{1}{2\pi} \int_{\mathbf{R}} \xi^2 |\hat{f}(\xi)|^2 d\xi.$$

L'inégalité de Heisenberg s'écrit alors

$$\Delta x \Delta p \geq \frac{1}{2},$$

où

$$(\Delta x)^2 := \langle x^2 \rangle - \langle x \rangle^2 \quad \text{et} \quad (\Delta p)^2 := \langle p^2 \rangle - \langle p \rangle^2,$$

et l'on retrouve ainsi la formule classique utilisée dans les cours de mécanique quantique, qui exprime le fait que le produit de l'incertitude sur la mesure de la position (écart type) et de l'incertitude sur la mesure de la quantité de mouvement est minoré par une constante strictement positive.

En fait, l'inégalité de Heisenberg peut aussi être vue comme un défaut de commutativité entre les deux opérateurs hermitiens A et B . En effet, on note

$$C := i(A B - B A),$$

dont on vérifie que c'est encore un opérateur hermitien (pour le produit hermitien L^2). En supposant que toutes les intégrales sont bien définies, on calcule, pour tout $t \in \mathbf{R}$

$$\int_{\mathbf{R}} |(A + itB)(f)|^2 dx = \int_{\mathbf{R}} |A(f)|^2 dx + t \int_{\mathbf{R}} (C(f)) \bar{f} dx + t^2 \int_{\mathbf{R}} |B(f)|^2 dx.$$

En utilisant le fait que cette quantité est positive pour tout $t \in \mathbf{R}$ et en poursuivant comme dans la démonstration de l'inégalité de Cauchy-Schwartz, on obtient

$$\left(\int_{\mathbf{R}} (C(f)) \bar{f} dx \right)^2 \leq 4 \left(\int_{\mathbf{R}} |A(f)|^2 dx \right) \left(\int_{\mathbf{R}} |B(f)|^2 dx \right).$$

Choisissons les opérateurs A et B tels que

$$A(f)(x) := xf(x) \quad \text{et} \quad B(f)(x) := i \frac{df}{dx}(x).$$

On vérifie que dans ce cas $C(f) = f$, et l'on retrouve alors

$$\left(\int_{\mathbf{R}} |f|^2 dx \right)^2 \leq 4 \left(\int_{\mathbf{R}} x^2 |f|^2 dx \right) \left(\int_{\mathbf{R}} \left| \frac{df}{dx} \right|^2 dx \right).$$

4. Spectre d'un opérateur

On rappelle qu'un endomorphisme A continu, défini sur un espace de Hilbert complexe H , est *hermitien* (on dira que c'est un *opérateur hermitien*) si

$$\langle A(x), y \rangle = \langle x, A(y) \rangle,$$

pour tous $x, y \in H$ (dans le cadre des espaces de Hilbert réels, on parlera d'opérateur symétrique). Rappelons également que les valeurs propres d'un opérateur hermitien sont nécessairement réelles. En effet, si $A(x) = \lambda x$ et si $x \neq 0$, alors

$$\lambda \langle x, x \rangle = \langle A(x), x \rangle = \langle x, A(x) \rangle = \bar{\lambda} \langle x, x \rangle,$$

donc $\lambda \in \mathbf{R}$.

Enfin, on rappelle que si A est une application hermitienne d'un espace vectoriel de *dimension finie* dans lui-même, il existe une base orthonormée de l'espace constituée de vecteurs propres de A . On rappelle aussi qu'en dimension finie un endomorphisme est injectif si et seulement s'il est surjectif. En particulier, les valeurs propres de A sont définies comme étant les $\lambda \in \mathbf{R}$ tels que $A - \lambda \text{Id}$ n'est pas inversible.

On se demande s'il existe une théorie semblable pour les endomorphismes définis sur des espaces de Hilbert de dimension infinie. On dira qu'un endomorphisme continu A d'un espace

de Hilbert est *inversible* s'il est bijectif et si A^{-1} est continu.¹ Remarquons qu'en dimension infinie, un endomorphisme d'un espace de Hilbert peut ne pas être bijectif tout en étant injectif. Cela nous amène à distinguer deux objets :

Définition 11.1. Soit A un endomorphisme continu d'un espace de Hilbert H . On appelle spectre de A et l'on note $\text{sp}(A)$, l'ensemble des $\lambda \in \mathbf{C}$ tels que $A - \lambda \text{Id}_H$ n'est pas inversible. Autrement dit, si $\lambda \notin \text{sp}(A)$ alors A est inversible et A^{-1} est un endomorphisme continu.

On appelle ensemble des valeurs propres de A et l'on note $\text{vp}(A)$, l'ensemble des $\lambda \in \mathbf{C}$ tels que $A - \lambda \text{Id}_H$ n'est pas injectif, ou encore, l'ensemble des $\lambda \in \mathbf{C}$ pour lesquels il existe $x \in H - \{0\}$ tel que $A(x) = \lambda x$.

Bien entendu, on a l'inclusion $\text{vp}(A) \subset \text{sp}(A)$ et l'exemple ci-dessous montre que l'égalité n'a pas nécessairement lieu.

Exemple 11.1. Soit A l'endomorphisme de $L^2(S^1; \mathbf{C})$ défini par

$$A(f)(t) := \cos t f(t).$$

Soit f un vecteur propre de A , alors $A(f) = \lambda f$ signifie que $\cos t f(t) = \lambda f(t)$, d'où $f = 0$ p.p. et donc $f \equiv 0$ dans $L^2(S^1; \mathbf{C})$. En particulier $\text{vp}(A) = \emptyset$. On vérifie que $A - \lambda \text{Id}$ est inversible si et seulement si $\lambda \in \mathbf{C} - [-1, 1]$. Donc que $\text{sp}(A) = [-1, 1]$.

Nous cherchons alors une classe d'opérateurs pour lesquels la situation est proche de celle de la dimension finie. Plus précisément, le résultat qui suit est une généralisation du théorème spectral au cas d'un espace de Hilbert de dimension quelconque :

Théorème 11.3. Soit A un opérateur hermitien compact d'un espace de Hilbert H , qui est séparable de dimension infinie, dans lui-même. Alors, il existe $(e_n)_{n \geq 0}$ une base hilbertienne de H , constituée de vecteurs propres de A , i.e. $A(e_n) = \lambda_n e_n$.

- (i) Les valeurs propres sont réelles et tendent vers 0 quand n tend vers l'infini.
- (ii) Les valeurs propres non nulles sont de multiplicité finie.
- (iii) On a $\text{sp}(A) = \text{vp}(A) \cup \{0\}$.

DÉMONSTRATION. Comme dans le cas de la dimension finie, on vérifie que deux sous-espaces propres de A sont orthogonaux et que l'orthogonal d'un espace invariant par A est invariant par A . On vérifie aussi que les valeurs propres sont nécessairement réelles.

Par exemple, si $A(x) = \lambda x$ avec $x \neq 0$ alors

$$\lambda \langle x, x \rangle = \langle A(x), x \rangle = \langle x, A^*(x) \rangle = \langle x, A(x) \rangle = \bar{\lambda} \langle x, x \rangle,$$

donc, nécessairement, $\lambda \in \mathbf{R}$. De même, si $A(x) = \lambda x$ et $A(y) = \mu y$ on a

$$\lambda \langle x, y \rangle = \langle A(x), y \rangle = \langle x, A^*(y) \rangle = \langle x, A(y) \rangle = \mu \langle x, y \rangle,$$

si bien que, si $\lambda \neq \mu$, on a $\langle x, y \rangle = 0$. Enfin, si E est invariant par A et si $x \in E^\perp$, on aura pour tout $y \in E$,

$$\langle A(x), y \rangle = \langle x, A^*(y) \rangle = \langle x, A(y) \rangle = 0,$$

puisque $A(y) \in E$. On en déduit que $A(x)$ est aussi dans E^\perp , ce qui montre que E^\perp est invariant par A .

Soit alors F l'adhérence de la somme de tous les espaces propres de A (i.e. l'adhérence des combinaisons linéaires finies de vecteurs propres de A). Par définition, c'est un sous-espace

1. En fait, on montre que si A est un endomorphisme continu sur un espace de Hilbert et si A est bijectif, alors A^{-1} est un endomorphisme continu, c'est une conséquence du *Théorème de l'application ouverte*.

vectorel fermé de H et donc un espace de Hilbert. Vérifions 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 11.2. *Soit A un opérateur hermitien, compact défini sur un espace de Hilbert E non réduit à $\{0\}$. On note $\lambda := \sup_{\|x\|=1} \langle A(x), x \rangle$ et $\mu := \inf_{\|x\|=1} \langle A(x), x \rangle$. Si $\lambda > 0$ (ou $\mu < 0$) alors λ (ou μ) est une valeur propre de A .*

DÉMONSTRATION. On suppose que $\lambda > 0$ et montrons que le supremum est atteint. Soit $(x_n)_{n \geq 0}$ une suite d'éléments de E de norme 1 tels que $\lim_{n \rightarrow +\infty} \langle A(x_n), x_n \rangle = \lambda$.

Par compacité de A , on peut extraire de la suite $(x_n)_{n \geq 0}$ une sous-suite telle que $(A(x_{\varphi(n)}))_{n \geq 0}$ converge vers $z \in E$. Comme

$$|\langle A(x_n) - z, x_n \rangle| \leq \|A(x_n) - z\| \|x_n\|,$$

on en déduit que $\langle x_{\varphi(n)}, z \rangle$ converge vers λ . Quitte à extraire une nouvelle sous-suite, on peut supposer que $x_{\varphi(n)}$ converge faiblement vers $x_\infty \in E$. D'après le résultat de la Proposition 10.10 page 214, on a $\|x_\infty\| \leq 1$.

De plus, la suite $(A(x_{\varphi(n)}))_{n \geq 0}$ converge faiblement vers $A(x_\infty)$ et fortement vers z , ce qui entraîne que $A(x_\infty) = z$, et donc $\langle A(x_\infty), x_\infty \rangle = \lambda$. Comme $\lambda \neq 0$ on a $x_\infty \neq 0$ et en posant $y = \frac{x_\infty}{\|x_\infty\|}$, on trouve que $\langle A(y), y \rangle \geq \lambda$ et donc le maximum est atteint en y .

Il reste à vérifier que $\langle A(y), w \rangle = 0$ pour tout w orthogonal à y . Mais, l'application $x \mapsto \langle A(x), x \rangle$ restreinte à la sphère unité atteint son maximum en y . Donc, pour tout w (de norme 1) orthogonal à y , nous pouvons considérer la fonction

$$t \mapsto \frac{1}{1+t^2} \langle A(y+tw), (y+tw) \rangle,$$

qui admet un maximum en $t = 0$. En dérivant par rapport à t en $t = 0$, on trouve que $\Re \langle A(y), w \rangle = 0$, puis, en remplaçant w par iw , on conclut que $\langle A(y), w \rangle = 0$. Donc $A(y)$ est colinéaire à y , ce qui permet montre que $A(y) = \lambda y$. \square

La démonstration de la première partie du théorème résulte immédiatement de ce que par hypothèse, A n'a pas de valeur propre sur F^\perp . On en déduit d'après le Lemme que, pour tout $x \in F^\perp$, $\langle A(x), x \rangle = 0$. Mais A étant hermitien, la formule de polarisation permet de dire que $\langle A(x), y \rangle = 0$ pour tout $x, y \in F^\perp$ et donc que $A \equiv 0$ sur F^\perp . Mais alors $F^\perp \subset \text{Ker} A$ ce qui contredit notre hypothèse.

Pour terminer la démonstration du Théorème il suffit de démontrer le :

Lemme 11.3. *Soit A un opérateur compact défini sur un espace de Hilbert H . Alors l'espace propre associé à une valeur propre λ non nulle est de dimension finie et l'ensemble des valeurs propres n'a pas d'autre point d'accumulation que 0.*

Autrement dit, pour tout $\varepsilon > 0$, une somme directe des espaces propres associés aux valeurs propres de module supérieur à $\varepsilon > 0$ est de dimension finie.

DÉMONSTRATION. En effet, soit $E_\lambda = \text{Ker}(A - \lambda \text{Id})$ l'espace propre associé à λ . Si E_λ est de dimension infinie, on peut trouver une suite $(x_n)_{n \geq 0}$ orthonormée de vecteurs de E_λ . La suite $(A(x_n))_{n \geq 0}$ n'admet pas de sous suite convergente, ce qui contredit la compacité de A . On en déduit que les espaces propres de A sont de dimensions finies. On démontre de même que la suite des valeurs propres n'a pas de point d'accumulation autre que 0. \square

Le Lemme ci-dessus permet de compléter la démonstration du Théorème. \square

On dit qu'un opérateur hermitien est positif si $\langle A(x), x \rangle \geq 0$ pour tout $x \in H$.

Proposition 11.2 (Caractérisation variationnelle des valeurs propres). *Soit A un opérateur hermitien, compact et positif sur l'espace de Hilbert H . Alors, les valeurs propres de A sont données par les formules suivantes*

$$\lambda_k = \sup_{\dim E=k} \inf_{\substack{x \in E \\ \|x\|=1}} \langle A(x), x \rangle.$$

DÉMONSTRATION. D'après le résultat précédent, $(\lambda_k)_{k \geq 1}$, les valeurs propres de A forment une suite positive, décroissante qui tend vers 0. On a alors une base hilbertienne $(e_k)_{k \geq 1}$ de H constituée des vecteurs propres de A associés aux valeurs propres de A . i.e. $A(e_k) = \lambda_k e_k$.

Soit E_0 l'espace engendré par (e_1, \dots, e_k) , on a

$$\begin{aligned} \langle A(x_1 e_1 + \dots + x_k e_k), x_1 e_1 + \dots + x_k e_k \rangle &= \lambda_1 x_1^2 + \dots + \lambda_k x_k^2 \\ &\geq \lambda_k (x_1^2 + \dots + x_k^2) \\ &\geq \lambda_k \|x_1 e_1 + \dots + x_k e_k\|^2. \end{aligned}$$

En particulier

$$\sup_{\dim E=k} \inf_{\substack{x \in E \\ \|x\|=1}} \langle A(x), x \rangle \geq \lambda_k.$$

De même on montre que

$$|\langle A(x), x \rangle| \leq \lambda_k \|x\|^2,$$

pour tout $x \in F_0 := E_0^\perp \oplus \mathbf{C}e_k$. Maintenant, tout sous-espace de H de dimension k doit rencontrer F_0 , car si P désigne la restriction à E de la projection orthogonale sur F_0^\perp , qui est de dimension $k-1$, l'application P a nécessairement un noyau. Par définition un élément de ce noyau est un élément $x \in E \cap F_0$. On a alors trouvé un vecteur $x \in E$ tel que $\langle A(x), x \rangle \leq \lambda_k \|x\|^2$. Cela montre que

$$\sup_{\dim E=k} \inf_{\substack{x \in E \\ \|x\|=1}} \langle A(x), x \rangle \leq \lambda_k.$$

D'où le résultat. □

Remarque 11.1. De manière analogue, on a

$$\lambda_k = \inf_{\dim E=k-1} \sup_{\substack{x \in E^\perp \\ \|x\|=1}} \langle A(x), x \rangle.$$

Nous reprenons notre étude des opérateurs à noyau. Rappelons que si K une fonction de $L^2([0, 1]^2; \mathbf{C})$ telle que $K(x, y) = \overline{K(y, x)}$, l'opérateur

$$A_K(f)(x) := \int_{[0,1]} K(x, y) f(y) dy,$$

défini de $L^2([0, 1]; \mathbf{C})$ dans lui-même est un opérateur continu, hermitien et compact. Par conséquent, on a la proposition suivante :

Proposition 11.3. *Sous les hypothèses ci-dessus, il existe une suite $((\varphi_n, \lambda_n))_{n \geq 0}$ de solutions de*

$$\int_{[0,1]} K(s, t) \varphi(t) dt = \lambda \varphi(s),$$

p.p. sur $[0, 1]$, telle que $(\varphi_n)_{n \geq 0}$ est une base hilbertienne de $L^2([0, 1]; \mathbf{C})$ et $(\lambda_n)_{n \geq 0}$ une suite de réels qui tend vers 0.

Le résultat suivant montre que la suite des valeurs propres de A_K tend vers 0 assez rapidement :

Lemme 11.4. *Sous les hypothèses ci-dessus, on a l'égalité*

$$\sum_{n \in \mathbf{N}} \lambda_n^2 = \int_{[0,1]^2} |K(x, y)|^2 dx dy.$$

DÉMONSTRATION. Le Théorème de Fubini nous assure que, la fonction $y \mapsto \overline{K(x, y)}$ appartient à $L^2([0, 1])$ pour presque tout $x \in [0, 1]$. On peut donc écrire, grâce à l'égalité de Parseval et grâce au fait que $(\varphi_n)_{n \geq 0}$ est une base hilbertienne de $L^2([0, 1])$ que

$$\int_{[0,1]} |K(x, y)|^2 dy = \sum_{n \geq 0} |\langle \overline{K(x, \cdot)}, \varphi_n \rangle_{L^2}|^2$$

La fonction $x \mapsto \int_{[0,1]} |K(x, y)|^2 dy$ est intégrable et, en utilisant le Théorème de Fubini, on obtient

$$\iint_{[0,1]^2} |K(x, y)|^2 dx dy = \int_{[0,1]} \left(\sum_{n \geq 0} |\langle \overline{K(x, \cdot)}, \varphi_n \rangle_{L^2}|^2 \right) dx.$$

Le Théorème de la convergence monotone de Beppo Levi permet d'intervertir l'intégrale et la somme dans le membre de droite pour conclure que

$$\iint_{[0,1]^2} |K(x, y)|^2 dx dy = \sum_{n \geq 0} \left(\int_{[0,1]} |\langle \overline{K(x, \cdot)}, \varphi_n \rangle_{L^2}|^2 dx \right).$$

Or, par construction

$$\langle \overline{K(x, \cdot)}, \varphi_n \rangle_{L^2} = \overline{A_K(\varphi_n)}(x) = \overline{\lambda_n \varphi_n(x)}.$$

Donc

$$\iint_{[0,1]^2} |K(x, y)|^2 dx dy = \sum_{n \geq 0} |\lambda_n|^2,$$

ce qui démontre l'égalité. □

Remarque 11.2. Sous les hypothèses ci-dessus, on montre également que

$$K(x, y) = \sum_{n \in \mathbf{N}} \lambda_n \varphi_n(x) \overline{\varphi_n(y)}.$$

5. L'espace de Sobolev $H^1(I; \mathbf{C})$

Dans cette section, on se donne $I :=]a, b[$ un intervalle ouvert, non vide, borné et $x_0 \in I$. L'espace de Sobolev $H^1(I; \mathbf{C})$ est l'espace des fonctions, à valeurs dans \mathbf{C} , qui sont des primitives de fonctions de $L^2(I; \mathbf{C})$. On définit de la même façon l'espace de Sobolev des fonctions à valeurs dans \mathbf{R} . Plus précisément :

Définition 11.2. *On définit l'espace de Sobolev $H^1(I; \mathbf{C})$ par*

$$H^1(I; \mathbf{C}) := \left\{ u \in L^2(I; \mathbf{C}) : \exists v \in L^2(I; \mathbf{C}), \exists c \in \mathbf{C}, u(x) = c + \int_{x_0}^x v(t) dt \quad \text{sur } I \right\}.$$

où, par convention,

$$\int_{x_0}^x v(t) dt := \int_{[x_0, x]} v(t) dt,$$

quand $x \geq x_0$ et

$$\int_{x_0}^x v(t) dt := - \int_{[x, x_0]} v(t) dt,$$

quand $x \leq x_0$.

Dans cette définition, l'égalité

$$u(x) = c + \int_{x_0}^x v(t) dt,$$

doit être comprise au sens de l'égalité de deux fonctions dans $L^2(I; \mathbf{C})$. Dans le même ordre d'idée, rappelons que les éléments de $L^2(I; \mathbf{C})$ sont en fait des classes d'équivalences de fonctions égales presque partout. Commençons par vérifier le :

Lemme 11.5. *Toute fonction $u \in H^1(I; \mathbf{C})$ admet un représentant continu.*

DÉMONSTRATION. Soit $u \in H^1(I; \mathbf{C})$. Par hypothèse, il existe $c \in \mathbf{C}$ et $v \in L^2(I; \mathbf{C})$ tels que

$$u(x) = c + \int_{x_0}^x v(t) dt,$$

pour presque tout $x \in I$. Or la fonction qui apparaît dans le membre de droite de cette égalité définit une fonction qui est continue. En effet, grâce à l'inégalité de Cauchy-Schwarz, nous pouvons écrire

$$\left| \int_{x_0}^y v(t) dt - \int_{x_0}^x v(t) dt \right| \leq \|v\|_{L^2(I)} |y - x|^{1/2},$$

pour tout $x, y \in I$ ce qui montre que la fonction u admet un représentant continu. \square

Dans la suite, on identifiera $u \in H^1(I; \mathbf{C})$ avec son représentant continu. Le résultat suivant est immédiat :

Lemme 11.6. *Dans la définition de $H^1(I; \mathbf{C})$, la décomposition de $u \in H^1(I; \mathbf{C})$ en*

$$u(x) = c + \int_{x_0}^x v(t) dt,$$

où $c \in \mathbf{C}$ et $v \in L^2(I; \mathbf{C})$, est unique.

DÉMONSTRATION. Supposons que

$$u(x) = c + \int_{x_0}^x v(t) dt = \tilde{c} + \int_{x_0}^x \tilde{v}(t) dt,$$

où $v, \tilde{v} \in L^2(I; \mathbf{C})$ et $c, \tilde{c} \in \mathbf{C}$. Alors

$$\tilde{c} - c = \int_{x_0}^x (v(t) - \tilde{v}(t)) dt,$$

pour tout $x \in I$. L'inégalité de Cauchy-Schwarz nous permet d'écrire que

$$|\tilde{c} - c| \leq \|\tilde{v} - v\|_{L^2} |x_0 - x|^{1/2},$$

et, en faisant tendre x vers x_0 , on conclut que $\tilde{c} = c$. Donc,

$$\int_{x_0}^x (v(t) - \tilde{v}(t)) dt = 0,$$

pour tout $x \in I$. En d'autres termes, $w := \tilde{v} - v$ est orthogonale à toutes les fonctions de la forme $\mathbf{1}_{[a,b]}$, où $[a,b] \subset I$, donc w est orthogonale à tous les éléments du sous-espace E engendré par ces fonctions. Or, on sait que E est dense dans $L^2(I; \mathbf{C})$. En particulier, il existe une suite $(w_n)_{n \geq 0}$ d'éléments de E qui converge vers w dans $L^2(I; \mathbf{C})$. Alors

$$\int_I |w(t)|^2 dt = \lim_{n \rightarrow +\infty} \int_I w_n(t) \overline{w}(t) dt = 0.$$

Donc $w = 0$ p.p. sur I ce qui termine la démonstration. \square

Remarquons que, le Théorème 7.4 qui permet d'affirmer que, si $v \in L^2(I; \mathbf{C})$, la fonction

$$x \mapsto \int_{x_0}^x v(t) dt,$$

est dérivable pour presque tout $x \in I$ et que sa dérivée est donnée par v . On peut s'appuyer sur ce résultat dans la démonstration ci-dessus pour conclure que la fonction $w = 0$ p.p. sur I .

Si u est le représentant continu d'un élément de $H^1(I; \mathbf{C})$, on conclut que $c = u(x_0)$ et ce même Théorème 7.4, plus permet d'affirmer que u est dérivable p.p. sur I et que sa dérivée (qui est définie p.p. sur I) est donnée par v . On dira que v est la *dérivée faible* de $u \in H^1(I; \mathbf{C})$ et l'on notera $u' := v$.

On peut munir l'espace $H^1(I; \mathbf{C})$ d'une structure hilbertienne induite par la structure hilbertienne de $L^2(I; \mathbf{C})$.

Proposition 11.4. *Muni du produit hermitien*

$$\langle u, v \rangle_{H^1(I; \mathbf{C})} := \int_I u(t) \overline{v}(t) dt + \int_I u'(t) \overline{v'}(t) dt,$$

l'espace $H^1(I; \mathbf{C})$ est un espace de Hilbert.

DÉMONSTRATION. Soit $(u_n)_{n \geq 0}$ une suite de Cauchy dans $H^1(I; \mathbf{C})$. On utilise le résultat du Lemme 11.7 pour conclure que la suite $(u_n)_{n \geq 0}$ est une suite de Cauchy dans $L^\infty(I; \mathbf{C})$, elle converge donc dans cet espace vers une limite notée u . De même, la suite $(u'_n)_{n \geq 0}$ est une suite de Cauchy dans $L^2(I; \mathbf{C})$, elle converge donc dans $L^2(I; \mathbf{C})$ vers une limite notée v .

Maintenant, on peut passer à la limite, quand n tend vers $+\infty$, dans l'égalité

$$u_n(x) = u_n(x_0) + \int_{x_0}^x u'_n(t) dt,$$

pour conclure que

$$u(x) = u(x_0) + \int_{x_0}^x v(t) dt,$$

ce qui montre que $u \in H^1(I; \mathbf{C})$. On peut donc identifier u' et v , ce qui montre que la suite $(u_n)_{n \geq 0}$ converge dans $H^1(I; \mathbf{C})$ vers u . \square

Nous définissons sur $H^1(I; \mathbf{C})$ la norme

$$\|u\|_{H^1(I; \mathbf{C})} := \left(\int_I (|u'(t)|^2 + |u(t)|^2) dt \right)^{1/2},$$

associée au produit hermitien $\langle \cdot, \cdot \rangle_{H^1(I; \mathbf{C})}$.

Nous avons vu que toute fonction $u \in H^1(I; \mathbf{C})$ avait un représentant continu. On montre que u se prolonge par continuité à $\bar{I} = [a, b]$. Ce résultat est l'objet du :

Lemme 11.7. *Une fonction $u \in H^1(I; \mathbf{C})$ se prolonge par continuité en une fonction définie sur $\bar{I} := [a, b]$. De plus, il existe une constante $C > 0$ (qui ne dépend que de $b - a$) telle que, pour toute $u \in H^1(I; \mathbf{C})$*

$$\|u\|_{L^\infty(I; \mathbf{C})} \leq C \|u\|_{H^1(I; \mathbf{C})}.$$

DÉMONSTRATION. Le Théorème de la convergence dominée permet d'affirmer que

$$\lim_{x \rightarrow b} u(x) = u(x_0) + \lim_{x \rightarrow b} \int_{x_0}^x u'(t) dt = u(x_0) + \int_{x_0}^b u'(t) dt,$$

ce qui montre que u peut être prolongée par continuité sur $[a, b]$.

Nous avons déjà remarqué que l'inégalité de Cauchy-Schwarz nous permet d'écrire

$$(11.2) \quad |u(x) - u(x_0)| \leq \|u'\|_{L^2} |x - x_0|^{1/2} \leq |b - a|^{1/2} \|u'\|_{L^2}.$$

Il nous reste à évaluer $u(x_0)$. Pour ce faire, on peut écrire, grâce à l'inégalité ci-dessus et grâce à l'inégalité triangulaire

$$|u(x_0)|^2 \leq 2(|u(x)|^2 + (b - a) \|u'\|_{L^2}^2).$$

En intégrant cette inégalité sur I et en utilisant le fait que I est un intervalle borné, on conclut que

$$(b - a) |u(x_0)|^2 \leq 2(\|u\|_{L^2}^2 + (b - a)^2 \|u'\|_{L^2}^2).$$

Donc, il existe une constante $C_1 > 0$, qui ne dépend que de $b - a$, telle que

$$|u(x_0)| \leq C_1 (\|u\|_{L^2} + \|u'\|_{L^2}).$$

En utilisant cette inégalité ainsi que (11.2), on conclut que

$$|u(x)| \leq C_2 (\|u\|_{L^2} + \|u'\|_{L^2}),$$

pour une constante $C_2 > 0$ qui ne dépend que de $b - a$, ce qui termine la démonstration. \square

L'application

$$H^1(I; \mathbf{C}) \rightarrow L^2(I; \mathbf{C})$$

$$u \mapsto u,$$

est appelée *injection de Sobolev de H^1 dans L^2* . C'est clairement une application continue.

De même, l'application

$$H^1(I; \mathbf{C}) \rightarrow L^\infty(I; \mathbf{C})$$

$$u \mapsto u,$$

est appelée *injection de Sobolev de H^1 dans L^∞* . Le résultat du Lemme 11.7 nous assure que c'est aussi une application continue.

Dans le Lemme suivant, nous démontrons que ces deux injections sont en fait des applications compactes.

Proposition 11.5 (Compacité des injections de Sobolev). *Les injections canoniques définies de $H^1(I; \mathbf{C}) \rightarrow L^\infty(I; \mathbf{C})$ et de $H^1(I; \mathbf{C}) \rightarrow L^2(I; \mathbf{C})$ sont compactes.*

DÉMONSTRATION. On commence par remarquer que l'injection canonique

$$\begin{aligned} L^\infty(I; \mathbf{C}) &\rightarrow L^2(I; \mathbf{C}) \\ u &\mapsto u, \end{aligned}$$

est continue. Étant donné que la composée d'une application linéaire compacte et d'une application continue est compacte, il suffit donc de démontrer que l'injection $H^1(I; \mathbf{C}) \rightarrow L^\infty(I; \mathbf{C})$ est compacte.

Soit $(u_n)_{n \geq 0}$ une suite de la boule unité fermée de $H^1(I; \mathbf{C})$. Par définition, il existe une constante $C > 0$ telle que

$$|u_n(x)| \leq C \|u_n\|_{H^1} = C,$$

et

$$|u_n(y) - u_n(x)| \leq C \|u_n\|_{H^1} |y - x|^{1/2} = C |y - x|^{1/2},$$

pour tous $x, y \in I$. On peut alors utiliser le Théorème d'Ascoli (voir l'exemple 4.6 dont on peut modifier facilement les arguments) pour démontrer directement que l'on peut extraire de la suite $(u_n)_{n \geq 0}$ une sous suite qui converge uniformément sur $[a, b]$. Ce qui montre que l'injection $H^1(I; \mathbf{C}) \rightarrow L^\infty(I; \mathbf{C})$ est bien une application compacte.

Donnons de ce résultat une autre démonstration qui indépendante du Théorème d'Ascoli. Pour simplifier les notations, supposons que $I =]a, b[=]-1, 1[$. On peut en utilisant le Théorème 10.9 page 214 extraire de la suite $(u_n)_{n \geq 0}$ une sous-suite telle que $u'_{\varphi(n)} \rightharpoonup v$ dans $L^2(I; \mathbf{C})$ et $(u_{\varphi(n)}(0))_{n \geq 0}$ converge vers $c \in \mathbf{R}$. Remarquons que (par exemple si $x \geq 0$)

$$u_n(x) = u_n(0) + \int_{[0, x]} u'_n(t) dt = u_n(0) + \langle u'_n, \mathbf{1}_{[0, x]} \rangle_{L^2},$$

ce qui nous permet de conclure que la suite $(u_{\varphi(n)})_{n \geq 0}$ converge simplement vers $u \in H^1(I; \mathbf{C})$ définie par

$$u(x) := c + \int_0^x v(t) dt, \quad \forall x \in [0, 1].$$

Fixons $\varepsilon > 0$ et choisissons $k \in \mathbf{N} - \{0\}$ tel que $4 \leq \varepsilon \sqrt{k}$. Choisissons $n_0 > 0$ tel que, pour tout $n \geq n_0$ et

$$|u_{\varphi(n)}(\frac{j}{k}) - u(\frac{j}{k})| \leq \varepsilon/2,$$

pour tout $j = -k, \dots, 0, \dots, k$. Alors, pour tout $x \in I$, il existe $j \in \{-k, \dots, 0, \dots, k-1\}$ tel que $x \in [\frac{j}{k}, \frac{j+1}{k}]$. En particulier

$$\begin{aligned} |u_n(x) - u(x)| &\leq |u_n(x) - u_n(\frac{j}{k})| + |u_n(\frac{j}{k}) - u(\frac{j}{k})| + |u(\frac{j}{k}) - u(x)| \\ &\leq \frac{2}{\sqrt{k}} + |u_n(\frac{j}{k}) - u(\frac{j}{k})| \\ &\leq \frac{\varepsilon}{2} + |u_n(\frac{j}{k}) - u(\frac{j}{k})|. \end{aligned}$$

On conclut que, pour tout $n \geq n_0$ et pour tout $x \in [0, 1]$

$$|u_{\varphi(n)}(x) - u(x)| \leq \varepsilon,$$

ce qui termine la démonstration. □

Remarquons que, grâce au Lemme 11.7, les fonctions de $H^1(I; \mathbf{C})$ se prolongent par continuité à \bar{I} et l'on peut donc parler de valeur de ces fonctions au points a et b . Un sous-espace de $H^1(I; \mathbf{C})$ sera utilisé dans les applications. On note

$$H_0^1(I; \mathbf{C}) := \{u \in H^1(I; \mathbf{C}) : u(a) = u(b) = 0\},$$

le sous-espace de $H^1(I; \mathbf{C})$ constitué des fonctions qui prennent la valeur 0 au bord de I .

Lemme 11.8. *Le sous-espace $H_0^1(I; \mathbf{C})$ de $H^1(I; \mathbf{C})$ est fermé. En particulier, muni du produit hermitien de $H^1(I; \mathbf{C})$, c'est un espace de Hilbert.*

DÉMONSTRATION. Pour vérifier que $H_0^1(I; \mathbf{C})$ est un sous-espace fermé de $H^1(I; \mathbf{C})$, il suffit de remarquer que, grâce au Lemme 11.7, l'application linéaire

$$H^1(I; \mathbf{C}) \ni u \mapsto (u(a), u(b)) \in \mathbf{C}^2,$$

est continue. En particulier, $H_0^1(I; \mathbf{C})$ est l'image réciproque de $\{(0, 0)\}$ par cette application, c'est donc un fermé de $H^1(I; \mathbf{C})$. \square

Exemple 11.2. Soit $\alpha \in \mathbf{R} - \{0\}$. On vérifie que la fonction $x \mapsto 1 - |x|^\alpha$ appartient à $H_0^1(I; \mathbf{C})$ si et seulement si $\alpha > 1/2$.

Proposition 11.6. *L'espace $\mathcal{C}^1(\bar{I}; \mathbf{C})$, où $\bar{I} = [a, b]$, est dense dans $H^1(I)$. De même, l'espace $\mathcal{C}_c^1(I; \mathbf{C})$ des fonctions de classe \mathcal{C}^1 sur I qui sont à support compact dans I , est dense dans $H_0^1(I; \mathbf{C})$.*

DÉMONSTRATION. Il suffit simplement d'utiliser la densité de $\mathcal{C}_c(I)$ dans $L^2(I)$ pour approcher u' pour la norme L^2 . \square

Nous aurons besoin du résultat suivant dont l'énoncé demande une attention particulière.

Proposition 11.7. *Soit $u \in L^2(I; \mathbf{C})$. On suppose qu'il existe $C > 0$ telle que*

$$\left| \int_I u(t) v'(t) dt \right| \leq C \|v\|_{L^2}, \quad \forall v \in \mathcal{C}_c^1(I; \mathbf{C}),$$

alors $u \in H^1(I; \mathbf{C})$.

DÉMONSTRATION. Pour tout $v \in \mathcal{C}_c^1(I; \mathbf{C})$, on pose

$$\Lambda(v) := - \int_I u(t) v'(t) dt.$$

Par hypothèse, il existe une constante $C > 0$ telle que

$$|\Lambda(v)| \leq C \|v\|_{L^2}.$$

Étant donné que $\mathcal{C}_c^1(I; \mathbf{C})$ est dense dans $L^2(I; \mathbf{C})$, on peut, grâce au Théorème 11.1 page 221, prolonger la forme linéaire Λ en une forme linéaire continue sur $L^2(I; \mathbf{C})$. Le Théorème de représentation de Riesz nous permet d'affirmer qu'il existe $h \in L^2(I; \mathbf{C})$ telle que

$$\Lambda(v) = - \int_I h(t) v(t) dt,$$

pour tout $v \in L^2(I, \mathbf{C})$. On définit

$$H(x) := \int_0^x h(t) dt.$$

Alors, par construction

$$\int_I (u(t) - H(t)) v'(t) dt = 0,$$

pour toute $v \in \mathcal{C}_c^1(I; \mathbf{C})$. Autrement dit la fonction $u - H$ est orthogonale à $\text{Vect}\{1\}^\perp$. Donc $u - H$ est une fonction constante, i.e. $u \in H^1(I; \mathbf{C})$ et $u' = h$ p.p. sur $[0, 1]$. \square

Les fonctions de $H^1(I; \mathbf{C})$ ne sont pas dérivables néanmoins, la formule d'intégration par parties reste valable pour les fonctions de cet espace. Plus précisément, nous avons le :

Lemme 11.9 (Intégration par parties). *Soient $u, v \in H^1(I; \mathbf{C})$, alors $uv \in H^1(I; \mathbf{C})$ et $(uv)' = u'v + uv'$. De plus*

$$(11.3) \quad \int_{[x,y]} (u'(t)v(t) + u(t)v'(t)) dt = (uv)(y) - (uv)(x).$$

pour tous $x, y \in [a, b]$, avec $x < y$.

DÉMONSTRATION. Si $u, v \in \mathcal{C}^1(I; \mathbf{C})$, on peut écrire

$$u(y)v(y) = u(x)v(x) + \int_x^y (u'(t)v(t) + u(t)v'(t)) dt,$$

pour tout $x \in I$. Pour étendre cette formule aux fonctions de $H^1(I; \mathbf{C})$ il suffit d'utiliser la densité des fonctions de $\mathcal{C}^1(\bar{I}; \mathbf{C})$ dans $H^1(I; \mathbf{C})$. On sait qu'il existe une suite $(u_n)_{n \geq 0}$ de fonctions de classe $\mathcal{C}^1(\bar{I}; \mathbf{C})$ telle que

$$\lim_{n \rightarrow +\infty} \|u - u_n\|_{H^1} = 0.$$

En utilisant le résultat du Lemme 11.7, on a aussi

$$\lim_{n \rightarrow +\infty} \|u - u_n\|_{L^\infty} = 0.$$

On peut maintenant facilement passer à la limite $n \rightarrow +\infty$ dans

$$u_n(y)v_n(y) = u_n(x)v_n(x) + \int_x^y (u'_n(t)v_n(t) + u_n(t)v'_n(t)) dt,$$

pour trouver 11.3. Enfin, on remarque que $u'v + uv' \in L^2(I; \mathbf{C})$ comme combinaison linéaire de produits de fonctions bornées par des fonctions de $L^2(I; \mathbf{C})$. \square

Attention, dans l'égalité $(uv)' = u'v + uv'$, les fonctions $(uv)'$, u' et v' sont les dérivées au sens faible des fonctions uv , u et v . En particulier, le résultat précédent nous assure que, si $u \in H^1(I; \mathbf{R})$ alors $u^2 \in H^1(I; \mathbf{R})$. Plus généralement, nous avons le résultat suivant dont la démonstration est laissée en exercice au lecteur (utiliser les arguments de la démonstration précédente) :

Lemme 11.10 (Composition). *Soient f une fonction de classe C^1 définie sur \mathbf{R} et $u \in H^1(I; \mathbf{R})$. Alors $f \circ u \in H^1(I; \mathbf{R})$ et*

$$(f \circ u)' = (f' \circ u) u'.$$

6. Spectre de l'opérateur de Sturm-Liouville

Les résultats de ce chapitre et ceux du chapitre précédent permettent d'étudier certaines équations différentielles ordinaires connues sous le nom *d'équations de Sturm-Liouville*.

Soit $t \mapsto q(t)$ une fonction continue sur $[0, 1]$ à valeurs réelles, plus généralement on pourra supposer que $q \in L^\infty([0, 1])$. On cherche les couples (u, λ) tels que $\lambda \in \mathbf{C}$ et u est une solution non identiquement nulle de l'équation différentielle ordinaire

$$(11.4) \quad \begin{cases} -u'' + q u = \lambda u & \text{sur } [0, 1] \\ u(0) = u(1) = 0, \end{cases}$$

où $'$ désigne la dérivation par rapport à la variable $t \in [0, 1]$. Dans ce cas, on dira que u est une *fonction propre* et que λ est une valeur propre de « l'opérateur de Sturm-Liouville » $u \mapsto -u'' + q u$ agissant sur l'ensemble des fonctions qui s'annulent au bord de l'intervalle I .

De telles équations apparaissent dans de nombreux modèles comme par exemple, dans l'étude de l'équation de Schrödinger stationnaire pour une particule quantique dans un champ dérivant du potentiel $t \mapsto q(t)$, nous renvoyons au cours de Mécanique quantique [2] pour plus de détails). La *fonction d'onde* $(x, t) \mapsto \psi(x, t)$ d'une particule quantique soumise à un potentiel $x \mapsto q(x)$, dépendant d'une variable unidimensionnelle $x \in \mathbf{R}$, vérifie l'équation de Schrödinger

$$-i\hbar \frac{\partial \psi}{\partial t}(x, t) = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2}(x, t) + q(x) \psi(x, t),$$

où $t > 0$ et $x \in \mathbf{R}$. Si l'on cherche la fonction d'onde ψ sous la forme

$$\psi(x, t) = \sum_{k \geq 0} \psi_k(x) e^{i \frac{E_k}{\hbar} t},$$

on constate que les fonctions $x \mapsto \psi_k(x)$ vérifient alors l'équation de Schrödinger stationnaire

$$(11.5) \quad -\frac{\hbar^2}{2m} \frac{d^2 \psi_k}{dx^2} + q \psi_k = E_k \psi_k,$$

où les constantes E_k représentent les *niveaux d'énergie* de la particule soumise au potentiel q . Une autre manière de voir les choses est de considérer que l'on recherche une base hilbertienne de $L^2(\mathbf{R}; \mathbf{C})$ telle que, si l'on décompose ψ sur cette base, l'équation de Schrödinger prenne une forme simple.

Or, si $(\psi_k)_{k \geq 0}$ est une telle base hilbertienne, on peut décomposer

$$\psi(x, t) = \sum_{k \geq 0} \psi_k(x) a_k(t),$$

et l'équation satisfaite par ψ s'écrit alors

$$\left(-\frac{\hbar^2}{2m} \frac{d^2 \psi_k}{dx^2}(x) + q(x) \psi_k(x) \right) a_k(t) = i \hbar \psi_k(x) \frac{da_k}{dt}(t).$$

Si les ψ_k vérifient l'équation de Schrödinger stationnaire, cette équation se ramène à

$$i \hbar \frac{da_k}{dt}(t) = E_k a_k(t).$$

Autrement dit

$$a_k(t) = a_k(0) e^{i \frac{E_k}{\hbar} t}.$$

Si l'on imagine que la particule est 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 le potentiel q

par le potentiel \tilde{q} tel que $\tilde{q}(x) = q(x)$ si $x \in [0, 1]$ et $\tilde{q}(x) = +\infty$ si $x \notin [0, 1]$. On admettra que ceci est équivalent à regarder les solutions de l'équation de Schrödinger sur $[0, 1]$ et à ajouter les conditions au bord $\psi(0, t) = \psi(1, t) = 0$. On est donc ramené à l'étude d'un problème de type Sturm-Liouville.

Grâce aux résultats démontrés dans ce chapitre et dans le chapitre précédent, nous disposons du :

Théorème 11.4. *Soit q une fonction continue sur $[0, 1]$. Les couples solutions (u, λ) de l'équation de Sturm-Liouville (11.4) forment une suite dénombrable $((u_n, \lambda_n))_{n \geq 0}$. La suite $(\lambda_n)_{n \geq 0}$ est une suite de réels qui tend vers $+\infty$ et la suite des fonctions $(u_n)_{n \geq 0}$ forme une base hilbertienne de $L^2([0, 1[; \mathbf{C})$. De plus, pour tout $n \geq 0$, u_n est de classe \mathcal{C}^2 sur $[0, 1]$ et est solution de*

$$-u_n'' + q u_n = \lambda_n u_n,$$

avec les conditions aux limites $u_n(0) = u_n(1) = 0$.

Remarque 11.3. Les λ_n étant réels, on peut supposer les fonctions u_n à valeurs réelles, car si u est solution de l'équation de Sturm-Liouville, il en est de même pour sa partie réelle $\Re u$ et sa partie imaginaire $\Im u$.

Exemple 11.3. Remarquons que, dans le cas où la fonction $q \equiv 0$, on a explicitement

$$u_n(t) = \sqrt{2} \sin(\pi n t),$$

et $\lambda_n = \pi^2 n^2$. On retrouve alors le fait que la suite $(\sqrt{2} \sin(\pi n \cdot))_{n \geq 1}$ est une base hilbertienne de $L^2([0, 1[; \mathbf{C})$.

DÉMONSTRATION. Nous allons supposer $1 \leq q(t)$ pour tout $t \in [0, 1]$, ce qui ne restreint pas la généralité du problème, car ajouter une constante c à la fonction q revient à décaler les valeurs de λ de la même constante. En effet, si (u, λ) est solution de (11.4), alors $(u, \lambda + c)$ est solution du même problème avec q remplacé par $q + c$. Comme mentionné ci-dessus, les fonctions propres sont à valeurs réelles, on se restreint donc ici aux espaces de fonctions à valeurs dans \mathbf{R} .

On définit sur $H_0^1([0, 1[)$ le produit hermitien

$$\langle u, v \rangle_q := \int_{[0, 1]} (u' v' + q u v) dt.$$

Remarquons que les normes $u \mapsto \|u\|_{H^1}$ et $u \mapsto \sqrt{\langle u, u \rangle_q}$ sont des normes équivalentes sur $H_0^1([0, 1[)$ car

$$\|u\|_{H^1([0, 1])}^2 \leq \sqrt{\langle u, u \rangle_q} \leq \|q\|_{L^\infty([0, 1])} \|u\|_{H^1([0, 1])}^2.$$

C'est pour démontrer l'inégalité de gauche que l'on utilise l'hypothèse $q \geq 1$. On peut donc, de manière interchangeable munir l'espace $H_0^1([0, 1[)$ du produit hermitien $\langle \cdot, \cdot \rangle$ ou $\langle \cdot, \cdot \rangle_q$.

Donnons nous $v \in L^2([0, 1[)$ et définissons

$$\Lambda(w) := \int_{[0, 1]} w v dt.$$

On vérifie, en utilisant l'inégalité de Cauchy-Schwarz, que Λ est une forme linéaire continue sur $H_0^1([0, 1[)$. Nous pouvons donc appliquer le Théorème de représentation de Riesz

(Théorème 10.3 page 198) dans l'espace $H_0^1(]0, 1[)$ muni de la forme sesquilinéaire $\langle \cdot, \cdot \rangle_q$. On obtient l'existence d'une fonction $u \in H_0^1(]0, 1[)$ telle que

$$\langle u, w \rangle_q = \Lambda(w), \quad \forall w \in H_0^1(]0, 1[).$$

Autrement dit

$$\int_{[0,1]} (u' w' + q u w) dt = \int_{[0,1]} v w dt, \quad \forall w \in H_0^1(]0, 1[).$$

En particulier, on déduit de la Proposition 11.7 que $u' \in H^1(]0, 1[)$ et par conséquent, on peut écrire que

$$\int_{[0,1]} (-u'' + q u - v) w dt = 0,$$

pour toute $w \in H_0^1(]0, 1[)$ et

$$\|u\|_{H^1(]0,1[)}^2 \leq \langle u, u \rangle_q = \Lambda(u) \leq \|v\|_{L^2(]0,1[)} \|u\|_{L^2(]0,1[)}.$$

Nous avons donc démontré que, pour tout $v \in L^2(]0, 1[)$, il existe une (unique) fonction $u \in H_0^1(]0, 1[)$ tel que

$$-u'' + q u = v, \quad \text{p.p. sur } [0, 1].$$

De plus $u' \in H^1(]0, 1[)$ et

$$\|u\|_{H^1(]0,1[)} + \|u'\|_{H^1(]0,1[)} \leq C \|v\|_{L^2(]0,1[)},$$

pour une constante $C > 0$ qui ne dépend pas de v (ni de u).

Nous pouvons donc définir l'application linéaire

$$\begin{aligned} A : L^2(]0, 1[) &\mapsto L^2(]0, 1[) \\ v &\mapsto u. \end{aligned}$$

D'après la Proposition 11.5 page 233, l'injection $H^1(]0, 1[) \rightarrow L^2(]0, 1[)$ est compacte, on conclut donc que A est une application compacte. De plus, le Lemme 11.3 page 236 implique que l'opérateur A est hermitien et l'on vérifie qu'il est positif grâce au fait que

$$\langle A(v), v \rangle_{L^2} = \langle v, v \rangle_q \geq 0.$$

On peut alors appliquer le Théorème 11.3 pour conclure que les valeurs propres de A forment une suite décroissante $(\alpha_n)_{n \geq 0}$ qui tend vers 0 et qu'il existe $(v_n)_{n \geq 0}$ une base hilbertienne de $L^2(]0, 1[)$ formée de fonctions propres de A .

Remarquons que, si $A(v_n) = \alpha_n v_n$, alors $u_n = A(v_n)$ appartient à $H_0^1([0, 1])$, $u'_n \in H^1(]0, 1[)$ et u_n est solution de

$$-u_n'' + q u_n = \frac{1}{\alpha_n} u_n \quad \text{avec} \quad u_n(0) = u_n(1) = 0.$$

En particulier, $u_n \in C^2([0, 1])$. En effet, $u_n \in H^1(]0, 1[)$ donc $u_n \in \mathcal{C}([0, 1])$ et donc, u_n'' est continue sur $[0, 1]$. Il en résulte que $u_n \in \mathcal{C}^2([0, 1])$, ce qui termine la démonstration. \square

Une remarque s'impose : l'opérateur qui apparaît naturellement dans le problème de Sturm-Liouville est l'opérateur $Lu := -u'' + q u$. Il serait tout naturel de travailler directement avec cet opérateur pour en déterminer les fonctions propres et les valeurs propres. Malheureusement, cet opérateur n'est pas défini sur $L^2(]0, 1[)$, ni sur $H^1(]0, 1[)$! L'opérateur L est un exemple d'opérateur *non borné* (voir [5] à ce sujet). De tels opérateurs apparaissent naturellement en physique (par exemple l'opérateur de position $f(x) \mapsto x f(x)$ ou l'opérateur

d'impulsion $f(x) \mapsto f'(x)$ que nous avons déjà mentionnés qui, par exemple, ne sont pas non plus bien définis sur $L^2(\mathbf{R})$. Dans la démonstration ci-dessus, nous avons contourné le problème remplaçant L par son « inverse » A ce qui permet de travailler avec un opérateur qui lui est bien défini de $L^2([0, 1])$ dans lui même.

Au passage, on remarque que, dans la démonstration du résultat précédent, nous avons démontré le :

Lemme 11.11. *On suppose que $q \in \mathcal{C}([0, 1])$ est une fonction strictement positive. Alors, pour toute $v \in L^2([0, 1])$, il existe une unique fonction $u \in H_0^1([0, 1])$ solution de*

$$-u'' + qu = v, \quad p.p. \text{ sur } [0, 1],$$

et telle que $u' \in H^1([0, 1])$.

Grâce à l'obtention d'une base hilbertienne $(u_n)_{n \geq 0}$ associée à l'opérateur

$$u \mapsto -u'' + qu,$$

il est maintenant facile de résoudre l'équation

$$-u'' + qu = v,$$

avec comme conditions aux limites $u(0) = 0, u(1) = 0$. En effet, si $v \in L^2([0, 1])$, on peut décomposer v sur les fonctions propres $(u_n)_{n \geq 0}$

$$v(t) = \sum_{n \geq 0} u_n v_n(t),$$

et nous pouvons alors définir la fonction

$$u(t) := \sum_{n \geq 0} \frac{1}{\lambda_n} v_n u_n(t),$$

qui sera solution de notre équation pourvu que les termes de la série décroissent suffisamment vite pour assurer que $u \in \mathcal{C}^2$. Cette formule peut se réécrire sous la forme

$$u(t) = \int_{[0, 1]} K(t, s) v(s) ds,$$

où par définition

$$K(t, s) := \sum_{n \in \mathbf{N}} \frac{1}{\lambda_n} u_n(s) u_n(t).$$

Le lecteur pourra comparer cette formule avec la formule qui apparaît dans la Remarque 11.2 page 230.

La démonstration ci-dessus, bien que dans un cadre élémentaire, met en place les principes généraux et les principaux outils permettant de résoudre les *équations aux dérivées partielles linéaires* : Théorème de représentation 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.

De nombreuses bases hilbertiennes sont obtenues à partir des fonctions propres de systèmes d'équations différentielles. Cependant, il s'agit souvent d'équations d'apparence un peu plus compliquée que celles que nous avons présentées ci-dessus. Par exemple,

$$(pu')' + qu = \lambda ru,$$

où p , q et r sont des fonctions continues. Pourtant, la théorie est la même que celle présentée ci-dessus. Il suffit simplement de définir

$$\langle f, g \rangle_{p,q} := \int_{[0,1]} (p(t) f'(t) g'(t) + q(t) f(t) g(t)) dt,$$

et

$$\Lambda_r(f, g) := \int_{[0,1]} f(t) g(t) r(t) dt.$$

Pourvu que p et r soient des fonctions continues et *strictement positives* sur $]0, 1[$, la théorie n'est pas modifiée et les résultats restent les mêmes.

Exemple 11.4. Les solutions de l'équation

$$-((1 - t^2)y')' = c y,$$

correspondent aux polynômes de Legendre. Ces polynômes forment un système orthogonal de $L^2([-1, 1]; \mathbf{C})$ muni du produit hermitien

$$(f, g) := \int_{[-1,1]} f(t) \overline{g(t)} dt.$$

Exemple 11.5. Les solutions de l'équation

$$-(ty'' + (1 - t)y') = c y,$$

correspondent aux polynômes de Laguerre, qui forment un système orthogonal de l'espace $L^2([0, +\infty[, e^{-x} dx)$ muni du produit hermitien

$$(f, g) := \int_{[0,+\infty[} f(t) \overline{g(t)} e^{-t} dt.$$

7. Harmoniques sphériques

Dans cette section nous expliquons comment les harmoniques sphériques, qui sont utilisées par exemple en physique dans la modélisation de l'atome d'hydrogène, apparaissent de manière naturelle comme base hilbertienne de l'espace des fonctions de carré intégrable sur la sphère unité de \mathbf{R}^N . Nous laissons le soin au lecteur de compléter les détails des différentes démonstrations.

Commençons par définir les harmoniques sphériques. On note \mathbf{P}_d le sous-espace des fonctions polynômes à N variables qui sont homogènes de degré $d \geq 0$ i.e. les fonctions polynômes de \mathbf{P}_d sont de la forme

$$P(x_1, \dots, x_N) = \sum_{p_1 + \dots + p_N = d} a_{p_1 \dots p_N} x_1^{p_1} \dots x_N^{p_N}$$

où les coefficients $a_{p_1 \dots p_N} \in \mathbf{R}$. On convient que $\mathbf{P}_d := \{0\}$ si $d < 0$.

Le Laplacien sur \mathbf{R}^N est défini par

$$\Delta := \sum_{j=1}^N \frac{\partial^2}{\partial x_j^2}.$$

On vérifie que l'image par Δ d'une fonction polynôme de degré d est une fonction polynôme de degré $d - 2$. De plus si P est une fonction polynôme homogène de degré d alors ΔP est une fonction polynôme homogène de degré $d - 2$. Autrement dit

$$\Delta : \mathbf{P}_d \rightarrow \mathbf{P}_{d-2}.$$

Enfin, nous définissons $\mathbf{H}_d := \{P \in \mathbf{P}_d : \Delta P = 0\}$ le sous-espace de \mathbf{P}_d formé des fonctions polynômes homogènes de degré d qui sont *harmoniques*.

Pour tout $N \geq 2$, on note $S^{N-1} := \{x \in \mathbf{R}^N : |x| = 1\}$ la sphère unité de \mathbf{R}^N . Nous avons alors la :

Définition 11.3. *Les harmoniques sphériques sont les restrictions à la sphère unité S^{N-1} des fonctions de \mathbf{H}_d .*

On notera $\tilde{\mathbf{H}}_d$, l'espace des restrictions à S^{N-1} des fonctions de \mathbf{H}_d . Le lecteur a sans doute déjà rencontré en physique l'expression du Laplacien en coordonnées sphériques. En dimension N , on a la formule

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{N-1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \Delta_{S^{N-1}},$$

où $r^2 = x_1^2 + \dots + x_N^2$. Cette formule généralise en toute dimension la formule bien connue, en dimension $N = 2$, du laplacien en coordonnées polaires

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2},$$

alors qu'en dimension $N = 3$, on a explicitement

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right),$$

où (θ, φ) sont les coordonnées sphériques définies en (6.1).

On dispose du résultat suivant qui montre que les restrictions à S^{N-1} des fonctions de \mathbf{H}_d font des fonctions propres de $\Delta_{S^{N-1}}$.

Lemme 11.12. *Si $P \in \mathbf{H}_d$ alors la restriction de P à S^{N-1} vérifie*

$$\Delta_{S^{N-1}} P = -d(d + N - 2) P_d.$$

DÉMONSTRATION. Il suffit d'utiliser le fait que

$$r \frac{\partial}{\partial r} P(x_1, \dots, x_N) = d P(x_1, \dots, x_N),$$

car P est homogène de degré d . □

Soit χ une fonction radiale, positive, à support compact dans $B(0, 2) - B_f(0, 1/2)$ telle que $\|\chi\|_{L^2(\mathbf{R}^N)} = 1$. On peut définir $L^2(S^{N-1})$ comme l'espace des fonctions $f : S^{N-1} \rightarrow \mathbf{R}$ telles que $x \mapsto \chi(x)f(x/|x|)$ appartient à $L^2(\mathbf{R}^N)$. Le produit hermitien (ou produit scalaire) dans $L^2(S^{N-1})$ peut alors être défini² par

$$\langle f, g \rangle_{L^2} := \int_{\mathbf{R}^N} \chi^2 f g dx,$$

2. Il existe bien entendu d'autres définitions directes du produit hermitien dans $L^2(S^{N-1})$ mais elles demandent de définir au préalable une *forme volume* sur S^{N-1} .

et la norme associée est donnée par

$$\|f\|_{L^2(S^{N-1})} := \|\chi f\|_{L^2(\mathbf{R}^N)}.$$

On vérifie que

$$(11.6) \quad \langle \Delta_{S^{N-1}} f, g \rangle_{L^2} = \langle f, \Delta_{S^{N-1}} g \rangle_{L^2},$$

pour toutes les fonctions f et g qui sont définies sur S^{N-1} et pour lesquelles χf et χg sont de classe \mathcal{C}^2 sur \mathbf{R}^N .

Nous avons le :

Théorème 11.5. *Il existe une base hilbertienne de $L^2(S^{N-1})$ qui est constituée de la réunion des bases hilbertiennes des $\tilde{\mathbf{H}}_d$, pour $d \in \mathbf{N}$. Autrement dit, il existe une base hilbertienne de $L^2(S^{N-1})$ dont les éléments sont des fonctions propres de $\Delta_{S^{N-1}}$.*

DÉMONSTRATION. On rappelle que $r^2 = |x|^2$. On commence par vérifier que

$$\Delta(r^{2k+2} \mathbf{H}_d) = r^{2k} \mathbf{H}_d,$$

pour tout $k \in \mathbf{N}$, i.e. l'image de $r^{2k+2} \mathbf{H}_d$ par l'opérateur Δ est égal à $r^{2k} \mathbf{H}_d$. En effet, si $P \in \mathbf{H}_d$, la fonction polynôme P est harmonique et on a donc

$$\Delta(r^{2k+2} P) = (2k+2)(N+2k+2d) r^{2k} P,$$

ce qui démontre l'assertion.

Grâce à ce résultat, on montre par récurrence sur $d \geq 0$ que

$$(11.7) \quad \mathbf{P}_d = \mathbf{H}_d \oplus r^2 \mathbf{H}_{d-2} \oplus r^4 \mathbf{H}_{d-4} \oplus \dots \oplus r^{2[d/2]} \mathbf{H}_{d-2[d/2]},$$

Supposons que le résultat soit vrai pour d . On utilise d'une part le fait que l'opérateur $\Delta : \mathbf{P}_{d+2} \rightarrow \mathbf{P}_d$ a pour noyau \mathbf{H}_{p+2} et d'autre part, le fait que nous venons de montrer que

$$\Delta : r^{2k+2} \mathbf{H}_{d+2} \rightarrow r^{2k} \mathbf{H}_d,$$

est une application bijective. Si $P \in \mathbf{P}_{d+2}$ alors $\Delta P \in \mathbf{P}_d$ et, par hypothèse de récurrence, il existe $Q \in \mathbf{H}_d \oplus r^2 \mathbf{H}_{d-2} \oplus r^4 \mathbf{H}_{d-4} \oplus \dots \oplus r^{2[d/2]} \mathbf{H}_{d-2[d/2]}$, tel que $\Delta(r^2 Q) = \Delta P$. En particulier, $\Delta(P - r^2 Q) = 0$, donc $P - r^2 Q \in \mathbf{H}_{d+2}$. Ce qui démontre déjà que tout élément de \mathbf{P}_{d+2} se décompose comme la somme d'un élément de \mathbf{H}_{d+2} et d'un élément de $r^2 \mathbf{H}_d \oplus r^4 \mathbf{H}_{d-2} \oplus \dots \oplus r^{2[d/2]+2} \mathbf{H}_{d-2[d/2]}$. Nous laissons le soin au lecteur de démontrer que ces deux espaces sont en somme directe.

Les espaces \mathbf{P}_d , pour $d \geq 0$ engendrent l'espace des polynômes et le Théorème de Stone-Weierstrass nous assure que les fonctions polynômes sont denses (pour la topologie de la convergence uniforme) dans l'espace des fonctions continues sur le compact $B_f(0, 2) - B(0, 1/2)$. En particulier, grâce à (11.7), nous pouvons affirmer que les restrictions à S^{N-1} des fonctions des \mathbf{H}_d , pour $d \geq 0$, engendrent un sous-espace dense de $\mathcal{C}(S^{N-1})$ (muni de la topologie de la convergence uniforme) donc de $L^2(S^{N-1})$.

Pour finir, on montre que, si $d \neq d'$, alors les restrictions de \mathbf{H}_d et $\mathbf{H}_{d'}$ à S^{N-1} sont orthogonales. Cette propriété est une simple conséquence de (11.6) et du Lemme 11.12 page ci-contre qui montre que les restrictions des fonctions de \mathbf{H}_d à S^{N-1} sont des fonctions propres de $\Delta_{S^{N-1}}$. En particulier, les espaces engendrés par les restrictions à S^{N-1} des fonctions de \mathbf{H}_d sont des espaces en somme directe.

Conclusion, on obtient une base hilbertienne de $L^2(S^{N-1})$ en réunissant des bases hilbertiennes des espaces des restrictions à S^{N-1} des fonctions de \mathbf{H}_d . \square

Remarque 11.4. Dans le cas où $N = 2$, on retrouve le fait que les fonctions $\theta \mapsto \cos(n\theta)$ et $\theta \mapsto \sin(n\theta)$, pour $n \in \mathbf{N}$, forment une base hilbertienne de l'espace $L^2(S^1)$.

Remarque 11.5. En dimension $N = 3$, on montre que l'espace des harmoniques sphériques de \mathbf{H}_d est de dimension $2d + 1$ qui est engendré par $Y_{-d,d}, \dots, Y_{0,d}, \dots, Y_{d,d}$ solutions de

$$\Delta_{S^2} Y_{m,d} = -d(d+1) Y_{m,d},$$

avec $d \geq 0$ et $m = -d, 1-d, \dots, 0, \dots, d-1, d$. Les premières harmoniques sphériques sont données par les formules suivantes : pour $d = 0$, on a simplement

$$\psi_{0,0}(x, y, z) = \frac{1}{\sqrt{4\pi}},$$

ensuite, pour $d = 1$, on trouve

$$\psi_{-1,1}(x, y, z) = \sqrt{\frac{3}{4\pi}} x, \quad \psi_{0,1}(x, y, z) = \sqrt{\frac{3}{4\pi}} z, \quad \text{et} \quad \psi_{1,1}(x, y, z) = \sqrt{\frac{3}{4\pi}} y.$$

etc ... Le lecteur pourra comparer ces formules avec les formules données au Chapitre 10 de [2].

On suppose maintenant que $N = 3$. Sans entrer dans les détails, rappelons que les solutions de l'équation

$$(11.8) \quad \frac{\hbar^2}{2m} \Delta \psi + \left(E + \frac{e^2}{r} \right) \psi = 0,$$

correspondent aux *ondes stationnaires* de l'équation de Schrödinger pour l'atome d'hydrogène, c'est-à-dire

$$\Psi(x, t) := \psi(x) e^{-i \frac{E}{\hbar} t}.$$

Grâce au résultat ci-dessus, on peut décomposer les solutions de (11.8) en somme de fonctions de la forme

$$\psi(x) = f(|x|) Y \left(\frac{x}{|x|} \right),$$

où Y est une harmonique sphérique, la décomposition ayant lieu dans $L^2(S^{N-1})$, à $|x|$ fixé. Ceci nous permet de réduire la résolution de (11.8) à la recherche d'une fonction f radiale, solution d'une équation différentielle ordinaire d'ordre 2 en r . En effet, si

$$\Delta_{S^{N-1}} Y = -d(d+1) \lambda Y,$$

alors, la fonction f est solution de

$$\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} - \frac{d(d+1)}{r^2} \right) f + \left(E + \frac{e^2}{r} \right) f = 0.$$

Nous renvoyons au Chapitre 11 de [2] pour plus de détails sur l'utilisation des harmoniques sphériques dans la modélisation de l'atome d'hydrogène.

8. Introduction à la résolution de problèmes variationnels

Dans cette section, nous nous proposons de donner deux exemples de résolution d'une équation différentielle non linéaire en utilisant la *structure variationnelle* du problème. Cette section constitue donc un premier contact avec la résolution de *problèmes variationnels* en dimension infinie³, résolution qui a occupé bon nombre de mathématiciens au cours du XX^e-ième siècle (voir le cours de MAP 431 *Optimisation et analyse numérique*).

Soit $N \geq 1$. On se donne une fonction $F : [0, 1] \times \mathbf{R}^N \rightarrow [0, +\infty[$ qui est de classe \mathcal{C}^1 sur $[0, 1] \times \mathbf{R}^N$ et l'on définit le gradient de la fonction F par

$$\text{grad } F(t, x) := \left(\frac{\partial F}{\partial x_1}(t, x), \dots, \frac{\partial F}{\partial x_N}(t, x) \right),$$

pour tous $(t, x) \in [0, 1] \times \mathbf{R}^N$. On cherche à résoudre le *système* d'équations différentielles ordinaires

$$(11.9) \quad -u''(t) + \text{grad } F(t, u(t)) = 0,$$

pour tout $t \in [0, 1]$, avec comme données aux limites $u(0) = U_0$ et $u(1) = U_1$. C'est-à-dire que l'on recherche une fonction $u \in \mathcal{C}^2([0, 1]; \mathbf{R}^N)$ solution de ce système d'équations. Comme dans le paragraphe précédent, ' désigne la dérivation par rapport à la variable t .

Remarque 11.6. Remarquons que, si F est de classe \mathcal{C}^2 et si l'on se donne donnée initiale $u(0)$ et une vitesse initiale $u'(0)$, le Théorème de Cauchy-Lipschitz (voir la section 6 page 64) assure l'existence d'une solution du système (11.9) pour des temps proches de 0. Mais le problème que nous nous posons maintenant est de nature différente car nous imposons une condition initiale au temps $t = 0$ et une condition terminale au temps $t = 1$.

Nous avons la :

Proposition 11.8. *Il existe une fonction $u \in \mathcal{C}^2([0, 1]; \mathbf{R}^N)$ solution de (11.9).*

La démonstration de ce résultat se fait en deux étapes. Dans un premier temps, nous utilisons la *structure variationnelle du problème* et nous recherchons un minimiseur d'une fonctionnelle définie sur un espace de fonctions, ensuite, on démontre dans un deuxième temps que le minimiseur de cette fonctionnelle est une solution de (11.9).

Commençons par préciser les notations utilisées. On dira que $u \in H^1(]0, 1[; \mathbf{R}^N)$ si les fonctions coordonnées $t \mapsto u_1(t), \dots, t \mapsto u_N(t)$ de u , appartiennent à $H^1(]0, 1[; \mathbf{R})$. Bien entendu, l'espace $H^1(]0, 1[; \mathbf{R}^N)$, s'il est muni du produit hermitien

$$\langle u, v \rangle := \sum_{i=1}^N \langle u_i, v_i \rangle_{H^1},$$

est un espace de Hilbert (exercice). On notera

$$H_0^1(]0, 1[; \mathbf{R}^N) := \{u \in H^1(]0, 1[; \mathbf{R}^N) : u(0) = u(1) = 0\},$$

qui est un sous-espace fermé de $H^1(]0, 1[; \mathbf{R}^N)$.

Définissons, pour tout $u \in H^1(]0, 1[; \mathbf{R}^N)$ l'énergie

$$\mathcal{E}(u) := \int_{[0, 1]} \left(\frac{1}{2} |u'(t)|^2 + F(t, u(t)) \right) dt.$$

3. Ce problème fait partie des 23 « problèmes de Hilbert » qui ont été présentés par David Hilbert lors du deuxième congrès international des mathématiciens tenu à Paris en 1900.

Enfin, si $U_0, U_1 \in \mathbf{R}^N$ sont fixés, on note

$$W := \{u \in H^1([0, 1[; \mathbf{R}^N) : u(0) = U_0, \quad u(1) = U_1\},$$

qui, cette fois-ci, est un sous-espace affine de $H^1([0, 1[; \mathbf{R}^N)$.

Nous avons le résultat suivant dont la démonstration est basée sur la Proposition 10.9 page 214 et sur la Proposition 11.5 page 233 :

Lemme 11.13. *Il existe $u \in W$ tel que*

$$\mathcal{E}(u) = \inf_{w \in W} \mathcal{E}(w).$$

Autrement dit, $\inf_{w \in W} \mathcal{E}(w)$ est atteint en $u \in W$.

DÉMONSTRATION. Soit $(u_n)_{n \geq 0}$ une suite d'éléments de W qui minimise \mathcal{E} , c'est-à-dire que

$$\lim_{n \rightarrow +\infty} \mathcal{E}(u_n) = \inf_{w \in W} \mathcal{E}(w).$$

On remarque que $\inf_{w \in W} \mathcal{E}(w) < +\infty$ et que

$$\mathcal{E}(w) \geq \frac{1}{2} \|w'\|_{L^2}^2.$$

On en déduit que la suite $(u'_n)_{n \geq 0}$ est bornée dans $L^2([0, 1[; \mathbf{R}^N)$. Étant donné que $u_n(0) = U_0$, on en déduit par intégration (et utilisation de l'inégalité de Cauchy-Schwarz) que la suite $(u_n)_{n \geq 0}$ est bornée dans $H^1([0, 1[; \mathbf{R}^N)$. Exploitions maintenant le résultat de la Proposition 11.5 page 233 et celui du Théorème 10.9 page 214 pour extraire, de la suite $(u_n)_{n \geq 0}$ une sous-suite $(u_{\varphi(n)})_{n \geq 0}$ telle que

$$(11.10) \quad u_{\varphi(n)} \rightarrow u \quad \text{dans} \quad L^\infty([0, 1[; \mathbf{R}^N),$$

et

$$(11.11) \quad u'_{\varphi(n)} \rightharpoonup v \quad \text{dans} \quad L^2([0, 1[; \mathbf{R}^N).$$

Pour tout $w \in H_0^1([0, 1[; \mathbf{R}^N)$, on peut écrire

$$\int_{[0,1]} u_{\varphi(n)} \cdot w' dt = - \int_{[0,1]} u'_{\varphi(n)} \cdot w dt,$$

où \cdot désigne le produit scalaire euclidien dans \mathbf{R}^N . En passant à la limite quand n tend vers $+\infty$ et en utilisant (11.10), on conclut que

$$\int_{[0,1]} u \cdot w' dt = - \int_{[0,1]} v \cdot w dt.$$

Du coup, le Lemme 11.7 nous assure que $u \in H^1([0, 1[; \mathbf{R}^N)$ et que $v = u'$. Enfin $u \in W$ par passage à la limite dans les égalités $u_{\varphi(n)}(0) = U_0$ et $u_{\varphi(n)}(1) = U_1$ grâce à (11.10). Enfin, grâce à (11.10)

$$\int_{[0,1]} F(t, u(t)) dt = \lim_{n \rightarrow +\infty} \int_{[0,1]} F(t, u_{\varphi(n)}(t)) dt,$$

et

$$\int_{[0,1]} |u'(t)|^2 dt \leq \liminf_{n \rightarrow +\infty} \int_{[0,1]} |u'_{\varphi(n)}(t)|^2 dt.$$

d'après le résultat de la Proposition 10.10 page 214. Conclusion, nous avons montré que

$$\inf_{w \in W} \mathcal{E}(w) \leq \mathcal{E}(u) \leq \liminf_{n \rightarrow +\infty} \mathcal{E}(u_n) = \inf_{w \in W} \mathcal{E}(w).$$

Ce qui termine la démonstration. \square

Le résultat qui suit permet de caractériser les minimiseurs de \mathcal{E} et cela termine ainsi la démonstration de la Proposition 11.8.

Lemme 11.14. *On suppose que $u \in W$ et que*

$$\mathcal{E}(u) = \inf_{w \in W} \mathcal{E}(w).$$

Alors, $u \in \mathcal{C}^2([0, 1]; \mathbf{R}^N)$, u est solution de (11.9) et $u(0) = U_0$ et $u(1) = U_1$.

DÉMONSTRATION. Donnons nous $w \in H_0^1([0, 1[; \mathbf{R}^N)$. Le lecteur montrera facilement que $s \mapsto \mathcal{E}(u + sw)$ est une fonction de classe \mathcal{C}^1 et aussi que

$$\frac{d}{ds} \mathcal{E}(u + sw)|_{s=0} = \int_{[0,1]} (u' \cdot w' + \text{grad } F(t, u(t)) \cdot w(t)) \, dt.$$

Pour ce faire, on pourra remarquer que les fonctions de $H^1([0, 1[; \mathbf{R}^N)$ sont continues. Maintenant, par hypothèse

$$\mathcal{E}(u + sw) \geq \mathcal{E}(u),$$

pour tous $w \in H_0^1([0, 1[; \mathbf{R}^N)$ et $s \in \mathbf{R}$. Donc, la fonction $s \mapsto \mathcal{E}(u + sw)$ admet un minimum local en $s = 0$ et sa dérivée en ce point doit être nulle. Conclusion, nous avons montré que

$$\int_{[0,1]} (u'(t) \cdot w'(t) + \text{grad } F(t, u(t)) \cdot w(t)) \, dt,$$

pour toute fonction $w \in H_0^1([0, 1[; \mathbf{R}^N)$. Utilisons une fois de plus le Lemme 11.7 page 235 pour conclure que $u' \in H^1([0, 1[; \mathbf{R}^N)$ et finalement que

$$\int_{[0,1]} (-u''(t) + \text{grad } F(t, u(t))) \cdot w(t) \, dt = 0.$$

Cette identité étant vraie pour toute fonction $w \in H_0^1([0, 1[; \mathbf{R}^N)$, on en déduit que

$$-u''(t) + \text{grad } F(t, u(t)) = 0,$$

pour presque tout $t \in [0, 1]$. Pour terminer, nous savons que $u, u' \in H^1([0, 1[; \mathbf{R}^N)$ donc $u, u' \in \mathcal{C}([0, 1]; \mathbf{R}^N)$ et l'on en déduit que $u \in \mathcal{C}^1([0, 1]; \mathbf{R}^N)$. En exploitant l'équation vérifiée par u , on conclut que $u \in \mathcal{C}^2([0, 1]; \mathbf{R}^N)$. Ceci termine la démonstration de ce résultat. \square

9. Géodésiques sur les graphes

Donons un deuxième exemple d'utilisation des outils que nous avons développés pour démontrer l'existence de *géodésiques sur un graphe*. Commençons par décrire le problème. On se donne une fonction $h \in \mathcal{C}^2(\mathbf{R}^2; \mathbf{R})$ et on note

$$\Sigma := \{(x, h(x)) \in \mathbf{R}^3 : x \in \mathbf{R}^2\},$$

le graphe de la fonction h (c'est donc une surface paramétrée de \mathbf{R}^3).

Définition 11.4. *On dit qu'une courbe $\gamma \in \mathcal{C}^2([0, 1]; \mathbf{R}^3)$ est une géodésique de Σ qui joint le point $\sigma_0 \in \Sigma$ au point $\sigma_1 \in \Sigma$ si les conditions suivantes sont vérifiées :*

- (1) $\gamma(t) \in \Sigma$ pour tout $t \in [0, 1]$;
- (2) $\gamma(0) = \sigma_0$ et $\gamma(1) = \sigma_1$;
- (3) $\gamma''(t)$ est un vecteur normal à Σ au point $\gamma(t)$, pour tout $t \in [0, 1]$.

En utilisant le caractère variationnel de ce problème, nous pouvons utiliser les outils développés dans ce chapitre et dans le chapitre précédent pour démontrer l'existence de géodésiques tracées sur Σ qui joignent les points σ_0 et σ_1 .

Théorème 11.6. *Pour tous $\sigma_0, \sigma_1 \in \Sigma$, il existe au moins une géodésique de Σ qui joint le point σ_0 au point σ_1 .*

DÉMONSTRATION. La démonstration utilise quelques notions de calcul différentiel. Rappelons que la différentielle de h au point $x \in \mathbf{R}^2$ est l'application linéaire Dh_x définie par

$$Dh_x(y) = y_1 \frac{\partial h}{\partial x_1}(x) + y_2 \frac{\partial h}{\partial x_2}(x),$$

et la différentielle seconde de h au point x est donnée par

$$D^2h_x(y, z) = \sum_{i,j=1}^2 \frac{\partial^2 h}{\partial x_i \partial x_j}(x) y_i z_j,$$

pour tous $y, z \in \mathbf{R}^2$.

Étape 1. La première partie de la démonstration suit les grandes lignes de la démonstration du Lemme 11.13. Choisissons $x^0, x^1 \in \mathbf{R}^2$, tels que $\sigma_i = (x^i, h(x^i))$. Cette fois-ci, on définit l'espace affine

$$W := \{u \in H^1(]0, 1[; \mathbf{R}^2) : u(0) = x^0, \text{ et } u(1) = x^1\}.$$

Si $u \in W$ alors $t \mapsto \gamma(t)$ où

$$\gamma(t) := (u(t), h(u(t))),$$

est une courbe continue, tracée sur Σ , qui relie σ_0 à σ_1 . Le Lemme 11.10 page 236 nous permet d'affirmer que $h \circ u \in H^1(]0, 1[; \mathbf{R})$. On peut alors considérer l'énergie

$$\mathcal{E}(u) := \frac{1}{2} \int_{[0,1]} |\gamma'(t)|^2 dt = \frac{1}{2} \int_{[0,1]} (|u'(t)|^2 + |(h \circ u)'(t)|^2) dt,$$

qui est définie pour toute $u \in W$.

Soit $(u_n)_{n \geq 0}$ une suite minimisante de \mathcal{E} , c'est-à-dire, que $u_n \in W$ et que

$$\lim_{n \rightarrow +\infty} \mathcal{E}(u_n) = \inf_{w \in W} \mathcal{E}(w).$$

La suite $(u_n)_{n \geq 0}$ est clairement bornée dans $H^1(]0, 1[; \mathbf{R}^2)$ (en fait, ce qui est clair, c'est que la suite $(u'_n)_{n \geq 0}$ est bornée dans $L^2(]0, 1[; \mathbf{R}^2)$ mais, étant donné que $u_n(0) = x^0$, on conclut facilement que $(u_n)_{n \geq 0}$ est bornée dans $H^1(]0, 1[; \mathbf{R}^2)$. De même la suite $((h \circ u_n)')_{n \geq 0}$ est bornée dans $L^2(]0, 1[; \mathbf{R})$.

En utilisant le résultat de la Proposition 10.9 page 214 ainsi que le résultat de la Proposition 11.5 page 233, on peut extraire de la suite $(u_n)_{n \geq 0}$, une sous-suite $(u_{\varphi(n)})_{n \geq 0}$ telle que

$$u_{\varphi(n)} \rightarrow u \quad \text{dans } L^\infty(]0, 1[; \mathbf{R}^2),$$

et

$$u'_{\varphi(n)} \rightharpoonup v \quad \text{dans } L^2(]0, 1[; \mathbf{R}^2).$$

On vérifie, comme dans la démonstration précédente, que $v = u'$ par passage à la limite quand n tend vers $+\infty$ dans l'égalité

$$\int_{[0,1]} u_{\varphi(n)} \cdot w' dt = - \int_{[0,1]} u'_{\varphi(n)} \cdot w dt.$$

De plus $u(0) = x^0$ et $u(1) = x^1$ par passage à la limite dans les égalités $u_{\varphi(n)}(0) = x^0$ et $u_{\varphi(n)}(1) = x^1$. En particulier, $u \in W$.

On prétend maintenant que

$$(h \circ u_{\varphi(n)})' \rightarrow (h \circ u)' \quad \text{dans} \quad L^2([0, 1[; \mathbf{R}).$$

En effet, on dispose de l'égalité

$$\int_{[0,1]} (h \circ u_{\varphi(n)})'(t) w(t) dt = - \int_{[0,1]} (h \circ u_{\varphi(n)})(t) w'(t) dt,$$

pour toute $w \in H_0^1([0, 1[; \mathbf{R}^2)$. D'après le Lemme 11.10 page 236,

$$(h \circ u_n)' = Dh_{u_n}(u_n'),$$

où Dh_x désigne la différentielle de l'application $x \mapsto h(x)$ au point x . On sait que $u_{\varphi(n)} \rightarrow u$ dans $L^\infty([0, 1[; \mathbf{R}^2)$ et l'on sait aussi que $u_{\varphi(n)}' \rightharpoonup u'$ dans $L^2([0, 1[; \mathbf{R}^2)$ et on conclut facilement (exercice) que

$$\lim_{n \rightarrow +\infty} \int_{[0,1]} (h \circ u_{\varphi(n)})'(t) w(t) dt = \int_{[0,1]} Dh_{u(t)}(u'(t)) w(t) dt.$$

(On pourra par exemple écrire $Dh_u(u') = ((\text{grad } h) \circ u) \cdot u'$). D'autre part, $u_{\varphi(n)} \rightarrow u$ dans $L^\infty([0, 1[; \mathbf{R}^2)$ ce qui nous permet d'affirmer que

$$\lim_{n \rightarrow +\infty} \int_{[0,1]} (h \circ u_{\varphi(n)})(t) w'(t) dt = \int_{[0,1]} (h \circ u)(t) w'(t) dt.$$

Finalement,

$$\int_{[0,1]} Dh_{u(t)}(u'(t)) w(t) dt = - \int_{[0,1]} (h \circ u)(t) w'(t) dt,$$

pour toute $w \in H_0^1([0, 1[; \mathbf{R}^2)$, ce qui termine la démonstration.

Pour poursuivre, on remarque que $u_{\varphi(n)}' \rightharpoonup u'$ et que $(h \circ u_{\varphi(n)})' \rightharpoonup (h \circ u)'$ et, grâce à la Proposition 10.10 page 214, on conclut que

$$\int_{[0,1]} |u'(t)|^2 dt \leq \liminf_{n \rightarrow \infty} \int_{[0,1]} |u_{\varphi(n)}'(t)|^2 dt,$$

et

$$\int_{[0,1]} |(h \circ u)'(t)|^2 dt \leq \liminf_{n \rightarrow \infty} \int_{[0,1]} |(h \circ u_{\varphi(n)})'(t)|^2 dt.$$

Nous avons donc démontré que

$$\mathcal{E}(u) \leq \liminf_{n \rightarrow +\infty} \mathcal{E}(u_n).$$

Autrement dit, le minimum de \mathcal{E} sur W est atteint en $u \in W$.

Étape 2. La démonstration se poursuit maintenant comme dans celle du Lemme 11.14. On vérifie que, pour tout $w \in H_0^1([0, 1[; \mathbf{R}^2)$ la fonction $s \mapsto \mathcal{E}(u + sw)$ est de classe \mathcal{C}^1 et que

$$\frac{d}{ds} \mathcal{E}(u + sw)|_{s=0} = \int_{[0,1]} (u' \cdot w' + Dh_u(u') Dh_u(w') + Dh_u(u') D^2 h_u(u', w)) dt.$$

Ensuite, on utilise le fait que u est un minimiseur de \mathcal{E} sur W et donc que $\mathcal{E}(u + sw) \geq \mathcal{E}(u)$ pour tout $s \in \mathbf{R}$, pour conclure que

$$\frac{d}{ds} \mathcal{E}(u + sw)|_{s=0} = 0,$$

pour tout $w \in H_0^1([0, 1[; \mathbf{R}^2)$. En particulier

$$(11.12) \quad \int_{[0,1]} (u' \cdot w' + Dh_u(u') Dh_u(w') + Dh_u(u') D^2 h_u(u', w)) dt = 0,$$

pour tout $w \in H_0^1([0, 1[; \mathbf{R}^N)$. En partant de (11.12) et en reprenant la démonstration du Lemme 11.3 (voir aussi le cours de F. Golse, *Distributions, analyse de Fourier et équations aux dérivées partielles*, MAT431), on conclut que

$$(11.13) \quad u' + Dh_u(u') (\text{grad } h) \circ u - \int_{[0,x]} Dh_u(u') D^2 h_u(u', \cdot) dy,$$

est une fonction constante. Donc $u' + Dh_u(u') (\text{grad } h) \circ u$ est continue (voir l'exercice 7.1 page 149) en tant que primitive de $Dh_u(u') D^2 h_u(u', \cdot)$, qui appartient à $L^1([0, 1[; \mathbf{R}^2)$. Remarquons que, pour tout $y \in \mathbf{R}^2$, l'application linéaire $x \mapsto x + (y \cdot x) y$, qui est définie de \mathbf{R}^2 dans \mathbf{R}^2 , est inversible d'inverse

$$x \mapsto x - \frac{(y \cdot x)}{1 + |y|^2} y.$$

En utilisant cette remarque et le fait que $u \in H^1([0, 1[; \mathbf{R}^2) \subset \mathcal{C}([0, 1]; \mathbf{R}^2)$, on en déduit que $u' \in \mathcal{C}([0, 1]; \mathbf{R}^2)$. Finalement, en reprenant (11.13), on trouve que $u' + Dh_u(u') (\text{grad } h) \circ u$ est de classe \mathcal{C}^1 , en tant que primitive de $Dh_u(u') D^2 h_u(u', \cdot)$ dont on peut maintenant dire que c'est un élément de $\mathcal{C}([0, 1]; \mathbf{R}^2)$. On poursuit pour conclure que $u \in \mathcal{C}^2([0, 1]; \mathbf{R}^2)$.

On vérifie ensuite que

$$(h \circ u)'' = Dh_u(u'') + D^2 h_u(u', u'),$$

et que

$$\begin{aligned} \int_{[0,1]} (u' \cdot w' + Dh_u(u') Dh_u(w')) dt &= \int_{[0,1]} (-u'' \cdot w - (h \circ u)'' Dh_u(w) \\ &\quad - Dh_u(u') D^2 h_u(u', w)) dt. \end{aligned}$$

On laisse le soin au lecteur de justifier proprement chaque étape. Finalement, on conclut en utilisant (11.12) que

$$\int_{[0,1]} (u'' \cdot w + (h \circ u)'' Dh_u(w)) dt = 0.$$

Donc

$$u'' + (h \circ u)'' (\text{grad } h) \circ u = 0,$$

sur $[0, 1]$.

Pour terminer, on définit

$$\gamma(t) := (u(t), h \circ u(t)).$$

L'image de $[0, 1]$ par γ est bien une courbe de classe \mathcal{C}^2 qui est tracée sur Σ et qui joint le point σ_0 au point σ_1 . Par construction, $\gamma''(t)$ est colinéaire au vecteur $-(\text{grad } h) \circ u(t), 1$ qui est un vecteur normal à Σ au point $\gamma(t)$. La démonstration de la proposition est terminée. \square

Le problème de la construction de géodésiques sur une surface ou plus généralement sur une variété, notamment la construction de géodésiques fermées, est un problème qui a fait – et qui fait toujours – couler beaucoup d'encre depuis les premiers travaux de H. Poincaré et G. Birkhoff⁴.

4. Le lecteur intéressé pourra consulter l'article de G. Birkhoff *Dynamical systems with two degrees of freedom* ou l'article de N. Anantharaman *On the existence of closed geodesics, The scientific legacy of Poincaré, 143-160, Hist. Math., 36, Amer. Math. Soc., Providence, RI, 2010.*

Bibliographie

- [1] L. Ambrosio et P. Tilli. *Topics on Analysis in Metric Spaces*. Oxford University Press, Oxford, New York, 2004.
- [2] J.-L. Basdevant et J. Dalibard. *Cours de Mécanique Quantique*. Éditions de l'École Polytechnique, Palaiseau, 2002.
- [3] J.-P. Bourguignon. *Calcul variationnel*. École Polytechnique, Palaiseau, 2007.
- [4] H. Brézis. *Analyse fonctionnelle : Théorie et applications*. Éditions Dunod. 1999.
- [5] H. Brézis. *Analisi Funzionale. Teoria e applicazioni. Con un appendice su Integrazione Astratta di Carlo Sbordone*. Liguori Editore, Naples, 1986.
- [6] H. Dym et H. P. McKean. *Fourier series and integrals*. Academic Press, 1972.
- [7] E. Lieb et M. Loss. *Analysis*. American Math. Soc., Providence RI, 2001.
- [8] P. Malliavin et H. Airault. *Intégration, analyse de Fourier, probabilités, analyse gaussienne*. Masson, Paris, Milan, Barcelone, 1994.
- [9] W. Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.
- [10] E. Stein et R. Shakarchi. *Real Analysis : Measure Theory, Integration, and Hilbert Spaces*. Princeton Lectures in Analysis, Princeton University Press 2005.
- [11] M. Willem. *Principes d'analyse fonctionnelle*, Cassini, Paris, 2003.

Index

- $H^k(\mathbf{N})$, 190
- $L^2(S^1, \mathbf{C})$, 191
- $\ell^2(\mathbf{N})$, 190
- Équation
 - de Sturm-Liouville, 237
- équation
 - de Schrödinger, 237
- Additivité dénombrable, 132
- Adhérence, 18
 - valeur d', 22
- Adjoint d'un opérateur, 200
- Anti-linéaire, 187
- Arithmétique dans $[0, +\infty]$, 77
- Banach, 45
- base hilbertienne, 202
- Bolzano-Weierstrass, 36
- Borel-Lebesgue, 40
- Boule
 - fermée, 10
 - ouverte, 10
- Cantor
 - argument diagonal de, 33
- Cauchy
 - suite de, 43
- Classe \mathcal{L}^+ , 74
- Compacité du cube de Hilbert, 216
- Compacité séquentielle faible de la boule unité, 214
- Compact, 35
 - opérateur, 211
- Complété d'un espace préhilbertien, 191
- Complétude
 - de L^1 , 159
 - de L^p , 164
- Complet
 - espace métrique, 43
- Concentration, 92
- Connexe, 28
 - composante, 29
 - espace, 28
- par arcs, 28
- Continuité
 - uniforme, 23
- Continuité L^1 des translations, 163
- Continuité uniforme, 39
- Contractante
 - application, 55
- Convergence faible, 213
- Coordonnées sphériques, 123
- Critère
 - de continuité pour les formes sesquilinéaires, 199
 - de densité, 202
- Cube de Hilbert, 216
- Décomposition en série de Fourier sur
 - $L^2([0, 1]; \mathbf{C})$, 205
- Décomposition orthogonale d'un espace de Hilbert, 197
- Dénombrabilité, 31
- Dénombrable
 - ensemble, 31
- Dérivation
 - sous le signe somme, 107
- Densité, 21
 - de \mathcal{C}_c dans L^1 , 162
 - de \mathcal{C}_c dans L^p , 164
 - gaussienne, 175
- Diamètre, 46
- Diffraction, 169
- Distance, 9
 - discrète, 9
 - produit, 17
 - somme, 17
- Distances
 - Lispitz-équivalentes, 17
- Dual topologique d'un espace de Hilbert, 195
- Egalité de Parseval, 202
- Ensemble
 - de Besicovich, 97
 - de Cantor, 96
- Ensembles
 - mesurables, 127

- négligeables, 79, 87, 93
- Equivalence des normes, 39
- Espace
 - L^1 , 157
 - L^∞ , 167
 - L^p , 164
 - \mathcal{L}^1 , 84
 - \mathcal{L}^p , 164
 - de Banach, 45
 - dense, 21
 - métrique, 9
 - complet, 43
 - métrique connexe, 28
- Espace de Hilbert, 189
- Espace métrique
 - compact, 35
- Espace préhilbertien, 188
- Espace séparable, 209
- Evanescence, 92
- Famille totale, 204
- Fermé, 15
- Fonction
 - de répartition complémentaire, 134
- fonction absolument continue, 148
- Fonction indicatrice
 - des rationnels, 70
- Fonctions
 - convexes, 109
 - et ensembles mesurables, 130
 - mesurables, 99
 - s.c.i., 76
 - sommables, 84
- Forme hermitienne, 187
- Forme linéaire
 - positive, 74
- Formule
 - du changement de variables, 120
- Formule de Poisson, 180
- Hermitien
 - opérateur, 226
 - produit, 187
- Hermitienne
 - forme, 187
- Inégalité
 - de Hölder, 112
 - de Hölder, 13
 - de Jensen, 111
 - de Markov, 134
 - de Minkowski, 113
- Inégalité de Heisenberg, 223
- Inégalité triangulaire, 9
- Intégrale
 - de Lebesgue, 84
 - des fonctions à valeurs complexes, 105
 - des fonctions continues, 73
 - des fonctions mesurables positives, 103, 104
 - semi-convergente, 86
 - sur un ensemble mesurable, 129
- Intérieur, 18
- Isomorphisme
 - entre espaces préhilbertiens, 192
 - isométrique, 192
- Lemme
 - de Fatou, 90
- Lipschitzienne
 - application, 24
- Métrique
 - induite, 9
- Masse de Dirac, 139
- Mesure
 - de comptage, 140
 - de Lebesgue, 131
 - de Radon, 139
- Nombre de Lebesgue, 37
- Non-compacité de la boule unité, 213
- Norme, 10
 - L^1 , 158
 - L^p , 164
 - de la convergence uniforme, 12, 49
 - matricielle subordonnée, 11
- Normes
 - équivalentes, 13
- Ondelettes de Haar, 207
- Opérateur
 - à noyau, 217
 - compact, 211
 - hermitien, 226
- Opérations sur les fonctions définies p.p., 85
- Orthogonal d'un sous-espace, 193
- Orthonormalisation de Schmidt, 204
- Ouvert, 15
- Paradoxe de Banach-Tarski, 125
- Parties positives/négatives, 75
- Poisson
 - formule, 180
- Préhilbertien, 188
- Presque partout (p.p.), 80
- Principe d'incertitude de Heisenberg, 185
- Principe de Cavalieri, 136
- Produit hermitien, 187
- Séparé
 - espace, 19
- Séparant
 - ensemble, 57

- Sesquilinéaire, 187
- Somme directe d'espaces de Hilbert, 192
- Sommes de Riemann, 71
- Sommes de Riemann dyadiques, 71
- Sous suite, 22
- Spectre d'un opérateur, 227
- Sturm-Liouville
 - équation, 237
- Suite de Levi, 74
- Suite extraite, 22
- Support d'une fonction, 19
- Théorème
 - de Banach-Steinhaus, 54
 - uniforme continuité, 39
 - d'existence des bases hilbertiennes, 209
 - d'inversion de Fourier dans L^1 , 176
 - d'Urysohn, 42
 - de Baire, 53
 - de Bolzano-Weierstrass, 36
 - de Borel-Lebesgue, 40
 - de Brouwer, 56
 - de comparaison, 103
 - de Dini, 41
 - de Fubini, 115
 - de Hahn-Banach, 201
 - de la convergence dominée, 89
 - de la convergence monotone, 104
 - de la convergence monotone dans \mathcal{L}^+ , 81
 - de la convergence monotone dans \mathcal{L}^1 , 89
 - de la projection, 197
 - de la projection sur un convexe, 196
 - de non-compacité de Riesz, 213
 - de point fixe de Banach, 55
 - de représentation de Riesz, 198
 - de représentation de Riesz (version
 - sequinéaire), 199
 - de Riemann-Lebesgue, 172
 - de Riesz-Fischer, 202
 - de Schauder, 56
 - de Stone-Weierstrass, 57
 - de Sturm-Liouville, 238
 - de Tonelli, 118
 - spectral, 227
- Topologie, 15
 - induite, 16
- Transformée de Fourier
 - dans L^2 , 223
 - normalisée, 223
- Transformation
 - de Fourier, 171
 - de Fourier et dérivation, 173
- Triangle de Sierpiński, 96
- Tribu, 127, 141
- Valeur propre d'un opérateur, 227
- Voisinage, 15