BASES D'ANALYSE FONCTIONNELLE

Jean-Yves CHEMIN
Laboratoire J.-L. Lions, Case 187
Université Pierre et Marie CURIE, 4 Place Jussieu
75230 Paris Cedex 05, France
Télécopie : 01 44 27 72 00, adresse électronique : chemin@ann.jussieu.fr

17 décembre 2017

Sommaire

Introduction

Ce texte est le support du cours "Bases d'Analyse fonctionnelle" de la première année du Master de Mathématiques de l'Université Paris et Marie Curie. Le but de ce cours est d'acquérir une maîtrise élémentraire mais solide d'outils qui sont fondamentaux pour la compréhension de mathématiques intervenant aussi bien dans le cœur de la discipline (géométrie, probabilités, équations aux dérivées partielles) qu'en physique, en mécanique, ou bien dans les applications des mathématiques à l'analyse des grands systèmes, l'analyse d'image, statistique. . ..

Tout d'abord, quelques remarques générales pour utiliser ces notes. Tout d'abord, il ne s'agit pas d'un traité. Certains résultats classiques sont absents comme le théorème de Cauchy-Lipschitz parce qu'exposés dans d'autres cours, ou bien traités dans un cas particulier comme le théorème de Stone-Weierstrass. Il arrive que des démonstrations faciles qui ne sont que des applications simples des définitions soient esquissées et même omises. Il est évident que leur rédaction détaillée constitue un excellent exercice d'apprentissage. D'une manière générale, le lecteur désireux d'acquérir de bonnes connaissances des concepts introduits devra se réapproprier les démonstrations du cours.

Des démonstrations qui, soit ne sont pas considérées comme centrales dans le cours soit sont considérées comme trop difficiles, sont présentées dans ces notes en petits caractères. Elles ne sont pas traitées en cours mais sont là pour satisfaire la curiosité d'auditeurs motivés.

La structure de ces notes est la suivante : Dans le chapitre 1, sont exposées les notions de base de la topologie des espaces métriques avec notamment les notions d'espaces complets et d'espaces compacts. Il s'agit d'un chapitre dont les résultats doivent absolument être maîtrisés.

Le chapitre 2 est consacré à l'étude des espaces vectoriels normés. L'exemple fondamental des espaces de fonctions y est traité. L'un des points clefs est la compréhension du changement induit sur la topologie par la dimension infinie (cas notamment des espaces de fonctions). Le théorème d'Ascoli qui donne un critère de compacité pour les parties des espaces de fonctions continues est une illustration des difficultés qui surgissent dans le cadre de la dimension infinie.

Le chapitre 3 est consacré à la notion de dualité. Bien que bref, il est fondamental. La notion de dualité est à la base de la théorie des distributions, qui a révolutionné l'analyse à l'orée de la seconde moitié du XXième siècle. Cette théorie sera étudiée aux chapitre 8 Outre le concept d'application linéaire transposée, on explique dans ce chapitre la procédure dite d'identification du dual d'un espace de Banach à un autre espace de Banach ainsi qu'une notion affaiblie de la convergence qui est définie dans le cadre du dual d'un espace de Banach : la convergence dite "faible étoile".

Le chapitre 4 est un classique : il est consacré à l'étude des espaces de Hilbert qui sont une extension à la dimension infinie des espaces euclidiens.

Le chapitre 5 est consacré à l'étude des espaces de puissance p ième intégrale par rapport à une mesure. On rappelle sans démonstration les résultats fondamentaux de la théorie de l'intégration. La notion fondamentale de convolution des fonctions est définie et étudiée, puis appliquée à la théorie de l'approximation.

Le chapitre 6 est consacré à l'étude du problème dit de Dirichlet dans un domaine borné. L'objectif de ce chapitre est de démontrer que pour toute fonction f de carré intégrable sur un ouvert Ω connexe borné de $\mathbb{R}^d$, il existe une unique solution u dans un espace fonctionnel que l'on définira (l'espace dit $H_0^1(\Omega)$ et appelé espace de Sobolev) telle que, en un sens élargi, on ait

$$-\sum_{j=1}^{d} \frac{\partial^2 u}{\partial {x_j}^2} = f$$

et tel que la fonction u soit nulle sur la frontière de l'ouvert Ω. On trouve la solution en cherchant si la borne inférieure (sur l'ensemble des fonctions continûment dérivable sur Ω et à support compact dans Ω) de la fonction

$$u \longmapsto \int_\Omega \sum_{j=1}^{d} \Big| \frac{\partial u}{\partial x_j} \Big|^2 dx - \int_\Omega f(x)u(x)dx.$$

est atteinte. Nous verrons à l'œuvre beaucoup de concepts et de résultats établis précédemment. On y trouve en germe beaucoup des idées de la théorie des distributions étudiée au chapitre 8

Le chapitre 7 est consacré à l'étude de la transformée de Fourier sur l'espace $\mathbb{R}^d$ des fonctions intégrables et à plusieurs de ces appications. Fondamental, ce chapitre est crucial pour les deux suivants.

Le chapitre 8 est consacré à la présentation de la théorie des distributions dites tempérées. Le choix de ne présenter que cette théorie et non la théorie générale des distributions tient à une volonté de simplicité. L'idée fondamentale est que lorsque l'on sait définir une opération sur les fonctions très régulières et très décroissantes (par exemple sur l'espace de Schwartz), on sait par dualité la définir sur l'espace des distributions tempérées qui généralise les fonctions et qui contient des objet très singuliers. Ce chapitre est bien sûr illustré d'exemples qui doivent être connus et maîtrisés sans quoi cette théorie ne peut être ni comprise et ni appliquée. Il se conclut par deux applications : la démonstration du fait que la transformée de Hilbert et, à une constante près, une isométrie de $L^2(\mathbb{R})$ et par la résolution explicite d'une équation différentielle sur $\mathbb{R}$.

Chapitre 1

Espaces métriques

Introduction

Ce chapitre condense des résultats de base sur les espaces métriques. La première section présente la notion très intuitive et naturelle de distance et montre comment elle permet une très grande généralisation des notions (familières dans les cas réel ou complexe) de suite convergente et de fonctions continues. Cette notion de distance permet aussi de définir les notions abstraites d'ouverts et de fermés qui sont utilisés constamment en analyse fonctionnelle.

Dans le deuxième section, on introduit la notion d'espace complet, espace dans lequel toute suite de Cauchy converge. Cette notion est fondamentale : ce sont dans ces espaces que l'on peut démontrer que des suites convergent sans avoir a priori aucune idée sur la limite. Il s'agit là d'un outil fondamental et d'usage très fréquent en analyse pour démontrer des théorèmes d'existence. Le théorème de Cauchy-Lipschitz en est l'une des illustrations.

Dans la troisième section, on introduit le concept d'espace compact. La pratique de l'analyse fonctionnelle nécessite de dépasser la représentation élémentaire des compacts dans les espaces $\mathbb{R}^N$.

L'ensemble de ce chapitre est assez abstrait. Les exemples, illustrations et applications de des notions fondamentales présentées dans ce chapitre seront fréquentes dans la suite du cours.

1.1 Définition des espaces métriques

Définition 1.1.1. *Soit X un ensemble, on appelle distance sur X toute application d de $X \times X$ dans $\mathbb{R}^+$ telle que*

$$\begin{array}{rcl} d(x,y) & = & 0 \Longleftrightarrow x = y \\ d(x,y) & = & d(y,x) \\ d(x,y) & \leq & d(x,z) + d(z,y) \end{array}$$

Le couple (X,d) est appelé un espace métrique.

Quelques exemples

— Prenons $X = \mathbb{R}$ et $d(x,y) = |x - y|$. Cela définit un espace métrique.

— Prenons $X = \mathbb{R}^N$ et choisissons les différentes distances suivantes :

$$\begin{aligned} d_e(x,y) &\stackrel{\text{déf}}{=} \Big(\sum_{j=1}^{N} (x_j - y_j)^2\Big)^{\frac{1}{2}} \\ d_\infty(x,y) &\stackrel{\text{déf}}{=} \max_{1\leq j\leq N} |x_j - y_j| \\ d_1(x,y) &\stackrel{\text{déf}}{=} \sum_{j=1}^{N} |x_j - y_j|. \end{aligned}$$

— Plus généralement, considérons une famille finie $(X_j, d_j)_{1\leq j\leq N}$ d'espaces métriques. On pose $X = \prod_{j=1}^{N} X_j$. On définit

$$D_\infty \left\{ \begin{array}{ccc} X\times X & \longrightarrow & \mathbb{R}^+ \\ (x,x') & \longmapsto & \max\limits_{1\leq j\leq N} d_j(x_j,y_j) \end{array} \right. \quad \text{et} \quad D_1 \left\{ \begin{array}{ccc} X\times X & \longrightarrow & \mathbb{R}^+ \\ (x,x') & \longmapsto & \sum\limits_{j=1}^{N} d_j(x_j,y_j) \end{array} \right. \tag{1.1}$$

Ces deux applications définissent des distances sur X.

— Prenons à nouveau $X = \mathbb{R}$, considérons une application injective f de $\mathbb{R}$ dans $\mathbb{R}$ et définissons

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

L'application d_f est une distance sur X.

Comme le montre l'exercice suivant, on peut définir une distance sur l'espace des suites d'un espace métrique.

Exercice 1.1.1. *Soit (X,d) une espace métrique et $(a_n)_{n\in\mathbb{N}}$ une suite de réels strictement positifs telle que*

$$\sum_{n\in\mathbb{N}} a_n < \infty.$$

On considère l'ensemble $X^{\mathbb{N}}$ des suites d'éléments de X. On définit alors

$$D \left\{ \begin{array}{ccc} X^{\mathbb{N}}\times X^{\mathbb{N}} & \longrightarrow & \mathbb{R}^+ \\ (x,y) & \longmapsto & \sum\limits_{n\in\mathbb{N}} \min\big\{a_n, d(x(n),y(n))\big\}. \end{array} \right.$$

L'application D est une distance sur $X^{\mathbb{N}}$.

Définition 1.1.2. *Soient (X,d) un espace métrique, x un point de X et α un réel strictement positif. On appelle boule ouverte (resp. fermée) de centre x et de rayon α, et l'on note $B(x,\alpha)$ (resp. $B_f(x,\alpha)$) l'ensemble des points y de X tels que $d(x,y) < \alpha$ (resp. $d(x,y) \leq \alpha$).*

La notion de distance permet de définir de manière très simple et générale, le concept de limite d'une suite et celui de fonction continue.

Définition 1.1.3 (convergence des suites)**.** *Soient $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments d'un espace métrique (X,d) et ℓ un point de X. On dit que la suite $(x_n)_{n\in\mathbb{N}}$ converge vers ℓ si et seulement si*

$$\forall \varepsilon > 0\,,\ \exists\, n_0 \,/\, n \geq n_0 \Longrightarrow d(x_n,\ell) < \varepsilon.$$

Exercice 1.1.2. *On considère un espace métrique (X, d) et la distance D_a définit sur $X^{\mathbb{N}}$ à l'exercice 1.1.1. Une suite $(x_p)_{p\in\mathbb{N}}$ d'éléments de $X^{\mathbb{N}}$ converge vers x au sens de D_a si et seulement si*

$$\forall n \in \mathbb{N}\,,\ \lim_{p\to\infty} d(x_p(n), x(n)) = 0.$$

Définition 1.1.4 (continuité des fonctions)**.** *Soient (X, d) et (Y, δ) deux espaces métriques. On considère une fonction f de X dans Y et un point x_0 de X. La fonction f est continue en x_0 si et seulement si*

$$\forall \varepsilon > 0\,,\ \exists\ \alpha > 0\,/\ d(x, x_0) < \alpha \Longrightarrow \delta(f(x), f(x_0)) < \varepsilon.$$

Proposition 1.1.1 (composition des fonctions continues)**.** *Soient (X, d), (Y, δ) et (Z, ρ) trois espaces métriques, f et g deux fonctions respectivement de X dans Y et de Y dans Z. Soit x_0 un point de X tel que f soit continue en x_0 et g le soit en $f(x_0)$. Alors la fonction $g \circ f$ est continue en x_0.*

Démonstration. La fonction g étant continue en $f(x_0)$, on a

$$\forall \varepsilon > 0\,,\ \exists \alpha > 0\,/\ \delta(y, f(x_0)) < \alpha \Longrightarrow \rho(g(y), (g \circ f)(x_0)) < \varepsilon.$$

La fonction f étant continue en x_0, on a

$$\exists \beta > 0\,/\ d(x, x_0) < \beta \Longrightarrow \delta(f(x), f(x_0)) < \alpha.$$

Le réel β étant ainsi choisi pour chaque ε, on en déduit alors que

$$\exists \beta > 0\,/\ d(x, x_0) < \beta \Longrightarrow \rho(g \circ f(x), (g \circ f)(x_0)) < \varepsilon.$$

Ceci conclue la démonstration. □

Proposition 1.1.2 (suite et fonctions continues)**.** *Soient (X, d) et (Y, δ) deux espaces métriques, f une fonction de X dans Y et $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de X. On suppose que $(x_n)_{n\in\mathbb{N}}$ converge vers ℓ et que la fonction f est continue en ℓ. Alors la suite d'éléments de Y définie par $(f(x_n))_{n\in\mathbb{N}}$ converge vers $f(\ell)$.*

La démonstration de cette proposition est analogue à celle de la proposition précédente et est laissée en exercice au lecteur.

Définition 1.1.5 (Intérieur et adhérence, ouvert et fermé)**.** *Soit (X, d) un espace métrique et A une partie quelconque de X.*

- *On appelle intérieur de A et l'on note $\overset{\circ}{A}$ l'ensemble des points x de X tels qu'il existe une boule ouverte $B(x, \alpha)$ (avec $\alpha > 0$) incluse dans A.*
- *On appelle adhérence de A et l'on note $\overline{A}$ l'ensemble des points x de X tels que, pour toute boule ouverte $B(x, \alpha)$ (avec $\alpha > 0$), l'intersection de $B(x, \alpha)$ avec A soit non vide.*
- *On dit que A est dense dans X si et seulement si $\overline{A} = X$.*
- *On dit que A est ouvert si et seulement si $\overset{\circ}{A} = A$.*
- *On dit que A est fermé si et seulement si $A = \overline{A}$.*

Remarque Il est clair d'après la définition que $\overset{\circ}{A} \subset A \subset \overline{A}$.

Exercice 1.1.3. *Soit A une partie finie d'un espace métrique (X, d). Alors $\overline{A} = A$.*

Proposition 1.1.3. *L'intérieur d'une boule ouverte est elle-même. L'adhérence d'une boule fermée est elle-même.*

Démonstration. Soit y un point de la boule ouverte $B(x, \alpha)$, considérons la boule ouverte de centre y et de rayon $\alpha - d(x, y)$ (qui est un nombre strictement positif). D'après l'inégalité triangulaire, on a

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + \alpha - d(x, y) = \alpha.$$

Considérons maintenant un point y de l'adhérence de la boule fermée de centre x et de rayon α. Par définition, pour tout réel strictement positif β, il existe un point z de $B(y, \beta) \cap B_f(x, \alpha)$. À nouveau d'après l'inégalité triangulaire, on a

$$d(x, y) \leq d(x, z) + d(z, y) < \beta + \alpha.$$

Ainsi donc, pour tout β strictement positif, $d(x, y) < \alpha + \beta$ et donc $d(x, y) \leq \alpha$, ce qui achève la démonstration. □

Remarques

— La proposition ci-dessus dit exactement que les boules ouvertes sont des ouverts et que les boules fermées sont des fermés (ce qui est rassurant quant à la cohérence de la terminologie).

— L'ensemble X est à la fois ouvert et fermé. On convient qu'il en est de même pour l'ensemble vide.

— La démonstration ci-dessus démontre que l'adhérence de la boule ouverte $B(x_0, \alpha)$ est incluse dans la boule fermée $B_f(x_0, \alpha)$. Il est par contre faux en général que l'adhérence de la boule ouverte soit la boule fermée. Par exemple, en prenant sur un ensemble X quelconque, la distance d définie par $d(x, y) = 1$ si x est différent de y et $d(x, x) = 0$, on a $B(x_0, 1) = \{x_0\}$ qui est fermé donc égal à son adhérence et $B_f(x_0, 1) = X$.

Proposition 1.1.4. *Soit A une partie d'un espace métrique (X, d). On a*

$$(\overset{\circ}{A})^c = \overline{A^c} \quad et \quad \overline{A}^c = (\overset{\circ}{A^c}).$$

Démonstration. Un point x de X appartient à $(\overset{\circ}{A})^c$ si et seulement si

$$\forall \alpha > 0\,,\ B(x, \alpha) \cap A^c \neq \emptyset,$$

ce qui signifie exactement que x appartient à l'adhérence du complémentaire de A. Un point x appartient à $(\overline{A})^c$ si et seulement si

$$\exists \alpha > 0\,,\ B(x, \alpha) \cap A = \emptyset,$$

c'est-à-dire $B(x, \alpha) \subset A^c$, ce qui signifie exactement que x appartient à $(\overset{\circ}{A^c})$. La proposition est démontrée. □

Remarque D'après la proposition 1.1.4, le complémentaire d'un ouvert est un fermé et réciproquement.

Proposition 1.1.5. *Soit A une partie d'un espace métrique (X, d). Un point x de X appartient à $\overline{A}$ si et seulement si il existe une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A tellle que $\lim_{n\to\infty} a_n = x$.*

Démonstration. Supposons que x soit limite d'une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A. Pour tout réel strictement positif α, il existe un entier n_0 tel que x_{n_0} appartienne à $B(x,\alpha)$ ce qui implique que $B(x,\alpha)\cap A\neq\emptyset$. Donc $x\in\overline{A}$.

Réciproquement, supposons que $x\in\overline{A}$. Alors, pour tout entier strictement positif n, il existe un élément a_n de X tel que

$$a_n\in A\cap B(x,n^{-1})$$

Soit $(a_n)_{n\in\mathbb{N}}$ la suite ainsi définie. La suite $(a_n)_{n\in\mathbb{N}}$ vérifie

$$d(a_n,x)\leq\frac{1}{n}\cdot$$

La suite $(a_n)_{n\in\mathbb{N}}$ converge donc vers x et la proposition est démontrée. □

Proposition 1.1.6. *Toute réunion d'ouverts en est un. Toute intersection finie d'ouverts en est un. Toute intersection de fermés en est un. Toute réunion finie de fermés en est un.*

Démonstration. Soit $(U_\lambda)_{\lambda\in\Lambda}$ une famille quelconque d'ouverts et x un point de $U\stackrel{\text{déf}}{=}\bigcup_{\lambda\in\Lambda}U_\lambda$. Soit $\lambda\in\Lambda$ tel que $x\in U_\lambda$. Comme U_λ est ouvert (ce qui signifie égal à son intérieur),

$$\exists\alpha>0\,/\;B(x,\alpha)\subset U_\lambda\subset U$$

et donc U est un ouvert. Soit $U=\bigcap_{j=1}^{N}$ où les U_j sont des ouverts. Pour tout j dans $\{1,\cdots N\}$, il existe un réel strictement positif α_j tel que $B(x,\alpha_j)\subset U_j$. Soit $\alpha\stackrel{\text{déf}}{=}\min\{\alpha_j\,,\ j\in\{1,\cdots N\}\}$. Pour tout j, la boule ouverte $B(x,\alpha)$ est incluse dans U_j et donc dans l'intersection U. □

Remarque Pour un ensemble X, on considère Θ une partie de $\mathcal{P}(X)$ l'ensemble des sous-ensembles de X telle que :

— L'ensemble vide et X appartiennent à Θ,

— Si $(U_j)_{1\leq j\leq N}$ est une famille finie d'éléments de Θ, alors $\bigcap_{j=1}^{N}U_j$ appartient à Θ,

— Si $(U_\lambda)_{\lambda\in\Lambda}$ est une famille quelconque d'éléments de Θ, alors $\bigcup_{\lambda\in\Lambda}$ appartient à Θ.

Ceci définit ce que l'on appelle une topologie sur X, les éléments de Θ étant par définition les ouverts de X, les fermés étant par définition les complémentaires des ouverts. Comme le montre la proposition suivante, les notions de suite convergente et d'application continue peuvent se définir en terme d'ouverts.

Proposition 1.1.7. *Soit (X,d) un espace métrique. Soit $(x_n)_{n\in\mathbb{N}}$ d'éléments de X et x un élément de X. La suite $(x_n)_{n\in\mathbb{N}}$ converge vers x si et seulement si pour tout ouvert U de X contenant x, il existe n_0 tel que*

$$\forall n\geq n_0\,,\ x_n\in U.$$

Soient (Y,δ) un espace métrique, f une fonction de X dans Y et x_0 un élément de X. La fonction f est continue en x_0 si et seulement si pour tout ouvert V de Y contenant $f(x_0)$, il existe un ouvert U de X contenant x_0 tel que

$$f(U)\subset V.$$

La démonstration de cette proposition est un exercice formateur vivement conseillé.

Théorème 1.1.1 (Caractérisation des applications continues)**.** *Soit f une application entre deux espaces métriques (X, d) et (Y, δ). Les trois assertions suivantes sont équivalentes.*
— *L'application f est continue en tout point de X,*
— *l'image réciproque d'un ouvert est un ouvert,*
— *l'image réciproque d'un fermé est un fermé.*

Démonstration. Supposons f continue en tout point de X et considérons un ouvert V de (Y, δ) et un point x de $f^{-1}(V)$. L'ensemble V étant ouvert, il existe par définition un réel strictement positif ε_0 tel que $B(f(x), \varepsilon_0) \subset U$. La fonction f étant continue en x,

$$\exists \alpha > 0\,,\; f(B(x, \alpha)) \subset B(f(x), \varepsilon_0).$$

Ainsi donc

$$B(x, \alpha) \subset f^{-1}\Big(f(B(x), \alpha)\Big) \subset f^{-1}(B(f(x), \varepsilon_0)) \subset f^{-1}(V).$$

Réciproquement, soit x_0 un point de X. Pour tout $\varepsilon > 0$, la boule $B(f(x_0), \varepsilon)$ est un ouvert de (Y, δ). Par hypothèse, $f^{-1}(B(f(x_0), \varepsilon))$ est un ouvert de (X, d). Donc il existe α strictement positif tel que $B(x_0, \alpha)$ soit inclus dans $f^{-1}(B(f(x_0), \varepsilon))$ et donc que

$$f(B(x_0, \alpha)) \subset f\Big(f^{-1}(B(f(x_0), \varepsilon))\Big) \subset B(f(x_0), \varepsilon).$$

L'équivalence entre les points deux et trois résultent du fait que les complémentaires des ouverts sont les fermés (et réciproquement) et que

$$\begin{aligned} f^{-1}(V) &= \{x \in X \,/\, f(x) \in V\} \\ &= \{x \in X \,/\, f(x) \in V^c\}^c \\ &= (f^{-1}(V^c))^c. \end{aligned}$$

Le théorème est ainsi démontré. □

On peut s'interroger sur les effets d'un changement de distance sur les propriétés d'un espace métrique.

Définition 1.1.6. *Soit X un ensemble, on considère deux distances d et δ sur X. On dit que les deux distances sont topologiquement équivalentes si et seulement si l'application* Id *est continue de (X, d) dans (X, δ) et de (X, δ) dans (X, d).*

Remarque Les ouverts associés à deux distances topologiquement équivalentes sont les mêmes ; ainsi donc les fonctions continues et les suites convergentes sont les mêmes.

Proposition 1.1.8. *Soient X un ensemble et d et δ deux distances sur X. Les propriétés suivantes sont alors satisfaites.*

Les distances d et δ sont topologiquement équivalentes si et seulement si

$$\forall x \in X\,, \forall \varepsilon > 0\,, \exists \eta > 0 \,/\, \forall y \in Y\,,$$
$$d(x, y) < \eta \Longrightarrow \delta(x, y) < \varepsilon \quad \text{et} \quad \delta(x, y) < \eta \Longrightarrow d(x, y) < \varepsilon.$$

La démonstration de cette proposition n'est qu'une application immédiate des définitions ; elle est laissée au lecteur.

Définition 1.1.7. *Soient A une partie d'un espace métrique (X,d) et x un point de X. On appelle distance du point x à l'ensemble A la quantité*

$$d(x,A) \stackrel{\text{déf}}{=} \inf_{a\in A} d(x,a).$$

Exercice 1.1.4. *Démontrez que $\overline{A}$ est l'ensemble des points x de X tels que $d(x,A)=0$.*

Proposition 1.1.9. *Soit A une partie d'un espace métrique (X,d). La fonction*

$$d_A \left\{ \begin{array}{ccc} X & \longrightarrow & \mathbb{R}^+ \\ x & \longmapsto & d(x,A) \end{array} \right.$$

est lipschitzienne de rapport 1, c'est-à-dire que

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

Démonstration. D'après l'inégalité triangulaire, on a, pour tout (x,y) de X^2 et tout point a de A,

$$d(x,a) \leq d(x,y)+d(y,a).$$

La borne inférieure étant un minorant, nous avons, pour tout (x,y) de X^2 et tout point a de A,

$$d(x,A) \leq d(x,y)+d(y,a) \quad \text{et donc} \quad d(x,A)-d(x,y) \leq d(y,a).$$

La borne inférieure étant le plus grand des minorants, nous en déduisons que

$$d(x,A)-d(x,y) \leq d(y,A) \quad \text{et donc} \quad d(x,A)-d(y,A) \leq d(x,y).$$

D'où la proposition en permutant le rôle de x et y. □

Nous allons maintenant introduire la notion de sous espaces métriques. Soit (X,d) un espace métrique et A une partie de X, il est naturel de considèrer l'espace métrique $(A, d_{|A\times A})$. Nous avons la propriété suivante.

Proposition 1.1.10. *Soit (X,d) un espace métrique et A une partie de X. Alors une partie B de A est un ouvert (resp. un fermé) de l'espace métrique $(A, d_{|A\times A})$ si et seulement si il existe une partie ouverte (resp. fermée) $\widetilde{B}$ de X tel que $B=\widetilde{B}\cap A$.*

Démonstration. Nous n'allons traiter que le cas des ouverts, celui des fermés s'en déduisant par passage au complémentaire [1]. Soit $\widetilde{B}$ un ouvert de X, démontrons que $\widetilde{B}\cap A$ est un ouvert de $(A, d_{A\times A})$. Soit a_0 un point de $\widetilde{B}\cap A$. Comme $\widetilde{B}$ est un ouvert de X, il existe une boule ouverte (pour l'espace métrique (X,d) tel que

$$B(a_0,\alpha) \subset \widetilde{B}.$$

Par intersection, on en déduit que $B(a_0,\alpha)\cap A$ est inclus dans B. Mais $B(a_0,\alpha)\cap A$ est exactement l'ensemble des a de A tels que $d(a_0,a)<\alpha$. Donc B est un ouvert de l'espace métrique $(A, d_{|A\times A})$.

Réciproquement, soit B un ouvert de l'espace métrique $(A, d_{|A\times A})$. Pour tout a dans B, il existe un réel strictement positif α_a tel que

$$B_A(a,\alpha_a) \subset A \quad \text{avec} \quad B_A(a,\alpha_a) = \{a' \in A \,/\, d(a,a')<\alpha\}.$$

1. Écrire en détail le cas des fermés est un exercice fortement recommandé

Posons $\widetilde{B} \stackrel{\text{déf}}{=} \bigcup_{a\in A} B_X(a,\alpha)$ qui est un ouvert comme réunion d'ensemble ouverts. Vu que l'on a $B_A(a,\alpha) = B_X(a,\alpha) \cap A$, on a $\widetilde{B} \cap A = B$ et la proposition est démontrée. $\square$

Pour conclure cette section introductive sur les espaces métriques, définissons la notion de diamètre dans un espace métrique.

Définition 1.1.8. *Soit (X,d) un espace métrique. On dit qu'une partie A de X est de diamètre finie si et seulement si il existe un réel strictement positif C tel que*

$$\forall (a,a') \in A^2\,,\ d(a,a') \leq C.$$

Lorsque A est de diamètre fini, on définit la diamètre comme étant la borne supérieure de l'ensemble des quantités $d(a,a')$ lorsque (a,a') parcourt l'ensemble $A \times A$.

1.2 Espaces complets

Définition 1.2.1. *Soit (X,d) un espace métrique, on appelle suite de Cauchy de X toute suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X telle que*

$$\forall \varepsilon > 0\,,\ \exists n_0 \in \mathbb{N} \,/\, \forall n \geq n_0\,,\ \forall m \geq n_0\,,\ d(x_n, x_m) < \varepsilon.$$

Remarquons que si $(x_n)_{n\in\mathbb{N}}$ est une suite d'éléments d'un espace métrique (X,d) qui converge vers un élément ℓ de X, alors comme

$$d(x_n, x_m) \leq d(x_n, \ell) + d(\ell, x_m)$$

c'est une suite de Cauchy. Les espaces complets sont précisément les espaces où la réciproque est vraie. Plus précisément on a la définiton suivante.

Définition 1.2.2. *Soit (X,d) un espace métrique, on dit que cet espace est complet si et seulement si toute suite de Cauchy est convergente.*

Donnons maintenant quelques exemples d'espaces complets. Un exemple fondamental est l'espace $\mathbb{R}$ muni de la distance $d(x,y) = |x-y|$ qui est complet par construction. Nous allons maintenant examiner comment fabriquer des espaces complets à partir d'espaces complets déjà connus. Dit autrement, cela signifie que l'on va rechercher les opérations sur les espaces métriques qui laissent stable la propriété d'espace complet.

Proposition 1.2.1. *Soit $(X_1,d_1),\cdots,(X_N,d_N)$ une famille de N espaces métriques complets. Si l'on pose*

$$X = X_1 \times \cdots \times X_N \quad \text{et} \quad d((x_1,\cdots,x_N),(y_1,\cdots,y_N)) = \max_{1\leq j\leq N} d_j(x_j,y_j),$$

alors l'espace (X,d) est un espace complet.

La démonstration est laissée en exercice. Il en résulte que l'espace $\mathbb{R}^N$ muni de l'une des distances d_e, d_1 ou d_∞ est un espace complet.

L'exercice suivant fournit un exemple intéressant d'espace complet.

Exercice 1.2.1. *Soit (X,d) un espace métrique complet. On considère sur $X^{\mathbb{N}}$ la distance D_a de l'exercice 1.1.1. L'espace métrique $(X^{\mathbb{N}}, D_a)$ est complet.*

Proposition 1.2.2. *Soit (X,d) un espace complet. On considère A une partie de X. L'espace métrique $(A, d_{|A\times A})$ est complet si et seulement si A est fermé.*

Démonstration. Supposons $(A, d_{|A\times A})$ complet. Soit $(a_n)_{n\in\mathbb{N}}$ une suite d'éléments de A qui converge vers x dans X. La suite $(a_n)_{n\in\mathbb{N}}$ est une suite de Cauchy de X formée d'éléments de A donc c'est une suite de Cauchy de $(A, d_{|A\times A})$ qui est complet. Il existe donc a dans A tel que la suite $(a_n)_{n\in\mathbb{N}}$ converge vers a au sens de $(A, d_{|A\times A})$ donc aussi au sens de (X,d). L'unicité de la limite assure que $a = x$ et donc que x appartient à A.

Réciproquement supposons A fermé et considérons une suite $(a_n)_{n\in\mathbb{N}}$ de Cauchy de l'espace métrique $(A, d_{|A\times A})$. C'est une suite de Cauchy de (X,d) qui est complet. Donc elle converge vers un élément x de X. Le fait que A soit fermé implique x est dans A et donc $(A, d_{|A\times A})$ est complet. □

Lorsqu'un espace (X,d) est complet, cela permet de démontrer l'existence de certains objets. Le théorème suivant en est l'illustration la plus spectaculaire.

Théorème 1.2.1 (de point fixe de Picard)**.** *Soit f une application d'un espace métrique complet (X,d) dans lui-même telle qu'il existe un réel k de l'intervalle $]0,1[$ vérifiant*

$$d(f(x), f(y)) \leq kd(x,y).$$

Il existe alors un unique point fixe z tel que $f(z) = z$.

Démonstration. Étant donné un élément x_0 de X, on considère la suite $(x_n)_{n\in\mathbb{N}}$ définie par la relation de récurrence $x_{n+1} = f(x_n)$. On peut écrire que

$$\begin{aligned} d(x_{n+1}, x_n) &= d(f(x_n), f(x_{n-1})) \\ &\leq kd(x_n, x_{n-1}). \end{aligned}$$

Par itération multiplicative, on trouve que

$$d(x_{n+1}, x_n) \leq k^n d(x_1, x_0).$$

Ainsi, pour tout couple d'entiers (n,p) , on a

$$\begin{aligned} d(x_{n+p}, x_n) &\leq \sum_{m=1}^{p} d(x_{n+m}, x_{n+m-1}) \\ &\leq d(x_1, x_0) \sum_{m=1}^{p} k^{n+m-1} \\ &\leq \frac{k^n}{1-k} d(x_1, x_0). \end{aligned}$$

La suite $(x_n)_{n\in\mathbb{N}}$ est donc de Cauchy. Soit z sa limite. Comme la fonction f est lipschitzienne, donc continue, on obtient, en passant à la limite dans la relation de définition de la suite $(x_n)_{n\in\mathbb{N}}$ que $z = f(z)$.

Il nous reste à démontrer l'unicité. Soient z_1 et z_2 deux solutions de $z = f(z)$. On a, d'après l'hypothèse faite sur f que

$$d(z_1, z_2) \leq kd(z_1, z_2).$$

Le fait que k soit strictement inférieur à 1 assure que $d(z_1, z_2) = 0$, donc que $z_1 = z_2$. Le théorème est ainsi démontré. □

Remarques Ce théorème est à la base d'innombrables théorèmes d'existence et d'unicité. Un exemple important est le théorème de Cauchy-Lipschitz d'existence et d'unicité pour les équations différentielles ordinaires.

Nous allons maintenant démontrer le classique théorème de Baire qui a de nombreuses applications en analyse fonctionnelle. Quelques exemples sont donnés au chapitre 2.

Théorème 1.2.2 (de Baire). *Soit (X,d) un espace métrique complet, on considère une suite $(U_n)_{n\in\mathbb{N}}$ d'ouverts denses dans X. Alors $\displaystyle\bigcap_{n\in\mathbb{N}} U_n$ est dense.*

Démonstration. Elle repose en grande partie sur le lemme suivant qui a son intérêt propre.

Lemme 1.2.1. *Soit (X,d) un espace métrique complet et $(F_n)_{n\in\mathbb{N}}$ une suite décroissante de fermés non vides de X dont le diamètre tend vers 0. Il existe alors un élément x de X tel que*

$$\bigcap_{n\in\mathbb{N}} F_n = \{x\}.$$

Démonstration. On considère une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X telle que pour tout entier x_n, appartienne à F_n. Comme la suite d'ensemble $(F_n)_{n\in\mathbb{N}}$ est décroissante (au sens de l'inclusion), on a

$$\forall p\in\mathbb{N}\,,\ d(x_n,x_{n+p})\leq\delta(F_n).$$

Le fait que le diamètre des F_n tend vers 0 implique que la suite $(x_n)_{n\in\mathbb{N}}$ est de Cauchy donc converge vers un point x de X. En faisant tendre p vers l'infini dans l'assertion ci-dessus, on trouve que toute suite $(x_n)_{n\in\mathbb{N}}$ telle x_n appartienne à F_n vérifie $d(x_n,x)\leq\delta(F_n)$. Ceci conclut la démonstration du lemme. □

Retour à la démonstration du théorème 1.2.2. Soit V un ouvert quelconque de X, nous allons démontrer que $\displaystyle\bigcap_n U_n\cap V\neq\emptyset$, ce qui assurera le théorème.

L'ouvert U_0 est dense, donc $U_0\cap V$ est un ouvert non vide. Donc il existe un réel strictement positif α_0 (que l'on peut supposer inférieur à 1) et un point x_0 de X tels que

$$B_f(x_0,\alpha_0)\subset U_0\cap V. \tag{1.2}$$

L'ouvert U_1 est dense, donc l'ensemble $U_1\cap B(x_0,\alpha_0)$ est un ouvert non vide. Donc il existe un réel strictement positif α_1 (que l'on peut supposer inférieur à $1/2$) et un point x_1 de X tels que

$$B_f(x_1,\alpha_1)\subset U_1\cap B(x_0,\alpha_0).$$

Nous allons procéder par récurrence et supposer construite une suite $(x_j)_{0\leq j\leq n}$ d'éléments de X et une suite $(\alpha_j)_{0\leq j\leq n}$ telles que, pour tout $j\leq n$, on ait

$$\alpha_j\leq\frac{1}{j+1}\quad\text{et}\quad B_f(x_j,\alpha_j)\subset U_j\cap B(x_{j-1},\alpha_{j-1}). \tag{1.3}$$

L'ouvert U_{n+1} est dense, donc l'ouvert $U_{n+1}\cap B(x_n,\alpha_n)$ est non vide. Il existe donc un réel strictement positif α_{n+1} (que l'on peut supposer inférieur à $1/(n+2)$) et un point x_{n+1} de X telles que

$$B_f(x_{n+1},\alpha_{n+1})\subset B(x_n,\alpha_n)\cap U_{n+1}.$$

Le lemme 1.2.1 appliquée à la suite $F_n=B_f(x_n,\alpha_n)$ implique l'existence d'un point x appartenant à l'intersection des boules fermées $B_f(x_n,\alpha_n)$. Vu les relations (1.2) et (1.3), on a

$$\forall n\in\mathbb{N}\,,\ x\in V\cap\bigcap_{j\leq n}U_j.$$

Le théorème de Baire est ainsi démontré. □

On utilise souvent l'énoncé suivant qui n'est rien d'autre que le théorème de Baire "passé au complémentaire".

Théorème 1.2.3. *Soit (X, d) un espace métrique complet, on considère une suite $(F_n)_{n\in\mathbb{N}}$ de fermés d'intérieur vide dans X. Alors $\bigcup_{n\in\mathbb{N}} F_n$ est d'intérieur vide.*

Le théorème de Baire est souvent utilisé sous la forme de ce corollaire, dont la démonstration, très facile, est laissée en exercice.

Corollaire 1.2.1. *Soit $(F_n)_{n\in\mathbb{N}}$ une suite de fermés d'un espace métrique complet (X, d) dont la réunion est X. Alors il existe un entier n_0 tel que $\overset{\circ}{F}_{n_0} \neq \emptyset$.*

Dit autrement, un espace métrique complet n'est pas réunion dénombrable de fermés d'intérieur vide, ce qui est le cas d'un espace notoirement non complet, l'espace $\mathbb{Q}$.

Introduisons maintenant la notion de fonction uniformément continue.

Définition 1.2.3. *Soient (X, d) et (Y, δ) deux espaces métriques ; on considère une fonction f de X dans Y. On dit que la fonction f est uniformément continue si et seulement si*

$$\forall \varepsilon > 0\,,\ \exists \alpha > 0 \,/\ d(x, x') < \alpha \Longrightarrow \delta(f(x), f(x')) < \varepsilon;$$

Donnons quelques exemples et contre exemples. La fonction $x \longmapsto x^2$ est uniformément continue sur tout intervalle borné $[a, b]$ mais n'est pas uniformément continue sur $\mathbb{R}$. La fonction $x \longmapsto \sqrt{|x|}$ est uniformément continue sur $\mathbb{R}$.

Théorème 1.2.4 (de prolongement des fonctions uniformément continues)**.** *Soient (X, d) et (Y, δ) deux espaces métriques, A une partie dense de X, et f une application uniformément continue de (A, d) dans (Y, δ). Si Y est complet, alors il existe une unique application uniformément continue $\widetilde{f}$ de (X, d) dans (Y, δ) telle que $\widetilde{f}_{|A} = f$.*

Démonstration. Considérons un élément x de X et essayons de définir $\widetilde{f}(x)$. L'ensemble A étant dense, la proposition 1.1.5 nous assure qu'il existe une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A convergeant vers x. La fonction f est uniformément continue, ce qui signifie que

$$\forall \varepsilon > 0\,,\ \exists \alpha \,/\ \forall (a, b) \in A^2\,,\ d(a, b) < \alpha \Rightarrow \delta(f(a), f(b)) < \varepsilon.$$

Mais la suite $(a_n)_{n\in\mathbb{N}}$ est de Cauchy car convergente. Donc il existe un entier n_0 tel que

$$\forall n \geq n_0\,,\ \forall p\,,\ d(a_n, a_{n+p}) < \alpha.$$

Donc la suite $(f(a_n))_{n\in\mathbb{N}}$ est une suite de Cauchy de Y, donc elle converge vers une limite y.

Une première chose à vérifier : cette limite y est indépendante de la suite $(a_n)_{n\in\mathbb{N}}$ choisie. En effet, soient $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ deux suites convergeant vers x. Par un raisonnement analogue au précédent, l'uniforme continuité de la fonction f sur A assure que

$$\lim_{n\to\infty} \delta(f(a_n), f(b_n)) = 0.$$

Donc la limite y est bien indépendante du choix de la suite $(a_n)_{n\in\mathbb{N}}$. On peut alors définir la fonction $\widetilde{f}$ par

$$\widetilde{f}(x) = \lim_{n\to\infty} f(a_n) \text{ pour toute suite } (a_n)_{n\in\mathbb{N}} \in A^{\mathbb{N}} \,/\ \lim_{n\to\infty} a_n = x.$$

Comme la fonction f est continue sur A, il est clair que $\widetilde{f}_{|A} = f$. Vérifions maintenant que la fonction $\widetilde{f}$ est uniformément continue sur X. Le fait que la fonction f soit uniformément continue sur A se traduit par

$$\forall \varepsilon > 0\,,\ \exists \alpha > 0 \,/\, \forall (a,b) \in A^2\,,\ d(a,b) < \alpha \Rightarrow \delta(f(a), f(b)) < \varepsilon. \tag{1.4}$$

Considérons maintenant un couple (x, y) d'éléments de X tels que $d(x, y) < \alpha$. Il existe alors deux suites $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ d'éléments de A telles que

$$\lim_{n\to\infty} a_n = x \quad \text{et} \quad \lim_{n\to\infty} b_n = y.$$

Il existe donc un entier n_0 tel que

$$n \geq n_0 \Longrightarrow d(a_n, b_n) < \alpha.$$

Donc, d'après la relation (1.4), on a

$$\delta(f(a_n), f(b_n)) < \varepsilon.$$

Par passage à la limite, on obtient que

$$d(x, y) < \alpha \Longrightarrow \delta(\widetilde{f}(x), \widetilde{f}(y)) \leq \varepsilon,$$

ce qui achève la démonstration du théorème. □

1.3 La notion de compacité

Nous allons utiliser de manière cruciale le concept de valeur d'adhérence d'une suite.

Définition 1.3.1. *Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments d'un espace métrique (X, d). On définit l'ensemble (éventuellement vide) des valeurs d'adhérence de $(x_n)_{n\in\mathbb{N}}$ que l'on note* $\mathrm{Adh}(x_n)$ *par*

$$\mathrm{Adh}(x_n) \stackrel{\text{déf}}{=} \bigcap_{n\in\mathbb{N}} \overline{A}_n \quad \text{avec} \quad A_n \stackrel{\text{déf}}{=} \{x_m\,, m \geq n\}.$$

Exemples Dans $\mathbb{R}$ muni de la distance $|x - y|$, la suite $(x_n)_{n\in\mathbb{N}}$ définie par $x_n = n$ est telle que $\mathrm{Adh}(x_n) = \emptyset$ et la suite $(y_n)_{n\in\mathbb{N}}$ définie par $y_n = (-1)^{n+1}$ est telle que $\mathrm{Adh}(x_n) = \{-1, 1\}$.

Proposition 1.3.1. *Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'un espace métrique (X, d) que l'on suppose convergeant vers un point ℓ de X. Alors* $\mathrm{Adh}(x_n) = \{\ell\}$.

Démonstration. Par définition de la limite d'une suite, pour tout $\varepsilon > 0$, il existe n_0 tel que l'ensemble A_{n_0} de la définition ci-dessus soit inclus dans la boule ouverte $B(\ell, \varepsilon)$. Donc soit x un point de X distinct de ℓ, en prenant ε strictement inférieur à $d(\ell, x)$, on trouve que $x \notin \overline{A}_{n_0}$ et donc que $x \notin \overline{A}_n$ pour tout $n \geq n_0$. D'où la proposition. □

Remarque Il est possible qu'une suite $(x_n)_{n\in\mathbb{N}}$ soit telle que $\mathrm{Adh}(x_n) = \{\ell\}$ et ne converge pas. Considérons par exemple la suite de nombres réels définie par

$$x_{2n} = 2n \quad \text{et} \quad x_{2n+1} = \frac{1}{2n+1}\,\cdot$$

Bien que cette suite ne converge pas, on a $\mathrm{Adh}(x_n) = \{0\}$.

Proposition 1.3.2. *Dans un espace métrique (X,d), toute suite de Cauchy ayant une valeur d'adhérence ℓ converge vers ℓ.*

Démonstration. Considérons une suite de Cauchy $(x_n)_{n\in\mathbb{N}}$ et supposons qu'elle ait une valeur d'adhérence ℓ. La suite étant de Cauchy, il existe un entier n_0 tel que

$$\forall n \geq n_0\,,\ \forall m \geq n_0\,,\ d(x_n,x_m) \leq \frac{\varepsilon}{2}\cdot$$

La suite ayant ℓ pour valeur d'adhérence, il existe un entier $n_1 \geq n_0$ tel que

$$d(x_{n_1},\ell) < \frac{\varepsilon}{2}\cdot$$

On en déduit alors que, pour tout entier $n \geq n_0$, on a

$$d(x_n,\ell) \leq d(x_n,x_{n_1}) + d(x_{n_1},\ell) < \varepsilon.$$

D'où la proposition. □

Proposition 1.3.3. *Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments d'un espace métrique (X,d). Un point ℓ de X appartient à $\mathrm{Adh}(x_n)$ si et seulement si il existe une fonction strictement croissante ϕ de $\mathbb{N}$ dans $\mathbb{N}$ telle que*

$$\lim_{n\to\infty} x_{\phi(n)} = \ell.$$

Démonstration. Avant d'entamer la démonstration proprement dite, remarquons qu'une fonction strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ vérifie

$$\forall n \in \mathbb{N}\,,\ \phi(n) \geq n. \tag{1.5}$$

Démontrons cette propriété par récurrence. Elle est bien sûr vraie pour $n = 0$. Supposons la réalisée pour n. Comme ϕ est strictement croissante, on a $\phi(n+1) > \phi(n) \geq n$. Ceci implique que $\phi(n+1) \geq n+1$. En particulier, la suite $(\phi(n))_{n\in\mathbb{N}}$ tend vers l'infini avec n.

Revenons à notre démonstration. Si ℓ appartient à $\mathrm{Adh}(x_n)$, on définit par récurrence la fonction de la manière suivante : on choisit $\phi(0) = 0$ et puis, on définit $\phi(n+1)$ à partir de $\phi(n)$ comme

$$\phi(n+1) \stackrel{\text{déf}}{=} \min\Big\{m > \phi(n)\,/ d(x_m,\ell) < \frac{1}{n+1}\Big\}\cdot$$

Par construction, la fonction ϕ est strictement croissante et nous avons, pour tout $n \geq 1$,

$$d(x_{\phi(n)},\ell) \leq \frac{1}{n} \quad \text{et donc} \quad \lim_{n\to\infty} x_{\phi(n)} = \ell.$$

Réciproquement, s'il existe une fonction strictement croissante ϕ telle que $\lim\limits_{n\to\infty} x_{\phi(n)} = \ell$, pour tout $\varepsilon > 0$, il existe n_0 tel que

$$\forall n \geq n_0\,,\ d(x_{\phi(n)},\ell) < \varepsilon.$$

Comme, d'après (1.5), $\phi(n) \geq n$, on en déduit que

$$\forall \varepsilon > 0\,, \forall n \in \mathbb{N}\,,\ \exists m \geq n\,/\ d(x_m,\ell) < \varepsilon,$$

ce qui signifie exactement que $\ell \in \mathrm{Adh}(x_n)$. D'où le théorème. □

Définition 1.3.2. *On appelle fonction d'extraction une fonction strictement croissante de* $\mathbb{N}$ *dans* $\mathbb{N}$. *Si* $(x_n)_{n\in\mathbb{N}}$ *est une suite d'éléments d'un ensemble* X, *on appelle suite extraite toute suite du type* $(x_{\phi(n)})_{n\in\mathbb{N}}$ *où* ϕ *est une fonction d'extraction.*

La notion d'espace compacts peut être vue de deux façons différentes : l'une avec les suites, l'autre avec des recouvrements par des boules. L'équivalence de deux points de vue est garanti par le théorème suivant.

Théorème 1.3.1. *Soit* (X, d) *un espace métrique. Les deux conditions suivantes sont équivalentes :*
i) toute suite d'éléments de (X, d) *admet une valeur d'adhérence*
ii) L'espace métrique (X, d) *est complet et*

$$\forall \varepsilon > 0\,,\ \exists (x_j)_{1\leq j\leq N} \,/\, X = \bigcup_{j=1}^{N} B(x_j, \varepsilon). \tag{1.6}$$

Définition 1.3.3. *Soit* (X, d) *un espace métrique. On dit que* X *est un espace compact si et seulement si l'une de deux conditions du théorème ci-dessus est satisfaite.*

Donnons des exemples d'espaces métriques compacts en commençant par remarquer que tout espace métrique fini est compact.

Théorème 1.3.2. *Les espaces métriques* $([a, b], |x - y|)$ *sont des espaces métriques compacts.*

Démonstration. Si l'on utilise le point *ii)*, ce que l'on peut faire puisque l'on sait que espace métrique $([a, b], |x - y|)$ est complet, il suffit d'observer que, pour tout ε strictement positif,

$$[a, b] = [a, a+\varepsilon[\cup\Bigg(\bigcup_{k=1}^{N_\varepsilon-1}]a + (k-1)\varepsilon, a + (k+1)\varepsilon[\Bigg)\cup]b - N_\varepsilon\varepsilon, b] \quad \text{avec} \quad N_\varepsilon \stackrel{\text{déf}}{=} \left[\frac{b-a}{\varepsilon}\right]\cdot$$

L'utilisation du critère *i)* demande un tout petit peu plus de travail. Soit $(x_n)_{n\in\mathbb{N}}$ une suite de l'intervalle $[a, b]$, on considère l'ensemble $A_n \stackrel{\text{déf}}{=} \{x_m\,,\ m \geq n\}$ qui est une partie majorée de réels qui admet donc une borne supérieure que l'on note M_n. La suite $(M_n)_{n\in\mathbb{N}}$ est décroissante donc convergente. Le lecteur se convaincra aisément que la limite de la suite $(M_n)_{n\in\mathbb{N}}$ est une valeur d'adhérence (c'est même la plus grande). □

Démonstration du théorème 1.3.1. D'après la proposition 1.3.2, pour démontrer que *i)* implique *ii)*, il suffit de démontrer que *i)* implique l'assertion (1.6). Nous allons procéder par contraposition. Supposons que l'assertion (1.6) ne soit pas satisfaite. Ceci se traduit par l'existence d'un réel strictement positif α tel que l'on ne puisse recouvrir X par un nombre fini de boules de rayon α. Soit x_0 un élément quelconque de X. Il existe un élément x_1 de X n'appartenant pas à $B(x_0, \alpha)$. Soit $(x_0, \cdots, x_p)$ un p-uplet d'éléments de X tel que, pour tout m différent de n, on ait $d(x_m, x_n) \geq \alpha$. Par hypothèse,

$$\bigcup_{n=1}^{p} B(x_n, \alpha) \neq X.$$

Donc, il existe un point x_{p+1} de X qui n'appartient pas à la réunion de boules ci-dessus. Par récurrence, nous construisons ainsi une suite $(x_n)_{n\in\mathbb{N}}$ telle que

$$m \neq n \Rightarrow d(x_m, x_n) \geq \alpha.$$

Une telle suite n'a bien sûr pas de valeur d'adhérence. Nous avons donc démontré que *i)* implique *ii)*.

Supposons maintenant *ii)* et considérons une suite infinie $(x_n)_{n\in\mathbb{N}}$ d'éléments de X. L'hypothèse de recouvrement implique l'existence d'un élément α_0 de X tel que l'ensemble $\mathcal{X}_0$ défini par

$$\mathcal{X}_0 \stackrel{\text{déf}}{=} \Big\{m \,/\, x_m \in F_0 \stackrel{\text{déf}}{=} B_f\Big(\alpha_0, \frac{1}{2}\Big)\Big\}$$

soit infini. On pose alors

$$\phi(0) \stackrel{\text{déf}}{=} \min \mathcal{X}_0.$$

De même, il existe un élément α_1 de X tel que l'ensemble $\mathcal{X}_1$ définie par

$$\mathcal{X}_1 \stackrel{\text{déf}}{=} \Big\{m \in \mathcal{X}_0 \,/\, x_m \in F_1 \stackrel{\text{déf}}{=} F_0 \cap B_f\Big(\alpha_1, \frac{1}{4}\Big)\Big\}$$

soit infini. On pose alors

$$\phi(1) \stackrel{\text{déf}}{=} \min\big\{m \in \mathcal{X}_1 \,/\, m > \phi(0)\big\}.$$

Supposons construit une suite de fermés $(F_m)_{1\leq m\leq n}$ de X, une suite strictement croissante d'entiers $(\phi(m))_{1\leq m\leq n}$ tels que $\delta(F_m) \leq 2^{-m}$ et telle que la suite $(\mathcal{X}_m)_{1\leq m\leq n}$ de sous ensemble de N définie par

$$\mathcal{X}_m \stackrel{\text{déf}}{=} \{m' \in \mathcal{X}_{m-1} /\, x_{m'} \in F_{m-1}\}$$

soit infini. Il existe alors un élément α_{n+1} de A tel que

$$\mathcal{X}_{n+1} \stackrel{\text{déf}}{=} \Big\{m \,/\, x_m \in F_{n+1} \stackrel{\text{déf}}{=} F_n \cap B_f\Big(\alpha_{n+1}, \frac{1}{2^{n+1}}\Big)\Big\}$$

soit infini. On pose alors

$$\phi(n+1) \stackrel{\text{déf}}{=} \min\{m \in \mathcal{X}_{n+1} \,/\, m > \phi(n)\}.$$

On a donc pour tout entier n,

$$x_{\phi(n)} \in F_n \quad \text{et} \quad \delta(F_n) \leq \frac{1}{2^n} \cdot$$

D'après le lemme 1.2.1, il existe un point ℓ appartenant à l'intersection de tous les F_n. Par définition du diamètre, on a

$$\forall n \in \mathbb{N}\,,\ d(x, x_{\phi(n)}) \leq \frac{1}{2^n} \cdot$$

Ceci conclut la démonstration du théorème. □

La proposition suivante permet de trouver beaucoup d'espaces compacts.

Proposition 1.3.4. *Soit* $(X_1, d_1), \cdots, (X_N, d_N)$ *une famille de* N *espaces métriques compacts. Posons*

$$X = X_1 \times \cdots \times X_N \quad \textit{et} \quad d((x_1, \cdots, x_N), (y_1, \cdots, y_N)) = \max_{1\leq j\leq N} d_j(x_j, y_j).$$

L'espace métrique (X, d) *est compact.*

Démonstration. Soit $(x_n)_{n\in\mathbb{N}}$ une suite déléments de X. Par définition, il existe pour chaque j une suite $(x_n^j)_{n\in\mathbb{N}}$ de X_j telle que $x_n = (x_n^1, \cdots, x_n^N)$. Par définition d'un espace métrique compacte, il existe une fonction d'extraction ϕ_1 et un point x^1 de X_1 tel que

$$\lim x^1_{\phi_1(n)} = x^1.$$

De même, il existe une fonction d'extraction ϕ_2 et un point a_2 de A_2 tel que

$$\lim x^2_{\phi_1\circ\phi_2(n)} = x^2.$$

En itérant N fois le processus, on construit une fonction d'extraction ϕ en posant

$$\phi \stackrel{\text{déf}}{=} \phi_1 \circ \cdots \circ \phi_N$$

telle que, pour tout j dans $\{1, \cdots, N\}$, on ait

$$\lim_{n\to\infty} x^j_{\phi(n)} = x^j.$$

Par définition de la distance sur X, cela implique que la suite $x_{\phi(n)}$ converge vers $(x^1, \cdots, x^N)$. D'où le résultat. $\square$

Définition 1.3.4. *Soit (X, d) un espace métrique. On dit que A est une partie compacte de X si et seulement si l'espace métrique $(A, d_{|A})$ est compact.*

Proposition 1.3.5. *Soit A une partie d'un espace métrique (X, d). Si A est compacte, alors A est fermée et de diamètre fini ce qui signifie que*

$$\delta(A) \stackrel{\text{déf}}{=} \sup_{(a,a')\in A^2} d(a, a') < \infty.$$

Démonstration. Soit x un point de $\overline{A}$. D'après la proposition 1.1.5, il existe une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A qui converge vers x. Comme A est supposée compacte, il existe une fonction d'extraction ϕ et un point ℓ de A tel que $\lim\limits_{n\to\infty} a_{\phi(n)} = \ell$. Comme la suite $(a_{\phi(n)})_{n\in\mathbb{N}}$ converge vers x, l'unicité de la limite implique que $\ell = x$ et donc que $x \in A$. Donc A est une partie fermée de X.

Démontrons maintenant qu'une partie compacte est de diamètre fini. Considérons $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ deux suites d'éléments de A telles que

$$\lim_{n\to\infty} d(a_n, b_n) = \delta(A).$$

Comme A est compact, il existe une fonction d'extraction ϕ_0 et un point a de A tel que la suite $(a_{\phi_0(n)})_{n\in\mathbb{N}}$ converge vers a. De même, il existe une fonction d'extraction ϕ_1 et un point b de A tel que la suite $(b_{\phi_0\circ\phi_1(n)})_{n\in\mathbb{N}}$ converge vers b. Par passage à la limite, on a

$$\delta(A) = \lim_{n\to\infty} d(a_{\phi_0\circ\phi_1(n)}, b_{\phi_0\circ\phi_1(n)}) = d(a, b).$$

Donc A est de diamètre fini. $\square$

Proposition 1.3.6. *Soit A une partie d'un espace métrique complet (X, d). Son adhérence $\overline{A}$ est compact si et seulement si pour tout réel strictement positif ε, il existe une famille finie $(x_j)_{1\leq j\leq N}$ de points de X telle que*

$$A \subset \bigcup_{j=1}^{N} B(x_j, \varepsilon)\,.$$

Démonstration. Si $\overline{A}$ est compact, c'est simplement la définition 1.3.3. Réciproquement, supposons la propriété de recouvrement pour une partie A. Montrons alors qu'elle est vrai pour son adhérence $\overline{A}$.

Par hypothèse, pour tout réel strictement positif ε, il existe une suite finie $(x_j)_{1\leq j\leq N}$ d'éléments de X telle que

$$A \subset \bigcup_{j=1}^{N} B\Big(x_j, \frac{\varepsilon}{2}\Big)\cdot$$

Soit x un point de $\overline{A}$. Par définition, il existe un point de A tel que $d(x,A)$ soit strictement inférieur à $\varepsilon/2$. Il existe donc un entier j dans $\{1,\cdots,N\}$ tel que a apartienne à $B(x_j,\varepsilon/2)$. Ainsi dont

$$d(x,x_j) \leq d(x,a) + d(a,x_j) < \varepsilon$$

et donc x appartient à $B(x_j,\varepsilon)$ ce qui assure la propriété de recouvrenemt pour $\overline{A}$. Comme l'espace métrique (X,d) est complet, il suffit pour démonter la compacité de $\overline{A}$ de démontrer que l'on a la propriété de recouvrement avec des x_j dans l'ensemble $\overline{A}$. On sait que pour tout réel strictement positif ε, il existe une suite finie $(x_j)_{1\leq j\leq N}$ d'éléments de X telle que

$$\overline{A} \subset \bigcup_{j=1}^{N} B\Big(x_j, \frac{\varepsilon}{2}\Big)\cdot$$

On ne retient bien sûr que les x_j telle que la boule ouverte de centre x_j et de rayon $\varepsilon/2$ rencontre $\overline{A}$. Désignons par a_j un point quelconque de cette intersection. Pour tout a dans $\overline{A}$, il existe j tel que

$$d(a,a_j) \leq d(a,x_j) + d(x_j,a_j) < \varepsilon$$

ce qui achève la démonstration. □

Du théorème 1.3.2 et des propositions 1.3.4 et 1.3.5, on déduit le corollaire suivant.

Corollaire 1.3.1. *Une partie fermée A de $\mathbb{R}^d$ est compacte si et seulement si*

$$\exists r > 0 \,/\, A \subset [-r,r]^d.$$

Exercice 1.3.1. *Soit A une partie d'un espace métrique (X,d). Le fermé $\overline{A}$ est compact si et seulement si toute suite d'éléments de A possède une valeur d'adhérence dans X.*

Proposition 1.3.7. *Soit A une partie compacte d'un espace métrique (X,d). Si B est un fermé de X inclus dans A, alors B est aussi compact.*

Démonstration. Soit $(b_n)_{n\in\mathbb{N}}$ une suite d'éléments de B. Comme B est inclus dans A qui est compacte, il existe une fonction d'extraction ϕ telle que $(b_{\phi(n)})_{n\in\mathbb{N}}$ converge vers ℓ qui appartient à A. Comme B est un fermé de X, alors $\ell \in \overline{B} = B$. Donc B est compacte. □

Théorème 1.3.3 (de Heine)**.** *Soient (X,d) et (Y,δ) deux espaces métriques et f une application continue de X dans Y. Alors, pour toute partie compacte A de X, $f(A)$ est une partie compacte de Y. De plus, si X est compact, alors f est une application uniformément continue de X dans Y.*

Démonstration. Soit $(y_n)_{n\in\mathbb{N}}$ une suite d'éléments de $f(A)$, il existe une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A telle que $y_n = f(a_n)$. Comme A est une partie compacte de X, il existe un point a' de A et une fonction d'extraction ϕ telle que

$$\lim_{n\to\infty} a_{\phi(n)} = a'$$

La fonction f étant continue, on a

$$\lim_{n\to\infty} y_{\phi(n)} = f(a') \in f(A).$$

Ceci démontre le premier point du théorème.

Pour démontrer le second point, procédons par contraposition. Supposons l'espace métrique (X,d) compact et considérons une fonction f sur X qui n'est pas uniformément continue. Ceci implique l'existence d'un réel strictement positif ε_0 et de deux suites $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ d'éléments de X telles que

$$\forall n \in \mathbb{N}\,,\ d(x_n,y_n) \leq \frac{1}{n+1} \quad \text{et} \quad \delta(f(x_n,y_n) \geq \varepsilon_0.$$

Comme l'espace métrique (X,d) est compact, il existe une fonction d'extraction ϕ et un point x de X tel que $\lim\limits_{n\to\infty} x_{\phi(n)}$ converge vers x. Comme la distance entre x_n et y_n tend vers 0, la suite $(y_{\phi(n)})_{n\in\mathbb{N}}$ converge elle aussi vers x. De plus l'inégalité triangulaire assure que

$$\delta(f(x_{\phi(n)}),f(x)) + \delta(f(x),f(y_{\phi(n)})) \geq \delta(f(x_{\phi(n)}),f(y_{\phi(n)})) \geq \varepsilon_0$$

ce qui implique que la fonction f n'est pas continue en x d'après la proposition 1.1.2. D'où le théorème. □

Corollaire 1.3.2. *Soit f une fonction continue de (X,d) dans $(\mathbb{R},|x-y|)$. Si (X,d) est compact, alors la fonction f admet un minimum et un maximum, c'est-à-dire qu'il existe x_m et x_M dans X tels que*

$$\forall x \in X\,,\ f(x_m) \leq f(x) \leq f(x_M).$$

Démonstration. D'après le théorème 1.3.3 de Heine, l'ensemble $f(X)$ est un compact de $\mathbb{R}$, et donc est un fermé inclus dans un intervalle $[a,b]$. Donc cet ensemble admet une borne inférieure et une borne supérieure qui sont des points de l'adhérence du fermé $f(X)$ donc qui appartiennent à $f(X)$. □

Exercice 1.3.2. *Soit f une fonction d'un espace métrique compact (X,d) à valeurs dans $\mathbb{R}$. On suppose que, pour tout x_0 de X, on ait la propriété suivante :*

$$\forall \varepsilon > 0\,,\ \exists \alpha > 0 \,/\, d(x,x_0) < \alpha \Longrightarrow f(y) > f(x) - \varepsilon.$$

Alors la fonction f est minorée et atteint son minimum sur X.

Nous allons maintenant donner une caractérisation de la compacité en terme de recouvrements par des ouverts.

Théorème 1.3.4. *Soit (X,d) un espace métrique, les trois propriétés suivantes sont équivalentes.*

i) Pour toute famille $(U_\lambda)_{\lambda\in\Lambda}$ d'ouverts de X recouvrant X, c'est-à-dire telle que

$$X = \bigcup_{\lambda\in\Lambda} U_\lambda,$$

on peut extraire un sous-recouvrement fini, c'est-à-dire qu'il existe une famille finie $(\lambda_j)_{1\leq j\leq N}$ telle que

$$X = \bigcup_{j=1}^{N} U_{\lambda_j}.$$

ii) Pour toute famille de fermés $(F_\lambda)_{\lambda\in\Lambda}$ on a

$$\forall N\,,\ \forall(\lambda_1\cdots,\lambda_N)\in\Lambda^N\,/\ \bigcap_{j=1}^N F_{\lambda_j}\neq\emptyset \Longrightarrow \bigcap_{\lambda\in\Lambda} F_\lambda\neq\emptyset.$$

iii) Toute suite d'éléments de X admet une valeur d'adhérence.

Démonstration. L'équivalence entre les points $i)$ et $ii)$ se démontrent par contraposition. En effet, par contraposition le point $ii)$ est équivalent à :

Pour toute famille de fermés $(F_\lambda)_{\lambda\in\Lambda}$,

$$\bigcap_{\lambda\in\Lambda} F_\lambda=\emptyset \Longrightarrow \exists(\lambda_1\cdots,\lambda_N)\in\Lambda^N\,/\ \bigcap_{j=1}^N F_{\lambda_j}=\emptyset.$$

Par passage au complémentaire, ceci est équivalent au point $i)$.

Démontrons maintenant que $ii)$ implique $iii)$. Par définition de l'ensemble des valeurs d'adhérence, on a

$$\mathrm{Adh}(x_n)=\bigcap_n \overline{A}_n \quad \text{avec} \quad A_n \overset{\text{déf}}{=} \{x_m\,, m\geq n\}.$$

La suite $\overline{A}_n$ est une suite décroissante de fermés non vides. Donc, d'après la propriéte $ii)$, l'intersection de tous les $\overline{A}_n$ (qui est l'ensemble des valeurs d'adhérence) est non vide. D'où le point $iii)$.

Démontrons maintenant que $iii)$ implique $i)$ ce qui constitue le point délicat de la démonstration. Il s'agit de déduire d'une propriété sur les suites, une propriété sur les recouvrements d'ouverts. Le lemme suivant est crucial, et peut aussi être utile utile pour démontrer d'autres propriétés des espaces compacts.

Lemme 1.3.1 (de Lebesgue)**.** *Soit (X,d) un espace métrique compact. Pour toute famille d'ouverts $(U_\lambda)_{\lambda\in\Lambda}$ recouvrant X, il existe un nombre réel α strictement positif tel que*

$$\forall x\in X\,,\ \exists\lambda\in\Lambda\,/\ B(x,\alpha)\subset U_\lambda.$$

Démonstration. Pour démontrer ce lemme, considérons la fonction δ définie par

$$\delta\left\{\begin{array}{ccl} X & \longrightarrow & \mathbb{R}_+^\star \\ x & \longmapsto & \sup\{\beta\,/\ \exists\lambda\,/\ B(x,\beta)\subset U_\lambda\}.\end{array}\right.$$

Démontrer que la fonction δ possède un minimum implique le lemme. Pour démontrer cela, établissons tout d'abord l'assertion suivante.

$$\forall\varepsilon\,,\ \forall x\in X\,,\ \exists\beta\,/\ d(x,y)<\beta\Longrightarrow \delta(y)>\delta(x)-\varepsilon. \tag{1.7}$$

Soient x un point de X et ε un réel strictement positif inférieur à $\delta(x)$. Supposons que y soit un point de X tel que $d(x,y)<\varepsilon/2$. L'inégalité triangulaire implique alors que

$$B(y,\delta(x)-\varepsilon)\subset B(x,\delta(x)-\varepsilon/2)\subset U_\lambda.$$

Et donc l'assertion (1.7) est démontrée.

Considérons maintenant une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X telle que $\lim_{n\to\infty}\delta(x_n)=\inf_{x\in X}\delta(x)$. Par hypothèse, on peut extraire de la suite $(x_n)_{n\in\mathbb{N}}$ une sous-suite convergeant vers un point x_∞ de X. On note toujours $(x_n)_{n\in\mathbb{N}}$ cette suite extraite. D'après l'assertion (1.7), pour tout $\varepsilon>0$, il existe un entier n_0 tel que

$$n\geq n_0\Longrightarrow \delta(x_n)>\delta(x_\infty)-\varepsilon.$$

Par passage à la limite, on en déduit que

$$\inf_{x\in X}\delta(x)\geq\delta(x_\infty)-\varepsilon,$$

et ce pour tout réel strictement positif ε. Par définition de la borne inférieure, il en résulte que $\delta(x_\infty)=\inf_{x\in X}\delta(x)$ et donc que $\inf_{x\in X}\delta(x)$ est strictement positif. Ceci conclut la démonstration du lemme. □

Conclusion de la démonstration du théorème 1.3.4

Pour montrer que iii) implique i), nous supposons que X est compact (de toute suite on peut extraire une suite convergente), et nous considérons une famille $(U_\lambda)_{\lambda\in\Lambda}$ d'ouverts de X recouvrant X, dont il s'agit d'extraire un sous-recouvrement fini. Soit $\alpha > 0$ donné par le lemme de Lebesgue. D'après la définition 1.3.3, il existe un nombre fini de boules $B_1, \ldots, B_N$ de rayon α qui recouvrent X. D'après la propriété de α donnée par le lemme de Lebesgue, chacune de ces boules B_j est incluse dans l'un des ouvert U_{λ_j} de la famille, et par conséquent

$$X \subset \bigcup_{j=1}^{N} B_j \subset \bigcup_{j=1}^{N} U_{\lambda_j}.$$

Ceci achève la démonstration du théorème. □

Exercice 1.3.3. *Soit (X, d) un espace métrique. On considère l'ensemble $X^{\mathbb{N}}$ des suites d'éléments de X et la distance sur $X^{\mathbb{N}}$ définie à la proposition 1.1.1.*

1) Démontrer que la suite $(x_p)_{p\in\mathbb{N}}$ est de Cauchy si et seulement si, pour tout entier n, la suite $(x_p(n))_{p\in\mathbb{N}}$ est une suite de Cauchy.

3) En déduire que si l'espace métrique (X, d) est complet, alors $(X^{\mathbb{N}}, d_a)$ aussi.

4) Démontrez que si l'espace métrique (X, d) est compact, alors $(X^{\mathbb{N}}, d_a)$ aussi.

5) Démontrez que U est un ouvert de $(X^{\mathbb{N}}, d_{\mathbb{N}})$ si et seulement si pour tout $x \in U$, il existe une partie finie J de $\mathbb{N}$ et un réel strictement positif α tel que

$$\forall y \in X^{\mathbb{N}}\,,\ \forall j \in J\,,\ d(y(j), x(j)) < \alpha \Longrightarrow y \in U.$$

Chapitre 2

Espaces normés, espaces de Banach

Ce chapitre est consacré à l'étude des espaces de Banach qui sont les espaces vectoriels normés complets pour la distance associée à la norme. Ce sont les objets de base de l'analyse fonctionnelle. Ce chapitre constitue la première utilisation sur des exemples concrets des concepts fondamentaux de la topologie métrique introduits au chapitre précédent.

Dans la première section, la notion de norme et de distance associée est introduite et outre les espaces $\mathbb{R}^N$, sont présentés et étudiés les espaces de fonctions bornées et continues ainsi que les espaces de suites de puissance $p$ième intégrable. Ce sont des exemples élémentaires mais importants que le lecteur devra garder à l'esprit comme modèle d'espaces de Banach.

Dans la seconde section, nous étudions l'espace des applications linéaires continues entre espaces de Banach. Il est important pour la suite de bien assimiler le critère de continuité des applications linéaires et la notion de norme d'une applications linéaire. Cette section se conclut par un critère de continuité pour les applications multilinéaires.

Dans la troisième section, nous étudions le cas des espaces vectoriels de dimension finie N et démontrons qu'ils sont tous identifiables à $\mathbb{R}^N$.

La quatrième section est consacrée au théorème d'Ascoli qui fournit un critère de compacité pour les parties de l'espace des fonctions continues d'un compact à valeurs dans un espace de Banach. Ce théorème permet de mesurer combien la compacité dans les espaces de fonctions continues est loin d'être équivalente à être fermé et borné comme dans le cas de la dimension finie.

La cinquième section concerne des résultats de densité toujours dans l'espaces des fonctions continues d'un compact et cette fois à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$. C'est alors l'occasion d'introduire la notion d'espace séparable qui jouera un rôle au chapitre suivant.

Enfin, convenons que dans toute la suite du cours, $\mathbb{K}$ désignera $\mathbb{R}$ ou $\mathbb{C}$.

2.1 Définition des espaces normés et des espaces de Banach

Définition 2.1.1. *Soit E un espace vectoriel sur $\mathbb{K}$. On dit qu'une application N de E dans $\mathbb{R}^+$ est une semi-norme sur E si et seulement si les deux conditions suivantes sont satisfaites*

— $N(x+y) \leq N(x) + N(y)$,

— *Pour tout λ appartenant à $\mathbb{K}$, on a $N(\lambda x) = |\lambda| N(x)$.*

On dit qu'une application N de E dans $\mathbb{R}^+$ est une norme sur E si et seulement si c'est une semi-norme telle que $N(x) = 0$ implique que $x = 0$.

Définition 2.1.2. *Soient E un espace vectoriel et N une norme sur E ; le couple (E, N) est appelé un espace normé.*

Définition 2.1.3. *Soit N_1 et N_2 deux normes sur un espace vectoriel E. On dit que les deux normes N_1 et N_2 sont équivalentes si et seulement si il existe une constante C telle que*

$$\forall x \in E\,,\ C^{-1}N_1(x) \leq N_2(x) \leq CN_1(x).$$

Notations Très souvent, on désigne une norme par $\|\cdot\|_E$ ou bien par $\|\cdot\|$.

Proposition 2.1.1. *Soit $(E, \|\cdot\|)$ un espace normé, alors l'application définie par*

$$\begin{cases} E\times E & \longrightarrow & \mathbb{R}^+ \\ (x,y) & \longmapsto & \|x-y\| \end{cases}$$

est une distance sur E, appelée distance associée à la norme $\|\cdot\|$.

L'évidente démonstration est laissée en exercice.

Convention : Sauf mention expresse du contraire, on considérera toujours l'espace E muni de la structure métrique ainsi définie.

Définition 2.1.4. *Soit $(E, \|\cdot\|)$ un espace normé. On dit que $(E, \|\cdot\|)$ est un espace de Banach si et seulement si l'espace métrique (E, d) où d est la distance associée à la norme $\|\cdot\|$ (i.e. $d(x, y) = \|x - y\|$) est un espace complet.*

Nous allons maintenant donner une suite d'exemple d'espaces de Banach.

Proposition 2.1.2. *Pour p dans l'intervalle $[1, \infty[$, on considère l'application définie par*

$$\|\cdot\|_{\ell^p}\ \begin{cases} \mathbb{K}^N & \longrightarrow & \mathbb{R}^+ \\ (x_j)_{1\leq j\leq N} & \longmapsto & \Big(\displaystyle\sum_{j=1}^N |x_j|^p\Big)^{\frac{1}{p}}. \end{cases}$$

C'est une norme sur $\mathbb{K}^N$ et $(\mathbb{K}^N, \|\cdot\|_{\ell^p})$ est un espace de Banach. De plus, on a, pour tout (p, q) de $[1, \infty[^2$ tels que $p \leq q$,

$$\|x\|_{\ell^\infty} \stackrel{\text{déf}}{=} \sup_{1\leq j\leq N} |x_j| \leq \|x\|_{\ell^q} \leq \|x\|_{\ell^p} \leq N^{1-\frac{1}{q}}\|x\|_{\ell^q} \leq N \sup_{1\leq j\leq N} \|x\|_{\ell^\infty}. \tag{2.1}$$

Démonstration. Commençons par démontrer l'inégalité dite de Hölder. Pour tout p dans l'intervalle $[1, \infty]$, on a, pour tout (a, b) dans $\mathbb{K}^N \times \mathbb{K}^N$,

$$\Big|\sum_{j=1}^N a_j b_j\Big| \leq \|a\|_{\ell^p}\|b\|_{\ell^{p'}} \quad \text{avec} \quad \frac{1}{p} + \frac{1}{p'} = 1. \tag{2.2}$$

Pour démontrer cette inégalité (qui est évidente si $p = 1$ ou $p = \infty$), observons que la concavité du logarithme implique que

$$\forall (a, b) \in]0, \infty[^2\,,\ \forall \theta \in [0, 1]\,,\ \theta \log a + (1 - \theta) \log b \leq \log(\theta a + (1 - \theta)b).$$

L'exponantielle étant une fonction croissante, on en déduit

$$a^\theta b^{1-\theta} \leq \theta a + (1 - \theta)b$$

ce qui est souvent utilisé sous la forme équivalente pour θ dans l'intervalle $]0,1[$,

$$ab \leq \theta a^{\frac{1}{\theta}} + (1-\theta) b^{\frac{1}{1-\theta}}. \tag{2.3}$$

Par division, on peut supposer que $\|a\|_{\ell^p} = \|b\|_{\ell^{p'}} = 1$. On écrit alors

$$\begin{aligned} \Big|\sum_{j=1}^N a_j b_j\Big| &\leq \sum_{j=1}^N |a_j|\,|b_j| \\ &\leq \frac{1}{p}\sum_{j=1}^N |a_j|^p + \Big(1-\frac{1}{p}\Big)|b_j|^{p'} \\ &\leq 1. \end{aligned}$$

L'inégalité (2.2) étant démontrée, on écrit que

$$\begin{aligned} \sum_{j=1}^N |x_j+y_j|^p &\leq \sum_{j=1}^N |x_j|\,|x_j+y_j|^{p-1} + \sum_{j=1}^N +|y_j|\,|x_j+y_j|^{p-1} \\ &\leq \bigg(\Big(\sum_{j=1}^N |x_j|^p\Big)^{\frac{1}{p}} + \Big(\sum_{j=1}^N |y_j|^p\Big)^{\frac{1}{p}}\bigg)\Big(\sum_{j=1}^N |x_j+y_j|^{(p-1)\frac{p}{p-1}}\Big)^{1-\frac{1}{p}} \\ &\leq \bigg(\Big(\sum_{j=1}^N |x_j|^p\Big)^{\frac{1}{p}} + \Big(\sum_{j=1}^N |y_j|^p\Big)^{\frac{1}{p}}\bigg)\Big(\sum_{j=1}^N |x_j+y_j|^p\Big)^{1-\frac{1}{p}}. \end{aligned}$$

Par simplification, on obtient le résultat pour p dans l'intervalle $]1,\infty[$, les deux cas $p=1$ ou $p=\infty$ étant évidents.

Démontrons maintenant (2.1). Pour tout (p,q) de $[1,\infty]^2$ tel que $p<q$ on a

$$\frac{x_j}{\|x\|_{\ell^q}} \leq 1.$$

Ceci implique que

$$\Big|\frac{x_j}{\|x\|_{\ell^q}}\Big|^q \leq \Big|\frac{x_j}{\|x\|_{\ell^q}}\Big|^p.$$

Par sommation, on obtient que

$$1 \leq \Big(\frac{\|x\|_{\ell^p}}{\|x\|_{\ell^q}}\Big)^p$$

ce qui assure le résultat. □

Proposition 2.1.3. *Soit X un ensemble et $(E,\|\cdot\|)$ un espace de Banach, on considère $\mathcal{B}(X,E)$ l'ensemble des fonctions f bornées de X dans E, c'est-à-dire telles que*

$$\|f\|_{\mathcal{B}(X,E)} \stackrel{\text{déf}}{=} \sup_{x\in X}\|f(x)\| < \infty.$$

Alors $(\mathcal{B}(X,E),\|\cdot\|_{\mathcal{B}(X,E)})$ est un espace de Banach. De plus si X est muni d'une structure d'espace métrique par une distance d, on définit $\mathcal{C}_b(X,E)$ comme étant l'ensemble des fonctions f de $\mathcal{B}(X,E)$ qui sont continues. Alors $(\mathcal{C}_b(X,E),\|\cdot\|_{\mathcal{B}(X,E)})$ est un espace de Banach.

Démonstration. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy de $\mathcal{B}(X,E)$. Par définition, on a

$$\forall\varepsilon\,,\ \exists n_0\,/\ \forall n\geq n_0\,,\ \forall p\,,\ \forall x\in X\,,\ \|f_n(x)-f_{n+p}(x)\|<\varepsilon. \tag{2.4}$$

En particulier, pour tout x, la suite $(f_n(x))_{n\in\mathbb{N}}$ est une suite de Cauchy de E. Comme cet espace est complet, cette suite est convergente. Donc, pour tout x de X, il existe un élément de E, noté $f(x)$, tel que l'on ait

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

Vérifions que f appartient à $\mathcal{B}(X,E)$. D'après l'inégalité (2.4) appliquée avec $\varepsilon=1$, on a

$$\forall p\,,\ \forall x\in X\,,\ \|f_{n_0}(x)-f_{n_0+p}(x)\|<1.$$

Par passage à la limite lorsque p tend vers l'infini, on obtient

$$\forall x\in X\,,\ \|f_{n_0}(x)-f(x)\|\leq 1.$$

La fonction f_{n_0} étant bornée, la fonction f l'est également. En effet, on a

$$\begin{aligned}\|f(x)-f(x')\| &\leq \|f(x)-f_{n_0}(x)\|+\|f_{n_0}(x)-f_{n_0}(x')\|+\|f_{n_0}(x')-f(x')\|\\ &\leq 2+\|f_{n_0}(x)-f_{n_0}(x')\|.\end{aligned}$$

Il faut maintenant vérifier que la suite $(f_n)_{n\in\mathbb{N}}$ converge vers f dans l'espace $\mathcal{B}(X,E)$. Pour ce faire, passons à la limite lorsque p tend vers l'infini dans l'inégalité (2.4), ce qui donne

$$\forall\varepsilon>0\,,\ \exists n_0\in\mathbb{N}\,/\ \forall n\geq n_0\,,\ \forall x\in X\,,\ \|f_n(x)-f(x)\|\leq\varepsilon,$$

ce qui démontre que $\mathcal{B}(X,E)$ est complet.

Pour démontrer que $\mathcal{C}_b(X,E)$ est complet, il suffit, d'après la proposition 1.2.2 de démontrer que $\mathcal{C}_b(X,E)$ est fermé dans $\mathcal{B}(X,E)$, c'est-à-dire qu'une limite uniforme de fonctions continues est continue. Pour cela, considérons une suite $(f_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{C}(X,E)$ converge vers f dans $\mathcal{B}(X,E)$, alors f est continue de X dans E. Par définition, pour tout $\varepsilon>0$, il existe un entier n_0 tel que

$$\forall x\in X\,,\ \|f_{n_0}(x)-f(x)\|<\frac{\varepsilon}{4}\cdot \tag{2.5}$$

L'utilisation répétée de l'inégalité triangulaire et l'inégalité (2.5) permettent d'écrire, pour tout couple (x_0,x) de $X\times X$,

$$\begin{aligned}\|f(x)-f(x_0)\| &\leq \|f(x)-f_{n_0}(x)\|+\|f_{n_0}(x)-f_{n_0}(x_0)\|+\|f_{n_0}(x_0)-f(x_0)\|\\ &\leq \frac{\varepsilon}{2}+\|f_{n_0}(x)-f_{n_0}(x_0)\|.\end{aligned}$$

La fonction f_{n_0} étant continue, pour tout x_0 de X, il existe $\alpha>0$ tel que

$$\forall x\in X,\ d(x,x_0)<\alpha\Longrightarrow\|f_{n_0}(x)-f_{n_0}(x_0)\|<\frac{\varepsilon}{2}\cdot$$

On en déduit que

$$\forall x\in X,\ d(x,x_0)<\alpha\Longrightarrow\|f(x)-f(x_0)\|<\varepsilon,$$

ce qui achève la démonstration du théorème. □

Remarque Lorsque $X = \mathbb{N}$ et $E = \mathbb{R}$ ou $\mathbb{C}$, on désigne l'espace $\mathcal{B}(\mathbb{N}; E)$ par $\ell^\infty(\mathbb{N}; \mathbb{K})$. Il arrive que l'on omette de mentionner $\mathbb{K}$ lorsque le fait que les suites soient à valeurs réelles ou complexes est sans importance.

Exercice 2.1.1. *On définit $\mathcal{C}_u(X, E)$ comme étant l'ensemble des fonctions f de $\mathcal{B}(X, E)$ qui sont uniformément continues. Démontrer que $(\mathcal{C}_u(X, E), \|\cdot\|_{\mathcal{B}(X,E)})$ est un espace de Banach.*

Théorème 2.1.1. *Pour p dans l'intervalle $[1, \infty[$, on considère l'espace $\ell^p(\mathbb{N})$ des suites x à valeurs dans $\mathbb{K}$ telles que*

$$\sum_{n\in\mathbb{N}} |x(n)|^p < \infty.$$

On pose alors

$$\|x\|_{\ell^p} \stackrel{\text{déf}}{=} \Big(\sum_{n\in\mathbb{N}} |x(n)|^p\Big)^{\frac{1}{p}}.$$

L'espace $(\ell^p(\mathbb{N}), \|\cdot\|_{\ell^p})$ est un espace de Banach. De plus, si $p \leq q$, alors $\ell^p(\mathbb{N})$ est inclus dans $\ell^q(\mathbb{N})$ et l'inclusion

$$\left\{\begin{array}{ccc} \ell^p(\mathbb{N}) & \longrightarrow & \ell^q(\mathbb{N}) \\ x & \longmapsto & x \end{array}\right.$$

est une application linéaire continue de norme 1.

Démonstration. Démontrons tout d'abord que $\ell^p(\mathbb{N})$ est un espace vectoriel. Considérons x et y deux éléments de $\ell^p(\mathbb{N})$. Pour tout entier N, on a, d'après la proposition 2.1.2,

$$\begin{aligned} \sum_{j=0}^{N} |x(j)+y(j)|^p &\leq \bigg(\Big(\sum_{j=0}^{N} |x(j)|^p\Big)^{\frac{1}{p}} + \Big(\sum_{j=0}^{N} |y(j)|^p\Big)^{\frac{1}{p}}\bigg)^p \\ &\leq \big(\|x\|_{\ell^p} + \|y\|_{\ell^p}\big)^p. \end{aligned}$$

Il en résulte que la série de terme général $|x(j) + y(j)|^p$ est convergente et que

$$\sum_{j\in\mathbb{N}} |x(j)+y(j)|^p \leq \big(\|x\|_{\ell^p} + \|y\|_{\ell^p}\big)^p.$$

ce qui assure que $(\ell^p(\mathbb{N}), \|\cdot\|_{\ell^p})$ est un espace normé. Démontrons qu'il est de Banach. Soit $(x_q)_{q\in\mathbb{N}}$ une suite de Cauchy de $\ell^p(\mathbb{N})$. Par définition

$$\forall \varepsilon > 0\,, \exists q_\varepsilon \,/\, \forall q' \geq q_\varepsilon\,,\ \forall q' \geq q_\varepsilon\,,\ \|x_q - x_{q'}\|_{\ell^p(\mathbb{N})} < \varepsilon\,. \tag{2.6}$$

Comme $|x(n)| \leq \|x\|_{\ell^p}$ pour tout p, pour tout entier n la suite $(x_q(n))_{q\in\mathbb{N}}$ est une suite de Cauchy de $\mathbb{K}$ qui donc converge vers un élément $x(n)$ de $\mathbb{K}$. Montrons que $x = (x(n))_{n\in\mathbb{N}}$ appartient à $\ell^p(\mathbb{N})$ et que la suite $(x_p)_{p\in\mathbb{N}}$ converge vers x dans $\ell^p(\mathbb{N})$.

L'assertion (2.6) appliquée avec $\varepsilon = 1$ implique que

$$\forall q \geq q_1\,,\ \|x_q\|_{\ell^p} \leq 1 + \|x_{q_1}\|_{\ell^p}.$$

Par définition de la norme ℓ^p, ceci implique que l'on a

$$\forall q \geq q_1\,,\ \forall N \in \mathbb{N}\,,\ \sum_{n=0}^{N} |x_q(n)|^p \leq \big(1 + \|x_{q_1}\|_{\ell^p}\big)^p.$$

Par passage à la limite lorsque q tend vers l'infini, on trouve que

$$\forall N \in \mathbb{N}\,,\ \sum_{n=0}^{N} |x(n)|^p \leq \big(1 + \|x_{q_1}\|_{\ell^p}\big)^p$$

ce qui assure que x appartient à $\ell^p(\mathbb{N})$. Pour démontrer la convergence, remarquons que (2.6) implique que

$$\forall \varepsilon > 0\,, \exists q_\varepsilon \,/\, \forall q \geq q_\varepsilon\,,\ \forall q' \geq q_\varepsilon\,, \forall N \in \mathbb{N}\,,\ \sum_{n=0}^{N} |x_q(n) - x_{q'}(n)|^p < \varepsilon^p\,.$$

En passant à la limite lorsque q' tend vers l'infini, on en déduit que

$$\forall \varepsilon > 0\,, \exists q_\varepsilon \,/\, \forall q \geq q_\varepsilon\,,\ \forall q' \geq q_\varepsilon\,, \forall N \in \mathbb{N}\,,\ \sum_{n=0}^{N} |x_q(n) - x(n)|^p \leq \varepsilon^p\,.$$

Ceci implique en faisant tendre N vers l'infini que

$$\forall \varepsilon > 0\,, \exists q_\varepsilon \,/\, \forall q \geq q_\varepsilon\,,\ \forall q' \geq q_\varepsilon\,,\ \sum_{n\in\mathbb{N}} |x_q(n) - x(n)|^p \leq \varepsilon^p\,.$$

D'où la convergence.

Pour conclure, observons que, d'après (2.1), pour tout x de $\ell^{p_1}(\mathbb{N})$, tout $p_2 \geq p_1$ et tout entier positif N,

$$\sum_{n=0}^{N} \Big(\frac{|x(n)|}{\|x\|_{\ell^{p_1}}}\Big)^{p_2} \leq \sum_{n=0}^{N} \Big(\frac{|x(n)|}{\|x\|_{\ell^{p_1}}}\Big)^{p_1} \leq 1.$$

On en déduit que x appartient à $\ell^{p_2}(\mathbb{N})$ et par passage à la limite que

$$\|x\|_{\ell^{p_2}} \leq \|x\|_{\ell^{p_1}}\,.$$

En appliquant ceci à $x - y$, on montre que l'inclusion est 1-lipschitzienne. Le théorème est démontré. $\square$

Exercice 2.1.2. *Démontrer que les suites qui n'ont qu'un nombre fini de termes non nuls sont denses dans $\ell^p(\mathbb{N})$ pour p dans l'intervalle $[1, \infty[$.*

Exercice 2.1.3. *Démontrer que la fonction $x \longmapsto \|x\|_{\ell^p}^p$ est différentiable en tout point de x de $\ell^p(\mathbb{N})$ lorsque p appartient à $]1, \infty[$ et qu'elle est deux fois différentiable lorsque p appartient à $[2, \infty[$.*

2.2 Les espaces d'applications linéaires continues

Comme nous avons une structure d'espace métrique, nous pouvons définir la notion de fonction continue. Sur un espace vectoriel, la classe des applications linéaires joue évidemment un rôle privilégié. Leur continuité se vérifie de manière très simple.

Théorème 2.2.1. *Soient E et F deux espaces vectoriels normés et ℓ une application linéaire de E dans F. Les trois conditions sont équivalentes :*

— *l'application ℓ est lipschitzienne,*

— *l'application ℓ est continue,*

— *il existe un ouvert U de E tel que ℓ soit bornée sur U c'est-à-dire tel que*

$$\sup_{x\in U} \|\ell(x)\|_F < \infty.$$

Démonstration. Il est clair que la seule chose à démontrer est que la troisième condition implique la première. Considérons un ouvert U sur lequel l'application linéaire ℓ est bornée. Soit x_0 appartenant à U. L'ensemble $U - x_0$ est un ouvert (exercice : démontrez-le !) contenant l'origine. Par définition d'un ouvert, il existe une boule ouverte de centre 0 et de rayon α incluse dans $U - x_0$. On peut alors écrire

$$\begin{aligned} M \stackrel{\text{déf}}{=} \sup_{x\in B(0,\alpha)} \|\ell(x)\| &\leq \sup_{x\in U-x_0} \|\ell(x)\| \\ &\leq \sup_{y\in U} \|\ell(y - x_0)\| \\ &\leq \sup_{y\in U} \|\ell(y)\| + \|\ell(x_0)\|. \end{aligned}$$

Donc l'application ℓ est bornée sur une boule de centre 0 et de rayon α. Pour tout y de $E \setminus \{0\}$, on a

$$\frac{\alpha}{2\|y\|_E} y \in B(0,\alpha).$$

On en déduit que

$$\left\| \ell\left(\frac{\alpha}{2\|y\|_E} y\right)\right\| \leq M.$$

Par linéarité de ℓ et homogénéité de la norme, on trouve que

$$\forall y \in E\,,\ \|\ell(y)\|_F \leq \frac{2M}{\alpha}\|y\|_E.$$

Cette inégalité étant vraie pour $y = x - x'$, la démonstration du théorème est achevée. □

Proposition 2.2.1. *Soient E et F deux espaces vectoriels normés ; on désigne par $\mathcal{L}(E,F)$ l'ensemble des applications linéaires continues de E dans F. L'application définie par*

$$\|L\|_{\mathcal{L}(E,F)} \stackrel{\text{déf}}{=} \sup_{\|x\|_E\leq 1} \|L(x)\|_F$$

est une norme sur $\mathcal{L}(E,F)$. Si $(F, \|\cdot\|_F)$ est un espace de Banach, alors $(\mathcal{L}(E,F), \|\cdot\|_{\mathcal{L}(E,F)})$ l'est aussi.

Démonstration. Il est très facile de démontrer que $\mathcal{L}(E,F)$ est un espace vectoriel. Vérifions les trois propriétés définissant une norme. Supposons que $\|L\|_{\mathcal{L}(E,F)} = 0$. Ceci implique que $L(x)$ vaut 0 pour tout x de la boule unité (c'est ainsi que l'on nomme usuellement la boule de centre 0 et de rayon 1). Alors, pour tout $x \neq 0$,

$$L\left(\frac{x}{2\|x\|_E}\right) = \frac{1}{2\|x\|_E} L(x) = 0\,.$$

On en déduit que $L(x) = 0$ et donc que $L = 0$. Soient L_1 et L_2 deux éléments de $\mathcal{L}(E,F)$. On a

$$\|L_1(x) + L_2(x)\|_F \leq \|L_1(x)\|_F + \|L_2(x)\|_F\,.$$

Ainsi donc, pour tout x de E de norme inférieure à 1, on a

$$\|L_1(x) + L_2(x)\|_F \leq \|L_1\|_{\mathcal{L}(E,F)} + \|L_2\|_{\mathcal{L}(E,F)}.$$

La borne supérieure étant le plus petit des majorants, on en déduit que

$$\|L_1 + L_2\|_{\mathcal{L}(E,F)} \leq \|L_1\|_{\mathcal{L}(E,F)} + \|L_2\|_{\mathcal{L}(E,F)}.$$

Enfin, soient $L \in \mathcal{L}(E,F)$ et $\lambda \in \mathbb{K}$, on a

$$\|\lambda L(x)\|_F = |\lambda| \, \|L(x)\|_F.$$

D'où il résulte que

$$\sup_{\|x\|_E \leq 1} \|\lambda L(x)\|_F = |\lambda| \sup_{\|x\|_E \leq 1} \|L(x)\|_F.$$

Ainsi donc $\|\cdot\|_{\mathcal{L}(E;F)}$ est une norme sur $\mathcal{L}(E;F)$.

Supposons maintenant que $(F, \cdot\|_F)$ soit complet ; considérons une suite de Cauchy $(L_n)_{n\in\mathbb{N}}$ de $\mathcal{L}(E,F)$. D'après la définition de la norme sur l'espace $\mathcal{L}(E,F)$, on en déduit que, pour tout x de E, la suite $(L_n(x))_{n\in\mathbb{N}}$ est une suite de Cauchy de F. L'espace F étant supposé complet, il existe, pour tout x de E, un élément de F, noté $L(x)$ tel que

$$\lim_{n\to\infty} L_n(x) = L(x).$$

Il suffit maintenant de démontrer que L appartient à $\mathcal{L}(E,F)$ et que

$$\lim_{n\to\infty} L_n = L \quad \text{dans} \quad \mathcal{L}(E,F).$$

L'unicité de la limite assure très facilement que L est une application linéaire. Comme la suite $(L_n)_{n\in\mathbb{N}}$ est de Cauchy, elle est bornée. Il existe donc une constante C telle que

$$\sup_{\substack{n\in\mathbb{N} \\ \|x\|_E \leq 1}} \|L_n(x)\|_F \leq C.$$

Par passage à la limite, on en déduit que

$$\sup_{\|x\|_E \leq 1} \|L(x)\|_F \leq C.$$

Donc, l'application linéaire L est continue. Démontrons la convergence de la suite $(L_n)_{n\in\mathbb{N}}$ vers L dans $\mathcal{L}(E,F)$. La suite $(L_n)_{n\in\mathbb{N}}$ étant de Cauchy, pour tout ε strictement positif, il existe un entier n_0 tel que l'on ait, pour tout entier $n \geq n_0$,

$$\sup_{\substack{\|x\|_E \leq 1 \\ p\in\mathbb{N}}} \|L_n(x) - L_{n+p}(x)\|_F < \varepsilon.$$

Par passage à la limite lorsque p tend vers l'infini, on obtient

$$\sup_{\|x\|_E \leq 1} \|L_n(x) - L(x)\|_F \leq \varepsilon,$$

ce qui achève la démonstration de la proposition. □

Proposition 2.2.2. *Soient E, F et G trois espaces vectoriels normés et (L_1, L_2) un élément de $\mathcal{L}(E,F) \times \mathcal{L}(F,G)$. La composée $L_2 \circ L_1$ appartient à $\mathcal{L}(E,G)$ et*

$$\|L_2 \circ L_1\|_{\mathcal{L}(E,G)} \leq \|L_1\|_{\mathcal{L}(E,F)} \|L_2\|_{\mathcal{L}(F,G)}.$$

Démonstration. La borne supérieure étant un majorant, on a, pour tout x dans la boule unité de E,

$$\begin{aligned} \|L_2 \circ L_1(x)\|_G &\leq \|L_2\|_{\mathcal{L}(F,G)} \|L_1(x)\|_F \\ &\leq \|L_2\|_{\mathcal{L}(F,G)} \|L_1\|_{\mathcal{L}(E,F)}. \end{aligned}$$

Le fait que la borne supérieure soit le plus petit des majorants assure le résultat. □

Remarque Même lorsque $F = \mathbb{K}$ (c'est le cas des formes linéaires), le supremum sur la boule peut ne pas être atteint comme le montre l'exercice suivant qu'il est vivement conseillé de faire.

Exercice 2.2.1. *Soit $(a_n)_{n\in\mathbb{N}}$ une suite d'éléments de l'intervalle $]0,1[$. On suppose que la suite $(a_n)_{n\in\mathbb{N}}$ tend vers 0. On définit alors la forme linéaire ℓ_a sur $\ell^1(\mathbb{N})$ par*

$$\langle \ell_a, x\rangle \stackrel{\text{déf}}{=} \sum_{n\in\mathbb{N}} (1 - a_n) x(n).$$

Démontrer que $\|\ell_a\|_{(\ell^1(\mathbb{N}))'} = 1$ et que pour tout x de la boule unité de $\ell^1(\mathbb{N})$, $|\langle \ell_a, x\rangle|$ est strictement inférieur à 1.

Lorsque $E = F$, on désigne $\mathcal{L}(E,F)$ par $\mathcal{L}(E)$. Nous allons étudier cet ensemble et en particulier ces éléments inversibles. L'un des résultats de base sur les éléments inversibles de $\mathcal{L}(E)$ est le suivant.

Théorème 2.2.2. *Soit $(E, \|\cdot\|)$ un espace de Banach ; l'ensemble des éléments de $\mathcal{L}(E)$ qui sont à distance strictement inférieure à 1 de Id sont inversibles dans $\mathcal{L}(E)$. Dit autrement,*

$$\forall A \in B_{\mathcal{L}(E)}(\mathrm{Id}, 1)\,,\ \exists !\, A^{-1} \in \mathcal{L}(E) \;/\; A \circ A^{-1} = A^{-1} \circ A = \mathrm{Id}\,.$$

Démonstration. Elle repose sur la proposition suivante.

Proposition 2.2.3. *Soit E est un espace de Banach et $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de E telle que*

$$\sum_{n\in\mathbb{N}} \|x_n\|_E < \infty.$$

Alors la suite $S_N \stackrel{\text{déf}}{=} \sum_{n=0}^{N} x_n$ converge.

Démonstration. Nous allons démontrer que la suite $(S_N)_{N\in\mathbb{N}}$ est de Cauchy ce qui résulte du fait que

$$\|S_{N+P} - S_N\| \leq \sum_{n=N+1}^{N+P} \|x_n\|.$$

La proposition est démontrée. □

Poursuite de la démonstration du théorème 2.2.2. Un élément A de $B_{\mathcal{L}(E)}(\mathrm{Id},1)$ s'écrit $A = \mathrm{Id} - L$ avec $\|L\|_{\mathcal{L}(E)}$ strictement inférieur à 1. Posons

$$S_N = \sum_{n=0}^{N} L^n.$$

D'après la proposition 2.2.3, la suite $(S_N)_{N\in\mathbb{N}}$ converge vers un élément A^{-1} de $\mathcal{L}(E)$ dès que $\|L\|_{\mathcal{L}(E)}$ est strictement inférieur à 1. Par ailleurs, on a

$$S_N A = S_N A = \mathrm{Id} - L^{N+1}.$$

En passant à la limite dans l'égalité ci-dessus, on conclut la démonstration de ce théorème. □

Corollaire 2.2.1. *L'ensemble* $U(E)$ *des éléments inversibles de* $\mathcal{L}(E)$ *est un ouvert et l'application* Inv *définie par*

$$\begin{cases} U(E) & \longrightarrow & U(E) \\ L & \longmapsto & L^{-1} \end{cases}$$

est continue.

Démonstration. Soit L_0 un élément de $U(E)$, on peut alors écrire que

$$L = L_0 - (L_0 - L) = \big(\mathrm{Id} - (L_0 - L)L_0^{-1}\big)L_0.$$

Si l'on suppose que

$$\|L - L_0\|_{\mathcal{L}(E)} < \frac{1}{\|L_0^{-1}\|_{\mathcal{L}(E)}},$$

on déduit du théorème 2.2.2 que $\mathrm{Id} - (L_0 - L)L_0^{-1}$ est inversible et donc que L aussi. L'ensemble des éléments inversibles et donc bien un ouvert puisque si L_0 appartient à $U(E)$, alors la boule $\mathcal{B}_{L_0}$ de centre L_0 et de rayon $\|L_0^{-1}\|^{-1}_{\mathcal{L}(E)}$ est incluse dans $U(E)$. Le théorème 2.2.2 implique que

$$L^{-1} = L_0^{-1} \sum_{n=0}^{\infty} \big((L_0 - L)L_0^{-1}\big)^n.$$

On en déduit que, si L appartient à $B(L_0, \rho_0)$ avec ρ_0 strictement inférieur à $\|L_0^{-1}\|^{-1}_{\mathcal{L}(E)}$, on a

$$\begin{aligned} \big\|L^{-1} - L_0^{-1}\|_{\mathcal{L}(E)} &\leq \|L_0^{-1}\|_{\mathcal{L}(E)} \sum_{n=1}^{\infty} \|L - L_0\|^n_{\mathcal{L}(E)} \|L_0^{-1}\|^n_{\mathcal{L}(E)} \\ &\leq \|L - L_0\|_{\mathcal{L}(E)} \|L_0^{-1}\|^2_{\mathcal{L}(E)} \sum_{n=0}^{\infty} \|L - L_0\|^n_{\mathcal{L}(E)} \|L_0^{-1}\|^n_{\mathcal{L}(E)} \\ &\leq \|L - L_0\|_{\mathcal{L}(E)} \|L_0^{-1}\|^2_{\mathcal{L}(E)} \frac{1}{1 - \rho\|L_0^{-1}\|_{\mathcal{L}(E)}} \cdot \end{aligned}$$

Ceci conclut la démonstration du corollaire. □

Remarque En fait, cette démonstration montre que, lorsque ρ_0 est strictement inférieur à $\|L_0^{-1}\|^{-1}_{\mathcal{L}(E)}$, l'application $L \mapsto L^{-1}$ est k lipschitzienne sur $B_{\mathcal{L}(E)}(L_0, \rho_0)$ avec

$$k = \frac{\|L_0^{-1}\|^2_{\mathcal{L}(E)}}{1 - \rho_0\|L_0^{-1}\|_{\mathcal{L}(E)}} \cdot$$

Exercice 2.2.2. *Démontrer que l'application* Inv *est une application de classe* C^1 *sur* $U(E)$.

Pour conclure cette section, nous allons donner une caractérisation des applications multilinéaires continues. Le théorème suivant est l'analogue pour les applications multilinéaires du théorème 2.2.1 relatif à la continuité des applications linéaires.

Théorème 2.2.3. *Soient* $((E_j, \|\cdot\|_j))_{1\leq j\leq N}$ *une famille d'espaces vectoriels normés et* $(F, \|\cdot\|)$ *un espace vectoriel normé. On considère une application* N*-linéaire de* $E_1 \times \cdots \times E_N$ *dans* F. *Cette application* f *est continue de* $E = E_1 \times \cdots \times E_N$ *muni de la norme*

$$\|(x_1, \cdots x_N)\|_E = \sum_{1\leq j\leq N} \|x_j\|_{E_j}$$

dans F *si et seulement si*

$$\sup_{\|(x_1,\cdots,x_N)\|_E\leq 1} \|f(x_1, \cdots, x_N)\|_F < \infty. \tag{2.7}$$

De plus, si f *est continue de* $E_1\times\cdots\times E_N$ *dans* F, *alors elle est lipschitztienne sur les ensembles bornés de* $E_1 \times \cdots \times E_N$.

Démonstration. Si f est continue en 0 alors, il existe un réel α strictement positif tel que

$$\sum_{j=1}^{N} \|x_j\|_{E_j} < \alpha \Longrightarrow \|f(x_1, \cdots, x_N)\|_F < 1.$$

On utilise alors un argument d'homogénéïté. Soit $x = (x_1, \cdots, x_N)$ un élément non nul de E de norme dans E plus petite que 1. Alors $\dfrac{\alpha}{2}x$ est de norme dans E strictement plus petite que α. Ainsi donc

$$\|f(x_1, \cdots, x_N)\|_F < \Big(\frac{2}{\alpha}\Big)^N.$$

Réciproquement, supposons l'inégalité (2.7). Soit $h = (h_1, \cdots, h_N)$ on pose $h_0 \stackrel{\text{déf}}{=} 0$ et pour j dans $\{1, \cdots, N\}$, $\widetilde{h}_j \stackrel{\text{déf}}{=} (h_1, \cdots, h_j, 0 \cdots, 0)$. On a alors, pour un x fixé dans E,

$$f(x + h) - f(x) = \sum_{j=1}^{N} f(x + h_j) - f(x + h_{j-1}).$$

L'inégalité (2.7) implique que, pour tout entier j compris entre 1 et N, on a

$$\|f(x + h_j) - f(x + h_{j-1})\|_F \leq C\Big(\prod_{k=1}^{j-1} \|x_k + h_k\|_{E_k}\Big)\|h_j\|_{E_j}\Big(\prod_{k=j+1}^{N} \|x_k\|_{E_k}\Big).$$

On peut supposer que $\|h\|_E \leq 1$. Ainsi donc, on a, pour tout entier j compris entre 1 et N,

$$\|f(x + h_j) - f(x + h_{j-1})\|_F \leq C_x\|h_j\|_{E_j}.$$

Par sommation sur j, on en déduit que

$$\|f(x + h) - f(x)\|_F \leq C_x\|h\|_E$$

et le théorème est démontré. $\square$

2.3 Espaces de Banach, compacité et dimension finie

Le fait que la dimension de l'espace soit finie ou non induit de grandes différences sur la topologie comme on peut le voir au travers de l'énoncé suivant.

Théorème 2.3.1. *Soit E un espace vectoriel normé. Si la dimension de E est finie alors la boule unité fermée est compacte et toutes les normes sur E sont équivalentes. Récipoquement, si la boule unité fermée est compacte, alors l'espace vectoriel E est de dimension finie.*

Démonstration. Soit E un espace vectoriel normé de dimension finie N. Nous allons démontrer qu'il existe une bijection linéaire continue d'inverse continue de $\mathbb{R}^N$ muni de la norme $\|\cdot\|_\infty$ sur E. Considérons une base $(\overrightarrow{e}_j)_{1\leq j\leq N}$ de E et l'application linéaire bijective I définie par

$$I\left\{\begin{array}{ccc}\mathbb{R}^N & \longrightarrow & E\\ x=(x_j)_{1\leq j\leq N} & \longmapsto & \displaystyle\sum_{j=1}^N x_j\overrightarrow{e}_j.\end{array}\right.$$

L'application $I(x)$ est une bijection linéaire. Montrons qu'elle est continue. On a

$$\begin{aligned}\|I(x)\|_E &\leq \sum_{j=1}^N |x_j|\,\|\overrightarrow{e}_j\|_E\\ &\leq M\|x\|_\infty \quad \text{avec} \quad M \stackrel{\text{déf}}{=} \sum_{j=1}^N \|\overrightarrow{e}_j\|_E.\end{aligned}$$

L'application I étant linéaire, on en déduit que

$$\|I(x)-I(y)\|_E \leq M\|x-y\|_\infty.$$

L'application

$$x \longmapsto \|I(x)\|_E$$

est donc continue de $\mathbb{R}^N$ dans $\mathbb{R}^+$. D'après le corollaire 1.3.1 page 23, la sphère $\mathbb{S}^{N-1}$, (i.e. l'ensemble des x de $\mathbb{R}^N$ tels que $\|x\|_\infty=1$) est compacte. De plus, I étant bijective, elle ne s'annule pas sur $\mathbb{S}^{N-1}$. D'après le corollaire 1.3.2 page 24, il existe un réel strictement positif m tel que

$$\forall x\in\mathbb{R}^N\setminus\{0\}\,,\ m\leq \Big\|I\Big(\frac{x}{\|x\|_\infty}\Big)\Big\|_E\leq M.$$

L'application I étant linéaire, on a

$$\forall x\in\mathbb{R}^N\,,\ m\|x\|_\infty\leq\|I(x)\|_E\leq M\|x\|_\infty. \tag{2.8}$$

Donc I et I^{-1} sont des bijections linéaires continues. Donc toutes les normes sur E sont équivalentes. Donc les ouverts (resp. les fermés) (resp. les compacts) de E sont exactement les images des ouverts (resp. des fermés) (resp. des compacts) de $\mathbb{R}^N$ par I. Donc la boule unité fermée de E est un compact car c'est un fermé inclus dans l'image par I dans la boule fermé de centre 0 et de rayon m^{-1}. Donc la boule unité de E est compacte.

La réciproque est basée sur le lemme géométrique suivant.

Lemme 2.3.1. *Soit $(E, \|\cdot\|)$ un espace vectoriel normé. On supose qu'il existe un réel δ de l'intervalle $]0,1[$ et une famille finie $(x_j)_{1\leq j\leq N}$ telle que*

$$\mathbb{S}_E \stackrel{\text{déf}}{=} \{x \in E\,, \|x\| = 1\} \subset \bigcup_{j=1}^{N} B(x_,\delta).$$

Alors la famille $(x_j)_{1\leq j\leq N}$ est une partie génératrice de E.

Si l'on admet ce lemme un instant, il suffit d'utiliser la caractérisation de la compacité donnée par le théorème 1.3.4 page 24 pour conclure. □

Démonstration du lemme 2.3.1 Soit F le sous espace vectoriel de E engendré par la famille des $(x_j)_{1\leq j\leq N}$. C'est un espace vectoriel de dimension finie donc d'après la partie directe du théorème, l'espace normé $(F, \|\cdot\|)$ est complet. Donc c'est un fermé de E.

Nous allons considérer un vecteur y de $E \setminus \{0\}$ et démontrer qu'il est dans l'adhérence de F ce qui assurera le lemme. Le point clef consiste à établir le fait suivant :

$$\forall y \in E\,,\ \exists y_1 \in F \,/\, \|y - y_1\| \leq \delta\|y\|. \tag{2.9}$$

Si $y = 0$, on prend $y_1 = 0$. Supposons y non nul. Par hypothèse, il existe un j de $\{1, \cdots, N\}$ tel que

$$\left\|\frac{y}{\|y\|} - x_j\right\| \leq \delta$$

ce qui s'écrit $\|y - \|y\|x_j\| \leq \delta\|y\|$ et qui démontre donc l'assertion (2.9).

Itérons le processus. L'assertion (2.9) implique

$$\exists y_2 \in F \,/\, \|y - y_1 - y_2\| \leq \delta\|y - y_1\| \leq \delta^2\|y\|.$$

Par récurrence, on construit ainsi une suite $(y_n)_{n\in\mathbb{N}}$ d'éléments de F telle que

$$\forall n \in \mathbb{N}\,,\ \left\|y - \sum_{j=1}^{n} y_j\right\| \leq \delta^n\|y\|.$$

Il est clair que l'on a

$$\lim_{n\to\infty} \sum_{j=1}^{n} y_j = y$$

ce qui implique que y appartient à F car F est fermé et le lemme est ainsi démontré. □

Corollaire 2.3.1. *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés tels que E soit de dimension finie. Alors toute application linéaire L de E dans F est continue.*

Démonstration. Soit $(\overrightarrow{e}_j)_{1\leq j\leq N}$ une base de E, on définit la norme

$$\|\vec{x}\|_{\widetilde{E}} \stackrel{\text{déf}}{=} \sum_{j=1}^{N} |x_j|$$

où $(x_j)_{1\leq j\leq N}$ désigne les coordonnées de x dans la base $(\overrightarrow{e}_j)_{1\leq j\leq N}$. Cette norme est équivalente à la norme $\|\cdot\|_E$. De plus

$$\begin{aligned} \|L(\overrightarrow{x})\|_F &\leq \sum_{j=1}^{N} |x_j|\|L(\overrightarrow{e}_j)\|_F \\ &\leq \max_{1\leq j\leq N} \|L(\overrightarrow{e}_j)\|_F \sum_{j=1}^{N} |x_j|. \end{aligned}$$

Le corollaire est ainsi démontré. □

2.4 Compacité dans les espaces de fonctions continues : le théorème d'Ascoli

Dans un espace vectoriel de dimension finie **et seulement dans ce cas**, les parties compactes sont les parties fermées et bornées. Comme l'affirme le théorème 2.3.1 ci-dessus, ceci est toujours faux dans les espaces vectoriels normés de dimension infinie. À titre d'illustration, considérons, dans l'espace E des fonctions continues de $[0,1]$ dans $\mathbb{R}$ muni de la norme de la convergence uniforme, la suite $(f_n)_{n\in\mathbb{N}}$ définie par $f_n(x) \stackrel{\text{déf}}{=} x^n$. Tous les éléments de cette suite sont de norme 1 et cette suite n'a pas de valeur d'adhérence **dans l'espace** E car

$$\forall x \in [0,1[\,,\ \lim_{n\to\infty} x^n = 0 \quad \text{et} \quad \lim_{n\to\infty} 1^n = 1.$$

L'objet de cette section est l'établissement d'un critère de compacité pour les fonctions continues d'un espace métrique compact à valeurs dans un espace de Banach.

Théorème 2.4.1 (d'Ascoli)**.** *Soient (X,d) un espace métrique compact et $(E,\|\cdot\|_E)$ un espace de Banach. On considère une partie A de $\mathcal{C}(X,E)$ qui vérifie les deux hypothèses suivantes :*

i) La partie A est uniformément équicontinue i.e.

$$\forall \varepsilon > 0\,,\ \exists \alpha > 0 \,/\, \forall (x,x') \in X^2\,,\ d(x,x') < \alpha \Longrightarrow \forall f \in A\,,\ \|f(x)-f(x')\|_E < \varepsilon\,; \qquad (2.10)$$

ii) pour tout $x \in X$, l'ensemble $\{f(x)\,,\ f \in A\}$ est d'adhérence compacte.

Alors A est d'adhérence compacte.

Démonstration. Pour démontrer ce théorème, nous allons utiliser le critère donné par le théorème 1.3.1 page 20. Soit ε un réel strictement positif, démontrons que l'on peut recouvrir A par un nombre fini de boules ouvertes de rayon ε ce qui d'après le thérème 1.3.1 page 20 assurera le résultat.

D'après l'hypothèse (2.10), il existe un réel strictement positif α tel que

$$\forall (x,x') \in X^2\,,\ d(x,x') < \alpha \Longrightarrow \forall f \in A\,,\ \|f(x)-f(x')\|_E < \frac{\varepsilon}{3}\cdot \qquad (2.11)$$

La compacité de X assure l'existence d'une suite finie $(x_j)_{1\leq j\leq N}$ d'éléments de X tel que

$$X = \bigcup_{j=1}^{N} B(x_j,\alpha).$$

Posons $A_j \stackrel{\text{déf}}{=} \{f(x_j)\,,\ f \in A\}$. Par hypothèse, $\overline{A}_j$ l'adhérence de A_j est un compact de E. D'après la proposition 1.3.4 page 21, le produit $\overline{A}_1 \times \cdots \times \overline{A}_N$ est un compact de l'espace de Banach E^N muni de la norme

$$\|(y_j)_{1\leq j\leq N}\|_\infty \stackrel{\text{déf}}{=} \max_{1\leq j\leq N} \|y_j\|_E.$$

Donc le produit $A_1 \times \cdots \times A_N$ est d'adhérence compacte dans l'espace de Banach E^N. Comme le sous-ensemble $\mathcal{A}_N$ de E_N défini

$$\mathcal{A} \stackrel{\text{déf}}{=} \big\{(f(x_1),\cdots,f(x_N))\,,\ f \in A\big\}$$

est inclus dans le produit $A_1 \times \cdots \times A_N$ et est donc d'adhérence compacte lui aussi. D'après le théorème 1.3.1 page 20, il existe une suite finie $(f_k)_{1\leq k\leq M}$ d'éléments de A telle que les boules (au sens de la norme $\|\cdot\|_\infty$ définie ci-dessus) de centre $(f_k(x_j))_{1\leq j\leq N}$ et de rayon $\varepsilon/3$ recouvre $\mathcal{A}$ ce qui signifie que

$$\forall f \in A\,,\ \exists k \in \{1,\cdots,M\}\,/\ \forall j \in \{1,\cdots,N\}\,,\ \|f(x_j)-f_k(x_j)\|_E < \frac{\varepsilon}{3}\cdot$$

Nous allons démontrer que les boules (dans $(\mathcal{C}(X,E))$ de centre f_k et de rayon ε recouvre A. Pour ce faire, considérons un élément f de A. Il existe un indice k de $\{1,\cdots,M\}$ tel que

$$\forall j \in \{1,\cdots,N\}\,,\ \|f(x_j)-f_k(x_j)\|_E < \frac{\varepsilon}{3}\cdot$$

Soit x un point quelconque de X. Il existe un indice j de $\{1,\cdots,N\}$ tel que $d(x,x_j)<\alpha$. On peut alors écrire que

$$\begin{aligned}\|f(x)-f_k(x)\|_E &\leq \|f(x)-f(x_j)\|_E+\|f(x_j)-f_k(x_j)\|_E+\|f_k(x_j)-f_k(x)\|_E\\ &< \varepsilon\,.\end{aligned}$$

D'où le théorème. □

Exercice 2.4.1. *Démontrez la réciproque du théorème d'Ascoli, à savoir que si une partie A de $\mathcal{C}(X,E)$ est compacte, alors elle vérifie les conditions i) et ii) de l'énoncé du théorème d'Ascoli.*

Exercice 2.4.2. *Soient (X,d) un espace métrique compact et Y une partie compacte d'un espace de Banach $(E,\|\cdot\|_E)$. Démontrez que l'ensemble des fonctions k-lipschitziennes de X dans Y est compact dans $\mathcal{C}(X,E)$.*

Exercice 2.4.3. *Exhiber une suite de fonctions $(f_n)_{n\in\mathbb{N}}$ de $[0,1]$ dans $\mathbb{R}$ qui soit une suite bornée de l'espace des fonctions Lipschitziennes, telle que la suite $(f_n)_{n\in\mathbb{N}}$ tend vers 0 dans l'espace $\mathcal{C}([0,1],\mathbb{R})$ mais $(f_n)_{n\in\mathbb{N}}$ ne tende pas vers 0 dans l'espace des fonctions lipschiztiennes de $[0,1]$ dans $\mathbb{R}$.*

2.5 Autour du théorème de Stone-Weierstrass

Il s'agit de trouver des sous-espaces vectoriels denses de l'espace des fonctions continues sur un espace métrique compact (X,d). Commençons par le cas où l'espace de départ est l'intervalle $[0,1]$.

Théorème 2.5.1 (de Bernstein). *Soit f une fonction continue sur l'intervalle $[0,1]$ à valeurs dans un espace de Banach E. La suite de fonctions $(S_n(f))_{n\in\mathbb{N}}$ définie par*

$$S_n(f)(x) \stackrel{\text{def}}{=} \sum_{k=0}^{n} f\Big(\frac{k}{n}\Big) C_n^k\, x^k(1-x)^{n-k}.$$

(où les C_n^k représent les cœfficients binomiaux) converge uniformément vers f sur l'intervalle $[0,1]$.

Démonstration. Remarquons tout d'abord que, comme pour tout x de l'intervalle $[0,1]$ et tout entier positif n, on a $1=(x+1-x)^n$, la formule du binôme implique que

$$\sum_{k=0}^{n} C_n^k\, C_n^k\, x^k(1-x)^{n-k}=1 \tag{2.12}$$

Par définition de la suite $S_n(f)$, on en déduit que

$$f(x) - S_n(f)(x) = \sum_{k=0}^{n} \Big(f(x) - f\Big(\frac{k}{n}\Big)\Big) C_n^k \, x^k (1-x)^{n-k}.$$

D'après la théorème de Heine 1.3.3 page 23, la fonction f est uniformément continue sur le compact $[0,1]$. Donc, si l'on considère un réel strictement positif arbitraire ε, il existe un réel strictement positif α_ε tel que

$$|x-y| < \alpha_\varepsilon \Longrightarrow \|f(x) - f(y)\|_E < \frac{\varepsilon}{2} \cdot$$

On écrit alors

$$\begin{aligned} \|f(x) - S_n(f)(x)\|_E &\leq \sum_{k=0}^{n} \Big\| f(x) - f\Big(\frac{k}{n}\Big)\Big\|_E C_n^k x^k (1-x)^{n-k} \\ &\leq \frac{\varepsilon}{2} + 2 \sup_{x\in[0,1]} \|f(x)\|_E \sum_{|k-nx|\geq n\alpha_\varepsilon} C_n^k \, x^k (1-x)^{n-k}. \end{aligned}$$

En multipliant et en divisant par $(k-nx)^2$, on trouve que

$$\begin{aligned} \|f(x) - S_n(f)(x)\|_E &\leq \frac{\varepsilon}{2} + \sup_{x\in[0,1]} \|f(x)\|_E \sum_{|k-nx|\geq n\alpha} \frac{(k-nx)^2}{(k-nx)^2} C_n^k \, x^k (1-x)^{n-k} \\ &\leq \frac{\varepsilon}{2} + \frac{2}{n^2\alpha^2} \sup_{x\in[0,1]} \|f(x)\|_E \sum_{|k-nx|\geq n\alpha_\varepsilon} (k-nx)^2 C_n^k \, x^k (1-x)^{n-k} \\ &\leq \frac{\varepsilon}{2} + \frac{2}{n^2\alpha^2} \sup_{x\in[0,1]} \|f(x)\|_E \sum_{k=0}^{n} (k-nx)^2 C_n^k \, x^k (1-x)^{n-k}. \end{aligned} \tag{2.13}$$

On peut calculer la somme intervenant dans le terme de droite. En dérivant la relation (2.12) et en la multipliant par x, on trouve que

$$\sum_{k=0}^{n} k C_n^k \, x^k (1-x)^{n-k} - nx \sum_{k=0}^{n-1} C_{n-1}^k x^k (1-x)^{n-1-k} = 0$$

ce qui en utilisant à nouveau (2.12) assure que

$$\sum_{k=0}^{n} k C_n^k \, x^k (1-x)^{n-k} = nx. \tag{2.14}$$

Ensuite, en dérivant et en multipliant par x l'identité ci-dessus, on trouve que

$$\sum_{k=0}^{n} k^2 C_n^k \, x^k (1-x)^{n-k} - nx \sum_{k=0}^{n-1} k C_{n-1}^k x^k (1-x)^{n-1-k} = nx$$

En appliquant l'identité (2.14), on trouve que

$$\sum_{k=0}^{n} k^2 C_n^k \, x^k (1-x)^{n-k} = nx(1 + (n-1)x)$$

ce qui implique que

$$\begin{aligned}\sum_{k=0}^{n}(k-nx)^2C_n^k\,x^k(1-x)^{n-k} &= \sum_{k=0}^{n}k^2C_n^k\,x^k(1-x)^{n-k}-2nx\sum_{k=0}^{n}kC_n^k\,x^k(1-x)^{n-k}+n^2x^2\\ &= nx(1+(n-1)x)-2n^2x^2+n^2x^2\\ &= nx(1-x).\end{aligned}$$

L'inégalité (2.13) devient alors

$$\|f(x)-S_n(f)(x)\|_E \leq \frac{\varepsilon}{2}+\frac{2}{n\alpha^2}\sup_{x\in[0,1]}\|f(x)\|_E.$$

Il suffit maintenant de choisir n tel que $n \geq n_\varepsilon \overset{\text{déf}}{=} \left[\sup_{x\in[0,1]}\|f(x)\|_E\dfrac{4}{\varepsilon\alpha_\varepsilon^2}\right]+1$ pour conclure la démonstration. □

Nous allons maintenant énoncer un critère de densité pour les sous-algèbres de $\mathcal{C}(X,\mathbb{K})$ où (X,d) est un espace métrique compact. Donnons tout d'abord une définition.

Définition 2.5.1. *Soient (X,d) un espace métrique et A une partie de $\mathcal{C}(X,\mathbb{K})$. On dit que A sépare les points de X si et seulement si pour tout couple (x,y) d'éléments de X tel que $x\neq y$, il existe une fonction f de A telle que $f(x)\neq f(y)$.*

Nous pouvons maintenant énoncer le théorème de densité suivant.

Théorème 2.5.2 (de Stone-Weierstrass)**.** *Soient (X,d) une espace métrique compact et A une sous algèbre de l'espace $C(X,\mathbb{R})$. Si A contient les fonctions constantes et si A sépare les points de X, alors A est dense dans $C(X,\mathbb{R})$.*

Le cas des fonctions à valeurs complexes est légèrement différent.

Théorème 2.5.3. *Soient (X,d) une espace métrique compact et A une sous-algèbre de l'espace $\mathcal{C}(X,\mathbb{C})$. Si A contient les fonctions constantes, si A sépare les points de X et si A est stable par conjugaison, (i.e. $f\in A\Rightarrow\overline{f}\in A$), alors A est dense dans $\mathcal{C}(X,\mathbb{C})$.*

On peut en l'énoncé suivant.

Corollaire 2.5.1. *Si X est un espace compact, l'espace de Banach $\mathcal{C}(X,\mathbb{K})$ est séparable, c'est-à-dire qu'il existe une suite de points de $\mathcal{C}(X,\mathbb{K})$ dense dans $\mathcal{C}(X,\mathbb{K})$.*

La démonstration du théorème général de Stone-Weierstrass est présenté ici à titre culturel.

Démonstration du théorème 2.5.2

L'une des étapes essentielles de la démonstration est le lemme suivant, que nous allons admettre pour le moment.

Lemme 2.5.1. *Soit A une sous-algèbre de $\mathcal{C}(X;\mathbb{R})$. Pour toute fonction f de $\overline{A}$, la fonction $|f|$ appartient à $\overline{A}$.*

Remarque Comme $\max\{f,g\}=\dfrac{1}{2}(f+g+|f-g|)$ et $\min\{f,g\}=\dfrac{1}{2}(f+g-|f-g|)$, si deux fonctions f et g appartiennent à A, alors leur minimum et leur maximum appartiennent à $\overline{A}$.

Retour à la démonstration du théorème 2.5.2 . Le fait que A est séparante implique que

$$(S^+) \qquad \forall (x_1,x_2)\in X^2\,/\,x_1\neq x_2\,,\ \forall(\alpha_1,\alpha_2)\in\mathbb{R}^2\,,\ \exists h\in A\,/\,h(x_j)=\alpha_j.$$

Comme A est séparante, il existe une fonction g de A telle que $g(x_1)\neq g(x_2)$. On pose alors

$$h(z)\stackrel{\text{déf}}{=}\alpha_1+(\alpha_2-\alpha_1)\frac{g(z)-g(x_1)}{g(x_2)-g(x_1)}\,\cdot$$

Considérons une fonction f de $\mathcal{C}(X;\mathbb{R})$. Soit x_0 un point de X et ε un réel strictement positif. Nous allons démontrer

$$\exists h_{x_0}\in\overline{A}\,/\,h_{x_0}(x_0)=f(x_0)\quad\text{et}\quad\forall z\in X\,,\ h_{x_0}(z)\leq f(z)+\varepsilon. \tag{2.15}$$

En effet, d'après l'assertion (S^+), pour tout y de X, il existe une fonction $h_{x_0,y}$ de A telle que h coïncide avec f aux points x_0 et y. Les deux fonctions f et $h_{x_0,y}$ sont continues. Il existe donc un réel strictement positif $\alpha_{x_0,y}$ tel que

$$d(z,y)<\alpha_{x_0,y}\Longrightarrow h_{x_0,y}(z)<f(y)+\varepsilon.$$

L'espace (X,d) étant supposé compact, on peut le recouvrir avec un nombre fini de boules du type ci-dessus, c'est-à-dire qu'il existe une suite finie $(y_j)_{1\leq j\leq N}$ telle que

$$X=\bigcup_{j=1}^N B(y_j,\alpha_{x_0,y_j}).$$

Posons $h_{x_0}\stackrel{\text{déf}}{=}\min\limits_{1\leq j\leq N}h_{x_0,y_j}$. D'après le lemme 2.5.1, la fonction h_{x_0} appartient à $\overline{A}$ et est telle que, pour tout z dans X, on ait $h_{x_0}(z)<f(z)+\varepsilon$. L'assertion (2.15) est donc démontrée.

Pour tout $x\in X$, il existe donc une fonction $h_x\in\overline{A}$ telle que

$$h_x(x)=f(x)\quad\text{et}\quad\forall y\in X\,,\ h_x(y)<f(y)+\varepsilon.$$

Les fonctions h_x et f étant continues, pour tout $y\in X$, il existe un réel strictement positif α_x tel que

$$\forall z\in B(x,\alpha_x)\,,\ h_x(z)>f(z)-\varepsilon.$$

À nouveau, on recouvre X par un nombre fini de telles boules, c'est-à-dire qu'il existe une famille finie $(x_j)_{1\leq j\leq N}$ de points de X telle que

$$X=\bigcup_{j=1}^N B(x_j,\alpha_{x_j}).$$

Posons $h\stackrel{\text{déf}}{=}\max\limits_{1\leq j\leq N}h_{x_j}$. D'après le lemme 2.5.1, la fonction h appartient à $\overline{A}$. De plus ; elle vérifie

$$\forall y\in X\,,\ f(y)-\varepsilon<h(y)<f(y)+\varepsilon,$$

c'est-à-dire $\|f-h\|_{\mathcal{C}(X,\mathbb{R})}<\varepsilon$. Le théorème est donc démontré pourvu bien sûr que nous démontrions le lemme 2.5.1. □

Démonstration du lemme 2.5.1 Remarquons tout d'abord que l'on peut supposer A fermé car A est une sous-algèbre de $\mathcal{C}(X,\mathbb{K})$, alors $\overline{A}$ aussi. (Exercice : démontrez-le !) La démonstration repose alors sur le lemme suivant, que nous admettons provisoirement.

Lemme 2.5.2. *Il existe une suite de polynômes $(P_n)_{n\in\mathbb{N}}$ telle que*

$$\lim_{n\to\infty}P_n=\sqrt{\cdot}\quad\textit{dans l'espace } C([0,1];\mathbb{R}).$$

Soit f un élément de A et posons

$$f_n \stackrel{\text{déf}}{=} P_n\Big(\frac{f^2}{\|f\|}\Big) \quad \text{avec bien sûr} \quad \|f\| = \sup_{x\in X} |f(x)|.$$

La suite $(f_n)_{n\in\mathbb{N}}$ est une suite d'éléments de A puisque A est une algèbre. D'après lemme 2.5.2, la suite de fonctions $(f_n)_{n\in\mathbb{N}}$ converge vers $\|f\|^{-1}\sqrt{f^2} = \|f\|^{-1}|f|$. Donc $|f| \in \overline{A}$. Le théorème 2.5.2 est démontré pourvu que l'on prouve le lemme 2.5.2. $\square$

Démonstration du lemme 2.5.2 Définissons par récurrence la suite de polynômes

$$P_{n+1}(t) = P_n(t) + \frac{1}{2}(t - P_n^2(t)) \quad \text{et} \quad P_0 = 0.$$

Montrons par récurrence que, pour tout $t \in [0,1]$, $0 \leq P_n(t) \leq \sqrt{t}$. C'est vrai pour $n = 0$. Supposons-le pour P_n. Le polynôme P_{n+1} est positif sur $[0,1]$ comme somme de deux fonctions positives. De plus, on a, comme P_n est positif sur $[0,1]$,

$$\begin{aligned} t - P_{n+1}^2(t) &= t - \Big(P_n(t) + \frac{1}{2}(t - P_n(t))^2\Big)^2 \\ &\geq t - P_n^2(t) \\ &\geq 0 \end{aligned}$$

Comme $P_{n+1}(t) - P_n(t) = 1/2(t - P_n^2(t))$, pour tout $t \in [0,1]$, on a $P_{n+1}(t) \geq P_n(t)$. Pour tout $t \in [0,1]$, la suite $(P_n(t))_{n\in\mathbb{N}}$ est croissante majorée par $\sqrt{t}$, donc elle converge. Soit $\ell(t)$ la limite. Par passage à la limite dans la relation de récurrence définissant la suite $(P_n)_{n\in\mathbb{N}}$, on trouve que $t - \ell^2(t) = 0$. Le fait que la suite $(P_n)_{n\in\mathbb{N}}$ converge uniformément est assuré par le théorème suivant.

Théorème 2.5.4 (de Dini). *Soit $(f_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{C}(X,\mathbb{R})$ telle que, pour tout x de X, la suite $(f_n(x))_{n\in\mathbb{N}}$ soit une suite croissante majorée. Si $g(x) \stackrel{\text{déf}}{=} \lim_{n\to\infty} f_n(x)$ est une fonction continue sur X, alors la suite $(f_n)_{n\in\mathbb{N}}$ converge vers g uniformément, c'est-à-dire*

$$\lim_{n\to\infty} \sup_{x\in X} |f_n(x) - g(x)| = 0.$$

Démonstration. Pour tout $x \in X$, la suite $(f_n(x))_{n\in\mathbb{N}}$ est croissante majorée ; pour tout $\varepsilon > 0$, il existe donc un entier n_x tel que

$$g(x) - \frac{\varepsilon}{2} < f_{n_x}(x).$$

Les deux fonctions g et f_{n_x} sont des fonctions continues sur le compact (X,d) ; elles sont donc uniformément continues sur (X,d). Ainsi, il existe un réel strictement positif α_x tel que

$$d(y,y') < \alpha_x \Longrightarrow |g(y) - g(y')| + |f_{n_x}(y) - f_{n_x}(y')| < \frac{\varepsilon}{2}\cdot$$

La famille $(B(x,\alpha_x))_{x\in X}$ recouvre le compact (X,d). On peut donc en extraire un sous-recouvrement fini, c'est-à-dire qu'il existe une suite $(x_j)_{1\leq j\leq N}$ telle que

$$X = \bigcup_{j=1}^{N} B(x_j, \alpha_{x_j}).$$

Posons $n_0 \stackrel{\text{déf}}{=} \max_{1\leq j\leq N} n_{x_j}$. La suite $(f_n(y))_{n\in\mathbb{N}}$ étant une suite croissante, on a alors, pour tout j dans $\{1,\cdots,N\}$,

$$\begin{aligned} f_{n_0}(y) &\geq f_{n_{x_j}}(y) \\ &\geq f_{n_{x_j}}(x_j) - |f_{n_{x_j}}(y) - f_{n_{x_j}}(x_j)| \\ &\geq g(x_j) - \frac{\varepsilon}{2} - |f_{n_{x_j}}(y) - f_{n_{x_j}}(x_j)| \\ &\geq g(y) - \frac{\varepsilon}{2} - |g(y) - g(x_j)| - |f_{n_{x_j}}(y) - f_{n_{x_j}}(x_j)|. \end{aligned}$$

Soit j tel que $y \in B(x_j, \alpha_{x_j})$. Il vient
$$f_{n_0}(y) \geq g(y) - \varepsilon.$$
La suite $(f_n(y))_{n\in\mathbb{N}}$ étant croissante, on a,
$$n \geq n_0 \Longrightarrow \forall y \in X\,,\ g(y) \geq f_n(y) > g(y) - \varepsilon.$$
Le théorème est alors démontré. □

2.6 Notions d'espaces séparables

Cette notion repose sur la notion d'ensemble dénombrable.

Définition 2.6.1. *Soit X un ensemble, on dit que X est dénombrable si et seulement si il existe une application injective de X dans $\mathbb{N}$.*

Remarquons que tout ensemble fini est dénombrable. De plus, on a la proposition suivante.

Proposition 2.6.1. *Soit X un ensemble infini dénombrable. Alors il existe une bijection de X sur $\mathbb{N}$.*

Démonstration. Par hypothèse, il existe une fonction ϕ injective de X dans $\mathbb{N}$. C'est une bijection de X sur $Y = \varphi(x)$. Il suffit donc de démontrer qu'une partie infinie Y de $\mathbb{N}$ est en bijection avec $\mathbb{N}$. On définit par récurrence
$$b(n) \stackrel{\text{déf}}{=} \min Y \setminus \{b(0), \cdots, b(n-1)\}$$
L'application b est strictement croissante donc injective. Elle est bijective car l'on prend les éléments de Y par ordre croissant. □

Proposition 2.6.2. *Tout produit fini d'ensembles dénombrables est dénombrable.*

Démonstration. Il suffit de la démontrer que le produit de deux ce qui d'après la proposition précédente, se ramène à construire une injection de $\mathbb{N} \times \mathbb{N}$ dans $\mathbb{N}$. On définit
$$\Phi \left\{ \begin{array}{ccc} \mathbb{N} \times \mathbb{N} & \Longrightarrow & \mathbb{N} \\ (p,q) & \longmapsto & p + \dfrac{(p+q+1)(p+q+2)}{2} \end{array} \right. .$$
Démontrons que Φ est injective. Pour ce faire, commençons par observer que si (p,q) et (p',q') sont deux couples de $\mathbb{N}^2$ tels que $p'+q'$ est strictement supérieur à $p+q$, alors $\Phi(p',q')$ est strictement supérieur à $\Phi(p,q)$. En effet, nous avons que $p'+q'$ est supérieur ou égal à $p+q+1$ et donc que
$$\begin{aligned} (p'+q'+1)(p'+q'+2) &\geq (p+q+3)(p+q+2) \\ &\geq 2(p+q+2) + (p+q+1)(p+q+2). \end{aligned}$$
On en déduit que
$$\Phi(p',q') \geq p' + q + 2 + \Phi(p,q).$$
Ainsi donc si $\Phi(p,q) = \Phi(p',q')$, alors $p+q = p'+q'$ ce qui implique immédiatement que $p = p'$ et donc que $q = q'$ ce qui démontre que l'application Φ est injective. □

Corollaire 2.6.1. *Les ensembles $\mathbb{Z}$ et $\mathbb{Q}$ sont dénombrables.*

Corollaire 2.6.2. *Une réunion dénombrable disjointe d'ensembles dénombrables est dénombrable. Plus précisement, soit A_n une suite d'ensembles dénombrables, on pose*

$$\mathcal{A} = \{(n,a) \quad \text{avec} \quad a \in A_n\}.$$

L'ensemble $\mathcal{A}$ est dénombrable.

Démonstration. Chaque ensemble A_n est dénombrable. Il existe donc pour tout n, une application injective ϕ_n de A_n dans $\mathbb{N}$. On définit alors

$$\Phi \left\{ \begin{array}{ccc} \mathcal{A} & \longrightarrow & \mathbb{N}^2 \\ a = (n, a_n) & \longmapsto & (n, \phi_n(a_n)). \end{array} \right.$$

Cette application est injective. En effet, si $\Phi(a) = \Phi(b)$, ce qui signifie que $(n, a_n) = (m, b_m)$ et donc que $n = m$ et $a_n = b_m$. Soit maintenant l'application ϕ de la proposition 2.6.1 ; l'application $\phi \circ \Phi$ est une application injective de $\mathcal{A}$ dans $\mathbb{N}$ comme composée de deux fonctions injectives. □

Nous avons le résultat négatif suivant.

Théorème 2.6.1. *Soit (X, d) un espace métrique complet infini tel que tout point soit d'intérieur vide. L'ensemble X n'est pas dénombrable.*

Démonstration. Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de X (tout ensemble dénombrable peut être représenté ainsi). Le théorème de Baire tel que traduit par le théorème 1.2.3 dit que l'ensemble

$$\bigcup_{n\in\mathbb{N}} \{x_n\}$$

est d'intérieur vide et est donc distinct de X. □

On peut maintenant définir la notion d'espace métrique séparable.

Définition 2.6.2. *Soit (X, d) un espace métrique. On dit que (X, d) est séparable si et seulement si il existe une suite dense dans X.*

Dit autrement, cette définition signifie que, dans un espace métrique séparable (X, d), il existe une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X telle que

$$\forall x \in X \, , \; \forall \varepsilon > 0 \, , \; \exists m \, / \; d(x, x_m) < \varepsilon.$$

Exemple. L'espace $\mathbb{R}$ muni de la distance $d(x, y) = |x - y|$ est séparable, car la suite des rationnels est dense dans $\mathbb{R}$.

Proposition 2.6.3. *Soit $(X_1, d_1), \cdots, (X_N, d_N)$ une famille de N espaces métriques séparables. Si l'on pose*

$$X = X_1 \times \cdots \times X_N \quad \text{et} \quad d((x_1, \cdots, x_N), (y_1, \cdots, y_N)) = \max_{1\leq j\leq N} d_j(x_j, y_j),$$

alors l'espace (X, d) est un espace séparable.

Démonstration. Pour chaque espace X_j, il existe une partie dénombrable D_j dense dans X_j. Considérons un point $x = (x_1, \cdots, x_N)$ quelconque de X et posons $D = D_1 \times \cdots \times D_N$. Cette partie D de X est dénombrable en tant que produit fini de parties dénombrables. De plus, comme D_j est dense dans X_j, pour tout ε strictement positif et tout j compris entre 1 et N, il existe un y_j appartenant à D_j tel que l'on ait

$$d_j(x_j, y_j) < \varepsilon.$$

Par définition de la distance d, on a $d(x, y) < \varepsilon$. La proposition est ainsi démontrée. □

Proposition 2.6.4. *Soient (X, d) un espace métrique séparable et A une partie quelconque de X. L'espace métrique (A, d) est séparable.*

Démonstration. Pour démontrer cette proposition, désignons par $\mathbb{D}$ une partie dénombrable dense de X. Pour tout entier n et tout x de $\mathbb{D}$, on choisit un élément $a_{x,n}$ de $B(x, n^{-1}) \cap A$ si cet ensemble est non vide. S'il est vide, on prend un quelconque élément de A. On a ainsi défini une partie dénombrable $\widetilde{\mathbb{D}} \stackrel{\text{déf}}{=} (a_{x,n})_{(x,n) \in \mathbb{D} \times \mathbb{N}}$ de A.

Démontrons que $\widetilde{\mathbb{D}}$ est dense dans A. Soit a un élément de A. Il existe un élément x de $\mathbb{D}$ tel que $d(x, a) < n^{-1}$. Par définition de $\widetilde{\mathbb{D}}$, l'élément $a_{x,n}$ appartient à $\widetilde{\mathbb{D}} \cap B(x, n^{-1})$. Donc, il en résulte que

$$d(a, a_{x,n}) \leq d(a, x) + d(x, a_{x,n,}) < \frac{2}{n},$$

d'où la proposition. □

Théorème 2.6.2. *Tout espace métrique (X, d) qui est compact est séparable.*

Démonstration. Soit $\mathcal{B}_n$ l'ensemble des boules de rayon n^{-1}. Cet ensemble constitue bien évidemment un recouvrement ouvert de X. L'hypothèse de compacité fait que, pour chaque entier n, par $(x_{j,n})_{1 \leq j \leq J(n)}$ une suite telle que

$$X = \bigcup_{j=1}^{J(n)} B(x_{j,n}, n^{-1}).$$

Soit D l'ensemble défini par

$$\{x_{j,n}\,,\ 1 \leq j \leq J(n)\,,\ n \in \mathbb{N}\}.$$

Le corollaire 2.6.2 implique que D est dénombrable. Démontrons qu'il est dense. Pour cela, considérons un réel strictement positif ε arbitraire et un point x quelconque de X. On choisit alors un entier n tel que n soit strictement supérieur à l'inverse de ε. Par définition des $(x_{j,n})$, il existe un entier j tel que

$$d(x, x_{j,n}) < \frac{1}{n}\cdot$$

Vu le choix de l'entier n, le théorème est démontré. □

Nous allons maintenant donner un exemple d'espace métrique non séparable.

Proposition 2.6.5. *Soit X l'ensemble des fonctions de $[0, 1]$ dans $[0, 1]$ muni de la distance*

$$d(f, g) = \sup_{x \in [0,1]} |f(x) - g(x)|.$$

L'espace métrique (X, d) ainsi défini n'est pas séparable.

Démonstration. Considérons l'ensemble A défini par

$$A = \{\mathbf{1}_P, P \in \mathcal{P}([0, 1])\},$$

où $\mathcal{P}$ désigne l'ensemble des parties de X. Remarquons deux choses. Tout d'abord, remarquons que, si $P \neq Q$, alors $d(\mathbf{1}_P, \mathbf{1}_Q) = 1$. Ensuite, l'ensemble $\mathcal{P}([0, 1])$, donc aussi l'ensemble A, n'est pas dénombrable. La démonstration de la proposition se conclut grâce au lemme suivant.

Lemme 2.6.1. *Soit (X, d) un espace métrique. S'il existe une partie non dénombrable A et un réel α strictement positif tels que*

$$\forall (x_1, x_2) \in A^2\,,\; x_1 \neq x_2 \Rightarrow d(x_1, x_2) \geq \alpha,$$

alors l'espace (X, d) n'est pas séparable.

Démonstration. Soit D une partie dense dans X. Pour tout a dans A, il existe un z dans D tel que

$$d(a, z) \leq \frac{\alpha}{3}.$$

Soient a_1 et a_2 deux éléments de A distincts. On leur associe deux éléments z_1 et z_2 de D vérifiant l'inégalité ci-dessus. Mais alors, on a

$$\begin{array}{rcl}\alpha & \leq & d(a_1, a_2) \\ & \leq & d(a_1, z_1) + d(a_2, z_2) + d(z_1, z_2) \\ & \leq & \dfrac{\alpha}{3} + \dfrac{\alpha}{3} + d(z_1, z_2) \\ & \leq & \dfrac{2\alpha}{3} + d(z_1, z_2).\end{array}$$

Donc z_1 est différent de z_2. Donc si une partie est dense, elle ne peut être dénombrable. Le lemme est ainsi démontré. □

Chapitre 3

Dualité dans les espaces de Banach

Introduction

Ce court chapitre est très important. Il étudie un type particulier d'espace de Banach : le dual (appelé parfois topologique) d'un espace de Banach qui est l'ensemble des formes linéaires continues sur cet espace. Tout d'abord, y est présenté le concept général de transposée d'une application linéaire. Insistons sur le fait qu'en dimension infinie, nous sommes très loin de la transposition des matrices. Le fait radicalement nouveau que nous allons découvrir dans ce chapitre est que le dual d'un espace de Banach peut-être très différent de l'espace lui-même. Ceci n'est bien sûr pas le cas en dimension finie où le dual d'un espace vectoriel de dimension finie est un espace vectoriel de même dimension.

Dans le deuxième section, nous donnerons un sens mathématique précis à l'expression "le dual d'un espace E est l'espace F". Cette définition est fondamentale. Elle permet de construire des isomorphismes entre le dual E' d'un espace de Banach et un espace de Banach F ce qui fournit une description de E'. L'exemple du dual des espaces de suite de puissance $p$ième intégrable est étudié en détail.

La troisième section introduit un concept crucial qui est une notion affaiblie de la convergence d'une suite d'éléments du dual d'un espace de Banach. Il s'agit essentiellement du concept de convergence dite faible $\star$ qui est essentiellement la convergence simple. Cette notion permet d'extraire de toute suite bornée du dual une sous suite qui converge en ce sens. D'où l'importance des théorèmes d'identification de la section précédente. Si un espace peut s'identifier à un dual, alors il jouira de cette propriété. Nous verrons dès le chapitre suivant une application de ce résultat lors de la démonstration du théorème 4.4.2 page 75 relatif à la structure des opérateurs autoadjoints compacts.

3.1 Présentation du concept de dualité

La notion de dualité est certainement bien connue du lecteur dans le cadre des espaces vectoriels. Néammoins, rappellons brièvement quelques notions fondamentales dans le cadre d'un espace vectoriel E quelconque. Désignons par $E^\star$ l'espace des formes linéaires définies sur E. De plus, lorsque ℓ est une forme linéaire sur E et x un vecteur de E , on note

$$\langle \ell, x\rangle \stackrel{\text{déf}}{=} \ell(x).$$

La première notion clef est celle de transposée d'une application linéaire. Considérons une application linéaire L de E dans E. On définit l'application linéaire transposée de E' (l'ensemble des formes linéaires sur E) dans E' par

$$\langle {}^tL(\ell), \overrightarrow{h} \rangle = \langle \ell, L\overrightarrow{h} \rangle.$$

Remarquons que si L_1 et L_2 sont deux applications linéaires de E dans E, alors on a

$${}^t(L_1 \circ L_2) = {}^tL_2 \circ {}^tL_1.$$

En effet, par définition de la transposition, on a

$$\begin{aligned}
\langle {}^t(L_1 \circ L_2)(\ell), \overrightarrow{h} \rangle &= \langle \ell, L_1 \circ L_2(\overrightarrow{h}) \rangle \\
&= \langle \ell, L_1(L_2(\overrightarrow{h})) \rangle \\
&= \langle {}^tL_1(\ell), L_2(\overrightarrow{h}) \rangle \\
&= \langle {}^tL_2({}^tL_1(\ell)), \overrightarrow{h} \rangle.
\end{aligned}$$

D'où le résultat.

Dans le cas où l'espace vectoriel E est de dimension finie, on peut alors choisir une base $(\overrightarrow{e}_j)_{1\leq j\leq N}$ de E. On démontre facilement que la famille $(e_j^\star)_{1\leq j\leq N}$ de formes linéaires définies par

$$\langle e_j^\star, \overrightarrow{e}_k \rangle = \delta_{j,k}$$

forme une base de l'espace vectoriel E' des formes linéaires sur E appelée base duale. De plus, il est aisé de démontrer que si la matrice de L dans une base $(\overrightarrow{e}_j)_{1\leq j\leq N}$ de E est $(L_j^i)_{1\leq i,j\leq N}$ alors la matrice de ${}^tL$ dans la base $(e_j^\star)_{1\leq j\leq N}$ est $(\widetilde{L}_j^i)_{1\leq i,j\leq N}$ avec $\widetilde{L}_j^i = L_i^j$. Ainsi donc dans le cas de la dimension finie, l'ensemble des formes linéaires continues (elles le sont toujours) est un espace vectoriel de même dimension.

Nous étudierons dans ce chapitre le cas où E est un espace de Banach général.

Définition 3.1.1. *On appelle dual topologique (ou simplement dual) d'un $\mathbb{K}$-espace vectoriel normé $(E, \|\cdot\|)$ et l'on note E' l'espace vectoriel des formes linéaires continues sur E. On le munit de la norme*

$$\|\ell\|_{E'} \stackrel{\text{déf}}{=} \sup_{\|x\|\leq 1} |\langle \ell, x \rangle|.$$

Par rapport au cas de la dimension finie, un phénomène fondamentalement nouveau apparait. Comme nous le verrons après la démonstration du théorème 3.2.1, l'espace de Banach $\ell^1(\mathbb{N})$ des séries convergentes est un espace séparable et son dual ne l'est pas. Ce qui peut apparaître à ce stade comme une difficulté va se révéler extrêmement fécond au travers de l'opération de transposition. Ces idées sont à la base de la notion de quasi dérivées que nous présenterons au chapitre 6 et plus généralement à la théorie des distributions que nous introduirons au chapitre 8

Il existe une forme bilinéaire naturelle définie sur l'espace $E' \times E$.

Proposition 3.1.1. *Soit E un espace vectoriel normé. Alors l'application $\langle \cdot, \cdot \rangle$ définie par*

$$\langle \cdot, \cdot \rangle \left\{ \begin{array}{ccc} E' \times E & \longrightarrow & \mathbb{K} \\ (\ell, x) & \longmapsto & \langle \ell, x \rangle. \end{array} \right.$$

est une forme bilinéaire continue.

Démonstration. Par définition de la norme sur E', on a $|\langle \ell, x \rangle| \leq \|\ell\|_{E'}\|x\|$. Le théorème 2.2.3 page 37 de caractérisation des applications multilinéaires continues assure alors le résultat. □

Nous allons maintenant introduire la notion d'application linéaire transposée. Cette notion jouera un rôle crucial au chapitre 8

Proposition 3.1.2. *Soient E et F deux espaces vectoriels normés et A un élément de l'espace $\mathcal{L}(E,F)$. L'application ${}^tA$ définie par*

$$\left\{\begin{array}{ccl} F' & \longrightarrow & E' \\ \ell & \longmapsto & {}^tA(\ell) : x \mapsto \langle \ell, Ax\rangle \end{array}\right.$$

est une application linéaire continue.

Démonstration. On écrit

$$\begin{array}{rcl} |\langle \ell, Ax\rangle| & \leq & \|\ell\|_{F'}\|Ax\|_E \\ & \leq & \|\ell\|_{F'}\|A\|_{\mathcal{L}(E,F)}\|x\|_E. \end{array}$$

Ainsi donc, on a

$$\begin{array}{rcl} \|{}^tA(\ell)\|_{E'} & = & \displaystyle\sup_{\|x\|_E\leq 1} |\langle \ell, Ax\rangle| \\ & \leq & \|\ell\|_{F'}\|A\|_{\mathcal{L}(E,F)}. \end{array}$$

D'où la proposition. □

Définition 3.1.2. *L'application linéaire définie ci-dessus s'appelle l'application linéaire transposée.*

Proposition 3.1.3. *Soient E, F et G trois espaces vectoriels normés et (A,B) un élément de $\mathcal{L}(E,F)\times\mathcal{L}(F,G)$. On a*

$${}^t(B\circ A) = {}^tA\circ {}^tB$$

De plus, si $E=F$ alors on a

$$A\in U(E) \Longleftrightarrow {}^tA\in U(E') \quad \text{et} \quad ({}^tA)^{-1} = {}^t(A^{-1}).$$

Démonstration. La première relation est purement algébrique. En effet, pour toute forme linéaire ℓ sur G et pour tout x de E, on a

$$\begin{array}{rcl} \langle {}^tA\circ {}^tB(\ell), x\rangle & = & \langle {}^tB(\ell), Ax\rangle \\ & = & \langle \ell, B\circ Ax\rangle \\ & = & \langle {}^t(B\circ A)(\ell), x\rangle. \end{array}$$

Enfin, si $A\in U(E)$, on a $A\circ A^{-1} = A^{-1}\circ A = \mathrm{Id}_E$. En appliquant l'opération de transposition à cette égalité, on trouve que

$${}^tA\circ {}^t(A^{-1}) = {}^t(A^{-1})\circ {}^tA = \mathrm{Id}_{E'}\,.$$

Donc ${}^tA$ est un élément inversible de $\mathcal{L}(E')$ et $({}^tA)^{-1} = {}^t(A^{-1})$. La proposition est ainsi démontrée. □

3.2 Identification d'un espace normé avec un dual

Cette section est très importante pour la suite du cours, notamment pour la théorie des distributions. Il arrive fréquemment de dire que tel espace "est le dual" de tel autre. Il s'agit de donner un sens précis à cette expression.

Proposition 3.2.1. *Soient E et F deux espaces de Banach et B une forme bilinéaire continue sur $F \times E$. L'application δ_B définie par*

$$\delta_B \left\{ \begin{array}{ccl} F & \longrightarrow & E' \\ y & \longmapsto & \delta_B(y) : \ \langle \delta_B(y), x \rangle = B(y, x) \end{array} \right. \tag{3.1}$$

est un élément de $\mathcal{L}(F, E')$.

Démonstration. Comme la forme bilinéaire B est continue, on peut écrire

$$|\langle \delta_B(y), x \rangle| \leq C \|y\|_F \|x\|_E.$$

Par définition de la norme sur E', on a

$$\|\delta_B(y)\|_{E'} \leq C \|y\|_F \ ;$$

d'où la proposition. □

Définition 3.2.1. *Soient E et F deux espaces de Banach et B une forme bilinéaire continue sur $F \times E$. On dit que B identifie E' et F si et seulement si l'application δ_B définie par (3.1) est un isomorphisme de F sur E'.*

Dit autrement, cela signifie qu'il existe une constante C telle que, pour toute forme linéaire continue ℓ sur E, il existe un unique élément y de F tel que,

$$\forall x \in E \, , \ \langle \ell, x \rangle = B(x, y) \quad \text{et} \quad C^{-1} \|y\|_F \leq \|\ell\|_{E'} \leq C \|y\|_F.$$

Théorème 3.2.1. *Soit p dans l'intervalle $[1, +\infty[$ et $F = \ell^{p'}(\mathbb{N})$ $p' = p/(p-1)$. On considère la forme bilinéaire B suivante*

$$B \left\{ \begin{array}{ccl} \ell^{p'}(\mathbb{N}) \times \ell^p(\mathbb{N}) & \longrightarrow & \mathbb{K} \\ (y, x) & \longmapsto & B(y, x) \stackrel{\text{déf}}{=} \displaystyle\sum_{n \in \mathbb{N}} x(n) y(n). \end{array} \right.$$

La forme bilinéaire B identifie $(\ell^p(\mathbb{N}))'$ et $\ell^{p'}(\mathbb{N})$. De plus l'application δ_B définie à la proposition 3.2.1 est une isométrie.

Démonstration. L'inégalité de Hölder dit que

$$|B(y, x)| \leq \|x\|_{\ell^p(\mathbb{N})} \|y\|_{\ell^{p'}(\mathbb{N})}. \tag{3.2}$$

Donc la forme bilinéaire B est continue. Démontrons que l'application δ_B est bijective. Soit T une forme linéaire sur $\ell^p(\mathbb{N})$, cherchons un élément y de $\ell^{p'}(\mathbb{N})$ tel que

$$\forall x \in \ell^p(\mathbb{N}) \, , \ \langle T, x \rangle = B(y, x). \tag{3.3}$$

Si l'identité ci-dessus est vérifiée, elle doit l'être en particulier pour $x = e_n$ où e_n désigne la suite dont tous les termes sont nuls, à l'exception du $n$ième qui vaut 1. Donc, nécessairement y doit vérifier

$$y(n) = \langle T, e_n \rangle. \tag{3.4}$$

L'application δ_B est donc injective. De plus par linéarité, on a que si V désigne l'ensemble des combinaisons linéaires (nécessairement) finies x d'éléments de $(e_n)_{n\in\mathbb{N}}$, on a

$$B(y,x) = \langle T, x\rangle. \tag{3.5}$$

Il s'agit maintenant de démontrer que la suite y définie par (3.4) est bien un élément de $\ell^{p'}(\mathbb{N})$.

Supposons tout d'abord que p vaille 1. Pour tout entier n, on a

$$\begin{aligned} |y(n)| &\leq \|T\|_{(\ell^1(\mathbb{N}))'}\|e_n\|_{\ell^1(\mathbb{N})} \\ &\leq \|T\|_{(\ell^1(\mathbb{N}))'}. \end{aligned}$$

Donc la suite $y = (y(n))_{n\in\mathbb{N}}$ appartient bien à $\ell^\infty(\mathbb{N})$ et l'on a bien $\|y\|_{\ell^\infty(\mathbb{N})} \leq \|T\|_{(\ell^1(\mathbb{N}))'}$. Les deux termes de l'égalité (3.5) sont continues et coïcident sur un sous espace dense de $\ell^1(\mathbb{N})$. Le théorème est donc démontré dans ce cas.

Supposons maintenant p strictement supérieur à 1. Considérons alors la suite $(x_r)_{r\in\mathbb{N}}$ d'éléments de $\ell^p(\mathbb{N})$ définie par

$$x_r \stackrel{\text{déf}}{=} \sum_{n\leq r} \frac{\overline{y}(n)}{|y(n)|}|y(n)|^{\frac{1}{p-1}}e_n.$$

Ceci implique que

$$x_r(n) = \frac{\overline{y}(n)}{|y(n)|}|y(n)|^{\frac{1}{p-1}} \quad \text{si} \quad n\leq r \quad \text{et} \quad 0 \quad \text{sinon.}$$

On a alors

$$\sum_{n\in\mathbb{N}} |x_r(n)|^p = \sum_{n\leq r}|y(n)|^{p'}.$$

$$\begin{aligned} B(y,x_r) &= \sum_{n=0}^{r} \frac{\overline{y}(n)}{|y(n)|}|y(n)|^{\frac{1}{p-1}}B(e_n,y) \\ &= \sum_{n=0}^{r} |y(n)|^{\frac{1}{p-1}+1} \\ &= \sum_{n=0}^{r} |y(n)|^{p'}. \end{aligned}$$

De plus, on sait que

$$B(y,x_r) = \sum_{n=0}^{r} \frac{\overline{y}(n)}{|y(n)|}|y(n)|^{\frac{1}{p-1}}T(e_n) = \langle T, x_r\rangle.$$

La forme linéaire T est continue ; on a donc

$$\begin{aligned} \sum_{n\leq r}|y(n)|^{p'} &\leq \|T\|_{(\ell^p(\mathbb{N}))'}\|x_r\|_{\ell^p(\mathbb{N})} \\ &\leq \|T\|_{(\ell^p(\mathbb{N}))'}\Big(\sum_{n\leq r}|y(n)|^{p'}\Big)^{\frac{1}{p}}. \end{aligned}$$

D'où il vient que

$$\Big(\sum_{n\leq r}|y(n)|^{p'}\Big)^{1-\frac{1}{p}} \leq \|T\|_{(\ell^p(\mathbb{N}))'}.$$

Cette inégalité étant vraie pour tout entier r, la suite y appartient bien à $\ell^{p'}(\mathbb{N})$ et l'on a

$$\|y\|_{\ell^{p'}(\mathbb{N})} \leq \|T\|_{(\ell^p(\mathbb{N}))'}. \tag{3.6}$$

L'espace V étant dense dans $\ell^p(\mathbb{N})$, on observe à nouveau que l'égalité (3.5) affirme que deux formes linéaires continues coïcindent sur un esapce dense et donc partout ce qui conclut la démonstrationdu théorème. $\square$

Exercice 3.2.1. *Soit $c_0(\mathbb{N})$ l'espace de Banach définie par*

$$c_0(\mathbb{N}) \stackrel{\text{déf}}{=} \{x=(x(n))_{n\in\mathbb{N}} \,/\, \lim_{n\to\infty} x(n)=0\} \quad \textit{muni de la norme} \quad \|x\|_{\ell^\infty(\mathbb{N})} \stackrel{\text{déf}}{=} \sup_{n\in\mathbb{N}}|x(n)|.$$

Démontrez que $c_0(\mathbb{N})$ est un sous-espace fermé de $\ell^\infty(\mathbb{N})$. Puis en reprenant la démonstration du théorème ci-dessus, démontrez que la forme bilinéaire B identifie $c_0(\mathbb{N})'$ à l'espace $\ell^1(\mathbb{N})$.

Remarque Au début de ce chapitre, nous avons remarqué qu'en dimension finie, le dual d'un espace vectoriel E était un espace vectoriel de même dimension, donc un espace isomorphe. Rien de tel en dimension infinie. Le théorème précédent nous dit que le dual de $\ell^1(\mathbb{N})$ est un espace isométrique à $\ell^\infty(\mathbb{N})$. Donc l'espace $(\ell^1(\mathbb{N}))'$ a les mêmes propriétés topologiques que l'espace $\ell^\infty(\mathbb{N})$. Mais cet espace n'est pas séparable. Or, on sait d'après un exercice que l'espace $\ell^1(\mathbb{N})$ l'est. En effet, l'espace vectoriel des suites nulles sauf un nombre fini de termes est dense dans $\ell^1(\mathbb{N})$, en fait dans tous les espaces $\ell^p(\mathbb{N})$ pour tout p dans $[1,\infty[$. (exercice : vérifiez-le !)

Exercice 3.2.2. *Soit s un réel quelconque, on définit l'espace $\ell^{2,s}(\mathbb{N})$ par*

$$\ell^{2,s}(\mathbb{N}) \stackrel{\text{déf}}{=} \big\{u=(u(n))_{n\in\mathbb{N}} \,/\, ((1+n^2)^{\frac{s}{2}}u(n))_{n\in\mathbb{N}} \in \ell^2(\mathbb{N})\big\}.$$

et on le munit de la norme

$$\|u\|_{\ell^{2,s}(\mathbb{N})} \stackrel{\text{déf}}{=} \left(\sum_{n\in\mathbb{N}}(1+n^2)^s|u(n)|^2\right)^{\frac{1}{2}}.$$

Soient B_1 et B_2 les deux formes bilinéaires définies ci-après par

$$B_1\left\{\begin{array}{ccl} \ell^{2,s}(\mathbb{N})\times\ell^{2,s}(\mathbb{N}) & \longrightarrow & \mathbb{K} \\ (y,x) & \longmapsto & B_1(y,x) \stackrel{\text{déf}}{=} \displaystyle\sum_{n\in\mathbb{N}}(1+n^2)^s x(n)y(n). \end{array}\right.$$

$$B_2\left\{\begin{array}{ccl} \ell^{2,-s}(\mathbb{N})\times\ell^{2,s}(\mathbb{N}) & \longrightarrow & \mathbb{K} \\ (y,x) & \longmapsto & B_2(y,x) \stackrel{\text{déf}}{=} \displaystyle\sum_{n\in\mathbb{N}} x(n)y(n). \end{array}\right.$$

1) Démontrez que B_1 identifie $(\ell^{2,s}(\mathbb{N}))'$ à $\ell^{2,s}(\mathbb{N})$ et que B_2 identifie $(\ell^{2,s}(\mathbb{N}))'$ à $\ell^{2,-s}(\mathbb{N})$.
2) Exhibez une isométrie linéaire surjective de $\ell^{2,s}(\mathbb{N})$ dans $\ell^{2,-s}(\mathbb{N})$.

3.3 Une définition affaiblie de la convergence dans E'

Définition 3.3.1. *Soit E un espace vectoriel normé. On considère une suite $(\ell_n)_{n\in\mathbb{N}}$ d'éléments de E' et ℓ un élément de E'. On dit que la suite $(\ell_n)_{n\in\mathbb{N}}$ converge $\star$-faiblement vers ℓ et l'on note $\operatorname*{limf\star}_{n\to\infty} \ell_n = \ell$ si et seulement si*

$$\forall x \in E\,,\ \lim_{n\to\infty} \langle \ell_n, x\rangle = \langle \ell, x\rangle.$$

Remarquons tout d'abord que la convergence faible $\star$ est impliquée par la convergence en norme. En effet, pourtout x de E, on a

$$\begin{aligned} \bigl|\langle \ell_n, x\rangle\bigr| &= \bigl|\langle \ell_n - \ell, x\rangle\bigr| \\ &\leq \|\ell_n - \ell\|_{E'}\|x\|_E. \end{aligned}$$

Donnons maintenant un exemple d'une suite $(\ell_n)_{n\in\mathbb{N}}$ de E' qui converge $\star$ faiblement et qui ne converge pas au sens de la norme. Plaçons nous dans l'espace $\ell^1(\mathbb{N})$ et considérons la suite $(\ell_n)_{n\in\mathbb{N}}$ de fonctions linéaires continues définie par

$$\langle \ell_n, x\rangle \stackrel{\text{déf}}{=} x(n).$$

Nous avons pour tout x de $\ell^1(\mathbb{N})$, $\lim_{n\to\infty} x(n) = 0$. Ainsi donc la suite $(\ell_n)_{n\in\mathbb{N}}$ de formes linéaires définies ci-dessus converge $\star$-faiblement vers 0. Comme $\|\ell_n\|_{(\ell^1(\mathbb{N}))'} = 1$, la suite $(\ell_n)_{n\in\mathbb{N}}$ ne converge pas en norme vers 0.

Nous allons maintenant étudier quelques propriétés de la convergence $\star$-faible.

Théorème 3.3.1. *Soient E un espace de Banach et $(\ell_n)_{n\in\mathbb{N}}$ une suite d'éléments de E'. Supposons que $\operatorname*{limf\star}_{n\to\infty} \ell_n = \ell$. Alors la suite $(\ell_n)_{n\in\mathbb{N}}$ est une suite bornée de E', ce qui implique en particulier que ℓ appartient à E'. De plus, on a*

$$\|\ell\|_{\mathcal{L}(E,\mathbb{K})} \leq \liminf_{n\to\infty} \|\ell_n\|_{\mathcal{L}(E,\mathbb{K})}.$$

Démonstration. Ce n'est qu'une traduction dans le cadre des formes linéaires d'un théorème classique : le théorème de Banach-Steinhaus dont voici l'énoncé et la démonstration.

Théorème 3.3.2 (de Banach-Steinhaus)**.** *Soient E et F deux espaces normés et $(L_n)_{n\in\mathbb{N}}$ une suite de $\mathcal{L}(E,F)$. Supposons que E soit complet. Supposons que, pour tout x de E, la limite de la suite $(L_n(x))_{n\in\mathbb{N}}$ existe ; désignons la par $L(x)$.*

Alors la suite $(L_n)_{n\in\mathbb{N}}$ est une suite bornée de $\mathcal{L}(E,F)$, ce qui implique en particulier que L appartient à $\mathcal{L}(E,F)$. De plus, on a

$$\|L\|_{\mathcal{L}(E,F)} \leq \liminf_{n\to\infty} \|L_n\|_{\mathcal{L}(E,F)}.$$

Démonstration. Elle repose sur le lemme suivant.

Lemme 3.3.1. *Soient E et F deux espaces normés et $(L_n)_{n\in\mathbb{N}}$ une suite de $\mathcal{L}(E,F)$. Supposons que E soit complet. Supposons que, pour tout x de E, la suite $(L_n(x))_{n\in\mathbb{N}}$ soit une suite bornée de F.*

Alors la suite $(L_n)_{n\in\mathbb{N}}$ est une suite bornée de $\mathcal{L}(E,F)$.

Démonstration. Considérons les ensembles $\mathcal{F}_{n,p}$ définis par

$$\mathcal{F}_{n,p} = \{x \in E \,/\, \|L_n(x)\|_F \leq p\}.$$

Ces ensembles $\mathcal{F}_{n,p}$ sont des fermés en tant qu'image réciproque de fermés par une application continue. Donc les ensembles $\mathcal{F}_p$ définis par

$$\mathcal{F}_p \stackrel{\text{déf}}{=} \bigcap_{n\in\mathbb{N}} \mathcal{F}_{n,p}$$

sont des ensembles fermés car ce sont des intersections de fermés. De plus, soit x un élément quelconque de E. La suite $(L_n(x))_{n\in\mathbb{N}}$ est supposée bornée. Donc, il existe un entier p tel que x appartienne à $\mathcal{F}_p$. Cela signifie que la réunion de tous les $\mathcal{F}_p$ est l'espace E tout entier. D'après le corollaire 1.2.1 du théorème 1.2.2 de Baire, il existe un entier p_0 tel que $\overset{\circ}{\mathcal{F}}_{p_0} \neq \emptyset$. Donc, il existe un point x_0 et un réel strictement positif α tel que $B(x_0,\alpha)$ soit inclus dans F_{p_0}. On a donc

$$\begin{aligned}
\sup_{\substack{n\in\mathbb{N}\\ \|x\|\leq\alpha}} \|L_n(x)\|_F &\leq \sup_{\substack{n\in\mathbb{N}\\ \|x\|\leq\alpha}} \|L_n(x+x_0)\|_F + \|L_n(x_0)\|_F\\
&\leq \sup_{\substack{n\in\mathbb{N}\\ x\in B(x_0,\alpha)}} \|L_n(x)\|_F + \|L_n(x_0)\|_F\\
&\leq 2p_0.
\end{aligned}$$

Ainsi, pour tout x de E de norme plus petite que 1, on peut écrire

$$\begin{aligned}
\|L_n(x)\|_F &\leq \alpha^{-1}\|L_n(\alpha x)\|_F\\
&\leq \frac{2p_0}{\alpha}\cdot
\end{aligned}$$

Donc la suite $(L_n)_{n\in\mathbb{N}}$ est une suite bornée de $\mathcal{L}(E,F)$. □

Retour à la démonstration du théorème 3.3.2. Par passage à la limite, on obtient alors que,

$$\forall x \in B(x_0,\alpha)\,,\ \|L(x)\|_F \leq 2p_0.$$

Démontrons maintenant la majoration de la norme de L. Soit x un élément de E de norme 1. On peut alors écrire

$$\begin{aligned}
\|L(x)\|_F &= \lim_{n\to\infty} \|L_n(x)\|_F\\
&= \liminf_{n\to\infty} \|L_n(x)\|_F\\
&\leq \liminf_{n\to\infty} \|L_n\|_{\mathcal{L}(E,F)} \quad (\text{car } \|x\|_E = 1).
\end{aligned}$$

Le théorème est ainsi complètement démontré. □

Il suffit d'appliquer ce théorème avec $F = \mathbb{K}$ pour obtenir le théorème 3.3.1. □

Le théorème suivant, bien que de démonstration très simple, sera utile dans la suite.

Théorème 3.3.3. *Soient E et F des espaces vectoriels normés, on considère un élément A de $\mathcal{L}(E,F)$, une suite $(\ell_n)_{n\in\mathbb{N}}$ de F' et un élément ℓ de F'. On a*

$$\operatorname*{limf\star}_{n\to\infty} \ell_n = \ell \quad \textit{dans} \quad F' \Longrightarrow \operatorname*{limf\star}_{n\to\infty} {}^t\!A\ell_n = {}^t\!A\ell \quad \textit{dans} \quad E'.$$

Démonstration. En effet, on a, pour tout x appartenant à E,

$$\langle {}^t A\ell_n, x\rangle = \langle \ell_n, Ax\rangle.$$

Par hypothèse, on a

$$\lim_{n\to\infty} \langle \ell_n, Ax\rangle = \langle \ell, Ax\rangle.$$

Mais, par définition de la transposée, $\langle \ell, Ax\rangle = \langle {}^t A\ell, x\rangle$. Ainsi donc

$$\lim_{n\to\infty} \langle {}^t A\ell_n, x\rangle = \langle {}^t A\ell, x\rangle.$$

Le théorème est démontré. □

Proposition 3.3.1. *Soient $(x_n)_{n\in\mathbb{N}}$ et $(\ell_n)_{n\in\mathbb{N}}$ deux suites de E et de E', E étant un espace de Banach.*

$$\left(\operatorname*{limf\star}_{n\to\infty} \ell_n = \ell \quad et \quad \lim_{n\to\infty} \|x_n - x\| = 0\right) \Longrightarrow \lim_{n\to\infty} \langle \ell_n, x_n\rangle = \langle \ell, x\rangle.$$

Démonstration. Commençons par écrire que

$$\langle \ell_n, x_n\rangle = \langle \ell_n, x_n - x\rangle + \langle \ell_n - \ell, x\rangle + \langle \ell, x\rangle.$$

D'après le théorème 3.3.1 de Banach-Steinhaus, la suite $(\ell_n)_{n\in\mathbb{N}}$ est bornée. Donc, il existe un réel M tel que l'on ait

$$|\langle \ell_n, x_n\rangle - \langle \ell, x\rangle| \leq M\|x_n - x\| + |\langle \ell_n - \ell, x\rangle|.$$

Soit ε un réel strictement positif arbitraire. Il existe un entier n_0 tel que

$$n \geq n_0 \Longrightarrow \|x_n - x\| < \frac{\varepsilon}{2M+1}\cdot$$

Le point x étant fixe, il existe un entier n_1 tel que l'on ait

$$n \geq n_1 \Longrightarrow |\langle \ell_n - \ell, x\rangle| < \frac{\varepsilon}{2}\cdot$$

Donc, on a

$$n \geq \max\{n_0, n_1\} \Longrightarrow |\langle \ell_n, x_n\rangle - \langle \ell, x\rangle| < \varepsilon.$$

D'où la proposition. □

Lorsque que l'on a identifié un espace à un dual, on peut définir une notion de convergence faible $\star$ sur cet espace de la manière suivante.

Définition 3.3.2. *Soient E et F deux espaces de Banach, B une forme biliéaire continue sur $F\times E$ qui identifie F à E'. On dit (par abus de language) qu'une suite $(y_n)_{n\in\mathbb{N}}$ converge $\star$-faiblement vers un élément y de F si et seulement si la suite $(\delta_B(y_n))_{n\in\mathbb{N}}$ converge $\star$- faiblement vers $\delta_B(y)$, ce qui signifie exactement que*

$$\forall x \in E\,,\ \lim_{n\to\infty} B(y_n, x) = B(y, x).$$

Donnons un exemple de cette extension. Étant donné q dans $]1,\infty]$, on considère une suite $(y_p)_{p\in\mathbb{N}}$ d'éléments de $\ell^q(\mathbb{N})$. D'après le théorème d'identification 3.2.1, la convergence de la suite $(y_p)_{p\in\mathbb{N}}$ vers un élément y de $\ell^p(\mathbb{N})$ signifie exactement que

$$\forall x \in \ell^{q'}(\mathbb{N})\,,\ \lim_{p\to\infty}\sum_{n=0}^{\infty} y_p(n)x(n) = 0.$$

Le théorème suivant est un résultat fondamental qui peut être compris comme un résultat de compacité faible $\star$ pour la boule unité d'un dual.

Théorème 3.3.4 (de compacité faible-$\star$)**.** *Soit E un espace séparable, on considère une suite $(\ell_n)_{n\in\mathbb{N}}$ bornée dans E'. Il existe un élément ℓ de E' et une fonction d'extraction ψ de $\mathbb{N}$ dans $\mathbb{N}$ telle que l'on ait*

$$\operatorname*{limf\star}_{n\to\infty} \ell_{\psi(n)} = \ell.$$

Démonstration. Soit $(a_p)_{p\in\mathbb{N}}$ une suite dense dans E. Nous allons tout d'abord extraire de la suite $(a_p)_{p\in\mathbb{N}}$ une famille libre et travailler sur l'espace vectoriel engendré. Plus précisément, on définit $\varphi(0)$ comme le plus petit entier p tel que a_p soit non nul. Soit $\varphi(1)$ le plus petit des entiers p tels que a_p n'appartienne pas à $\mathbb{K}a_{\varphi(0)}$. Supposons construit $(\varphi(0),\cdots,\varphi(p))$ tels que

$$\forall p' \leq p\,,\ a_{\varphi(p)} \notin V_{p'} \stackrel{\text{déf}}{=} \mathrm{Vect}\{a_{\varphi(0)},\cdots a_{\varphi(p')}\}.$$

Considérons $\varphi(p+1)$ comme le plus petit des entiers tels que a_q n'appartienne pas à V_p. On définit

$$V = \bigcup_{p\in N} V_p = \mathrm{Vect}\big\{a_{\varphi(p)}\,,\ p\in\mathbb{N}\big\}.$$

La famille $(a_{\varphi(p)})_{p\in\mathbb{N}}$ est une base (au sens de l'algèbre linéaire) de V. Dorénavant, on omet de noter l'extraction et $(a_p)_{p\in\mathbb{N}}$ désigne maintenant une suite de vecteurs de E linéairement indépendant telle que $\mathrm{Vect}\{a_p\,,\ p\in\mathbb{N}\}$ est dense dans E.

Nous allons utiliser le procédé diagonal de Cantor pour construire une fonction d'extraction ψ telle que

$$\forall j\in\mathbb{N}\,,\ \lim_{n\to\infty}\langle \ell_{\psi(n)}, a_j\rangle = \lambda_j. \tag{3.7}$$

La suite $(\langle \ell_n, x_0\rangle)_{n\in\mathbb{N}}$ est une suite bornée de $\mathbb{K}$ donc il existe un élément λ_0 de $\mathbb{K}$ et une fonction d'extraction ϕ_0 de $\mathbb{N}$ dans $\mathbb{N}$ telle que l'on ait

$$\lim_{n\to\infty}\langle \ell_{\phi_0(n)}, a_0\rangle = \lambda_0.$$

Supposons construites une suite finie $(\phi_j)_{0\leq j\leq m}$ de fonctions strictement croissantes de $\mathbb{N}$ dans $\mathbb{N}$ et une suite finie $(\lambda_j)_{1\leq j\leq m}$ de scalaires telles que, pour tout $j\leq m$, on ait

$$\lim_{n\to\infty}\langle \ell_{\phi_0\circ\cdots\circ\phi_j(n)}, a_j\rangle = \lambda_j.$$

La suite $(\langle \ell_{\phi_0\circ\cdots\circ\phi_m(n)}, a_{m+1}\rangle)_{n\in\mathbb{N}}$ est une suite bornée de $\mathbb{K}$. Il existe donc une fonction strictement croissante ϕ_{m+1} de $\mathbb{N}$ dans $\mathbb{N}$ et un élément λ_{m+1} de $\mathbb{K}$ tels que

$$\forall j\leq m+1\,,\ \lim_{n\to\infty}\langle \ell_{\phi_0\circ\cdots\circ\phi_{m+1}(n)}, a_j\rangle = \lambda_j.$$

Posons $\psi(n) = \phi_0 \circ \cdots \circ \phi_n(n)$. C'est une fonction strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$. En effet, observons que toute fonction μ strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ vérifie $\mu(n) \geq n$. Donc, si $n < m$, on a

$$\psi(n) = \varphi_0 \circ \cdots \circ \varphi_n(n) < \varphi_0 \circ \cdots \circ \varphi_n \circ \varphi_{n+1} \circ \varphi_m(m) = \psi(m).$$

De plus, pour tout entier j, on a, pour tout n strictement plus grand que j

$$\langle \ell_{\psi(n)}, x_j \rangle = \langle \ell_{\varphi_0 \circ \varphi_1 \cdot \circ \varphi_j(\varphi_{j+1} \circ \varphi_n(n)}, x_j \rangle.$$

Ainsi, une sous-suite d'une suite convergente étant convergente, on a

$$\forall j \in \mathbb{N}\,,\ \lim_{n\to\infty} \langle \ell_{\psi(n)}, x_j \rangle = \lambda_j.$$

Nous allons maintenant démontrer que, pour tout x de E, la suite $\big(\langle \ell_{\psi(n)}, x \rangle\big)_{n\in\mathbb{N}}$ est une suite de Cauchy. Soit x dans E ; la suite $(x_j)_{j\in\mathbb{N}}$ étant dense,

$$\forall \varepsilon > 0\,,\ \exists j_\varepsilon \in \mathbb{N} \,/\, |x - x_{j_\varepsilon}\|_E < \frac{\varepsilon}{4M} \quad \text{avec} \quad M \stackrel{\text{def}}{=} \sup_{n\in\mathbb{N}} \|\ell_n\|_{E'}.$$

Écrivons maintenant que

$$\begin{aligned}
\big|\langle \ell_{\psi(n+p)}, x \rangle - \langle \ell_{\psi(n)}, x \rangle\big| \;&\leq\; \big|\langle \ell_{\psi(n+p)}, x \rangle - \langle \ell_{\psi(n+p)}, x_{j_\varepsilon} \rangle\big| + \big|\langle \ell_{\psi(n+p)}, x_{j_\varepsilon} \rangle - \langle \ell_{\psi(n)}, x_{j_\varepsilon} \rangle\big| \\
&\quad + \big|\langle \ell_{\psi(n)}, x_{j_\varepsilon} \rangle - \langle \ell_{\psi(n)}, x \rangle\big| \\
&\leq\; \|\ell_{\psi(n+p)}\|_{E'} \|x - x_{j_\varepsilon}\|_E + \big|\langle \ell_{\psi(n+p)}, x_{j_\varepsilon} \rangle - \langle \ell_{\psi(n)}, x_{j_\varepsilon} \rangle\big| \\
&\quad + \|\ell_{\psi(n)}\|_{E'} \|x_{j_\varepsilon} - x\|_E \\
&\leq\; \frac{\varepsilon}{2} + \big|\langle \ell_{\psi(n+p)}, x_{j_\varepsilon} \rangle - \langle \ell_{\psi(n)}, x_{j_\varepsilon} \rangle\big|.
\end{aligned}$$

La suite $\big(\langle \ell_{\psi(n)}, x_{j_\varepsilon} \rangle\big)_{n\in\mathbb{N}}$ étant convergente, elle est de Cauchy. Ainsi donc

$$\exists n_\varepsilon \,/\, \forall n \geq n_\varepsilon\,,\ \forall p \in \mathbb{N}\,,\ \big|\langle \ell_{\psi(n+)}, x_{j_\varepsilon} \rangle\big| < \frac{\varepsilon}{2} \cdot$$

On en déduit donc que

$$\forall \varepsilon > 0,\ \exists n_\varepsilon \,/\, \forall n \geq n_\varepsilon\,,\ \forall p \in \mathbb{N}\,,\ \big|\langle \ell_{\psi(n+p)}, x \rangle - \langle \ell_{\psi(n)}, x \rangle\big| < \varepsilon$$

La suite $\big(\langle \ell_{\psi(n)}, x \rangle\big)_{n\in\mathbb{N}}$ est une suite de Cauchy et donc elle converge vers une limite que l'on note $\ell(x)$. Par unicité de la limite, on montre que ℓ ainsi définie est une application linéaire (exercice : démontrez-le). De plus, on a

$$\big|\ell(x)\big| \leq M\|x\|_E.$$

Ceci conclut la démonstration du théorème.

□

Dans le cadre de la définition 3.3.2, le théorème 3.3.4 peut s'énoncer de la manière suivante.

Théorème 3.3.5. *Soient E un espace de Banach séparable et F un espace de Banach, on considère une forme bilinéaire B qui identifie F à E'. Alors, pour toute suite bornée $(y_n)_{n\in\mathbb{N}}$ d'éléments de F, il existe un élément y de F et une fonction d'extraction ϕ telle que*

$$\forall x \in E\,,\ \lim_{n\to\infty} B(y_{\phi(n)}, x) = B(y, x).$$

Ce théorème est très important. Il signifie que lorsqu'un espace F peut s'identifier à un dual, alors de toute suite bornée, on peut en extraire une sous-suite qui converge au sens affaibli de la définition 3.3.2. Ce fait est extrêmement utilisé. On verra un exemple important d'application au chapitre qui suit lorsque nous démontrerons un théorème de diagonalisation dans les espaces de Hilbert (le théorème 4.4.2 page 75).

Chapitre 4

Espaces de Hilbert

Introduction

Ce chapitre est consacré à l'étude des espaces dits de Hilbert qui sont des espaces de Banach dont la norme est de type euclidienne c'est-à-dire qu'elle est associé à un produit scalaire. Ces espaces sont la généralisation à la dimension infinie des espaces euclidiens de dimension finie. Le point fondamendal est que l'on peut y faire de la géométrie. On sait par exemple ce que signifie deux vecteurs (qui peuvent être des fonctions) orthogonaux. Insistons sur l'un des points clefs de l'analyse fonctionnelle : un élément d'un espace de Banach (très souvent une fonction) est vraiment vu comme un point (ou un vecteur) d'un espace géométrique.

La première section présente le concept d'espace de Hilbert notamment ce que signifie précisément qu'une norme est associée à un produit scalaire.

La deuxième section illustre comment cette structure généralise à la dimension infinie la notion d'espace euclidien. Deux exemples :

— il existe une notion de projection orthogonale sur un fermé convexe;

— il existe une notion de *base hilbertienne* qui généralise à la dimension infinie celle familière de base orthonormée d'un espace euclidien de dimension finie.

Dans la troisième section, on montre que le dual d'un espace de Hilbert $\mathcal{H}$ peut s'identifier à l'espace $\mathcal{H}$ lui-même. Nous verrons au chapitre 6 comment cette identification permet de démontrer de façon spectaculairement simple un théorème d'existence et d'unicité de solutions d'une équation aux dérivées partielles. Notons enfin que le phénomème mis en évidence dans la section 3.2 n'a pas lieu dans le cadre des espaces de Hilbert.

Dans la quatrième et dernière section de ce chapitre, on introduit la notion d'adjoint d'un opérateur continue d'un espace de Hilbert dans lui-même ainsi que la notion plus délicate et nouvelle d'opérateur compact. Le résultat important de cette section est la généralisation du théorème classique de diagonalisation des opérateurs symétriques en dimension finie. Ce théorème sera appliqué au chapitre 6.

4.1 Le concept d'orthogonalité

Dans toute la suite, nous conviendrons que si $\mathbb{K} = \mathbb{R}$, $\overline{\lambda} = \lambda$ et si $\mathbb{K} = \mathbb{C}$, c'est l'opérateur de conjugaison usuelle. Rappelons tout d'abord la définition d'un produit scalaire.

Définition 4.1.1. *Soient E et F deux espaces vectoriels sur $\mathbb{K}$. On appelle application antilinéaire de E dans F toute application ℓ telle que, pour tout couple (x, y) de $E \times E$, et tout* *scalaire λ de $\mathbb{K}$, on ait*

$$\ell(\lambda x + y) = \overline{\lambda}\ell(x) + \ell(y).$$

Une application f de $E \times E$ dans $\mathbb{K}$ est dite sesquilinéaire[1] *si elle est linéaire par rapport à la première variable et antilinéaire par rapport à la seconde. Autrement dit*

$$\forall \lambda \in \mathbb{K}\,,\ \forall (x,y,z) \in E^3 \quad \begin{array}{rcl} f(\lambda x + y, z) & = & \lambda f(x,y) + f(y,z) \\ f(z, \lambda x + y) & = & \overline{\lambda} f(z,x) + f(z,y). \end{array}$$

Une application f de $E \times E$ dans $\mathbb{K}$ est appelée produit scalaire si et seulement si elle est sesquilinéaire et hermitienne définie positive, c'est-à-dire vérifie

$$\begin{gathered} \forall (x,y) \in E^2\,,\ f(y,x) = \overline{f(x,y)} \quad \text{(hermitienne)} \\ \forall x \in E\,,\ f(x,x) \in \mathbb{R}^+, \quad \text{(positive)} \\ f(x,x) = 0 \Longleftrightarrow x = 0 \quad \text{(définie)}. \end{gathered}$$

Remarque On note très souvent $(\cdot|\cdot)$ un produit scalaire et $\|\cdot\|^2 = (\cdot|\cdot)$.

Proposition 4.1.1. *Soient E un espace vectoriel sur $\mathbb{K}$ et $(\cdot|\cdot)$ un produit scalaire sur E. L'application $x \mapsto (x|x)^{\frac{1}{2}}$ est une norme sur E et l'on a*

$$\forall (x,y) \in E^2\,,\ |(x|y)| \leq \|x\|\,\|y\| \qquad \text{(Inégalité de Cauchy-Schwarz)}.$$

$$\forall (x,y) \in E^2\,,\ \|x\|^2 + \|y\|^2 = 2\left\|\frac{x+y}{2}\right\|^2 + 2\left\|\frac{x-y}{2}\right\|^2 \qquad \text{(Relation de la médiane)}.$$

Démonstration. On observe que

$$0 \leq \left\|\overline{(x|y)}\frac{x}{\|x\|} - \|x\|y\right\|^2.$$

En développant le second membre, on obtient

$$0 \ \leq \ |\overline{(x|y)}|^2 - 2\,\Re e\Big(\overline{(x|y)}(x|y)\Big) + \|x\|^2\|y\|^2.$$

Pour démontrer que $x \mapsto (x|x)^{\frac{1}{2}}$ est une norme, il suffit d'écrire que

$$\begin{array}{rcl} (x+y|x+y) & \leq & (x|x) + 2\,\Re e(x|y) + (y|y) \\ & \leq & (x|x) + 2(x|x)^{\frac{1}{2}}(y|y)^{\frac{1}{2}} + (y|y) \\ & \leq & \Big((x|x)^{\frac{1}{2}} + (y|y)^{\frac{1}{2}}\Big)^2 \quad \text{et que} \\ (\lambda x|\lambda x) & \leq & \lambda\overline{\lambda}\|x\|^2. \end{array}$$

Pour démontrer l'identité de la médiane, il suffit d'observer que

$$\begin{array}{rcl} 2\left\|\frac{x+y}{2}\right\|^2 + 2\left\|\frac{x-y}{2}\right\|^2 & = & \frac{1}{2}\|x+y\|^2 + \frac{1}{2}\|x-y\|^2 \\ & = & \frac{1}{2}(\|x\|^2 + 2\,\Re e(x|y) + \|y\|^2 + \|x\|^2 - 2\,\Re e(x|y) + \|y\|^2) \\ & = & \|x\|^2 + \|y\|^2. \end{array}$$

D'où la proposition. $\square$

1. du latin "sesqui" qui signifie un demi en plus

Exercice 4.1.1. *Soit $(E, \|\cdot\|)$ un espace vectoriel normé sur $\mathbb{K}$. Supposons que*

$$\forall (x,y) \in E^2\,,\ \|x\|^2 + \|y\|^2 = 2\left\|\frac{x+y}{2}\right\|^2 + 2\left\|\frac{x-y}{2}\right\|^2.$$

Démontrez qu'il existe une forme sesquilinéaire telle que l'on ait $\|x\|^2 = (x|x)$. Il est conseillé de distinguer le cas où $\mathbb{K} = \mathbb{R}$, plus facile, de celui où $\mathbb{K} = \mathbb{C}$.

Le produit scalaire permet de définir la notion d'orthogonalité. Ce concept, familier dans le plan ou dans l'espace à trois dimensions, se révèlera extrêmement fécond dans les espaces de dimension infinie, notamment dans les espaces de fonctions.

Définition 4.1.2. *Soient E un espace vectoriel muni d'un produit scalaire et x et y deux éléments de E. On dit que x et y sont orthogonaux (et l'on note $x \perp y$) si et seulement si $(x|y) = 0$.*

Rappelons pour mémoire la fameuse relation de Pythagore :

$$x \perp y \Longrightarrow \|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

Proposition 4.1.2. *Soit E un espace vectoriel sur $\mathbb{K}$ muni d'un produit scalaire et A une partie de E. Soit*

$$A^\perp \stackrel{\text{déf}}{=} \{x \in E / \ \forall a \in A\,,\ x \perp a\}.$$

La partie $A^\perp$ est un sous-espace vectoriel fermé de E.

Démonstration. L'application linéaire $\mathcal{L}_a$ définie par $x \mapsto (x|a)$ est une application linéaire continue de E dans $\mathbb{K}$. Donc le noyau de $\mathcal{L}_a$ est un sous-espace vectoriel fermé de E. Il est clair que

$$A^\perp = \bigcap_{a \in A} \ker \mathcal{L}_a.$$

Donc $A^\perp$ est un sous-espace vectoriel fermé en tant qu'intersection de sous-espaces vectoriels fermés, d'où la proposition. $\square$

Exercice 4.1.2. *Soit E un espace vectoriel sur $\mathbb{K}$ muni d'un produit scalaire hermitien et A une partie de E. Démontrez que $\overline{A}^\perp = A^\perp = \mathrm{Vect}(A)^\perp$.*

Définition 4.1.3. *Soit $\mathcal{H}$ un espace vectoriel sur $\mathbb{K}$ muni d'un produit scalaire $(\cdot|\cdot)$. On dit que $\mathcal{H}$ est un espace de Hilbert si et seulement si l'espace $\mathcal{H}$ muni de la norme associée au produit scalaire est un espace complet.*

Cette structure d'espace de Hilbert est tout à fait fondamentale. On y retrouvera beaucoup de méthodes et de concepts simples de géométrie que l'on utilise dans les espaces de dimension finie.

4.2 Les propriétés des espaces de Hilbert

La relation de Pythagore reçoit la généralisation suivante.

Théorème 4.2.1. *Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments deux à deux orthogonaux d'un espace de Hilbert $\mathcal{H}$. La série $\sum_n x_n$ converge si et seulement si la série $\sum_n \|x_n\|^2$ converge et l'on a alors*

$$\Big\|\sum_{n\in\mathbb{N}} x_n\Big\|^2 = \sum_{n\in\mathbb{N}} \|x_n\|^2$$

Démonstration. Posons $S_q = \sum_{p\leq q} x_p$. La relation de Pythagore dit que

$$\sum_{p\leq q} \|x_p\|^2 = \|S_q\|^2. \tag{4.1}$$

Le membre de droite de l'inégalité ci-dessus converge donc est majoré indépendemment de q. Donc la série $\sum_n \|x_n\|^2$ converge. Réciproquement, si la série $\sum_n \|x_n\|^2$ converge, alors on a

$$\begin{aligned} \|S_{q+q'} - S_q\|^2 &= \sum_{q+1}^{q+q'} \|x_p\|^2 \\ &\leq \sum_{q+1}^{\infty} \|x_p\|^2. \end{aligned}$$

En passant à la limite dans l'inégalité (4.1), on achève la démonstration du théorème. □

Exercice 4.2.1. *Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments d'un espace de Hilbert $\mathcal{H}$ telle que*

$$\sum_0^\infty \|x_n\|^2 < \infty.$$

On suppose qu'il existe un entier N_0 tel que, si $|n-m| \geq N_0$, alors x_n et x_m sont orthogonaux. Démontrez qu'alors la série $\sum_n x_n$ est convergente et qu'il existe une constante C, ne dépendant que de N_0 telle que l'on ait

$$\Big\|\sum_{n\in\mathbb{N}} x_n\Big\|^2 \leq C \sum_{n\in\mathbb{N}} \|x_n\|^2.$$

Beaucoup des propriétés remarquables des espaces de Hilbert repose sur le théorème de projection sur les convexes.

Théorème 4.2.2 (de projection sur un convexe fermé)**.** *Soit Γ une partie convexe fermée d'un espace de Hilbert $\mathcal{H}$. Pour tout point x de $\mathcal{H}$, il existe un unique point de Γ, noté $p_\Gamma(x)$ et appelé projection de x sur Γ, tel que*

$$\|x - p_\Gamma(x)\| = \inf_{g\in\Gamma} \|x-g\|.$$

Démonstration. Elle repose sur l'identité de la médiane. Soient γ et γ' deux poins de Γ, on a

$$\frac{1}{2}\|\gamma-\gamma'\|^2 = \|x-\gamma\|^2 + \|x-\gamma'\|^2 - 2\left\|x - \frac{\gamma+\gamma'}{2}\right\|^2. \tag{4.2}$$

Remarquons que comme l'ensemble Γ est convexe, le point $\dfrac{\gamma+\gamma'}{2}$ appartient aussi à Γ. Démontrons tout d'abord l'unicité. Supposons qu'il existe deux points γ_1 et γ_2 de Γ réalisant le minimum. D'après la relation de la médiane, on a

$$\begin{aligned} \frac{1}{2}\|\gamma_1-\gamma_2\|^2 &= \|x-\gamma_1\|^2 + \|x-\gamma_2\|^2 - 2\left\|x - \frac{\gamma_1+\gamma_2}{2}\right\|^2 \\ &= 2d^2 - 2\left\|x - \frac{\gamma_1+\gamma_2}{2}\right\|^2 \\ &\leq 2d^2 - 2d^2 = 0. \end{aligned}$$

Ceci assure l'unicité. Démontrons maintenant l'existence. Par définition de la borne inférieure, il existe une suite minimisante, c'est-à-dire une suite $(\gamma_n)_{n\in\mathbb{N}}$ d'éléments de Γ telle

$$\lim_{n\to\infty} \|x-\gamma_n\| = d \stackrel{\text{déf}}{=} \inf_{\gamma\in\Gamma} \|x-\gamma\|.$$

On utilise à nouveau la relation (4.2) pour dire que, pour tout couple d'entiers (n,m), on a

$$\frac{1}{2}\|\gamma_n-\gamma_m\|^2 = \|x-\gamma_n\|^2 + \|x-\gamma_m\|^2 - 2\left\|x-\frac{\gamma_n+\gamma_m}{2}\right\|^2.$$

La partie Γ étant convexe, le milieu de γ_n et de γ_m appartient aussi à Γ. Ainsi, l'on a

$$\frac{1}{2}\|\gamma_n-\gamma_m\|^2 \leq \|x-\gamma_n\|^2 + \|x-\gamma_m\|^2 - 2d^2.$$

Vu la définition de la suite $(\gamma_n)_{n\in\mathbb{N}}$, ceci implique que cette suite est de Cauchy. L'espace $\mathcal{H}$ étant de Hilbert, il est complet. Comme la partie Γ est fermée, elle est complète et donc la suite $(\gamma_n)_{n\in\mathbb{N}}$ converge dans Γ, ce qui conclut la démonstration de ce théorème. □

Proposition 4.2.1. *Soient Γ une partie convexe fermée d'un espace de Hilbert $\mathcal{H}$ et x un point de $\mathcal{H}$, on désigne par $p_\Gamma(x)$ l'unique point de Γ tel que $d(x,\Gamma) = \|x-p_\Gamma(x)\|$. On a alors*

$$\Re e(x-p_\Gamma(x)|p_\Gamma(x)-\gamma) \geq 0 \quad \textit{et} \quad \|p_\Gamma(x)-p_\Gamma(x')\| \leq \|x-x'\|.$$

Démonstration. Pour démontrer la première inégalité, observons que, comme Γ est convexe, on a pour tout λ in $[0,1]$ et pour tout γ in Γ,

$$\big\|x-\big((1-\lambda)p_\Gamma(x)+\lambda\gamma\big)\big\|^2 = \big\|x-p_\Gamma(x)+\lambda(p_\Gamma(x)-x)\big\|^2 \geq \|x-p_\Gamma(x)\|^2.$$

Un calcul immédiat implique que

$$\forall\lambda\in[0,1]\,,\ \forall\gamma\in\Gamma\ \ \lambda\big(2\,\Re e(x-p_\Gamma(x)|p_\Gamma(x)-\gamma)+\lambda\|p_\Gamma(x)-\gamma\|^2\big)\geq 0$$

ce qui assure la première inégalité. Pour démontrer la seconde, utilisons la première avec respectivement x et $\gamma = p_\Gamma(x')$ et x' et $\gamma = p_\Gamma(x)$. Ceci assure que

$$\begin{aligned} \Re e\big(x-p_\Gamma(x)|p_\Gamma(x)-p_\Gamma(x')\big) &\geq 0 \quad \text{et} \\ \Re e\big(x'-p_\Gamma(x')|p_\Gamma(x')-p_\Gamma(x)\big) &\geq 0. \end{aligned}$$

Par addition, il vient

$$\Re e\big(x-x'-(p_\Gamma(x)-p_\Gamma(x'))\big|p_\Gamma(x)-p_\Gamma(x')\big)\geq 0.$$

Ceci implique, grâce à l'inégalité de Cauchy-Schwarz que

$$\|p_\Gamma(x)-p_\Gamma(x')\|^2 \leq \Re e\big(x-x'\big|p_\Gamma(x)-p_\Gamma(x')\big) \leq \|x-x'\|\,\|p_\Gamma(x)-p_\Gamma(x')\|$$

ce qui conclut la démonstration du thérème. □

Presque toutes les propriétés des espaces de Hilbert peuvent être vues comme des corollaires du théorème 4.2.2 ci-dessus.

Corollaire 4.2.1. *Soit F un sous-espace vectoriel fermé d'un espace de Hilbert $\mathcal{H}$. On a alors*

$$\mathcal{H} = F\oplus F^\perp \quad \textit{et} \quad x = p_F(x)+p_{F^\perp}(x).$$

Démonstration. Considérons un élément x quelconque de $\mathcal{H}$. Pour tout point f de F, et pour tout réel λ, on a

$$\|x - p_F(x) + \lambda f\|^2 = \lambda^2 \|f\|^2 + 2\lambda \,\Re e(x - p_F(x)|f) + \|x - p_F(x)\|^2.$$

Par définition de la projection, il faut que

$$\|x - p_F(x) + \lambda f\|^2 \geq \|x - p_F(x)\|^2.$$

Ceci impose donc que, pour tout f appartenant à F, on ait

$$\Re e(x - p_F(x)|f) = 0.$$

Si $\mathbb{K} = \mathbb{R}$, il n'y a rien de plus à dire. Si $\mathbb{K} = \mathbb{C}$, il suffit de changer f en if pour obtenir que

$$x - p_F(x) \in F^\perp.$$

Nous venons donc de démontrer que $\mathcal{H} = F + F^\perp$. Mais, si $x \in F \cap F^\perp$, alors $(x|x) = 0$ et donc $x = 0$. Le corollaire est ainsi démontré. □

Corollaire 4.2.2. *Soit A une partie quelconque d'un espace de Hilbert $\mathcal{H}$. Cette partie A est totale si et seulement si $A^\perp = \{0\}$.*

Démonstration. Utilisons l'exercice 4.1.2 qui dit que $A^\perp = (\overline{\mathrm{Vect}(A)})^\perp$. Si $A^\perp = \{0\}$, cela signifie, d'après le corollaire 4.2.1, que $\overline{\mathrm{Vect}(A)}$ est égal à $\mathcal{H}$, ce qui signifie exactement que la partie A est totale.

Réciproquement, si la partie A est totale, alors $\overline{\mathrm{Vect}(A)}$ est égal à $\mathcal{H}$, d'où $A^\perp = \{0\}$. Le corollaire est donc démontré. □

Exercice 4.2.2. *Soit A une partie quelconque d'un espace de Hilbert $\mathcal{H}$. Démontrez que*

$$(A^\perp)^\perp = \overline{\mathrm{Vect}}(A).$$

Définition 4.2.1. *Soit $\mathcal{H}$ un espace de Hilbert séparable de dimension infinie. On appelle base hilbertienne ou base orthonormale de $\mathcal{H}$ toute suite $(e_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{H}$ qui est totale et telle que*

$$(e_n|e_m) = \delta_{n,m} \quad \text{avec} \quad \delta_{n,m} = 1 \quad \textit{si} \quad n = m \quad \textit{et} \quad 0 \quad \textit{sinon.} \tag{4.3}$$

Remarque Une base hilbertienne n'est pas une base algébrique.

Exercice 4.2.3. *On considère l'espace $\ell^2(\mathbb{N})$ la suite $(e_n)_{n\in\mathbb{N}}$ d'éléments de $\ell_2(\mathbb{N})$ définie par*

$$e_n(k) = \delta_{n,k}.$$

Démontrer que $(e_n)_{n\in\mathbb{N}}$ est une base hilbertienne de $\ell^2(\mathbb{N})$. Quel est l'espace vectoriel engendré par les e_n ?

Théorème 4.2.3. *Dans un espace de Hilbert séparable $\mathcal{H}$, il existe des bases hilbertiennes.*

Démonstration. Considérons une partie dénombrable totale $(a_n)_{n\in\mathbb{N}}$. On peut supposer, quitte à extraire une sous-suite, que la famille $(a_n)_{n\in\mathbb{N}}$ est libre (au sens de l'algèbre linéaire). En effet, on définit la fonction d'extraction suivante

$$\phi(0) = \min\{n \,/\, a_n \neq 0\} \quad \text{et} \quad \phi(n) = \min\left\{m \,/\, a_m \notin \mathrm{Vect}\{a_0, \ldots, a_{\phi(n)}\rangle\right\}.$$

C'est un exercice laissé au lecteur que de vérifier que l'espace vectoriel engendré par $a_{\phi(n)}$ est égal à celui engendré par les a_n. Nous allons maintenant utiliser le procécé d'orthogonalisation de Schmidt. Rappelons ce procédé. On pose

$$e_1 = \frac{a_1}{\|a_1\|}\cdot$$

Supposons définis les termes $(e_j)_{1\leq j\leq n}$ vérifiant la relation (4.3) et telle que

$$\mathrm{Vect}\{a_1, \cdots, a_n\} = \mathrm{Vect}\{e_1, \cdots, e_n\}.$$

Posons

$$e_{n+1} = \frac{f_{n+1}}{\|f_{n+1}\|} \quad \text{avec} \quad f_{n+1} = a_{n+1} - \sum_{j=1}^{n}(a_{n+1}|e_j)e_j.$$

La vérification des relations (4.3) est immédiate. Le théorème est ainsi démontré. □

Théorème 4.2.4. *Soit $\mathcal{H}$ un espace de Hilbert séparable et $(e_n)_{n\in\mathbb{N}}$ une base hilberitienne de $\mathcal{H}$. L'application $\mathcal{I}$ définie par*

$$\left\{\begin{array}{ccc} \mathcal{H} & \longrightarrow & \ell^2(\mathbb{N}) \\ x & \longmapsto & ((x|e_n))_{n\in\mathbb{N}} \end{array}\right.$$

est une bijection linéaire isométrique. Ceci contient en particulier les identités dites de Bessel et Parseval

$$x = \sum_{n\in\mathbb{N}}(x|e_n)e_n \quad \textit{et} \quad \|x\|^2 = \sum_{n\in\mathbb{N}}|(x|e_n)|^2.$$

Démonstration. Son point principal est le fait que l'application $\mathcal{I}$ envoie bien $\mathcal{H}$ dans $\ell^2(\mathbb{N})$. Pour le prouver, posons

$$x_q \stackrel{\text{déf}}{=} \sum_{p\leq q}(x|e_p)e_p.$$

Il est clair que

$$\begin{aligned} (x|x_q) &= \sum_{p\leq q}\overline{(x|e_p)}(x|e_p) \\ &= \sum_{p\leq q}|(x|e_p)|^2 \\ &= (x_q|x_q). \end{aligned}$$

L'inégalité de Cauchy-Schwarz affirme que $\|x_q\|^2 \leq \|x\| \times \|x_q\|$. On en déduit que, pour tout entier q, on a

$$\sum_{p=0}^{q}|(x|e_p)|^2 \leq \|x\|^2.$$

Donc $\mathcal{I}(x)$ appartient à $\ell^2(\mathbb{N})$.

L'injectivité de l'application $\mathcal{I}$ résulte simplement du corollaire 4.2.2 qui affirme que l'orthogonal d'une partie totale est réduit à 0. La surjectivité et le fait que $\mathcal{I}$ est une isométrie est exactement le théorème 4.2.1. Le théorème 4.2.3 est ainsi démontré. □

4.3 Dualité des espaces de Hilbert

Soit $\mathcal{H}$ un espace de Hilbert de dimension finie d, c'est-à-dire un espace euclidien ou hermitien. On sait que l'application δ définie par

$$\delta \left\{ \begin{array}{ccl} \mathcal{H} & \longrightarrow & \mathcal{H}' \\ x & \longmapsto & \delta(x) \,:\, h \mapsto (h|x) \end{array} \right.$$

est une bijection antilinéaire et isométrique de $\mathcal{H}$ sur $\mathcal{H}'$. Pour un espace de Hilbert de dimension infinie, la situation diffère assez peu.

Théorème 4.3.1 (de représentation de Riesz)**.** *Soit $\mathcal{H}$ un espace de Hilbert, on considère l'application δ définie par*

$$\delta \left\{ \begin{array}{ccl} \mathcal{H} & \longrightarrow & \mathcal{H}' \\ x & \longmapsto & \delta(x) \,:\, h \mapsto (h|x) \end{array} \right.$$

C'est une bijection antilinéaire isométrique.

Démonstration. Le fait que $\|\delta(x)\|_{\mathcal{H}'} \leq \|x\|$ (donc en particulier que δ est une forme linéaire continue) résulte de l'inégalité de Cauchy-Schwarz. Le fait que δ est antilinéaire est évident. De plus,

$$\langle \delta(x), \frac{x}{\|x\|} \rangle = \|x\|\,;$$

donc δ est une isométrie. Une isométrie étant injective, il reste à démontrer la surjectivité de δ. Pour ce faire, considérons un élément ℓ de $\mathcal{H}' \setminus \{0\}$. Son noyau est un sous-espace fermé de $\mathcal{H}$, distinct de $\mathcal{H}$. Soit x un vecteur non nul de $(\ker \ell)^\perp$ tel que $\langle \ell, x\rangle = \|x\|^2$. Les deux formes linéaires ℓ et $\delta(x)$ ont même noyau et il existe un h_0 non nul tel que $\langle \ell, h_0\rangle = \langle \delta(x), h_0\rangle$, donc $\delta(x) = \ell$. Le théorème est ainsi démontré. $\square$

Le théorème ci-dessus peut être vu d'une autre manière dans le cas où $\mathbb{K} = \mathbb{R}$. Cette vision nous sera utile au chapitre 5 pour la description du dual des espaces L^p.

Théorème 4.3.2. *Soit $\mathcal{H}$ un espace de Hilbert réel, on considère une forme linéaire ℓ continue sur $\mathcal{H}$ et on définit la fonction*

$$F \left\{ \begin{array}{ccl} \mathcal{H} & \longrightarrow & \mathbb{R} \\ u & \longmapsto & \dfrac{1}{2}\|u\|_{\mathcal{H}}^2 - \langle \ell, u\rangle. \end{array} \right.$$

La fonction F est minorée et il existe un unique v dans $\mathcal{H}$ tel que $F(v) = \inf_{u\in\mathcal{H}} F(u)$ qui vérifie

$$\forall v \in \mathcal{H}\,,\ (v|h)_{\mathcal{H}} = \langle \ell, h\rangle.$$

Démonstration. Comme la forme linéaire ℓ est continue, on peut écrire

$$\begin{aligned} F(u) &\geq \frac{1}{2}\|u\|_{\mathcal{H}}^2 - \|\ell\|_{\mathcal{H}'}\|u\| \\ &\geq \frac{1}{2}(\|u\|_{\mathcal{H}} - \|\ell\|_{\mathcal{H}'})^2 - \frac{1}{2}\|\ell\|_{\mathcal{H}'}^2. \end{aligned}$$

Posons $m \stackrel{\text{déf}}{=} \inf_{u\in\mathcal{H}} F(u)$. On considère une suite minimisante $(u_n)_{n\in\mathbb{N}}$ c'est-à-dire une suite d'éléments de $\mathcal{H}$ telle

$$F(u_n) = m + \varepsilon_n \quad \text{avec} \quad \lim_{n\to\infty} \varepsilon_n = 0. \tag{4.4}$$

Grâce à la relation de la médiane, écrivons que

$$\Big\|\frac{u_n-u_m}{2}\Big\|_{\mathcal{H}}^2=\frac{1}{2}\|u_n\|_{\mathcal{H}}^2+\frac{1}{2}\|u_m\|_{\mathcal{H}}^2-\Big\|\frac{u_n+u_m}{2}\Big\|_{\mathcal{H}}^2.$$

Comme ℓ est une application linéaire, on a

$$\begin{aligned}\Big\|\frac{u_n-u_m}{2}\Big\|_{\mathcal{H}}^2 &= \frac{1}{2}\|u_n\|_{\mathcal{H}}^2-\langle\ell,u_n\rangle+\frac{1}{2}\|u_m\|_{\mathcal{H}}^2-\langle\ell,u_m\rangle\\ &\qquad -2\Big(\frac{1}{2}\Big\|\frac{u_n+u_m}{2}\Big\|_{\mathcal{H}}^2-\langle\ell,\frac{u_n+u_m}{2}\rangle\Big).\end{aligned}$$

Par définition de m et d'après (4.4), on a

$$\Big\|\frac{u_n-u_m}{2}\Big\|_{\mathcal{H}}^2\leq\varepsilon_n+\varepsilon_m.$$

Ainsi donc la suite $(u_n)_{n\in\mathbb{N}}$ est de Cauchy et donc converge vers un élément v de $\mathcal{H}$. La fonction F étant continue, on a $F(v)=m$. Comme on a, pour tout h de $\mathcal{H}$, et pour tout réel t,

$$m\leq F(v+th)=\frac{1}{2}\|v\|_{\mathcal{H}}^2+t(v|h)+\frac{1}{2}t^2\|h\|_{\mathcal{H}}^2-t\langle\ell,h\rangle,$$

on en déduit que

$$\forall h\in\mathcal{H}\,,\ \forall t\in\mathbb{R}\,,\ t(v|h)+\frac{1}{2}t^2\|h\|_{\mathcal{H}}^2-t\langle\ell,h\rangle\geq 0$$

ce qui implique que

$$\forall h\in\mathcal{H}\,,\ (v|h)=\langle\ell,h\rangle.$$

Un tel v est bien sûr unique car si $(v_1-v_2|h)=0$ pour tout h, on obtient l'unicité e prenant $h=v_1-v_2$. □

On peut en déduire les deux corollaires ci-dessous.

Définition 4.3.1. *Soient $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments d'un espace de Hilbert $\mathcal{H}$ et x un élément de $\mathcal{H}$. On dit que la suite $(x_n)_{n\in\mathbb{N}}$ converge faiblement vers x et l'on note $\operatorname*{limf}_{n\to\infty} x_n=x$ ou bien $(x_n)_{n\in\mathbb{N}}\rightharpoonup x$ si et seulement si*

$$\forall h\in\mathcal{H}\,,\ \lim_{n\to\infty}(h|x_n)=(h|x).$$

La définition ci-dessus peut se formuler en disant que

$$\operatorname*{limf}_{n\to\infty} x_n=x\Longleftrightarrow\operatorname*{limf}_{n\to\infty}\star\,\delta(x_n)=\delta(x).$$

Le théorème suivant dit, entre autre chose, pourquoi cette convergence est dite faible.

Théorème 4.3.3. *Soient $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments d'un espace de Hilbert $\mathcal{H}$ et x un élément de $\mathcal{H}$. On a alors*

$$\textit{si}\quad\lim_{n\to\infty}\|x_n-x\|=0\quad\textit{alors}\quad\operatorname*{limf}_{n\to\infty}x_n=x\,;\tag{4.5}$$

$$\textit{si}\quad\operatorname*{limf}_{n\to\infty}x_n=x\ \textit{et}\ \lim_{n\to\infty}\|x_n\|=\|x\|,\quad\textit{alors}\quad\lim_{n\to\infty}\|x_n-x\|=0.\tag{4.6}$$

De plus, si $(x_n)_{n\in\mathbb{N}}\rightharpoonup x$, alors la suite $(x_n)_{n\in\mathbb{N}}$ est bornée.

Démonstration. Le premier point du théorème résulte simplement du fait que

$$|(h|x_n) - (h|x)| \leq \|h\| \times \|x_n - x\|.$$

Quant au second point, il suffit d'écrire que

$$\|x_n - x\|^2 = \|x_n\|^2 - 2\,\Re e(x|x_n) + \|x\|^2;$$

Comme la suite $(x_n)_{n\in\mathbb{N}}$ tend faiblement vers x, on a

$$-2\lim_{n\to\infty} \Re e(x|x_n) = -2\|x\|^2.$$

Le deuxième point est ainsi démontré. Le troisième point est un corollaire du théorème de Banach-Steinhaus. En effet, l'hypothèse de convergence faible signifie que, pour tout y de $\mathcal{H}$, on a

$$\lim_{n\to\infty} \delta(x_n)(y) = \lim_{n\to\infty} (y|x_n) = (y|x).$$

Le théorème 3.3.2 de Banach-Steinhaus dit que la suite $(\delta(x_n))_{n\in\mathbb{N}}$ est une suite bornée de $\mathcal{H}'$. D'après le théorème 4.3.1 de représentation de Riesz, ceci signifie que la suite $(x_n)_{n\in\mathbb{N}}$ est bornée. D'où le théorème.

Exemple Soit $\mathcal{H}$ un espace de Hilbert de dimension infinie et séparable. Considérons une base hilbertienne $(e_n)_{n\in\mathbb{N}}$ de $\mathcal{H}$. D'après le théorème 4.3.1 de représentation de Riesz, pour tout x de $\mathcal{H}$, la suite $((x|e_n))_{n\in\mathbb{N}}$ appartient à $\ell^2(\mathbb{N})$. Donc elle tend vers 0. Voilà donc un exemple de suite faiblement convergente vers 0 qui ne converge pas au sens de la norme vers 0 puisque chaque e_n est de norme 1.

Proposition 4.3.1. *Soient $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ deux suites d'éléments de $\mathcal{H}$ telles que*

$$\lim_{n\to\infty} x_n = x \quad et \quad \operatorname{limf}_{n\to\infty} y_n = y.$$

Alors, on a $\lim_{n\to\infty} (x_n|y_n) = (x|y)$.

Démonstration. Il suffit d'écrire que

$$\begin{aligned} |(x_n|y_n) - (x|y)| &\leq |(x_n - x|y_n)| + |(x|y_n - y)| \\ &\leq \|x_n - x\| \times \|y_n\| + |(x|y_n - y)|. \end{aligned}$$

Le théorème 4.3.3 affirme que la suite $(y_n)_{n\in\mathbb{N}}$ est bornée. Donc, on a

$$|(x_n|y_n) - (x|y)| \leq C\|x_n - x\| + |(x|y_n - y)|,$$

D'où la proposition.

Exercice 4.3.1. *Trouvez deux suites $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{H}$ telles que*

$$\operatorname{limf}_{n\to\infty} x_n = x\,,\ \operatorname{limf}_{n\to\infty} y_n = y \quad et \quad \lim_{n\to\infty} (x_n|y_n) \neq (x|y).$$

Exercice 4.3.2. *Démontrez que, dans un espace de Hilbert $\mathcal{H}$, si une suite $(x_n)_{n\in\mathbb{N}}$ tend faiblement vers x, alors,*

$$\|x\| \leq \liminf_{n\to\infty} \|x_n\|$$

Le théorème suivant est une théorème de compacité faible de la boule unité d'un espace de Hilbert. Lorsque celui-ci n'est pas de dimension finie, on sait, d'après le théorème 2.3.1 que la boule unité n'est pas compacte au sens de la topologie définie à l'aide de la norme. Cependant, on a le théorème ci-après qui n'est qu'une traduction du théorème 3.3.4.

Théorème 4.3.4 (de compacité faible)**.** *Soit $(x_n)_{n\in\mathbb{N}}$ une suite bornée d'un espace de Hilbert séparable $\mathcal{H}$, on peut en extraire une sous-suite faiblement convergente.*

4.4 Adjoint d'un opérateur et opérateurs auto-adjoints

La notion d'adjoint est bien connue en dimension finie. Étant donné un produit scalaire sur un espace $\mathcal{H}$ de dimension finie d, l'adjoint $A^\star$ d'un opérateur est défini par la relation

$$(Ax|y) = (x|A^\star y).$$

Il est connu que, si $(A_{i,j})_{1\leq i,j\leq d}$ est la matrice de A dans une base orthonormée, alors la matrice de $A^\star$ dans cette même base est donnée par $A^\star_{i,j} = \overline{A}_{j,i}$. De plus, c'est un théorème classique d'algèbre linéaire que si u est autoadjoint (i.e. $A = A^\star$), l'opérateur A est diagonalisable dans une base orthonormée.

Nous allons étudier une généralisation de ces concepts et de ces résultats au cas des espaces de Hilbert de dimension infinie.

Théorème 4.4.1. *Soit A un opérateur linéaire continu sur un espace de Hilbert $\mathcal{H}$. Il existe un unique opérateur linéaire $A^\star$ continu sur $\mathcal{H}$ tel que*

$$\forall (x,y) \in \mathcal{H}\times\mathcal{H}\,,\ (Ax|y) = (x|A^\star y).$$

De plus, l'application qui, à l'opérateur A, associe $A^\star$, est une application antilinéaire isométrique de $\mathcal{L}(\mathcal{H})$ dans lui-même.

On déduit du théorème ci-dessus la définition suivante.

Définition 4.4.1. *On appelle l'opérateur $A^\star$ ci-dessus l'opérateur adjoint de A. On dit qu'un élément A de $\mathcal{L}(\mathcal{H})$ est autoadjoint si et seulement si $A = A^\star$.*

Démonstration du théorème 4.4.1 L'unicité de l'opérateur $A^\star$ résulte immédiatement du fait que $\mathcal{H}^\perp = \{0\}$. Soit $\mathcal{L}_A$ l'application définie par

$$\mathcal{L}_A \left\{ \begin{array}{rcl} \mathcal{H} & \longrightarrow & \mathcal{H}' \\ y & \longmapsto & \mathcal{L}_A(y)\ :\ x \longmapsto (Ax|y). \end{array} \right.$$

Il est visible que l'application $\mathcal{L}_A$ est une application antilinéaire continue de $\mathcal{H}$ dans $\mathcal{H}'$. Posons $A^\star \stackrel{\text{déf}}{=} \delta^{-1}\circ\mathcal{L}_A$. Par définition de $A^\star$, on a

$$(x|A^\star y) = (\mathcal{L}_A y|x) = (Ax|y).$$

Donc l'opérateur A vérifie les propriétés voulues. De plus, on a

$$\begin{aligned} \|A\| &= \sup_{\|x\|\leq 1, \|y\|\leq 1} |(Ax|y)| \\ &= \sup_{\|x\|\leq 1, \|y\|\leq 1} |(x|A^\star y)| \\ &= \|A^\star\|. \end{aligned}$$

Le théorème 4.4.1 est donc démontré.

Exercice 4.4.1. *Soit $\mathcal{H}$ un espace de Hilbert réel et Ω un ouvert de $\mathcal{H}$. On définit l'opérateur "gradient" par*

$$\mathrm{grad} \left\{ \begin{array}{rcl} C^1(\Omega;\mathbb{R}) & \longrightarrow & \mathcal{C}(\Omega,\mathcal{H}) \\ g & \longmapsto & \delta^{-1}Dg \end{array} \right.$$

Soient g une application C^1 d'un ouvert Ω d'un espace de Hilbert réel $\mathcal{H}$ dans un ouvert Ω' de $\mathcal{H}$ et f une application C^1 de Ω' dans $\mathbb{R}$. Calculez $\mathrm{grad}\,(f\circ g)$.

Proposition 4.4.1. *Soit A une application linéaire continue sur un espace de Hilbert $\mathcal{H}$. On a les propriétés suivantes. L'opération * est une involution, c'est-à-dire que $A = A^{\star\star}$. On a aussi*

$$(\ker A)^\perp = \overline{\operatorname{Im} A^\star} \quad et \quad \ker A = (\operatorname{Im} A^\star)^\perp. \tag{4.7}$$

Enfin, si la suite $(x_n)_{n\in\mathbb{N}}$ converge faiblement vers x, alors la suite $(Ax_n)_{n\in\mathbb{N}}$ converge faiblement vers Ax.

Démonstration. Le premier point résulte simplement du fait que si l'on a $(Ax|y) = (x|A^\star y)$, alors $(A^\star y|x) = (y|Ax)$. Pour démontrer les relations (4.7), écrivons que

$$\forall y \in \mathcal{H}\,,\ (Ax|y) = 0 \Longleftrightarrow \forall y \in \mathcal{H}\,,\ (x|A^\star y) = 0.$$

La traduction ensembliste de l'équivalence ci-dessus est exactement l'égalité entre les deux ensembles $\ker A$ et $(\operatorname{Im} A^\star)^\perp$. L'autre relation s'en déduit par passage à l'orthogonal.

Soit $(x_n)_{n\in\mathbb{N}}$ une suite faiblement convergente vers x. Pour tout y de $\mathcal{H}$, on a

$$(y|Ax_n) = (A^\star y|x_n).$$

La convergence faible de la suite $(x_n)_{n\in\mathbb{N}}$ entraîne celle de la suite $(Ax_n)_{n\in\mathbb{N}}$. □

Proposition 4.4.2. *Soit A un opérateur autoadjoint de $\mathcal{L}(\mathcal{H})$. On a alors*

$$\|A\|_{\mathcal{L}(\mathcal{H})} = \sup_{\|u\|=1} |(Au|u)|.$$

Démonstration. Tout d'abord, nous avons

$$\|A\|_{\mathcal{L}(\mathcal{H})} = \sup_{\substack{\|u\|=1\\ \|v\|=1}} \Re e(Au|v).$$

Il est clair que $M \stackrel{\text{déf}}{=} \sup_{\|u\|=1} |(Au|u)|$ est inférieur ou égal à $\|A\|_{\mathcal{L}(\mathcal{H})}$. Soient u et v deux vecteurs de $\mathcal{H}$ de norme 1. On a

$$\begin{aligned} \Re e(Au|v) &= \frac{1}{4}\big((A(u+v)|u+v) - (A(u-v)|u-v)\big)\\ &\leq \frac{M}{4}(\|u+v\|^2 + \|u-v\|^2). \end{aligned}$$

Comme $\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2)$ et comme les vecteurs u et v sont de norme plus petite de 1, on trouve que

$$\Re e(Au|v) \leq M$$

ce qui démontre la proposition. □

Nous souhaitons maintenant étudier une généralisation du théorème de diagonalisation des opérateurs auto-adjoints en dimension finie. À notre niveau, ceci nécessite un hypothèse supplémentaire sur l'opérateur, celle de compacité. Introduisons la notion d'opérateurs compacts. Cette notion peut être définie pour une application linéaire continue entre espaces de Banach.

Définition 4.4.2. *Soient E et F deux espaces de Banach. Un élément u de $\mathcal{L}(E, F)$ est dit compact si et seulement si pour toute partie bornée A, $u(A)$ est d'adhérence compacte.*

Exemple Si $u(E)$ est un espace vectoriel de dimension finie, alors l'opérateur u est compact. En effet, si A est une partie bornée de A, l'image de A est une partie bornée de l'espace vectorielde dimension finie $u(E)$. C'est donc une partie d'adhérence compacte.

Exercice 4.4.2. *Soit E l'espace vectoriel des fonctions f de $[0,1]$ à valeurs dans $\mathbb{R}$ telles que*

$$\sup_{x\neq x'} \frac{|f(x)-f(x')|}{|x-x'|^{\frac{1}{2}}} < \infty$$

1) Démontrer que

$$\|f\|_E \stackrel{\text{déf}}{=} |f(0)| + \sup_{x\neq x'} \frac{|f(x)-f(x')|}{|x-x'|^{\frac{1}{2}}}$$

définit une norme sur E et que $(E, \|\cdot\|_E)$ est un espace de Banach.

2) Démontrer que l'application

$$\iota \left\{ \begin{array}{ccc} E & \longrightarrow & \mathcal{C}([0,1],\mathbb{R}) \\ f & \longmapsto & f \end{array} \right.$$

est compacte.

Exercice 4.4.3. *Soit $(A(n,m))_{(n,m)\in\mathbb{N}^2}$ une suite de réels telle que $\displaystyle\sum_{(n,m)\in\mathbb{N}^2} A^2_{n,m}$ soit finie.*

1) Démontrer que

$$(\mathcal{A}x)(n) \stackrel{\text{déf}}{=} \sum_{m\in\mathbb{N}} A(n,m)x(m)$$

définit un opérateur continu de $\ell^2(\mathbb{N})$ dans lui-même et que $\displaystyle\|\mathcal{A}\|_{\mathcal{L}(\ell^2)} \leq \Big(\sum_{(n,m)\in\mathbb{N}^2} A^2_{n,m}\Big)^{\frac{1}{2}}$.

2) Soit $\mathcal{A}_N$ l'opérateur défini par $(\mathcal{A}_N x)(n) = \mathbf{1}_{\{0,\cdots,N\}}(n)(\mathcal{A}x)(n)$. Démontrer que

$$\lim_{N\to\infty} \|\mathcal{A}_N - \mathcal{A}\|_{\mathcal{L}(\ell^2)} = 0.$$

3) En déduire que $\mathcal{A}$ est compact.

Énonçons maintenant le théorème de diagonalisation des opérateurs auto-adjoints compacts sur un espace de Hilbert.

Théorème 4.4.2. *Soit A un opérateur compact autoadjoint d'un espace de Hilbert $\mathcal{H}$ de dimension infinie séparable. Il existe une suite $(\lambda_j)_{j\in\mathbb{N}}$ de nombres réels telle que la suite $(|\lambda_j|)_{j\in\mathbb{N}}$ soit une suite décroissante tendant vers 0 et telle que*

- *Pour tout j tel que λ_j soit non nul, λ_j est une valeur propre de A et $E_j \stackrel{\text{déf}}{=} \ker A - \lambda_j \,\mathrm{Id}$ est de dimension finie. De plus, si λ_j et λ'_j sont différents, les sous espaces propres sont orthogonaux.*
- *Si $E \stackrel{\text{déf}}{=} \mathrm{Vect} \displaystyle\bigcup_{\substack{j\in\mathbb{N}\\ \lambda_j\neq 0}} E_j$, alors $\ker A = E^\perp$.*
- *On a $\|A\|_{\mathcal{L}(\mathcal{H})} = \max_{j\in\mathbb{N}} |\lambda_j|$.*

Avant de nous lancer dans la démonstration de ce théorème, nous allons rappeler (et démontrer) le théorème de diagonalisation des iopérateurs auto-adjoints dans un espace euclidien (ou hermitien) de dimension finie.

Théorème 4.4.3. *Soit A un opérateur auto-adjoint dans un espace de Hilbert de dimension finie N. Il existe une suite finie $(\lambda_j)_{1\leq j\leq N}$ et une base orthonormée $(e_j)_{1\leq j\leq N}$ telle que, pour tout j, e_j soit un vecteur propre associé à la valeur propre λ_j.*

Démonstration. Considérons la forme quadratique (ou hermitienne) associée à A, c'est-à-dire

$$Q_A \quad \begin{cases} \mathcal{H} & \longrightarrow \quad \mathbb{R} \\ x & \longmapsto \quad (Ax|x) \end{cases}$$

C'est une fonction continue sur $\mathcal{H}$. Désignons par $\mathbb{S}_{\mathcal{H}}$ la sphère unité de $\mathcal{H}$. Vu que nous sommes en dimension finie, le corollaire 1.3.2 page 24 du théorème de Heine (voir le théorème 1.3.3 page 23) assure l'existence d'un point x_M de $\mathbb{S}_{\mathcal{H}}$ telle que

$$|(Ax_M|x_M)| = M_0 \stackrel{\text{déf}}{=} \sup_{x\in\mathbb{S}_{\mathcal{H}}} |(Ax|x)| \tag{4.8}$$

Il s'agit de démontrer que x_M est un vecteur propre de A associée à la valeur propre $\pm M_0$. Ce fait est général que la dimension soit finie ou infinie. Plus précisement, on a le lemme suivant.

Lemme 4.4.1. *Soit $\mathcal{H}$ un espace de Hilbert (de dimension finie ou infinie). Soit x_0 un vecteur de $\mathbb{S}_{\mathcal{H}}$ telle que*

$$|(Ax_0|x_0)| = M_0 \stackrel{\text{déf}}{=} \sup_{x\in\mathbb{S}_{\mathcal{H}}} |(Ax|x)|$$

Alors x_0 est vecteur propre associée à la valeur propre M_0 ou $-M_0$.

Démonstration. Quitte à changer A en $-A$, on peut supposer que $(Ax_0|x_0) = M_0$. Observons tout d'abord que, pour tout y non nul dans $\mathcal{H}$, on a

$$\Big(A\Big(\frac{y}{\|y\|}\Big)\Big|\frac{y}{\|y\|}\Big) \leq M_0$$

ce qui implique que

$$\forall y \in \mathcal{H}\,,\ F(y) \stackrel{\text{déf}}{=} M_0\|y|^2 - (Ay|y) \geq 0. \tag{4.9}$$

Le point x_0 pour lequel $F(x_0)$ est nul est donc un minimum de la fonction F. On peut invoquer le théorème disant que la différentielle de F en x_0 est nulle. On peut aussi redémontrer l'assertion (4.9) directement en observant que

$$\forall \lambda \in \mathbb{R}\,,\ \forall h \in \mathcal{H}\,,\ M_0\|x_0 + \lambda h\|^2 - \big(A(x_0 + \lambda h|x_0 + \lambda h\big) \geq 0$$

Comme $F(x_0) = 0$, on trouve en développant que

$$\forall \lambda \in \mathbb{R}\,,\ \forall h \in \mathcal{H}\,,\ 2\lambda \Re e\big(M_0x_0 - Ax_0\big|h\big) + \lambda^2\big(M_0\|h\|^2 - A(h|h)\big) \geq 0.$$

Le discriminant d'un polynôme positif étant négatif ou nul, il en résulte que

$$\forall h \in \mathcal{H}\,,\ \Re e\big(M_0x_0 - Ax_0\big|h\big) = 0$$

Dans le cas hermitien, en appliquant la relation ci-dessus avec ih, on trouve que, pour tout h dans $\mathcal{H}$, $\big(M_0x_0 - Ax_0\big|h\big) = 0$ et le lemme est ainsi démontré. $\square$

Poursuite de la démonstration du théorème 4.4.3 Nous allons maintenant démontrer un lemme d'algèbre linéaire très simple qui lui aussi est valable indépendemment du fait que la dimension doit finie ou non.

Lemme 4.4.2. *Soit A un opérateur auto-adjoint sur un espace de Hilbert $\mathcal{H}$. Deux vecteurs propres de A associés à des valeurs propres différentes sont orthogonaux.*

Démonstration. Soient u et v deux vecteurs propres associées à deux valeurs propres différentes λ et λ'. On a

$$\lambda(u|v) = (Au|v) = (Av|u) = \lambda'(u|v).$$

Ainsi donc $(\lambda - \lambda')(u|v) = 0$ ce qui assure le lemme. $\square$

Poursuite de la démonstration du théorème 4.4.3 Définissons

$$E_0^{\pm} \stackrel{\text{déf}}{=} \ker(A \pm M_0 \,\mathrm{Id}) \quad \text{et} \quad \mathcal{H}_1 \stackrel{\text{déf}}{=} (E_0^+ \oplus E_0^-)^\perp = (E_0^+)^\perp \cap (E_0^-)^\perp. \tag{4.10}$$

Le lemme 4.7 assure en particulier que

$$(A \pm M_0 \,\mathrm{Id})(\mathcal{H}_1) \subset (E_0^{\pm})^\perp \quad \text{et donc que} \quad A(\mathcal{H}_1) \subset \mathcal{H}_1 \tag{4.11}$$

On applique le même procédé à $\mathcal{H}_1$ et l'on trouve une ou deux valeurs propres $\pm M_1$ ou la valeur propre 0. Comme la dimension est finie, la suite des espaces $\mathcal{H}_j$ ainsi construits est finie et le théorème 4.4.3 est démontré. □

Démonstration du théorème 4.4.2 Elle suit le même schéma avec quelques difficultés supplémentaires dues à la dimension infinie. Trois points (classés par ordre croissant de difficulté) méritent l'attention :

— le fait que les espaces propres associés aux valeurs propres non nulles soient de dimension finie ;

— le fait que la suite des valeurs propres tende vers 0 ;

— l'existence d'un point x_M réalisant la borne supérieure M_0.

Démontrons le premier point. Pour cela, nous allons démontrer que la boule unité de l'espace propre E_j est compacte ce qui, d'après le théorème 2.3.1 page 38 de Riesz assurera que la dimension de E_j est finie.

Soit $(x_n)_{n\in\mathbb{N}}$ une suite bornée de E_j. L'opérateur A étant compact, on peut en extraire une sous-suite $(x_{\varphi(n)})_{n\in\mathbb{N}}$ telle que la suite $(Ax_{\varphi(n)})_{n\in\mathbb{N}}$ converge. Mais $Ax_{\varphi(n)} = \lambda_j x_{\varphi(n)}$ et λ_j est non nulle. La boule unité de E_j est donc compacte.

Pour démontrer le deuxième point, considérons une suite infinie $(\lambda_j)_{j\in\mathbb{N}}$ de valeurs propres deux à deux distinctes. Soit $(e_j)_{j\in\mathbb{N}}$ une suite de vecteurs de norme 1 qui soient pour chaque j un vecteur propreassocié à la valeur propre λ_j. La suite $(e_j)_{j\in\mathbb{N}}$ tend faiblement vers 0 (Exercice : démontrez-le !). Quitte à extraire, on peut supposer que la suite $(Ae_j)_{j\in\mathbb{N}}$ tend fortement vers 0. Mais comme $Ae_j = \lambda_j e_j$, on en déduit que

$$\lim_{j\to\infty} |\lambda_j|\,\|e_j\| = 0.$$

Les vecteurs e_j étant de norme 1, la suite $(\lambda_j)_{j\in\mathbb{N}}$ tend vers 0.

Il s'agit maintenant de construire la suite des λ_j. Le point crucial est le lemme suivant.

Lemme 4.4.3. *Soient $\mathcal{H}$ un espace de Hilbert séparable de dimension infinie et A un opérateur autoadjoint compact sur $\mathcal{H}$. Il existe un vecteur x_M de la sphère unité de $\mathcal{H}$ tel que*

$$|(Ax_M|x_M)| = M_0 \stackrel{\text{def}}{=} \sup_{\|x\|=1} |(Ax|x)|.$$

De plus $Ax_M = \pm M_0 x_M$.

Démonstration. Si $M_0 = 0$, d'après la proposition 4.4.2, $A = 0$ et il n'y a rien à démontrer. Supposons que M_0 soit strictement positif. Comme $(Ax|x)$ est réel, quitte à changer A en $-A$, on peut supposer que

$$M_0 = \sup_{\|x\|=1} (Ax|x).$$

Par définition, il existe une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de la boule unité B telle que

$$\lim_{n\to\infty} (Ax_n|x_n) = M_0$$

L'opérateur A étant compact, on peut, quitte à extraire une sous-suite, dire qu'il existe un y dans $\mathcal{H}$ tel que

$$\lim_{n\to\infty} Ax_n = y.$$

À nouveau après extraction, on peut supposer, d'après le théorème 4.3.4, qu'il existe un élément x de $\mathcal{H}$ tel que

$$\operatorname*{limf}_{n\to\infty} x_n = x.$$

Démontrons que $Ax = y$. L'opérateur A étant autoadjoint, on a, pour tout z dans $\mathcal{H}$,

$$\lim_{n\to\infty} (x_n|Az) = (x|Az) = (Ax|z).$$

De plus, il est clair que

$$\lim_{n\to\infty} (Ax_n|z) = (y|z).$$

Il en résulte que, pour tout z dans $\mathcal{H}$, on a $(y|z) = (Ax|z)$, ce qui implique que $y = Ax$. Mais, d'après la proposition 4.3.1, on a

$$\lim_{n\to\infty} (Ax_n|x_n) = (Ax|x).$$

Le maximum M_0 est donc atteint en un point x qui bien sûr est différent de 0 puisque M_0 est strictement positif. De plus, si l'on avait $\|x\| < 1$, alors on aurait

$$\left(A\frac{x}{\|x\|}\middle|\frac{x}{\|x\|}\right) = \frac{M_0}{\|x\|^2} > M_0,$$

ce qui contredit la maximalité de M_0 car M_0 est strictement positif. Donc $\|x\| = 1$, ce qui implique en particulier que la convergence de $(x_n)_{n\in\mathbb{N}}$ vers x est forte.

On peut maintenant appliquer le lemme 4.4.3 et ainsi travailler dans l'espace

$$\mathcal{H}_1 \stackrel{\text{déf}}{=} \Big(\ker(A - M_0\,\mathrm{Id}) + \ker(A + M_0\,\mathrm{Id})\Big)^\perp = \ker(A - M_0\,\mathrm{Id})^\perp \cap \ker(A + M_0\,\mathrm{Id})^\perp.$$

Démontrons que $\mathcal{H}_1$ est stable par A. Pour ce faire, il suffit de démontrer le lemme suivant.

Lemme 4.4.4. *Soit B un élément de $\mathcal{L}(\mathcal{H})$. Alors $B^\star$ laisse stable* $(\ker B)^\perp$.

Démonstration. Soit $x \in (\ker B)^\perp$ et $y \in \ker B$. On a

$$(y|B^\star x) = (By|x) = 0$$

Ainsi donc $B^\star x$ est dans l'orthogonal de $\ker B$. D'où le lemme. □

Poursuite de la démonstration du théorème 4.4.2 Si l'on applique le lemme ci-dessus à $B = A \pm M_0\,\mathrm{Id}$, on trouve, comme A est autoadjoint, que

$$(A \pm M_0\,\mathrm{Id})\big(\ker(A \pm M_0\,\mathrm{Id})\big)^\perp \subset \big(\ker(A \pm M_0\,\mathrm{Id})\big)^\perp.$$

Ceci implique immédiatement que

$$A\big(\ker(A \pm M_0\,\mathrm{Id})\big)^\perp \subset \big(\ker(A \pm M_0\,\mathrm{Id})\big)^\perp$$

et donc que $A\mathcal{H}_1 \subset \mathcal{H}_1$. Étudions maintenant l'opérateur A restreint à l'espace de Hilbert $\mathcal{H}_1$. Cet opérateur $A_{|\mathcal{H}_1}$ est un opérateur autoadjoint compact de $\mathcal{H}_1$ (Exercice : vérifiez-le). Posons

$$M_1 \stackrel{\text{déf}}{=} \sup_{\substack{\|x\|=1 \\ x\in\mathcal{H}_1}} (Ax|x).$$

D'après l'étude précédente, si $M_1 \neq 0$, le maximum est atteint en au moins un point de norme 1 de $(\ker(A - M_1 \operatorname{Id}) + \ker(A + M_1 \operatorname{Id}))^\perp$. Ceci implique que $M_1 < M_0$. On itère le processus de la manière suivante. Posons .

$$\mathcal{H}_j \stackrel{\text{déf}}{=} \big(E_0 + \cdots + E_{j-1}\big)^\perp \quad \text{et} \quad \sup_{\substack{\|u\|=1 \\ u\in\mathcal{H}_{j+1}}} |(Au|u)| = M_{j+1}.$$

Si $M_{j+1} = 0$, la proposition 4.4.2 implique que $\mathcal{H}_{j+1} = \ker A$ et le processus s'arrête. Si M_{j+1} est srtictement positif, on poursuit l'algorithme.

Si la suite des $(M_j)_{j\in\mathbb{N}}$ est une suite de réels strictement positifs, on considère

$$E^\perp = \bigcap_{j\in\mathbb{N}} \mathcal{H}_j.$$

Ainsi donc

$$\forall j \in \mathbb{N}\,,\ \sup_{\substack{\|u\|=1 \\ u\in E^\perp}} |(Au|u)| \leq |\lambda_j|$$

et donc

$$\sup_{\substack{\|u\|=1 \\ u\in E^\perp}} |(Au|u)| = 0$$

ce qui implique d'après la proposition 4.4.2 que $A_{|E^\perp} = 0$. Le théorème 4.4.2 est complètement démontré. $\square$

Chapitre 5

Espaces L^p

Introduction

Ce chapitre est dévolu à l'étude des fonctions de puissance p ième intégrable définies sur un sensemble X muni d'une tribu $\mathcal{B}$ et d'une mesure positive μ sur cette tribu. Deux exemples fondamentaux sont étudiés.

Tout d'abord, le cas où $X = \mathbb{N}$ et où la tribu $\mathcal{B}$ est l'ensemble des parties de $\mathbb{N}$ et μ la mesure de décompte, c'est-à-dire que la mesure d'une partie de $\mathbb{N}$ est le nombre (fini ou infini) d'éléments de cette partie. Les espaces de fonctions de puissance p ième intégrables sont alors les espaces $\ell^p(\mathbb{N})$ étudiés aux chapitres 2 et 3.

L'autre exemple important est le cas où X est l'espace $\mathbb{R}^d$ ou une partie de l'espace $\mathbb{R}^d$ muni de la tribu borélienne et de la mesure de Lebesgue.

La première section est constituée du rappel sans démonstration d'un certain nombre d'énoncés fondamentaux de la théorie de l'intégration qui seront d'usage constant dans ce chapitre. Ces énoncés doivent absolument être connus.

Dans le deuxième section, on démontre que les espaces (de classes de) fonctions de puissance p ième intégrable sont des espaces de Banach. Le fait que les fonctions soient définies modulo des ensembles négligeables constitue une difficulté supplémentaire par rapport au cas des espaces $\ell^p(\mathbb{N})$ traité au chapitre 2 (voir le théorème 2.1.1 page 31 et sa démonstration). L'inégalité de Hölder joue un rôle important dans la démonstration. Dans cette section, nous verrons également un critère d'appartenance aux espaces de fonctions de puissance p ième inégrable qui s'exprime au travers d'une inégalité sur une collection de moyenne pondérées. C'est le fondamental lemme 5.2.2. Cette idée qu'une collection de moyennes pondérées caractérise une fonction est à la base de la notion de dérivée faible qui sera introduite au chapitre 6 et plus généralement de la notion de distributions (voir le chapitre 8).

La troisième section est consacrée à la démonstration du théorème de densité des fonctions continues à support compact dans les espaces de fonctions de puissance p ième intégrable dans le cas où X est un ouvert de $\mathbb{R}^d$.

La quatrième section est consacrée à la définition du produit de convolution de deux fonctions définies sur $\mathbb{R}^d$. Cette opération est d'un usage constant en analyse et fournit en particulier des approximations "explicites" des fonctions de puissance p ième intégrable par des fonctions indéfiniment différentiables à support compact.

La cinquième et dernière section de ce chapitre traite de l'identification du dual des fonctions de puissance p ième intégrable avec les fonctions de puissance p' intégrable lorsque p est réel et p' est l'exposant conjugé de p. Ceci a déjà été étébli au chapitre 3 (voir le théorème 3.2.1 page 54). Nous ne le présentons ici que dans le cas où p est dans l'intervalle $[1, 2]$.

5.1 Rappel sur la théorie de la mesure et définition des espaces L^p

Dans cette section, nous allons définir les espaces L^p et rappeler les principaux théorèmes de base de la théorie de la mesure. Soit X muni d'une tribu $\mathcal{B}$ et μ une mesure positive sur $\mathcal{B}$. On suppose, comme il est classique, que la mesure est σ-finie, c'est-à-dire que X est une réunion dénombrable d'ensembles de mesure finie.

Définition 5.1.1. *Soit p dans l'intervalle $[1, +\infty[$, on désigne par $L^p(X, d\mu)$ l'espace des (classes d'équivalence modulo la relation de coïncidence presque partout) des fonctions mesurables telles que*

$$\int_X |f(x)|^p d\mu < \infty.$$

Si $p = \infty$, on désigne par $L^p(X, d\mu)$ l'espace des (classes d'équivalence modulo la relation de coïncidence presque partout) des fonctions mesurables telles que

$$\|f\|_{L^\infty} \stackrel{\text{déf}}{=} \sup\big\{\lambda \,/\, \mu\{x \,/\, |f(x)| > \lambda\} > 0\big\} = \inf\big\{M \,/\, \mu\{x \,/\, |f(x)| > M\}\big\} = 0.$$

Démontrons l'égalité entre les deux quantités qui définissent $\|f\|_{L^\infty}$. Remarquons tout d'abord que pour tout couple (λ, M) tel que $\mu\{x \,/\, |f(x)| > M\} = 0$ et $\mu\{x \,/\, |f(x)| > \lambda\} > 0$ est tel que λ et strictement inférieur à M. Ainsi donc

$$\sup\big\{\lambda \,/\, \mu\{x \,/\, |f(x)| > \lambda\} > 0\big\} \leq \inf\big\{M \,/\, \mu\{x \,/\, |f(x)| > M\}.$$

Soit M_1 un nombre réel strictement supérieur à $\sup\big\{\lambda \,/\, \mu\{x \,/\, |f(x)| > \lambda\} > 0\big\}$. Par définition de la borne supérieure, on a $\mu\{x \,/\, |f(x)| > M_1\} = 0$. Ainsi donc

$$\inf\big\{M \,/\, \mu\{x \,/\, |f(x)| > M\} \leq \sup\big\{\lambda \,/\, \mu\{x \,/\, |f(x)| > \lambda\} > 0\big\}.$$

Nous commencerons par rappeler les principaux énoncés qui fondent la théorie de la mesure.

Théorème 5.1.1 (de convergence monotone)**.** *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables positives sur X. On suppose que, pour tout x, la suite $(f_n(x))_{n\in\mathbb{N}}$ est croissante. Soit f la fonction définie par*

$$f \left\{ \begin{array}{ccc} X & \longrightarrow & \mathbb{R}^+ \cup \{+\infty\} \\ x & \longmapsto & \displaystyle\lim_{n\to\infty} f_n(x). \end{array} \right.$$

Alors on a, au sens des égalités de $\mathbb{R}^+ \cup \{+\infty\}$,

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu$$

Lemme 5.1.1 (de Fatou)**.** *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables de X dans $[0, +\infty]$. On a*

$$\int_X (\liminf_{n\to\infty} f_n) d\mu \leq \liminf_{n\to\infty} \int f_n d\mu.$$

Les rapports entre la convergence dominée presque partout et la convergence en norme L^p peuvent être décrits par les deux théorème suivants.

Théorème 5.1.2. *Soit p un nombre réel supérieur ou égal à 1, on considère une suite $(f_n)_{n\in N}$ de L^p et une fonction f telles que*

$$\lim_{n\to\infty}\int_X |f(x - f_n(x)|^p d\mu(x) = 0.$$

Il existe une fonction d'extraction ψ telle que

$$\forall\, p.p\, d\mu(x)\,,\ \lim_{n\to\infty} f_{\psi(n)}(x) = f(x).$$

Théorème 5.1.3 (de convergence dominée de Lebesgue)**.** *Soient p un réel supérieur à 1 et $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de $L^p(X, d\mu)$. Si, pour presque tout x de X,*

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

et s'il existe une fonction g de $L^p(X, d\mu)$ telle que, pour presque tout x de X, on ait

$$|f_n(x)| \leq g(x),$$

alors, on a

$$f \in L^p \quad \text{et} \quad \lim_{n\to\infty} \|f_n - f\|_{L^p} = 0.$$

Théorème 5.1.4 (de dérivation sous l'intégrale)**.** *Soient (X, μ) un espace mesuré, Ω un ouvert de $\mathbb{R}^d$ et f une application mesurable de $\Omega\times X$ dans $\mathbb{K}$. Si, pour presque tout x de X, la fonction*

$$z \longmapsto f(z, x)$$

est différentiable sur Ω, si pour tout z de Ω,

$$x \longmapsto f(z, x) \quad \text{et} \quad x \longmapsto Df(z, x)$$

sont intégrables sur X et si enfin, pour presque tout x de X et pour tout z de Ω, on a

$$|Df(z, x)| \leq g(z) \quad \text{avec} \quad g \in L^1(X, d\mu),$$

alors l'application

$$z \longmapsto \int_X f(z, x)d\mu(x)$$

est différentiable sur Ω et l'on a

$$D\int_X f(z, x)d\mu(x) = \int_X Df(z, x)d\mu(x).$$

Théorème 5.1.5 (de Fubini)**.** *Soient (X_1, μ_1) et (X_2, μ_2) deux espaces mesurés et F une fonction mesurable positive de $X_1 \times X_2$ dans $[0, +\infty]$. On a*

$$\begin{aligned}\int_{X_1\times X_2} F(x_1, x_2)d\mu_1(x_1) \otimes d\mu_2(x_2) &= \int_{X_1} d\mu_1(x_1)\int_{X_2} F(x_1, x_2)d\mu_2(x_2)\\ &= \int_{X_2} d\mu_2(x_2)\int_{X_1} F(x_1, x_2)d\mu_1(x_1).\end{aligned}$$

En appliquant ce théorème, on résout l'exercice suivant.

Exercice 5.1.1. *Soit f une fonction mesurable sur un espace mesuré (X, μ). On a*

$$\int_X |f(x)|^p d\mu(x) = p \int_0^{+\infty} \lambda^{p-1} \mu(|f| > \lambda) d\lambda.$$

Théorème 5.1.6 (de Fubini-Tonelli)**.** *Soient (X_1, μ_1) et (X_2, μ_2) deux espaces mesurés et F une fonction mesurable sur $X_1 \times X_2$. Les deux assertions suivantes sont équivalentes.*

i) La fonction F appartient à $L^1(X_1 \times X_2, d\mu_1 \otimes d\mu_2)$,

ii) Pour presque tout point x_1 de X_1, la fonction $F(x_1, \cdot)$ appartient à $L^1(X_2, d\mu_2)$, la fonction

$$\int_{X_2} F(x_1, x_2) d\mu_2(x_2) \in L^1(X_1, d\mu_1) \quad et$$

$$\left\| \int_{X_2} F(x_1, x_2) d\mu_2(x_2) \right\|_{L^1(X_1, d\mu_1)} \leq \|F\|_{L^1(X_1 \times X_2, d\mu_1 \otimes d\mu_2)}.$$

De même, pour presque tout point x_2 de X_2, la fonction $F(\cdot, x_2)$ appartient à $L^1(X_1, d\mu_1)$, la fonction

$$\int_{X_1} F(x_1, x_2) d\mu_1(x_1) \in L^1(X_2, d\mu_2) \quad et$$

$$\left\| \int_{X_1} F(x_1, x_2) d\mu_1(x_1) \right\|_{L^1(X_2, d\mu_2)} \leq \|F\|_{L^1(X_1 \times X_2, d\mu_1 \otimes d\mu_2)}.$$

De plus, lorsque i) est vérifié, on a

$$\begin{aligned} \int_{X_2} \Big(\int_{X_1} F(x_1, x_2) d\mu_1(x_1) \Big) d\mu(x_2) &= \int_{X_1} \Big(\int_{X_2} F(x_1, x_2) d\mu_2(x_2) \Big) d\mu(x_1) \\ &= \int_{X_1 \times X_2} F(x_1, x_2) d\mu_1 \otimes d\mu_2(x_1, x_2). \end{aligned}$$

5.2 Les espaces L^p comme espaces de Banach

L'étude de ces espaces requiert la notion d'exposant conjugué :

$$\text{Si} \quad p \in]1, \infty[\,, \ p' \stackrel{\text{déf}}{=} \frac{p}{p-1}, \quad \text{si} \quad p = 1, \ p' \stackrel{\text{déf}}{=} +\infty, \quad \text{et si} \quad p = +\infty\,, \ p' \stackrel{\text{déf}}{=} 1.$$

Les exposants p et p' sont dits conjugués et, en convenant que $\dfrac{1}{\infty} = 0$, on a

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

Le premier théorème important est le suivant.

Théorème 5.2.1. *Pour tout p dans $[1, \infty]$, $(L^p(X, d\mu), \| \cdot \|_{L^p})$ est un espace de Banach.*

Démonstration. Commençons par traiter le cas où p est égal à l'infini. Tout d'abord si l'on a $\|f\|_{L^\infty} = 0$, cela signifie qu'il existe une suite de réels strictement positifs $(\lambda_n)_{n \in \mathbb{N}}$ tendant vers 0 telle que, pour tout n, l'ensemble $\{x \,/\, |f(x)| > \lambda_n\}$ soit négligeable. Une réunion dénombrable de négligeables l'étant aussi, l'ensemble des x tels que $f(x)$ n'est pas nulle est négligeable. Ainsi donc $f = 0$ dans L^∞. La démonstration de l'homogénéïté est laissée en exercice et repose sur le fait que

$$\{x \,/\, |\lambda f(x)| > M\} = \Big\{x \,/\, |f(x)| > \frac{M}{|\lambda|}\Big\}.$$

Considérons f et g deux (classes de) fonctions de L^∞ et considérons deux nombres réels strictement positifs M_f et M_g tels que les deux ensembles $\{x \,/\, |f(x)| > M_f\}$ et $\{x \,/\, |g(x)| > M_g\}$ soit de mesure nulle. Remarquons que

$$\big\{x \,/\, |f(x)| + g(x) > M_f + M_g\big\} \subset \{x \,/\, |f(x)| > M_f\} \cup \{x \,/\, |g(x) > M_g\}.$$

L'espace L^∞ est donc stable par somme. De plus, la borne inférieure étant un minorant, on a $\|f+g\|_{L^\infty} \leq M_f + M_g$. La borne inférieure étant le plus grand des minorants, on a donc $\|f+g\|_{L^\infty} \leq \|f\|_{L^\infty} + \|g\|_{L^\infty}$. Ainsi donc, $(L^\infty(X, d\mu), \|\cdot\|_{L^\infty})$ est un espace vectoriel normé.

Pour démontrer que l'espace $(L^\infty, \|\cdot\|_{L^\infty})$ est un espace de Banach nous allons nous ramener à la proposition 2.1.3 page 29. Pour ce faire, considérons une suite de Cauchy $(f_n)_{n\in\mathbb{N}}$ de $(L^\infty, \|\cdot\|_{L^\infty})$ et introduisons les ensembles $E_{m,p,n}$ définis par

$$E_{m,p,n} \stackrel{\text{déf}}{=} \Big\{x \in X \,/\, |f_m(x) - f_{m+p}(x)| > \frac{1}{n+1}\Big\}.$$

Ici nous notons les classes de fonctions f_n et l'un de leur représantant de la même manière. Comme la suite $(f_n)_{n\in\mathbb{N}}$ est de Cauchy, on peut définir la fonction ϕ de $\mathbb{N}$ dans $\mathbb{N}$ par

$$\phi(n) = \min\Big\{m \geq \phi(n-1) + 1 \,/\, \forall m' \geq m \,, \forall p \geq 0\,,\ \|f_{m'} - f_{m'+p}\|_{L^\infty} < \frac{1}{n+1}\Big\}\cdot$$

Par définition de la borne supérieure essentielle, pour tout $m \geq \phi(n)$, la mesure de $E_{m,p,n}$ est nulle. Ainsi donc l'ensemble E défini par

$$E \stackrel{\text{déf}}{=} \bigcup_{\substack{(n,p)\in\mathbb{N}^2 \\ m\geq\phi(n)}} E_{m,p,n}$$

est de mesure nulle. On peut donc choisir de représenter les (classes de) fonctions f_n par des fonctions nulles sur l'ensemble négligeable E (que l'on persiste à noter f_n). Ainsi donc on a

$$\forall n\,,\ \forall m \geq \phi(n)\,,\ \forall p\,,\ \forall x \in X\,,\ |f_m(x) - f_{m+p}(x)| < \frac{1}{n+1}\cdot$$

Ainsi donc, la suite $(f_n)_{n\in\mathbb{N}}$ est de Cauchy dans $\mathcal{B}(X, \mathbb{K})$. La proposition 2.1.3 page 29 permet de conclure que $L^\infty(X, d\mu)$ est complet.

Étudions maintenant le cas où p est fini. La première étape consiste à démontrer que l'espace $(L^p(X, d\mu), \|\cdot\|_{L^p})$ est un espace vectoriel normé avec

$$\|f\|_{L^p} \stackrel{\text{déf}}{=} \Big(\int_X |f(x)|^p d\mu(x)\Big)^{\frac{1}{p}}.$$

C'est clair pour $p = 1$. Lorsque p appartient à $]1, \infty[$, observons tout d'abord que, comme

$$|f(x) + g(x)|^p \leq 2^p(|f(x)|^p + |g(x)|^p),$$

l'espace $L^p(X, d\mu)$ est un espace vectoriel. Démontrons que c'est un espace normé. Ce fait repose sur l'inégalité de Hölder énoncé dans la proposition suivante.

Proposition 5.2.1 (Inégalité de Hölder)**.** *Soient (X,μ) un espace mesuré, f une fonction de $L^p(X,d\mu)$ et g une fonction de $L^{p'}(X,d\mu)$. Alors, le produit fg appartient à $L^1(X,d\mu)$ et*

$$\int_X |f(x)g(x)|d\mu(x) \leq \|f\|_{L^p}\|g\|_{L^{p'}}.$$

Démonstration. Remarquons tout d'abord que si $p=1$ ou $p=\infty$, il n'y a rien à démontrer. De plus, on peut supposer que $\|f\|_{L^p} = \|g\|_{L^{p'}} = 1$ et que les deux fonctions f et g sont positives. En utilisant l'inégalité de convexité (2.3) page 29 on peut écrire que

$$f(x)g(x) = (f(x)^p)^{\frac{1}{p}}(g(x)^{p'})^{\frac{1}{p'}} \leq \frac{1}{p}f(x)^p + \frac{1}{p'}g(x)^{p'}$$

ce qui assure l'inégalité de Hölder. Par intégration, il vient

$$\begin{aligned}
\int_X |fg|d\mu &\leq \frac{1}{p}\int_X |f|^p d\mu + \frac{1}{p'}\int_X |g|^{p'}d\mu\\
&\leq \frac{1}{p} + \frac{1}{p'} = 1.
\end{aligned}$$

Ceci conclut la démonstration de l'inégalité de Hölder. □

Poursuite de la démonstration du théorème 5.2.1 Comme $f+g$ appartient à L^p, l'inégalité de Hölder implique que

$$\begin{aligned}
\int_X |f+g|^p d\mu &\leq \int_X |f||f+g|^{p-1}d\mu + \int_X |g||f+g|^{p-1}d\mu\\
&\leq \Big(\int_X |f|^p d\mu\Big)^{\frac{1}{p}}\Big(\int_X |f+g|^{(p-1)\frac{p}{p-1}}d\mu\Big)^{\frac{1}{p'}}\\
&\qquad + \Big(\int_X |g|^p d\mu\Big)^{\frac{1}{p}}\Big(\int_X |f+g|^{(p-1)\frac{p}{p-1}}d\mu\Big)^{\frac{1}{p'}}\\
&\leq \bigg(\Big(\int_X |f|^p d\mu\Big)^{\frac{1}{p}} + \Big(\int_X |g|^p d\mu\Big)^{\frac{1}{p}}\bigg)\Big(\int_X |f+g|^{(p-1)\frac{p}{p-1}}d\mu\Big)^{1-\frac{1}{p}}.
\end{aligned}$$

Ceci implique que l'on a $\|f+g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p}$ et donc que $L^p(X,d\mu)$ est un espace normé.

Démontrons maintenant la complétude de ces espaces. Cela repose sur le lemme suivant.

Lemme 5.2.1. *Si E est un espace normé tel que pour toute suite $(x_n)_{n\in\mathbb{N}}$, on ait*

$$\sum_{n\in\mathbb{N}} \|x_n\|_E < \infty \Longrightarrow S_N \stackrel{\text{déf}}{=} \sum_{n=0}^{N} x_n \quad \textit{converge,}$$

alors l'espace E est de Banach.

Démonstration. Rappelons tout d'abord que la proposition 1.3.2 page 19 affirme que toute suite de Cauchy admettant une valeur d'adhérence cenverge. Soit $(a_n)_{n\in\mathbb{N}}$ une suite de Cauchy de E, nous allons extraire de $(a_n)_{n\in\mathbb{N}}$ une suite convergente assurera le résultat. Définissons l'application ϕ de $\mathbb{N}$ dans $\mathbb{N}$ par $\phi(0)=0$ et

$$\phi(n+1) \stackrel{\text{déf}}{=} \min\left\{m \geq \phi(n)+1\,,\ \sup_{p\geq 0}\|a_m - a_{m+p}\| \leq \frac{1}{(2+n)^2}\right\}.$$

Ainsi donc, on a, pour tout entier n,

$$\|a_{\phi(n+1)} - a_{\phi(n)}\| \leq \frac{1}{(1+n)^2}.$$

Posons $x_n \stackrel{\text{déf}}{=} a_{\phi(n+1)} - a_{\phi(n)}$. On a

$$\sum_{n\in\mathbb{N}} \|x_n\| \leq \sum_{n\in\mathbb{N}} \frac{1}{(1+n)^2} < \infty.$$

Donc par hypothèse, la suite $\displaystyle\sum_{n=0}^{N} x_n = a_{\phi(N)} - a_0$ converge. La proposition 1.3.2 page 19 affirme que totue suite de Cauchy ayant une valeur d'adhérence converge. Le lemme est ainsi démontré. □

Poursuite de la démonstration du théorème 5.2.1 Soit $(f_n)_{n\in\mathbb{N}}$ une suite d'éléments de L^p telle que

$$\sum_{n\in\mathbb{N}} \|f_n\|_{L^p} < \infty.$$

Introduisons les fonctions

$$S_N(x) = \sum_{n=0}^{N} f_n(x) \quad \text{et} \quad S_N^+(x) = \sum_{n=0}^{N} |f_n(x)|.$$

Pour tout entier n, on a

$$\|S_N^+\|_{L^p} \leq \sum_{n\in\mathbb{N}} \|f_n\|_{L^p}$$

Ceci s'écrit

$$\forall n \in \mathbb{N}\,, \int_X S_N^+(x)^p d\mu(x) \leq \Big(\sum_{n\in\mathbb{N}} \|f_n\|_{L^p}\Big)^p.$$

La suite $(S_N^+(x))_{N\in\mathbb{N}}$ est une suite croissante. Le théorème de convergence monotone 5.1.1 implique que

$$\int_X S^+(x)^p d\mu(x) \leq \Big(\sum_{n\in\mathbb{N}} \|f_n\|_{L^p}\Big)^p \quad \text{avec} \quad S^+(x) = \sum_{n\in\mathbb{N}} |f_n(x)|.$$

La fonction S^+ est finie presque partout et appartient à L^p. Ainsi donc, pour presque tout x, la série de terme général $f_n(x)$ converge dans $\mathbb{C}$. Désignons par S sa somme. C'est un élément de L^p. On a

$$\Big\|S - \sum_{n=0}^{N} f_n\Big\|_{L^p}^p \quad = \quad \int_X \Big|S(x) - \sum_{n=0}^{N} f_n(x)\Big|^p d\mu(x).$$

On sait que

$$\Big|S(x) - \sum_{n=0}^{N} f_n(x)\Big| \leq 2S^+(x).$$

Comme S^+ est dans L^p, le théorème de convergence dominée assure que

$$\lim_{N\to\infty} \Big\|S - \sum_{n=0}^{N} f_n\Big\|_{L^p}^p = 0.$$

Le théorème est ainsi démontré. □

Corollaire 5.2.1. *Soit* $(f_n)_{n\in\mathbb{N}}$ *une suite convergente vers* f *dans* $L^1(X,d\mu)$*. Il existe une fonction d'extraction* ϕ *telle que*

$$\forall \text{p.p.}\, x\in X\,,\ \lim_{n\to\infty} f_{\phi(n)}(x)=f(x).$$

Nous allons maintenant faire quelques remarques sur l'inégalité de Hölder. Donnons tout d'abord le corollaire suivant.

Corollaire 5.2.2. *Soient* p *et* q *deux réels strictement supérieurs à* 1 *tels que*

$$\frac{1}{p}+\frac{1}{q}\leq 1.$$

Alors l'application

$$\begin{cases} L^p\times L^q & \longrightarrow & L^r\\ (f,g) & \longmapsto & fg\end{cases}$$

est une application bilinéaire continue si

$$\frac{1}{r}=\frac{1}{p}+\frac{1}{q}\cdot$$

Démonstration. On applique l'inégalité de Hölder avec $s=r/p$. D'où il vient que

$$\int_X |f(x)g(x)|^r dx\leq \|f\|_{L^p}^r\|g\|_{L^q}^r,$$

ce qui signifie que

$$\|fg\|_{L^r}\leq \|f\|_{L^p}\|g\|_{L^q},$$

d'où le corollaire. □

Corollaire 5.2.3. *Si* X *est de mesure finie, alors* $L^p(X,d\mu)\subset L^q(X,d\mu)$ *si* $p\geq q$*.*

Démonstration. Comme la fonction $\mathbf{1}$ appartient visiblement à l'espace $L^\infty(X,d\mu)\cap L^1(X,d\mu)$, donc à $L^p(X,d\mu)$ pour tout p. On a même

$$\forall p\in[1,\infty]\,,\ \|\mathbf{1}\|_{L^p}=\mu(X)^{\frac{1}{p}}.$$

Le Corollaire 5.2.2 implique que

$$\|f\|_{L^q}\leq \mu(X)^{\frac{1}{s}}\|f\|_{L^p}\quad \text{avec}\quad \frac{1}{s}=\frac{1}{q}-\frac{1}{p}\cdot$$

D'où le corollaire. □

L'inégalité de Hölder est fondamentale. Nous allons voir qu'elle est optimale au sens suivant.

Lemme 5.2.2. *Soit* (X,μ) *un espace mesuré. Soit* f *une fonction mesurable et* p *un élément de* $[1,\infty]$*. Suppsons que*

$$\sup_{\substack{\|g\|_{L^{p'}}\leq 1\\ g\geq 0}}\int_X |f(x)|g(x)d\mu(x)<+\infty. \tag{5.1}$$

Alors, f *appartient à* L^p *et*

$$\|f\|_{L^p}=\sup_{\|g\|_{L^{p'}}\leq 1}\left|\int_X f(x)g(x)d\mu(x)\right|.$$

Démonstration. Commençons par le cas où $p=1$. En choisissant pour g la fonction constante 1. La fonction g est bornée de norme L^∞ égale à 1, on obtient grâce à l'inégalité (5.1) que

$$\int_X |f(x)|g(x)d\mu(x) < \infty.$$

Considérons maintenant la fonction g définie par

$$g(x) = \frac{\overline{f}(x)}{|f(x)|} \quad \text{si} \quad f(x) \neq 0 \quad \text{et} \quad 0 \text{ sinon.}$$

La fonction g est bornée de norme égale à 1. De plus,

$$\int_X f(x)g(x)d\mu(x) = \int_X |f(x)|d\mu(x).$$

D'où le lemme dans ce cas.

Supposons maintenant que p soit réel strictement supérieur à 1 et considérons alors une suite croissante d'ensembles de mesure finie $(K_n)_{n\in\mathbb{N}}$ dont la réunion est X et posons

$$f_n^+(x) = \mathbf{1}_{K_n\cap(|f|\leq n)}|f| \quad \text{et} \quad g_n(x) = \frac{f_n^+(x)^{p-1}}{\|f_n^+\|_{L^p}^{\frac{p}{p'}}}\cdot$$

Il est clair que la fonction f_n est positive et appartient à $L^1\cap L^\infty$ donc à L^p pour tout p et que l'on a

$$\|g_n\|_{L^{p'}}^{p'} = \frac{1}{\|f_n^+\|_{L^p}^p}\int_X f_n^+(x)^{(p-1)\frac{p}{p-1}}d\mu(x) = 1.$$

La définition des fonctions f_n et g_n assure que

$$\begin{aligned}
\int_X |f(x)|\mathbf{1}_{K_n\cap(f\leq n)}g_n(x)d\mu(x) &= \int_X f_n^+(x)g_n(x)d\mu(x)\\
&= \left(\int_X f_n^+(x)^p d\mu(x)\right)\|f_n\|_{L^p}^{-\frac{p}{p'}}\\
&= \|f_n^+\|_{L^p}^{p-\frac{p}{p'}}\\
&= \|f_n^+\|_{L^p}.
\end{aligned}$$

On a donc

$$\int_X f_n^+(x)^p d\mu(x) \leq \Bigg(\sup_{\substack{\|g\|_{L^{p'}}\leq 1\\ g\geq 0}}\int_X |f(x)|g(x)d\mu(x)\Bigg)^p.$$

Le théorème de convergence monotone appliqué à la suite croissante $\big((f_n^+)^p\big)_{n\in\mathbb{N}}$ implique immédiatement que

$$\|f\|_{L^p} \leq \sup_{\substack{\|g\|_{L^{p'}}\leq 1\\ g\geq 0}}\int_X f(x)g(x)d\mu(x).$$

Ainsi donc, si l'hypothèse (5.1) est vérifiée, alors la fonction f appartient à L^p. Supposons maintenant que f appartienne à l'espace L^p. Alors en posant

$$g(x) = \frac{\overline{f}(x)|f(x)|^{p-1}}{|f(x)|\times\|f\|_{L^p}^{\frac{p}{p'}}}$$

on a

$$\|g\|_{L^{p'}}^{p'} = \frac{1}{\|f\|_{L^p}^p} \int_X |f(x)|^{(p-1)\frac{p}{p-1}} d\mu(x) = 1 \quad \text{et} \quad \|f\|_{L^p} = \int_X f(x)g(x)d\mu(x).$$

D'où le lemme pour p appartient à $[1, \infty[$. Dans le cas où $p = +\infty$, considérons un réel strictement positif λ tel que $\mu(|f| \geq \lambda)$ soit strictement positif. On pose $E_\lambda \stackrel{\text{déf}}{=} (|f| \geq \lambda)$. Soit g_0 une fonction de L^1, positive, supportée dans E_λ, et d'intégrale 1. On pose

$$g(x) = \frac{\overline{f}(x)}{|f(x)|} g_0 \,.$$

On a alors

$$\int_X |f(x)g(x)| d\mu(x) = \int_X |f(x)| g_0(x) d\mu(x) \geq \lambda \int_X g_0 d\mu(x) \geq \lambda \,.$$

Le lemme est ainsi démontré. □

5.3 Densité dans les espaces L^p

Dans toute la suite de ce chapitre, Ω désignera un ouvert de $\mathbb{R}^d$ muni de la distance euclidienne usuelle et μ désignera une mesure positive finie sur tous les compacts. Le résultat principal de cette section est le suivant.

Commençons par une digression pour la topologie des ouverts de $\mathbb{R}^d$.

Théorème 5.3.1 (de caractérisation des compacts d'un ouvert de $\mathbb{R}^d$)**.** *Soit A une partie fermée d'un ouvert Ω de $\mathbb{R}^d$ telle qu'il existe un réel strictement positif r tel que $A \subset [-r, r]^d$. Alors*

$$A \text{ est compacte} \iff \inf_{x \in A} d(x, X \setminus \Omega) > 0.$$

Démonstration. Supposons A compacte. Comme Ω est ouvert, pour tout x de A, $d(x, X \setminus \Omega)$ est strictement positif. D'après l'exercice 1.1.4, on sait que la fonction distance d'un point à un ensemble est continue. Le corollaire 1.3.2 nous dit que l'infimum est atteint, donc il est strictement positif.

Réciproquement, soit A une partie fermée de Ω telle qu'il existe un réel strictement positif r tel que $A \subset [-r, r]^d$ et vérifiant

$$\delta \stackrel{\text{déf}}{=} \inf_{x \in A} d(x, X \setminus \Omega) > 0.$$

Pour démontrer que A est compacte, il suffit de démontrer que A est une partie fermée de $\mathbb{R}^d$. Comme A est fermée dans Ω, cela signifie qu'il existe un fermé B de $\mathbb{R}^d$ tel que $A = B \cap \Omega$. Soit $\Omega_{\delta/2}$ le fermé défini par

$$\Omega_{\delta/2} \stackrel{\text{déf}}{=} \{x \,/\, d(x, X \setminus \Omega) \geq \delta/2\}.$$

On a les relations ensemblistes suivantes

$$A = A \cap \Omega_{\delta/2} = B \cap \Omega \cap \Omega_{\delta/2} = B \cap \Omega_{\delta/2}.$$

L'ensemble $B \cap \Omega_{\delta/2}$ est un fermé de $\mathbb{R}^d$ en tant qu'intersection de deux fermés, donc A est une partie fermée de $\mathbb{R}^d$. Le théorème est démontré. □

Le résultat suivant nous sera également utile

Théorème 5.3.2. *Soit Ω un ouvert de $\mathbb{R}^d$, il existe une suite de compacts $(K_n)_{n\in\mathbb{N}}$ telle que*

$$\bigcup_{n\in\mathbb{N}} K_n = \Omega\,,\ K_n \subset \overset{\circ}{K}_{n+1} \quad \textit{et} \quad \forall K \textit{ compact de } \Omega\,,\ \exists n \,/\, K \subset \overset{\circ}{K}_n\,.$$

Démonstration. Posons

$$K_n \overset{\text{déf}}{=} B(0,n) \cap \left\{ x \in \mathbb{R}^d \,/\ d(x, \mathbb{R}^d \setminus \Omega) \geq \frac{1}{n} \right\}\cdot$$

Comme Ω est un ouvert, pour tout x de X, $d(x, \mathbb{R}^d \setminus \Omega)$ est strictement positif. Donc, il existe un entier n tel que $x \in K_n$. De plus,

$$K_n \subset \overset{\circ}{B}(0, n+1) \cap \left\{ x \in \mathbb{R}^d \,/\ d(x, \mathbb{R}^d \setminus \Omega) > \frac{1}{n} \right\} \subset \overset{\circ}{K}_{n+1}\,.$$

D'où le premier point du théorème. Quant au second, il suffit d'observer que $\displaystyle\bigcup_{n\in\mathbb{N}} \overset{\circ}{K}_{n+1} = \Omega$ puisque $K_n \subset \overset{\circ}{K}_{n+1}$. Le théorème est démontré. □

Définition 5.3.1. *Soit Ω un ouvert de $\mathbb{R}^d$, on appelle suite exhaustive de compacts toute suite de compacts qui vérifie les conclusions du théorème 5.3.2 ci-dessus.*

Revenons aux espaces L^p.

Théorème 5.3.3. *Soit p un élément de $[1, +\infty[$, l'espace $C_c(\Omega)$ des fonctions continues à support compact dans Ω est dense dans $L^p(\Omega, d\mu)$.*

La démonstration de ce théorème, longue et délicate, est présentée ici à titre culturel. Pour démontrer ce théorème, nous allons tout d'abord démontrer deux résultats de densité valables pour les fonctions définies sur un espace mesuré quelconque. Démontrons tout d'abord la proposition suivante.

Proposition 5.3.1. *Si p appartient à $[1, \infty[$, $L^1(X, d\mu) \cap L^\infty(X, d\mu)$ est dense dans $L^p(X, d\mu)$.*

Démonstration. Considérons une suite croissante $(K_n)_{n\in\mathbb{N}}$ d'ensembles de mesure finie dont la réunion est l'ensemble X. Posons

$$f_n = \mathbf{1}_{K_n \cap (|f| \leq n)} f.$$

Il est clair que, pour presque tout x, on a

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

De plus, on a clairement l'inégalité

$$|f_n(x) - f(x)|^p \leq 2^p |f(x)|^p.$$

Le théorème de la convergence dominée assure alors le résultat. □

Nous allons établir maintenant un second résultat de densité.

Proposition 5.3.2. *Si p appartient à $[1, \infty[$, alors l'espace des fonctions étagées intégrables (i.e. ne prenant qu'un nombre fini de valeurs non nulles sur des ensembles de mesure finie) est dense dans $L^p(X, d\mu)$.*

Démonstration. Nous allons utiliser le fait suivant dont la démonstration est laissée en exercice au lecteur : si (Y, d) est un espace métrique, si A est dense dans Y et si tout élément de A est limlite d'une suite d'éléments de B alors B est aussi dense dans Y. Grâce à la proposition ci-dessus, il suffit de démontrer que tout fonction de $L^1(X, d\mu) \cap L^\infty(X, d\mu)$ est limite d'une suite de fonctions étagées intégrables. Pour ce faire, une fonction f de $L^1(X, d\mu) \cap L^\infty(X, d\mu)$ étant donnée, considérons la suite de fonctions $(f_n)_{n\in\mathbb{N}}$ définie par

$$f_n \stackrel{\text{déf}}{=} \sum_{j=0}^{+\infty} \frac{j}{n+1} \mathbf{1}_{f^{-1}([\frac{j}{n+1}, \frac{j+1}{n+1}[)} + \sum_{j=-\infty}^{-1} \frac{j+1}{n+1} \mathbf{1}_{f^{-1}([\frac{j}{n+1}, \frac{j+1}{n+1}[)} \cdot \tag{5.2}$$

Il est clair que

$$\forall x \in X\,, |f_n(x) - f(x)| \leq \frac{1}{n+1} \quad \text{et} \quad |f_n(x)| \leq |f(x)|.$$

Le théorème de la convergence dominée assure que $f_n \to \widetilde{f}$ dans L^p. Donc, d'après (5.4), il existe un entier n_0 tel que

$$\|f_n - f\|_{L^p} < \frac{\varepsilon}{2} \cdot \tag{5.3}$$

De plus, comme la fonction f est essentiellement bornée, les sommes apparaissant dans (5.2) sont finies, puisqu'il existe un réel λ tel que

$$\frac{|j|}{n+1} \geq \lambda \Longrightarrow \mu\Big(\widetilde{f}^{-1}\Big(\Big[\frac{j}{n+1}, \frac{j+1}{n+1}\Big[\Big)\Big) = 0.$$

Chaque fonction f_n ne prend qu'un nombre fini de valeurs. De plus, la fonction f étant intégrable, les ensembles $f^{-1}\left(\left[\frac{j}{n+1}, \frac{j+1}{n+1}\right[\right)$ sont de mesure finie ce qui achève la démonstration de la proposition.
□

Démonstration du théorème 5.3.3 Soit f une fonction de $L^p(\Omega)$. Il s'agit de démontrer que, pour tout ε strictement positif, il existe une fonction g continue à support compact telle que

$$\|f - g\|_{L^p} < \varepsilon. \tag{5.4}$$

D'après la proposition 5.3.1, il existe une fonction étagée intégrable $\widetilde{f}$ telle que

$$\|f - \widetilde{f}\|_{L^p} < \frac{\varepsilon}{2} \quad \text{avec} \quad \widetilde{f} = \sum_{j=1}^{N} \alpha_j \mathbf{1}_{A_j} \tag{5.5}$$

où les α_j sont des nombres réels et où les A_j sont des ensembles boréliens de mesure finie. Admettons un instant le lemme suivant.

Lemme 5.3.1. *Pour tout p dans $[1, \infty[$, pour tout borélien A de mesure finie, pour tout réel strictement positif ε, il existe une fonction h continue à support compact telle que*

$$\|h - \mathbf{1}_A\|_{L^p} < \varepsilon.$$

Nous pouvons donc affirmer l'existence, pour tout j de $\{1, \cdots, N\}$ l'existence d'une fonction g_j continue à support compact dans Ω telle que

$$\|\mathbf{1}_{A_j} - g_j\|_{L^p} \leq \frac{\varepsilon}{2N|\alpha_j|} \cdot$$

Définissons maintenant la fonction g par Soit $g \stackrel{\text{def}}{=} \sum_{j=1}^{N} \alpha_j g_j$. Nous avons

$$\begin{aligned} \|\widetilde{f} - g\|_{L^p} &\leq \sum_{j=1}^{N} |\alpha_j|\, \|\mathbf{1}_{A_j} - g_j\|_{L^p} \\ &\leq \frac{\varepsilon}{2} \cdot \end{aligned}$$

Le théorème 5.3.3 est alors démontré, pourvu que l'on démontre le lemme 5.3.1. □

Démonstration du Lemme 5.3.1 Soit $(K_n)_{n\in\mathbb{N}}$ une suite exhaustive de compacts (voir Théorème 5.3.2 et Définition 5.3.1 page 91). Comme A est de mesure finie, la fonction caractéristique de A appartient à L^p pour tout p. De plus, pour tout x de Ω, on a

$$\lim_{n\infty} \mathbf{1}_{K_n\cap A}(x) = \mathbf{1}_A(x) \quad \text{et} \quad 0 \leq \mathbf{1}_{K_n\cap A}(x) \leq \mathbf{1}_A(x).$$

Donc, pour tout ε strictement positif, il existe n_0 tel que

$$\|\mathbf{1}_{K_n\cap A} - \mathbf{1}_A\|_{L^p} < \frac{\varepsilon}{2} \cdot \tag{5.6}$$

Nous avons maintenant besoin d'un théorème fondamental de la théorie de la mesure dit théorème de régularité des boréliens. Nous allons admettre provisoirement ce théorème.

Théorème 5.3.4. *Pour tout ensemble A appartenant à la tribu borélienne complétée $\mathcal{B}$ et de mesure finie,*

$$\forall \varepsilon > 0\,,\ \exists K \text{ compact}\,,\ \exists U \text{ ouvert} /\ K \subset A \subset U \quad \text{et} \quad \mu(U \setminus K) < \varepsilon. \tag{5.7}$$

Conclusion de la démonstration du Lemme 5.3.1 Ainsi donc il existe un ouvert U et un compact K tel que

$$K \subset K_n \cap A \subset U \quad \text{et} \quad \mu(U \setminus K) < \Big(\frac{\varepsilon}{2}\Big)^p \cdot$$

On peut supposer que l'adhérence de U est compacte par exemple en remplaçant l'ouvert U par l'ouvert $U \cap K_n + B(0,\varepsilon)$ que l'on persistera à noter U. Posons

$$h(x) \stackrel{\text{déf}}{=} \frac{d(x,U^c)}{d(x,U^c) + d(x,K)} \cdot$$

La fonction h est continue, à valeurs dans $[0,1]$ et est telle que

$$h_{|X\setminus U} = 0 \quad \text{et} \quad h_{|K} = 1.$$

On en déduit facilement que

$$|h - \mathbf{1}_A| \leq \mathbf{1}_{U\setminus K} \quad \text{et donc que} \quad |h - \mathbf{1}_A|^p \leq \mathbf{1}_{U\setminus K} \cdot$$

D'où il vient que

$$\|h - \mathbf{1}_A\|_{L^p} < \frac{\varepsilon}{2} \cdot$$

Le lemme est ainsi démontré puisque la fonction g est nulle en dehors de l'adhérence de U, qui est compacte. □

Démonstration du théorème 5.3.4 Elle est présentée ici à titre culturel. La première étape consiste à se ramener au cas où l'ouvert Ω est de mesure finie. Pour ce faire, supposons le théorème démontré dans ce cas. Considérons alors une partie A de la tribu complétée de Ω et de mesure finie. On définit alors les suites

$$\Omega_p \stackrel{\text{déf}}{=} \Omega \cap B(0,p)\,,\ A_p \stackrel{\text{déf}}{=} A \cap B(0,p) \quad \text{et} \quad B_0 \stackrel{\text{déf}}{=} A_0 \quad \text{et} \quad B_p = A_p \setminus A_{p-1} \quad \text{pour} \quad p \geq 1.$$

Les ensembles B_p vérifient

$$\bigcup_{p'=0}^{p} B_{p'} = \bigcup_{p'=0}^{p} A_{p'}.$$

Comme ils sont deux à deux disjoints, on a

$$\sum_{p=0}^{\infty} \mu(B_p) = \mu(A) < \infty. \tag{5.8}$$

L'assertion (5.7) implique alors que

$$\forall \varepsilon > 0\,,\ \forall p\,,\ \exists K_{p,\varepsilon} \text{ compact}\,,\ \exists U_{p,\varepsilon} \text{ ouvert de } \Omega \,/\, K_{p,\varepsilon} \subset A_p \subset U_{p,\varepsilon} \quad \text{et} \quad \mu(U_{p,\varepsilon} \setminus K_{p,\varepsilon}) < \varepsilon 2^{-p-2}.$$

L'assertion (5.8) implique que la série de terme général $(\mu(K_{p,\varepsilon}))_{p\in\mathbb{N}}$ converge. Il existe donc un entier p_ε tel que

$$\sum_{p=p_\varepsilon+1}^{\infty} \mu(B_p) < \frac{\varepsilon}{2}\cdot \tag{5.9}$$

Posons alors

$$U_\varepsilon \stackrel{\text{déf}}{=} \bigcup_{p=0}^{\infty} U_{p,\varepsilon} \quad \text{et} \quad K_\varepsilon \stackrel{\text{déf}}{=} \bigcup_{p=0}^{p_\varepsilon} K_{p,\varepsilon}$$

On a

$$A = \bigcup_{p=0}^{\infty} B_p \subset U_\varepsilon \quad \text{et} \quad K_\varepsilon \subset \bigcup_{p=0}^{p_\varepsilon} B_p \subset A.$$

De plus, U_ε est un ouvert de Ω en tant que réunion d'ouverts et K_ε est un compact en tant que réunion finie de compacts. Enfin on a

$$\begin{aligned} U_\varepsilon \setminus K_\varepsilon &= \bigcup_{p=0}^{\infty} U_{p,\varepsilon} \cap \Big(\bigcup_{p=0}^{p_\varepsilon} K_{p,\varepsilon}\Big)^c \\ &\subset \bigcup_{p=0}^{\infty} U_{p,\varepsilon} \cap \bigcap_{p=0}^{p_\varepsilon} K_{p,\varepsilon}^c \\ &\subset \Big(\bigcup_{p=0}^{\infty} U_{p,\varepsilon} \cap \bigcap_{p=0}^{\infty} K_{p,\varepsilon}^c\Big) \cup \bigcup_{p=p_\varepsilon+1}^{\infty} K_{p,\varepsilon}. \end{aligned}$$

Comme l'on a

$$\bigcup_{p=0}^{\infty} U_{p,\varepsilon} \cap \bigcap_{p=0}^{\infty} K_{p,\varepsilon}^c \subset \bigcup_{p=0}^{\infty} (U_{p,\varepsilon} \setminus K_{p,\varepsilon}),$$

on en déduit que

$$U_\varepsilon \setminus K_\varepsilon \subset \Big(\bigcup_{p=0}^{\infty} (U_{p,\varepsilon} \setminus K_{p,\varepsilon})\Big) \cup \bigcup_{p=p_\varepsilon+1}^{\infty} K_{p,\varepsilon}.$$

Comme $K_{p,\varepsilon}$ est inclus dans B_P, l'inégalité (5.12) assure que l'on a

$$\begin{aligned} \mu(U_\varepsilon \setminus K_\varepsilon) &\leq \sum_{p=0}^{\infty} \mu(U_{p,\varepsilon} \setminus K_{p,\varepsilon}) + \sum_{p=p_\varepsilon+1}^{\infty} \mu(K_{p,\varepsilon} \\ &\leq \sum_{p=0}^{\infty} \varepsilon 2^{-p-2} + \frac{\varepsilon}{2} \\ &\leq \varepsilon. \end{aligned}$$

Nous allons démontrer ce théorème dans le cas où Ω est de mesure finie. Désignons par $\mathcal{A}$ l'ensemble des parties de Ω qui vérifie (5.7). Nous allons démontrer tout d'abord que les ouverts appartiennent à $\mathcal{A}$ puis que $\mathcal{A}$ est une tribu, ce qui assurera le résultat.

Si $A = U$ est un ouvert, alors la suite (croisssante) de compacts $(K_n)_{n\in\mathbb{N}}$ définie par

$$K_N \stackrel{\text{def}}{=} \big\{x \in U \,/\, \|x\| \leq n \text{ et } d(x, U^c) \geq 2^{-n}\big\}$$

est telle que

$$\forall x \in U\,,\ \lim_{n\to\infty} \mathbf{1}_{K_n}(x) = \mathbf{1}_U(x).$$

Le théorème 5.1.1 de convergence monotone implique alors que

$$\lim_{n\to\infty} \mu(U \setminus K_n) = 0$$

ce qui assure que U vérifie appartient à $\mathcal{A}$. Considérons maintenant un élément A de $\mathcal{A}$, ε un réel strictement positif, U un ouvert et K un compact tel que

$$\mu(U \setminus K) < \frac{\varepsilon}{2}\cdot$$

Si C est inclus dans B, alors on a $B \setminus C = B \cap C^c$, et donc $C^c \setminus B^c = B \setminus C$. On a alors

$$U^c \subset A^c \subset K^c \quad \text{et} \quad \mu(K^c \setminus U^c) = \mu(U \setminus K) < \frac{\varepsilon}{2}\cdot$$

Mais U^c n'est pas nécessairement compact. Pour contourner cette difficulté, introduisons une suite de compacts K_n donné par le théorème 5.3.2 page 91. Il est clair que

$$\forall x \in \Omega\,,\ \mathbf{1}_{K^c\setminus(U^c\cap K_n)}(x) = \mathbf{1}_{K^c}(x) - \mathbf{1}_{U^c\cap K_n}(x) = \mathbf{1}_{K^c}(x) - \mathbf{1}_{U^c}(x)$$

Le théorème de convergence monotone permet de conclure que A^c appartient à $\mathcal{A}$.

Il reste maintenant à démontrer que, si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{A}$, alors la réunion des A_n appartient à $\mathcal{A}$. Commençons par le démontrer pour la réunion d'un nombre fini. Soient A_1 et A_2 deux éléments de $\mathcal{A}$. Pour tout réel ε strictement positif, il existe deux ouverts U_1 et U_2 et deux compacts K_1 et K_2 tels que

$$K_j \subset A_j \subset U_j \quad \text{et} \quad \mu(U_j \setminus K_j) < \frac{\varepsilon}{2}\,\cdot$$

$$K_1 \cup K_2 \subset A_1 \cup A_2 \subset U_1 \cup U_2. \tag{5.10}$$

L'ensemble $U_1 \cup U_2$ est un ouvert et l'ensemble $K_1 \cup K_2$ est un compact. De plus, on a

$$\begin{aligned}(U_1 \cup U_2) \setminus (K_1 \cup K_2) &= (U_1 \cup U_2) \cap (K_1 \cup K_2)^c\\ &= \big(U_1 \cap (K_1 \cup K_2)^c\big) \cup \big(U_2 \cap (K_1 \cup K_2)^c\big)\\ &\subset (U_1 \cap K_1^c) \cup (U_2 \cap K_2^c).\end{aligned}$$

D'où il résulte que

$$\begin{aligned}\mu\big((U_1 \cup U_2) \setminus (K_1 \cup K_2)\big) &\leq \mu(U_1 \setminus K_1) + \mu(U_2 \setminus K_2)\\ &\leq \varepsilon.\end{aligned}$$

Avec (5.8), ceci assure que $U_1 \cup U_2$ appartient à $\mathcal{A}$. Comme $\mathcal{A}$ est stable par complémentaire, si A_1 et A_2 appartiennent à $\mathcal{A}$, alors

$$A_1 \cap A_2 = (A_1^c \cup A_2^c)^c \in \mathcal{A}.$$

Et donc $\mathcal{A}$ est stable par réunion finie.

Soit maintenant une suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{A}$. On suit une démarche analogue à celle utilisée pour la réduction au cas où Ω est de mesure finie. Définissons

$$B_n \stackrel{\text{déf}}{=} \bigcup_{k=0}^{n} A_k \setminus \bigcup_{j=0}^{k-1} A_k.$$

Les ensembles B_n sont deux à deux disjoints et leur réunion est égale à celle des A_n et l'on a

$$\sum_{n\in\mathbb{N}} \mu(B_n) = \mu\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) < \infty.$$

Comme les ensembles B_n appartiennent à $\mathcal{A}$, il existe une suite $(U_n)_{n\in\mathbb{N}}$ d'ouverts et une suite $(K_n)_{n\in\mathbb{N}}$ de compacts telles que

$$\forall n\in\mathbb{N}\,,\ K_{n,\varepsilon}\subset B_n\subset U_{n,\varepsilon}\quad\text{et}\quad \mu(U_{n,\varepsilon}\setminus K_{n,\varepsilon})<\varepsilon 2^{-n-2}. \tag{5.11}$$

Il existe un entier n_ε tel que

$$\sum_{n=n_e+1}^{\infty}\mu(B_n)\leq\frac{\varepsilon}{2}\,\cdot \tag{5.12}$$

On pose alors

$$U_\varepsilon\stackrel{\text{déf}}{=}\bigcup_{n\in\mathbb{N}}U_{n,\varepsilon}\quad\text{et}\quad K_\varepsilon\stackrel{\text{déf}}{=}\bigcup_{n=0}^{n_\varepsilon}K_{n,\varepsilon}$$

L'ensemble U_ε est un ouvert en tant que réunion d'ouverts et K_ε est un compact en tant que réunion finie de compacts. Il est clair que

$$K_\varepsilon\subset\bigcup_{n\in\mathbb{N}}A_n\subset U_\varepsilon. \tag{5.13}$$

De plus, on a

$$K_\varepsilon^c=\Big(\bigcup_{n=0}^{n_\varepsilon}K_{n,\varepsilon}\Big)^c=\bigcap_{n=0}^{n_\varepsilon}K_{n,\varepsilon}^c.$$

Ainsi donc, on en déduit

$$\begin{aligned}
U_\varepsilon\setminus K_\varepsilon &= \Big(\bigcup_{n\in\mathbb{N}}U_{n,\varepsilon}\Big)\cap\Big(\bigcap_{n=0}^{n_\varepsilon}K_{n,\varepsilon}^c\Big)\\
&\subset \bigg(\Big(\bigcup_{n\in\mathbb{N}}U_{n,\varepsilon}\Big)\cap\Big(\bigcap_{n\in\mathbb{N}}K_{n,\varepsilon}^c\Big)\bigg)\cup\bigcup_{n=n_\varepsilon+1}^{\infty}K_{n,\varepsilon}\\
&\subset \Big(\bigcup_{n\in\mathbb{N}}U_{n,\varepsilon}\cap K_{n,\varepsilon}^c\Big)\cup\bigcup_{n=n_\varepsilon+1}^{\infty}K_{n,\varepsilon}\\
&\subset \Big(\bigcup_{n\in\mathbb{N}}U_{n,\varepsilon}\setminus K_{n,\varepsilon}\Big)\cup\bigcup_{n=n_\varepsilon+1}^{\infty}K_{n,\varepsilon}.
\end{aligned}$$

D'après (5.11) et (5.12), on en déduit que

$$\begin{aligned}
\mu(U_\varepsilon\setminus K_\varepsilon) &\leq \sum_{n\in\mathbb{N}}\mu(U_{n,\varepsilon}\setminus K_{n,\varepsilon})+\sum_{n=n_\varepsilon+1}^{\infty}\mu(K_{n,\varepsilon})\\
&\leq \varepsilon\sum_{n\in\mathbb{N}}2^{-n-2}+\sum_{n=n_\varepsilon+1}^{\infty}\mu(B_n)\\
&\leq \varepsilon.
\end{aligned}$$

D'où le théorème. □

De ce théorème, que l'on admettra, on déduit l'énoncé fondamental suivant.

Corollaire 5.3.1. *Soit p un réel de l'intervalle $[1,\infty[$; l'espace $L^p(\Omega)$ est séparable.*

Démonstration. Soit $(K_n)_{n\in\mathbb{N}}$ une suite exaustive de compacts (voir la définition 5.3.1 page 91). Il suffit de trouver une partie démonbrable A de $L^p(\Omega)$ telle que, pour toute fonction continue à support compact dans K_n, on ait

$$\forall\varepsilon\,,\ \exists g\in A\ /\ \|f-g\|_{L^\infty(K_n)}\leq\varepsilon. \tag{5.14}$$

Soit $(\psi_n)_{n\in\mathbb{N}}$ une suite de fonctions continues valant 1 près de K_n et dont le support est inclus dans $\overset{\circ}{K}_{n+1}$. On prend pour A l'ensemble des fonctions de type $\psi_n P$ où P est un polynôme à d indéterminées à coefficients rationnels. Démontrer que cet ensemble satisfait à la propriété (5.14) est un excellent exercice que le lecteur est fortement invité à faire. $\square$

Exercice 5.3.1. *Soient $x \in \mathbb{R}^d$ et τ_x l'application définie par*

$$\tau_x \begin{cases} L^p & \longrightarrow & L^p \\ f & \longmapsto & \tau_x(f) \,:\, y \mapsto f(x-y). \end{cases}$$

L'application τ_x est une isométrie de L^p dans L^p et, si p est réel, alors, on a

$$\forall f \in L^p\,,\ \lim_{x\to 0} \|\tau_x f - \check{f}\|_{L^p} = 0 \quad \text{avec} \quad \check{f}(y) \stackrel{\text{déf}}{=} f(-y).$$

L'exercice ci-après montre comment l'on peut "régulariser" linéairement les fonctions de L^p.

Exercice 5.3.2. *1) Soit K une partie compacte de $\mathbb{R}^d$, démontrer que, pour tout ε strictement positif, il existe une famille finie $(A_{k,\varepsilon})_{k\in\mathbb{N}}$ d'ensembles disjoints, d'adhérence compacte dont le diamètre est inférieur à ε et telle que*

$$\bigcup_{k=0}^{\infty} A_{k,\varepsilon} = K.$$

On considère l'application P_n définie par

$$P_n \begin{cases} L^1 & \longrightarrow & L^\infty \\ f & \longmapsto & P_n(f) \stackrel{\text{déf}}{=} \displaystyle\sum_{j=1}^{\infty} \frac{1}{\mu(A_{j,\frac{1}{n+1}})} \Big(\int_{A_{j,\frac{1}{n+1}}} f(x) d\mu(x) \Big) \mathbf{1}_{A_{j,\frac{1}{n+1}}}\,. \end{cases}$$

2) Démontrez que P_n est une application linéaire continue de $L^1(K)$ dans $L^\infty(K)$.
3) Démontrez que, pour tout p, on a $\|P_n\|_{\mathcal{L}(L^p;L^p)} \leq 1$.
4) Démontrez que, si f appartient à L^p et g à $L^{p'}$, alors

$$\int_K f(x) P_n g(x) d\mu(x) = \int_K P_n f(x) g(x) d\mu(x)\,.$$

5) Démontrez que

$$\forall f \in L^p\,,\ \lim_{n\to\infty} \|P_n f - f\|_{L^p} = 0\,.$$

5.4 Convolution et régularisation

Dans toute cette section, nous travaillerons dans l'espace $\mathbb{R}^d$. L'opérateur de convolution est un opérateur crucial dans l'étude des fonctions sur $\mathbb{R}^d$.

Théorème 5.4.1. *Soient f et g deux fonctions de L^1. Alors, pour presque tout x de $\mathbb{R}^d$, la fonction*

$$y \longmapsto f(x-y)g(y)$$

est intégrable et la fonction F définie par

$$F(x) \stackrel{\text{déf}}{=} \int_{\mathbb{R}^d} f(x-y)g(y)dy$$

appartient à L^1. Elle est appelée la convoluée de f et de g et notée $f \star g$. L'opération $\star$ ainsi définie est une application bilinéaire continue de $L^1 \times L^1$ dans L^1. On a

$$\|f\star g\|_{L^1} \leq \|f\|_{L^1}\|g\|_{L^1} \quad \text{et} \quad \int_{\mathbb{R}^d} (f\star g)(x)dx = \Big(\int_{\mathbb{R}^d} f(x)dx\Big)\Big(\int_{\mathbb{R}^d} g(x)dx\Big).$$

Démonstration. Comme les fonctions f et g sont supposées être dans L^1, on a

$$|f(x-y)| \times |g(y)| \in L^1(\mathbb{R}^d \times \mathbb{R}^d) \quad \text{et} \quad \int_{\mathbb{R}^d \times \mathbb{R}^d} |f(x-y)| \times |g(y)| dxdy = \|f\|_{L^1} \|g\|_{L^1}.$$

Les conclusions du théorème 5.1.5 de Fubini entraînent immédiatement la première partie du théorème. Pour la seconde, il suffit d'appliquer le théorème de Fubini-Tonelli. $\square$

On peut aussi définir la convolution d'une fonction de L^p par une fonction de $L^{p'}$.

Théorème 5.4.2. *Soient f une fonction de L^p et g une fonction de $L^{p'}$. La formule*

$$(f \star g)(x) \stackrel{\text{déf}}{=} \int_{\mathbb{R}^d} f(x-y)g(y)dy$$

définit une application bilinéaire continue de $L^p \times L^{p'}$ dans L^∞.

Démonstration. D'après l'inégalité de Hölder, pour presque tout x dans $\mathbb{R}^d$, l'application

$$y \longmapsto f(x-y)g(y)$$

est intégrable et

$$\left| \int_{\mathbb{R}^d} f(x-y)g(y)dy \right| \leq \|f\|_{L^p} \|g\|_{L^{p'}}.$$

D'où le théorème. $\square$

Les deux théorèmes ci-dessus se généralisent. On peut définir la convolution de deux fonctions si elles appartiennent à des espaces L^p convenables. Plus précisément :

Théorème 5.4.3 (Inégalités de Young)**.** *Soit (p, q, r) un triplet de réels tel que*

$$1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q} \cdot \tag{5.15}$$

Considérons un couple de fonctions (f, g) dans $L^p \times L^q$. Alors la formule

$$(f \star g)(x) \stackrel{\text{déf}}{=} \int_{\mathbb{R}^d} f(x-y)g(y)dy$$

définit une fonction de L^r et l'on a

$$\|f \star g\|_{L^r} \leq \|f\|_{L^p} \|g\|_{L^q}.$$

Démonstration. C'est un exemple d'application du lemme 5.2.2. Supssons tout d'abord que f et g soient positives. Soit φ une fonction positive de $L^{r'}$, nous allons majorer

$$I(f, g, \varphi) \stackrel{\text{déf}}{=} \int_{\mathbb{R}^d \times \mathbb{R}^d} f(x-y)g(y)\varphi(x)dxdy.$$

Supposons démontré que

$$I(f, g, \varphi) \leq \|f\|_{L^p} \|g\|_{L^q} \|\varphi\|_{L^{r'}}. \tag{5.16}$$

Ceci signifie que

$$(x, y) \longmapsto f(x-y)g(y)\varphi(x)$$

appartient à $L^1(\mathbb{R}^d\times\mathbb{R}^d)$. Le théorème de Fubini-Tonelli 5.1.5 assure que, pour presque tout x,

$$y \longmapsto f(x-y)g(y)$$

appartient à $L^1(\mathbb{R}^d, dy)$. Ainsi donc $f\star g$ est bien définie pour presque tout x de $\mathbb{R}^d$. De plus, le théorème de Fubini assure que

$$\begin{aligned}\int_{\mathbb{R}^d} f\star g(x)\,\varphi(x)dx &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(x-y)g(y)dy\varphi(x)dx\\ &\leq I(f,g,\varphi)\,.\end{aligned}$$

L'inégalité (5.16) et le lemme 5.2.2 impliquent que $f\star g$ appartient à L^r et que

$$\|f\star g\|_{L^r}\leq \|f\|_{L^p}\|g\|_{L^p}.$$

Si les fonctions f et g sont de signe quelconque (ou à valeurs complexes), il suffit d'appliquer ce qui procède à $|f|$ et $|g|$ et d'observer que

$$\Big|\int_{\mathbb{R}^d} f(x-y)g(y)dy\Big|\leq \int_{\mathbb{R}^d}|f(x-y)|\,|g(y)|dy.$$

Démontrons l'inégalité (5.16). Remarquons tout d'abord que l'on ne restreint pas la généralité de la démonstration en supposant que $\|f\|_{L^p}=\|g\|_{L^q}=1$. Considérons α et β deux réels de l'intervalle $]0,1[$. On écrit

$$I(f,g,\varphi)=\int_{\mathbb{R}^d\times\mathbb{R}^d} f(x-y)^{1-\alpha}g(y)^{1-\beta}\varphi(x)\times f(x-y)^{\alpha}g(y)^{\beta}dxdy\,.$$

Appliquons l'inégalité de Hölder avec la mesure positive

$$d\mu(x,y)=f(x-y)^{\alpha}g(y)^{\beta}dxdy$$

et le couple de réels (r,r'). Il en résulte que

$$\begin{aligned}I(f,g,\varphi) &\leq I^{(1)}(f,g,\varphi)^{\frac1r}I^{(2)}(f,g,\varphi)^{1-\frac1r}\quad\text{avec}\\ I^{(1)}(f,g,\varphi) &\stackrel{\text{déf}}{=} \int_{\mathbb{R}^d\times\mathbb{R}^d}|f(x-y)|^{(1-\alpha)r+\alpha}|g(y)|^{(1-\beta)r+\beta}dxdy\quad\text{et}\\ I^{(2)}(f,g,\varphi) &\stackrel{\text{déf}}{=} \int_{\mathbb{R}^d\times\mathbb{R}^d}\varphi(x)^{r'}f(x-y)^{\alpha}g(y)^{\beta}dxdy\,.\end{aligned}$$

Majorons $I^{(1)}(f,g,\varphi)$. Le théorème de Fubini pour les fonctions positives et l'invariance par translation de la mesure de Lebesgue impliquent que

$$I^{(1)}(f,g,\varphi)=\Big(\int_{\mathbb{R}^d} f(x)^{(1-\alpha)r+\alpha}dx\Big)\Big(\int_{\mathbb{R}^d} g(y)^{(1-\beta)r+\beta}dy\Big)$$

Il est naturel de choisir α tel que $(1-\alpha)r+\alpha=p$ et β tel que $(1-\beta)r+\beta=q$, ce qui donne

$$\alpha=\frac{r-p}{r-1}\quad\text{et}\quad\beta=\frac{r-q}{r-1}\cdot$$

Comme $\|f\|_{L^p}=\|g\|_{L^q}=1$, on trouve que $I^{(1)}(f,g,\varphi)=1$. Pour majorer $I^{(2)}(f,g)$, observons que, par définition de α et β, on a

$$\frac{\alpha}{p}+\frac{\beta}{q}=1.$$

L'inégalité de Hölder entraîne alors que

$$\begin{aligned}\int_{\mathbb{R}^d} f(x-y)^\alpha |g(y)|^\beta dy &\leq \Big(\int_{\mathbb{R}^d} f(x-y)^{\alpha\times\frac{p}{\alpha}} dy\Big)^{\frac{\alpha}{p}} \Big(\int_{\mathbb{R}^d} g(y)^{\beta\times\frac{q}{\beta}} dy\Big)^{\frac{\beta}{q}}\\ &\leq \|f\|_{L^p}^\alpha \|g\|_{L^q}^\beta\\ &\leq 1.\end{aligned}$$

D'où il résulte que

$$I^{(2)}(f,g,\varphi) \leq \|\varphi\|_{L^{r'}}.$$

L'inégalité 5.16 et alors le théorème 5.4.3 sont démontrés. □

Théorème 5.4.4. *Soient f et g deux fonctions respectivement dans L^p et L^q telles que*

$$\frac{1}{p}+\frac{1}{q}\geq 1.$$

Alors, on a

$$\text{Supp }(f\star g)\subset \text{Adh }(\text{Supp }\ f+\text{Supp }\ g).$$

Démonstration. Soit x un point de $\mathbb{R}^d$ et ρ un réel strictement positif tels que

$$B(x,\rho)\cap(\text{Supp }\ f+\text{Supp }\ g)=\emptyset.$$

Pour toute fonction φ bornée et nulle en dehors de $B(x,\rho)$, on a

$$\begin{aligned}\int_{\mathbb{R}^d\times\mathbb{R}^d}\varphi(x)f(x-y)g(y)dxdy &= \int_{\mathbb{R}^d\times\mathbb{R}^d}\varphi(x+y)f(x)g(y)dxdy\\ &= 0\,.\end{aligned}$$

D'où le théorème en appliquant le lemme 5.2.2 pour $X=B(x,\rho)$. □

La convolution est une opération cruciale, car elle permet la définition d'une procédure explicite d'approximation et de régularisation.

Théorème 5.4.5. *Soient φ une fonction de $L^1(\mathbb{R}^d)$ d'intégrale 1 et p un réel supérieur ou égal à 1. Posons*

$$\varphi_\varepsilon(x)=\varepsilon^{-d}\varphi\left(\frac{x}{\varepsilon}\right)\cdot$$

On a alors, pour toute fonction f appartenant à L^p,

$$\lim_{\varepsilon\to 0}\|\varphi_\varepsilon\star f-f\|_{L^p}=0.$$

Démonstration. Soient f une fonction de L^p et η un réel strictement positif. Il existe deux fonctions g et ψ continues à support compact telles que

$$\|f-g\|_{L^p}<\frac{\eta}{8\|\varphi\|_{L^1}}\quad\text{et}\quad\|\varphi-\psi\|_{L^1}<\frac{\eta}{8\|g\|_{L^p}+1}\cdot$$

Comme la fonction φ, donc aussi la fonction φ_ε, est d'intégrale 1, on peut écrire que

$$\begin{aligned}\varphi_\varepsilon\star f-f &= \varphi_\varepsilon\star f-\Big(\int_{\mathbb{R}^d}\varphi_\varepsilon(x)dx\Big)f\\ &= \varphi_\varepsilon\star(f-g)-\Big(\int_{\mathbb{R}^d}\varphi_\varepsilon(x)dx\Big)(f-g)\\ &\quad+(\varphi_\varepsilon-\psi_\varepsilon)\star g-\Big(\int_{\mathbb{R}^d}(\varphi_\varepsilon-\psi_\varepsilon)(x)dx\Big)g\\ &\quad+\psi_\varepsilon\star g-\Big(\int_{\mathbb{R}^d}\psi_\varepsilon(x)dx\Big)g.\end{aligned}$$

Mais, vu le choix de la fonction g, on a

$$\|\varphi_\varepsilon\star(f-g)\|_{L^p}<\frac{\eta}{8}\quad\text{et}\quad\Big\|\Big(\int_{\mathbb{R}^d}\varphi_\varepsilon(x)dx\Big)(f-g)\Big\|_{L^p}<\frac{\eta}{8}\cdot$$

Vu le choix de ψ, on a

$$\|(\varphi_\varepsilon-\psi_\varepsilon)\star g\|_{L^p}<\frac{\eta}{8}\quad\text{et}\quad\Big\|\Big(\int_{\mathbb{R}^d}(\varphi_\varepsilon-\psi_\varepsilon)(x)dx\Big)g\Big\|_{L^p}<\frac{\eta}{8}\cdot$$

Finalement, on a

$$\|\varphi_\varepsilon\star f-f\|_{L^p}<\frac{\eta}{2}+\Big\|\psi_\varepsilon\star g-\Big(\int_{\mathbb{R}^d}\psi_\varepsilon(x)dx\Big)g\Big\|_{L^p}\cdot$$

Mais, par définition de la convolution, on a, grâce à un changement de variables,

$$(\psi_\varepsilon\star g)(x)-\Big(\int_{\mathbb{R}^d}\psi_\varepsilon(x)dx\Big)g(x)=\int_{\mathbb{R}^d}\psi(z)(g(x-\varepsilon z)-g(x))dz\,.$$

On en déduit immédiatement que

$$|(\psi_\varepsilon\star g)(x)-\Big(\int_{\mathbb{R}^d}\psi_\varepsilon(x)dx\Big)g(x)|\leq\int_{\mathbb{R}^d}|\psi(z)|\,|g(x-\varepsilon z)-g(x)|dz\,.$$

La fonction g est uniformément continue, donc, pour tout η, il existe un α strictement positif tel que

$$|y-y'|\leq\alpha\Rightarrow|g(y)-g(y')|<\frac{\eta}{2\|\psi\|_{L^1}|\operatorname{Supp}\ g+B(0,\varepsilon R)|^{\frac{1}{p}}+1}\cdot$$

Posons

$$R\stackrel{\text{déf}}{=}\sup_{z\in\operatorname{Supp}\ \psi}|z|.$$

On a alors

$$\varepsilon<\frac{\alpha}{R}\Rightarrow\|\psi_\varepsilon\star g-g\|_{L^\infty}\leq\frac{\eta}{2|\operatorname{Supp}\ g+B(0,\varepsilon R)|^{\frac{1}{p}}}\cdot$$

D'où l'on déduit que

$$\|\varphi_\varepsilon\star f-f\|_{L^p}<\eta.$$

Le théorème est ainsi démontré. □

Remarque importante Comme le montre l'exercice suivant, ce théorème ne s'applique pas pour p vaut ∞.

Exercice 5.4.1. *Considérer pour f la fonction de Heavyside H (la fonction caractérisique des réels positifs) et pour φ une fonction continue à support compact, paire et d'intégrale* 1*. Démontrer alors que $\|\varphi_\varepsilon\star f-f\|_{L^\infty}\geq 1/2$.*

Comme nous allons le voir, ce théorème est extrêmement important pour la régularisation des fonctions. En effet, le théorème de dérivation sous l'intégrale assure que si φ est une fonction indéfiniment différentiable à support compact, alors la fonction $\varphi_\varepsilon\star f$ l'est aussi. Ainsi l'on aura approximé toute fonction de L^p, pour p réel, par une fonction indéfiniment différentiable à support compact. L'existence de telles fonctions doit être démontrée.

Proposition 5.4.1. *Soit f la fonction de $\mathbb{R}$ dans $\mathbb{R}$ définie par*

$$f(x) = e^{\frac{1}{x(x-1)}} \quad si \quad x \in [0,1] \quad et \quad f(x) = 0 \quad sinon.$$

Cette fonction est indéfiniment différentiable à support compact.

Démonstration. Il suffit d'observer que

$$f^{(k)}(x) = \frac{P_k(x)}{x^{k+1}(1-x)^{k+1}} e^{\frac{1}{x(x-1)}},$$

où P_k est un polynôme de degré δ_k. Les détails sont laissés en exercice.

Remarque Les familles $(\varphi_\varepsilon)_\varepsilon$ sont appelées suites régularisantes ou bien approximations de l'identité.

Corollaire 5.4.1. *Soient Ω un ouvert de $\mathbb{R}^d$, on désigne par $\mathcal{D}(\Omega)$ l'ensemble des fonctions indéfiniment différentiables et à support compact inclus dans Ω. L'espace $\mathcal{D}(\Omega)$ est différent de $\{0\}$.*

Démonstration. Il suffit pour cela de considérer $\varphi(x) = f(|x|)$ où f est la fonction définie dans la proposition 5.4.1 ci-dessus et de faire une translation et une homothétie. □

Corollaire 5.4.2. *Pour tout réel $p \geq 1$, l'espace $\mathcal{D}(\mathbb{R}^d)$ est dense dans $L^p(\mathbb{R}^d)$*

Démonstration. D'après le corollaire 5.4.1 il existe une fonction φ indéfiniment différentiable à support compact d'intégrale 1. Considérons la famille $(\varphi_\varepsilon)_{\varepsilon>0}$

$$\varphi_\varepsilon(x) = \varepsilon^{-d}\varphi\left(\frac{x}{\varepsilon}\right)\cdot$$

Considérons alors une fonction f de L^p. On peut trouver une fonction g continue à support compact telle g soit arbitrairement proche de f dans L^p. posons

$$g_\varepsilon = \varphi_\varepsilon \star g.$$

D'après la proposition 5.4.4, la fonction g_ε est à support compact. Comme la fonction φ, donc aussi la fonction φ_ε est indéfiniment différentiable, le théorème de dérivation sous l'intégrale assure g_ε est indéfiniment différentiable à support compact. D'où le corollaire. □

Corollaire 5.4.3. *Soit Ω un ouvert de $\mathbb{R}^d$ et p un réel de l'intervalle $[1, \infty[$. L'espace $\mathcal{D}(\Omega)$ est dense dans $L^p(\Omega)$.*

Démonstration. Soit $(K_n)_{n\in\mathbb{N}}$ une suite de exhaustive de compacts (voir la définition 5.3.1 page 91). On pose alors $f_n \stackrel{\text{déf}}{=} \mathbf{1}_{K_n} f$. D'après le théorème 5.1.3, on a

$$\lim_{n\to\infty} \|f_n - f\|_{L^p(\Omega)} = 0.$$

Posons alors $f_{n,\varepsilon} \stackrel{\text{déf}}{=} \varphi_\varepsilon \star f_n$. Cette fonction est indéfiniment différentiable. De plus, son support est le compact $K_n + B(0,\varepsilon)$, c'est-à-dire l'ensemble des points à distance au plus ε de K_n. D'après la proposition 5.3.1 page 90, pour $\varepsilon \leq \varepsilon_n$, alors le compact $K_n + B(0,\varepsilon)$ est inclus dans Ω. D'où le corollaire. □

Corollaire 5.4.4. *Soit K un compact d'un ouvert Ω de $\mathbb{R}^d$; il existe une fonction ψ de $\mathcal{D}(\Omega)$ telle que ψ vaille 1 au voisinage de K.*

Démonstration. Soit δ un réel strictement positif tel que, si

$$K_r \stackrel{\text{def}}{=} \{x \in \Omega \,/\, d(x,K) \leq 2r\},$$

alors $K_{2\delta}$ soit inclus dans Ω. (Un tel réel existe d'après le théorème 5.3.1 page 90). On considère alors une approximation de l'identité $(\varphi_\varepsilon)_\varepsilon$ et l'on pose

$$\psi_\varepsilon(x) \stackrel{\text{def}}{=} \varphi_\varepsilon \star \mathbf{1}_{K_\delta}(x) = \int_{K_\delta} \varphi_\varepsilon(x-y)dy\,.$$

Il existe une constante C telle que

$$x \notin K_\delta + B(0, C\varepsilon) \Longrightarrow \varphi_\varepsilon \star 1_{K_\delta}(x) = 0\,.$$

Donc, si $C\varepsilon < \delta$, alors $\psi_\varepsilon \in \mathcal{D}(\Omega)$ car $K_{2\delta} \subset \Omega$. Enfin, si x est dans $K_{\delta - C\varepsilon}$ (qui est non vide puisque $\delta > C\varepsilon$), alors, pour tout $y \in B(x, C\varepsilon)$, y dans K_δ. Ainsi donc

$$\begin{aligned} \varphi_\varepsilon \star \mathbf{1}_{K_\delta}(x) &= \int_{\mathbb{K}_\delta} \varphi_\varepsilon(x-y)dy \\ &= \int_{\mathbb{R}^d} \varphi_\varepsilon(x-y)dy \\ &= 1. \end{aligned}$$

D'où le corollaire. □

5.5 Dualité entre L^p et $L^{p'}$

Le théorème principal de cette section est le théorème suivant, qui dit comment l'on peut représenter une forme linéaire continue sur les espaces L^p lorque p est réel.

Théorème 5.5.1. *Soient p un réel supérieur à 1 et B la forme bilinéaire définie par*

$$B \left\{ \begin{array}{ccc} L^{p'} \times L^p & \longrightarrow & \mathbb{K} \\ (g,f) & \longmapsto & \displaystyle\int_X f(x)g(x)d\mu(x). \end{array} \right.$$

Alors la forme bilinéaire B identifie le dual de $L^p(\Omega, d\mu)$ à $L^{p'}(\Omega, d\mu)$.

Avant de démontrer ce théorème, nous allons en donner une importante conséquence et faire quelques commentaires autour de ce résultat.

Corollaire 5.5.1. *Soit p dans $[1, \infty[$. On considère une suite $(g_n)_{n\in\mathbb{N}}$ bornée dans $L^{p'}$. Il existe une extraction ϕ et une fonction g dans $L^{p'}$ telles que*

$$\forall f \in L^p\,,\ \lim_{n\to\infty} \int_X g_{\phi(n)}(x)f(x)d\mu(x) = \int_X g(x)f(x)d\mu(x). \tag{5.17}$$

Démonstration. On applique le théorème 3.3.4 page 60 à la suite de formes linéaires $(L_n)_{n\in\mathbb{N}}$ sur L^p définie par

$$\langle L_n, f\rangle \stackrel{\text{déf}}{=} \int_X g_n(x)f(x)d\mu(x).$$

Ceci implique l'existence d'une fonction d'extraction ϕ et d'une forme linéaire L telle que

$$\forall f \in L^p\,,\ \lim_{n\to\infty} \int_X g_{\phi(n)}(x) f(x) d\mu(x) = \langle L, f\rangle.$$

Le théorème 5.5.1 ci-dessus assure alors l'existence d'une fonction g de $L^{p'}$ telle que

$$\forall f \in L^p\,,\ \lim_{n\to\infty} \int_X g(x) f(x) d\mu(x) = \langle L, f\rangle.$$

Le corollaire est ainsi démontré. □

Plaçons nous dans le cas X est un ouvert de $\mathbb{R}^d$ et μ la mesure de Lebesgue. Remarquons que comme p est fini, l'espace $C_c(\Omega)$ de fonctions continues est dense dans $L^p(\Omega)$. C'est un exercice laissé au lecteur que de démontrer qu'alors l'assertion (5.17) est équivalente à l'assertion

$$\forall \varphi \in C_c(\Omega)\,,\ \lim_{n\to\infty} \int_X g_{\phi(n)}(x) \varphi(x) d\mu(x) = \int_X g(x)\varphi(x) d\mu(x). \tag{5.18}$$

Nous allons observer que cette propriété est fausse lorsque p' vaut 1 c'est-à-dire lorsque $p$vaut l'infini. En effet, soit χ une fonction de $\mathcal{D}(\mathbb{R}^d)$ d'intégrale 1. Comme nous l'avons montré dans la démonstration du théorème 5.4.5, si $(\varepsilon_n)_{n\in\mathbb{N}}$ est une suite de réels positifs, nous avons

$$\forall \phi \lim_{n\to\infty} \int_{\mathbb{R}^d} \frac{1}{\varepsilon_n^d} \chi\Big(\frac{x}{\varepsilon_n}\Big) \varphi(x) dx = \varphi(0).$$

Or, la forme linéaire δ_0 définie sur $C_c(\Omega)$ par $\langle \delta_0, \varphi\rangle = \varphi(0)$ correspond à la mesure ponctuelle unité placée en 0. Ceci ne peut pas se représenter par une fonction g de L^1 à l'aide de la forme bilinéaire B car, si l'on a

$$\int_\Omega g(x)\varphi(x) dx = \varphi(0),$$

alors la fonction g est nulle sauf en 0, et donc nulle dx-presque partout.

Démonstration du théorème 5.5.1 lorsque p appartient à $]1,2]$. Nous avons déjà démontré ce théorème dans le cas où $p = 2$ au chapitre 4 sur les espaces de Hilbert. L'espace X est supposé être réunion dénombrable de parties K_n de mesure finie. On peut sans perte de généralité supposer la suite des K_n croissante au sens de l'inclusion. Soit Φ une forme linéaire sur $L^p(X, d\mu)$ et K une partie quelconque de mesure finie de X, on désigne par Φ_K la restriction de Φ à K c'est-à-dire la forme linéaire

$$\langle \Phi_K, f\rangle \stackrel{\text{déf}}{=} \langle \Phi, \mathbf{1}_K f\rangle.$$

Nous allons maintenant travailler dans l'espace $L^p(K, d\mu)$. Considérons une forme linéaire Φ sur $L^p(K, d\mu)$. Supposons démontré qu'il existe une unique fonction g_K sur K telle que l'on ait

$$\forall f \in L^p(K, d\mu)\,,\ \langle \Phi, f\rangle = B(g_K, f) = \int_K g_K(x) f(x) d\mu(x) \tag{5.19}$$

et telle que l'on ait

$$\|g_K\|_{L^{p'}} = \|\Phi_K\|_{(L^p)'}. \tag{5.20}$$

On considère alors une suite croissante $(K_n)_{n\in\mathbb{N}}$ de parties de X de mesure finie telle que la réunion des K_n soit l'ensemble X tout entier. Il existe donc une suite de fonctions $(g_n)_{n\in\mathbb{N}}$ vérifiant

$$\forall f \in L^p(K_n, d\mu)\,,\ \langle \Phi_{K_n}, f\rangle = B(g_n, f).$$

D'après le lemme 5.2.2, on sait que

$$g_n \in L^{p'} \quad \text{et que} \quad \|g_n\|_{L^{p'}} = \|\Phi_{K_n}\|_{(L^p)'} \leq \|\Phi\|_{(L^p)'}.$$

De plus, l'unicité de la fonction g_n vérifiant (5.19) et (5.20) assure que si $m \geq n$, alors

$$\mathbf{1}_{K_n} g_m = g_n$$

car la fonction $\mathbf{1}_{K_n} g_m$ vérifie les relations (5.19) et (5.20) pour la forme linéaire Φ_{K_n}. La suite $(|g_n(x)|^{p'})_{n\in\mathbb{N}}$ est donc une suite croissante de fonctions telle que

$$\sup_n \int_\Omega |g_n(x)|^{p'} d\mu(x) \leq \|\Phi\|^{p'}_{(L^p)'}.$$

Le théorème de convergence monotone assure que la fonction

$$g(x) = \lim_{n\to\infty} g_n(x) \in L^{p'} \quad \text{et que} \quad \|g\|_{L^{p'}} \leq \|\Phi\|_{(L^p)'}.$$

Le lemme 5.2.2 assure que $\|g\|_{L^{p'}} = \|\Phi\|_{(L^p)'}$. De plus, pour toute fonction f de L^p et pour tout entier n, on a

$$\langle \Phi, \mathbf{1}_{K_n} f\rangle = \int_X g(x) \mathbf{1}_{K_n} f(x) d\mu(x).$$

Le théorème de convergence dominée assure alors la conclusion de la démonstration. Mais il nous reste à démontrer les relations (5.19) et (5.20). D'après l'inégalité de Hölder, on sait que

$$\begin{aligned}
\|f\|_{L^p} &= \left(\int_K |f(x)|^p d\mu(x)\right)^{\frac{1}{p}} \\
&\leq \left(\int_K |f(x)|^{p\frac{2}{p}} d\mu(x)\right)^{\frac{p}{2}\times\frac{1}{p}} \mu(K)^{\frac{1}{p}\left(1-\frac{p}{2}\right)} \\
&\leq \|f\|_{L^2}\mu(K)^{\frac{1}{p}-\frac{1}{2}}.
\end{aligned}$$

Ainsi, la forme linéaire Φ apparaît comme une forme linéaire continue sur L^2. Il existe donc une fonction g appartenant à L^2 telle que

$$\forall f \in L^2\,,\ \langle \Phi, f\rangle = \int_X f(x) g(x) d\mu(x).$$

On en déduit que, pour toute fonction f de L^2, on a

$$\int_K |f(x)g(x)| d\mu(x) \leq \|\Phi\|_{(L^p)'} \|f\|_{L^p}.$$

Mais, d'après le corollaire 5.3.1, l'espace $L^\infty \cap L^p$ est dense dans L^p. Comme ici p est compris entre 1 et 2 et que K est compact, L^2 est dense dans L^p. D'après le lemme 5.2.2, on en déduit que g appartient à $L^{p'}$ et que $\|g\|_{L^{p'}} = \|\Phi\|_{(L^p)'}$. D'où les relations (5.19) et (5.20).

Démonstration du théorème 5.5.1 lorsque p appartient à $]2,\infty[$. Commençons par démontrer, pour p dans l'intervalle $]2,\infty[$ l'inégalité

$$\forall (a,b) \in \mathbb{R}^2\,,\ \frac{|a|^p + |b|^p}{2} \geq \left|\frac{a+b}{2}\right|^p + \frac{p(p-1)}{2^{p-2}}\left|\frac{a-b}{2}\right|^p. \tag{5.21}$$

Comme la fonction $t \mapsto |t|^p$ est deux fois continûment dérivable sur $\mathbb{R}$, la formule de Taylor avec reste intégral à l'ordre 2 entre a et $(a+b)/2$ et b et $(a+b)/2$ respectivement assure que l'on a

$$
\begin{aligned}
|a|^p &= \Big|\frac{a+b}{2}\Big|^p + p\,\frac{a-b}{2}\,\frac{a+b}{2}\,\Big|\frac{a+b}{2}\Big|^{p-2} \\
&\qquad + p(p-1)\Big(\frac{a-b}{2}\Big)^2\int_0^1 (1-\theta)\Big|\frac{a+b}{2}+\theta\frac{a-b}{2}\Big|^{p-2}d\theta \quad \text{et} \\
|b|^p &= \Big|\frac{a+b}{2}\Big|^p + p\,\frac{b-a}{2}\,\frac{a+b}{2}\,\Big|\frac{a-b}{2}\Big|^{p-2} \\
&\qquad + p(p-1)\Big(\frac{a-b}{2}\Big)^2\int_0^1 (1-\theta)\Big|\frac{a+b}{2}+\theta\frac{b-a}{2}\Big|^{p-2}d\theta\,.
\end{aligned}
$$

La demi-somme de ces deux inégalités implique que

$$
\begin{aligned}
\frac{|a|^p+|b|^p}{2} &= \Big|\frac{a+b}{2}\Big|^p \\
&+ \frac{p(p-1)}{2}\Big(\frac{a-b}{2}\Big)^2\int_0^1 (1-\theta)\bigg(\Big|\frac{a+b}{2}+\theta\frac{a-b}{2}\Big|^{p-2}+\Big|\frac{b-a}{2}+\theta\frac{a+b}{2}\Big|^{p-2}\bigg)d\theta\,.
\end{aligned}
$$

Si l'on a

$$
\Big|\frac{a+b}{2}+\theta\frac{a-b}{2}\Big| \leq \frac{1}{4}|b-a| \quad \text{et} \quad \Big|\frac{a+b}{2}+\theta\frac{b-a}{2}\Big| \leq \frac{1}{4}|b-a| \tag{5.22}
$$

alors on a

$$
\begin{aligned}
\theta|b-a| &= \Big|\theta\frac{a-b}{2}-\theta\frac{b-a}{2}\Big| \\
&= \Big|\frac{a+b}{2}+\theta\frac{a-b}{2}-\frac{a+b}{2}-\theta\frac{b-a}{2}\Big| \\
&\leq \Big|\frac{a+b}{2}+\theta\frac{a-b}{2}\Big| + \Big|\frac{a+b}{2}-\theta\frac{b-a}{2}\Big| \\
&\leq \frac{1}{2}|b-a|.
\end{aligned}
$$

Vu que a et b sont supposés différents (sinon il n'y a rien à démontrer), il en résulte que l'assertion (5.22) ne peut être satisfaite que si θ est inférieur ou égal à $1/2$. D'où il vient

$$
\begin{aligned}
\frac{|a|^p+|b|^p}{2} &\geq \Big|\frac{a+b}{2}\Big|^p + \frac{p(p-1)}{2}\Big(\frac{b-a}{2}\Big)^2\Big|\frac{b-a}{4}\Big|^{p-2}\int_{\frac{1}{2}}^1(1-\theta)d\theta \\
&\geq \Big|\frac{a+b}{2}\Big|^p + \frac{p(p-1)}{2^{p-2}}\Big|\frac{b-a}{2}\Big|^p
\end{aligned}
$$

ce qui démontre l'inégalité (5.21). La démonstration va suivre maintenant l'idée de la démonstration du théorème 4.3.2 page 70. Soit ℓ une forme linéaire continue sur $L^p(X, d\mu)$, considérons la fonction

$$
F_p \quad \left\{ \begin{array}{ccc} L^p(X,d\mu) & \longrightarrow & \mathbb{R} \\ f & \longmapsto & \dfrac{1}{p}\|f\|_{L^p}^p - \langle \ell, f\rangle \end{array}\right.
$$

La fonction F_p est minorée car

$$
F_p(f) \geq \frac{1}{p}\|f\|_{L^p}^p - \|\ell\|_{(L^p)'}\|f\|_{L^p}.
$$

Soit $m_p \stackrel{\text{def}}{=} \inf_{f\in L^p} F_p(f)$. On considère une suite minimisante $(f_n)_{n\in\mathbb{N}}$ c'est-à-dire une suite telle que

$$F_p(f_n) = m_p + \varepsilon_n \quad \text{avec} \quad \lim_{n\to\infty} \varepsilon_n = 0.$$

Nous allons démontrer que cette suite est de Cauchy. Appliquons tout d'abord l'inégalité (5.21) avec $a = f(x)$ et $b = g(x)$ puis intégrons par rapport à la mesure $d\mu$; on trouve que

$$\forall (f,g) \in \bigl(L^p(X,d\mu)\bigr)^2\,,\ \|f\|_{L^p}^p + \|g\|_{L^p}^p \geq 2\Big\|\frac{f+g}{2}\Big\|_{L^p}^p + \frac{p(p-1)}{2^{p-1}}\Big\|\frac{f-g}{2}\Big\|_{L^p}^p. \tag{5.23}$$

En appliquant cette inégalité avec $f = f_n$ et $g = f_m$ on peut écrire

$$\frac{1}{p}\|f_n\|_{L^p}^p + \frac{1}{p}\|f_m\|_{L^p}^p \geq \frac{2}{p}\Big\|\frac{f_n+f_m}{2}\Big\|_{L^p}^p + \frac{(p-1)}{2^{p-1}}\Big\|\frac{f_n-f_m}{2}\Big\|_{L^p}^p.$$

En retranchant $\langle \ell, f_n + f_m\rangle$ à cette inégalité on en déduit que

$$\begin{aligned}\frac{1}{p}\|f_n\|_{L^p}^p - \langle \ell, f_n\rangle + \|f_m\|_{L^p}^p - \langle \ell, f_m\rangle& \\ \geq 2\biggl(\Big\|\frac{f_n+f_m}{2}\Big\|_{L^p}^p - \langle \ell, \frac{f_n+f_m}{2}\rangle\biggr) &+ \frac{(p-1)}{2^{p-1}}\Big\|\frac{f_n-f_m}{2}\Big\|_{L^p}^p.\end{aligned}$$

Par définition de F_p, ceci s'écrit

$$\frac{(p-1)}{2^{p-1}}\Big\|\frac{f_n-f_m}{2}\Big\|_{L^p}^p \leq F_p(f_n) + F_p(f_m) - 2F_p\Bigl(\frac{f_n+f_m}{2}\Bigr)$$

La suite $(f_n)_{n\in\mathbb{N}}$ étant une suite minimisante, on en déduit que

$$\frac{(p-1)}{2^{p-1}}\Big\|\frac{f_n-f_m}{2}\Big\|_{L^p}^p \leq \varepsilon_n + \varepsilon_m$$

ce qui assure que la suite $(f_n)_{n\in\mathbb{N}}$ est de Cauchy. Soit f sa limite. La fonction F_p étant continue, on en déduit que $F_p(f) = m$ c'est-à-dire que la borne inférieure de F_p est un minimum.

Considérons maintenant, pour une fonction h de L^p donnée, la fonction g_h de $\mathbb{R}$ dans $\mathbb{R}$ définie par

$$g_h(t) \stackrel{\text{def}}{=} F_p(f + th).$$

D'après le théorème de dérivation sous l'intégrale, la fonction g_p est dérivable et

$$g_h'(t) = \int_X (f(x) + th(x))|f(x) + th(x)|^{p-2}h(x)d\mu(x) - \langle \ell, h\rangle.$$

Le point $t = 0$ est un minimum de la fonction g_h et donc la dérivée s'annule en ce point ce qui signifie que

$$\forall h \in L^p(X,d\mu)\,,\ \int_X f(x)|f(x)|^{p-2}h(x)d\mu(x) = \langle \ell, h\rangle$$

ce qui achève la démonstration du théorème 5.5.1. □

Il est possible d'interpréter en terme de dualité les résultats du théorème 5.4.5 sur les familles dites d'approximation de l'identité : l'espace $L^1(\Omega, dx)$ ne s'identifie, par l'application bilinéaire

$$B(f,g) = \int_\Omega f(x)g(x)dx,$$

au dual d'un espace normé de fonctions contenant les fonctions continues à support compact dans Ω. En effet, soit $(\varepsilon_n)_{n\in\mathbb{N}}$ une suite de nombres réels positifs convergeant vers 0, on considère la suite $(\varphi_{\varepsilon_n})_{n\in\mathbb{N}}$ d'approximation de l'identité définie dans l'énoncé du théorème 5.4.5. On a, pour toute fonction g continue à support compact,

$$\lim_{n\to\infty}\int_\Omega \varphi_{\varepsilon_n}(x)g(x)dx = g(0).$$

Chapitre 6

Le problème de Dirichlet

Introduction

Le but de ce chapitre est de résoudre le problème de Dirichlet sur un domaine borné de $\mathbb{R}^d$. Nous allons voir comment un problème qui se formule en terme de fonctions de classe C^1 va naturellement conduire à l'introduction du concept de dérivée faible (ou de quasi-dérivées). Ce concept conduira à celui plus général de dérivées au sens des distributions.

Pour formuler le problème de Dirichlet, nous allons tout dabord introduire les notations suivantes. Soient Ω un ouvert borné de $\mathbb{R}^d$, on désigne par $C_0^1(\Omega)$ désigne l'ensemble des fonctions continûment différentiables sur Ω et à support dans un compact dand $\overline{\Omega}$. Pour tout x de Ω et toute fonction u de classe C^1 sur Ω, on note

$$\nabla u(x) = \operatorname{grad} u(x) = \big(\partial_1 u(x), \cdots, \partial_d u(x)\big) \in \mathbb{R}^d.$$

De plus on note

$$|\nabla u(x)|^2 = \sum_{j=1}^d |\partial_j u(x)|^2 \quad \text{et} \quad \|\nabla u\|_{L^2(\Omega)}^2 \stackrel{\text{def}}{=} \int_\Omega |\nabla u(x)|^2 dx = \sum_{j=1}^d \int_\Omega |\partial_j u(x)|^2 dx.$$

En l'absence d'ambiguïté, on omet de noter la dépendance en Ω dans les normes. Soit f une fonction de $L^2(\Omega)$, on considère la fonctionnelle

$$F \left\{ \begin{array}{ccl} C_0^1(\Omega) & \longrightarrow & \mathbb{R} \\ u & \longmapsto & \dfrac{1}{2}\|\nabla u(x)\|_{L^2}^2 - \displaystyle\int_\Omega f(x)u(x)dx = \frac{1}{2}\sum_{j=1}^d \int_\Omega |\partial_j u(x)|^2 dx - \int_\Omega f(x)u(x)dx \end{array} \right.$$

et l'on cherche s'il existe un minimum pour F c'est-à-dire s'il existe une fonction u de $C_0^1(\Omega)$ telle que

$$F(u) = \inf_{v \in C_0^1(\Omega)} F(v).$$

Remarquons que F est une fonction continue (exercice : démontrez-le !).

Dans la première section de ce chapitre, nous présentons une approche très classique du problème, à savoir l'étude d'une suite minimisante dont on démontre qu'à extraction près, elle converge faiblement vers une fonction u et que les dérivées partielles converge elle aussi faiblement.

Ceci nous conduit à introduire dans la deuxième section le concept de quasi-dérivée d'une fonction de L^2. Introduit par le mathématicien français Jean Leray en 1934 pour résoudre le problème de l'existence globale de solutions pour l'équation régissant les écoulements des fluides visqueux incompressibles. Il s'agit de la première émancipation mathématiquement fondée du cadre des fonctions dérivables au sens classique.

Ce concept permet de définir à la troisième section l'espace de Sobolev qui est un espace de Hilbert adapté au problème. Une fois cet espace défini, la résolution du problème de Dirichlet $-\Delta u = f$ et u "nulle au bord" est une simple application du théorème de Riesz (voir le théorème 4.3.1 page 70). Ceci illustre la puissance de l'analyse fonctionnelle et aussi à quel point un problème semblant compliqué peut devenir simple et de résolution élégante pour autant que le cadre conceptuel soit bien choisi. Enfin, nous appliquons le théorème de diagonalisation des opérateurs autoadjoints compacts (voir le théorème 4.4.2 page 75) pour construire une "diagonalisation" de l'opérateur de Laplace Δ.

6.1 Une approche classique du problème

Il s'agit tout d'abord de démontrer que F est minorée. Ceci repose sur l'inégalité de Poincaré.

Proposition 6.1.1 (Inégalité de Poincaré)**.** *Soit Ω un ouvert borné. Il existe une constante C (dépendant de Ω) telle que,*

$$\forall u \in C_0^1(\Omega)\,,\ \|u\|_{L^2} \leq C\|\nabla u\|_{L^2}.$$

Démonstration. Soit u une fonction de $C_0^1(\Omega)$. Observons tout d'abord que la fonction qui vaut 0 en dehors de Ω et coïncide avec u sur Ω est une fonction C^1 à support compact sur $\mathbb{R}^d$ que l'on désigne toujours par u dans le but de simplifer les notations. Écrivons que

$$\forall (x_1, x') \in \Omega\,,\ u(x_1, x') = \int_{-\infty}^{x_1} \partial_1 u(y_1, x')dy_1.$$

D'après l'inégalité de Cauchy-Schwarz, on a

$$\forall (x_1, x') \in \Omega\,,\ |u(x_1, x')|^2 \leq \delta(\Omega) \int_{-\infty}^{+\infty} |\partial_1 u(y_1, x')|^2 dy_1.$$

En intégrant cette inégalité par rapport à la variable $x = (x_1, x')$, on trouve que

$$\int_\Omega |u(x_1, x')|^2 dx \leq \delta(\Omega)^2 \int_\Omega |\partial_1 u(y_1, x')|^2 dy_1 dx'.$$

L'inégalité de Poincaré est ainsi démontrée. □

On en déduit le corollaire suivant.

Corollaire 6.1.1. *La fonctionnelle F est minorée. De plus, on a*

$$\forall A > 0\,,\ \exists B \,/\, \|\nabla u\|_{L^2} \geq B \Longrightarrow F(u) \geq A. \tag{6.1}$$

Démonstration. Il suffit d'écrire que, d'après l'inégalité de Poincaré, on a

$$\begin{aligned} F(u) &\geq \frac{1}{2}\|\nabla u\|_{L^2}^2 - \|f\|_{L^2}\|u\|_{L^2} \\ &\geq \frac{1}{2}\|\nabla u\|_{L^2}^2 - C\|f\|_{L^2}\|\nabla u\|_{L^2} \\ &\geq \frac{1}{2}\left(\|\nabla u\|_{L^2} - C\|f\|_{L^2}\right)^2 - \frac{1}{2}C^2\|f\|_{L^2}^2. \end{aligned}$$

Donc $F(u)$ est minorée. De plus, si A est un réel strictement positif, en posant

$$B = \left(2A + C^2\|f\|_{L^2}^2\right)^{\frac{1}{2}} + C\|f\|_{L^2},$$

on conclut la démonstration du corollaire. □

On considère maintenant une suite $(u_n)_{n\in\mathbb{N}}$ de $C_0^1(\Omega)$ minimisante, c'est-à-dire une suite telle que

$$\lim_{n\to\infty} F(u_n) = m = \inf_{v\in C_0^1(\Omega)} F(v).$$

Comme on cherche à savoir si le minimum est atteint, on cherche une limite à cette suite. L'inégalité (6.1) du corollaire 6.1.1 implique que la suite $(\|\nabla u_n\|_{L^2})_{n\in\,\mathbb{N}}$ est bornée. L'inégalité de Poincaré implique que la suite $(u_n)_{n\in\mathbb{N}}$ est bornée dans $L^2(\Omega)$. Ainsi donc les suites $(u_n)_{n\in\mathbb{N}}$ et $(\partial_j u_n)_{n\in\mathbb{N}}$ sont des suites bornées de L^2. Le théorème du compacité faible 4.3.4 page 72 implique l'existence d'une fonction u et de fonctions $u^{(j)}$ telle que, après extraction, on ait

$$\operatorname*{limf}_{n\to\infty} u_n = u \quad \text{et} \quad \operatorname*{limf}_{n\to\infty} \partial_j u_n = u^{(j)}.$$

Supposons un instant que u soit une fonction de $C_0^1(\Omega)$ et considérons une fonction φ de $C_0^1(\Omega)$. Par définition de la convergence faible dans un espace de Hilbert, on a

$$\int_\Omega u^{(j)}(x)\varphi(x)dx = \lim_{n\to\infty}\int_\Omega \partial_j u_n(x)\varphi(x)dx.$$

Les fonctions u_n et φ étant des fonctions de $C_0^1(\Omega)$, on a, par intégration par parties,

$$\int_\Omega u^{(j)}(x)\varphi(x)dx = -\lim_{n\to\infty}\int_\Omega u_n(x)\partial_j\varphi(x)dx.$$

À nouveau d'après la définition de la convergence faible dans L^2, on en déduit que

$$\int_\Omega u^{(j)}(x)\varphi(x)dx = -\int_\Omega u(x)\partial_j\varphi(x)dx.$$

D'après le corollaire 5.4.3 page 102, on sait que $C_0^1(\Omega)$ est dense dans $L^2(\Omega)$. Ainsi donc, nous avons $u^{(j)} = \partial_j u$.

Il n'y a aucune raison pour que u soit toujours de classe C^1 ; c'est d'ailleurs faux en général. Cependant, le calcul ci-dessus met en lumière un lien entre les dérivées de u et les fonctions $u^{(j)}$. Ce lien va être fondé mathématiquement par la section suivante.

6.2 Le concept de quasi-dérivée

Voici ce dont il s'agit.

Définition 6.2.1. *Soit u une fonction de $L^2(\Omega)$. On dit que u admet une quasi-dérivée partielle dans la direction j si et seulement si il existe une constante C telle que*

$$\forall\varphi\in C_0^1(\Omega)\,,\ \left|\int_\Omega u(x)\partial_j\varphi(x)dx\right| \leq C\|\varphi\|_{L^2}$$

Remarque fondamentale D'après le théorème 5.3.1 page 91, l'espace $C^1_0(\Omega)$ est dense dans $L^2(\Omega)$. Le théorème de prolongement 1.2.4 page 17 implique que l'on peut prolonger la forme linéaire

$$\ell_j \left\{ \begin{array}{ccc} C^1_0(\Omega) & \longrightarrow & \mathbb{R} \\ \varphi & \longmapsto & -\displaystyle\int_\Omega u(x)\partial_j\varphi(x)dx. \end{array} \right.$$

à l'espace $L^2(\Omega)$ tout entier. Désignons par $\widetilde{\ell}_j$ cette forme linéaire. Le théorème 4.3.1 page 70 de représentation de Riesz assure l'existence d'une fonction $u^{(j)}$ de L^2 telle que

$$\forall f \in L^2(\Omega)\,, \int_\Omega u^{(j)}(x)f(x)dx = \langle \widetilde{\ell}_j, f\rangle.$$

En particulier pour $f = \varphi \in C^1_0(\Omega)$, ceci donne

$$\forall \varphi \in C^1_0(\Omega)\,, \int_\Omega u^{(j)}(x)\varphi(x)dx = -\int_\Omega u(x)\partial_j\varphi(x)dx.$$

On peut alors donner une définition plus complète de la quasi-dérivée.

Définition 6.2.2. *Soit u une fonction de $L^2(\Omega)$. On dit que u admet une quasi-dérivée partielle dans la direction j si et seulement si il existe une constante C telle que*

$$\forall \varphi \in C^1_0(\Omega)\,, \left|\int_\Omega u(x)\partial_j\varphi(x)dx\right| \leq C\|\varphi\|_{L^2}.$$

De plus, la fonction $u^{(j)}$ de $L^2(\Omega)$ telle que

$$\forall \varphi \in C^1_0(\Omega)\,, \int_\Omega u^{(j)}(x)\varphi(x)dx = -\int_\Omega u(x)\partial_j\varphi(x)dx.$$

est appelée la quasi-dérivée de u dans la $j$ième direction.

Ppur illuster cette notion, nous allons traiter l'exemple suivant.

Proposition 6.2.1. *Soit $\Omega =]-1, 1[$. On considère la fonction $u(x) = |x|^\alpha$ avec α dans $]1/2, 1]$. Cette fonction admet une quasi-dérivée qui est la fonction*

$$x \longmapsto \alpha \mathrm{sg}(x)|x|^{\alpha-1}.$$

Démonstration. Écrivons que, pour toute fonction φ de classe C^1 et à support compact dans l'intervalle $]-1, 1[$,

$$\int_{-1}^1 |x|^\alpha\varphi(x)dx = \lim_{\varepsilon\to 0} I_\varepsilon(\varphi) \quad \text{avec} \quad I_\varepsilon(\varphi) \stackrel{\text{déf}}{=} \int_{-1}^{-\varepsilon} (-x)^\alpha\varphi'(x)dx + \int_\varepsilon^1 x^\alpha\varphi'(x)dx.$$

Par intégration par parties, on trouve que

$$I_\varepsilon(\varphi) = \varepsilon^\alpha\big(\varphi(-\varepsilon) - \varphi(\varepsilon)\big) + \int_{-1}^{-\varepsilon} \alpha(-x)^{\alpha-1}\varphi(x)dx - \int_\varepsilon^1 \alpha x^{\alpha-1}\varphi(x)dx$$

Comme α est strictement supérieur à $1/2$, la fonction $|x|^{\alpha-1}$ est de carré intégrable. On peut donc passer à la limite dans les intégrales ci-dessus ce qui donne

$$\int_{-1}^1 |x|^\alpha\varphi(x)dx = -\int_{-1}^1 \alpha\mathrm{sg}(x)|x|^{\alpha-1}\varphi(x)dx.$$

La proposition est ainsi démontrée. □

La proposition suivante montre que l'on a bien généralisé le concept de dérivées tout au moins en ce qui concerne la dérivation des fonctions C^1.

Proposition 6.2.2. *Si u appartient à $C_0^1(\Omega)$, alors u admet des quasi-dérivées dans toutes les directions et les quasi-dérivées coïncident avec les dérivées partielles.*

Démonstration. Par intégration par parties et en utilisant l'inégalité de Cauchy-Schwarz, on a

$$\begin{aligned} \left|\int_\Omega u(x)\partial_j\varphi(x)dx\right| &= \left|\int_\Omega \partial_j u(x)\varphi(x)dx\right| \\ &\leq \|\partial_j u\|_{L^2}\|\varphi\|_{L^2} \\ &\leq C\|\partial_j u\|_{L^\infty}\|\varphi\|_{L^2}. \end{aligned}$$

Soit $u^{(j)}$ quasi-dérivée de u dans la $j$ième direction. Par définition, cette fonction de $L^2(\Omega)$ vérifie

$$\forall\varphi\in C_0^1(\Omega)\,,\ \int_\Omega u^{(j)}(x)\varphi(x)dx = -\int_\Omega u(x)\partial_j\varphi(x)dx.$$

De plus, par intégration par parties, on a

$$\forall\varphi\in C_0^1(\Omega)\,,\ \int_\Omega \partial_j u(x)\varphi(x)dx = -\int_\Omega u(x)\partial_j\varphi(x)dx.$$

Ceci implique alors que

$$\int_\Omega \left(\partial_j u - u^{(j)}\right)(x)\varphi(x)dx = 0$$

L'espace $C_0^1(\Omega)$ étant dense dans $L^2(\Omega)$, on en déduit que $u^{(j)} = \partial_j u$. D'où la proposition. □

Notation Dorénavant, la quasi-dérivée de u dans la $j$ième direction sera notée $\partial_j u$.

6.3 L'espace $H_0^1(\Omega)$ et le problème de Dirichlet

Nous allons dans cette section introduire l'espace de fonctions adapté au problème de minimisation introduit au début de ce chapitre.

Définition 6.3.1. *Soit Ω un ouvert de $\mathbb{R}^d$. On désigne par $H_0^1(\Omega)$ l'espace des fonctions u de $L^2(\Omega)$ telles qu'il existe une suite $(u_n)_{n\in\mathbb{N}}$ de $C_0^1(\Omega)$ telle que*

$$\lim_{n\to\infty}\|u_n - u\|_{L^2(\Omega)} = 0$$

et les suites $(\partial_j u_n)_{n\in\mathbb{N}}$ sont de Cauchy dans $L^2(\Omega)$.

Les propriétés de cet espace sont décrites par les propositions suivantes.

Proposition 6.3.1. *Si $u\in H_0^1(\Omega)$, alors u admet des quasi-dérivées dans toutes les directions.*

Démonstration. Soit $(u_n)_{n\in\mathbb{N}}$ une suite de fonctions de $C_0^1(\Omega)$ qui converge en norme $L^2(\Omega)$ vers u et telles que les suites $(\partial_j u_n)_{n\in\mathbb{N}}$ soient de Cauchy dans $L^2(\Omega)$. L'espace $L^2(\Omega)$ étant complet, il existe une fonction $u^{(j)}$ de $L^2(\Omega)$ telle que

$$\lim_{n\to\infty}\|\partial_j u_n - u^{(j)}\|_{L^2} = 0.$$

On peut alors écrire, grâce à une intégration par parties, que

$$\begin{aligned} -\int_\Omega u(x)\partial_j\varphi(x)dx &= -\lim_{n\to\infty}\int_\Omega u_n(x)\partial_j\varphi(x)dx \\ &= \lim_{n\to\infty}\int_\Omega \partial_j u_n(x)\varphi(x)dx \\ &= \int_\Omega u^{(j)}(x)\varphi(x)dx. \end{aligned}$$

Ceci démontre la proposition. □

Proposition 6.3.2. *Muni de la norme définie par*

$$N(u) \stackrel{\text{déf}}{=} \Big(\|u\|^2_{L^2(\Omega)} + \sum_{j=1}^d \|\partial_j u\|^2_{L^2(\Omega)}\Big)^{\frac{1}{2}},$$

l'espace $H^1_0(\Omega)$ *est un espace de Hilbert.*

Démonstration. Le fait que N soit une norme associée au produit scalaire

$$(u|v) \stackrel{\text{déf}}{=} \int_\Omega u(x)v(x)dx + \sum_{j=1}^d \int_\Omega \partial_j u(x)\partial_j v(x)dx$$

est un exercice facile laissé au lecteur. Démontrons que l'espace $H^1_0(\Omega)$ est complet. Considérons une suite de Cauchy $(u_n)_{n\in\mathbb{N}}$ d'éléments de $H^1_0(\Omega)$. L'espace $L^2(\Omega)$ étant complet, il existe des fonctions u et $(u^{(j)})_{1\leq j\leq d}$ telles que

$$\lim_{n\to\infty} \|u_n - u\|_{L^2(\Omega)} = \lim_{n\to\infty} \|\partial_j u_n - u^{(j)}\|_{L^2(\Omega)} = 0. \tag{6.2}$$

Par définition de $H^1_0(\Omega)$, pour tout n, il existe une fonction v_n de $C^1_0(\Omega)$ telle que

$$\|u_n - v_n\|_{L^2(\Omega)} \leq \frac{1}{n} \quad \text{et} \quad \|\partial_j u_n - \partial_j v_n\|_{L^2(\Omega)} \leq \frac{1}{n}\cdot$$

Ainsi donc, il existe une suite $(v_n)_{n\in\mathbb{N}}$ de $C^1_0(\Omega)$ telle que

$$\lim_{n\to\infty} \|v_n - u\|_{L^2(\Omega)} = \lim_{n\to\infty} \|\partial_j v_n - u^{(j)}\|_{L^2(\Omega)} = 0.$$

Donc $u \in H^1_0(\Omega)$ et $u^{(j)} = \partial_j u$. D'après (6.2), on a

$$\lim_{n\to\infty} N(u_n - u) = 0.$$

La proposition est alors démontrée. $\square$

Nous allons maintenant examiner les propriétés particulières de l'espace $H^1_0(\Omega)$ lorsque l'ouvert Ω est un ouvert borné de $\mathbb{R}^d$.

Proposition 6.3.3. *Muni de la norme définie par*

$$\|u\|^2_{H^1_0(\Omega)} \stackrel{\text{déf}}{=} \|\nabla u\|^2_{L^2(\Omega)} \stackrel{\text{déf}}{=} \Big(\sum_{j=1}^d \|\partial_j u\|^2_{L^2(\Omega)}\Big)^{\frac{1}{2}},$$

l'espace $H^1_0(\Omega)$ *est un espace de Hilbert.*

Démonstration. Par rapport à la proposition 6.3.2, la seule chose à démontrer est que $\|\cdot\|_{H^1_0(\Omega)}$ est une norme et qu'elle est équivalente à N. Il suffit pour cela de démontrer que, pour toute u de $H^1_0(\Omega)$, on a

$$\|u\|_{L^2(\Omega)} \leq C\|\nabla u\|_{L^2(\Omega)}.$$

Soit $(u_n)_{n\in\mathbb{N}}$ une suite de $C^1_0(\Omega)$ telle que

$$\lim_{n\to\infty} \|u_n - u\|_{L^2(\Omega)} = \lim_{n\to\infty} \|\partial_j u_n - \partial_j u\|_{L^2(\Omega)} = 0$$

pour tout j appartenant à $\{1, \cdots d\}$. L'inégalité de Poincaré implique que, pour tout n, on a

$$\|u_n\|_{L^2(\Omega)} \leq C\|\nabla u_n\|_{L^2(\Omega)}.$$

Par passage à la limite, on trouve que, pour toute fonction u de $H^1_0(\Omega)$,

$$\|u\|_{L^2(\Omega)} \leq C\|\nabla u\|_{L^2(\Omega)}.$$

$\square$

Nous avons la proposition suivante, dont la démonstration est laissée en exercice.

Proposition 6.3.4. *La fonctionnelle F se prolonge continûment à l'espace $H^1_0(\Omega)$ tout entier et l'on a*

$$m = \inf_{v\in C^1_0(\Omega)} F(v) = \inf_{v\in H^1_0(\Omega)} F(v)$$

Le problème de minimisation posé rentre alors dans le cadre abstrait suivant : soit $\mathcal{H}$ un espace de Hilbert, ℓ une forme linéaire continue sur $\mathcal{H}$, peut-on trouver le minimum de F définie par

$$F_\ell \left\{ \begin{array}{ccc} \mathcal{H} & \longrightarrow & \mathbb{R} \\ v & \longmapsto & \dfrac{1}{2}\|v\|^2_{\mathcal{H}} - \langle \ell, v\rangle. \end{array} \right.$$

Posé ainsi, ce problème se résout immédiatement en utilisant le théorème 4.3.2.

Remarque Revenons au problème initial. L'unique minimum de F définie sur $H^1_0(\Omega)$ est la fonction u de $H^1_0(\Omega)$ qui vérifie

$$\forall v \in H^1_0(\Omega)\,,\ \sum_{j=1}^d \int_\Omega \partial_j u(x)\partial_j v(x)dx = (u|v)_{H^1_0(\Omega)} = \int_\Omega f(x)v(x)dx = (f|v)_{L^2}. \qquad (6.3)$$

Supposons que la solution u trouvée soit en outre de classe C^2 sur Ω. Alors, pour toute fonction φ de classe C^2 et à support compact dans Ω, on a

$$\sum_{j=1}^d \int_\Omega \partial_j u(x)\partial_j \varphi(x)dx = \int_\Omega f(x)\varphi(x)dx = (f|v)_{L^2}.$$

Par intégration par parties, on trouve que

$$\sum_{j=1}^d \int_\Omega \partial_j u(x)\partial_j \varphi(x)dx = \int_\Omega f(x)\varphi(x)dx = -\int_\Omega \Delta u(x)\varphi(x)dx.$$

Lorsque u est seulement dans $H^1_0(\Omega)$, nous dirons que $-\Delta u = f$ "au sens faible" ou "au sens des distributions".

Le propriété de compacité suivante qui sera démontrée au chapitre suivant, est déterminante pour démontrer l'existence de la suite des valeurs propres.

Théorème 6.3.1 (de Rellich)**.** *Toute partie bornée de $H^1_0(\Omega)$ est d'adhérence compacte dans l'espace $L^2(\Omega)$.*

Nous allons maintenant donner un résultat qui décrit plus finement la structure du laplacien et son action sur $H^1_0(\Omega)$.

Théorème 6.3.2. *Il existe une suite croissante $(\lambda_j)_{j\in\mathbb{N}}$ de nombres réels strictement positifs qui tend vers l'infini et une base hilbertienne de l'espace $L^2(\Omega)$ notée $(e_j)_{j\in\mathbb{N}}$ telle que la suite $(\lambda_j^{-1}e_j)_{j\in\mathbb{N}}$ soit une base orthonormée de $H_0^1(\Omega)$ et telles que*

$$-\Delta e_j = \lambda_j^2 e_j \quad \text{"au sens faible"}$$

c'est-à-dire que, pour toute fonction de classe C^1 à support compact dans Ω, on a

$$\sum_{k=1}^{d}\int_\Omega \partial_k e_j(x)\partial_k\varphi(x)dx = \lambda_j^2\int_\Omega e_j(x)\varphi(x)dx.$$

Démonstration. Définissons l'opérateur B par

$$B\left\{\begin{array}{ccl} L^2 & \longrightarrow & H_0^1(\Omega)\subset L^2(\Omega)\\ f & \longmapsto & u\end{array}\right.$$

tel que u soit la solution dans $H_0^1(\Omega)$ du problème de Dirichlet. L'opérateur B est continu de $L^2(\Omega)$ dans $H_0^1(\Omega)$. D'après le théorème 6.3.1 ci-dessus, l'opérateur B est un opérateur compact de $L^2(\Omega)$ dans $L^2(\Omega)$. Démontrons que B est un opérateur autoadjoint. D'après (6.3), on a, pour tout f dans L^2,

$$\forall v\in H_0^1(\Omega)\,,\ (Bf|v)_{H_0^1(\Omega)} = (f|v)_{L^2}.$$

En appliquant cette relation avec $v = Bg$ pour un g quelconque dans L^2, on en déduit que

$$(f|Bg)_{L^2} = (Bf|Bg)_{H_0^1(\Omega)} \quad \text{et donc que} \quad (f|Bg)_{L^2} = (f|Bg)_{L^2}.$$

L'opérateur B est donc autoadjoint. Démontrons qu'il est injectif. Considérons une fonction f de L^2 telle que alors $Bf = 0$. D'après la relation (6.3), ceci signifie que

$$\forall v\in H_0^1(\Omega)\,,\int_\Omega f(x)v(x)dx = 0.$$

Le corollaire (5.4.2) affirme la densité de $\mathcal{D}(\Omega)$, donc de $H_0^1(\Omega)$ qui le contient, dans $L^2(\Omega)$. On en déduit que si $Bf = 0$, alors

$$\forall g\in L^2\,,\ \int_\Omega f(x)v(x)dx = 0$$

ce qui implique que $f = 0$. Enfin observons que si λ est une valeur propre non nulle, elle est positive. En effet, il existe alors une fonction f_λ de L^2 non nulle telle que

$$\lambda\|f\|_{L^2}^2 = (Bf|f)_{L^2} = \|f\|_{H_0^1}^2.$$

Le théorème 4.4.2 page 75 assure l'existence d'une suite $(\mu_j)_{j\in\mathbb{N}}$ de réels décroissante tendant vers 0 et d'une base hilbertienne $(e_j)_{j\in\mathbb{N}}$ de $L^2(\Omega)$ telles que

$$Be_j = \mu_j^2 e_j.$$

En application le laplacien, on trouve que

$$e_j = -\Delta Be_j = \mu_j^2(-\Delta)e_j,$$

ce qui s'écrit

$$-\Delta e_j = \frac{1}{\mu_j^2}e_j\,.$$

Pour démontrer que la suite $(\lambda_j^{-1}e_j)_{j\in\mathbb{N}}$ est une base orthonormée de $H_0^1(\Omega)$, il suffit d'observer que l'on a

$$\lambda_j^2(e_j|e_k)_{L^2} = (e_j|e_k)_{H_0^1}.$$

Le théorème 6.3.2 est démontré. □

Chapitre 7

La transformation de Fourier

Introduction

Bien que bref, ce chapitre est fondamental. La transformation de Fourier est une opération très générale des fonctions intégrables par rapport à la mesure de Haar (c'est-à-dire une mesure invariante par translation) sur un groupe commutatif locallement compact G. Nous nous contenterons ici d'étudier le cas de l'espace $\mathbb{R}^d$ muni de la mesure de Lebesgue.

Dans la première section de ce chapitre, nous définissons la transformation de Fourier d'une fonction intégrable et établissons les principales propriétés de la transformation, à savoir

— elle transforme la dérivation en multiplication et réciproquement,

— elle transforme le produit de convolution en produit numérique

Outre ses propriétés de base, nous présentons quelques exemples de calcul de transformées de Fourier en particulier l'exemple fondamental de la transformée des gaussiennes.

Dans la seconde section, nous démontrons le théorème d'inversion de Fourier qui affirme que lorsque la transformée de Fourier d'une fonction intégrable est intégrable, alors on peut retrouver la fonction à partir de sa transformation de Fourier. On déduit de ce théorème le théorème de Fourier-Plancherel qui établit que l'on peut prolonger la transformation de Fourier à l'espace des fonctions de carré intégrable ; ce prologement est, à une constante près, une isométrie de l'espace de fonctions de carré intégrable.

La troisième section est une application de ces résultats à la démonstration du théorème de Rellich utilisé au chapitre précédent (voir le théorème 6.3.1 page 115).

7.1 La transformée de Fourier sur $L^1(\mathbb{R}^d)$

Définition 7.1.1. *Soit f une fonction de $L^1(\mathbb{R}^d)$. On appelle transformée de Fourier de f et l'on note $\widehat{f}$ ou bien $\mathcal{F}(f)$ la fonction définie par*

$$\mathcal{F}(f) \quad \begin{cases} \mathbb{R}^d & \longrightarrow \quad \mathbb{C} \\ \xi & \longmapsto \quad \displaystyle\int_{\mathbb{R}^d} e^{-i\langle \xi, x\rangle} f(x)dx \end{cases} \qquad \text{avec} \quad \langle \xi, x\rangle \stackrel{\text{déf}}{=} \sum_{j=1}^{d} \xi_j x_j$$

Nous allons donner quelques exempes de calculs de transformées de Fourier.

Proposition 7.1.1. *Soit a un nombre réel strictement positif. Nous avons les résultats suivants :*

$$\mathcal{F}(\mathbf{1}_{[-a,a]})(\xi) = \frac{2\sin a\xi}{\xi}\,,$$
$$\mathcal{F}\big(e^{-a|\cdot|}\mathbf{1}_{\mathbb{R}_\pm}\big)(\xi) = \frac{1}{a \pm i\xi} \quad \textit{et donc} \quad \mathcal{F}\big(e^{-a|\cdot|}\big)(\xi) = \frac{2a}{a^2+\xi^2}\,,$$
$$\mathcal{F}\Big(\frac{a}{a^2+|\cdot|^2}\Big)(\xi) = \pi e^{-a|\xi|} \quad \textit{et}$$
$$\mathcal{F}\big(e^{-a|\cdot|^2}\big)(\xi) = \Big(\frac{\pi}{a}\Big)^{\frac{d}{2}} e^{-\frac{|\xi|^2}{4a}}.$$

Démonstration. Par définition de la transformation de Fourier, on a

$$\begin{aligned}
\mathcal{F}(\mathbf{1}_{[-a,a]})(\xi) &= \int_{-a}^{a} e^{-i\xi x}dx \\
&= -\frac{1}{i\xi}\big(e^{-i\xi a} - e^{i\xi a}\big) \\
&= \frac{2\sin a\xi}{\xi}\,.
\end{aligned}$$

À nouveau par définition de la transformation de Fourier, on a

$$\begin{aligned}
\mathcal{F}\big(e^{-a|\cdot|}\mathbf{1}_{\mathbb{R}_+}\big)(\xi) &= \int_{0}^{\infty} e^{-(i\xi+a)x}dx \\
&= \frac{1}{a+i\xi}\,.
\end{aligned}$$

De même on a $\mathcal{F}\big(e^{-a|\cdot|}\mathbf{1}_{\mathbb{R}_+}\big)(\xi) = \dfrac{1}{a-i\xi}\cdot$ Par addtion on a déduit que

$$\mathcal{F}\big(e^{-a|\cdot|}\big)(\xi) = \frac{2a}{a^2+\xi^2}\,\cdot$$

Toujours par définition de la transformation de Fourier, on a

$$\begin{aligned}
\mathcal{F}\big(e^{-a|\cdot|^2}\big)(\xi) &= \lim_{R\to\infty}\int_{-R}^{R} e^{-i\xi x - ax^2}dx \\
&= e^{-\frac{\xi^2}{4a}} \lim_{R\to\infty}\int_{-R}^{R} e^{-a\left(x+i\frac{\xi}{2a}\right)^2}dx.
\end{aligned}$$

La fonction $z \mapsto e^{-az^2}$ étant une fonction holomorphe sur $\mathbb{C}$, on a

$$-\int_{-R}^{R} e^{-a\left(x+i\frac{\xi}{2a}\right)^2}dx - \int_{0}^{\frac{\xi}{2a}} e^{-a(R-iy)^2}dy + \int_{-R}^{R} e^{-ax^2}dx + \int_{0}^{\frac{\xi}{2a}} e^{-a(R+iy)^2}dy = 0$$

En faisant tendre R vers l'infini, on trouve que

$$\int_{-\infty}^{\infty} e^{-a\left(x+i\frac{\xi}{2a}\right)^2}dx = \int_{-\infty}^{\infty} e^{-ax^2}dx$$

Il est classique que

$$\int_{\mathbb{R}} e^{-ax^2}dx = \Big(\frac{\pi}{a}\Big)^{\frac{1}{2}}.$$

Le résultat est démontré en dimension 1. En dimension d quelconque, il suffit d'observer que

$$e^{-|x|^2} = \prod_{j=1}^{d} e^{-x_j^2}$$

et que, si les f_j sont des fonctions de $L^1(\mathbb{R})$, on a

$$\mathcal{F}(f_1 \otimes \cdots \otimes f_d) = \widehat{f}_1 \otimes \cdots \otimes \widehat{f}_d$$

ce qui conclut la démonstration de cette proposition $\square$

Remarquons que le théorème de continuité sous l'intégrale implique que la transformation de Fourier est une fonction continue (et bien sûr bornée) sur $\mathbb{R}^d$.

L'une des caractéristiques de la transformation de Fourier est, comme le montre le théorème suivant, de transformer la multiplication en dérivation et réciproquement.

Théorème 7.1.1. *Soit f une fonction sur $\mathbb{R}^d$ telle que $(1+|x|)f(x)$ définisse une fonction intégrable sur $\mathbb{R}^d$. Alors la transformée de Fourier de f est une fonction de classe C^1 et*

$$\partial_{\xi_j}\mathcal{F}(f) = -i\mathcal{F}(M_j f) \quad \text{avec} \quad (M_j f)(x) \stackrel{\text{déf}}{=} x_j f(x). \tag{7.1}$$

Soit f une fonction de classe C^1 telle que f ainsi que ses dérivées partielles soient intégrables sur $\mathbb{R}^d$. Alors on a

$$\mathcal{F}(\partial_{x_j} f) = iM_j\mathcal{F}(f). \tag{7.2}$$

Démonstration. Les deux résultats reposent sur le fait que

$$\partial_{\xi_j} e^{-i\langle\xi,x\rangle} = -ix_j e^{-i\langle\xi,x\rangle} \quad \text{et} \quad \partial_{x_j} e^{-i\langle\xi,x\rangle} = -i\xi_j e^{-i\langle\xi,x\rangle}. \tag{7.3}$$

Dans le premier cas, l'hypothèse et le théorème de dérivation sous l'intégrale assurent que f admet des dérivées partielles ainsi que la relation (7.1). Le second nécessite une procédure d'approximation. Soient χ une fonction de classe C^1 à support compact vaillant 1 en 0 et $(\varepsilon_n)_{n\in\mathbb{N}}$ une suite de réels strictement positifs tendant vers 0. En utilisant (7.3), et en faisant une intégration par parties, on trouve que

$$\begin{aligned} \xi_j\mathcal{F}(\chi(\varepsilon_n\cdot)f)(\xi) &= -i\int_{\mathbb{R}^d} \partial_{x_j}\big(e^{-i\langle\xi,x\rangle}\big)\chi(\varepsilon_n x)f(x)dx \\ &= -i\int_{\mathbb{R}^d} e^{-i\langle\xi,x\rangle}\partial_{x_j}\big(\chi(\varepsilon_n x)f(x)\big)dx. \end{aligned}$$

La formule de Leibnitz puis le théorème de convergence dominée assure la seconde formule. $\square$

Ce théorème permet de redémontrer la formule de la proposition 7.1.1 sur la transformée de Fourier des gausiennes.

Démonstration. D'après la proposition 7.1.2 appliqué au cas où l'application linéaire A est une homothétie de rapport $\sqrt{a}$, il suffit de traiter le cas où $a=1$. Supposons tout d'abord que la dimension d vaille 1 et posons

$$g(\xi) \stackrel{\text{déf}}{=} \int_{\mathbb{R}} e^{-i\xi x}e^{-x^2}dx.$$

La relation (7.1) du théorème 7.1.1 implique que la fonction g est dérivable et que l'on peut écrire

$$\begin{aligned} g'(\xi) &= \int_{\mathbb{R}} -ixe^{-i\xi x}e^{-x^2}dx \\ &= \int_{\mathbb{R}} ie^{-i\xi x}\frac{1}{2}(e^{-x^2})'dx. \end{aligned}$$

En appliquant la relation (7.2) du théorème 7.1.1 on trouve que

$$g'(\xi) = -\frac{\xi}{2}g(\xi).$$

L'équation différentielle est aisée à résoudre. On a $g(\xi) = g(0)e^{-\frac{\xi^2}{4}}$. Il est classique que

$$g(0) = \int_{\mathbb{R}} e^{-x^2}dx = \sqrt{\pi}.$$

Le résultat est démontré en dimension 1. En dimension d quelconque, il suffit d'observer que

$$e^{-|x|^2} = \prod_{j=1}^{d} e^{-x_j^2}$$

et que, si les f_j sont des fonctions de $L^1(\mathbb{R})$, on a

$$\mathcal{F}(f_1 \otimes \cdots \otimes f_d) = \widehat{f_1} \otimes \cdots \otimes \widehat{f_d}$$

ce qui conclut la démonstration de cette proposition □

Nous allons maintenant donner un corollaire qui décrit les propriétés de la transformée de Fourier des fonctions indéfiniment différentiables à support compact. Sa démonstration (omise) n'est que l'application répétée du théorème ci-dessus.

Corollaire 7.1.1. *Soit f une fonction de $\mathcal{D}(\mathbb{R}^d)$. Sa transformation de Fourier $\widehat{f}$ est indéfiniment différentiable et vérifie*

$$\forall N \in \mathbb{N}\,,\ \forall \alpha \in \mathbb{N}^d\,,\ \sup_{\xi \in \mathbb{R}^d} (1+|\xi|)^N |\partial^\alpha \widehat{f}(\xi)| < \infty. \quad \text{avec} \quad \partial^\alpha \stackrel{\text{déf}}{=} \partial_{x_1}^{\alpha_1} \cdots \partial_{x_d}^{\alpha_d}.$$

Nous allons maintenant étudier l'influence des transformations linéaires sur la transformation de Fourier.

Proposition 7.1.2. *Soient f une fonction de $L^1(\mathbb{R}^d)$ et A une application linéaire inversible de $\mathbb{R}^d$ dans $\mathbb{R}^d$. On a*

$$\mathcal{F}(f \circ A)(\xi) = |\det A|^{-1} \widehat{f} \circ {}^tA^{-1}.$$

Démonstration. Faisons le changement de variable $y = Ax$ dans l'intégrale définissant la transformation de Fourier ; ceci donne

$$\mathcal{F}(f \circ A)(\xi) = |\det A|^{-1} \int_{\mathbb{R}^d} e^{-i\langle \xi, A^{-1}x\rangle} f(x)dx.$$

Par définition de la transposée, nous avons $\langle \xi, A^{-1}x\rangle = \langle {}^tA^{-1}\xi, x\rangle$. D'où la proposition. □

Du théorème 7.1.1 et de la proposition 7.1.2 ci-dessus, nous allons déduire la proposition suivante, qui explicite le calcul de la transformée de Fourier d'une fonction gaussienne. Ce résultat jouera un rôle crucial dans la démonstration de la formule d'inversion de Fourier.

Proposition 7.1.3. *Soit a un réel strictement positif. On a alors*

$$\mathcal{F}\left(e^{-a|\cdot|^2}\right)(\xi) = \left(\frac{\pi}{a}\right)^{\frac{d}{2}} e^{-\frac{|\xi|^2}{4a}}.$$

La proposition suivante décrit partiellement les rapports entre transformée de Fourier et convolution.

Proposition 7.1.4. *Soient f et g deux fonctions de $L^1(\mathbb{R}^d)$. On a*

$$\mathcal{F}(f\star g) = \widehat{f}\,\widehat{g}.$$

Démonstration. La fonction F_ξ définie de $\mathbb{R}^d\times\mathbb{R}^d$ dans $\mathbb{C}$ par

$$F_\xi(x,y) \stackrel{\text{déf}}{=} e^{-i\langle\xi,x\rangle} f(x-y)g(y) = e^{-i\langle\xi,x-y\rangle} f(x-y)e^{-i\langle\xi,y\rangle}g(y)$$

est, pour tout ξ dans $\mathbb{R}^d$, une fonction de $L^1(\mathbb{R}^d\times\mathbb{R}^d)$. Le théorème de Fubini assure que

$$\int_{\mathbb{R}^d} e^{-i\langle\xi,x\rangle}\left(\int_{\mathbb{R}^d} f(x-y)g(y)dy\right)dx = \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} e^{-i\langle\xi,x-y\rangle} f(x-y)dx\right)e^{-i\langle\xi,x\rangle}g(y)dy,$$

ce qui signife exactement que la transformée de Fourier du produit de convolution est le produit (usuel) des transformées de Fourier. □

La transformation de Fourier est symétrique au sens de la proposition suivante.

Proposition 7.1.5. *Pour tout couple (f,g) de fonctions de $L^1(\mathbb{R}^d)$, nous avons*

$$\int_{\mathbb{R}^d}\widehat{f}(x)g(x)dx = \int_{\mathbb{R}^d} f(\xi)\widehat{g}(\xi)d\xi.$$

Démonstration. Le résulte découle immédiatement du théorème de Fubini appliqué à la fonction intégrable sur $\mathbb{R}^d\times\mathbb{R}^d$ définie par $(x,\xi)\longmapsto f(\xi)e^{-i\langle\xi,x\rangle}g(x)$. □

7.2 La formule d'inversion et le théorème de Fourier-Plancherel

Nous allons maintenant démontrer le théorème fondamental de l'analyse de Fourier.

Théorème 7.2.1. *Soit f une fonction $L^1(\mathbb{R}^d)$ dont la transformée est aussi une fonction de $L^1(\mathbb{R}^d)$. Alors la fonction f est continue et l'on a*

$$\forall x\in\mathbb{R}\,,\ f(x) = (2\pi)^{-d}\int_{\mathbb{R}^d} e^{i\langle\xi,x\rangle}\widehat{f}(\xi)d\xi \tag{7.4}$$

ainsi que la relation de Fourier-Plancherel

$$\|\widehat{f}\|_{L^2}^2 = (2\pi)^d\|f\|_{L^2}^2. \tag{7.5}$$

Remarques

— On peut écrire ce résultat sous la forme

$$\mathcal{F}^2 f = (2\pi)^d\check{f} \quad \text{avec} \quad \check{f}(x) \stackrel{\text{déf}}{=} f(-x). \tag{7.6}$$

— Ce théorème signifie qu'une fonction f dans $\mathcal{S}$ s'écrit comme une superposition d'oscillations – les fonctions $e^{-i\langle\xi,x\rangle}$ –, la transformée de Fourier apparaissant alors comme la densité d'oscillations.

Démonstration du théorème 7.2.1 D'après la proposition 7.1.3, le théorème est déjà démontré pour les fonctions gaussiennes. Nous allons bien sûr utiliser ce résultat. Posons

$$G(x) \stackrel{\text{déf}}{=} \pi^{-\frac{d}{2}} e^{-|x|^2} \quad \text{et} \quad G_\varepsilon(x) = \frac{1}{\varepsilon^d} G\Big(\frac{x}{\varepsilon}\Big)\cdot$$

La proposition 7.1.3 nous dit que $\widehat{G}_\varepsilon(\xi) = e^{-\frac{\varepsilon^2|\xi|^2}{4}}$. D'après le théorème de convergence dominée, on déduit de l'appartenance de $\widehat{f}$ à L^1 que

$$(2\pi)^{-d} \int_{\mathbb{R}^d} e^{i\langle \xi, x\rangle} \widehat{f}(\xi) d\xi = \lim_{\varepsilon \to 0} (2\pi)^{-d} \int_{\mathbb{R}^d} e^{i\langle \xi, x\rangle} \widehat{G}_\varepsilon(\xi) \widehat{f}(\xi) d\xi. \tag{7.7}$$

Comme la fonction f appartient à L^1, on sait que $\widehat{G}_\varepsilon \otimes f$ est une fonction de $L^1(\mathbb{R}^d \times \mathbb{R}^d)$. D'après le théorème de Fubini, on a

$$(2\pi)^{-d} \int_{\mathbb{R}^d} e^{i\langle \xi, x\rangle} \widehat{G}_\varepsilon(\xi) \widehat{f}(\xi) d\xi \quad = \quad (2\pi)^{-d} \int_{\mathbb{R}^d \times \mathbb{R}^d} e^{i\langle \xi, x\rangle} \widehat{G}_\varepsilon(\xi) e^{-i\langle \xi, y\rangle} f(y) d\xi dy.$$

La proposition 7.1.3 appliquée avec $a = \varepsilon^2/4$ affirme en particulier que

$$(2\pi)^{-d} \int_{\mathbb{R}^d} e^{i\langle \xi, y-x\rangle} \widehat{G}_\varepsilon(\xi) d\xi = (2\pi)^{-d} \Big(\frac{4\pi}{\varepsilon^2}\Big)^{\frac{d}{2}} e^{-\frac{|x-y|^2}{\varepsilon^2}} = G_\varepsilon(x-y).$$

Ainsi donc on trouve que

$$\forall \varepsilon > 0\,,\ (G_\varepsilon \star f)(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{i(x|\xi)} \widehat{G}_\varepsilon(\xi) \widehat{f}(\xi) d\xi.$$

Le théorème 5.4.5 page 100 dit que si $(\varepsilon_n)_{n\in\mathbb{N}}$ est une suite de réels strictement positifs tendant vers 0, alors $\lim_{n\to\infty} \|G_{\varepsilon_n} \star f - f\|_{L^1(\mathbb{R}^d)} = 0$. Le théorème 5.1.2 page 83 assure que qu'à une extraction près que nous omettons de noter, pour presque tout x dans $\mathbb{R}^d$,

$$\lim_{n\to\infty} (G_{\varepsilon_n} \star f)(x) = f(x).$$

L'assertion (7.7) assure alors que pour presque tout x de $\mathbb{R}^d$,

$$f(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{i\langle \xi, x\rangle} \widehat{f}(\xi) d\xi.$$

Le théorème de continuité sous l'intégrale affirmant que le terme de droite de l'égalité ci-dessus est continue permet de conclure la démonstration de (7.4) et du fait que la fonction f est alors continue.

Pour démontrer (7.5), observons tout d'abord que le fait que f et $\widehat{f}$ soit dans $L^1(\mathbb{R}^d)$ implique que f est une fonction bornée et ainsi donc f et $\widehat{f}$ sont toutes deux de carré intégrable. En remarquant que $\overline{\mathcal{F}(f)} = \mathcal{F}(\overline{\check{f}})$ et en utilisant la proposition 7.1.5, on peut écrire

$$\begin{aligned}(2\pi)^{-d} \|\widehat{f}\|_{L^2}^2 &= \int_{\mathbb{R}^d} (2\pi)^{-d} (\mathcal{F}f)(\xi) (\mathcal{F}(\overline{\check{f}})(\xi) d\xi \\ &= \int_{\mathbb{R}^d} (2\pi)^{-d} (\mathcal{F}^2 f)(x) \overline{f}(-x) dx.\end{aligned}$$

La formule d'inversion de Fourier telle qu'exprimée dans (7.6) assure alors que

$$(2\pi)^{-d} \|\widehat{f}\|_{L^2}^2 = \int_{\mathbb{R}^d} f(-x) \overline{f}(-x) dx = \|f\|_{L^2}^2$$

ce qui conclut la démonstration du théorème fondamental de l'analyse de Fourier. □

Ce théorème, fondement de l'analyse de Fourier, a d'innombrables applications. Donnons tout d'abord deux corollaires.

Corollaire 7.2.1. *La transformation de Fourier s'étend en une application linéaire continue inversible de $L^2(\mathbb{R}^d)$ dans $L^2(\mathbb{R}^d)$ qui vérifie*

$$\forall f \in L^2(\mathbb{R}^d)\,,\ \|\widehat{f}\|_{L^2(\mathbb{R}^d)} = (2\pi)^{\frac{d}{2}}\|f\|_{L^2(\mathbb{R}^d)}.$$

Démonstration. La transformation de Fourier est bien définie sur $\mathcal{D}(\mathbb{R}^d)$ qui d'après le corollaire 5.4.2 page 102 est dense dans $L^2(\mathbb{R}^d)$. Le corollaire 7.1.1 nous autorise à appliquer la relation de Fourier-Plancherel (7.5) aux fonctions de $\mathcal{D}(\mathbb{R}^d)$. Le théorème de prolongement 1.2.4 page 17 permet d'étendre $\mathcal{F}$ en une application linéaire continue de $L^2(\mathbb{R}^d)$ dans lui-même, la relation de Fourier-Plancherel s'étendant bien sûr à tout l'espace L^2. La seule chose restant à démontrer que la surjectivité de $\mathcal{F}$. La proposition 7.1.5 implique en particulier que

$$\forall (f,g) \in \mathcal{D}(\mathbb{R}^d)\,,\ (\mathcal{F}f|g)_{L^2(\mathbb{R}^d)} = (f|\check{\mathcal{F}}g)_{L^2(\mathbb{R}^d)}. \tag{7.8}$$

Pour démontrer que $\mathcal{F}$ est surjective, nous allons tout d'abord observer que l'identité de Fourier-Plancherel implique que $\mathcal{F}(L^2)$ est fermé, puis nous allons démontrer que $\mathcal{F}(L^2)$ est dense en utilisant la critère 4.2.2 page 68 qui affirme qu'il suffit de démontrer que $\mathcal{F}(L^2)^\perp$ est réduit à la fonction nulle. En effet, si g est une fonction de $L^2(\mathbb{R}^d)$ telle que

$$\forall f \in L^2(\mathbb{R}^d)\,,\ (\mathcal{F}(f)|g)_{L^2(\mathbb{R}^d)} = 0,$$

alors la relation (7.8) étendue à tout L^2 implique que $\mathcal{F}g$ est nulle et donc aussi g puisque $\mathcal{F}$ est injective. □

Corollaire 7.2.2. *Soient f et g deux fonctions de $L^1(\mathbb{R}^d)$ dont la transformée de Fourier aussi dans $L^1(\mathbb{R}^d)$. On a*

$$\mathcal{F}(fg) = (2\pi)^{-d}\widehat{f} \star \widehat{g}.$$

Démonstration. Appliquons la proposition 7.1.4 à $\widehat{f}$ et $\widehat{g}$. Ceci donne

$$\mathcal{F}(\widehat{f} \star \widehat{g}) = \mathcal{F}^2 f \mathcal{F}^2 g = (2\pi)^{2d}\check{f}\check{g}.$$

Ceci peut s'écrire

$$(\widehat{f} \star \widehat{g}) = (2\pi)^{2d}\mathcal{F}^{-1}(\check{f}\check{g})$$

La formule d'inversion de Fourier du théorème 7.2.1 assure le résultat. □

7.3 Démonstration du théorème de Rellich

Le but de cette section est de démontrer le théorème de compacité 6.3.1 dit de Rellich que nous rapellons. La démonstration n'est présenté ici qu'à titre culturelle.

Théorème 7.3.1. *Toute partie bornée de $H_0^1(\Omega)$ est d'adhérence compace dans $L^2(\Omega)$.*

Démonstration. Nous allons démontrer que, pour tout réel strictement positif α, on peut recouvrir une partie bornée A de $H_0^1(\Omega)$ par un nombre fini de boules de rayon α pour la norme L^2. Pour ce faire, commençons par démontrer que si χ est une fonction de $\mathcal{D}(B(0,1))$ d'intégrale 1, alors

$$\Big\|\frac{1}{\varepsilon^d}\chi\Big(\frac{\cdot}{\varepsilon}\Big) \star a - a\Big\|_{L^2} \leq C_\chi\,\varepsilon\|\nabla a\|_{L^2} \tag{7.9}$$

où C_χ est une constante dépendant de χ. D'après la relation de Fourier-Plancherel (théorème 7.2.1 page 121 et d'après le théorème 8.2.5 on a

$$\begin{aligned}\Big\|\frac{1}{\varepsilon^d}\chi\Big(\frac{\cdot}{\varepsilon}\Big)\star a-a\Big\|_{L^2} &= (2\pi)^{\frac{d}{2}}\Big\|\mathcal{F}\Big(\frac{1}{\varepsilon^d}\chi\Big(\frac{\cdot}{\varepsilon}\Big)\star a-a\Big)\Big\|_{L^2}\\ &= (2\pi)^{\frac{d}{2}}\big\|\widehat{\chi}(\varepsilon\cdot)\widehat{a}-\widehat{a}\big\|_{L^2}.\end{aligned}$$

L'inégalité des accroissements finis et le théorème 7.1.1 implique que

$$\begin{aligned}|\widehat{\chi}(\zeta)-1| &\leq |\zeta|\|D\widehat{\chi}\|_{L^\infty}\\ &\leq |\zeta|\max_{1\leq j\leq d}\|x_j\chi\|_{L^1(\mathbb{R}^d)}.\end{aligned}$$

Ainsi donc, on en déduit que

$$\Big\|\frac{1}{\varepsilon^d}\chi\Big(\frac{\cdot}{\varepsilon}\Big)\star a-a\Big\|_{L^2}\leq C_\chi\,\varepsilon\|\,|\cdot|\widehat{a}\|_{L^2}.$$

La proposition 7.1.1 et la relation de Fourier-Plancherel assurent alors l'inégalité (7.9). Nous allons maintenant démontrer que, pour tout ε dans l'intervalle $]0,1[$, l'ensemble

$$\mathcal{B}_\varepsilon\stackrel{\text{déf}}{=}\Big\{\frac{1}{\varepsilon^d}\chi\Big(\frac{\cdot}{\varepsilon}\Big)\star a\,,\ a\in B_{H^1_0(\Omega)}(0,1)\Big\}$$

est une partie d'adhérence compacte dans $\mathcal{C}(\overline{\Omega}+B(0,1))$. Pour ce faire, observons tout d'abord que, d'après l'inégalité de Poincaré, on a, pour tout x de $\overline{\Omega}+B(0,1)$,

$$\begin{aligned}\Big|\frac{1}{\varepsilon^d}\chi\Big(\frac{\cdot}{\varepsilon}\Big)\star a(x)\Big| &\leq \Big\|\frac{1}{\varepsilon^d}\chi\Big(\frac{\cdot}{\varepsilon}\Big)\Big\|_{L^2}\|a\|_{L^2}\\ &\leq \frac{C}{\varepsilon^{\frac{d}{2}}}\|\chi\|_{L^2}\|\nabla a\|_{L^2}. \end{aligned}\tag{7.10}$$

De plus, d'après l'inégalité des accroissements finis, on a

$$\begin{aligned}\Big|\frac{1}{\varepsilon^d}\chi\Big(\frac{\cdot}{\varepsilon}\Big)\star a(x)-\frac{1}{\varepsilon^d}\chi\Big(\frac{\cdot}{\varepsilon}\Big)\star a(x')\Big| &\leq |x-x'|\,\Big\|\frac{1}{\varepsilon^d}D\Big(\chi\Big(\frac{\cdot}{\varepsilon}\Big)\star a\Big)\Big\|_{L^\infty}\\ &\leq |x-x'|\,\varepsilon^{-\frac{d}{2}}\|\chi\|_{L^2}\|\nabla a\|_{L^2}.\end{aligned}$$

Pour tout ε de l'intervalle $]0,1[$, l'ensemble $\mathcal{B}_\varepsilon$ est une partie équicontinue de $\mathcal{C}(\overline{\Omega}+B(0,1))$. D'après l'inégalité (7.10) et le théorème d'Ascoli (voir le théorème 2.4.1 page 40), l'ensemble $\mathcal{B}_\varepsilon$ est une partie relativement compacte de $\mathcal{C}(\overline{\Omega}+B(0,1))$. Elle peut donc être recouverte par une famille finie de boules (pour la norme de la convergence uniforme sur $\overline{\Omega}+B(0,1)$) de rayon

$$\frac{\alpha}{\big(\mu(\overline{\Omega}+B(0,1))\big)^{\frac{1}{2}}}\,\cdot$$

Comme on a, pour toute fonction f bornée supportée dans $\overline{\Omega}+B(0,1)$,

$$\|f\|_{L^2}\leq\big(\mu(\overline{\Omega}+B(0,1))\big)^{\frac{1}{2}}\|f\|_{L^\infty},$$

On a démontré grâce à (7.9) que $\mathcal{B}_\varepsilon$ était d'adhérence compacte dans $L^2(\Omega)$. □

Chapitre 8

Les distributions tempérées en une dimension

Introduction

Ce chapitre constitue l'un des aboutissements de ce cours. Comme nous l'avons vu au chapitre 3, le fait de savoir si un espace de Banach pouvait être identifié, via une forme bilinéaire continue, au dual d'un autre (voir la définition 3.2.1 page 54) était un point important. En fait, ceci permettait d'extraire de toute suite bornée une suite convergeant en un sens affaibli : la convergence dite faible$\star$.

Lorsque au chapitre 5 nous avons étudié les espaces de fonctions de puissance p ième intégrable, nous avons vu que la connaissance des familles de moyennes pondérées

$$\left(\int_\Omega f(x)\varphi(x)dx\right)_{\varphi\in A}$$

pour φ variant par exemple dans une partie dense A était une information parfois plus parlante et facile à manier que la connaissance de toutes les valeurs ponctuelles de la fonction f. Ainsi, les fonctions apparaissent comme des formes linéaires continues sur certains espaces de fonctions. C'est ainsi que nous allons généraliser ici la notion de fonction. Le point de vue que nous adaptons est celui des distributions tempérées qui est un point de vue global sur tout l'espace $\mathbb{R}$. Il est particulièrement adapté à la généralisation de la tranformation de Fourier à une large classe de fonctions et pas seulement les fonctions L^1 ou L^2 comme au chapitre précédent. Pour ce premier contact avec les distributions tempérées, nous avons, dans un souci de simplicité, fait le choix de se restreindre à la dimension 1.

Dans la première section, nous allons définir le concept de distribution tempérée et montrer comment il généralise celui de fonction localement intégrable (avec une petite condition à l'infini près). À ce titre, le théorème 8.1.1 montre bien comment la connaissance de

$$\left(\int_\Omega f(x)\varphi(x)dx\right)_{\varphi\in E}$$

pour un certaine espace vectoriel E équivaut à la connaissance de la fonction (au sens usuel des fonctions c'est-à-dire au sens des valeurs ponctuelles). Ensuite, plusieurs exemples sont données de distributions tempérées qui ne sont pas des fonctions. Enfin, on montre comment la notion de convergence faible$\star$ conduit à réviser les notions de convergence (voir la proposition 8.1.6).

Dans la seconde section, nous définissons des opérations sur les distributions par transposition. Ceci introduit une nouveauté radicale : toutes les distributions admettent une dérivée ! et l'on peut définir la transformation de Fourier de toute distribution tempérée. La suite de la section montre diverses formules qui peuvent parfois surprendre un peu.

Dans la troisième section, nous donnons deux applications de cette théorie ; l'un à un problème de l'analyse classique dont la résolution dans le cadre strict des fonctions est difficile et qui ici ne prend que quelques lignes. Ensuite, nous établirons une formule intégrale pour la résolution de l'équation

$$u - u'' = f$$

avec des conditions de nullité pour u en $-\infty$ et $+\infty$.

8.1 Définition des distributions tempérées ; Exemples

Comme nous l'avons déjà écrit dans l'introduction de ce chapitre, le concept de distributions prend sa source dans le fait q'une fonction peut être décrite au travers d'une collection de ses moyennes pondérées. On peut ainsi penser à généraliser la fonction de fonction en terme de forme linéaires sur un espace de fonctions : l'espace que l'on appelera espace des fonctions de test. Comment choisir cet espace de fonctions de test ? Si l'on se refère aux espaces L^p, le plus "petit" espace dense rencontré, c'est-à-dire l'espace des fonctions indéfiniment dérivables à support compact. En effet, nous savons par exemple que si f est dans $L^2(\mathbb{R})$ et que

$$\forall \varphi \in \mathcal{D}(\mathbb{R}) \, , \; \int_{\mathbb{R}} f(x)\varphi(x)dx = 0$$

alors la fonction f est nulle en tant que fonctions de $L^2(\mathbb{R})$, c'est-à-dire qu'elle est nulle presque partout. Ce choix est bon et conduit à la théorie générale des distributions.

Pour des raisons qui apparaitront bientôt, on souhaite en outre que cet espace de fonction de test soit stable par la transformation de Fourier. Comme nous allons le voir, ce n'est pas du tout le cas pour l'espace $\mathcal{D}(\mathbb{R})$ des fonctions indéfiniment dérivables à support compact. En effet soit φ une fonction de $\mathcal{D}(\mathbb{R})$, alors la fonction

$$\left\{ \begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{C} \\ \zeta & \longmapsto & \displaystyle\int_{\mathbb{R}} e^{-i\zeta x} f(x)dx \end{array} \right.$$

est bien défine car φ est à support compact, et le théorème de dérivation complexe pour les intégrales assure que cette fonction es tune fonction holomorphe (c'est-à-dire dérivable au sens complexe) sur $\mathbb{C}$. D'après le principe des zéros isolés, cette fonction ne peut coîncider sur les réels avec une fonction à support compact que si elle est nulle ce qui d'après la formule d'inversion de Fourier entraîne la nullité de φ.

Observons quelles sont les propriétés de la transformée de Fourier d'une fonction de $\mathcal{D}(\mathbb{R})$. Le corollaire 7.1.1 page 120 implique que

$$\sup_{\xi \in \mathbb{R}^d} (1 + |\xi|)^N \Big| \frac{d}{d\xi} \widehat{f}(\xi) \Big| < \infty.$$

Ceci conduit naturellement à la définition suivante.

Définition 8.1.1. *On désigne par $\mathcal{S}(\mathbb{R})$ (ou bien par $\mathcal{S}$) l'ensemble des fonctions f indéfiniment différentiables sur $\mathbb{R}$ telles que*

$$\forall n \in \mathbb{N}\,,\ \|f\|_{n,\mathcal{S}} \stackrel{\text{déf}}{=} \max_{k\leq n} \sup_{x\in\mathbb{R}} (1+|x|)^n |f^{(k)}(x)| < \infty.$$

Remarquons tout d'abord que l'espace $\mathcal{S}(\mathbb{R})$ contient l'espace $\mathcal{D}(\mathbb{R})$ des fonctions indéfiniment dérivables à support compact. Remarquons aussi que les gaussiennes, c'est-à-dire les fonctions

$$x \longmapsto e^{-\frac{(x-x_0)^2}{a^2}}$$

sont des éléments de $\mathcal{S}$.

Les distributions tempérées vont appaître comme des formes linéaires (vérifiant une propriété de "continuité") sur cet espace. Plus précisément, on pose la définition suivante.

Définition 8.1.2. *On appelle distribution tempérée sur $\mathbb{R}$ une forme linéaire définie sur l'espace $\mathcal{S}(\mathbb{R})$ et qui vérifie*

$$\exists n \in \mathbb{N}\,,\ \exists C \,/\, \forall \varphi \in \mathcal{S}(\mathbb{R})\,,\ |\langle T, \varphi\rangle| \leq C\|\varphi\|_{n,\mathcal{S}}. \tag{8.1}$$

On note $\mathcal{S}'(\mathbb{R})$ l'ensemble de ces formes linéaires.

Nous allons donner une suite d'exemples de distributions tempérées.

Définition 8.1.3. *On désigne par $L^1_{\mathcal{M}}(\mathbb{R})$ l'ensemble des fonctions localement intégrables telles qu'il existe un entier N tel que $(1+|x|)^{-N} f(x)$ soit dans $L^1(\mathbb{R})$.*

Exercice 8.1.1. *Démontrer que les espaces $L^p(\mathbb{R})$ sont inclus dans $L^1_{\mathcal{M}}(\mathbb{R})$.*

Exercice 8.1.2. *Démontrer que les fonctions*

$$x \longmapsto e^x \quad \text{et} \quad x \longmapsto e^{x^2}$$

n'appartiennent pas à $L^1_{\mathcal{M}}(\mathbb{R})$

Proposition 8.1.1. *La forme linéaire définie par*

$$\iota \left\{ \begin{array}{ccc} \mathcal{S}(\mathbb{R}) & \longrightarrow & \mathbb{C} \\ \varphi & \longmapsto & \displaystyle\int_{\mathbb{R}} f(x)\varphi(x)dx \end{array} \right.$$

définit une distribution sur $\mathbb{R}$.

Démonstration.

□

Proposition 8.1.2. *Soit a un point de $\mathbb{R}$. Les formes linéaires*

$$\left\{ \begin{array}{ccc} \mathcal{S}(\mathbb{R}) & \longrightarrow & \mathbb{C} \\ \phi & \longmapsto & \phi(a) \end{array} \right. \quad \text{et} \quad \left\{ \begin{array}{ccc} \mathcal{S}(\mathbb{R}) & \longrightarrow & \mathbb{C} \\ \phi & \longmapsto & \displaystyle\sum_{n\in\mathbb{N}} \phi(n) \end{array} \right.$$

sont des distributions tempérées.

Démonstration.

□

Proposition 8.1.3. *Soit ϕ une fonction de $\mathcal{S}$. Alors la fonction*

$$\begin{cases} \mathbb{R} & \longrightarrow & \mathbb{C} \\ \varphi & \longmapsto & \dfrac{\varphi(x)-\varphi(-x)}{x} \end{cases}$$

est intégrable et la forme linéaire définie par

$$\left\langle \mathrm{vp}\,\frac{1}{x},\, \varphi \right\rangle = \frac{1}{2}\int_{-\infty}^{\infty} \frac{\varphi(x)-\varphi(-x)}{x} dx$$

est une distribution tempérée.

Démonstration. D'après l'inégalité des accroissements finis, on a

$$\big|\varphi(x)-\varphi(-x)\big| \leq 2|x| \sup_{x\in\mathbb{R}} |\varphi'(x)|.$$

Par définition des semi-normes sur $\mathcal{S}$, on a

$$\big|\varphi(x)-\varphi(-x)\big| \leq \frac{2}{1+|x|}\|\varphi\|_{1,\mathcal{S}}.$$

Ainsi donc,

$$\begin{array}{lcll} \forall x \in \mathbb{R}\,/\,|x| \leq 1, & \left|\dfrac{\varphi(x)-\varphi(-x)}{x}\right| & \leq & 2\|\varphi\|_{1,\mathcal{S}} \quad \text{et} \\ \forall x \in \mathbb{R}\,/\,|x| \geq 1, & \left|\dfrac{\varphi(x)-\varphi(-x)}{x}\right| & \leq & \dfrac{2}{|x|^2}\|\varphi\|_{1,\mathcal{S}}. \end{array}$$

La fonction $x \longmapsto \dfrac{\varphi(x)-\varphi(-x)}{x}$ est donc intégrable sur $\mathbb{R}$ et la forme linéaire ainsi définie est continue sur $\mathcal{S}$. □

Remarque Si φ est une fonction de test identiquement nulle au voisinage de 0, on a alors

$$\left\langle \mathrm{vp}\,\frac{1}{x},\, \varphi \right\rangle = \int_{\mathbb{R}} \frac{\varphi(x)}{x} dx.$$

Ceci signifie que l'on a "étendu" la fonction x^{-1}.

Proposition 8.1.4. *Soient a dans $\mathbb{R}$ et k dans $\mathbb{N}$. La forme linéaire définie par*

$$\langle \delta_a^{(k)}, \varphi\rangle \stackrel{\text{def}}{=} (-1)^k \varphi^{(k)}(a)$$

est une distribution tempérée.

Démonstration. Il suffit d'observer que par définition des semi-normes sur $\mathcal{S}$, on a

$$|\varphi^{(k)}(a)| \leq C_a \|\varphi\|_{k,\mathcal{S}}.$$

ce qui assure le résultat. □

Proposition 8.1.5. *La forme linéaire définie par*

$$\left\langle \mathrm{Pf}\,\frac{1}{x^2},\, \varphi \right\rangle = \frac{1}{2}\int_{-\infty}^{\infty} \frac{\varphi(x)+\varphi(-x)-2\varphi(0)}{x^2}\, dx$$

est une distribution tempérée.

Démonstration. L'inégalité de Taylor à l'ordre deux assure que

$$\big|\varphi(x)+\varphi(-x)-2\varphi(0)\big| \leq \frac{|x|^2}{2}\sup_{x\in\mathbb{R}}|\varphi''(x)|.$$

Ainsi donc on a

$$\frac{\varphi(x)+\varphi(-x)-2\varphi(0)}{x^2} \leq \min\Big\{1, \frac{1}{|x|^2}\Big\}\, \|\varphi\|_{2,\mathcal{S}}.$$

Ceci assure que $\mathrm{Pf}\,\dfrac{1}{x^2}$ est une distribution tempérée. □

Remarque Si φ est une fonction de test identiquement nulle au voisinage de 0, on a alors

$$\Big\langle \mathrm{Pf}\,\frac{1}{x^2}, \varphi\Big\rangle = \int_{\mathbb{R}} \frac{\varphi(x)}{x^2}dx.$$

Ceci signifie que l'on a "étendu" la fonction x^{-2}.

Le théorème suivant montre que que l'espace $L^1_{\mathcal{M}}(\mathbb{R})$ peut être identifié à un sous espace des distributions tempérées.

Théorème 8.1.1. *Soit ι l'application définie par*

$$\iota\left\{\begin{array}{ccl} L^1_{\mathcal{M}}(\mathbb{R}) & \longrightarrow & \mathcal{S}'(\mathbb{R}) \\ f & \longmapsto & \iota(f) : \varphi \mapsto \displaystyle\int_{\mathbb{R}} f(x)\varphi(x)dx \end{array}\right.$$

est une injection linéaire. De plus, on a la propriété suivante. Soit N tel que $(1+|x|)^{-N}f(x)$ soit dans $L^1(\mathbb{R})$, alors

$$|\langle \iota(f), \varphi\rangle| \leq C\big\|(1+|\cdot|)^{-N}f\big\|_{L^1}\|\varphi\|_{N,\mathcal{S}}.$$

Démonstration. Soit f une fonction de $L^1_{\mathcal{M}}(\mathbb{R})$ telle que $\iota(f)=0$. Nous allons démontrer que, pour tout couple de réels (a,b) tel que a soit strictement inférieur à b, on a

$$\int_a^b |f(x)|dx = 0$$

ce qui assurera le théorème. On se donne une fonction χ de $\mathcal{D}([-1,1[)$ paire et d'intégrale 1 et on considère l'approximation de l'identité associée, c'est-à-dire la famille

$$\chi_\varepsilon(x) \stackrel{\text{déf}}{=} \varepsilon^{-d}\chi\Big(\frac{x}{\varepsilon}\Big).$$

On définit enfin la fonction g par

$$g(x) = \frac{\overline{f}(x)}{|f(x)|} \quad \text{si} \quad f(x)\neq 0 \quad \text{et 0 sinon.}$$

C'est une fonction de $L^\infty(\mathbb{R})$. La fonction

$$\varphi_\varepsilon \stackrel{\text{déf}}{=} \chi_\varepsilon \star (\mathbf{1}_{[a,b]}g)$$

est une fonction indéfiniment différentiable à support compact sur $\mathbb{R}$, donc une fonction de $\mathcal{S}(\mathbb{R})$. Par hypothèse, on a

$$I_\varepsilon \stackrel{\text{déf}}{=} \int_{\mathbb{R}} f(x)\varphi_\varepsilon(x)dx = 0.$$

Pour tout ε dans l'intervalle $]0,\delta[$, la fonction

$$(x,y) \longmapsto \varepsilon^{-d}\chi\Big(\frac{x-y}{\varepsilon}\Big)\mathbf{1}_{[a,b]}(y)g(y)f(x) = \varepsilon^{-d}\chi\Big(\frac{x-y}{\varepsilon}\Big)\mathbf{1}_{[a,b]}(y)g(y)f(x)\mathbf{1}_{[a-\delta,b+\delta]}(x)$$

est une fonction de $L^1(\mathbb{R}\times\mathbb{R})$ car elle est à support compact et bornée. En effet, la fonction est bornée en module par 1 ; d'après le théorème de Fubini 5.1.5 page 83, pour tout ε plus petit que δ,

$$\int_{\mathbb{R}\times\mathbb{R}} \varepsilon^{-d}\Big|\chi\Big(\frac{x-y}{\varepsilon}\Big)\Big|\mathbf{1}_{[a,b]}(y)|g(y)|\mathbf{1}_{B(0,\delta)}(x)|f(x)|dxdy \leq \|\chi\|_{L^1}\|\mathbf{1}_{[a-\delta,b+\delta]}f\|_{L^1}.$$

Le théorème de Fubini-Tonelli 5.1.6 page 84 joint à la parité de la fonction χ assure alors que

$$\begin{aligned}
I_\varepsilon &= \int_{\mathbb{R}\times\mathbb{R}} \chi_\varepsilon(x-y)\mathbf{1}_{[a,b]}(y)g(y)f(x)dxdy\\
&= \int_{\mathbb{R}\times\mathbb{R}} \chi_\varepsilon(x-y)\mathbf{1}_{[a,b]}(y)g(y)\mathbf{1}_{K+B(0,\delta)}(x)f(x)dxdy\\
&= \int_{\mathbb{R}}\Big(\int_{\mathbb{R}} \chi_\varepsilon(y-x)\mathbf{1}_{[a-\delta,b+\delta]}(x)f(x)dx\Big)\mathbf{1}_K(y)g(y)dy\\
&= \int_{\mathbb{R}}\big(\chi_\varepsilon\star(\mathbf{1}_{[a-\delta,b+\delta])}f)\big)(y)\mathbf{1}_{[a,b]}(y)g(y)dy.
\end{aligned}$$

Le théorème 5.4.5 page 100 dit que

$$\lim_{\varepsilon\to 0}\big\|(\chi_\varepsilon\star(\mathbf{1}_{[a-\delta,b+\delta])}f)-\mathbf{1}_{[a-\delta,b+\delta])}f\big\|_{L^1}=0$$

Comme g appartient à $L^\infty(\Omega)$, on a

$$\begin{aligned}
0 &= \lim_{\varepsilon\to 0} I_\varepsilon\\
&= \int \mathbf{1}_{[a-\delta,b+\delta])}(y)f(y)\mathbf{1}_{[a,b]}(y)g(y)dy\\
&= \int_a^b f(y)\frac{\overline{f}(y)}{|f(y)|}dy\\
&= \int_a^b |f(y)|dy.
\end{aligned}$$

D'où le théorème. □

Remarque Toute fonction de $L^1_{\mathcal{M}}$ s'identifie ainsi à une distribution tempérée et l'on peut maintenant donner la définition suivante.

Définition 8.1.4. *On dit qu'une distribution tempérée T est une fonction si et seulement si il existe une fonction f de $L^1_{\mathcal{M}}$ telle que $T=\iota(f)$.*

Nous allons maintenant donner la définition de la convergence d'une suite de distributions tempérées. IL s'agit d'ue convergence de type faible-$\star$ tell que définie au chapitre 3.

Définition 8.1.5. *De plus, soient une suite $(T_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{S}'(\mathbb{R})$ et T un élément de $\mathcal{S}(\mathbb{R})$. On dit que la suite $(T_n)_{n\in\mathbb{N}}$ converge vers T si et seulement si*

$$\forall\varphi\in\mathcal{S}(\mathbb{R})\,,\ \lim_{n\to\infty}\langle T_n,\varphi\rangle=\langle T,\varphi\rangle.$$

Nou allons donner maintenant un exemple de suite convergente qui montre combien la notion de suite convergente au sens des distributions tempérées est loin de celle au sens des fonctions.

Proposition 8.1.6. *La suite de fonctions* $S_N(x) \stackrel{\text{déf}}{=} \sum_{n=1}^{N} \sin(nx)$ *converge dans* $\mathcal{S}'(\mathbb{R})$.

Démonstration. Il agit de trouver une distribution tempérée S telle que, pour toute fonction ϕ de $\mathcal{S}$, on ait

$$\lim_{N\to\infty} \sum_{n=1}^{N} \int_{\mathbb{R}} \sin(nx)\varphi(x)dx = \langle S, \varphi \rangle.$$

En intégrant par parties deux fois, il vient, pour tout entier n supérieur ou égal à 1

$$\int_{\mathbb{R}} \sin(nx)\varphi(x)dx = -\frac{1}{n^2} \int_{\mathbb{R}} \sin(nx)\varphi''(x)dx. \tag{8.2}$$

Définissons la fonction Σ par

$$\Sigma(x) = -\sum_{n=1}^{\infty} \frac{1}{n^2} \sin(nx)$$

C'est une fonction continue et bornée sur $\mathbb{R}$. On pose alors, pour ϕ dans $\mathcal{S}$,

$$\langle S, \varphi \rangle \stackrel{\text{def}}{=} \int_{\mathbb{R}} \Sigma(x)\varphi''(x)dx.$$

Ceci définit une distribution tempérée. De plus pour tout fonction ϕ, d'après (8.2), on a

$$\lim_{N\to\infty} \sum_{n=1}^{N} \int_{\mathbb{R}} \sin(nx)\varphi(x)dx = \int_{\mathbb{R}} \Sigma(x)\varphi''(x)dx = \langle S, \varphi \rangle.$$

La proposition est démontrée. □

8.2 Opérations sur les distributions tempérées.

L'idée fondamentale est la suivante : comme nous venons de l'observer dans la section précédente, l'espace $\mathcal{S}'$ est beaucoup plus "gros" que l'espace $\mathcal{S}$; sur cet espace $\mathcal{S}$, on peut définir beaucoup d'opération, en particulier la dérivation, la transformation de Fourier, la convolution par une fonction dont la croissance est contrôlée. Pour cela, il importe de définir une notion d'application linéaire continue de $\mathcal{S}$ dans $\mathcal{S}$.

Définition 8.2.1. *Une application linéaire* A *de* $\mathcal{S}$ *dans lui-même est dite continue si et seulement si*

$$\forall k \in \mathbb{N}, \exists (C_k, n_k) \in]0, \infty[\times \mathbb{N} \,/\, \forall \phi \in \mathcal{S},\ \|A\phi\|_{k,\mathcal{S}} \leq C_k \|f\|_{n_k,\mathcal{S}}.$$

Introduisons les deux espaces de fonctions suivants.

Définition 8.2.2. *On appelle espace des fonctions à croissance modérée (ou lente) et l'on désigne par* $\mathcal{O}_M$ *l'espace des fonctions* f *indéfiniment différentiables sur* $\mathbb{R}$ *telles que*

$$\forall k \in \mathbb{N},\ \exists N \in \mathbb{N},\ \exists C \in \mathbb{R}^+ \,/\, \forall x \in \mathbb{R},\ |f^{(k)}(x)| \leq (1+|x|)^N.$$

On désigne par $L^1_{\mathcal{S}}$ *l'espace des fonctions localement intégrables telles que, pour tout entier* N, *la fonction* $(1+|x|)^N f(x)$ *soit dans* L^1.

Les polynômes sont d'excellents exemples de fonctions de $\mathcal{O}_M$. Les fonctions localement intégrables décroissant à l'infini plus vite que toute puissance négative de $|x|$ sont d'excellents exemples de fonctions de $L^1_{\mathcal{S}}$. Par exemple, la fonction

$$x \longmapsto \frac{1}{x^\alpha} e^{-|x|^\beta}$$

avec α dans $]0,1[$ et β strictement positif appartient à $L^1_{\mathcal{S}}$.

Proposition 8.2.1. *Les applications linéaires suivantes sont continues de $\mathcal{S}$ dans $\mathcal{S}$ au sens suivant :*
- *l'application $\phi \mapsto \check{\varphi}(x) \stackrel{\text{déf}}{=} \varphi(-x)$;*
- *l'application $\phi \mapsto (-1)^k \phi^{(k)}$;*
- *la transformation de Fourier définie à la définition 7.1.1 page 117 ;*
- *si f désigne une fonction de $\mathcal{O}_M$, l'application M_f qui désigne la multiplication par f ;*
- *si f désigne une fonction de $L^1_{\mathcal{S}}$, l'application $f\star$ qui désigne la convolution par f, c'est-à-dire que*

$$(f \star \varphi)(x) \stackrel{\text{déf}}{=} \int_{\mathbb{R}} f(y)\varphi(x-y)dy\,.$$

Démonstration. La première continuité est évidente. Pour la seconde, observons que, par définition des normes $\|\cdot\|_{k,\mathcal{S}}$, il est immédiat que

$$\|\varphi^{(\ell)}\|_{k,\mathcal{S}} \leq C\|\varphi\|_{k+\ell,\mathcal{S}}.$$

Pour la continuité de la multiplication, il suffit d'appliquer la formule de Leibnitz qui dit que

$$(f\varphi)^{(k)} = \sum_{\ell\leq k} C_k^\ell f^{(k-\ell)}\varphi^{(\ell)}.$$

Comme la fonction f appartient à $\mathcal{O}_M$, il existe un entier N tel que, pour tout entier ℓ plus petit que k, on ait

$$\forall x \in \mathbb{R}\,,\ |f^{(\ell)}(x)| \leq C(1+|x|)^N.$$

On en déduit alors que

$$\|f\varphi\|_{k,\mathcal{S}} \leq C_f\|\varphi\|_{k+N,\mathcal{S}}.$$

Démontrons la continuité de la transformation de Fourier. Obervons que le théorème 7.1.1 page 119 implique que

$$M(\widehat{\varphi})' = -IM\mathcal{F}(Mf) = -\mathcal{F}\big((M\phi)'\big) = -\widehat{\phi} - \mathcal{F}(M\varphi').$$

Il en résulte que

$$(1+|\xi|)|\widehat{\varphi}(\xi)| \leq 2\|\varphi\|_{L^1(\mathbb{R})} + \|M\varphi'\|_{L^1(\mathbb{R})} \leq 2\|f\|_{3,\mathcal{S}}.$$

Nous amettons la démonstration générale pour les semi-normes d'indice k sur $\mathcal{S}(\mathbb{R})$.

Étudions maintenant la convolution. Par dérivation sous le signe somme, on a

$$(f\star\phi)^{(k)}(x) = \int_{\mathbb{R}} f(y)\phi^{(k)}(x-y)dy.$$

Comme $|x| \leq 2\max\{|x-y|, |y|\}$, on a

$$(1+|x|)^k \leq 2^k\left((1+|x-y|)^k + (1+|y|)^k\right).$$

Ainsi donc, on peut écrire que

$$(1+|x|)^k|(f\star\phi)^{(k)}(x)| \leq 2^k\int_{\mathbb{R}}|f(y)|(1+|x-y|)^k|\partial^\alpha\phi(x-y)|dy$$
$$+\,2^k\int_{\mathbb{R}}(1+|y|)^k|f(y)|\,|\varphi^{k)}(x-y)|dy.$$

On en déduit alors que

$$\|f\star\phi\|_{k,\mathcal{S}} \leq C\|(1+|\cdot|)^kf\|_{L^1}\|\varphi\|_{k,\mathcal{S}}.$$

D'où la continuité de la convolution □

Nous pouvons donc définir leur transposée (qui leur rassemble souvent beaucoup) sur l'espace $\mathcal{S}'$. C'est une extension spectaculaire. Tout repose sur le théorème suivant.

Théorème 8.2.1. *Soit A une application linéaire continue de $\mathcal{S}(\mathbb{R})$ dans $\mathcal{S}(\mathbb{R})$. Alors l'application linéaire ${}^tA$*

$${}^tA\left\{\begin{array}{ccc}\mathcal{S}'(\mathbb{R}) & \longrightarrow & \mathcal{S}'(\mathbb{R})\\ T & \longmapsto & {}^tAT \quad \textit{définie par} \quad \langle {}^tAT,\varphi\rangle = \langle T, A\varphi\rangle.\end{array}\right.$$

est bien définie. De plus, elle est continue au sens suivant : si $(T_n)_{n\in\mathbb{N}}$ est une suite de distributions tempérées qui converge vers une distribution tempérée T, alors la suite $({}^tAT_n)_{n\in\mathbb{N}}$ converge vers ${}^tAT$.

Démonstration. Par définition T est une forme linéaire continue sur $\mathcal{S}(\mathbb{R})$ et A une application linéaire continue de $\mathcal{S}(\mathbb{R})$ dans lui-même. Donc la composée ${}^tAT \stackrel{\text{déf}}{=} T\circ A$ est une forme linéaire continue sur $\mathcal{S}(\mathbb{R})$ donc une distribution tempérée.

De plus, soit $(T_n)_{n\in\mathbb{N}}$ est une suite de distributions tempérées convergeant vers T au sens de la définition 8.1.2. Par définition de la convergence, ceci signifie que, pour toute fonction θ de $\mathcal{S}(\mathbb{R})$,

$$\lim_{n\infty}\langle T_n,\theta\rangle = \langle T,\theta\rangle.$$

En appliquant cela avec $\theta = A\phi$, on trouve que

$$\forall\phi\in\mathcal{S}(\mathbb{R})\,,\ \lim_{n\infty}\langle {}^tAT_n,\varphi\rangle = \langle {}^tAT,\varphi\rangle.$$

D'où le théorème. □

Comme première application de ce théorème, on peut définir étendre l'opération ˇ aux distributions tempérées par la formule

$$\forall\phi\in\mathcal{S}\,,\ \langle\check{T},\varphi\rangle \stackrel{\text{déf}}{=} \langle T,\check{\varphi}\rangle$$

Définissons maintenant les dérivées d'une distribution tempérée.

Définition 8.2.3. *Soit k un élément de $\mathbb{N}$. On appelle dérivation d'ordre k sur $\mathcal{S}'$ la transposée de l'application linéaire continue de $\mathcal{S}(\mathbb{R})$ dans lui-même $(-1)^k\dfrac{d^k}{dx^k}$, ce qui s'écrit, pour une distribution tempérée T,*

$$\forall\phi\in\mathcal{S}(\mathbb{R})\,,\ \langle T^{(k)},\varphi\rangle \stackrel{\text{def}}{=} \langle T,(-1)^k\phi^{(k)}\rangle.$$

Notons tout d'abord que pour toute fonction f de classe C^1 à support compact sur $\mathbb{R}$, on a, par intégration par parties,

$$\forall \phi \in \mathcal{S}(\mathbb{R})\,, \quad \int_{\mathbb{R}} f'(x)\varphi(x)dx = -\int_{\mathbb{R}} f(x)\varphi'(x)dx\,.$$

Donc pour la fonction de classe C^1, les deux notions coïncident.

Remarquons que la fonction Σ qui apparaît dans la démonstration de la proposition 8.1.6 n'est autre que

$$-\Big(\frac{d}{dx}\Big)^2 \sum_{n=1}^{\infty} \frac{1}{n^2}\sin(n\cdot)$$

c'est-à-dire la dérivée seconde au sens des distributions de la fonctions bornée et uniformément continue

$$x \longmapsto -\sum_{n=1}^{\infty} \frac{1}{n^2}\sin(nx).$$

Nous allons voir sur un exemple très simple comment la dérivée au sens des distributions prend mieux en compte les variations de la fonction que la notion classique de dérivée lorsque les fonctions ne sont pas de classe C^1, par exemple lorsque qu'une fonction est dérivable sur $\mathbb{R}$ sauf en un point. Nous avons la proposition suivante.

Proposition 8.2.2. *On a* $\dfrac{d}{dx}\mathbf{1}_{\mathbb{R}^+} = \delta_0$.

Démonstration. Il s'agit de d'étudier, pour ϕ dans $\mathcal{S}(\mathbb{R})$, l'intégrale

$$-\int_{\mathbb{R}} \mathbf{1}_{\mathbb{R}^+}(x)\varphi(x)dx.$$

Une intégration par parties assure que

$$\begin{aligned} -\int_{\mathbb{R}} \mathbf{1}_{\mathbb{R}^+}(x)\varphi'(x)dx &= -\lim_{A\to\infty}\int_0^A \varphi'(x)dx \\ &= \lim_{A\to\infty}(\phi(0)-\phi(A)) \\ &= \phi(0). \end{aligned}$$

Par définition de la masse de Dirac en 0, ceci démontre la proposition. □

Nous allons maintenant à titre d'illustration calculer la dérivée de la fonction $x \mapsto \log|x|$.

Proposition 8.2.3. *Soit* $\mathrm{vp}\dfrac{1}{x}$ *la distribution tempérée définie à la proposition 8.1.3. On a, au sens des distributions,*

$$\frac{d}{dx}\log|x| = \mathrm{vp}\frac{1}{x}\cdot$$

Démonstration. Il s'agit d'étudier, pour une fonction φ de $\mathcal{S}(\mathbb{R})$, l'intégrale

$$I(\varphi) = -\int_{\mathbb{R}} \log|x|\varphi'(x)dx.$$

D'après le théorème de la convergence dominée, on a

$$I(\varphi) = \lim_{\varepsilon\to 0} I_\varepsilon(\varphi) \stackrel{\text{def}}{=} -\int_{[-\frac{1}{\varepsilon},\frac{1}{\varepsilon}]\setminus[-\varepsilon,\varepsilon]} \log|x|\varphi'(x)dx.$$

Par intégration par parties, on trouve que

$$
\begin{aligned}
I_\varepsilon(\varphi) = -\log\varepsilon\,\varphi(-\varepsilon) + \log\Big(\frac{1}{\varepsilon}\Big)\varphi\Big(-\frac{1}{\varepsilon}\Big) + \int_{-\frac{1}{\varepsilon}}^{-\varepsilon} \frac{\varphi(x)}{x}dx \\
+ \log\varepsilon\,\varphi(\varepsilon) - \log\Big(\frac{1}{\varepsilon}\Big)\varphi\Big(\frac{1}{\varepsilon}\Big) + \int_{\varepsilon}^{\frac{1}{\varepsilon}} \frac{\varphi(x)}{x}dx \\
= \log\varepsilon\,\big(\varphi(\varepsilon) - \varphi(-\varepsilon)\big) - \log\varepsilon\,\varphi\Big(-\frac{1}{\varepsilon}\Big) + \log\varepsilon\,\varphi\Big(\frac{1}{\varepsilon}\Big) + \int_{-\frac{1}{\varepsilon}}^{-\varepsilon} \frac{\varphi(x)}{x}dx + \int_{\varepsilon}^{\frac{1}{\varepsilon}} \frac{\varphi(x)}{x}dx.
\end{aligned}
$$

Par le changement de variable $x = -y$, on trouve que

$$
\int_{[-\frac{1}{\varepsilon},\frac{1}{\varepsilon}]\setminus[-\varepsilon,\varepsilon]} \frac{\varphi(x)}{x}dx = -\int_{[-\frac{1}{\varepsilon},\frac{1}{\varepsilon}]\setminus[-\varepsilon,\varepsilon]} \frac{\varphi(x)}{x}dx
$$

ce qui donne

$$
I_\varepsilon(\varphi) = \log\varepsilon\,\big(\varphi(\varepsilon) - \varphi(-\varepsilon)\big) - \log\varepsilon\,\varphi\Big(-\frac{1}{\varepsilon}\Big) + \log\varepsilon\,\varphi\Big(\frac{1}{\varepsilon}\Big) + \frac{1}{2}\int_{[-\frac{1}{\varepsilon},\frac{1}{\varepsilon}]\setminus[-\varepsilon,\varepsilon]} \frac{\varphi(x)-\varphi(-x)}{x}dx.
$$

Vu que la fonction φ appartient à $\mathcal{S}$, on a

$$
\lim_{\varepsilon\to 0} \log\varepsilon\,\big(\varphi(\varepsilon) - \varphi(-\varepsilon)\big) - \log\varepsilon\,\varphi\Big(-\frac{1}{\varepsilon}\Big) + \log\varepsilon\,\varphi\Big(\frac{1}{\varepsilon}\Big) = 0.
$$

De plus, la proposition 8.1.3 affirme en particulier que la fonction

$$
x \longmapsto \frac{\varphi(x)-\varphi(-x)}{x}
$$

est intégrable sur $\mathbb{R}$. Le théorème de convergence dominée assure que

$$
-\int_{\mathbb{R}} \log|x|\varphi'(x)dx = \frac{1}{2}\int_{\mathbb{R}} \frac{\varphi(x)-\varphi(-x)}{x}dx
$$

ce qui est l'énoncé recherché. □

Exercice 8.2.1. *Démontrer que*

$$
\frac{d}{dx}\mathrm{vp}\frac{1}{x} = -\mathrm{Pf}\frac{1}{x^2}\cdot
$$

Nous allons maintenant définir la notion de primitive (sur $\mathbb{R}$) d'une distribution tempérée. Avant cela, démontrons le résultat suivant.

Théorème 8.2.2. *Soit T une distribution tempérée telle que $T' = 0$. Alors T est une fonction constante au sens où*

$$
\langle T,\varphi\rangle = c\int_{\mathbb{R}} \varphi(x)dx.
$$

Démonstration. Le résultat repose sur le lemme suivant.

Lemme 8.2.1. *Soit $\mathcal{S}_0(\mathbb{R})$ l'ensemble des fonctions de $\mathcal{S}(\mathbb{R})$ d'intégrale nulle. L'application*

$$
\mathbb{P}\left\{\begin{array}{ccc} \mathcal{S}_0(\mathbb{R}) & \longrightarrow & \mathcal{S}(\mathbb{R}) \\ \varphi & \longmapsto & x \mapsto \displaystyle\int_{-\infty}^{x} \varphi(y)dy \end{array}\right.
$$

est une application linéaire continue de $\mathcal{S}_0(\mathbb{R})$ dans $\mathcal{S}(\mathbb{R})$ telle que pour tout φ de $\mathcal{S}_0(\mathbb{R})$, on a $(\mathbb{P}\varphi)' = \phi$.

Démonstration. Si x est inférieur ou égal à -1, alors on a

$$
\begin{aligned}
|(\mathbb{P}\varphi)(x)| &\leq \|\varphi\|_{N+2,\mathcal{S}} \int_{-\infty}^{x} \frac{1}{(1+|y|)^{N+2}} dy \\
&\leq |x|^{-N} \|\varphi\|_{N+2,\mathcal{S}} \int_{-\infty}^{0} \frac{1}{(1+|y|)^{2}} dy.
\end{aligned}
$$

Si x est supérieur ou égal à 1, on utilise la nullité de l'intégrale de ϕ pour écrire que

$$
(\mathbb{P}\varphi)(x) = -\int_{x}^{+\infty} \varphi(y)dy
$$

et l'on fait le même raisonnement. Ensuite, on observe que $(\mathbb{P}\varphi)' = \varphi$ pour conclure. □

Poursuite de la démonstration du théorème 8.2.2 On considère dans toute la suite de cette démonstration une fonction φ_1 de $\mathcal{S}(\mathbb{R})$ d'intégrale 1. On considère alors la projection p de $\mathcal{S}(\mathbb{R})$ sur $\mathcal{S}_0(\mathbb{R})$ parallèlement à la droite engendré par φ_1, c'est-à-dire l'application

$$
p(\varphi) = \varphi - \Big(\int_{-\infty}^{+\infty} \varphi(y)dy\Big)\varphi_1. \tag{8.3}
$$

D'après le lemme 8.2.1, on a $(\mathbb{P}(p(\phi))' = p(\phi)$. Comme on a supposé que T' était nulle, on en déduit que

$$
\langle T, p(\varphi)\rangle = 0 = \langle T, \varphi\rangle - \Big(\int_{-\infty}^{+\infty} \varphi(y)dy\Big)\langle T, \varphi_1\rangle.
$$

D'où le théorème. □

Nous allons maintenant appliquer ces idées à la recherche de primitive de distributions en dimension 1.

Proposition 8.2.4. *Soit T une distribution tempérée sur $\mathbb{R}$. Il existe une distribution tempérée S telle que $S' = T$ et deux distributions tempérées vérifiant cette propriété diffèrent d'une constante.*

Démonstration. Soit φ_1 une function de $\mathcal{S}(\mathbb{R})$ d'intégrale 1. Montrons que $S \stackrel{\text{déf}}{=} -{}^t\mathbb{P}\,T$ vérifie $S' = T$. Observons que, pour toute fonction φ dans $\mathcal{S}(\mathbb{R})$, φ' est d'intégrale nulle et donc que

$$
\mathbb{P} \circ \frac{d}{dx} = \mathrm{Id}_{\mathcal{S}(\mathbb{R})}\,.
$$

On en déduit par transposition que

$$
{}^t\Big(\mathbb{P} \circ \frac{d}{dx}\Big) = \mathrm{Id}_{\mathcal{S}'(\mathbb{R})}\,.
$$

Comme ${}^t(AB) = {}^tB{}^tA$, on en déduit que

$$
{}^t\Big(\frac{d}{dx}\Big) \circ {}^t\mathbb{P} = \mathrm{Id}_{\mathcal{S}'(\mathbb{R})}\,.
$$

Le fait que ${}^t\Big(\dfrac{d}{dx}\Big) = -\dfrac{d}{dx}$ assure le résultat. □

Remarque Une démonstration plus explicite peut être faite en définissant simplement la distribution tempérée S par

$$\langle S,\varphi\rangle \stackrel{\text{def}}{=} -\langle T,(\mathbb{P}\circ p)\varphi\rangle \quad \text{avec} \quad (\mathbb{P}\circ p)\varphi(x) \stackrel{\text{def}}{=} \int_{-\infty}^{x}\Big(\varphi(y)-\Big(\int_{-\infty}^{\infty}\varphi(t)dt\Big)\varphi_1(y)\Big)dy$$

où φ_1 est une fonction de $\mathcal{S}(\mathbb{R})$ d'intégrale 1, par exemple la fonction $\sqrt{\pi}e^{-x^2}$.

Le lien entre primitive et dérivée peut être précisé de la manière suivante.

Théorème 8.2.3. *Soit T une distribution tempérée sur $\mathbb{R}$. Si sa dérivée (au sens des distributions) T' est une fonction $L^1_{\mathcal{M}}(\mathbb{R})$, alors T est une fonction continue et l'on a*

$$T(x)=\int_0^x T'(y)dy+C.$$

Démonstration. Elle repose sur le lemme suivant que nous admettrons un court instant.

Lemme 8.2.2. *Soit f une fonction de $L^1_{\mathcal{M}}(\mathbb{R})$. Alors*

$$\frac{d}{dx}\int_0^x f(y)dy=f(x)$$

au sens des distributions.

Posons $F(x)=\displaystyle\int_a^x T'(y)dy$. D'après le lemme 8.2.2, on a $F'=T'$. D'après la proposition 8.2.2, on a

$$T-F=C.$$

D'où le théorème, pourvu bien sûr que nous démontrions le lemme admis. □

Démonstration du lemme 8.2.2 D'après la définition de la dérivation au sens des distributions, il s'agit de démontrer que, si

$$F(x)\stackrel{\text{déf}}{=}\int_0^x f(y)dy,$$

alors on a

$$\forall\varphi\in\mathcal{S}(\mathbb{R})\,,\ -\int_{\mathbb{R}}F(x)\varphi'(x)dx=\int_{\mathbb{R}}f(x)\varphi(x)dx.$$

Observons que si

$$\mathcal{I}^+\stackrel{\text{def}}{=}\big\{(x,y)\in\mathbb{R}^2\,,\ 0\le y\le x\big\} \quad \text{et} \quad \mathcal{I}^-\stackrel{\text{def}}{=}\big\{(x,y)\in\mathbb{R}^2\,,\ x\le y\le 0\big\},$$

les fonctions

$$\mathbf{1}_{\mathcal{I}^+}(x,y)f(y)\varphi'(x) \quad \text{et} \quad \mathbf{1}_{\mathcal{I}^-}(x,y)f(y)\varphi'(x)$$

sont intégrables sur $\mathbb{R}^2$. En effet, par définition de $L^1_{\mathcal{M}}(\mathbb{R})$,

$$\begin{aligned} I(f,\varphi) &\stackrel{\text{def}}{=} \int_{\mathbb{R}^2}\mathbf{1}_{\mathcal{I}^+}(x,y)|f(y)|\,|\varphi'(x)|dxdy\\ &\le \int_{\mathbb{R}^2}\mathbf{1}_{\mathcal{I}^+}(x,y)(1+|y|)^M|\widetilde{f}(y)|\,|\varphi'(x)|dxdy \quad \text{avec } \widetilde{f}\in L^1(\mathbb{R}).\end{aligned}$$

D'après le théorème de Fubini sur les fonctions positives, on a

$$I(f,\phi)\le C_{M,a}\int_{\mathbb{R}}(1+|x|)^{M+1}|\varphi'(x)|dx\le C_{M,a}\|\varphi\|_{M+1,\mathcal{S}}.$$

Le fait que $\mathbf{1}_{\mathcal{I}^-}(x,y)f(y)\varphi'(x)$ se démontre de manière strictement analogue. Observons que, grâce au théorème de Fubini, on peut écrire que

$$\begin{aligned} -\int_{\mathbb{R}} F(x)\varphi'(x)dx &= -\int_{\mathbb{R}}\Big(\int_y^{\infty}\varphi'(x)dx\Big)\mathbf{1}_{[0,\infty[}(y)f(y)dy \\ &\qquad\qquad + \int_{\mathbb{R}}\Big(\int_{\infty}^{y}\varphi'(x)dx\Big)\mathbf{1}_{]-\infty,0]}(y)f(y)dy \\ &= \int_{\mathbb{R}}\varphi(y)\mathbf{1}_{[0,\infty[}(y)f(y)dy + \int_{\mathbb{R}}\varphi(y)\mathbf{1}_{]-\infty,0]}(y)f(y)dy \\ &= \int_{\mathbb{R}}\varphi(y)f(y)dy. \end{aligned}$$

Le lemme 8.2.2 et donc le théorème 8.2.3 est démontré. □

Définissons maintenant la multiplication par une fonction de $\mathcal{O}_M$.

Définition 8.2.4. *Soit θ une fonction de $\mathcal{O}_M$. La multiplication par θ sur $\mathcal{S}'$ est la transposée de la multiplication par θ sur $\mathcal{S}$, ce qui s'écrit pour T dans $\mathcal{S}'$,*

$$\langle \theta T, \varphi\rangle \stackrel{\text{def}}{=} \langle T, \theta\varphi\rangle.$$

Il est clair que si T est une fonction de L^1_M, alors la multiplication ainsi définie coïncide avec la multiplication usuelle des fonctions.

Nous allons maintenant démontrer une généralisation de la formule de Liebnitz.

Proposition 8.2.5. *Si θ appartient à $\mathcal{O}_M$ et T à $\mathcal{S}'$, alors on a, pour tout entier k,*

$$(\theta T)^{(k)} = \sum_{\ell\leq k} C_k^{\ell}\theta^{(k-\ell)}T^{(\ell)}.$$

Démonstration. Il suffit de démontrer que $(\theta T)' = \theta' T + \theta T'$ et ensuite d'itérer comme dans le cas classique. Par définition de la dérivation et de la multiplication, on a

$$\begin{aligned} \langle (\theta T)', \varphi\rangle &= -\langle \theta T, \varphi'\rangle \\ &= -\langle T, \theta\varphi'\rangle. \end{aligned}$$

La formule de Leibnitz pour les fonctions régulières implique alors que

$$\langle (\theta T)', \varphi\rangle = -\langle T, (\theta\phi)'\rangle + \langle T, \theta'\varphi\rangle.$$

À nouveau par définition de la dérivation et de la multiplication, on en déduit que

$$\begin{aligned} \langle (\theta T)', \varphi\rangle &= \langle T', \theta\varphi\rangle + \langle T, \theta'\varphi\rangle \\ &= \langle \theta T' + \theta' T, \varphi\rangle. \end{aligned}$$

D'où la proposition. □

Définissons maintenant la convolution par une fonction de L^1_S.

Définition 8.2.5. *Soit f une fonction de L^1_S. On définit la convolution par f sur les distributions tempérées comme la transposée de la convolution par $\check{f}$ sur $\mathcal{S}$, ce qui s'écrit*

$$\langle f\star T, \varphi\rangle \stackrel{\text{def}}{=} \langle T, \check{f}\star\varphi\rangle.$$

Vérifions que l'on a bien généralisé la convolution définie grâce au théorème 5.4.1 page 97. Soit g une fonction de $L^1(\mathbb{R})$, on a

$$\begin{aligned}\langle \iota(g), \check{f} \star \varphi\rangle &= \int_{\mathbb{R}} g(x)(\check{f} \star \phi)(x)dx \\ &= \int_{\mathbb{R}} g(x)\Big(\int_{\mathbb{R}} \check{f}(x-y)\varphi(y)dy\Big)dx.\end{aligned}$$

Le théorème de Fubini implique que

$$\begin{aligned}\langle \iota(g), \check{f} \star \varphi\rangle &= \int_{\mathbb{R}\times\mathbb{R}} g(x)f(y-x)\varphi(x)dx \\ &= \int_{\mathbb{R}} g(x)\Big(\int_{\mathbb{R}} f(x-y)\varphi(y)dy\Big)dx \\ &= \int_{\mathbb{R}} (g \star f)(y)\varphi(y)dy.\end{aligned}$$

D'où le résultat.

Les rapports entre convolution et dérivation sont décrits par la proposition suivante.

Proposition 8.2.6. *Soit f dans L^1_S et T dans $\mathcal{S}'$. Pour tout entier k, on a*

$$(f \star T)^{(k)} = f \star T^{(k)}.$$

Démonstration. Par définition de la dérivation et de la convolution sur $\mathcal{S}'$, on a

$$\begin{aligned}\langle (f \star T)^{(k)}, \varphi\rangle &= \langle f \star T, (-1)^k \varphi^{(k)}\rangle \\ &= \langle T, \check{f} \star (-1)^k \varphi^{(k)}\rangle.\end{aligned}$$

D'après la théorème de dérivation des intégrales, nous avons

$$\check{f} \star \varphi^{(k)} = (\check{f} \star \varphi)^{(k)}.$$

À nouveau d'après la définition de la dérivation et de la convolution sur $\mathcal{S}'$, on en déduit que

$$\begin{aligned}\langle (f \star T)^{(k)}, \varphi\rangle &= \langle T^{(k)}, \check{f} \star \varphi\rangle \\ &= \langle f \star T^{(k)}, \varphi\rangle.\end{aligned}$$

D'où la proposition. □

Un point fondamental de la théorie est l'extension de la transformation de Fourier aux distributions tempérées. Au chapitre précédent, nous avons vu que la transformation de Fourier peut être définie sur $L^1(\mathbb{R})$ mais peut être étendue à $L^2(\mathbb{R})$. Le fait que la transformation de Fourier envoie continûment $\mathcal{S}(\mathbb{R})$ dans $\mathcal{S}(\mathbb{R})$ (voir le théorème 7.2.1 page 121) autorise à étendre la transformation de Fourier à $\mathcal{S}'(\mathbb{R})$.

Définition 8.2.6. *On appelle transformation de Fourier sur $\mathcal{S}'$ la transposée de la transformation de Fourier sur $\mathcal{S}$, ce qui s'écrit*

$$\forall \phi \in \mathcal{S}\,,\ \langle \mathcal{F}_{\mathcal{S}'} T, \varphi\rangle = \langle T, \widehat{\varphi}\rangle.$$

La proposition suivante montre qu'il s'agit bien d'une généralisation de la transformation de Fourier.

Proposition 8.2.7. *Pour tout fonction f de $L^1(\mathbb{R})$, on a*

$$\mathcal{F}_{\mathcal{S}'}\big(\iota(f)\big) = \iota(\widehat{f}).$$

Démonstration. Avec les notations du théorème 8.1.1, on a, pour toute fonction ϕ de $\mathcal{S}$,

$$\begin{aligned} \langle \mathcal{F}_{\mathcal{S}'}\iota(f), \varphi\rangle &= \langle \iota(f), \widehat{\varphi}\rangle \\ &= \int_{\mathbb{R}} f(x)\widehat{\varphi}(x)dx. \end{aligned}$$

Vu que la fonction

$$(x, y) \longmapsto f(x)e^{-i(x|\xi)}\varphi(\xi)$$

est intégrable sur $\mathbb{R}\times\mathbb{R}$, on en déduit que

$$\begin{aligned} \langle \mathcal{F}_{\mathcal{S}'}\iota(f), \varphi\rangle &= \int_{\mathbb{R}} \widehat{f}(\xi)\varphi(\xi)d\xi \\ &= \langle \iota(\widehat{f}), \varphi\rangle. \end{aligned}$$

D'où la proposition. □

Nous allons maintenant généraliser les formules de base sur la transformation de Fourier.

Théorème 8.2.4. *Pour toute distribution tempérée T, on a*

$$\mathcal{F}^2_{\mathcal{S}'}T = 2\pi\check{T}\,,\ \big(\mathcal{F}_{\mathcal{S}'}(T)\big)' = -i\mathcal{F}_{\mathcal{S}'}(MT) \quad \textit{et} \quad \mathcal{F}_{\mathcal{S}'}(T') = iM\mathcal{F}_{\mathcal{S}'}T.$$

Démonstration. Pour tout φ de $\mathcal{S}$, on a $\mathcal{F}^2\varphi = 2\pi\check{\varphi}$. Par transposition, nous avons donc, pour tout T de $\mathcal{S}'$,

$$({}^t\mathcal{F})^2T = \mathcal{F}^2_{\mathcal{S}'}T = 2\pi\check{T}.$$

D'où la première formule. Pour φ dans $\mathcal{S}$, on sait d'après le théorème 7.1.1 que

$$\mathcal{F}(i\varphi') = M\widehat{\varphi} \quad \text{et} \quad \mathcal{F}(M\varphi) = -i\frac{d}{d\xi}\widehat{\varphi}.$$

Les deux dernières formules s'en déduisent par transposition. □

Remarque Dorénavant, nous cesserons de distinguer par les notations la transformation de Fourier des fonctions de $L^1(\mathbb{R})$ et la transformation de Fourier des distributions. Pour tout T de $\mathcal{S}'(\mathbb{R})$, nous noterons donc $\widehat{T}$ ou bien $\mathcal{F}(T)$ sa transformée de Fourier.

Comme applications du théorème 8.2.4 ci-dessus, nous allons tout d'abord calculer la transformée de Fourier de la fonction constante 1, puis celle de la distribution tempérée $\mathrm{vp}\frac{1}{x}$ de la proposition 8.1.3.

Proposition 8.2.8. *Nous avons les formules suivantes*

$$\mathcal{F}(\delta_0) = \mathbf{1}\,,\ \mathcal{F}(\mathbf{1}) = 2\pi\delta_0 \quad \textit{et} \quad \mathcal{F}\Big(\mathrm{vp}\frac{1}{x}\Big)(\xi) = -i\pi\frac{\xi}{|\xi|}\,\cdot$$

Démonstration. La première formule résulte simplement du fait que

$$\langle \mathcal{F}(\delta_0), \varphi\rangle = \widehat{\varphi}(0) = \int_{\mathbb{R}} \varphi(x)dx.$$

La seconde résulte de la formule d'inversion de Fourier sur $\mathcal{S}'$ qui donne ici

$$\mathcal{F}^2\delta_0 = \mathcal{F}(\mathbf{1}) = 2\pi\delta_0.$$

Étudions maintenant le cas de la valeur principale. Par définition de la valeur principale, on a, pour toute fonction ϕ de $\mathcal{S}(\mathbb{R})$,

$$\begin{aligned}
\left\langle \mathrm{vp}\frac{1}{x}, x\varphi \right\rangle &= \frac{1}{2}\int_{-\infty}^{\infty} \frac{x\varphi(x) + x\varphi(-x)}{x} dx \\
&= \frac{1}{2}\int_{\mathbb{R}} \big(\varphi(x) + \varphi(-x)\big) dx \\
&= \int_{\mathbb{R}} \varphi(x) dx \\
&= \langle \mathbf{1}, \varphi \rangle.
\end{aligned}$$

Ainsi donc $x\,\mathrm{vp}\frac{1}{x} = 1$. En utilisant le théorème 8.2.4, on en déduit que

$$-\frac{1}{i}\frac{d}{d\xi}\mathcal{F}\Big(\mathrm{vp}\frac{1}{x}\Big) = \mathcal{F}(\mathbf{1}).$$

Or, on sait que $\mathcal{F}\mathbf{1} = 2\pi\delta_0$. Ainsi donc, nous avons

$$\frac{d}{d\xi}\widehat{T} = -2i\pi\delta_0.$$

Ainsi donc, on a, d'après la proposition 8.2.2 et le lemme 8.2.2, nous avons

$$\widehat{T} = -2i\pi H(\xi) + C$$

où C est un scalaire à déterminer. Pour ce faire, observons que la valeur principale est une distribution impaire, c'est-à-dire que

$$\widecheck{\mathrm{vp}\frac{1}{x}} = -\mathrm{vp}\frac{1}{x}\,\cdot$$

De plus, nous laissons au lecteur le soin de vérifier que la transformée de Fourier d'une distribution tempérée impaire est une distribution tempérée impaire. Ainsi donc la distribution $\widehat{T}$ doit être impaire et donc $C = i\pi$. D'où le résultat. □

Étudions maintenant les relations entre convolution, produit et transformation de Fourier. Il s'agit en fait de généraliser (c'est-à-dire transposer) la proposition 7.1.4 page 121 et le corollaire 7.2.2 page 123.

Théorème 8.2.5. *Soit θ une fonction de $\mathcal{S}$ et T une distribution tempérée. On a*

$$\mathcal{F}(\theta \star T) = \widehat{\theta}\widehat{T} \quad \textit{et} \quad \mathcal{F}(\theta T) = (2\pi)^{-1}\widehat{\theta} \star \widehat{T}.$$

Démonstration. Par définition de la transformée de Fourier et de la convolution, on a, pour toute fonction ϕ de $\mathcal{S}$,

$$\begin{aligned}
\langle \mathcal{F}(\theta \star T), \varphi \rangle &= \langle \theta \star T, \widehat{\varphi} \rangle \\
&= \langle T, \check{\theta} \star \widehat{\varphi} \rangle \\
&= \langle \widehat{T}, \mathcal{F}^{-1}(\check{\theta} \star \widehat{\varphi}) \rangle.
\end{aligned}$$

D'après la formule d'inversion de Fourier et la formule de calcul de la transformée de Fourier pour la convoluée des fonctions (voir Proposition 7.1.4 page 121), on a

$$\begin{aligned}
\mathcal{F}^{-1}(\check{\theta} \star \widehat{\varphi})(\xi) &= (2\pi)^{-1}\mathcal{F}(\check{\theta} \star \widehat{\varphi})(-\xi) \\
&= (2\pi)^{-1} \int_{\mathbb{R}\times\mathbb{R}} e^{i(x|\xi)}\theta(y-x)\widehat{\varphi}(y)dydx \\
&= (2\pi)^{-1} \int_{\mathbb{R}\times\mathbb{R}} e^{-i(y-x|\xi)}\theta(y-x)e^{i(y|\xi)}\widehat{\varphi}(y)dydx \\
&= \widehat{\theta}(\xi)\mathcal{F}^{-1}(\widehat{\varphi})(\xi) \\
&= \widehat{\theta}(\xi)\varphi(\xi).
\end{aligned}$$

Par définition de la multiplication, on en déduit que

$$\langle \mathcal{F}(T \star \theta), \varphi \rangle = \langle \widehat{T}, \widehat{\theta}\varphi \rangle = \langle \widehat{\theta}\widehat{T}, \varphi \rangle.$$

D'où la première formule. Pour établir la seconde, appliquons la première à $\mathcal{F}(\check{\theta})$ et $\mathcal{F}(\check{T})$, ce qui donne

$$\mathcal{F}\big(\mathcal{F}(\check{\theta}) \star \mathcal{F}(\check{T})\big) = \mathcal{F}^2\check{\theta}\mathcal{F}^2\check{T} = (2\pi)^2\theta T.$$

En appliquant la transformation de Fourier à cette relation, on trouve que

$$\mathcal{F}^2\big(\mathcal{F}(\check{\theta}) \star \mathcal{F}(\check{T})\big) = (2\pi)^2\mathcal{F}(\theta T).$$

La formule d'inversion de Fourier assure le résultat. □

8.3 Deux exemples d'applications

Comme première application de ce résultat, on peut citer le théorème suivant, qui est une question classique d'analyse harmonique.

Théorème 8.3.1. *Il existe une constante C telle que, pour toute fonction ϕ de $\mathcal{S}(\mathbb{R})$, on ait*

$$\Big\|\varphi \star \mathrm{vp}\frac{1}{x}\Big\|_{L^2} = \pi\|\varphi\|_{L^2}.$$

Démonstration. Utilisons la transformation de Fourier. D'après la proposition 8.2.5, on a

$$\mathcal{F}\Big(\varphi \star \mathrm{vp}\frac{1}{x}\Big) = \widehat{\varphi}\mathcal{F}\Big(\mathrm{vp}\frac{1}{x}\Big) = -i\pi\frac{\xi}{|\xi|}\widehat{\varphi}(\xi).$$

Le théorème de Fourier-Plancherel 7.2.1 page 121 assure que

$$\begin{aligned}
\Big\|\varphi \star \mathrm{vp}\frac{1}{x}\Big\|_{L^2} &= (2\pi)^{\frac{1}{2}}\Big\|\mathcal{F}\Big(\varphi \star \mathrm{vp}\frac{1}{x}\Big)\Big\|_{L^2} \\
&= (2\pi)^{\frac{1}{2}}\Big\|i\pi\frac{\xi}{|\xi|}\widehat{\varphi}(\xi)\Big\|_{L^2} \\
&= \pi\|\varphi\|_{L^2}.
\end{aligned}$$

D'où le théorème. □

Donnons une application à la résolution explicite d'une équation différentielle.

Théorème 8.3.2. *Soit f une distribution tempérée sur $\mathbb{R}$, il existe une unique distribution tempérée u solution de*

$$u - u'' = f$$

dans l'espace $\mathcal{S}'(\mathbb{R})$. Elle est donnée par la formule

$$u = \frac{1}{2} e^{-|\cdot|} \star f.$$

Démonstration. Pour démontrer cela, utilisons le théorème 8.2.4 pour affirmer que

$$\mathcal{F}((\mathrm{Id} - \Delta)u) = (1 + |\xi|^2)\widehat{u}.$$

Le théorème 8.2.4 d'inversion de Fourier permet d'écrire que

$$(\mathrm{Id} - \Delta)u = f \Longleftrightarrow (1 + |\xi|^2)\widehat{u} = \widehat{f}.$$

La fonction $\xi \mapsto (1 + |\xi|^2)^{\pm 1}$ est une fonction indéfiniment différentiable à croissance lente. D'où il vient que

$$\begin{aligned} \widehat{u} &= \frac{1 + |\xi|^2}{1 + |\xi|^2} \widehat{u} \\ &= \frac{1}{1 + |\xi|^2} \widehat{f}. \end{aligned}$$

Donc la solution de l'équation s'écrit

$$u = \mathcal{F}^{-1} \left(\frac{1}{1 + |\xi|^2} \widehat{f} \right).$$

Lorsque $d = 1$, on utilise l'exercice 7.1 page 119 qui affirme que

$$\frac{1}{2} \mathcal{F}(e^{-|\cdot|})(\xi) = \frac{1}{1 + |\xi|^2} \cdot$$

Ceci conclut la démonstration du théorème. □