Notes de cours
Outils de base en analyse appliquée

Hervé Le Dret

19 décembre 2011

Table des matières

Avant-propos

Les notes de cours : mode d'emploi

Les notes qui suivent sont à bien des égards trop fournies. Le cours en amphi n'en traitera qu'une partie, parfois sans entrer dans les détails. Les lecteurs et lectrices curieux et curieuses pourront souvent trouver les détails en question dans ces notes et pourront également, du moins on l'espère, conserver celles-ci pour s'y référer plus tard, dans une étape ultérieure de leur carrière mathématique.

Certains points particulièrement épineux, ou qui peuvent être passés en première voire deuxième lecture, sont indiqués par un signe de danger radioactif dans la marge. Certaines sections sont même totalement radioactives. On ne s'en approchera que muni des protections appropriées.

Chapitre 1

Topologie

1.1 Espaces topologiques

La notion d'espace topologique a mis un certain temps à prendre sa forme axiomatique actuelle. Le but est de formaliser les concepts intuitifs de proximité et de continuité de la façon la plus générale et la plus simple possible. Il n'est pas nécessairement évident que ce but soit atteint quand on lit la définition qui suit pour la première fois de sa vie. L'expérience montre pourtant que c'est bien le cas.

Définition 1.1.1 *Soit X un ensemble. On appelle* topologie *sur X la donnée d'un sous-ensemble $\mathscr{O}$ de l'ensemble $\mathscr{P}(X)$ des parties de X, $\mathscr{O}$ devant vérifier les trois conditions suivantes :*

- *Les ensembles $\emptyset$ et X appartiennent à $\mathscr{O}$,*
- *Soit $(U_\lambda)_{\lambda\in\Lambda}$ une famille quelconque d'éléments de $\mathscr{O}$, alors*

$$\bigcup_{\lambda\in\Lambda} U_\lambda \in \mathscr{O},$$

- *Soit $(U_j)_{1\leq j\leq N}$ une famille finie d'éléments de $\mathscr{O}$, alors*

$$\bigcap_{j=1}^{N} U_j \in \mathscr{O}.$$

Les éléments de l'ensemble $\mathscr{O}$ sont appelés ouverts *(pour la topologie $\mathscr{O}$). Le complémentaire d'un ouvert est appelé un* fermé.

Un couple $(X,\mathscr{O})$ tel que $\mathscr{O}$ vérifie les trois propriétés ci-dessus est appelé espace topologique.

Commentaires. Dans la deuxième propriété portant sur la réunion, il faut bien noter que l'ensemble des indices Λ qui sert à indexer la famille (U_λ) est un ensemble absolument *quelconque*. À ce propos, on rappelle qu'une famille $(U_\lambda)_{\lambda\in\Lambda}$ n'est rien d'autre qu'une application de Λ à valeurs dans $\mathscr{P}(X)$, mais que l'on a décidé de noter U_λ l'image de λ par cette application, et que $\bigcup_{\lambda\in\Lambda} U_\lambda = \{x\in X; \exists\, \lambda \in \Lambda, x\in U_\lambda\}$. Par contre, dans la troisième propriété qui porte sur l'intersection, on ne considère que des familles *finies*, c'est-à-dire $\Lambda = \{1,\dots,N\}$ pour un certain $N\in\mathbb{N}$ [1]. On paraphrase souvent la définition 1.1.1 en disant que l'on a affaire à une topologie sur X si l'ensemble vide et l'ensemble X tout entier sont des ouverts, si une réunion quelconque d'ouverts est un ouvert et si une intersection finie d'ouverts est un ouvert. On dit encore parfois que $\mathscr{O}$ contient l'ensemble vide et X et est stable par réunion quelconque et intersection finie.

Remarquons tout de suite en passant aux complémentaires que l'ensemble vide et X sont aussi des fermés, qu'une intersection quelconque de fermés est un fermé et qu'une réunion finie de fermés est un fermé.

Exemples d'espaces topologiques. Prenons $X=\mathbb{R}$ et désignons par $\mathscr{O}_0$ l'ensemble des parties U de $\mathbb{R}$ telles que, pour tout point x de U, il existe un intervalle ouvert (au sens usuel, c'est-à-dire de la forme $]a,b[$) contenant x et inclus dans U. On a ainsi défini une topologie sur $\mathbb{R}$ (exercice : vérifiez-le !) qui est la topologie usuellement utilisée sur $\mathbb{R}$.

Deux curiosités : soit X un ensemble quelconque, on définit $\mathscr{O}' = \{\emptyset, X\}$ et $\widetilde{\mathscr{O}} = \mathscr{P}(X)$. Dans les deux cas, nous avons défini une topologie sur X. La première topologie s'appelle la *topologie grossière* sur X et la seconde la *topologie discrète* sur X. Ces deux topologies n'ont guère d'intérêt pratique en général, si ce n'est celui de nous convaincre que la définition 1.1.1 n'est pas absurde, puisqu'il existe toujours au moins une topologie sur tout ensemble X ! La vraie question pour les applications est de savoir s'il existe une topologie intéressante sur X, une topologie qui soit bien adaptée à tel ou tel problème que l'on a à résoudre, ce qui est une autre histoire.

Exercice 1.1.1 *Prenons à nouveau $X=\mathbb{R}$.*

1) On définit $\mathscr{O}_d$ comme l'ensemble U des parties de $\mathbb{R}$ telles que, pour tout point x de U, il existe un intervalle $[a,b[$ inclus dans U et contenant x. Démontrez que l'on a bien défini une topologie sur $\mathbb{R}$.

2) On définit $\mathscr{O}_g$ comme l'ensemble U des parties de $\mathbb{R}$ telles que, pour tout point x de U, il existe un intervalle $]a,b]$ inclus dans U et contenant x. Démontrez que l'on a bien défini une topologie sur $\mathbb{R}$.

---

1. Notons que $\Lambda=\emptyset$ définit également une famille finie, à zéro éléments, et que $\cap_{i\in\emptyset} U_i = X$, donc ce cas apparemment particulier, et qui n'apparaît pas dans la définition telle qu'elle est écrite plus haut parce qu'il ne présente pas grand intérêt, n'en est pas un. Dans le même ordre d'idées, $\cup_{\lambda\in\emptyset} U_\lambda = \emptyset$, et la première propriété de topologie peut être considérée comme redondante, à une petite réécriture près. Mais on l'écrit traditionnellement à part et nous suivons la tradition.

3) Les trois ensembles $\mathscr{O}_0$, $\mathscr{O}_d$ et $\mathscr{O}_g$ sont-ils égaux ? Peut-on les comparer au sens de l'inclusion ?

Le langage des voisinages est souvent utilisé dans la littérature. Voici la définition d'un voisinage d'un point dans le cadre des espaces topologiques.

Définition 1.1.2 *Soit x un point d'un espace topologique $(X,\mathscr{O})$. On dit qu'une partie V de X est un* voisinage *de x si et seulement si il existe un ouvert contenant x et inclus dans V.*

Remarquons qu'un voisinage de x n'est jamais vide, que si V est un voisinage de x alors tout ensemble qui contient V est aussi un voisinage de x et que l'intersection d'un nombre fini de voisinages de x est un voisinage de x (on dit que l'ensemble de voisinages d'un point forme un *filtre*). Les voisinages d'un même point sont donc naturellement partiellement ordonnés par l'inclusion. La notion de voisinage permet alors de quantifier l'idée intuitive pour un point y de X d'être « plus ou moins près » d'un élément donné x, au sens de la topologie considérée naturellement, un voisinage inclus dans un autre traduisant une plus grande proximité. Il faut faire attention à ce qu'un voisinage n'est pas en général un ouvert. Il est possible d'axiomatiser la topologie uniquement en termes de voisinages, plutôt qu'en termes d'ouverts, cette façon équivalente de voir les choses étant juste un peu moins courante.

On définit également un voisinage d'une partie quelconque A de X en remplaçant le point x de la définition ci-dessus par la partie A.

Remarque de vocabulaire. On utilise couramment l'expression « telle ou telle propriété a lieu au voisinage de x ou de A » pour signifier qu'il existe un voisinage de x ou de A où la propriété en question a lieu.

À l'aide du langage des voisinages, on peut caractériser les ouverts de la manière suivante.

Proposition 1.1.1 *Soit $(X,\mathscr{O})$ un espace topologique. Une partie U de X est un ouvert si et seulement si, pour tout x appartenant à U, U est un voisinage de x.*

Démonstration. Soit U un ouvert de X et x un point de U, alors U est évidemment un voisinage de x, on prend comme ouvert contenant x l'ensemble U lui-même. Réciproquement, soit $U \subset X$ un ensemble qui est voisinage de chacun de ses éléments. Soit $\mathscr{V}$ l'ensemble des ouverts inclus dans U. On pose $\tilde{U} = \cup_{V\in\mathscr{V}} V$, c'est un ouvert puisque réunion d'ouverts et il est inclus dans U puisque chaque V l'est. Or pour tout $x \in U$, il existe par définition d'un voisinage un ouvert $V \subset U$ tel que $x \in V$. Par conséquent, $x \in \tilde{U}$, ce qui implique que $\tilde{U} = U$ et U est donc ouvert. $\square$

Nous allons maintenant définir la notion de sous-espace topologique. Ce concept essentiel est fondé sur la proposition suivante.

Proposition 1.1.2 *Soit A une partie d'un espace topologique $(X, \mathcal{O})$, on considère l'ensemble $\mathcal{O}_A$ des parties V de A pour lesquelles il existe U appartenant à $\mathcal{O}$ tel que $V = A \cap U$. Le couple $(A, \mathcal{O}_A)$ est un espace topologique.*

Démonstration. Tout d'abord, il est clair que $\emptyset = \emptyset \cap A$ et $A = X \cap A$ appartiennent à $\mathcal{O}_A$. Considérons ensuite une famille quelconque $(V_\lambda)_{\lambda \in \Lambda}$ de $\mathcal{O}_A$. Par définition, pour tout $\lambda \in \Lambda$, il existe U_λ dans $\mathcal{O}$ tel que $V_\lambda = A \cap U_\lambda$. Mais l'on a

$$\bigcup_{\lambda \in \Lambda} U_\lambda \in \mathcal{O} \quad \text{donc} \quad \bigcup_{\lambda \in \Lambda} V_\lambda = \bigcup_{\lambda \in \Lambda} (A \cap U_\lambda) = A \cap \bigcup_{\lambda \in \Lambda} U_\lambda \in \mathcal{O}_A.$$

Soit maintenant une famille finie $(V_j)_{1 \leq j \leq N}$ d'éléments de $\mathcal{O}_A$. Il existe une famille finie $(U_j)_{1 \leq j \leq \mathbb{N}}$ d'éléments de $\mathcal{O}$ telle que $V_j = U_j \cap A$. Mais l'on a

$$\bigcap_{j=1}^{N} U_j \in \mathcal{O} \quad \text{donc} \quad \bigcap_{j=1}^{N} V_j = \bigcap_{j=1}^{N} (A \cap U_j) = A \cap \bigcap_{j=1}^{N} U_j \in \mathcal{O}_A.$$

Les trois propriétés requises pour parler de topologie sont donc bien satisfaites. □

Convention : Sauf mention expresse du contraire, on considère toujours une partie d'un espace topologique comme munie de cette topologie, dite topologie induite. Attention, les ouverts de A pour la topologie induite ne sont manifestement pas nécessairement des ouverts de X ! Plus précisément :

Exercice 1.1.2 *Soient $(X, \mathcal{O})$ un espace topologique et A une partie de X. Démontrez que $\mathcal{O}_A$ est inclus dans $\mathcal{O}$ si et seulement si A est un ouvert de X. En déduire que si A est fermé, les fermés de A sont des fermés de X.*

Définition 1.1.3 *On dit qu'un espace topologique $(X, \mathcal{O})$ est* séparé *si et seulement si pour tout couple (x, y) de points distincts de X, il existe deux ouverts U et V tels que*

$$x \in U\,,\ y \in V \quad \textit{et} \quad U \cap V = \emptyset.$$

Commentaires. Un espace topologique séparé est aussi dit être de Hausdorff. Il existe toute une zoologie de propriétés de séparation plus ou moins fortes. Celle qui nous intéresse ici signifie que l'on peut toujours inclure deux points distincts de X dans deux ouverts disjoints, qui les séparent donc au sens de la topologie considérée.

À titre d'exemple, le lecteur vérifiera que la topologie grossière $\mathcal{O}' = \{\emptyset, X\}$ n'est pas séparée dès que X a plus d'un élément alors que la topologie discrète est séparée. Dans un espace topologique séparé, tout singleton $\{x\}$ est fermé, et le complémentaire d'un singleton est ouvert. Dans les applications, nous ne considérerons essentiellement jamais d'espace topologique qui ne soit pas séparé, ce qui ne signifie pas que ces derniers soient dénués d'intérêt.

Exercice 1.1.3 *Démontrer que les trois topologies de l'exercice* 1.1.1 *sont séparées.*

Définition 1.1.4 *Soient* $(X,\mathscr{O})$ *un espace topologique et* A *une partie de* X*. On appelle* adhérence *de* A *et l'on note* $\mathrm{Adh}(A)$ *ou bien* $\bar{A}$ *l'ensemble*

$$\bar{A} = \{x \in X; \forall U \in \mathscr{O} \text{ tel que } x \in U, A \cap U \neq \emptyset\}.$$

On appelle intérieur *de* A *et l'on note* $\mathrm{Int}(A)$ *ou bien* $\mathring{A}$ *l'ensemble*

$$\mathring{A} = \{x \in A; \exists\, U \in \mathscr{O} \text{ tel que } x \in U \text{ et } U \subset A\}.$$

On appelle frontière *d'une partie* A *de* X *l'ensemble* $\partial A = \bar{A} \setminus \mathring{A}$.
On dit enfin qu'une partie A *de* X *est* dense *si et seulement si* $\bar{A} = X$.

Commentaires. L'adhérence d'une partie A est donc l'ensemble des points qui « collent » à A au sens où tout voisinage d'un de ces points contient aussi au moins un point de A, alors que l'intérieur de A est l'ensemble des points bien « calés » dans A puisqu'au moins un de leurs voisinages est inclus dans A.

Exemples. L'intérieur des intervalles $]a,b[$, $[a,b[$, $]a,b]$ et $[a,b]$ pour la topologie usuelle $\mathscr{O}_0$ sur $\mathbb{R}$ est l'intervalle $]a,b[$. Leur adhérence est l'intervalle $[a,b]$. Leur frontière est l'ensemble $\{a,b\}$. Qu'en est-il pour les deux autres topologies $\mathscr{O}_d$ et $\mathscr{O}_g$?

Cet exemple ne doit pas faire croire que l'intérieur d'une partie A « remplit » toujours pratiquement A et que ∂A est toujours un ensemble « petit ». Ainsi, on sait qu'un intervalle ouvert non vide de $\mathbb{R}$ contient toujours au moins un nombre rationnel et un nombre irrationnel (en fait une infinité de chaque). Par conséquent, l'ensemble $\mathbb{Q}$ des nombres rationnels est d'intérieur vide pour la topologie usuelle de $\mathbb{R}$, il est dense dans $\mathbb{R}$ et sa frontière est égale à $\mathbb{R}$, et la même chose est vraie pour son complémentaire, l'ensemble $\mathbb{R} \setminus \mathbb{Q}$ des nombres irrationnels.

Proposition 1.1.3 *L'adhérence de* A *est l'intersection de tous les fermés contenant* A *et c'est le plus petit (au sens de l'inclusion) fermé qui contient* A*. L'intérieur de* A *est la réunion de tous les ouverts inclus dans* A *et c'est le plus grand ouvert inclus dans* A*.*

Démonstration. Montrons d'abord que $\bar{A}$ est un fermé car son complémentaire est ouvert. Si $x \in X \setminus \bar{A}$, il vient en niant la définition de $\bar{A}$ qu'il existe un ouvert U contenant x et tel que $A \cap U = \emptyset$. Donc $U \subset X \setminus \bar{A}$, lequel est par conséquent ouvert.

Soit F l'intersection de tous les fermés contenant A. C'est un fermé comme intersection de fermés, et comme il contient A, c'est le plus petit d'entre eux. En particulier, $F \subset \bar{A}$ car $\bar{A}$ contient évidemment A. Montrons que $\bar{A} \subset F$. Le complémentaire de F est un ouvert et si $x \in X \setminus F$, il existe un ouvert U contenant x tel que $F \cap U = \emptyset$. A fortiori, $A \cap U = \emptyset$, et donc $x \neq \bar{A}$, c'est-à-dire $\bar{A} \subset F$.

De même, $\mathring{A}$ est un ouvert, puisque si $x \in \mathring{A}$, il existe un ouvert U contenant x et inclus dans A, donc tout autre point y de U est aussi dans $\mathring{A}$. Soit V la réunion de tous les ouverts inclus dans A. C'est un ouvert comme réunion d'ouverts, et comme il est contenu dans A, c'est le plus grand d'entre eux. En particulier, $\mathring{A} \subset V$ car $\mathring{A}$ est évidemment contenu dans A. Montrons que $V \subset \mathring{A}$. C'est évident car si $x \in V \subset A$, on prend $U = V$ dans la définition de l'intérieur, et $x \in \mathring{A}$. □

Commentaires. On aurait tout aussi bien pu définir l'adhérence de A comme le plus petit fermé contenant A et son intérieur comme le plus grand ouvert inclus dans A, et en déduire les caractérisations de la définition 1.1.4. On a aussi clairement, $\mathring{A} = X \setminus \mathrm{Adh}(X \setminus A)$, $\bar{A} = \mathring{A} \cup \partial A = X \setminus \mathrm{Int}(X \setminus A)$ et $\partial A = \partial(X \setminus A)$. Par ailleurs, la frontière d'une partie A est toujours un fermé car $\partial A = \bar{A} \cap (X \setminus \mathring{A})$. Enfin une partie est dense si et seulement si son complémentaire est d'intérieur vide.

L'intérêt du concept de topologie tel qu'il a été présenté ici, est qu'il permet de définir de manière très abstraite, donc très générale et simple des notions familières.

Définition 1.1.5 *Soient $(X, \mathscr{O})$ et $(Y, \widetilde{\mathscr{O}})$ deux espaces topologiques et f une application de X dans Y. Considérons un point x_0 de X. On dit que 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 contenant x_0 tel que*

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

Commentaires. En d'autres termes, f est continue en x_0 si pour tout voisinage V de $f(x_0)$, on peut trouver un voisinage de x_0 dont l'image par f est incluse dans le voisinage V, c'est-à-dire, intuitivement, qu'au voisinage de x_0, l'application f prend des valeurs aussi proches que l'on veut de $f(x_0)$ au sens de la topologie de Y à condition de se placer suffisamment près de x_0 au sens de la topologie de X.

Théorème 1.1.1 *Soient $(X, \mathscr{O})$ et $(Y, \widetilde{\mathscr{O}})$ deux espaces topologiques et f une application de X dans Y. L'application f est continue en tout point x de X si et seulement si*

$$\forall V \in \widetilde{\mathscr{O}},\ f^{-1}(V) \in \mathscr{O},$$

c'est-à-dire si et seulement si l'image réciproque de tout ouvert est un ouvert. De même, f est continue en tout point de X si et seulement si l'image réciproque de tout fermé de Y est un fermé de X. On dit alors que f est continue *sur X.*

Démonstration. Supposons que l'image réciproque de tout ouvert de $(Y, \widetilde{\mathscr{O}})$ soit un ouvert de $(X, \mathscr{O})$. Considérons alors un point x de X et un ouvert V de Y contenant $f(x)$. L'ensemble $f^{-1}(V)$ est un ouvert de X qui n'est pas vide, puisqu'il contient bien sûr x. De plus, on a toujours (pour n'importe quelle application f et n'importe quel ensemble V, indépendamment de toute considération de topologie) que $f(f^{-1}(V)) \subset V$. Donc f est continue en x.

Réciproquement, supposons que l'application f soit continue en tout point de X. Soit V un ouvert de Y quelconque. Alors soit $f^{-1}(V)$ est vide, auquel cas c'est bien un ouvert de X, soit il n'est pas vide et l'on en considère un point arbitraire x, pour lequel on a naturellement $f(x) \in V$. L'application f étant continue en x, il existe donc un ouvert U de X contenant x tel que $f(U) \subset V$. On en déduit que

$$U \subset f^{-1}(f(U)) \subset f^{-1}(V).$$

Par conséquent, $f^{-1}(V)$ est un voisinage de x et donc $f^{-1}(V)$ est ouvert par la proposition 1.1.1. Le cas des fermés se traite aisément par passage au complémentaire (exercice : écrivez-le en détail !). □

Remarque. Quelles que soient les topologies choisies sur X et Y, il est clair que les fonctions constantes sont toujours continues.

Proposition 1.1.4 *Soit f une application continue de $(X, \mathscr{O})$ dans $(Y, \widetilde{\mathscr{O}})$ et $A \subset X$. Alors la restriction $f_{|A}$ de f à A est continue de $(A, \mathscr{O}_A)$ dans $(Y, \widetilde{\mathscr{O}})$.*

Démonstration. On utilise la caractérisation précédente. Soit V un ouvert de Y. On a

$$\left(f_{|A}\right)^{-1}(V) = \{x \in A; f(x) \in V\} = f^{-1}(V) \cap A.$$

Comme $f^{-1}(V)$ est un ouvert de X, on en déduit que $\left(f_{|A}\right)^{-1}(V)$ est un élément de $\mathscr{O}_A$, c'est-à-dire un ouvert de la topologie induite. □

Théorème 1.1.2 *Soient $(X, \mathscr{O})$, $(Y, \widetilde{\mathscr{O}})$ et $(Z, \mathscr{O}')$ trois espaces topologiques. On considère une application f de X dans Y et une application g de Y dans Z. Soit x_0 un point de X. Si f est continue en x_0 et si g l'est en $f(x_0)$, alors $g \circ f$ est continue en x_0.*

Démonstration. Soit W un ouvert de Z contenant $g \circ f(x_0)$. Comme g est continue en $f(x_0)$, il existe un ouvert V de Y contenant $f(x_0)$ tel que $g(V) \subset W$. Comme f est continue en x_0, il existe un ouvert U de X contenant x_0 tel que $f(U) \subset V$. Prenant son image par g, il vient que

$$g \circ f(U) \subset g(V) \subset W,$$

d'où le théorème. □

Commentaire. On en déduit immédiatement que si f est continue sur X à valeurs dans Y et g est continue sur Y à valeurs dans Z (munis de leur topologie respective), alors $g \circ f$ est continue sur X à valeurs dans Z. Ce résultat est fondamental en analyse.

Exercice 1.1.4 *1) Démontrez que les fonctions continues de* $(\mathbb{R}, \mathscr{O}_0)$ *dans* $(\mathbb{R}, \mathscr{O}_0)$ *sont les fonctions continues usuelles.*

2) Démontrez que les fonctions continues de $(\mathbb{R}, \mathscr{O}_g)$ *dans* $(\mathbb{R}, \mathscr{O}_0)$ *sont les fonctions continues à gauche. Même chose pour la droite.*

3) Que dire des fonctions continues de $(\mathbb{R}, \mathscr{O}_0)$ *dans* $(\mathbb{R}, \mathscr{O}_g)$ *?*

Définition 1.1.6 *Soient* $(X, \mathscr{O})$ *et* $(Y, \widetilde{\mathscr{O}})$ *deux espaces topologiques et f une application de X dans Y bijective. On dit que f est un* homéomorphisme *de X sur Y si f et son inverse f^{-1} sont continues. On dit que* $(X, \mathscr{O})$ *et* $(Y, \widetilde{\mathscr{O}})$ *sont* homéomorphes *s'il existe un homéomorphisme de X sur Y.*

Commentaire. Les images directe et réciproque d'un ouvert par un homéomorphisme sont des ouverts de leurs espaces respectifs. On peut donc « lire » la topologie de Y sur la topologie de X à travers l'application f. Par conséquent, du strict point de vue de la topologie, deux espaces homéomorphes $(X, \mathscr{O})$ et $(Y, \widetilde{\mathscr{O}})$ se comportent exactement de la même façon. Ils partagent les mêmes propriétés topologiques.

Les topologies sur X étant des sous-ensembles particuliers de l'ensemble des parties de X, elles sont naturellement ordonnées par l'inclusion.

Définition 1.1.7 *Soient $\mathscr{O}_1$ et $\mathscr{O}_2$ deux topologies sur le même ensemble X. On dit que $\mathscr{O}_1$ est* plus forte *que $\mathscr{O}_2$ (ou encore est* plus fine*) si $\mathscr{O}_2 \subset \mathscr{O}_1$.*

Commentaires. Dans ce cas, $\mathscr{O}_1$ possède plus d'ouverts que $\mathscr{O}_2$. Donc toute application de X à valeurs dans un autre espace topologique Y qui est continue pour $\mathscr{O}_2$ est aussi continue pour $\mathscr{O}_1$, alors que l'inverse n'est pas forcément vrai. De même, toute application d'un autre espace topologique Y à valeurs dans X qui est continue pour $\mathscr{O}_1$ est aussi continue pour $\mathscr{O}_2$. La topologie discrète est plus forte que toute topologie et la topologie grossière est moins forte que toute topologie sur X. Enfin, deux topologies peuvent parfaitement ne pas être ordonnées au sens ci-dessus car il s'agit d'une relation d'ordre partiel. Néanmoins, lorsqu'en analyse on aura à considérer deux topologies différentes sur un même ensemble, l'une sera (presque) toujours plus forte que l'autre.

De nombreuses notions de topologie ont une expression en termes de suites, lesquelles sont parfois plus faciles à manipuler que les ouverts. Néanmoins, les suites sont un concept fondamentalement dénombrable, ce qui ne permet pas de couvrir toutes les situations possibles en topologie générale. Dans la pratique de l'analyse, les suites sont bien souvent suffisantes. On rappelle qu'une suite d'éléments de X n'est rien d'autre qu'une application de $\mathbb{N}$ dans X, mais que l'on utilise la notation indicielle plutôt que la notation fonctionnelle.

Définition 1.1.8 *Soit $(X,\mathscr{O})$ un espace topologique et $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de X. On dit que la suite $(x_n)_{n\in\mathbb{N}}$* converge *vers un élément ℓ de X si et seulement si*

$$\forall V \in \mathscr{O} \text{ tel que } \ell \in V\,, \exists\, n_0 \text{ tel que } \forall\, n \geq n_0\,,\ x_n \in V. \tag{1.1}$$

On dit que ℓ est limite *de la suite (x_n) et l'on note la convergence*

$$\ell = \lim_{n\to+\infty} x_n \quad \text{ou bien} \quad x_n \underset{n\to+\infty}{\longrightarrow} \ell,$$

la topologie étant en général sous-entendue, sauf quand on a affaire à plusieurs topologies sur un même espace.

Commentaire. Une suite converge donc vers une limite ℓ si et seulement si pour tout voisinage de cette limite, à partir d'un certain rang, la suite ne sort jamais plus de ce voisinage.

Proposition 1.1.5 *Soit $(X,\mathscr{O})$ un espace topologique séparé et $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de X. Il existe au plus un élément ℓ vérifiant la condition* (1.1) *ci-dessus.*

Démonstration. On raisonne par l'absurde. Supposons que $\ell_1 \neq \ell_2$ vérifient la condition (1.1). Comme l'espace X est séparé, nous pouvons choisir deux ouverts U_1 et U_2 tels que $\ell_j \in U_j$ et $U_1 \cap U_2 = \emptyset$. Mais, il existe n_1 tel que, pour tout $n \geq n_1$, on ait $x_n \in U_1$ et n_2 tel que, pour tout $n \geq n_2$, on ait $x_n \in U_2$. Par conséquent, pour tout $n \geq n_0 = \max(n_1,n_2)$, $x_n \in U_1 \cap U_2$, ce qui est impossible puisque cet ensemble est vide. Contradiction. □

Commentaire. Dans un espace topologique séparé, on a donc *unicité* de la limite d'une suite convergente. Ce n'est pas forcément le cas dans un espace topologique qui n'est pas séparé.

Proposition 1.1.6 *Soient $(X,\mathscr{O})$ et $(Y,\widetilde{\mathscr{O}})$ deux espaces topologiques. On considère une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X convergeant vers ℓ et une application f de X dans Y continue au point ℓ. Alors*

$$\lim_{n\to+\infty} f(x_n) = f(\ell).$$

Démonstration. Soit V un ouvert de Y qui contient $f(\ell)$. L'image réciproque $U = f^{-1}(V)$ est un ouvert de X qui contient ℓ. Par conséquent, il existe $n_0 \in \mathbb{N}$ tel que pour tout $n \geq n_0$, $x_n \in U$. Par conséquent, $f(x_n) \in V$ pour tout $n \geq n_0$. □

Attention, la réciproque n'est pas vraie en général, c'est-à-dire qu'une application qui satisfait cette propriété sur les suites — on dit qu'elle est *séquentiellement continue* — n'est pas forcément continue, en raison de questions de dénombrabilité comme indiqué plus haut. Cette réciproque est quand même vraie dans des cas importants, comme nous le verrons plus loin.

Définition 1.1.9 *Soient $(X,\mathscr{O})$ un espace topologique et $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de X. On appelle* valeur d'adhérence *de la suite $(x_n)_{n\in\mathbb{N}}$ tout élément y de X tel que*

$$\forall U \in \mathscr{O} \text{ tel que } y \in U, \forall n \in \mathbb{N}, \exists\, m \geq n \text{ tel que } x_m \in U.$$

Commentaires. Intuitivement, cela signifie que, si y est une valeur d'adhérence de la suite $(x_n)_{n\in\mathbb{N}}$, alors cette suite repasse une infinité de fois aussi près que l'on veut de y.

Proposition 1.1.7 *Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments d'un espace topologique $(X,\mathscr{O})$, l'ensemble des valeurs d'adhérence de la suite $(x_n)_{n\in\mathbb{N}}$ est le fermé*

$$\mathscr{A} = \bigcap_{n\in\mathbb{N}} \mathrm{Adh}\{x_m,\ m \geq n\}.$$

Démonstration. Pour démontrer cela, il suffit d'écrire que

$$\begin{aligned} y \in \mathscr{A} &\iff \forall n,\ y \in \mathrm{Adh}\{x_m,\ m \geq n\} \\ &\iff \forall n,\ \forall U \in \mathscr{O} \text{ tel que } y \in U, \exists\, m \geq n \text{ tel que } x_m \in U, \end{aligned}$$

par définition de l'adhérence d'un ensemble. □

Nous allons maintenant introduire une classe particulière très importante d'espaces topologiques.

1.2 Espaces métriques

On commence par définir la notion de distance.

Définition 1.2.1 *Soit X un ensemble, on appelle* distance *sur X toute application d de $X \times X$ dans $\mathbb{R}_+$ telle que*

$$\begin{aligned} d(x,y) &= 0 \Leftrightarrow x = y, \\ d(x,y) &= d(y,x), && \textit{(symétrie)} \\ d(x,y) &\leq d(x,z) + d(z,y,) && \textit{(inégalité triangulaire)} \end{aligned}$$

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

Commentaires. Les trois axiomes de distance sont la formalisation de propriétés intuitives de la distance « physique » de l'espace dans lequel nous vivons, à savoir que la distance entre deux points est nulle si et seulement si ces deux points sont confondus, que la distance entre deux points ne dépend pas de l'ordre dans lequel on considère ces deux points et qu'il est plus court d'aller directement du point x au point y que d'y aller en passant par z. Ceci dit, en effectuant cette formalisation, on a considérablement étendu la généralité de la notion, dont on verra qu'elle s'applique très largement.

Définition 1.2.2 *Soient (X,d) un espace métrique, x un point de X et α un réel positif.*

On appelle boule ouverte (*resp.* fermée) *de centre x et de rayon α, et l'on note $B(x,\alpha)$ (resp. $\bar{B}(x,\alpha)$) l'ensemble des points y de X tels que $d(x,y) < \alpha$ (resp. $d(x,y) \leq \alpha$).*

Remarque. Par définition, $B(x,0) = \emptyset$ et $\bar{B}(x,0) = \{x\}$. De même, il est clair que $B(x,\alpha) \subset \bar{B}(x,\alpha)$.

Définition 1.2.3 *Soit A une partie de X. On dit que A est une partie* bornée *de X si et seulement si il existe un réel strictement positif α et un point x de X tels que A soit inclus dans $B(x,\alpha)$.*

Exercice 1.2.1 *Soit A une partie non vide et bornée d'un espace métrique (X,d). Démontrez qu'il existe un plus petit réel δ tel que*

$$\forall (x,y) \in A \times A\,,\; d(x,y) \leq \delta.$$

On note $\delta(A)$ ce réel et on l'appelle le diamètre de A.

Quelques exemples (à vérifier)

- Soit X un ensemble quelconque et $d\colon X^2 \to \mathbb{R}_+$ définie par $d(x,x) = 0$ et $d(x,y) = 1$ dès que $x \neq y$. Alors, (X,d) est un espace métrique. La distance d est appelée distance discrète. Quelles en sont les boules ouvertes et fermées ?
- Prenons $X = \mathbb{R}$ et $d(x,y) = |x-y|$. Cela définit un espace métrique. La distance d est la distance usuellement utilisée sur $\mathbb{R}$ en analyse.
- Prenons $X = \mathbb{R}^N$. Les expressions suivantes définissent différentes distances sur X :

$$d_p(x,y) \stackrel{\text{déf}}{=} \Bigl(\sum_{j=1}^{N} |x_j - y_j|^p\Bigr)^{\frac{1}{p}}$$

$$d_\infty(x,y) \stackrel{\text{déf}}{=} \max_{1\leq j\leq N} |x_j - y_j|$$

pour $1 \leq p < +\infty$. Ces distances sur $\mathbb{R}^N$ sont couramment utilisées en analyse. Pour $p = 2$ et $N = 1, 2$ ou 3, la première distance est la distance « physique » mentionnée plus haut (pour $N = 1$, toutes ces distances coïncident en fait avec la distance usuelle). On l'appelle la distance euclidienne.

– 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.
– Les espaces vectoriels normés sont des espaces métriques que nous étudierons plus en détail au chapitre 2.

Un espace métrique est automatiquement muni d'une topologie naturelle qui est induite par sa structure métrique.

Théorème 1.2.1 *Soit (X,d) un espace métrique. Soit $\mathscr{O}$ l'ensemble des parties U de X telles que, pour tout $x \in U$, il existe un réel strictement positif α tel que $B(x,\alpha)$ soit incluse dans U. Alors $\mathscr{O}$ est une topologie sur X.*

Démonstration. Le premier point de la définition d'une topologie est évidemment vérifié. En effet, si $U = \emptyset$, la condition à vérifier est vide, donc vérifiée.

Soient $(U_\lambda)_{\lambda\in\Lambda}$ une famille quelconque d'éléments de $\mathscr{O}$ et un point x de X tel que
$$x \in \bigcup_{\lambda\in\Lambda} U_\lambda.$$
Par définition de la réunion, il existe un indice $\lambda \in \Lambda$ tel que $x \in U_\lambda$. Par définition de $\mathscr{O}$, il existe une boule ouverte $B(x,\alpha)$ incluse dans U_λ, donc *a fortiori* dans la réunion des U_λ. Par conséquent, $\bigcup_{\lambda\in\Lambda} U_\lambda \in \mathscr{O}$.

Soient maintenant une famille finie $(U_j)_{1\leq j\leq N}$ d'éléments de $\mathscr{O}$ et x un point de X tel que
$$x \in \bigcap_{j=1}^{N} U_j.$$
Par définition de $\mathscr{O}$, pour tout j, il existe un réel strictement positif α_j tel que $B(x,\alpha_j)$ soit incluse dans U_j. Considérons $\alpha = \min_{1\leq j\leq N} \alpha_j$. Comme c'est le minimum d'un nombre fini de nombres strictement positifs, on a $\alpha > 0$. Comme $\alpha \leq \alpha_j$, on a $B(x,\alpha) \subset B(x,\alpha_j)$, et la boule ouverte $B(x,\alpha)$ est incluse dans chacun des U_j donc dans leur intersection. On a donc bien défini une topologie sur X. $\square$

Convention : Le plus souvent et sauf mention expresse du contraire, on munira tout espace métrique de la topologie ainsi définie. On dit que c'est la *topologie métrique* associée à la distance d.

Remarques : On vérifie facilement que la topologie associée à la distance discrète sur X n'est autre que la topologie discrète.

On vérifie également facilement que si $A \subset X$ est une partie d'un espace métrique, alors la topologie induite sur A par la topologie métrique de X n'est autre que la topologie métrique associée à la restriction de la distance d à $A \times A$ et que les boules de A sont les intersections des boules de X avec A.

Dans la démonstration ci-dessus, on n'a pas utilisé les axiomes de distance, mais seulement le fait que si deux boules ouvertes sont centrées en un même point, alors la boule de plus petit rayon est incluse dans la boule de plus grand rayon. Pour cela, n'importe quelle application de $X \times X$ dans $\mathbb{R}_+$ aurait fait l'affaire. La proposition cruciale suivante fait par contre fondamentalement appel aux deux derniers axiomes de distance, symétrie et inégalité triangulaire.

Proposition 1.2.1 *Soit (X,d) un espace métrique. Pour tout x de X et pour tout réel $\alpha \geq 0$, la boule $B(x,\alpha)$ est un ouvert et la boule $\bar{B}(x,\alpha)$ est un fermé.*

Démonstration. Commençons par les boules ouvertes. Le cas $\alpha = 0$ ne pose pas de difficulté puisque $\emptyset$ est un ouvert. Soit donc $\alpha > 0$. Soit y un élément de $B(x,\alpha)$, on choisit un réel strictement positif β tel que

$$\beta < \alpha - d(x,y).$$

On a alors, pour tout élément z de X tel que $d(z,y) < \beta$

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

ce qui signifie que $B(y,\beta) \subset B(x,\alpha)$. Par conséquent, $B(x,\alpha)$ est un ouvert.

Pour les boules fermées, on prend $\alpha \geq 0$ et y un élément du complémentaire de $\bar{B}(x,\alpha)$. Comme $d(x,y) > \alpha$, on peut choisir un réel strictement positif γ tel que

$$\gamma < d(x,y) - \alpha.$$

On a alors, pour tout élément z de X tel que $d(z,x) < \gamma$,

$$d(x,z) \geq d(x,y) - d(z,y) > \alpha,$$

ce qui veut dire que $B(y,\gamma) \subset X \setminus \bar{B}(x,\alpha)$. Par conséquent, l'ensemble $X \setminus \bar{B}(x,\alpha)$ est un ouvert. □

Commentaires. Il est facile de voir que $B(x,\alpha) \subset \mathrm{Int}(\bar{B}(x,\alpha))$, $\mathrm{Adh}(B(x,\alpha)) \subset \bar{B}(x,\alpha)$. L'ensemble $\{y \in X; d(x,y) = \alpha\}$ s'appelle la sphère de centre x et de rayon α. Les deux inclusions précédentes peuvent être strictes, la sphère peut être vide et elle n'est pas forcément la frontière de l'une ou l'autre boule de même rayon, *cf.* la topologie discrète.

Le premier axiome de distance intervient dans la proposition suivante.

Théorème 1.2.2 *La topologie d'un espace métrique est séparée.*

Soient x et y deux points distincts de X, on a donc $d(x,y) > 0$. Par l'inégalité triangulaire, on a pour tout z tel que $d(x,z) < \frac{d(x,y)}{2}$,

$$d(z,y) \geq d(x,y) - d(x,z) > \frac{d(x,y)}{2}.$$

De même, si $d(y,z) < \frac{d(x,y)}{2}$ alors $d(x,z) > \frac{d(x,y)}{2}$ et par conséquent,

$$B\left(x, \frac{d(x,y)}{2}\right) \cap B\left(y, \frac{d(x,y)}{2}\right) = \emptyset.$$

Comme les boules ouvertes sont des ouverts de $\mathscr{O}$, l'espace topologique $(X,\mathscr{O})$ est donc séparé. $\square$

Définition 1.2.4 *On dit qu'un espace topologique $(X,\mathscr{O})$ est* métrisable *s'il existe une distance sur X dont la topologie métrique associée coïncide avec $\mathscr{O}$.*

Commentaire. L'intérêt, pas forcément apparent au premier coup d'œil, de cette définition est qu'il existe certaines situations dans lesquelles une topologie est définie d'une façon naturelle, en en décrivant les ouverts ou les voisinages par exemple, et où il existe une distance qui engendre la même topologie, mais qui ne s'écrit pas de façon naturelle. Dans les applications, on n'utilisera donc jamais la distance explicitement, mais par contre on saura que l'espace topologique considéré a les mêmes propriétés topologiques qu'un espace métrique, ce qui peut être utile. Il existe naturellement des topologies qui ne sont pas métrisables et qui sont utilisées dans la pratique.

Exercice 1.2.2 *Soient (X,d) un espace métrique et f une fonction strictement croissante de $\mathbb{R}^+$ dans lui-même, continue en 0 et telle que*

$$f(0) = 0 \quad \textit{et} \quad f(x+y) \leq f(x) + f(y).$$

Montrer que l'application d_f définie par

$$d_f(x,y) = f(d(x,y))$$

est une distance sur X et que les topologies définies par d et d_f sont égales.

Dans les espaces métriques, la convergence des suites peut ne s'exprimer qu'en termes de boules, c'est-à-dire de distance, comme en témoigne la proposition suivante.

Proposition 1.2.2 *Soit (X,d) un espace métrique, considérons une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X et ℓ un point de X. Alors*

$$\begin{aligned}\lim_{n\to\infty} x_n = \ell &\iff \forall \varepsilon > 0,\ \exists\, n_0 \text{ tel que } \forall n \geq n_0,\ x_n \in B(\ell,\varepsilon),\\ &\iff \forall \varepsilon > 0,\ \exists\, n_0 \text{ tel que } \forall n \geq n_0,\ d(x_n,\ell) < \varepsilon,\\ &\iff d(x_n,\ell) \to 0 \text{ quand } n \to +\infty,\end{aligned}$$

la dernière convergence ayant lieu dans $\mathbb{R}$ muni de sa topologie usuelle.

À titre d'exercice, la démonstration est laissée au lecteur qui aura certainement reconnu dans le cas où $X = \mathbb{R}$ la définition usuelle de la limite d'une suite à l'aide de la distance usuelle.

Proposition 1.2.3 *Soient (X,d) un espace métrique et A une partie de X. Alors l'adhérence de A est égale à l'ensemble des limites de toutes les suites convergentes d'éléments de A*

$$\bar{A} = \{x \in X, \exists\, (a_n)_{n\in\mathbb{N}} \text{ avec } a_n \in A \text{ telle que } \lim_{n\to+\infty} a_n = x\}.$$

Démonstration. Supposons que x soit limite d'une suite d'éléments de A. Pour tout ouvert U contenant x, il existe un n_0 tel que x_{n_0} appartienne à U. Donc $x \in \bar{A}$. Réciproquement, supposons que $x \in \bar{A}$. Alors, pour tout entier strictement positif n, l'ensemble $A \cap B(x,n^{-1})$ n'est pas vide et l'on choisit un élément, que l'on note a_n, dans cet ensemble. Soit $(a_n)_{n\in\mathbb{N}}$ la suite ainsi définie. Par construction, $a_n \in A$ et $d(a_n,x) < 1/n$, d'où $d(a_n,x) \to 0$ quand $n \to +\infty$, c'est-à-dire $\lim_{n\to+\infty} a_n = x$.□

La notion de valeur d'adhérence d'une suite s'exprime également de façon plus simple dans le cadre des espaces métriques.

Proposition 1.2.4 *Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments d'un espace métrique (X,d). Un point y de X est une valeur d'adhérence de la suite $(x_n)_{n\in\mathbb{N}}$ si et seulement si il existe une application ϕ de $\mathbb{N}$ dans $\mathbb{N}$ strictement croissante telle que*

$$\lim_{p\to+\infty} x_{\phi(p)} = y.$$

Démonstration. La proposition 1.1.7 dit que

$$y \in \bigcap_{n\in\mathbb{N}} \mathrm{Adh}\{x_m,\ m \geq n\}.$$

Ceci signifie que, pour tout entier n,

$$\exists\, m \geq n \text{ tel que } d(x_m,y) \leq \frac{1}{n}\cdot \tag{1.2}$$

Soit l'application $\phi : \mathbb{N} \to \mathbb{N}$ définie par récurrence par

$$\begin{aligned} \phi(0) &= 0 \quad \text{et} \\ \phi(p) &= \min\{m \in \mathbb{N} \text{ qui satisfont (1.2) avec } n = \phi(p-1)+1\}. \end{aligned}$$

L'application ϕ est bien définie, elle est par construction strictement croissante de $\mathbb{N}$ dans lui-même, donc $\phi(p) \to +\infty$ quand $p \to +\infty$. De plus, on a

$$d(x_{\phi(p)}, y) \leq \frac{1}{\phi(p-1)+1} \to 0 \text{ quand } p \to +\infty.$$

Réciproquement, si y est limite d'une suite de la forme $(x_{\phi(p)})_{p\in\mathbb{N}}$, alors y est trivialement une valeur d'adhérence de la suite initiale. □

Commentaire. Une suite de la forme $(x_{\phi(p)})_{p\in\mathbb{N}}$ avec ϕ strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ s'appelle *sous-suite* ou *suite extraite* de la suite originale. En effet, elle s'obtient en sélectionnant certains éléments de la suite originale et en les numérotant en ordre croissant de 0 à $+\infty$. On utilise parfois la notation $(x_{n_p})_{p\in\mathbb{N}}$ pour désigner une sous-suite, $(n_p)_{p\in\mathbb{N}}$ étant une suite strictement croissante d'entiers.

La continuité des applications peut, elle aussi, s'exprimer uniquement en termes de boules ou de distance.

Proposition 1.2.5 *Soient (X,d) et (Y,δ) deux espaces métriques. On considère une application f de X dans Y et un point x_0 de X.*

i) L'application f est continue en x_0 si et seulement si

$$\forall \varepsilon > 0, \exists\, \alpha > 0 \text{ tel que } d(x,x_0) < \alpha \Rightarrow \delta(f(x), f(x_0)) < \varepsilon.$$

ii) L'application f est continue en x_0 si et seulement si, pour toute suite $(x_n)_{n\in\mathbb{N}}$ convergeant vers x_0, on a

$$f(x_n) \underset{n\to+\infty}{\longrightarrow} f(x_0).$$

Démonstration. En effet, supposons la propriété i) ci-dessus. Soit V un ouvert contenant $f(x_0)$. Par définition des ouverts, il existe une boule ouverte de centre $f(x_0)$ et de rayon ε incluse dans V. On choisit alors comme ouvert U la boule de centre x_0 et de rayon α donnée par l'hypothèse.

Réciproquement, supposons f continue au point x_0. La boule $B(f(x_0),\varepsilon)$ étant un ouvert, il existe un ouvert U contenant x_0 tel que $f(U)$ soit inclus dans $B(f(x_0),\varepsilon)$. Par définition des ouverts, il existe un réel α strictement positif tel que $B(x_0,\alpha)$ soit incluse dans U. Le premier point de la proposition est alors démontré.

Pour le point ii), on sait déjà que f continue entraîne la convergence annoncée. Pour la réciproque, on raisonne par contraposition en considérant une application f qui n'est pas continue en x_0. Il existe donc un ouvert V de Y contenant $f(x_0)$ tel que pour tout ouvert U de X contenant X_0, $f(U) \not\subset V$. Prenons pour U la boule ouverte $B(x_0, 1/n)$. Comme $f(B(x_0, 1/n)) \not\subset V$, il existe $x_n \in B(x_0, 1/n)$ tel que $f(x_n) \notin V$. Par construction, cette suite converge dans X vers x_0, mais son image $f(x_n)$ ne converge pas vers $f(x_0)$ dans Y. $\square$

Commentaire. On reconnaîtra dans la caractérisation i) la définition usuelle de la continuité d'une fonction quand $X = Y = \mathbb{R}$ et $d = \delta$ est la distance usuelle sur $\mathbb{R}$. Grâce aux notions topologiques introduites, on a donc considérablement étendu le champ d'application de cette définition.

L'une des richesses de la structure d'espace métrique est de pouvoir décrire de manière beaucoup plus précise les multiples façons dont une application peut être continue.

Définition 1.2.5 *Soient (X,d) et (Y,δ) deux espaces métriques ; on considère une application f de X dans Y.*

– *On dit que l'application f est* uniformément continue *si et seulement si*

$$\forall \varepsilon > 0, \exists\, \alpha > 0 \text{ tel que } d(x,x') < \alpha \Rightarrow \delta(f(x), f(x')) < \varepsilon;$$

– *On dit que l'application f est* localement hölderienne d'exposant $r \in]0,1]$ *si et seulement si*

$$\forall x \in X, \exists\, m,k > 0, d(x,x') \leq m \Rightarrow \delta(f(x), f(x')) \leq k d(x,x')^r.$$

Lorsque $r = 1$, on dit que l'application f est localement k-lipschitzienne.

– *On dit que l'application f est* (globalement) hölderienne d'exposant $r \in]0,1]$ *si et seulement si*

$$\exists\, k > 0, \forall (x,x') \in X \times X, \delta(f(x), f(x')) \leq k d(x,x')^r.$$

Lorsque $r = 1$, on dit que f est (globalement) k-lipschitzienne. On dit aussi simplement f est lipschitzienne si l'on ne souhaite pas préciser la constante de Lipschitz k de l'application f.

Commentaire. Noter la ressemblance entre la continuité en x_0 du i) de la proposition 1.2.5 et la définition de la continuité uniforme. La différence est qu'ici x' n'est pas fixé. Pour un ε donné, la constante α est la même pour tous les x', d'où le nom de continuité uniforme. Par ailleurs, on a clairement f k-lipschitzienne $\Rightarrow$ f localement k-lipschitzienne $\Rightarrow$ f localement hölderienne $\Rightarrow$ f uniformément continue $\Rightarrow$ f continue. Aucune des implications réciproques n'est vraie (trouver des exemples).

Exercice 1.2.3 *Soit x_0 un point d'un espace métrique (X,d). Démontrez que l'application définie par*

$$x \mapsto d(x, x_0)$$

est 1-lipschitzienne de X dans $\mathbb{R}$.

Exercice 1.2.4 *Démontrez que l'application définie sur* $\mathbb{R}$ *par*

$$x \mapsto |x|^r \quad avec \quad r \in]0,1]$$

est hölderienne d'exposant r.

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

Définition 1.2.6 *Soit X un ensemble, on considère deux distances d et δ sur X. On dit que les deux distances sont topologiquement (resp. uniformément, resp. fortement) équivalentes si et seulement si l'application* Id *est continue (resp. uniformément continue, resp. lipschitzienne) de (X,d) dans (X,δ) et de (X,δ) dans (X,d).*

Remarque. Pour deux distances topologiquement équivalentes, l'application identité est un homéomorphisme de (X,d) dans (X,δ). Les topologies métriques associées, donc les ouverts, les applications continues et les suites convergentes sont les mêmes. Par contre, deux distances qui ne sont pas topologiquement équivalentes n'engendrent pas les mêmes topologies sur X.

Proposition 1.2.6 *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

$$\forall x, \forall \varepsilon, \exists\, \eta \text{ tel que } d(x,y) < \eta \Longrightarrow \delta(x,y) < \varepsilon \text{ et } \delta(x,y) < \eta \Longrightarrow d(x,y) < \varepsilon.$$

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

$$\forall \varepsilon, \exists\, \eta \text{ tel que } d(x,y) < \eta \Longrightarrow \delta(x,y) < \varepsilon \text{ et } \delta(x,y) < \eta \Longrightarrow d(x,y) < \varepsilon.$$

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

$$\exists\, C > 0 \text{ tel que } \forall (x,y) \in X \times X, \quad C^{-1} d(x,y) \leq \delta(x,y) \leq C d(x,y).$$

La démonstration de cette proposition n'est qu'une application immédiate des définitions ; elle est laissée au lecteur. D'après la caractérisation de la continuité à l'aide des suites, pour vérifier que deux distances sont topologiquement équivalentes, il suffit encore de montrer que

$$d(x_n,x) \to 0 \Longleftrightarrow \delta(x_n,x) \to 0.$$

On pourra également vérifier avec profit que les deux distances définies précédemment sur $\mathbb{R}^N$ sont fortement équivalentes. Elles définissent donc une seule et même topologie sur $\mathbb{R}^N$.

Théorème 1.2.3 *Soient (X,d) un espace métrique et f et g deux fonctions continues de X à valeurs dans $\mathbb{R}$ muni de sa distance usuelle. Alors les fonctions $f+g$ et fg sont continues. Si de plus f est à valeurs dans $\mathbb{R}^*$, alors $1/f$ est continue.*

Démonstration. On définit une application $F\colon X \to \mathbb{R}^2$ par $F(x) = (f(x), g(x))$. Il est facile de voir, en utilisant par exemple la proposition 1.2.5, ii), que F est continue quand on munit $\mathbb{R}^2$ de la topologie métrique usuelle. Par ailleurs, il est tout aussi facile de voir que les applications

$$S\colon \begin{array}{l} \mathbb{R}^2 \to \mathbb{R} \\ (u_1,u_2) \mapsto u_1+u_2 \end{array} \qquad P\colon \begin{array}{l} \mathbb{R}^2 \to \mathbb{R} \\ (u_1,u_2) \mapsto u_1u_2 \end{array}$$

sont également continues. Par conséquent, les applications $x \mapsto f(x)+g(x)$ et $x \mapsto f(x)g(x)$ sont continues de X dans $\mathbb{R}$ comme composées d'applications continues, respectivement $S\circ F$ et $P \circ F$, *cf.* théorème 1.1.2. De même, l'application $t \mapsto 1/t$ est continue de $\mathbb{R}^*$ dans $\mathbb{R}$, donc si f est continue et ne s'annule pas, $1/f$ est continue de X dans $\mathbb{R}$. □

Commentaires. Le fait que X soit un espace métrique n'a en fait aucune importance. Le résultat reste vrai si X est un espace topologique quelconque, comme on peut s'en convaincre aisément. Par ailleurs, on déduit immédiatement du théorème que si on a un nombre fini $f_1, f_2, \dots f_p$ de fonctions continues de X dans $\mathbb{R}$, alors les fonctions $\sum_{i=1}^p f_i$ et $\prod_{i=1}^p f_i$ sont également continues. Enfin, on voit aisément grâce au théorème que l'espace $C(X;\mathbb{R})$ des fonctions continues de X à valeurs dans $\mathbb{R}$ est un espace vectoriel pour l'addition des applications et leur multiplication par un scalaire. C'est aussi un anneau et en fait une algèbre en y ajoutant la multiplication des applications. Ceci fait entrevoir la raison de l'importance cruciale de l'algèbre linéaire en analyse.

Nous allons maintenant exposer quelques propriétés supplémentaires spécifiques des espaces métriques, au sens où elles ne sont pas nécessairement vraies dans un espace topologique général.

Définition 1.2.7 *Soient A une partie non vide d'un espace métrique (X,d) et x un point de X. On appelle distance du point x à l'ensemble A la quantité*

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

Exercice 1.2.5 *Soit A une partie non vide d'un espace métrique (X,d). Démontrez que l'application*

$$x \mapsto d(x,A)$$

est 1-lipschitzienne. Démontrez que $\bar{A}$ est l'ensemble des points x de X tels que $d(x,A) = 0$.

Une conséquence de ce résultat est que si A est fermé, alors $x \in A$ si et seulement si $d(x,A) = 0$.

Dans un espace métrique, on peut séparer deux fermés disjoints au sens suivant.

Proposition 1.2.7 *Soient A et B deux parties fermées non vides disjointes d'un espace métrique (X,d). Il existe une fonction continue de (X,d) dans $[0,1]$ telle que $f_{|A} = 0$ et $f_{|B} = 1$ et $0 < f(x) < 1$ si $x \notin A \cup B$.*

Démonstration. Il suffit de vérifier que la fonction f définie par

$$f(x) \stackrel{\text{déf}}{=} \frac{d(x,A)}{d(x,A) + d(x,B)}$$

convient bien. □

Une conséquence importante de ce résultat est la possibilité de construire des partitions de l'unité. Voici ce dont il s'agit.

Théorème 1.2.4 *Soient (X,d) un espace métrique et $(U_n)_{n\in\mathbb{N}}$ une famille d'ouverts localement finie (ce qui signifie que, pour tout point x de X, il existe un voisinage V de x tel que $U_n \cap V = \emptyset$ sauf pour un nombre fini d'indices n) tel que*

$$\bigcup_{n\in\mathbb{N}} U_n = X.$$

Alors il existe une famille de fonctions continues $(f_n)_{n\in\mathbb{N}}$ de (X,d) dans l'intervalle $[0,1]$ telle que

$$\forall x \in X, \sum_{n\in\mathbb{N}} f_n(x) = 1 \quad \text{et} \quad f_{n|X\setminus U_n} = 0.$$

Démonstration. Supposons d'abord qu'il existe un n_0 tel que $X \setminus U_{n_0} = \emptyset$, c'est-à-dire $U_{n_0} = X$. On note que, pour l'indice n_0, la deuxième condition est vide et il suffit de prendre $f_{n_0} = 1$ et $f_n = 0$ pour tout $n \neq n_0$, ce qui répond à la question dans ce cas particulier.

On se place maintenant dans le cas plus intéressant où $X \setminus U_n \neq \emptyset$ pour tout n. On définit la famille de fermés $(F_n)_{n\in\mathbb{N}}$ par

$$F_n \stackrel{\text{déf}}{=} X \setminus \bigcup_{m\neq n} U_m.$$

Comme $\bigcup_{n\in\mathbb{N}} U_n = X$, on a clairement $F_n \cap (X \setminus U_n) = \emptyset$. Si $F_n \neq \emptyset$, les ensembles F_n et $X \setminus U_n$ sont deux fermés non vides disjoints auxquels on peut appliquer la proposition précédente, ce qui fournit une famille de fonctions continues $(\varphi_n)_{n\in\mathbb{N}}$ à valeurs dans $[0,1]$ telles que

$$\varphi_n = 1 \text{ sur } F_n \quad \text{et} \quad \varphi_n = 0 \text{ sur } X \setminus U_n.$$

Enfin, si $F_n = \emptyset$ on pose $\varphi_n = 0$.

Soit x un élément de X, il existe un ouvert V contenant x et tel que l'ensemble $\{n \in \mathbb{N}; U_n \cap V \neq \emptyset\}$ soit un ensemble fini. La série de fonctions $\sum_n \varphi_n$ est donc en fait, sur l'ouvert V, une somme finie de fonctions continues, donc une fonction continue sur V. Par conséquent, x étant quelconque, on a ainsi défini une fonction continue S sur X par

$$S(x) \stackrel{\text{déf}}{=} \sum_{n\in\mathbb{N}} \varphi_n(x),$$

la somme ne contenant qu'un nombre fini de termes non nuls en tout point x (nombre qui peut bien sûr varier avec x).

Soit $I = \{n \in \mathbb{N}; F_n = \emptyset\}$ l'ensemble des indices pour lesquels $\bigcup_{m\neq n} U_m = X$. On voit que $\bigcup_{n\in\mathbb{N}\setminus I} U_n = X$ car pour tout $n \in I$ et tout point x, $x \in U_m$ pour un certain $m \neq n$. On en déduit que

$$\forall x \in X, \quad S(x) > 0.$$

En effet, on a aussi $S(x) = \sum_{n\in\mathbb{N}\setminus I} \varphi_n(x)$ et si x est tel que $S(x) = 0$, cela implique que $x \in X \setminus U_n$ pour tout $n \in \mathbb{N} \setminus I$. Or

$$\bigcap_{n\in\mathbb{N}\setminus I} (X \setminus U_n) = X \setminus \Big(\bigcup_{n\in\mathbb{N}\setminus I} U_n\Big) = \emptyset,$$

donc il n'existe aucun point x où S s'annule. On pose alors pour tout $n \in \mathbb{N}$

$$f_n \stackrel{\text{déf}}{=} \frac{\varphi_n}{S},$$

ce qui définit une famille de fonctions continues qui satisfont les conditions du théorème. $\square$

Commentaires. Le terme partition de l'unité signifie donc que l'on a découpé la fonction constante 1 en une somme de fonctions continues à valeurs dans $[0,1]$ et telles que f_n s'annule en dehors de U_n. On dit encore que f_n est *à support dans* U_n. On utilise les résultats de partition de l'unité de la façon suivante. Soit g une fonction continue sur X à valeurs dans $\mathbb{R}$. Comme $g = g \times 1 = g \times (\sum_{n=0}^{+\infty} f_n) = \sum_{n=0}^{+\infty} g_n$, avec $g_n = gf_n$ qui est une fonction continue sur X et à support dans U_n. On a de la sorte *localisé* la fonction g en la découpant en morceaux à support dans des ouverts donnés, ce qui est souvent très utile.

Notons enfin que les théorèmes de séparation des fermés et de partition de l'unité sont des exemples de propriétés purement topologiques que possèdent les espaces métriques. Par conséquent, si l'on sait qu'un espace est métrisable, alors on sait qu'il possède ces mêmes propriétés sans avoir besoin d'avoir recours explicitement à la distance.

1.3 Espaces complets

Nous allons aborder l'étude d'une classe particulièrement importante d'espaces métriques.

Définition 1.3.1 *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 qui satisfait la condition*

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

Commentaire. Une suite de Cauchy est donc telle que deux éléments quelconques de la suite sont arbitrairement proches l'un de l'autre, à condition de se placer suffisamment loin dans la suite.

Proposition 1.3.1 *Toute suite de Cauchy d'un espace métrique est bornée.*

Démonstration. En effet, si la suite $(x_n)_{n\in\mathbb{N}}$ est de Cauchy, il existe un entier n_0 tel que, pour tout $n \geq n_0$ et tout $m \geq n_0$, on ait $d(x_n, x_m) \leq 1$ (on a pris $\varepsilon = 1$). Ainsi donc, on a

$$d(x_{n_0}, x_m) \leq \max\{1, \max_{m<n_0} d(x_0, x_m)\} \stackrel{\text{déf}}{=} \alpha.$$

D'où $x_m \in B(x_{n_0}, \alpha)$ pour tout $m \in \mathbb{N}$. □

Proposition 1.3.2 *Soient d et δ deux distances uniformément équivalentes sur X. Toute suite de Cauchy pour l'une est une suite de Cauchy pour l'autre.*

Démonstration. En effet, considérons une suite $(x_n)_{n\in\mathbb{N}}$ de Cauchy pour la distance d. Pour tout réel strictement positif ε, il existe un réel strictement positif η tel que

$$d(x,y) < \eta \Longrightarrow \delta(x,y) < \varepsilon \quad \text{et} \quad \delta(x,y) < \eta \Longrightarrow d(x,y) < \varepsilon.$$

La suite $(x_n)_{n\in\mathbb{N}}$ étant de Cauchy, il existe un entier n_0 tel que

$$\forall\, n \geq n_0,\ \forall\, m \geq n_0,\ d(x_n, x_m) < \eta.$$

Par l'équivalence uniforme des distances, il existe ainsi un entier n_0 tel que

$$\forall\, n \geq n_0, \forall\, m \geq n_0, \delta(x_n, x_m) < \varepsilon,$$

d'où la proposition. □

Attention, si l'on a affaire à deux distances qui sont seulement topologiquement équivalentes, alors une suite de Cauchy pour l'une des distances peut parfaitement ne pas être de Cauchy pour l'autre distance (à titre d'exercice, trouver un exemple de ceci sur $X = \mathbb{R}$). La notion de suite de Cauchy n'est donc pas une notion purement topologique (on parle plus généralement de *structure uniforme* pour axiomatiser la continuité uniforme et introduire des objets qui jouent le rôle des suites de Cauchy et de la complétude que nous verrons dans un moment, dans un cadre ensembliste plus étendu que le cadre métrique).

Les deux notions de suite convergente et de suite de Cauchy sont étroitement liées.

Proposition 1.3.3 *i) Dans un espace métrique (X,d), toute suite convergente est de Cauchy.*

ii) Toute suite de Cauchy ayant une valeur d'adhérence ℓ converge vers ℓ.

Démonstration. Point i) : considérons une suite $(x_n)_{n\in\mathbb{N}}$ convergente vers une limite ℓ. Pour tout ε strictement positif, il existe un entier n_0 tel que l'on ait

$$\forall n \geq n_0\,,\ d(x_n,\ell) < \frac{\varepsilon}{2}\cdot$$

Donc, si n et m sont deux entiers plus grands que n_0, on a

$$\begin{aligned} d(x_n,x_m) &\leq d(x_n,\ell)+d(\ell,x_m)\\ &\leq \varepsilon. \end{aligned}$$

Point ii) : considérons maintenant 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. □

Si une suite convergente est toujours de Cauchy, la réciproque n'est pas vraie en général, c'est-à-dire qu'il peut exister des suites de Cauchy sans valeur d'adhérence. On est donc amené à poser la définition suivante, qui est extrêmement importante comme on pourra s'en rendre compte dans la suite.

Définition 1.3.2 *Soit* (X,d) *un espace métrique, on dit que cet espace est* complet *si et seulement si toute suite de Cauchy est convergente.*

Commentaire. L'intérêt majeur de la notion d'espace complet vient du fait que l'on sait y démontrer la convergence d'une suite en utilisant une propriété qui ne fait pas intervenir la limite, sur laquelle en général on ne sait rien *a priori*, ni même qu'elle existe ! La définition de la convergence d'une suite fait par contre elle intervenir la limite. Par conséquent, démontrer qu'une suite dans un espace complet est de Cauchy est un moyen très puissant de démontrer *l'existence même* de cette limite. C'est pourquoi cette idée est à la base d'un très grand nombre de résultats de convergence, après avoir démontré une fois pour toutes que l'espace dans lequel on travaille est complet.

Proposition 1.3.4 *Soient* d *et* δ *deux distances uniformément équivalentes sur* X. *Si* (X,d) *est complet, alors* (X,δ) *l'est aussi.*

Démonstration. En effet, soit $(x_n)_{n\in\mathbb{N}}$ une suite de Cauchy de (X,δ), d'après la proposition 1.3.2, c'est une suite de Cauchy pour la distance d. Mais comme (X,d) est complet, c'est une suite convergente pour d, donc pour δ puisque les distances d et δ étant uniformément équivalentes, elles sont topologiquement équivalentes. □

Un exemple fondamental d'espace complet est l'espace $\mathbb{R}$ muni de la distance usuelle $d(x,y)=|x-y|$, ce qui découle de la propriété de borne supérieure satisfaite par $\mathbb{R}$ (le démontrer). Un exemple simple d'espace non complet est son sous-espace $\mathbb{Q}$, muni de la même distance (considérer une suite de nombres rationnels qui converge vers $\sqrt{2}$, par exemple la suite des nombres décimaux 1, 1,4, 1,41, 1,414,...).

Nous allons maintenant examiner comment fabriquer des espaces complets à partir d'espaces complets déja connus. Autrement dit, on va chercher des opérations sur les ensembles qui laissent stable la propriété d'espace complet.

Proposition 1.3.5 *Soit* $(X_1,d_1),\cdots,(X_N,d_N)$ *une famille de* N *espaces métriques complets. Soit* $X=X_1\times\cdots\times X_N$ *muni de la distance*

$$d((x_1,\cdots,x_N),(y_1,\cdots,y_N))=\max_{1\le j\le N} d_j(x_j,y_j),$$

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

La démonstration est laissée en exercice.

Exercice 1.3.1 *Démontrer le même résultat en prenant*

$$\delta((x_1,\cdots,x_N),(y_1,\cdots,y_N))=\sum_{j=1}^{N} d_j(x_j,y_j)$$

ou bien encore

$$\widetilde{d}((x_1,\cdots,x_N),(y_1,\cdots,y_N)) = \Big(\sum_{j=1}^{N} d_j(x_j,y_j)^2\Big)^{\frac{1}{2}}.$$

En corollaire, on voit que $\mathbb{R}^N$ est complet pour chacune des trois distances que l'on y a introduites.

On dit qu'une application f d'un ensemble X à valeurs dans un espace métrique (Y,δ) est bornée si l'ensemble $f(X)$ est borné dans Y.

Proposition 1.3.6 *Soient X un ensemble quelconque et (Y,δ) un espace métrique, désignons par $\mathscr{B}(X,Y)$ (resp. si X est un espace topologique $\mathscr{C}(X,Y)$) l'espace des applications bornées (resp. bornées et continues) de X dans Y et définissons une distance D sur $\mathscr{B}(X,Y)$ par*

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

Si l'espace (Y,δ) est complet, alors les espaces $(\mathscr{B}(X,Y),D)$ et $(\mathscr{C}(X,Y),D)$ le sont aussi.

Démonstration. On vérifie tout d'abord aisément que l'application D satisfait les trois axiomes de distance, et donc que les espaces $(\mathscr{B}(X,Y),D)$ et $(\mathscr{C}(X,Y),D)$ sont bien des espaces métriques.

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

$$\forall \varepsilon > 0, \exists\, n_0 \text{ tel que } \forall n \geq n_0, \forall m \geq n_0, \forall x \in X,\ \delta(f_n(x),f_m(x)) \leq D(f_n,f_m) < \varepsilon. \tag{1.3}$$

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

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

Vérifions que $f \in \mathscr{B}(X,Y)$. D'après l'inégalité (1.3) appliquée avec $\varepsilon = 1$, on a

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

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

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

L'application f_{n_0} étant bornée, on en déduit que f l'est également par l'inégalité triangulaire. Il faut maintenant vérifier que la suite $(f_n)_{n\in\mathbb{N}}$ converge vers f dans l'espace $(\mathscr{B}(X,Y),D)$. Pour ce faire, passons à la limite lorsque p tend vers l'infini dans l'inégalité (1.3), ce qui donne

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

ce qui achève la démonstration de la proposition pour les applications bornées en prenant la borne supérieure par rapport à x.

Dans le cas où X est un espace topologique, on a $\mathscr{C}(X,Y) \subset \mathscr{B}(X,Y)$, avec la même distance D, donc on sait déjà que toute suite de Cauchy $(f_n)_{n\in\mathbb{N}}$ dans $\mathscr{C}(X,Y)$ converge dans $\mathscr{B}(X,Y)$ vers un certain f. Il nous reste à montrer que sa limite f est continue. Soit donc $x_0 \in X$ et $B(f(x_0), \alpha)$ une boule ouverte de Y avec $\alpha > 0$. Par l'inégalité triangulaire

$$\delta(f(x), f(x_0)) \leq \delta(f(x), f_n(x)) + \delta(f_n(x), f_n(x_0)) + \delta(f_n(x_0), f(x_0))$$

pour tout x et n. Comme $f_n \to f$, on peut choisir $n \in \mathbb{N}$ tel que $D(f_n, f) < \frac{\alpha}{3}$. Pour cette valeur de n, il vient donc

$$\delta(f(x), f(x_0)) < \frac{2\alpha}{3} + \delta(f_n(x), f_n(x_0)).$$

La fonction f_n est continue. Il existe donc un ouvert U de X contenant x_0 tel que pour tout $x \in U$, $\delta(f_n(x), f_n(x_0)) < \frac{\alpha}{3}$. On en déduit donc que pour tout $x \in U$,

$$\delta(f(x), f(x_0)) < \alpha,$$

c'est-à-dire que f est continue en x_0. Comme x_0 est arbitraire, on a bien $f \in \mathscr{C}(X,Y)$. □

Commentaire. Le lecteur attentif aura reconnu la généralisation aux espaces métriques de la notion de convergence uniforme d'une suite de fonctions et du fait qu'une limite uniforme de fonctions continues est continue.

Proposition 1.3.7 *Soient (X,d) un espace complet et A une partie de X. Alors l'espace (A,d) est complet si et seulement si A est une partie fermée de X.*

Démonstration. Supposons que A soit fermée. Soit $(a_n)_{n\in\mathbb{N}}$ une suite de Cauchy d'éléments de A. C'est aussi une suite de Cauchy d'éléments de X. Donc, puisque X est complet, elle converge vers un élément a_∞ de X. Cet élément de X est limite d'une suite d'éléments de A qui est un fermé de X. Donc a_∞ appartient à A, ce qui prouve que A est complet.

Réciproquement, supposons que A soit complet. Considérons un élément x de X tel qu'il existe une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A dont la limite soit x. La suite $(a_n)_{n\in\mathbb{N}}$ est une suite convergente (dans X), c'est donc une suite de Cauchy dans X, donc aussi dans A. Comme A est complet, la suite $(a_n)_{n\in\mathbb{N}}$ converge dans A. L'unicité de la limite dans l'espace X entraîne que $x \in A$, d'où la proposition. □

Une propriété fort utile des espaces complets est la possibilité de prolonger des applications uniformément continues définies sur une partie dense.

Théorème 1.3.1 *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* $\tilde{f}$ *de* (X,d) *dans* (Y,δ) *telle que* $\tilde{f}_{|A} = f$*.*

Démonstration. Considérons un élément x de X et tâchons de définir $\tilde{f}(x)$. L'ensemble A étant dense, la proposition 1.2.4 nous assure qu'il existe une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A convergeant vers x. L'application f est uniformément continue, ce qui signifie que

$$\forall\,\varepsilon > 0, \exists\,\alpha \text{ tel que } \forall\,(a,b) \in A^2\,,\; d(a,b) < \alpha \Rightarrow \delta(f(a),f(b)) < \varepsilon. \qquad (1.4)$$

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

$$\forall\,n \geq n_0, \forall\,m \geq n_0,\; d(a_n,a_m) < \alpha.$$

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

La première chose à vérifier est que 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 l'application 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 $(a_n)_{n\in\mathbb{N}}$. Définissons alors sans ambiguïté l'application $\tilde{f}$ par

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

Pour tout $x \in A$, prenant la suite constante $a_n = x$, on en déduit immédiatement que $\tilde{f}(x) = f(x)$, c'est-à-dire la propriété de prolongement.

Vérifions maintenant que l'application $\tilde{f}$ est uniformément continue sur X. Considérons un couple (x,y) d'éléments de X tels que $d(x,y) < \alpha$. Par la densité de A dans X, il existe 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

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

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

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

ce qui achève la démonstration du théorème. □

Commentaire. Il est aisé de construire des fonctions continues sur une partie dense et non prolongeables par continuité. Considérons par exemple $f\colon \mathbb{R}^* \to \mathbb{R}$ définie par $f(x) = 0$ si $x < 0$ et $f(x) = 1$ si $x > 0$. Cette fonction est trivialement continue sur $\mathbb{R}^*$ (le vérifier), mais visiblement pas prolongeable par continuité à $\mathbb{R}$, et bien sûr non uniformément continue sur $\mathbb{R}^*$.

Le théorème suivant, dit de point fixe de Picard, est important, car il permet de montrer l'existence et l'unicité de la solution de certaines équations de manière assez générale et constructive.

Théorème 1.3.2 *Soit f une application d'un espace métrique complet (X,d) dans lui-même telle qu'il existe un réel k de l'intervalle* $]0,1[$ *tel que*

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

(*une telle application est dite* strictement contractante). *Il existe un unique élément $z \in X$ point fixe de f, c'est-à-dire 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

$$x_{n+1} = f(x_n).$$

On peut écrire que

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

En itérant cette inégalité, on obtient

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

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

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

Comme $0 < k < 1$, la suite $(x_n)_{n\in\mathbb{N}}$ est donc de Cauchy. Comme X est complet, elle est convergente. Soit z sa limite. L'application f étant 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 $z = f(z)$, donc z est un point fixe de f.

Il nous reste à démontrer l'unicité. Soient z_1 et z_2 deux points fixes de f. Comme f est strictement contractante

$$d(z_1,z_2) \leq kd(z_1,z_2) \Leftrightarrow (1-k)d(z_1,z_2) \leq 0.$$

Comme $k < 1$, on a $1-k > 0$, donc $d(z_1,z_2) = 0$, c'est-à-dire $z_1 = z_2$. Le théorème est ainsi démontré. □

Le théorème de Baire qui suit est un théorème profond, dont nous verrons par exemple aux chapitres 2 et 4 qu'il implique des propriétés que l'on serait bien en peine de démontrer autrement.

Théorème 1.3.3 *Soit (X,d) un espace métrique complet, on considère une suite $(U_n)_{n\in\mathbb{N}}$ d'ouverts denses dans X. Alors $\bigcap_{n\in\mathbb{N}} U_n$ est dense.*

Démonstration. Soit V un ouvert de X, nous allons démontrer que $\left(\bigcap_n U_n\right) \cap V \neq \emptyset$, ce qui assurera le théorème.

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

$$\bar{B}(x_0,\alpha_0) \subset U_0 \cap V, \tag{1.5}$$

(bien noter la boule fermée au membre de gauche). 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

$$\bar{B}(x_1,\alpha_1) \subset U_1 \cap B(x_0,\alpha_0).$$

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

$$\alpha_j \leq \frac{1}{j+1} \quad \text{et} \quad \bar{B}(x_j,\alpha_j) \subset U_j \cap B(x_{j-1},\alpha_{j-1}). \tag{1.6}$$

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

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

On a ainsi construit par récurrence une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X et une suite $(\alpha_n)_{n\in\mathbb{N}}$ de réels strictement positifs qui tend vers 0 et telle que

$$\forall m \geq n\,,\; x_m \in B(x_n, \alpha_n). \tag{1.7}$$

Comme la suite $(\alpha_n)_{n\in\mathbb{N}}$ tend vers 0, la suite $(x_n)_{n\in\mathbb{N}}$ est de Cauchy. Donc elle converge dans X qui est complet. Appelons x sa limite. D'après la relation (1.7), le point x appartient à $\bar{B}(x_n, \alpha_n)$ pour tout entier n. Vu les relations (1.5) et (1.6), on a

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

Le théorème de Baire est ainsi démontré. □

Une utilisation courante du théorème de Baire consiste à dire que dans un espace complet, une intersection dénombrable d'ouverts denses n'est pas vide (puisqu'elle est dense !). En s'inspirant de la démonstration que nous venons de faire, on obtient aisément le lemme suivant, qui nous sera utile dans la suite

Lemme 1.3.1 *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 choisit dans chaque F_n un élément que l'on appelle x_n. Comme la suite $(F_n)_{n\in\mathbb{N}}$ est décroissante, pour tout entier p, on sait que x_{n+p} appartient à F_n, donc que $d(x_n, x_{n+p}) \leq \delta(F_n)$. Donc la suite $(x_n)_{n\in\mathbb{N}}$ est de Cauchy et converge vers un élément x de X qui appartient bien sûr à l'intersection des F_n. Soit y un élément de cette intersection. On a alors

$$\forall n\,,\; d(x,y) \leq \delta(F_n)$$

et donc $d(x,y) = 0$, d'où la proposition. □

Remarquons que l'on utilise souvent l'énoncé suivant qui n'est rien d'autre que le théorème de Baire « passé au complémentaire ».

Théorème 1.3.4 *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 $\displaystyle\bigcup_{n\in\mathbb{N}} F_n$ est d'intérieur vide.*

Nous utiliserons souvent le théorème de Baire sous la forme de ce corollaire, dont la démonstration, très facile, est laissée en exercice.

Corollaire 1.3.1 *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* $\mathrm{Int}(F_{n_0}) \neq \emptyset$.

Autrement dit, 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 $\mathbb{Q}$.

A ce propos, justement qu'en est-il des espaces non complets ? Est-il possible de pallier leur non complétude ? La réponse est oui, au sens du théorème suivant, que nous ne démontrerons pas.

Théorème 1.3.5 *Soit (X,d) un espace métrique. On peut construire un espace métrique complet $(\hat{X},\hat{d})$, qui contient X comme sous-espace dense et dont la distance est un prolongement de la distance sur X. On appelle $(\hat{X},\hat{d})$* un complété *de (X,d). Il est unique au sens suivant : si $(\hat{X}_1,\hat{d}_1)$ et $(\hat{X}_2,\hat{d}_2)$ sont deux complétés de (X,d), alors il existe une unique bijection entre $\hat{X}_1$ et $\hat{X}_2$ conservant les distances* (*une* isométrie) *et égale à l'identité sur X.*

Commentaires. Naturellement, si (X,d) est lui-même déjà complet, on peut prendre $(\hat{X},\hat{d}) = (X,d)$. La construction du théorème est une construction abstraite : on considère l'ensemble des suites de Cauchy de X que l'on quotiente par une relation d'équivalence permettant d'identifier deux suites de Cauchy qui « aimeraient bien converger vers la même limite », et que l'on munit d'une distance qui prolonge celle de X (identifié aux classes d'équivalence des suites constantes).

On rencontrera dans la pratique des situations dans lesquelles on aura sous la main d'autres réalisations du complété, par des espaces concrets. C'est souvent le cas pour des espaces de fonctions dont le complété s'identifiera isométriquement à un autre espace de fonctions plus grand. Dans l'exemple ci-dessus, c'est aussi du concret car on a $\hat{\mathbb{Q}} = \mathbb{R}$.

1.4 Espaces compacts

La notion de compacité est absolument essentielle en analyse. Il s'agit d'une notion topologique, indépendante de la notion de distance. Les compacts métriques ont en fait des propriétés particulières que ne possèdent pas nécessairement les compacts en général, comme on le verra dans la suite.

Définition 1.4.1 *Soit $(X,\mathscr{O})$ un espace topologique. On dit que X est* compact *si et seulement si de tout recouvrement de X par des ouverts, on peut extraire un sous-recouvrement fini. Autrement dit, pour toute famille $(U_\lambda)_{\lambda\in\Lambda}$ d'ouverts de X telle que*

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

il existe une famille finie $(\lambda_j)_{1\leq j\leq N}$ [2] *telle que*

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

On dit qu'une partie $A \subset X$ est une partie compacte de X, lorsqu'elle est compacte pour la topologie induite.

On ajoute parfois l'hypothèse de séparation à la définition de la compacité, un compact au sens précédent sans hypothèse de séparation étant alors qualifié de quasi-compact.

La compacité d'une partie A s'exprime aussi de la façon suivante.

Proposition 1.4.1 *Soient $(X,\mathscr{O})$ un espace topologique et A une partie de X. La partie A est un compact de X si et seulement si pour toute famille $(U_\lambda)_{\lambda\in\Lambda}$ d'ouverts de X telle que $A \subset \bigcup_{\lambda\in\Lambda} U_\lambda$, il existe une famille finie $(\lambda_j)_{1\leq j\leq N}$ telle que $A \subset \bigcup_{j=1}^{N} U_{\lambda_j}$.*

La démonstration est un exercice très formateur sur la notion de sous-espace topologique ; le lecteur est fortement incité à le faire.

La proposition suivante est une définition équivalente de la compacité qui est souvent utilisée.

Proposition 1.4.2 *Un espace topologique $(X,\mathscr{O})$ est compact si et seulement si pour toute famille de fermés $(F_\lambda)_{\lambda\in\Lambda}$ telle que*

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

alors on a

$$\bigcap_{\lambda\in\Lambda} F_\lambda \neq \emptyset.$$

Démonstration. On passe au complémentaire. Soit $(F_\lambda)_{\lambda\in\Lambda}$ une famille de fermés de X, posons $U_\lambda = X\setminus F_\lambda$, qui est un ouvert. On a alors

$$\bigcap_{\lambda\in\Lambda} F_\lambda = \emptyset \Longleftrightarrow \bigcup_{\lambda\in\Lambda} U_\lambda = X.$$

De même, on peut écrire que, pour toute famille finie $(\lambda_1,\cdots,\lambda_N)\in\Lambda^N$,

$$\bigcap_{j=1}^{N} F_{\lambda_j} = \emptyset \Longleftrightarrow \bigcup_{j=1}^{N} U_{\lambda_j} = X.$$

On conclut par contraposition. □

2. On laisse également ici de côté le cas de $X=\emptyset$ et d'une famille à 0 éléments. L'ensemble vide est naturellement compact pour son unique topologie.

La propriété en hypothèse dans la proposition 1.4.2 est appelée *propriété d'intersection finie.*

Proposition 1.4.3 *Soit $(X, \mathscr{O})$ un espace topologique compact. Toute partie fermée de X est compacte.*

Démonstration. Soit A une partie fermée de X. On considère un recouvrement ouvert $(U_\lambda)_{\lambda\in\Lambda}$ de A. La famille formée des U_λ à laquelle on adjoint l'ouvert $X \setminus A$ forme un recouvrement ouvert de X. Comme X est compact, on en extrait un sous-recouvrement fini, auquel on peut toujours adjoindre $X \setminus A$, tel que

$$(X \setminus A) \cup \bigcup_{j=1}^{N} U_{\lambda_j} = X.$$

Il est alors clair que

$$A \subset \bigcup_{j=1}^{N} U_{\lambda_j},$$

d'où la proposition. □

Proposition 1.4.4 *Soit $(X, \mathscr{O})$ un espace topologique séparé. Si A est une partie compacte de X, alors A est un fermé de X.*

Démonstration. Soit A un compact de X. Considérons un point x du complémentaire de A et $\mathscr{U}$ l'ensemble des ouverts U tel qu'il existe un ouvert V contenant x et tel que $U \cap V = \emptyset$. Comme X est séparé, tout point y de A appartient à un élément de $\mathscr{U}$, qui constitue donc un recouvrement ouvert de A. On peut en extraire un sous-recouvrement fini $(U_j)_{1\leq j\leq N}$. Pour chacun de ces indices, on a donc un ouvert V_j contenant x et d'intersection vide avec U_j. Posons

$$V = \bigcap_{j=1}^{N} V_j.$$

C'est un ouvert comme intersection finie d'ouverts. Il contient le point x par construction. De plus, il est clair que l'on a

$$V \cap A \subset V \cap \bigcup_{j=1}^{N} U_j = \emptyset.$$

Donc $X \setminus A$ est ouvert ou encore A est fermé. □

Commentaire. La séparation de l'espace est ici cruciale. En fait, il est facile de construire un espace topologique non séparé contenant une partie compacte mais non fermée.

Exercice 1.4.1 *On considère l'espace topologique* $]0,1[$ *muni de la topologie métrique usuelle. Démontrez que cet espace topologique n'est pas compact.*

Exercice 1.4.2 *Soit* $[a,b]$ *muni de la distance* $d(x,y) = |x-y|$*. Démontrez que cet espace topologique est compact. Pour cela, considérez un recouvrement* $(U_\lambda)_{\lambda\in\Lambda}$ *de* $[a,b]$ *et étudiez la borne supérieure de l'ensemble des* x *tels que l'intervalle* $[a,x]$ *puisse être recouvert par un nombre fini d'éléments* U_λ*.*

Exercice 1.4.3 *On considère l'espace topologique* $([0,1],\mathscr{O}_d)$ *(voir l'exercice 1.1.1 pour la définition). Démontrez que cet espace topologique n'est pas compact.*

Le théorème suivant est extrêmement simple mais vraiment fondamental.

Théorème 1.4.1 *Soient* $(X,\mathscr{O})$ *et* $(Y,\widetilde{\mathscr{O}})$ *deux espaces topologiques et* f *une fonction continue de* X *dans* Y*. Si* A *est une partie compacte de* X*, alors* $f(A)$ *est une partie compacte de* Y*.*

Démonstration. Pour démontrer cette proposition, considérons un recouvrement $(V_\lambda)_{\lambda\in\Lambda}$ de $f(A)$ par des ouverts. La famille $(f^{-1}(V))_{\lambda\in\Lambda}$ est un recouvrement de A par des ouverts puisque f est continue. Extrayons-en un sous-recouvrement fini $(f^{-1}(V_{\lambda_j}))_{1\leq j\leq N}$. Mais alors, la famille $(f(f^{-1}(V_{\lambda_j})))_{1\leq j\leq N}$ recouvre $f(A)$. On conclut alors en observant que $f(f^{-1}(V_{\lambda_j}))\subset V_{\lambda_j}$. □

Attention, l'image réciproque d'un compact par une application continue n'est en général pas compacte (trouvez un exemple avec une fonction simple de $\mathbb{R}$ dans $\mathbb{R}$!). Il ne faut donc pas confondre cette propriété avec le fait que l'image réciproque d'un ouvert (ou d'un fermé) est un ouvert (ou un fermé).

Donnons une première propriété particulière aux compacts métriques.

Proposition 1.4.5 *Toute partie compacte d'un espace métrique est bornée.*

Démonstration. Soit (X,d) un espace métrique, $A\subset X$ un compact de X et $x_0\in X$. Comme $X=\bigcup_{n=1}^{+\infty} B(x_0,n)$, on a aussi $A\subset\bigcup_{n=1}^{+\infty} B(x_0,n)$. On peut en extraire un sous recouvrement fini et si N est le plus grand indice de ce recouvrement, il vient $A\subset B(x_0,N)$. □

Du théorème 1.4.1 et des propositions 1.4.4 et 1.4.5, on déduit très aisément le

Corollaire 1.4.1 *Soient* $(X, \mathscr{O})$ *un espace topologique compact et* f *une fonction continue de* X *dans* $(\mathbb{R}, d)$ *muni de sa distance usuelle. Alors* $f(X)$ *est un ensemble borné de* $\mathbb{R}$ *et la fonction* f *atteint ses bornes, c'est-à-dire qu'il existe deux éléments* x *et* y *de* X *tels que*

$$\forall z \in X, f(x) = \inf_X f \leq f(z) \leq \sup_X f = f(y).$$

Démonstration. En effet, $f(X)$ est compact donc fermé et borné (pour la distance usuelle $d(x,y) = |x-y|$). Comme il est borné, il admet une borne inférieure et une borne supérieure finies. Ces bornes sont des valeurs d'adhérence, elles appartiennent donc à $f(X)$ puisque celui-ci est fermé. □

En particulier, $\inf_X f > -\infty$ et $\sup_X f < +\infty$. Si les fonctions continues définies sur un compact ont, comme nous venons de le voir, des propriétés remarquables, il en va de même des suites à valeurs dans un espace compact.

Théorème 1.4.2 *Soient* $(X, \mathscr{O})$ *un espace topologique compact et* $(x_n)_{n\in\mathbb{N}}$ *une suite d'éléments de* X*. Alors l'ensemble des valeurs d'adhérence de* $(x_n)_{n\in\mathbb{N}}$ *est non vide. De plus, si cet ensemble est réduit à un seul élément* ℓ*, alors on a* $\lim_{n\to\infty} x_n = \ell$

Démonstration. Posons

$$F_n = \mathrm{Adh}\{x_m\,,\ m \geq n\}.$$

La suite $(F_n)_{n\in\mathbb{N}}$ est une suite décroissante de fermés non vides. Toute intersection finie d'éléments de la suite $(F_n)_{n\in\mathbb{N}}$ est par conséquent un ensemble F_p, où p est l'indice maximum apparaissant dans la famille finie en question. Cette intersection est donc non vide. Comme X est compact, on en déduit que

$$\bigcap_{n\in\mathbb{N}} F_n \neq \emptyset,$$

par la proposition 1.4.2. Or la proposition 1.1.7 affirme que l'ensemble des valeurs d'adhérence est justement l'intersection de tous les ensembles F_n, d'où le premier point de ce théorème.

Pour démontrer le second point, considérons un ouvert quelconque U contenant ℓ et posons

$$\widetilde{F}_n = \mathrm{Adh}\{x_m\,, m \geq n\} \cap (X \setminus U).$$

Le fait que ℓ soit la seule valeur d'adhérence implique que l'intersection des $\widetilde{F}_n$ est vide. Donc, comme la suite $(\widetilde{F}_n)_{n\in\mathbb{N}}$ est une suite décroissante, il existe un entier n_0 tel que $\widetilde{F}_{n_0}$ soit vide, encore par la proposition 1.4.2 ou plutôt sa contraposée. Cela signifie en particulier que

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

ce qui termine la démonstration de ce théorème. □

Dans le cadre des espaces métriques, *et seulement dans ce cadre, au moins dans ces notes*, cette propriété est caractéristique des compacts. C'est la propriété de Bolzano-Weierstrass, un point fondamental.

Théorème 1.4.3 *Soit (X,d) un espace métrique. Il est compact si et seulement si toute suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X possède une valeur d'adhérence.*

Il s'agit de déduire d'une propriété sur les suites, une propriété sur les recouvrements d'ouverts. Une première étape consiste à le faire dans le cas des boules.

Lemme 1.4.1 *Soit (X,d) un espace métrique dans lequel toute suite $(x_n)_{n\in\mathbb{N}}$ possède une valeur d'adhérence. Alors, pour tout réel strictement positif α, on peut recouvrir X par un nombre fini de boules de rayon α.*

Démonstration. Démontrons ce lemme par contraposition. Supposons qu'il existe 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 $B(x_0,\alpha)$, sinon X serait recouvert par une boule de rayon α. Supposons construit $(x_0,\cdots,x_p)$ un $(p+1)$-uplet d'éléments de X tel que, pour tout $m\neq n$, on ait $d(x_m,x_n)\geq\alpha$. Par hypothèse,

$$\bigcup_{n=0}^{p} B(x_n,\alpha)\neq X.$$

Donc, il existe un point x_{p+1} de X qui n'appartient pas à la réunion de boules ci-dessus ce qui implique que la distance de x_{p+1} à chacun des points antérieurement construits est supérieure à α. 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 (exercice, vérifiez-le !). □

Le lemme de Lebesgue suivant, crucial pour la présente démonstration, sera utile dans d'autres contextes.

Lemme 1.4.2 *Soit (X,d) un espace métrique et F une partie fermée de X. On suppose que toute suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de F possède une valeur d'adhérence. Alors, pour toute famille $(U_\lambda)_{\lambda\in\Lambda}$ d'ouverts de X recouvrant F, il existe un nombre α strictement positif tel que*

$$\forall x\in F,\exists\,\lambda\in\Lambda \text{ tel que } B(x,\alpha)\subset U_\lambda.$$

Démonstration. On va procéder par l'absurde en supposant que cette propriété est fausse, c'est-à-dire que l'assertion suivante est vraie

$$\forall\,\alpha\in\mathbb{R}_+^*,\exists\,x\in F,\forall\lambda\in\Lambda,\quad B(x,\alpha)\not\subset U_\lambda.$$

On choisit $\alpha = 1/n$ pour $n \in \mathbb{N}^*$. Ceci nous donne une suite $x_n \in F$ telle que

$$\forall n \in \mathbb{N}^*, \forall \lambda \in \Lambda, \quad B(x_n, 1/n) \not\subset U_\lambda.$$

Par hypothèse, cette suite admet une valeur d'adhérence, ce qui implique qu'il existe une sous-suite notée x_{n_p} et un $x \in F$ tels que $d(x_{n_p}, x) \to 0$ quand $p \to +\infty$.

Comme les U_λ recouvrent F, il existe $\mu \in \Lambda$ tel que $x \in U_\mu$. Soit $\beta > 0$ tel que $B(x,\beta) \subset U_\mu$. Choisissons alors p assez grand pour que $d(x_{n_p}, x) < \beta/2$ et $1/n_p < \beta/2$. Pour tout $y \in B(x_{n_p}, 1/n_p)$, on en déduit que $d(y,x) \leq d(y, x_{n_p}) + d(x_{n_p}, x) < \beta$, c'est-à-dire

$$B(x_{n_p}, 1/n_p) \subset B(x,\beta) \subset U_\mu,$$

ce qui contredit la définition de x_n. □

Démonstration du théorème 1.4.3. Soit (X,d) un espace métrique tel que toute suite admette une valeur d'adhérence et soit $(U_\lambda)_{\lambda \in \Lambda}$ un recouvrement ouvert de X. Soit $\alpha > 0$ le nombre associé à ce recouvrement fourni par le lemme 1.4.2 appliqué à $F = X$. D'après le lemme 1.4.1, on peut recouvrir X par un nombre fini de boules ouvertes de rayon α, $X = \cup_{i=1}^p B(x_i, \alpha)$. Le lemme 1.4.2 nous donne pour chaque i un indice $\lambda_i \in \Lambda$ tel que $B(x_i, \alpha) \subset U_{\lambda_i}$. On en déduit immédiatement que $X = \cup_{i=1}^p U_{\lambda_i}$. □

Ce théorème est très important. Il est fréquent qu'il soit plus facile de vérifier cette propriété que la propriété sur les ouverts (ou celle sur les fermés). En particulier, on en déduit le théorème fondamental suivant.

Théorème 1.4.4 *Les compacts de l'espace* $(\mathbb{R}, \mathscr{O}_0)$ *sont les parties fermées et bornées pour la distance usuelle.*

Démonstration. Le seul point à démontrer est que toute suite bornée $(x_n)_{n\in\mathbb{N}}$ admet une valeur d'adhérence. Toute famille majorée de réels admet une borne supérieure. Posons donc

$$M_n = \sup_{m \geq n} x_m.$$

La suite $(M_n)_{n\in\mathbb{N}}$ est décroissante et minorée 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, on l'appelle la limite supérieure de la suite). □

ATTENTION ! Cette propriété est tout à fait particulière à $\mathbb{R}$ (ou à $\mathbb{R}^d$ comme nous le verrons). Si les compacts d'un espace métrique sont toujours fermés et bornés, il est presque toujours faux que les fermés bornés d'un espace métrique soient compacts. Les deux exemples suivants sont révélateurs.

Considérons l'espace $]0,1[$ muni de la distance usuelle. C'est un ensemble borné. L'intervalle $]0,1[$ est fermé dans lui-même, comme tout espace topologique. Nous avons vu lors de l'exercice 1.4.1 ci-dessus que cet espace n'est pas compact.

Considérons maintenant l'espace $\mathbb{R}$ muni de la distance définie par

$$\widetilde{d}(x,y) = \frac{|x-y|}{1+|x-y|}.$$

Cette distance définit sur $\mathbb{R}$ la même topologie que la distance $d(x,y) = |x-y|$ (exercice : vérifiez-le !). L'ensemble $\mathbb{R}$ est borné pour la distance $\widetilde{d}$. Il est fermé dans lui-même. Le lecteur se convaincra (et démontrera) que l'espace $(\mathbb{R},d)$ n'est pas compact, donc l'espace $(\mathbb{R},\widetilde{d})$ non plus.

Conclusion : Si dans un espace métrique on a toujours

$$\text{compact} \Longrightarrow \text{fermé et borné},$$

il faut bien réaliser que sauf dans le cas de $\mathbb{R}^d$, muni d'une de ses distances usuelles, ainsi qu'un certain nombre d'autres cas particuliers,

$$\textbf{fermé et borné} \not\Longrightarrow \textbf{compact !!!}$$

Corollaire 1.4.2 *Soit A une partie d'un espace métrique (X,d). Le fermé $\bar{A}$ est compact si et seulement si toute suite d'éléments de A possède une valeur d'adhérence dans X.*

La démonstration de ce corollaire est laissée en exercice.

Théorème 1.4.5 *Soient $(X_1,d_1),\cdots,(X_N,d_N)$ une famille finie d'espaces métriques compacts. Le produit $X_1\times\cdots\times X_N$ muni de l'une quelconque des distances définies dans l'énoncé de la proposition* 1.3.5 *est métrique compact.*

La preuve de ce théorème est un exercice laissé au lecteur qui pourra utiliser le critère des suites puis ensuite celui des ouverts et comparer ainsi les deux méthodes. Notons qu'il s'agit d'un cas particulier d'un théorème infiniment plus général, le théorème de Tychonoff, qui affirme qu'un produit quelconque d'espaces topologiques compacts est compact, pour la topologie dite produit que nous n'avons pas introduite précédement.

On déduit également du théorème 1.4.5 l'important résultat suivant, déjà mentionné plusieurs fois.

Corollaire 1.4.3 *Les compacts de $\mathbb{R}^d$ sont les fermés bornés (pour une des distances usuelles) de $\mathbb{R}^d$.*

En effet, on peut toujours inclure un tel borné de $\mathbb{R}^d$ dans un produit cartésien d'intervalles compacts.

Les relations entre le concept d'espaces complets et d'espaces compacts sont décrites par les deux énoncés suivants.

Proposition 1.4.6 *Soit (X,d) un espace métrique compact, il est alors complet.*

Démonstration. Soit $(x_n)_{n\in\mathbb{N}}$ une suite de Cauchy de X. Comme l'espace X est compact, cette suite admet une valeur d'adhérence et la proposition 1.3.3 nous dit qu'une suite de Cauchy admettant une valeur d'adhérence converge, d'où la proposition. $\square$

La réciproque de cette proposition est bien sûr fausse. Par exemple, $\mathbb{R}$ muni de la distance $d(x,y) = |x-y|$ est un espace métrique complet qui n'est pas compact. Cependant, dans le cas des espaces métriques complets, on peut trouver une caractérisation des ensembles compacts uniquement en termes de recouvrements finis par des boules de même rayon, mais arbitrairement petit.

Théorème 1.4.6 *Soit (X,d) un espace métrique complet, on considère une partie A de X. L'adhérence de A est compacte si et seulement si*

$$\forall \varepsilon > 0, \exists\, x_j \in A, 1 \leq j \leq N \text{ tels que } A \subset \bigcup_{j=1}^{N} B(x_j,\varepsilon). \tag{1.8}$$

Démonstration. La condition nécessaire est facile, puisque $A \subset \bar{A} \subset \bigcup_{x\in A} B(x,\varepsilon)$.

Pour la condition suffisante, considérons une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de A. Recouvrons A par un nombre fini de boules ouvertes de rayon $1/2$. Au moins une de ces boules doit contenir une infinité d'éléments de la suite $(x_n)_{n\in\mathbb{N}}$. Appelons l'adhérence de cette boule F_0. Recouvrons alors A par un nombre fini de boules de rayon $1/4$. Clairement, au moins une de ces boules doit avoir une intersection non vide avec F_0 et cette intersection doit contenir une infinité d'éléments de la suite $(x_n)_{n\in\mathbb{N}}$, sinon F_0 n'en contiendrait qu'un nombre fini. Appelons F_1 l'adhérence de l'intersection de cette deuxième boule avec F_0. Clairement, $F_1 \subset F_0$, le diamètre de F_1 est inférieur à $2 \times 1/4 = 1/2$ et F_1 contient une infinité d'éléments de la suite $(x_n)_{n\in\mathbb{N}}$.

On procède maintenant par récurrence. Soit $k \geq 1$ et supposons que, pour tout $0 \leq p \leq k$, on ait construit un fermé F_p de diamètre inférieur à 2^{-p}, tel que $F_p \subset F_{p-1}$ et qui contient une infinité d'éléments de la suite $(x_n)_{n\in\mathbb{N}}$. Recouvrons A par un nombre fini de boules de rayon $2^{-(k+2)}$. Comme précédemment, au moins une de ces boules a une intersection non vide avec F_k, intersection qui contient aussi une infinité d'éléments de la suite $(x_n)_{n\in\mathbb{N}}$, sinon F_k n'en contiendrait qu'un nombre fini. On note F_{k+1} l'adhérence de cette intersection. Par construction, $F_{k+1} \subset F_k$ et le diamètre de F_{k+1} est inférieur à $2 \times 2^{-(k+2)} = 2^{-(k+1)}$ et F_{k+1} contient une infinité d'éléments de la suite (x_n).

On a ainsi construit une suite décroissante de fermés $(F_k)_{k\in\mathbb{N}}$, dont le diamètre tend vers 0 et telle que chaque fermé F_k contienne une infinité d'éléments de la suite $(x_n)_{n\in\mathbb{N}}$. Comme l'espace métrique (X,d) est complet, d'après le lemme 1.3.1, il existe un élément x de X tel que x appartienne à l'intersection des F_k.

Démontrons que x est valeur d'adhérence de la suite $(x_n)_{n\in\mathbb{N}}$. C'est clair car, étant donné $\varepsilon > 0$, on choisit k tel que $\delta(F_k) < \varepsilon$. Par construction des F_k, il existe une infinité d'éléments de $(x_n)_{n\in\mathbb{N}}$ dont la distance à x est inférieure à ε. La suite $(x_n)_{n\in\mathbb{N}}$ admet donc bien x comme une valeur d'adhérence. C'est un exercice facile et laissé au lecteur que de démontrer que $\bar{A}$ est compact si et seulement si toute suite d'éléments de A admet une valeur d'adhérence, ce qui démontre le théorème. □

Remarques. Une partie A dont l'adhérence est compacte est dite *relativement compacte*. Une partie satisfaisant l'hypothèse (1.8) est dite *précompacte*. L'hypothèse de complétude de l'espace métrique (X,d) est essentielle pour le théorème 1.4.6. Considérons l'intervalle $]0,1[$ dans $\mathbb{Q}$. Il est clair qu'il est précompact, mais que l'intervalle $[0,1]$ dans $\mathbb{Q}$ n'est pas compact.

Exercice 1.4.4 *Soit (X,d) un espace métrique compact. On considère l'ensemble $X^{\mathbb{N}}$ des suites d'éléments de X.*

1) On définit l'application $d_{\mathbb{N}}$ de $X^{\mathbb{N}} \times X^{\mathbb{N}}$ dans $\mathbb{R}_+$ par

$$d_{\mathbb{N}}((x_n)_{n\in\mathbb{N}},(y_n)_{n\in\mathbb{N}}) = \sum_{n=0}^{\infty} \frac{1}{2^n} \min\{d(x_n,y_n),1\}.$$

Démontrer que $d_{\mathbb{N}}$ est une distance sur $X^{\mathbb{N}}$.

2) Soit $(x^{(p)})_{p\in\mathbb{N}}$ une suite d'éléments de $X^{\mathbb{N}}$ et x un élément de $X^{\mathbb{N}}$. Démontrez que

$$\lim_{p\to\infty} x^{(p)} = x \quad \textit{si et seulement si} \quad \forall n \in \mathbb{N},\ \lim_{p\to\infty} x_n^{(p)} = x_n.$$

3) Démontrez que $(X^{\mathbb{N}},d_{\mathbb{N}})$ est compact.

4) 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

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

Le résultat suivant est très important en analyse.

Théorème 1.4.7 *Soient (X,d) et (Y,δ) deux espaces métriques. Supposons que X soit compact. Alors, toute application continue de X dans Y est uniformément continue.*

Démonstration. Pour tout $x \in X$, comme f est continue au point x, pour tout $\varepsilon > 0$, il existe $\alpha_x > 0$ telle que l'on ait

$$d(x,x') < \alpha_x \Rightarrow \delta(f(x), f(x')) < \frac{\varepsilon}{2}.$$

L'ensemble des boules $B(x, \alpha_x/2)$ recouvre X. On en extrait un sous-recouvrement fini de la forme $(B(x_j, \alpha_{x_j}/2))_{1 \leq j \leq N}$. On pose alors

$$\alpha = \inf_{1 \leq j \leq N} \alpha_{x_j} > 0.$$

Soit $(x,x') \in X^2$ tel que $d(x,x') < \alpha/2$. Comme la famille $(B(x_j, \alpha_{x_j}/2))_{1 \leq j \leq N}$ recouvre X, il existe un indice $j \in \{1, \cdots, N\}$ tel que $B(x, \alpha/2) \subset B(x_j, \alpha_{x_j})$. Mais alors, le point x' appartient à $B(x_j, \alpha_{x_j})$ et l'on a ainsi

$$\begin{aligned} \delta(f(x), f(x')) &\leq \delta(f(x), f(x_j)) + \delta(f(x_j), f(x')) \\ &< \varepsilon, \end{aligned}$$

d'où le théorème. □

Exercice 1.4.5 *Démontrez ce théorème en utilisant le lemme* 1.4.2 *de Lebesgue.*

Voici un autre théorème classique important, le théorème d'Ascoli, qui permet de caractériser les parties compactes de l'espace des fonctions continues sur un compact.

Théorème 1.4.8 *Soient (X,d) et (Y,δ) deux espaces métriques. Supposons que (X,d) soit compact et (Y,δ) complet. On considère une partie A de $\mathscr{C}(X,Y)$, l'ensemble des fonctions continues de X dans Y muni de la distance*

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

Faisons les deux hypothèses suivantes :

i) La partie A est équicontinue*, i.e.*

$$\forall x \in X, \forall \varepsilon > 0, \exists\, \alpha > 0 \text{ tel que } \forall f \in A, \quad d(x,x') < \alpha \Rightarrow \delta(f(x), f(x')) < \varepsilon,$$

ii) pour tout $x \in X$, l'ensemble $\{f(x), f \in A\}$ est relativement compact dans Y.

Alors A est relativement compacte dans $(\mathscr{C}(X,Y), D)$.

Pour démontrer ce théorème, nous allons tout d'abord observer que la première hypothèse se traduit par l'équicontinuité uniforme

$$\forall \varepsilon > 0, \exists\, \alpha > 0 \text{ tel que } \forall (x,x') \in X^2, \forall f \in A, d(x,x') < \alpha \Rightarrow \delta(f(x), f(x')) < \varepsilon. \tag{1.9}$$

La démonstration de cette assertion est exactement la même que celle du théorème précédent (exercice : écrivez-la !). En d'autres termes, sur un compact, toute partie équicontinue est uniformément équicontinue. Remarquons qu'une famille finie est automatiquement équicontinue et qu'une réunion de deux familles équicontinues est équicontinue.

Nous allons maintenant énoncer un lemme montrant que les parties A qui sont uniformément équicontinues ont une propriété très spéciale vis-à-vis de la convergence.

Lemme 1.4.3 *Soient A une partie uniformément équicontinue de $\mathscr{C}(X,Y)$, $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de A, $f \in \mathscr{C}(X,Y)$ et Δ une partie dense de X. On a l'équivalence suivante*

$$\forall x \in \Delta,\ f_n(x) \underset{n\to\infty}{\longrightarrow} f(x) \Leftrightarrow D(f_n, f) \underset{n\to\infty}{\longrightarrow} 0.$$

En d'autres termes, sur une partie uniformément équicontinue, la convergence simple d'une suite de fonctions sur une partie dense de X est équivalente à la convergence uniforme sur X !

Démonstration. Il n'y a quelque chose à démontrer que de gauche à droite. Soit ε un réel strictement positif arbitraire. Par hypothèse, il existe un réel α strictement positif tel que

$$d(x,x') < \alpha \Rightarrow \forall n \in \mathbb{N},\ \delta(f_n(x), f_n(x')) < \frac{\varepsilon}{3} \text{ et } \delta(f(x), f(x')) < \frac{\varepsilon}{3},$$

(la partie $A \cup \{f\}$ est également uniformément équicontinue).

Les boules centrées sur les points de Δ et de rayon α recouvrent X par densité de Δ. On en extrait un recouvrement fini de boules $(B(x_j, \alpha))_{1\leq j\leq N_\varepsilon}$ avec $x_j \in \Delta$. Ainsi, par la propriété de recouvrement,

$$\forall x \in X, \exists\, k \leq N_\varepsilon, d(x, x_k) \leq \alpha,$$

d'où

$$\begin{aligned} \forall x \in X,\ \delta(f_n(x), f(x)) &\leq \delta(f_n(x), f_n(x_k)) + \delta(f_n(x_k), f(x_k)) + \delta(f(x_k), f(x)) \\ &\leq \frac{2\varepsilon}{3} + \max_{1\leq j\leq N_\varepsilon} \delta(f_n(x_j), f(x_j)). \end{aligned}$$

Par hypothèse, et comme on n'a affaire qu'à un nombre fini de points, il existe n_0 tel que

$$\forall n \geq n_0\,,\ \max_{1\leq j\leq N_\varepsilon} \delta(f_n(x_j), f(x_j)) < \frac{\varepsilon}{3},$$

ce qui montre que f_n converge uniformément vers f, d'où le lemme. □

Démonstration du théorème d'Ascoli. Considérons une suite $(f_n)_{n\in\mathbb{N}}$ d'éléments de A. Nous allons en extraire une sous-suite convergente.

Nous démontrerons très bientôt que dans tout espace métrique compact, il existe une partie dénombrable dense (voir théorème 1.5.1) que l'on note $\Delta = (x_p)_{p\in\mathbb{N}}$. Par hypothèse, l'ensemble

$$\{f(x_0)\,,\ f\in A\}$$

est d'adhérence compacte. Donc il existe un point de Y que nous notons $\ell(x_0)$ et une application φ_0 strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ telle que

$$\lim_{n\to\infty} f_{\varphi_0(n)}(x_0) = \ell(x_0).$$

De même, la suite $f_{\varphi_0(n)}(x_1)$ reste dans un ensemble relativement compact, donc il existe un point $\ell(x_1)$ de Y et une application φ_1 strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ telle que

$$\lim_{n\to\infty} f_{\varphi_0\circ\varphi_1(n)}(x_1) = \ell(x_1).$$

Par construction, on a toujours

$$\lim_{n\to\infty} f_{\varphi_0\circ\varphi_1(n)}(x_0) = \ell(x_0),$$

puisqu'il s'agit d'une suite extraite de la précédente (on sélectionne à l'aide de φ_1 des indices qui appartiennent déjà à la suite $\varphi_0(n)$).

On définit ainsi par récurrence une suite $(\varphi_p)_{p\in\mathbb{N}}$ d'applications strictement croissantes de $\mathbb{N}$ dans $\mathbb{N}$ et des points $\ell(x_k)$ de Y, c'est-à-dire en fin de compte une application ℓ de Δ dans Y, tels que

$$\forall\, p\in\mathbb{N},\ \forall\, k\le p,\ \lim_{n\to\infty} f_{\varphi_0\circ\cdots\circ\varphi_p(n)}(x_k) = \ell(x_k).$$

On définit alors une application ψ de $\mathbb{N}$ dans $\mathbb{N}$ par

$$\psi(n) = \varphi_0\circ\cdots\circ\varphi_n(n).$$

Cette application est strictement croissante. En effet, toute application μ strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ vérifie $\mu(n)\ge n$ (immédiat par récurrence). Donc, pour tout n, on a

$$\begin{aligned}\psi(n+1) &= \varphi_0(\cdots(\varphi_n(\varphi_{n+1}(n+1)))\cdots) \ge \\ &\ge \varphi_0(\cdots(\varphi_n(n+1))\cdots) > \varphi_0(\cdots(\varphi_n(n))\cdots) = \psi(n),\end{aligned}$$

car $\varphi_n(n+1) > \varphi_n(n)$ et chacune des applications composées est strictement croissante.

Par construction, la suite extraite $(f_{\psi(n)})_{n\in\mathbb{N}}$ vérifie

$$\forall p \in \mathbb{N}, \lim_{n\to\infty} f_{\psi(n)}(x_p) = \ell(x_p). \tag{1.10}$$

En effet, dès que $n \geq p$, c'est une suite extraite de la suite $f_{\varphi_0\circ\cdots\circ\varphi_p}(x_p)$, laquelle converge bien vers $\ell(x_p)$.

Démontrons maintenant que l'application ℓ est uniformément continue sur l'ensemble Δ. Soit ε un réel strictement positif, on considère un réel α vérifiant l'assertion (1.9). Pour tout couple d'entiers (p,q) tel que $d(x_p,x_q) < \alpha$, on a

$$\begin{aligned}\delta(\ell(x_p),\ell(x_q)) &\leq \delta(\ell(x_p),f_{\psi(n)}(x_p)) + \delta(f_{\psi(n)}(x_p),f_{\psi(n)}(x_q))\\ &\qquad\qquad\qquad\qquad + \delta(f_{\psi(n)}(x_q),\ell(x_q))\\ &\leq \delta(\ell(x_p),f_{\psi(n)}(x_p)) + \varepsilon + \delta(f_{\psi(n)}(x_q),\ell(x_q)).\end{aligned}$$

L'inégalité ci-dessus étant vraie pour tout n, on obtient, en passant à la limite quand $n\to\infty$,

$$d(x_p,x_q) < \alpha \Rightarrow d(\ell(x_p),\ell(x_q)) \leq \varepsilon.$$

L'application ℓ étant uniformément continue sur la partie dense Δ et l'espace Y étant complet, on peut appliquer le théorème 1.3.1 qui affirme que l'on peut prolonger ℓ en une application uniformément continue sur X tout entier. On utilise alors le lemme 1.4.3 pour conclure la démonstration du théorème d'Ascoli. □

Commentaire. Le procédé d'extraction de sous-suite qui aboutit à la propriété 1.10 est un procédé courant en analyse, appelé procédé diagonal. La suite finale extraite s'appelle suite diagonale car si l'on range les valeurs $f_n(x_k)$ dans un tableau indexé par $\mathbb{N}\times\mathbb{N}$, on voit que les indices de la suite extraite s'en vont chacun vers l'infini diagonalement dans le tableau. Il s'agit d'un exemple typique d'utilisation de la dénombrabilité en analyse.

Exercice 1.4.6 *Démontrez la réciproque du théorème d'Ascoli, à savoir que si une partie A de $\mathscr{C}(X,Y)$ est relativement compacte, alors elle vérifie les conditions i) et ii) de l'énoncé du théorème d'Ascoli.*

Exercice 1.4.7 *Soient (X,d) et (Y,δ) deux espaces métriques compacts. Démontrez que l'ensemble des fonctions k-lipschitziennes de (X,d) dans (Y,δ) est compact dans $\mathscr{C}(X,Y)$.*

Mentionnons un dernier théorème classique très important, le théorème de Stone-Weierstrass. On dira qu'un ensemble A de fonctions sur X *sépare les points* si pour tout couple (x,y) d'éléments de X tel que $x \neq y$, il existe une fonction f de A telle que $f(x) \neq f(y)$.

Théorème 1.4.9 *Soient $(X,\mathscr{O})$ un espace compact et A une sous-algèbre de $C(X,\mathbb{R})$ muni de l'addition et du produit de fonctions usuels. Si A contient les fonctions constantes et si A sépare les points de X, alors A est dense dans $C(X,\mathbb{R})$.*

Démonstration. On commence par noter que pour tout $\varepsilon > 0$ et tout $M > 0$, il existe un polynôme P tel que $\big|P(t) - |t|\big| \leq \varepsilon$ sur $[-M,M]$. Pour cela, il suffit de considérer le cas $M = 1$, par homothétie. On construit la suite de polynômes suivante :

$$Q_0(t) = 0, \quad Q_{n+1}(t) = Q_n(t) + \frac{1}{2}\big(t - Q_n(t)^2\big).$$

Pour $t \geq 0$, on a l'identité

$$\sqrt{t} - Q_{n+1}(t) = \big(\sqrt{t} - Q_n(t)\big)\big(1 - \frac{1}{2}(\sqrt{t} + Q_n(t))\big).$$

On en déduit immédiatement par récurrence que pour t dans l'intervalle $[0,1]$, la suite $Q_n(t)$ est croissante et que $0 \leq Q_n(t) < \sqrt{t}$ pour $0 < t \leq 1$. On déduit de l'identité ci-dessus que

$$0 \leq \frac{\sqrt{t} - Q_{n+1}(t)}{\sqrt{t} - Q_n(t)} \leq 1 - \frac{1}{2}\sqrt{t},$$

pour tout $0 < t \leq 1$, d'où

$$0 \leq \sqrt{t} - Q_n(t) \leq (1 - \frac{1}{2}\sqrt{t})^n\sqrt{t}$$

pour tout $0 \leq t \leq 1$. Il est facile de voir que le membre de droite de cette inégalité tend uniformément vers 0 quand n tend vers $+\infty$. Par conséquent, la suite de polynômes Q_n tend uniformément vers $\sqrt{t}$ sur $[0,1]$, ce qui implique immédiatement que la suite de polynômes $P_n(t) = Q_n(t^2)$ converge uniformément vers $|t|$ sur $[-1,1]$.

Soit maintenant $f \in \bar{A}$ quelconque. Alors pour tout polynôme P, la fonction $x \mapsto P(f(x))$ appartient aussi à $\bar{A}$ (il suffit de noter que l'application $f \mapsto f^k$ est continue de $C(X,\mathbb{R})$ dans $C(X,\mathbb{R})$, par le critère des suites par exemple et d'utiliser le fait que A est une algèbre). Comme X est compact, si l'on prend $M = \max_{x\in X}|f(x)|$ et la suite de polynômes du paragraphe précédent sur $[-M,M]$, on en déduit que

$$\forall f \in \bar{A}, \quad |f| = \lim_{n\to+\infty} P_n(f) \in \bar{A}.$$

Par conséquent,

$$\forall f \in \bar{A}, \quad f_+ = \sup(0,f) = \frac{1}{2}(f + |f|) \in \bar{A},$$

et

$$\forall f,g \in \bar{A}, \quad \sup(f,g) = f + (g-f)_+ \in \bar{A} \text{ ainsi que } \inf(f,g) \in \bar{A}.$$

On considère alors $f \in C(X,\mathbb{R})$ quelconque. Pour tout couple (x,y) de points distincts de X, il existe un élément $g_{x,y}$ de A tel que $g_{x,y}(x) = f(x)$ et $g_{x,y}(y) = f(y)$. En effet, A contient les constantes, est stable par multiplication par un scalaire et sépare les points.

On se donne $\varepsilon > 0$ et l'on fixe x. Pour tout $y \in X$, il existe un voisinage $V(y)$ tel que pour tout $y' \in V(y)$, $g_{x,y}(y') \leq f(y') + \varepsilon$, par continuité de $g_{x,y}$ et de f. Comme X est compact, on peut le recouvrir par un nombre fini de tels voisinages $V(y_1)$, $V(y_2)$, ..., $V(y_p)$. On pose alors $g_x = \inf_{1\leq i\leq p} g_{x,y_i}$. Par construction,

$$g_x(x) = f(x),\ g_x(y) \leq f(y) + \varepsilon \text{ pour tout } y \in X \text{ et } g_x \in \bar{A}.$$

On recommence alors la construction en notant que pour tout $x \in X$, il existe un voisinage $W(x)$ tel que pour tout $x' \in W(x)$, $f(x') - \varepsilon \leq g_x(x')$. Par compacité, on trouve m points x_k tels que les $W(x_k)$ recouvrent X et l'on pose $g = \sup_{1\leq k\leq m} g_{x_k}$, ce qui fournit une fonction $g \in \bar{A}$ telle que pour tout x dans X, $f(x) - \varepsilon \leq g(x) \leq f(x) - \varepsilon$. Par conséquent, $f \in \bar{\bar{A}} = \bar{A}$, d'où le théorème. □

Signalons le corollaire suivant, qui est le théorème original de Weierstrass.

Corollaire 1.4.4 *Soit f une fonction continue sur une partie compacte X de $\mathbb{R}^d$ à valeurs réelles. Pour tout ε strictement positif, il existe un polynôme à d indéterminées P et à coefficients réels tel que*

$$\sup_{x\in X} |f(x) - P(x)| < \varepsilon.$$

Autrement dit, les polynômes à coefficients réels sont denses dans l'espace des fonctions réelles continues sur un compact.

Le cas des fonctions à valeurs complexes est légèrement différent.

Théorème 1.4.10 *Soient (X,d) un espace métrique compact et A une sous-algèbre de $C(X,\mathbb{C})$. Si A contient les fonctions constantes, si A sépare les points de X et si $f \in A \Rightarrow \overline{f} \in A$, alors A est dense dans $C(X,\mathbb{C})$.*

On a un corollaire analogue.

Corollaire 1.4.5 *Soit f une fonction continue sur une partie compacte X de $\mathbb{R}^d$ à valeurs complexes. Pour tout ε strictement positif, il existe un polynôme P à coefficients complexes à d indéterminées tel que*

$$\sup_{x\in X} |f(x) - P(x)| < \varepsilon.$$

1.5 Espaces séparables, espaces connexes

Ces notions sont un peu moins fondamentales que celles d'espace complet ou d'espace compact. Cependant, elles sont souvent fort utiles.

Définition 1.5.1 *Soit $(X,\mathscr{O})$ un espace topologique. On dit que $(X,\mathscr{O})$ est* séparable *s'il existe une partie dénombrable dense dans X.*

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

$$\forall x \in X, \forall \varepsilon > 0, \exists\, m \text{ tel que } d(x,x_m) < \varepsilon.$$

Attention : la terminologie classique est un peu malheureuse dans la mesure où les notions d'espace séparé et d'espace séparable n'ont strictement rien à voir l'une avec l'autre. C'est regrettable, mais ainsi le veut l'usage.

Exemple. L'espace $\mathbb{R}$ muni de sa topologie usuelle est séparable, car l'ensemble dénombrable des rationnels est dense dans $\mathbb{R}$.

Proposition 1.5.1 *L'image d'un espace topologique séparable par une application continue est séparable.*

Démonstration. Soit $(X,\mathscr{O})$ un espace topologique séparable et $f\colon X \to Y$ une application continue à valeurs dans un autre espace topologique. On note $(x_n)_{n\in\mathbb{N}}$ une partie dénombrable dense dans X. Soit U un ouvert de Y tel que $f(X)\cap U \neq \emptyset$. Alors $f^{-1}(U)$ est un ouvert non vide de X, ce qui implique qu'il existe n tel que $x_n \in f^{-1}(U)$. Par conséquent, $f(x_n) \in U$ et la partie dénombrable $(f(x_n))_{n\in\mathbb{N}}$ est dense dans $f(X)$. □

Proposition 1.5.2 *Soit $(X_1,d_1),\cdots,(X_N,d_N)$ une famille de N espaces métriques séparables. Si l'on pose*

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

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

Démonstration. Pour chaque espace X_j, on dispose d'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 d, on a donc $d(x,y) < \varepsilon$ ce qui établit la proposition. □

Proposition 1.5.3 *Soient (X,d) un espace métrique séparable et A une partie quelconque de X. L'espace métrique (A,d) est séparable.*

Démonstration. Désignons par D une partie dénombrable dense de X. Pour tout entier n et tout x de 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 $\tilde{D} \stackrel{\text{déf}}{=} (a_{x,n})_{(x,n)\in D\times\mathbb{N}}$ de A.

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

Attention, la proposition 1.5.3 est fausse dans un espace topologique général au sens où l'on peut construire un espace topologique (non métrisable) séparable qui contient un sous-espace non séparable pour la topologie induite.

Théorème 1.5.1 *Tout espace métrique compact est séparable.*

Démonstration. Soit (X,d) un espace métrique compact. Construisons une suite dense. Soit $\mathscr{B}_n$ l'ensemble des boules ouvertes de rayon n^{-1}. Cet ensemble constitue bien évidemment un recouvrement ouvert de X. L'hypothèse de compacité fait que nous pouvons en extraire un sous-recouvrement fini. Pour chaque entier n, notons $(x_{j,n})_{1\leq j\leq J(n)}$ une famille finie de points telle que

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

On pose alors

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

Cet ensemble 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 $\frac{1}{n}\leq\varepsilon$. Par définition des points $x_{j,n}$, il existe un entier j tel que $d(x,x_{j,n}) < \frac{1}{n} \leq \varepsilon$ d'où la densité de D. □

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

Proposition 1.5.4 *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

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

où $\mathscr{P}$ désigne l'ensemble des parties de X et $\mathbf{1}_P$ la fonction caractéristique de P qui vaut 1 si $x \in P$ et 0 si $x \notin P$. Remarquons deux choses. Tout d'abord, si $P \neq Q$, alors $d(\mathbf{1}_P, \mathbf{1}_Q) = 1$. Ensuite, l'ensemble $\mathscr{P}([0,1])$ qui est en bijection avec A, n'est pas dénombrable. La démonstration de la proposition se conclut grâce au lemme 1.5.1 suivant. □

Lemme 1.5.1 *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 $a \in A$, il existe un élément z de D tel que

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

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

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

Donc z_1 est différent de z_2. L'application $a \mapsto z$ ainsi définie étant injective, si une partie est dense, elle ne peut être dénombrable. □

Notons que l'on peut par contre déduire immédiatement du théorème de Stone-Weierstrass l'énoncé suivant.

Proposition 1.5.5 *Si X est un compact de $\mathbb{K}^N$, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, alors l'espace $C(X,\mathbb{K})$ est séparable.*

Démonstration. Il suffit de remarquer que l'ensemble des polynômes à coefficients rationnels, lequel est dénombrable, est dense dans l'ensemble des polynômes pour la distance de $C(X,\mathbb{K})$. On en conclut qu'il est dense dans $C(X,\mathbb{K})$ par un argument de double limite. □

Passons maintenant à la notion de connexité. Il s'agit de formaliser l'idée qu'un espace topologique peut être formé de un ou plusieurs « morceaux d'un seul tenant ».

Définition 1.5.2 *On dit qu'un espace topologique* $(X,\mathscr{O})$ *est* connexe *s'il satisfait l'une des trois propriétés équivalentes suivantes :*

i) Il n'existe pas de partition de X *en deux parties ouvertes non vides.*

ii) Il n'existe pas de partition de X *en deux parties fermées non vides.*

iii) Les seules parties à la fois ouvertes et fermées de X *sont* $\emptyset$ *et* X*.*

Si A *est une partie de* X*, on dira que* A *est une* partie connexe *si elle est connexe pour la topologie induite.*

L'équivalence entre i), ii) et iii) est claire. La caractérisation fort utile suivante des connexes est parfois incluse dans leur définition.

Proposition 1.5.6 *Un espace topologique* $(X,\mathscr{O})$ *est connexe si et seulement toute application continue de* X *dans* $\{0,1\}$ *muni de la topologie discrète est constante.*

Démonstration. Soit $(X,\mathscr{O})$ un espace topologique connexe et $f\colon X\to\{0,1\}$ une application continue. Comme l'espace d'arrivée est discret, $\{0\}$ en est un ouvert et aussi un fermé. Par conséquent, $f^{-1}(\{0\})$ est un ouvert-fermé de X. C'est donc soit l'ensemble vide, auquel cas f est constante et vaut 1, soit X tout entier, auquel cas f est constante et vaut 0.

Réciproquement, supposons que X ne soit pas connexe. Il existe alors un ouvert-fermé A distinct de $\emptyset$ et de X. La fonction caractéristique de A, $f=\mathbf{1}_A$, est continue à valeurs dans $\{0,1\}$. En effet, $f^{-1}(\{1\})=A$ est ouvert et $f^{-1}(\{0\})=X\setminus A$ également. Elle est par ailleurs non constante car il existe $x\in A$, donc $f(x)=1$ et il existe aussi $y\in X\setminus A$, donc $f(y)=0$. $\square$

Donnons tout de suite un exemple important.

Proposition 1.5.7 *Les parties connexes de* $\mathbb{R}$ *sont les intervalles.*

Démonstration. Soit A une partie connexe de $\mathbb{R}$. Montrons que si x et y appartiennent à A, alors l'intervalle fermé $[x,y]$ est inclus dans A. On raisonne par l'absurde en supposant qu'il existe $z\in[x,y]$ tel que $z\notin A$. Dans ce cas, $A=(A\cap]-\infty,z[)\cup(A\cap]z,+\infty[)$ est une partition de A en deux ouverts non vides, contradiction. Appelons $a=\inf A\in[-\infty,+\infty[$ et $b=\sup A\in]-\infty,+\infty]$. D'après les propriétés des bornes supérieures et inférieures des parties de $\mathbb{R}$, pour tout $n>0$, il existe $a_n,b_n\in A$ tels que $a\le a_n\le a+1/n$ et $b-1/n\le b_n\le b$ (si a et b sont finis, avec les modifications évidentes dans le cas contraire). Par conséquent, A contient l'intervalle $[a_n,b_n]$ pour tout n et est donc l'un des quatre intervalles $[a,b]$, $]a,b]$, $[a,b[$ ou $]a,b[$.

Réciproquement, soit A un intervalle de $\mathbb{R}$ et $B \subset A$ une partie ouverte et fermée, non vide. On va montrer que $B = A$. Soit $c \in B$ et $D = \{x \in A; [c,x] \subset B\}$. On note $d = \sup D \in]-\infty, +\infty]$. Supposons que $d < b$. En particulier, d est fini, et il existe une suite croissante $x_n \in D$ telle que $x_n \to d$. Comme B est fermée, $d \in B$. Par ailleurs, B est ouverte dans A, donc il existe un intervalle $]d-\eta, d+\eta[$ avec $\eta > 0$ tel que $(]d-\eta, d+\eta[\cap A) \subset B$. Comme $d < b$, on peut choisir η assez petit pour que $d + \eta < b$ ce qui contredit le fait que d est la borne supérieure de D. Par conséquent, $b = d$ et si $b \in A$ alors $d \in B$. On recommence alors le même raisonnement à droite de c, pour en conclure que $B = A$. □

Comme toute bonne propriété topologique, la connexité se marie bien avec la continuité.

Théorème 1.5.2 *L'image d'un espace topologique connexe par une application continue est connexe.*

Démonstration. Soient $(X, \mathscr{O})$ et $(Y, \tilde{\mathscr{O}})$ deux espaces topologiques avec X connexe et soit f continue de X dans Y. On considère une application $g\colon f(X) \to \{0,1\}$ continue. Alors $g \circ f$ est continue de X dans $\{0,1\}$, donc constante par la proposition 1.5.6. Par conséquent, g est aussi constante et $f(X)$ est connexe par cette même proposition. □

Le corollaire suivant généralise le théorème des valeurs intermédiaires.

Corollaire 1.5.1 *Si f est continue de $(X, \mathscr{O})$ connexe à valeurs dans $\mathbb{R}$, alors l'image de f est un intervalle.*

Démonstration. On vient de voir qu'un connexe de $\mathbb{R}$ est un intervalle. □

Proposition 1.5.8 *L'adhérence d'une partie connexe A d'un espace topologique $(X, \mathscr{O})$ est connexe.*

Démonstration. Considérons une application g continue de $\bar{A}$ à valeurs dans $\{0,1\}$. La restriction de g à A est donc constante, mettons qu'elle vaille 0. Soit maintenant $x \in \bar{A} \setminus A$. L'ensemble $\{g(x)\}$ est un ouvert de $\{0,1\}$, donc son image réciproque est un ouvert U de X qui contient x. Mais par définition de l'adhérence, il existe $y \in A$ tel que $y \in U$. On a donc $y \in g^{-1}(\{g(x)\})$, ce qui n'est qu'une façon compliquée de dire que $0 = g(y) = g(x)$, et g est constante sur $\bar{A}$. □

A contrario, l'intérieur d'une partie connexe n'a aucune raison d'être connexe, comme le montre l'exemple de deux disques fermés disjoints du plan reliés par un segment.

Lemme 1.5.2 *La réunion de deux parties connexes A et B de X d'intersection non vide est connexe.*

Démonstration. Soit $f : A \cup B \to \{0,1\}$ continue. Alors $f_{|A}$ et $f_{|B}$ sont constantes. Comme $A \cap B$ est non vide, elles prennent donc la même valeur sur A et sur B, ce qui implique que f est constante sur $A \cup B$ et que $A \cup B$ est connexe. □

La réunion de deux parties connexes d'intersection vide peut être ou ne pas être connexe (ce n'est donc pas une question), comme le montrent les exemples $[0,1[\cup[1,2]$ et $[0,1[\cup]1,2]$.

Définition 1.5.3 *On dit qu'un point x d'un espace topologique $(X,\mathscr{O})$ est* connecté à *un autre point y de ce même espace s'il existe une partie connexe de X qui les contient tous les deux.*

Théorème 1.5.3 *La relation « x est connecté à y » est une relation d'équivalence sur X.*

Démonstration. Il est clair que cette relation est réflexive — x est connecté à x pour tout x — et symétrique — si x est connecté à y alors y est connecté à x. Montrons qu'elle est transitive. Soient trois points x, y et z tels que x est connecté à y et y est connecté à z. Il existe donc une partie connexe A de X qui contient x et y et une partie connexe B qui contient y et z. Mais $A \cap B$ est non vide puisqu'elle contient y. Par conséquent, le lemme 1.5.2 montre que $A \cup B$ est connexe. Cette partie contenant x et z, on en déduit que x est connecté à z. □

Définition 1.5.4 *Une classe d'équivalence pour cette relation s'appelle* une composante connexe de *X. L'espace X est réunion disjointe de ses composantes connexes. On appelle* composante connexe de *x la composante connexe de X qui contient x.*

Commentaire. On a évidemment une définition analogue pour les composantes connexes d'une partie A de X. On montre facilement que la composante connexe de x est la plus grande partie connexe de A qui contient x. C'est donc un fermé dans A par la proposition 1.5.8. Les composantes connexes de A sont les « morceaux d'un seul tenant » auxquels on a fait allusion plus haut.

Exemple. Les composantes connexes de $\mathbb{Q}$ muni de la topologie induite de $\mathbb{R}$ sont réduites à des points. On dit que $\mathbb{Q}$ est complètement discontinu.

1.6 Quelques remarques sur la topologie des ouverts de $\mathbb{R}^d$

Le but de cette section est essentiellement de démontrer quelques résultats qui nous seront utiles dans la suite. Ici, lorsque nous ferons référence à des propriétés métriques, on supposera toujours $\mathbb{R}^d$ muni de la distance associée à la norme euclidienne.

Rappelons à nouveau que, sauf entre autres dans le cas très particulier de $\mathbb{R}^d$,

$$\textbf{fermé et borné} \quad \neq \quad \textbf{compact}.$$

Il est cependant possible de décrire les compacts d'un ouvert de $\mathbb{R}^d$ grâce au théorème suivant.

Théorème 1.6.1 *Soit A une partie fermée (pour la topologie induite) et bornée d'un ouvert Ω de $\mathbb{R}^d$. Alors*

$$A \text{ est compacte } \Longleftrightarrow \inf_{x\in A} d(x, \mathbb{R}^d \setminus \Omega) > 0.$$

Démonstration. Supposons A compacte. Comme Ω est ouvert, pour tout x de A, $d(x, \mathbb{R}^d \setminus \Omega)$ est strictement positif. D'après l'exercice 1.2.5, on sait que la fonction distance d'un point à un ensemble est continue. Le corollaire 1.4.1 nous dit que l'infimum est atteint, il est donc strictement positif.

Réciproquement, soit A une partie fermée et bornée de Ω vérifiant

$$\delta = \inf_{x\in A} d(x, \mathbb{R}^d \setminus \Omega) > 0.$$

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 $\bar{\Omega}_{\delta/2}$ le fermé de $\mathbb{R}^d$ défini par

$$\bar{\Omega}_{\delta/2} \stackrel{\text{déf}}{=} \{x \in \mathbb{R}^d ; d(x, \mathbb{R}^d \setminus \Omega) \geq \delta/2\},$$

(c'est l'image réciproque d'un fermé par une application continue). Comme $\bar{\Omega}_{\delta/2} \subset \Omega$ et que $A \subset \bar{\Omega}_{\delta/2}$, on a

$$A = A \cap \bar{\Omega}_{\delta/2} = (B \cap \Omega) \cap \bar{\Omega}_{\delta/2} = B \cap (\Omega \cap \bar{\Omega}_{\delta/2}) = B \cap \bar{\Omega}_{\delta/2}.$$

L'ensemble $B \cap \bar{\Omega}_{\delta/2}$ 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 résultat suivant nous sera également utile.

Théorème 1.6.2 *Soit Ω un ouvert de $\mathbb{R}^d$, il existe une suite de compacts $(K_n)_{n\in\mathbb{N}}$ de Ω telle que $\bigcup_{n\in\mathbb{N}} K_n = \Omega$, $K_n \subset \operatorname{Int} K_{n+1}$ et pour tout compact K de Ω, il existe un entier n tel que $K \subset K_n$.*

Démonstration. Il suffit de poser

$$K_n \stackrel{\text{déf}}{=} \bar{B}(0,n) \cap \Big\{ x \in \mathbb{R}^d ; d(x, \mathbb{R}^d \setminus \Omega) \geq \frac{1}{n} \Big\} \cdot$$

D'après le théorème précédent, c'est un compact de Ω. Comme Ω est un ouvert, pour tout x de Ω, $d(x, \mathbb{R}^d \setminus \Omega)$ est strictement positif. Donc, il existe un entier n tel que $x \in K_n$, d'où $\bigcup_{n \in \mathbb{N}} K_n = \Omega$. De plus,

$$K_n \subset B(0, n+1) \cap \Big\{ x \in \mathbb{R}^d ; d(x, \mathbb{R}^d \setminus \Omega) > \frac{1}{n} \Big\} \subset \operatorname{Int} K_{n+1},$$

d'où le premier point du théorème. Quant au second, il suffit d'observer que

$$\bigcup_{n \in \mathbb{N}} \operatorname{Int} K_{n+1} = \Omega$$

puisque $K_n \subset \operatorname{Int} K_{n+1}$. Les ouverts $(\operatorname{Int} K_{n+1})_{n \in \mathbb{N}}$ recouvrent donc K, on en extrait un sous-recouvrement fini et l'on prend l'adhérence du plus grand de ces ouverts, qui est bien un certain K_n. □

Définition 1.6.1 *Soit Ω un ouvert de $\mathbb{R}^d$, on appelle* suite exhaustive *de compacts toute suite de compacts qui vérifie les conclusions du théorème* 1.6.2 *ci-dessus.*

Pour ce qui concerne la connexité, on a le résultat suivant.

Proposition 1.6.1 *Tout ouvert de $\mathbb{R}^d$ est réunion disjointe au plus dénombrable d'ouverts connexes. En particulier, tout ouvert de $\mathbb{R}$ est réunion disjointe au plus dénombrable d'intervalles ouverts.*

Démonstration. Soit Ω un ouvert de $\mathbb{R}^d$. Il est égal à la réunion disjointe de ses composantes connexes. Or, toute boule ouverte B (disons pour la distance euclidienne pour fixer les idées) de $\mathbb{R}^d$ est connexe. En effet, supposons que l'on puisse écrire $B = U_1 \cup U_2$ avec U_1, U_2 ouverts non vides et $U_1 \cap U_2 = \emptyset$. Choisissons $x \in U_1, y \in U_2$ et définissons $f : I = [0,1] \to B$ par $f(t) = tx + (1-t)y$. L'application f est continue (c'est le segment qui joint x à y). Son image est donc connexe, mais $f(I) = B \cap f(I) = (U_1 \cup U_2) \cap f(I) = (U_1 \cap f(I)) \cup (U_2 \cap f(I))$, $(U_1 \cap f(I)) \cap (U_2 \cap f(I)) = \emptyset$ et $x \in U_1 \cap f(I) \neq \emptyset$ et $y \in U_2 \cap f(I) \neq \emptyset$, contradiction.

On déduit de ceci que les composantes connexes de Ω sont ouvertes. En effet, soit C_x la composante connexe de x. Il existe une boule centrée en x et incluse dans Ω. Comme cette boule est connexe, elle est incluse dans C_x qui est la plus grande partie connexe de Ω contenant x. Donc C_x est ouvert.

Pour conclure, il suffit de remarquer que Ω est séparable, et donc on définit une injection de l'ensemble des composantes connexes dans $\mathbb{N}$ en prenant, pour chaque composante connexe, le numéro du premier élément de la partie dénombrable dense lui appartenant. □

Commentaire. On appelle parfois *domaine* un ouvert connexe de $\mathbb{R}^d$. Dans les problèmes d'équations aux dérivées partielles, il est souvent suffisant de considérer des domaines, car les équations posées dans des composantes connexes disjointes sont en général totalement découplées les unes des autres.

Ainsi s'achève notre survol de la topologie générale. Il convient de se rendre compte que nous n'avons fait qu'effleurer le sujet, lequel est infiniment plus vaste que ce qui peut transparaître de ces notes. Nous avons essentiellement tenté d'en dégager ce qui est le plus utile pour les besoins ultérieurs de l'analyse appliquée.

Chapitre 2

Espaces vectoriels normés, espaces de Banach

Nous allons maintenant nous intéresser à l'interaction entre la topologie et l'algèbre linéaire. Nous nous restreindrons à la classe des espaces vectoriels normés. Il existe bien d'autres classes d'espaces vectoriels topologiques, utiles en analyse, mais dont l'étude dépasse le cadre de ce cours. Dans la suite, $\mathbb{K}$ désignera $\mathbb{R}$ ou $\mathbb{C}$ et la notation $|\cdot|$ désignera suivant le cas la valeur absolue ou le module. Tous les espaces vectoriels considérés seront des $\mathbb{K}$-espaces vectoriels. La plupart des notions introduites peuvent être étendues à des corps $\mathbb{K}$ plus généraux.

2.1 Définition des espaces vectoriels normés

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* norme *sur E si les trois conditions suivantes sont satisfaites*

- $N(x+y) \leq N(x)+N(y)$ (*inégalité triangulaire*).
- *Pour tout λ appartenant à $\mathbb{K}$, on a $N(\lambda x) = |\lambda|N(x)$ (homogénéité positive),*
- $N(x) = 0$ *implique que* $x = 0$.

Une application N de E dans $\mathbb{R}_+$ qui satisfait les deux premières propriétés est une semi-norme *sur E.*

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

Exemples. On considère sur $\mathbb{K}^d$ les trois quantités suivantes

$$\|x\|_2 \stackrel{\text{déf}}{=} \Big(\sum_{j=1}^{d} |x_j|^2\Big)^{\frac{1}{2}}, \quad \|x\|_\infty \stackrel{\text{déf}}{=} \max_{1\leq j\leq d} |x_j|, \quad \|x\|_1 \stackrel{\text{déf}}{=} \sum_{j=1}^{d} |x_j|.$$

Ce sont trois normes sur $\mathbb{K}^d$ (qui coïncident avec la valeur absolue ou le module pour $d = 1$). La première d'entre elles est la norme euclidienne ou hermitienne usuelle.

Soit X un ensemble, on considère le $\mathbb{K}$-espace vectoriel $\mathscr{B}(X,\mathbb{K})$ des fonctions bornées de X à valeurs dans $\mathbb{K}$. On définit alors l'application

$$f \mapsto \sup_{x\in X} |f(x)|.$$

C'est une norme sur l'espace $\mathscr{B}(X,\mathbb{K})$ au sens ci-dessus.

Parmi tous les exemples que l'on peut donner, il en est de très importants, notamment celui qui suit.

Théorème 2.1.1 *Soit (X,μ) un espace mesuré. Pour tout réel $p \in [1,+\infty]$, on considère l'espace $L^p(X,d\mu)$. L'application de $L^p(X,d\mu)$ dans $\mathbb{R}_+$ définie*

$$\text{pour } p < +\infty \text{ par } f \mapsto \|f\|_{L^p} \stackrel{\text{déf}}{=} \left(\int_X |f(x)|^p \, d\mu(x)\right)^{\frac{1}{p}} \quad \text{et}$$
$$\text{pour } p = +\infty \text{ par } f \mapsto \sup\text{ess}\,(|f|)$$

est une norme.

On rappelle que, pour $p \in [1,+\infty[$, l'espace $L^p(X,d\mu)$ est l'espace des classes d'équivalences des fonctions sur X à valeurs dans $\mathbb{K}$ qui sont μ-mesurables et de puissance p-ième μ-intégrable, modulo l'égalité μ-presque partout. Pour $p = +\infty$, on remplace « de puissance p-ième μ-intégrable », par μ-essentiellement bornée (*i.e.* bornée en dehors d'un ensemble de μ-mesure nulle). Nous reverrons tout ceci en détail au chapitre 5.

Définition 2.1.2 *Soient E un espace vectoriel sur $\mathbb{K}$ et N une norme sur E. Le couple (E,N) est appelé un* espace vectoriel normé *ou plus rapidement espace normé.*

La proposition suivante va munir naturellement un espace normé d'une structure d'espace métrique.

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

$$\begin{aligned} E\times E &\to \mathbb{R}_+ \\ (x,y) &\mapsto \|x-y\| \end{aligned}$$

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

La démonstration est laissée en exercice simple.

Convention : Sauf mention expresse du contraire, on considérera toujours l'espace E muni de la structure métrique ainsi définie. On appelle alors *boule unité* de E la boule $B(0,1) = \{x \in E; \|x\| < 1\}$ pour cette distance. C'est le prototype de toutes les boules de E qui s'en déduisent par une translation et une homothétie (le vérifier). Dans le cas des espaces vectoriels normés, on a $\mathrm{Int}\bar{B}(0,1) = B(0,1)$, $\mathrm{Adh}\,B(0,1) = \bar{B}(0,1)$ et $\partial B(0,1) = \partial\bar{B}(0,1) = S(0,1) = \{x \in E; \|x\| = 1\}$.

Comme dans le cas des espaces métriques, on peut avoir à considérer plusieurs normes sur un même espace vectoriel et on est donc souvent amené à les comparer entre elles.

Définition 2.1.3 *On dit que deux normes N_1 et N_2 sur E sont équivalentes s'il existe deux constantes strictement positives α et β telles que*

$$\forall x \in E, \quad \alpha N_1(x) \leq N_2(x) \leq \beta N_1(x).$$

Commentaires. Si c'est le cas, alors on a aussi $\beta^{-1}N_2(x) \leq N_1(x) \leq \alpha^{-1}N_2(x)$, c'est une relation symétrique entre normes. Il est facile de voir que si deux normes sont équivalentes, alors les distances qui leur sont associées sont fortement équivalentes. Elles définissent donc la même topologie métrique sur E et les mêmes suites de Cauchy.

Nous pouvons dès maintenant percevoir de très grandes différences entre la dimension finie et la dimension infinie.

Proposition 2.1.2 *Soit E un espace vectoriel de dimension finie. Toutes les normes sur E sont équivalentes.*

Démonstration. Soit d la dimension de E et choisissons $(e_1, e_2, \dots, e_d)$ une base de E. On munit E de la norme $\|x\| = \max_{1\leq i\leq d} |x_i|$, où les x_i sont les composantes de x sur la base choisie (vérifier que c'est bien une norme). Ceci fournit une structure d'espace métrique sur E et il est facile de vérifier — mais encore une fois faut-il le faire — que l'application $x \mapsto (x_i)_{1\leq i\leq d}$ réalise un homéomorphisme entre E et $\mathbb{K}^d$ muni de la norme $\|\cdot\|_\infty$, c'est même une isométrie. Par conséquent, la sphère unité $S(0,1)$ de E est compacte (si $\mathbb{K} = \mathbb{C}$, il suffit de noter que $\mathbb{C}^d$ est trivialement homéomorphe à $\mathbb{R}^{2d}$).

Considérons maintenant une seconde norme N sur E. Par l'inégalité triangulaire, on a pour tout couple $(x,y) \in E^2$

$$|N(x) - N(y)| \leq N(x-y), \tag{2.1}$$

ainsi que

$$0 \leq N(x) \leq \sum_{i=1}^{d} |x_i| N(e_i) \leq \Big(\sum_{i=1}^{d} N(e_i)\Big) \|x\|. \tag{2.2}$$

On déduit immédiatement de ces deux inégalités que la fonction N est continue pour la topologie métrique induite par la première norme.

Elle atteint par conséquent ses bornes sur le compact $S(0,1)$ c'est-à-dire qu'il existe z_1 et z_2 dans $S(0,1)$ tels que pour tout z dans $S(0,1)$,

$$N(z_1) \leq N(z) \leq N(z_2).$$

Comme $z_1 \neq 0$, on voit que $\alpha = N(z_1) > 0$. Posons également $\beta = N(z_2)$.

Pour $x = 0$, on a trivialement $\alpha\|x\| \leq N(x) \leq \beta\|x\|$. Pour $x \neq 0$, on a $x/\|x\| \in S(0,1)$, d'où

$$\alpha \leq N(x/\|x\|) = \frac{1}{\|x\|}N(x) \leq \beta,$$

ce qui implique le résultat. □

La situation en dimension infinie est radicalement différente. Considérons par exemple l'espace $E = \mathscr{C}([0,1],\mathbb{R})$ muni des deux normes $\|f\|_\infty = \max_{x\in[0,1]} |f(x)|$ et $\|f\|_1 = \int_0^1 |f(x)|\,dx$. Il n'existe aucune constante strictement positive telle que $\|f\|_\infty \leq \beta\|f\|_1$. En effet, la suite $f_n(x) = 1 - nx$ pour $0 \leq x \leq 1/n$ et $f_n(x) = 0$ pour $1/n < x \leq 1$ est telle que $\|f_n\|_\infty = 1$ mais $\|f_n\|_1 = 1/2n$.

Remarquons que l'on a utilisé la compacité des fermés bornés de $\mathbb{R}$, qui repose sur la complétude de $\mathbb{R}$. La proposition 2.1.2 est par exemple fausse pour un $\mathbb{Q}$-espace vectoriel.

Proposition 2.1.3 *Soient E et F deux $\mathbb{K}$-espaces vectoriels normés tels que E soit de dimension finie. Toute application linéaire de E dans F est continue.*

Démonstration. On note $\|\cdot\|_E$ et $\|\cdot\|_F$ les normes respectives de E et F. Soit d la dimension de E et $(e_1, e_2, \ldots, e_d)$ une base de E. Comme toutes les normes sur E sont équivalentes, on a $\max_{1\leq i\leq d} |x_i| \leq \beta\|x\|_E$ pour une certaine constante $\beta > 0$, où les x_i sont les composantes de x sur la base choisie. On considère une application linéaire quelconque f de E dans F. Pour tout $x \in E$, on a

$$x = \sum_{i=1}^{d} x_i e_i \quad \text{d'où} \quad f(x) = \sum_{i=1}^{d} x_i f(e_i).$$

Comme $f(x) - f(y) = f(x-y)$ par linéarité, il suffit de montrer que f est continue en 0. Prenant les normes dans l'égalité ci-dessus et utilisant l'inégalité triangulaire, il vient

$$\|f(x)\|_F \leq \sum_{i=1}^{d} |x_i|\|f(e_i)\|_F \leq \beta\Big(\sum_{i=1}^{d} \|f(e_i)\|_F\Big)\|x\|_E.$$

Donc, pour toute suite x_n telle que $\|x_n\|_E \to 0$, on a bien $\|f(x_n)\|_F \to 0$, ce qui montre la continuité de f. □

Situation encore une fois radicalement différente si l'espace de départ E est de dimension infinie (la dimension de l'espace d'arrivée ne joue absolument aucun rôle dans cette affaire). Par exemple, soit E l'espace des fonctions C^1 de l'intervalle $[0,1]$ dans $\mathbb{R}$ muni de la norme $\|f\|_\infty$. Considérons l'application δ_0' définie par

$$\delta_0' \colon E \to \mathbb{R},\\ f \mapsto f'(0).$$

Cette forme linéaire n'est pas continue (exercice : vérifiez-le !). On peut alors montrer que son noyau, qui est un hyperplan, n'est pas fermé. En dimension infinie en général, un sous-espace vectoriel peut ne pas être fermé.

Proposition 2.1.4 *Tout sous-espace de dimension finie d'un espace vectoriel normé est fermé.*

Démonstration. En effet, soit F un sous-espace vectoriel de dimension finie de E et soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de F qui converge dans vers x dans E. Elle est donc bornée. On a vu que la boule unité de F est relativement compacte, donc il existe une sous-suite x_{n_p} et un $y \in F$ tels que $x_{n_p} \to y$ quand $p \to +\infty$. Par unicité de la limite, on voit que $x = y \in F$ ce qui montre que F est fermé. □

Une autre différence entre dimension finie et infinie, qui s'avérera capitale, est que la boule unité d'un espace de dimension infinie n'est jamais compacte. C'est ce qu'affirme le théorème de Riesz.

Théorème 2.1.2 *Soit E un $\mathbb{K}$-espace vectoriel normé. La dimension de E est finie si et seulement si la boule unité fermée de E est compacte.*

Démonstration. On a déjà vu que la boule unité fermée d'un espace de dimension finie est compacte. Réciproquement, soit E un espace normé dont la boule unité fermée est compacte. Choisissons un nombre $0 < \alpha < 1$ et recouvrons $\bar{B}(0,1)$ par un nombre fini de boules ouvertes de rayon α centrées en des points x_i, $1 \le i \le N$, de $\bar{B}(0,1)$. On note $M = \operatorname{vect}(x_1, x_2, \dots, x_N)$. C'est un sous-espace vectoriel de E de dimension finie. On va montrer que $E = M$.

On procède par récurrence. Soit $x \in \bar{B}(0,1)$. Par la propriété de recouvrement, il existe un indice i_0 et un point $y_1 \in B(0,1)$ tels que

$$x = x_{i_0} + \alpha y_1.$$

Posons $z_1 = x_{i_0}$. C'est un élément de M et $(x - z_1) \in \alpha B(0,1)$. Faisons l'hypothèse de récurrence que nous ayons déjà construit $z_k \in M$ et $y_k \in B(0,1)$ tels que $x - z_k = \alpha^k y_k$. Comme $y_k \in \bar{B}(0,1)$, par la propriété de recouvrement, il existe un indice i_k et un point $y_{k+1} \in B(0,1)$ tels que

$$y_k = x_{i_k} + \alpha y_{k+1}.$$

On en déduit que

$$x - z_k = \alpha^k (x_{i_k} + \alpha y_{k+1}) = \alpha^k x_{i_k} + \alpha^{k+1} y_{k+1},$$

d'où en posant $z_{k+1} = z_k + \alpha^k x_{i_k} \in M$, on obtient bien $x - z_{k+1} = \alpha^{k+1} y_{k+1}$.

On déduit du raisonnement précédent que pour tout $x \in \bar{B}(0,1)$, il existe une suite $(z_k)_{k\in\mathbb{N}}$ de M telle que

$$\|x - z_k\|_E \leq \alpha^k \longrightarrow 0 \text{ quand } k \to +\infty.$$

La suite $(z_k)_{k\in\mathbb{N}}$ converge donc vers x et comme M est fermé, on en déduit que $x \in M$, c'est-à-dire $\bar{B}(0,1) \cap M = \bar{B}(0,1)$. Ceci implique immédiatement par homothétie que $E = M$. □

Commentaires. Un corollaire immédiat du résultat précédent est qu'un compact d'un espace normé de dimension infinie est nécessairement d'intérieur vide. En d'autres termes, un point n'admet aucun voisinage compact (on dit qu'un tel espace n'est pas *localement compact*). Un autre corollaire est qu'une suite bornée dans un espace normé de dimension infinie peut parfaitement n'avoir aucune valeur d'adhérence et qu'une fonction continue sur la boule unité fermée n'a aucune raison d'y être bornée, ou d'atteindre ses bornes si elle est bornée.

Définition 2.1.4 *Soit A une partie d'un espace vectoriel normé $(E, \|\cdot\|)$. On dit que A est* totale *si l'espace vectoriel engendré par A est dense dans E.*

Proposition 2.1.5 *Dans un espace vectoriel normé séparable, il existe une partie totale libre au plus dénombrable.*

Démonstration. Soit $(x_n)_{n\in\mathbb{N}}$ une partie dénombrable dense de E. Il existe au moins un x_n non nul. On note x_{n_0} le premier d'entre eux. On procède ensuite par récurrence. Supposons extraite de la suite $(x_n)_{n\in\mathbb{N}}$ une sous-suite libre $(x_{n_p})_{0\leq p\leq q}$ telle que $\operatorname{vect}\{x_i, 0 \leq i \leq n_q\} = \operatorname{vect}\{x_{n_p}, 0 \leq p \leq q\}$. Alors, soit $\operatorname{vect}\{x_{n_p}, 0 \leq p \leq q\} = E$, auquel cas E est de dimension finie et l'on a terminé, soit il existe $n > n_q$ tel que $x_n \notin \operatorname{vect}\{x_{n_p}, 0 \leq p \leq q\}$. On prend pour $x_{n_{q+1}}$ le premier de ces vecteurs. La famille $(x_{n_p})_{p\in\mathbb{N}}$ répond à la question. □

Exercice 2.1.1 *Soit $(e_n)_{n\in\mathbb{N}}$ la famille définie par $e_n(p) = \delta_{n,p}$. Démontrez que cette famille est totale dans $\ell^p(\mathbb{N})$.*

Sur un espace vectoriel, les applications linéaires jouent évidemment un rôle privilégié parmi toutes les applications. Dans le cas des espaces normés, leur continuité se vérifie de manière très simple.

Théorème 2.1.3 *Soient E et F deux espaces vectoriels normés et f une application linéaire de E dans F. L'application f est continue si et seulement si il existe un ouvert U de E tel que f soit bornée sur U, c'est-à-dire tel que*

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

Démonstration. Soit f une application linéaire de E dans F qui ne soit bornée sur aucun ouvert. En particulier, elle n'est pas bornée sur la boule unité et il existe une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de E telle que

$$\forall n \in \mathbb{N},\ x_n \in B(0,1) \quad \text{et} \quad \|f(x_n)\|_F \longrightarrow +\infty \text{ quand } n \to +\infty.$$

Posons $y_n = x_n / \|f(x_n)\|_F$. Alors

$$\|y_n\|_E = \frac{\|x_n\|_E}{\|f(x_n)\|_F} \leq \frac{1}{\|f(x_n)\|_F} \to 0 \text{ quand } n \to +\infty,$$

c'est-à-dire que $y_n \to 0$ dans E d'une part et

$$\|f(y_n)\|_F = \left\| f\Big(\frac{x_n}{\|f(x_n)\|_F}\Big)\right\|_F = \left\|\frac{f(x_n)}{\|f(x_n)\|_F}\right\|_F = \frac{\|f(x_n)\|_F}{\|f(x_n)\|_F} = 1$$

d'autre part, si bien que $f(y_n) \not\to 0 = f(0)$ dans F. Donc f n'est pas continue. Par contraposition, on en déduit que si f est linéaire continue, alors elle est bornée sur au moins un ouvert (à savoir ici la boule unité).

Réciproquement, soit f linéaire bornée sur un ouvert U. Pour montrer qu'elle est continue, il suffit de montrer qu'elle est continue en 0, par linéarité. Soit x_0 appartenant à U. Par définition d'un ouvert, il existe une boule ouverte de centre x_0 et de rayon $\alpha > 0$ incluse dans U. On peut alors écrire

$$\begin{aligned}
\sup_{x \in B(0,\alpha)} \|f(x)\|_F &= \sup_{y \in B(x_0,\alpha)} \|f(y - x_0)\|_F \\
&\leq \sup_{y \in U} \|f(y - x_0)\|_F \\
&\leq \sup_{y \in U} \|f(y)\|_F + \|f(x_0)\|_F.
\end{aligned}$$

Donc l'application f est bornée sur la boule de centre 0 et de rayon α. Appelons M sa borne supérieure sur cette boule. Pour tout y de $E \setminus \{0\}$, on a

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

On en déduit que

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

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

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

On en déduit immédiatement que f est continue en 0. $\square$

Attention, on a bien pris garde de ne pas utiliser d'argument de compacité : une application linéaire continue est bornée sur la boule unité (et réciproquement) même en dimension infinie ! On tire aisément de la démonstration précédente le critère de continuité qui suit et que l'on utilise *systématiquement* dans la pratique.

Corollaire 2.1.1 *Une application linéaire f de E dans F est continue si et seulement si il existe une constante C telle que*

$$\forall x \in E, \quad \|f(x)\|_F \leq C\|x\|_E.$$

En particulier, si f est linéaire continue, elle est 1-lipschitzienne. Le théorème 2.1.3 se généralise assez facilement aux applications multilinéaires.

Théorème 2.1.4 *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 $E_1 \times \cdots \times E_N$ dans F si et seulement si*

$$\sup_{\|x_1\|_1\leq 1, \cdots, \|x_N\|_N \leq 1} \|f(x_1, \cdots, x_N)\|_F < +\infty.$$

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

La démonstration de ce théorème est un bon exercice. De même une application multilinéaire est continue si et seulement si on peut trouver une constante C telle que

$$\forall (x_1, \cdots, x_N) \in E_1 \times \cdots \times E_N, \quad \|f(x_1, \cdots, x_N)\|_F \leq C \prod_{i=1}^{N} \|x_i\|_{E_i}.$$

Proposition 2.1.6 *Soient E et F deux espaces vectoriels normés. On désigne par $\mathscr{L}(E,F)$ l'espace vectoriel des applications linéaires continues de E dans F. L'application*

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

est une norme sur $\mathscr{L}(E,F)$.

Démonstration. Il est très facile de montrer que $\mathscr{L}(E,F)$ est un espace vectoriel. Vérifions les trois propriétés définissant une norme.

Supposons que $\|f\|_{\mathscr{L}(E,F)} = 0$. Ceci implique que $f(x) = 0$ pour tout x de la boule unité fermée. Alors, pour tout $x \neq 0$, $y = x/\|x\|_E$ est dans cette boule, d'où

$$f(x) = f(\|x\|_E y) = \|x\|_E f(y) = 0.$$

Par conséquent, $f = 0$.

Soient $f \in \mathscr{L}(E,F)$ et $\lambda \in \mathbb{K}$, on a

$$\|\lambda f(x)\|_F = |\lambda| \|f(x)\|_F.$$

Il en résulte que

$$\sup_{\|x\|_E \leq 1} \|\lambda f(x)\|_F = |\lambda| \sup_{\|x\|_E \leq 1} \|f(x)\|_F,$$

d'où l'homogénéité positive.

Enfin, soient f_1 et f_2 deux éléments de $\mathscr{L}(E,F)$. On a

$$\|f_1(x) + f_2(x)\|_F \leq \|f_1(x)\|_F + \|f_2(x)\|_F$$

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

$$\|f_1(x) + f_2(x)\|_F \leq \|f_1\|_{\mathscr{L}(E,F)} + \|f_2\|_{\mathscr{L}(E,F)}.$$

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

$$\|f_1 + f_2\|_{\mathscr{L}(E,F)} \leq \|f_1\|_{\mathscr{L}(E,F)} + \|f_2\|_{\mathscr{L}(E,F)},$$

d'où l'inégalité triangulaire. □

Une conséquence immédiate des résultats qui précèdent.

Corollaire 2.1.2 *Soit $f \in \mathscr{L}(E,F)$, alors*

$$\forall x \in X, \quad \|f(x)\|_F \leq \|f\|_{\mathscr{L}(E,F)} \|x\|_E,$$

et $\|f\|_{\mathscr{L}(E,F)}$ est la plus petite constante telle que cette inégalité ait lieu.

On voit donc également que

$$\|f\|_{\mathscr{L}(E,F)} = \sup_{x \in E \setminus \{0\}} \frac{\|f(x)\|_F}{\|x\|_E} = \sup_{\|x\|_E = 1} \|f(x)\|_F.$$

Proposition 2.1.7 *Soient E, F et G trois espaces vectoriels normés, $f \in \mathscr{L}(E,F)$ et $g \in \mathscr{L}(F,G)$. La composée $g \circ f$ appartient à $\mathscr{L}(E,G)$ et*

$$\|g \circ f\|_{\mathscr{L}(E,G)} \leq \|f\|_{\mathscr{L}(E,F)} \|g\|_{\mathscr{L}(F,G)}.$$

Démonstration. On sait déjà que $g \circ f$ est linéaire continue. Il suffit alors d'observer que pour tout $x \in E$.

$$\begin{aligned}\|g \circ f(x)\|_G &\le \|g\|_{\mathscr{L}(F,G)} \|f(x)\|_F \\ &\le \|g\|_{\mathscr{L}(F,G)} \|f\|_{\mathscr{L}(E,F)} \|x\|_E,\end{aligned}$$

et l'on prend la borne supérieure pour x dans la boule unité de E. □

Remarque. Si $E = F = G$, on note simplement $\mathscr{L}(E) \stackrel{\text{déf}}{=} \mathscr{L}(E,E)$. Pour tout $(f,g) \in \mathscr{L}(E)^2$,

$$\|g \circ f\|_{\mathscr{L}(E)} \le \|f\|_{\mathscr{L}(E)} \|g\|_{\mathscr{L}(E)}. \tag{2.3}$$

On dit que $\mathscr{L}(E)$ muni de sa structure d'espace vectoriel et de la composition est une *algèbre normée*.

Définition 2.1.5 *Soient E et F deux espaces vectoriels normés. On dit qu'une application linéaire f de E dans F est un* isomorphisme *si elle est bijective et si f et f^{-1} sont continues. On dit que E et F sont* isomorphes *s'il existe un isomorphisme entre E et F.*

D'après ce qui précède, il est clair que f est un isomorphisme si et seulement si elle est surjective et il existe une constante $C > 0$ telle que

$$\forall x \in E, \quad C^{-1} \|x\|_E \le \|f(x)\|_F \le C\|x\|_E.$$

Deux espaces normés isomorphes ont donc essentiellement la même structure algébrique, la même structure topologique et la même structure uniforme.

2.2 Espaces de Banach

Nous avons vu que la complétude est une notion fondamentale pour les espaces métriques. Reprenons-la dans le contexte des espaces normés.

Définition 2.2.1 *Soit $(E, \|\cdot\|)$ un espace normé. On dit que $(E, \|\cdot\|)$ est un* espace de Banach *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.*

Exemples. L'espace $\mathbb{R}^d$ muni de la norme euclidienne est un espace complet. Plus généralement, tout sous-espace vectoriel de dimension finie d'un espace vectoriel normé sur $\mathbb{R}$ ou $\mathbb{C}$ est un espace de Banach. Soient E un espace de Banach et X un ensemble ; l'ensemble $\mathscr{B}(X,E)$ des fonctions bornées de X dans E muni de la norme

$$\|f\|_{\mathscr{B}(X,E)} \stackrel{\text{déf}}{=} \sup_{x \in X} \|f(x)\|_E$$

est un espace de Banach.

Soit $C^k(S^1)$ l'espace vectoriel des fonctions 2π-périodiques et k fois continûment dérivables à valeurs réelles ou complexes. Muni de la norme

$$\|f\|_k \stackrel{\text{déf}}{=} \sup_{\substack{x\in[0,2\pi] \\ j\leq k}} |f^{(j)}(x)|,$$

c'est un espace de Banach.

Remarque. Les espaces normés que nous rencontrerons seront presque toujours complets. Le théorème suivant (voir chapitre 5) est une des raisons essentielles du succès de la théorie de l'intégration de Lebesgue.

Théorème 2.2.1 *Soit* (X,μ) *un espace mesuré. Pour tout réel p appartenant à l'intervalle* $[1,+\infty]$, *l'espace* $(L^p(X,d\mu),\|\cdot\|_{L^p})$ *est un espace de Banach.*

La proposition suivante va nous fournir un nouvel exemple très important.

Proposition 2.2.1 *Soient E un espace normé et F un espace de Banach. Alors l'espace* $\mathscr{L}(E,F)$ *muni de la norme définie dans la proposition* 2.1.6 *est un espace de Banach.*

Démonstration. Soit une suite de Cauchy $(f_n)_{n\in\mathbb{N}}$ de $\mathscr{L}(E,F)$. D'après la définition de la norme sur l'espace $\mathscr{L}(E,F)$, on en déduit que, pour tout x de E, la suite $(f_n(x))_{n\in\mathbb{N}}$ est une suite de Cauchy de F. L'espace F étant supposé complet, il existe, pour tout x de E, un élément de F, que nous notons $f(x)$ tel que

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

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

$$\lim_{n\to\infty} f_n = f \quad \text{dans} \quad \mathscr{L}(E,F).$$

L'unicité de la limite assure très facilement que f est une application linéaire. Comme la suite $(f_n)_{n\in\mathbb{N}}$ est de Cauchy, elle est bornée (proposition 1.3.1). Il existe donc une constante C telle que

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

Par conséquent, pour tout x dans la boule unité de E

$$\sup_{n\in\mathbb{N}} \|f_n(x)\|_F \leq C,$$

d'où en passant à la limite en n,

$$\|f(x)\|_F \leq C.$$

Ceci étant vrai pour tout x de la boule unité de E, en prenant la borne supérieure, il vient

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

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

$$\sup_{\|x\|_E \leq 1} \|f_n(x) - f_m(x)\|_F < \varepsilon.$$

Par un passage à la limite analogue à celui effectué plus haut lorsque m tend vers l'infini, on obtient

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

ce qui achève la démonstration de la proposition. □

Théorème 2.2.2 *Soient E et F deux espaces vectoriels normés, G un sous-espace dense de E et f une application linéaire continue de G dans F. Supposons que F soit complet. Il existe alors un unique élément $\widetilde{f}$ de $\mathscr{L}(E,F)$ tel que $\widetilde{f}_{|G} = f$.*

Pour démontrer ce théorème, il suffit d'appliquer le théorème 1.3.1 et de vérifier que le prolongement ainsi obtenu est linéaire, ce qui est un exercice facile laissé au lecteur.

Proposition 2.2.2 *Soit E un espace de Banach et $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de E. Si la série à termes positifs $(\|x_n\|_E)_{n\in\mathbb{N}}$ est convergente, alors la suite des sommes partielles*

$$S_N = \sum_{n=0}^{N} x_n \quad \textit{est convergente.}$$

De plus, $\big\|\lim\limits_{N\to+\infty} S_N\big\|_E \leq \sum\limits_{n\in\mathbb{N}} \|x_n\|_E$.

Démonstration. En effet, on a pour tout couple d'entiers N et P,

$$\begin{aligned} \|S_{N+P} - S_N\|_E &= \Big\|\sum_{n=N+1}^{N+P} x_n\Big\|_E \\ &\leq \sum_{n=N+1}^{N+P} \|x_n\|_E, \end{aligned}$$

par l'inégalité triangulaire. L'hypothèse de sommabilité de la série des normes implique que, pour tout ε strictement positif, il existe un entier N_0 tel que

$$\forall N \geq N_0\,,\ \forall P\,,\ \sum_{n=N+1}^{N+P} \|x_n\|_E < \varepsilon.$$

La suite $(S_N)_{N\in\mathbb{N}}$ est donc de Cauchy, ce qui, vu que l'espace E est de Banach, montre qu'elle converge.

Par ailleurs, par l'inégalité triangulaire

$$\|S_N\|_E \leq \sum_{n=0}^{N} \|x_n\|_E.$$

Passant à la limite dans les deux termes quand $N \to +\infty$, on obtient la majoration de la norme. □

Commentaire. Dans un espace vectoriel normé tout comme dans le cas réel, on dit qu'une série est *convergente* si la suite des sommes partielles est convergente. Dans ce cas, on appelle *somme de la série*, et l'on note $\sum_{n=0}^{+\infty} x_n$ ou $\sum_{n\in\mathbb{N}} x_n$, la limite de cette suite des sommes partielles. Le théorème précédent nous dit que, dans un espace de Banach, si la série des normes est convergente dans $\mathbb{R}$, alors la série elle-même est convergente. On dit que la série est *normalement convergente*.

Théorème 2.2.3 *Si E est un espace normé tel que toute série normalement convergente est convergente. Alors E est un espace de Banach.*

Démonstration. Soit $(x_n)_{n\in\mathbb{N}}$ une suite de Cauchy de E, nous allons en extraire une suite convergente, ce qui, d'après la proposition 1.3.3, assurera le résultat.

On définit l'application ϕ de $\mathbb{N}$ dans $\mathbb{N}$ par $\phi(0) = 0$ et

$$\phi(n+1) \stackrel{\text{déf}}{=} \min\Big\{m \geq \phi(n)+1\,;\ \sup_{p\geq 0} \|x_m - x_{m+p}\|_E \leq \frac{1}{(n+2)^2}\Big\}.$$

Ainsi donc, on a, pour tout entier n,

$$\|x_{\phi(n+1)} - x_{\phi(n)}\|_E \leq \frac{1}{(n+1)^2}.$$

Posons $y_n \stackrel{\text{déf}}{=} x_{\phi(n+1)} - x_{\phi(n)}$. On a

$$\sum_{n\in\mathbb{N}} \|y_n\|_E \leq \sum_{n\in\mathbb{N}} \frac{1}{(n+1)^2} < \infty.$$

Donc par hypothèse, la suite

$$S_N = \sum_{n=0}^{N} y_n = x_{\phi(N)} - x_0$$

converge, d'où le théorème. □

Exercice 2.2.1 *Soient E et F deux espaces normés tels que F soit complet. On considère une suite bornée $(f_n)_{n\in\mathbb{N}}$ dans $\mathscr{L}(E,F)$. On suppose qu'il existe un sous-espace vectoriel G de E, dense dans E et une application linéaire f de G dans F telle que*

$$\forall x \in G,\ \lim_{n\to\infty} f_n(x) = f(x).$$

Démontrez que l'on peut prolonger f en un élément $\widetilde{f}$ de $\mathscr{L}(E,F)$ et que

$$\forall x \in E,\ \lim_{n\to\infty} f_n(x) = \widetilde{f}(x).$$

Exercice 2.2.2 *Soient E et F deux espaces vectoriels normés tels que E soit un espace de Banach. Si ι est une isométrie de E dans F, alors $\iota(E)$ est fermé dans F.*

Les deux théorèmes suivants sont des conséquences importantes du théorème 1.3.3 de Baire. Le premier est le théorème de Banach-Steinhaus ou principe de la borne uniforme.

Théorème 2.2.4 *Soient E et F deux espaces normés avec E complet, et $(f_\lambda)_{\lambda\in\Lambda}$ une famille quelconque d'élements de $\mathscr{L}(E,F)$. Supposons que, pour tout x de E, la famille $(f_\lambda(x))_{\lambda\in\Lambda}$ soit bornée dans F. Alors la famille $(f_\lambda)_{\lambda\in\Lambda}$ est bornée dans $\mathscr{L}(E,F)$.*

Démonstration. Considérons les ensembles $F_{\lambda,p}$ définis par

$$F_{\lambda,p} = \{x \in E \text{ tels que } \|f_\lambda(x)\|_F \leq p\}.$$

Ces ensembles $F_{\lambda,p}$ sont des fermés de E en tant qu'images réciproques de fermés par une application continue. Donc les ensembles F_p définis par

$$F_p \stackrel{\text{déf}}{=} \bigcap_{\lambda\in\Lambda} F_{\lambda,p}$$

sont des ensembles fermés car ce sont des intersections de fermés.

Soit x un élément quelconque de E. La famille $(f_\lambda(x))_{\lambda\in\Lambda}$ est par hypothèse bornée. Donc, il existe un entier p tel que x appartienne à F_p, c'est-à-dire que la réunion de tous les F_p est l'espace E tout entier.

Comme E est complet, d'après le corollaire 1.3.1 du théorème 1.3.3 de Baire, il existe un entier p_0 tel que $\overset{\circ}{F}_{p_0} \neq \emptyset$. Par conséquent, il existe un point x_0 et un réel strictement positif α tel que $B(x_0,\alpha)$ soit incluse dans F_{p_0}. On a donc, pour tout $\lambda \in \Lambda$,

$$\begin{aligned}
\sup_{\|x\|_E\leq\alpha} \|f_\lambda(x)\|_F &\leq \sup_{\|x\|_E\leq\alpha} \|f_\lambda(x+x_0)\|_F + \|f_\lambda(x_0)\|_F \\
&\leq \sup_{y\in B(x_0,\alpha)} \|f_\lambda(y)\|_F + \|f_\lambda(x_0)\|_F \\
&\leq 2p_0.
\end{aligned}$$

Ceci implique que

$$\|f_\lambda\|_{\mathscr{L}(E,F)} \leq \frac{2p_0}{\alpha}$$

et la famille $(f_\lambda)_{\lambda\in\Lambda}$ est donc bornée dans $\mathscr{L}(E,F)$. $\square$

Commentaires. Le théorème de Banach-Steinhaus est tout à fait remarquable dans la mesure où il montre que, pour une famille d'applications linéaires continues sur un espace de Banach, une borne ponctuelle implique une borne uniforme, ce qui est bien sûr violemment faux si l'on considère des familles d'applications continues quelconques.

Il admet même une version plus forte aux termes de laquelle on a l'alternative frappante suivante : soit il existe un ensemble dense de points $x \in E$ tels que $(f_\lambda(x))_{\lambda\in\Lambda}$ n'est pas bornée (cet ensemble est même un G_δ, c'est-à-dire une intersection dénombrable d'ouverts), soit la famille est uniformément bornée comme précédemment.

Corollaire 2.2.1 *Soient E et F deux espaces normés, E complet, et soit $(f_n)_{n\in\mathbb{N}}$ une suite de $\mathscr{L}(E,F)$. Supposons que,0 pour tout x de E, la suite $(f_n(x))_{n\in\mathbb{N}}$ soit convergente et désignons sa limite par $f(x)$. Alors la suite $(f_n)_{n\in\mathbb{N}}$ est une suite bornée de $\mathscr{L}(E,F)$, f appartient à $\mathscr{L}(E,F)$ et l'on a*

$$\|f\|_{\mathscr{L}(E,F)} \leq \liminf_{n\to\infty} \|f_n\|_{\mathscr{L}(E,F)}.$$

Démonstration. Tout d'abord, il est clair par unicité de la limite que f est une application linéaire.

Comme pour tout x de E, la suite $(f_n(x))_{n\in\mathbb{N}}$ est convergente par hypothèse, elle est bornée dans F. On peut donc lui appliquer le théorème de Banach-Steinhaus. Par passage à la limite dans la borne uniforme, on obtient alors que,

$$\forall x \in \bar{B}(0,1)\,,\ \|f(x)\|_F \leq \frac{2p_0}{\alpha},$$

et f est donc continue.

Démontrons maintenant la majoration de la norme de f. Soit x un élément de la boule unité de E. On peut alors écrire

$$\begin{aligned} \|f(x)\|_F &= \lim_{n\to\infty} \|f_n(x)\|_F \\ &= \liminf_{n\to\infty} \|f_n(x)\|_F \\ &\leq \liminf_{n\to\infty} \|f_n\|_{\mathscr{L}(E,F)} \quad (\text{car } \|x\|_E \leq 1). \end{aligned}$$

Prenant le sup sur x dans cette inégalité, on obtient le corollaire. $\square$

Remarque. L'inégalité majorant la norme de f peut être stricte comme le montre l'exemple suivant.

On considère l'espace $\ell^1(\mathbb{N})$ des suites à valeurs complexes telles que la série qui leur est associée est absolument convergente. C'est bien sûr un espace de Banach pour la norme $\sum_{n\in\mathbb{N}} |x_n|$. On considère la suite de formes linéaires $(e_n)_{n\in\mathbb{N}}$ définie par

$$e_n((x_p)_{p\in\mathbb{N}}) \stackrel{\text{déf}}{=} x_n.$$

C'est un exercice facile de démontrer que

$$\forall x \in \ell^1(\mathbb{N}), \ \lim_{n\to\infty} e_n(x) = 0.$$

Il est très facile d'observer que $\|e_n\|_{\mathscr{L}(\ell^1(\mathbb{N}),\mathbb{C})} = 1$.

Le théorème suivant, connu sous le nom de théorème de Banach, est une autre conséquence importante du théorème de Baire.

Théorème 2.2.5 *Soient E et F deux espaces de Banach et f un élément de $\mathscr{L}(E,F)$. Si f est bijective, alors $f^{-1} \in \mathscr{L}(F,E)$.*

En d'autres termes, une application linéaire continue bijective entre deux espaces de Banach est automatiquement un isomorphisme : il n'est nul besoin de vérifier que f^{-1} est continue. Le théorème de Banach est une conséquence facile du théorème de l'application ouverte suivant.

Théorème 2.2.6 *Soient E et F deux espaces de Banach et f un élément de $\mathscr{L}(E,F)$. Si f est surjective, alors il existe une constante $\alpha > 0$ telle que*

$$f(B_E(0,1)) \supset B_F(0,\alpha).$$

Démonstration. Posons $F_n = \mathrm{Adh}(f(B_E(0,n))$. Comme f est surjective, on a $F = \bigcup_{n\in\mathbb{N}} F_n$. Grâce au théorème de Baire, on en déduit qu'il existe n_0 tel que $\mathrm{Int}\, F_{n_0} \neq \emptyset$. Comme $F_1 = n_0^{-1} F_{n_0}$, on en déduit que $\mathrm{Int}\, F_1 \neq \emptyset$.

Il existe donc $y_0 \in \mathrm{Adh}(f(B_E(0,1)))$ et un nombre $\alpha > 0$ tels que $B_F(y_0, 4\alpha) \subset \mathrm{Adh}(f(B_E(0,1)))$. Par linéarité de f, on aussi $-y_0 \in \mathrm{Adh}(f(B_E(0,1)))$. Or pour tout $y \in B_F(0,4\alpha)$, nous pouvons écrire $y = y + y_0 - y_0$ et il existe deux suites x_k et x'_k dans $B_E(0,1)$ telles que

$$f(x_k) \longrightarrow y + y_0 \quad \text{et} \quad f(x'_k) \longrightarrow -y_0.$$

On en déduit que $x''_k = x_k + x'_k \in B(0,2)$ est telle que

$$f(x''_k) \longrightarrow y \quad \Longrightarrow \quad B_F(0,4\alpha) \subset \mathrm{Adh}(f(B_E(0,2))),$$

soit encore par linéarité de f et homogénéité des normes

$$\forall r > 0, \quad B_F(0,r) \subset \mathrm{Adh}\Big(f\Big(B_E\Big(0,\frac{r}{2\alpha}\Big)\Big)\Big). \tag{2.4}$$

Soit maintenant $y \in B_F(0,\alpha)$. Pour conclure, nous devons construire un $x \in B_E(0,1)$ tel que $y = f(x)$. D'après l'inclusion (2.4) avec $r = \alpha$, il existe x_1 tel que

$$x_1 \in B_E(0,1/2) \quad \text{et} \quad \|y - f(x_1)\|_F < \alpha/2.$$

Supposons construite une suite x_k, $k = 1,\dots,n$, telle que

$$x_k \in B_E(0,2^{-k}) \quad \text{et} \quad \Big\|y - f\Big(\sum_{k=1}^{n} x_k\Big)\Big\|_F < 2^{-n}\alpha.$$

Posant $y_n = y - f\big(\sum_{k=1}^n x_k\big)$, l'inclusion (2.4) avec $r = 2^{-n}\alpha$ nous fournit un x_{n+1} tel que

$$x_{n+1} \in B_E(0,2^{-(n+1)}) \quad \text{et} \quad \Big\|y_n - f(x_{n+1})\Big\|_F < 2^{-(n+1)}\alpha.$$

Nous voyons que la série $\sum_{k=1}^{+\infty} x_k$ ainsi construite est normalement convergente. Comme E est complet, elle converge vers un certain $x \in E$. De plus,

$$\|x\|_E \le \sum_{k=1}^{+\infty} \|x_k\|_E < \sum_{k=1}^{+\infty} 2^{-k} = 1,$$

et par construction $\|y - f(x)\|_F \le 0$, soit $y = f(x)$. □

Un corollaire important de ce qui précède est le théorème du graphe fermé.

Théorème 2.2.7 *Soient E et F deux espaces de Banach et f une application linéaire de E dans F dont le graphe est fermé dans $E \times F$. Alors $f \in \mathscr{L}(E,F)$.*

Démonstration. On munit $E \times F$ de la norme $\|(x,y)\|_{E\times F} = \|x\|_E + \|y\|_F$ qui en fait espace vectoriel normé et l'on a $(x_n,y_n) \to (x,y)$ dans $E \times F$ si et seulement si $x_n \to x$ dans E et $y_n \to y$ dans F. Le graphe de g est l'ensemble

$$G(f) = \{(x,y) \in E \times F; y = f(x)\}.$$

Considérons une nouvelle norme sur E

$$\|x\|_2 = \|x\|_E + \|f(x)\|_F,$$

et soit x_n une suite de Cauchy pour cette norme. Il s'ensuit que c'est une suite de Cauchy pour la norme de E et que $f(x_n)$ est une suite de Cauchy dans F. Par conséquent, il existe $(x,y) \in E \times F$ tel que $x_n \to x$ dans E et $f(x_n) \to y$ dans F. Or $(x_n, f(x_n)) \in G(f)$ et le graphe $G(f)$ est fermé dans $E \times F$. Donc $(x,y) \in G(f)$, c'est à dire que $y = f(x)$. Comme

$$\|x_n - x\|_2 = \|x_n - x\|_E + \|f(x_n - x)\|_F = \|x_n - x\|_E + \|f(x_n) - f(x)\|_F,$$

on voit que $(E, \|\cdot\|_2)$ est complet.

L'application identité de $(E, \|\cdot\|_2)$ dans $(E, \|\cdot\|_E)$ est bijective, continue puisque $\|x\|_E \leq \|x\|_2$ et les deux espaces sont de Banach. Par le théorème de Banach, l'identité est un isomorphisme. En particulier il existe une constante C telle que $\|x\|_2 \leq C\|x\|_E$, soit

$$\|f(x)\|_F \leq \|x\|_E + \|f(x)\|_F \leq C\|x\|_E,$$

et f est continue. □

Remarques. La seconde norme sur E est souvent appelée « norme du graphe ». Notons que si f est continue, alors son graphe est trivialement fermé, et qu'il existe de nombreuses applications non linéaires non continues dont le graphe est fermé, par exemple, $E = F = \mathbb{R}$, $f(x) = \frac{1}{x}$ si $x \neq 0$ et $f(0) = 0$.

2.3 Inverses et spectre dans $\mathscr{L}(E)$

L'un des résultats de base sur les éléments inversibles de $\mathscr{L}(E)$ est le suivant.

Théorème 2.3.1 *Soit E un espace de Banach. Les éléments de $\mathscr{L}(E)$ qui sont à distance strictement inférieure à 1 de Id sont inversibles dans $\mathscr{L}(E)$,*

$$\forall u \in B_{\mathscr{L}(E)}(\mathrm{Id}, 1), \exists!\, u^{-1} \in \mathscr{L}(E) \text{ tel que } u \circ u^{-1} = u^{-1} \circ u = \mathrm{Id}.$$

Démonstration. Posons

$$S_N = \sum_{n=0}^{N} (-1)^n (u - \mathrm{Id})^n.$$

D'après la proposition 2.2.2 et l'inégalité (2.3), la suite $(S_N)_{N\in\mathbb{N}}$ converge vers un élément g de $\mathscr{L}(E)$ dès que $\|u - \mathrm{Id}\|_{\mathscr{L}(E)} < 1$, ce qui implique aussi que $(u - \mathrm{Id})^{N+1} \to 0$. Par ailleurs, on a

$$\begin{aligned} S_N \circ u &= S_N \circ (\mathrm{Id} + u - \mathrm{Id}) \\ &= \mathrm{Id} - (-1)^{N+1}(u - \mathrm{Id})^{N+1} \\ &= u \circ S_N. \end{aligned}$$

En passant à la limite dans l'égalité ci-dessus, on en déduit que

$$g \circ u = u \circ g = \mathrm{Id},$$

ce qui conclut la démonstration de ce théorème. □

Corollaire 2.3.1 *L'ensemble $U(E)$ des éléments inversibles de $\mathscr{L}(E)$ est un ouvert et l'application* Inv *définie par*

$$\begin{cases} U(E) \rightarrow U(E) \\ \quad u \mapsto u^{-1} \end{cases}$$

est continue.

Démonstration. Soit u_0 un élément de $U(E)$. Considérons alors un élément u de $\mathscr{L}(E)$ tel que

$$\|u - u_0\|_{\mathscr{L}(E)} < \frac{1}{\|u_0^{-1}\|_{\mathscr{L}(E)}}.$$

D'après l'inégalité (2.3), il vient

$$\|u \circ u_0^{-1} - \mathrm{Id}\|_{\mathscr{L}(E)} < 1.$$

D'après le théorème précédent, $u \circ u_0^{-1}$ est inversible, donc u aussi. L'ensemble $U(E)$ est donc ouvert.

Montrons ensuite la continuité en Id. Comme $\mathscr{L}(E)$ est un espace métrique, il suffit de considérer une suite u_n telle que $\|u_n - \mathrm{Id}\,\|_{\mathscr{L}(E)} \rightarrow 0$. Grâce à la représentation en série entière vue précédemment, on peut supposer n assez grand pour que $u_n^{-1} = \sum_{k=0}^{+\infty}(-1)^k(u_n - \mathrm{Id})^k$. Par conséquent, $u_n^{-1} - \mathrm{Id} = \sum_{k=1}^{+\infty}(-1)^k(u_n - \mathrm{Id})^k$ et

$$\begin{aligned} \|u_n^{-1} - \mathrm{Id}\,\|_{\mathscr{L}(E)} &\leq \sum_{k=1}^{+\infty} \|(u_n - \mathrm{Id})^k\|_{\mathscr{L}(E)} \\ &\leq \sum_{k=1}^{+\infty} \|u_n - \mathrm{Id}\,\|_{\mathscr{L}(E)}^k = \frac{\|u_n - \mathrm{Id}\,\|_{\mathscr{L}(E)}}{1 - \|u_n - \mathrm{Id}\,\|_{\mathscr{L}(E)}} \rightarrow 0 \text{ quand } n \rightarrow +\infty. \end{aligned}$$

Pour achever la démonstration du corollaire, il reste à observer que, si $u_n \rightarrow u \in U(E)$, alors $g_n = u_n \circ u^{-1} \rightarrow \mathrm{Id}$, car $\|g_n - \mathrm{Id}\,\|_{\mathscr{L}(E)} \leq \|u_n - u\|_{\mathscr{L}(E)}\|u^{-1}\|_{\mathscr{L}(E)} \rightarrow 0$. Donc $g_n^{-1} = u \circ u_n^{-1} \rightarrow \mathrm{Id}$ d'après ce qui précède, et donc $u_n^{-1} = u^{-1} \circ g_n^{-1} \rightarrow u^{-1}$ par le même argument. □

Exercice 2.3.1 *Démontrer que l'application* Inv *est analytique sur $U(E)$.*

Exercice 2.3.2 *Soient E un espace de Banach et I un intervalle de $\mathbb{R}$, on désigne par $\mathscr{C}(I,E)$ l'ensemble des fonctions continues de I dans E. On se donne une fonction continue A de I dans $\mathscr{L}(E)$ et l'on fixe un réel t_0 dans l'intervalle I. On définit alors, pour tout réel λ, l'espace F_λ des fonctions f de $\mathscr{C}(I,E)$ telles que*

$$\|f\|_\lambda \stackrel{\text{déf}}{=} \sup_{t \in I} e^{-\lambda \int_{[t_0,t]} \|A(\tau)\|_{\mathscr{L}(E)} d\tau} \|f(t)\|_E < \infty.$$

1) Démontrer que, muni de la norme définie ci-dessus, F_λ est un espace de Banach.

2) Démontrer qu'il existe un réel strictement positif λ_0 tel que, pour tout $\lambda > \lambda_0$, l'application définie par

$$\mathscr{A}\left\{\begin{array}{l} F_\lambda \to F_\lambda \\ f \mapsto t \mapsto f(t) + \int_{t_0}^t A(\tau)f(\tau)\,d\tau \end{array}\right.$$

appartient à $U(F_\lambda)$.

3) Soit $f_0 \in \mathscr{C}(I,E)$, posons $f = \mathscr{A}^{-1} f_0$. Calculez $\mathscr{A}'(t) - \mathscr{A}(t)f(t)$. Quel théorème connu a-t-on ainsi redémontré ?

Définition 2.3.1 *Soient E un espace de Banach sur $\mathbb{K}$ et u un élément de $\mathscr{L}(E)$. On appelle* spectre *de u et l'on note* $\mathrm{Sp}(u)$ *l'ensemble des éléments λ de $\mathbb{K}$ tels que* $u - \lambda\,\mathrm{Id}$ *ne soit pas inversible. On dit que $\lambda \in \mathbb{K}$ est une valeur propre de u si* $\ker(u - \lambda\,\mathrm{Id}) \neq \{0\}$.

Il est clair que toute valeur propre appartient au spectre. Par contre, le spectre n'est pas toujours réduit aux seules valeurs propres.

En dimension finie, il n'existe qu'une seule façon de ne pas être inversible pour une application linéaire. En effet, soit u une application linéaire de E dans E, avec $\dim E < +\infty$. L'algèbre linéaire élémentaire dit que

$$u \text{ bijective} \Leftrightarrow u \text{ surjective} \Leftrightarrow u \text{ injective.}$$

L'application réciproque u^{-1} est linéaire donc continue puisque nous sommes en dimension finie. Le spectre est donc dans ce cas égal à l'ensemble des valeurs propres.

Il en va tout autrement en dimension infinie. Par exemple, soit $E = \ell^\infty(\mathbb{N})$ muni de la norme

$$\|(x_n)_{n\in\mathbb{N}}\|_{\ell^\infty(\mathbb{N})} \stackrel{\text{déf}}{=} \sup_{n\in\mathbb{N}} |x_n|.$$

C'est un espace de Banach. Considérons le décalage à droite u_1 défini par

$$\begin{array}{rl} u_1 \colon E & \to E \\ (x_n)_{n\in\mathbb{N}} & \mapsto (y_n)_{n\in\mathbb{N}},\ y_0 = 0, y_n = x_{n-1} \quad \text{si } n \geq 1. \end{array}$$

L'application u_1 est une isométrie, c'est-à-dire que $\|u_1(x)\| = \|x\|$, donc injective, mais elle n'est pas surjective. Dans ce cas, $0 \in \mathrm{Sp}(u_1)$ mais 0 n'est pas valeur propre.

On a aussi le décalage à gauche u_2 défini par

$$\begin{array}{rl} u_2 \colon E & \to E \\ (x_n)_{n\in\mathbb{N}} & \mapsto (y_n)_{n\in\mathbb{N}},\ y_n = x_{n+1}. \end{array}$$

L'application u_2 est surjective, mais pas injective. Mais cette fois-ci 0 est une valeur propre.

Proposition 2.3.1 *Soient E un espace de Banach sur $\mathbb{K}$ et u un élément de $\mathscr{L}(E)$. Le spectre de u est un compact inclus dans la boule fermée de centre 0 et de rayon $\|u\|_{\mathscr{L}(E)}$.*

Démonstration. En effet, l'application L_u définie par

$$\begin{aligned} L_u \colon \mathbb{K} &\to \mathscr{L}(E) \\ \lambda &\mapsto u - \lambda \operatorname{Id} \end{aligned}$$

est une application continue de $\mathbb{K}$ dans $\mathscr{L}(E)$. Or, par définition de $\operatorname{Sp}(u)$, le spectre est l'image réciproque par l'application continue L_u du fermé $\mathscr{L}(E) \setminus U(E)$, donc le spectre est fermé. De plus, si $|\lambda| > \|u\|_{\mathscr{L}(E)}$, écrivons

$$u - \lambda \operatorname{Id} = -\lambda \left(\operatorname{Id} - \frac{1}{\lambda} u \right).$$

Comme $\operatorname{Id} - \lambda^{-1} u \in B(\operatorname{Id}, 1)$, elle est inversible, et donc $u - \lambda \operatorname{Id}$ aussi, c'est-à-dire $\lambda \notin \operatorname{Sp}(u)$. □

Théorème 2.3.2 *Soient E un espace de Banach sur $\mathbb{C}$ et u un élément de $\mathscr{L}(E)$. Le spectre de u n'est pas vide.*

Démonstration. Supposons que le spectre de u soit vide. On considère alors l'application L_u^{-1} définie par

$$\begin{aligned} L_u^{-1} \colon \mathbb{C} &\to \mathscr{L}(E) \\ \lambda &\mapsto (u - \lambda \operatorname{Id})^{-1}. \end{aligned}$$

D'après l'exercice 2.3.1, l'application L_u^{-1} est analytique de $\mathbb{C}$ dans $\mathscr{L}(E)$. En particulier, pour tout $x \in E$ et toute forme linéaire continue v sur E (voir chapitre suivant), l'application $\lambda \mapsto v(L_u^{-1}(x))$ est analytique de $\mathbb{C}$ dans $\mathbb{C}$. De plus, on a

$$L_u^{-1}(\lambda) = \frac{1}{\lambda} \left(\frac{1}{\lambda} u - \operatorname{Id} \right)^{-1}.$$

Il en résulte que

$$\lim_{|\lambda| \to +\infty} L_u^{-1}(\lambda) = 0.$$

Donc l'application $\lambda \mapsto v(L_u^{-1}(x))$ est une fonction entière bornée sur $\mathbb{C}$, donc constante d'après le théorème de Liouville, donc nulle d'après la limite ci-dessus. On en déduit que $L_u^{-1} = 0$, ce qui est absurde. Il en résulte que le spectre de u n'est pas vide. □

Attention ! Le corps de base est très important dans les questions de spectre et ceci n'a rien à avoir avec la dimension infinie. Pour s'en convaincre, il suffit d'observer qu'une rotation de $\mathbb{R}^2$ d'angle non congru à 0 modulo π a un spectre vide sur $\mathbb{R}$ (ses valeurs propres sont complexes).

2.4 Opérateurs compacts

On parle souvent d'opérateur pour désigner une application linéaire — continue ou non — sur un espace vectoriel de dimension infinie.

Définition 2.4.1 *Soient E et F deux espaces de Banach. Un élément u de $\mathscr{L}(E,F)$ est dit* compact *si et seulement si pour toute partie bornée A, $u(A)$ est relativement compacte.*

Exercice 2.4.1 *Démontrez que si l'image de u est de dimension finie, alors u est compact.*

Il existe une caractérisation des opérateurs compacts en termes de suites.

Théorème 2.4.1 *Soient E et F deux espaces de Banach. Un élément u de $\mathscr{L}(E,F)$ est compact si et seulement si pour toute suite bornée $(x_n)_{n\in\mathbb{N}}$ d'éléments de E, la suite $(u(x_n))_{n\in\mathbb{N}}$ possède une valeur d'adhérence.*

Démonstration. Supposons que u soit compact et considérons une suite bornée quelconque $(x_n)_{n\in\mathbb{N}}$ d'éléments de E. Posant $M \stackrel{\text{déf}}{=} \sup_n \|x_n\|_E$, la suite $(u(x_n))_{n\in\mathbb{N}}$ est incluse dans $u(\bar{B}_E(0,M))$, ensemble relativement compact dans F. On peut donc extraire de la suite $(u(x_n))_{n\in\mathbb{N}}$ une sous-suite convergente.

Réciproquement, soient A une partie bornée de E et $(y_n)_{n\in\mathbb{N}}$ une suite d'éléments de $u(A)$. Il existe donc une suite $x_n \in A$ telle que $y_n = u(x_n)$. Cette suite étant bornée, on en déduit que la suite $(y_n)_{n\in\mathbb{N}}$ admet une valeur d'adhérence dans F, ce qui montre que $u(A)$ est relativement compacte. □

Proposition 2.4.1 *Soient E et F deux espaces de Banach. Un élément u de $\mathscr{L}(E,F)$ est compact si et seulement si $u(B_E(0,1))$ est relativement compacte.*

Démonstration. En effet, si A est une partie bornée de E, il existe $\lambda > 0$ tel que $A \subset \lambda B_E(0,1)$. Mais alors l'adhérence de $u(A)$ est incluse dans $\lambda\,\mathrm{Adh}(u(B_E(0,1)))$ qui est compacte, d'où la proposition. □

Théorème 2.4.2 *Soient E, F et G trois espaces de Banach.*

- *L'ensemble* $\mathscr{K}(E,F)$ *des éléments de* $\mathscr{L}(E,F)$ *qui sont compacts est un sous-espace vectoriel fermé de* $\mathscr{L}(E,F)$*.*
- *Soient* $u \in \mathscr{L}(E,F)$ *et* $v \in \mathscr{L}(F,G)$ *; alors l'opérateur* $v \circ u$ *est compact dès que u ou v l'est.*

Démonstration. Montrons d'abord le premier point. Soient u et v deux opérateurs compacts de E dans F et λ un scalaire. Considérons une suite bornée $(x_n)_{n\in\mathbb{N}}$ de E. Comme u est compact, on peut en extraire une sous-suite $(x_{n_p})_{p\in\mathbb{N}}$ telle que $(u(x_{n_p}))_{p\in\mathbb{N}}$ est convergente. Comme v est compact, on peut de cette sous-suite extraire une seconde sous-suite $(x_{n_{p_q}})_{q\in\mathbb{N}}$ telle que $(v(x_{n_{p_q}}))_{q\in\mathbb{N}}$ est convergente. La suite $((\lambda u+v)(x_{n_{p_q}}))_{q\in\mathbb{N}}$ est donc convergente. On en déduit que l'opérateur $\lambda u+v$ est compact.

Démontrons maintenant que $\mathscr{K}(E,F)$ est fermé. Soient u un élément de $\overline{\mathscr{K}(E,F)}$ et A une partie bornée de E. D'après le théorème 1.4.6, il suffit de démontrer que, pour tout réel strictement positif ε, on peut recouvrir $u(A)$ par un nombre fini de boules de rayon ε centrées en des points de $u(A)$. Or, il existe un opérateur compact v tel que

$$\|u-v\|_{\mathscr{L}(E,F)} < \frac{\varepsilon}{4M(A)} \quad \text{avec} \quad M(A) \stackrel{\text{déf}}{=} \sup_{x\in A} \|x\|.$$

Comme v est compact, il existe un ensemble fini $(x_j)_{1\leq j\leq N}$ d'éléments de A tel que

$$v(A) \subset \bigcup_{j=1}^{N} B_F(v(x_j), \varepsilon/2).$$

Ainsi donc, pour tout y de $u(A)$, il existe $x \in A$ tel que $y = u(x)$ et donc un indice j tel que

$$\begin{aligned} \|y-u(x_j)\| &\leq \|u(x)-v(x)\| + \|v(x)-v(x_j)\| + \|v(x_j)-u(x_j)\| \\ &< \varepsilon, \end{aligned}$$

ce qui implique que $u(A) \subset \bigcup_{j=1}^{N} B_F(u(x_j), \varepsilon)$.

Passons au deuxième point. Supposons v compact et considérons une suite bornée $(x_n)_{n\in\mathbb{N}}$ de E. Comme u est continue, la suite $(u(x_n)_{n\in\mathbb{N}})$ est bornée dans F et, comme v est supposée compacte, la suite $((v\circ u)(x_n))_{n\in\mathbb{N}}$ admet une valeur d'adhérence dans G. Donc $v \circ u$ est compact.

Supposons maintenant u compact et considérons une suite bornée $(x_n)_{n\in\mathbb{N}}$ de E. On peut en extraire une sous-suite $(x_{n_p})_{p\in\mathbb{N}}$ telle que $(u(x_{n_p}))_{p\in\mathbb{N}}$ est convergente dans F. Comme v est continue, la suite $(v \circ u(x_{n_p}))_{p\in\mathbb{N}}$ est convergente dans G. Donc $v \circ u$ est compact. $\square$

Exercice 2.4.2 *Soit* (X,d) *un espace métrique compact. On considère sur l'espace* $C^\alpha(X,\mathbb{K})$ *des fonctions hölderiennes d'exposant* $\alpha \in]0,1]$ *de* X *dans* $\mathbb{K}$*, la norme*

$$\|f\|_\alpha = \sup_{x\in X} |f(x)| + \sup_{\substack{(x,y)\in X^2\\ x\neq y}} \frac{|f(x)-f(y)|}{d(x,y)^\alpha}.$$

1) *Démontrez que cette norme munit* $C^\alpha(X,\mathbb{K})$ *d'une structure d'espace de Banach.*

2) *Démontrez que, pour tout* α*, l'inclusion de* $C^\alpha(X,\mathbb{K})$ *dans l'espace de Banach des fonctions continues de* X *dans* $\mathbb{K}$ *est compacte.*

Chapitre 3

Dualité dans les espaces de Banach

3.1 Le concept de dualité

Le concept de dualité, c'est-à-dire la considération de l'espace vectoriel des formes linéaires sur une espace vectoriel donné, est certainement bien connu du lecteur dans le cadre des espaces vectoriels de dimension finie. En termes matriciels, il se traduit simplement par la transposition des matrices de coordonnées, les matrices colonnes devenant des matrices lignes. L'espace des formes linéaires sur un espace de dimension d, le dual de cet espace, est également un espace vectoriel de dimension d. Un espace vectoriel de dimension finie est donc toujours isomorphe à son dual, quoique de façon non canonique, et il n'y a guère de mystère.

Nous étudierons ici le concept en dimension infinie. Des phénomènes fondamentalement nouveaux apparaissent. Tout d'abord, il existe des formes linéaires non continues. Nous nous empressons de les ignorer. Plus dépaysant, l'ensemble des formes linéaires continues sur un espace vectoriel normé E pourra apparaître comme très différent de l'espace E.

Définition 3.1.1 *On appelle* dual algébrique *d'un $\mathbb{K}$-espace vectoriel E et l'on note E^* l'ensemble des formes linéaires définies sur E.*

On appelle dual topologique *d'un $\mathbb{K}$-espace vectoriel normé $(E, \|\cdot\|)$ et l'on note E' l'espace vectoriel des formes linéaires continues sur E. On munit $E' = \mathscr{L}(E, \mathbb{K})$ de sa norme d'espace d'applications linéaires continues*

$$\|\ell\|_{E'} = \sup_{\|x\|_E \leq 1} |\ell(x)|.$$

Remarquons que, comme le montre l'exemple de la forme linéaire δ_0' définie page 61, les deux espaces ci-dessus définis ne coïncident pas en dimension infinie. En fait, le dual algébrique E^* d'un espace vectoriel normé de dimension infinie est un espace vectoriel « énorme » dont on ne sait en général pas grand-chose et qui se révèle peu utile dans la pratique. Le dual topologique est par contre extrêmement utile dans bien des situations. Dans la suite, lorsque nous dirons simplement dual, il s'agira toujours du dual topologique.

Notons d'abord que, d'après la proposition 2.2.1, le dual E' de n'importe quel espace vectoriel normé E, est un espace de Banach.

L'étude et la description de cet espace E' donne parfois lieu à des problèmes redoutables. Toutefois, dans de nombreux cas, nous verrons qu'il est possible de représenter très concrètement le dual d'un espace vectoriel normé.

Il existe une forme bilinéaire naturelle sur le produit d'un espace et de son dual.

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

$$\begin{aligned} \langle \cdot, \cdot \rangle : E' \times E &\to \mathbb{K} \\ (\ell, x) &\mapsto \langle \ell, x \rangle \stackrel{\text{déf}}{=} \ell(x). \end{aligned}$$

est une forme bilinéaire continue. On l'appelle le crochet de dualité.

Démonstration. Il est bien clair que nous avons affaire à une forme bilinéaire sur $E' \times E$. De plus, par définition de la norme sur E', on a

$$|\langle \ell, x \rangle| = |\ell(x)| \leq \|\ell\|_{E'} \|x\|_E.$$

Le théorème 2.1.4 de caractérisation des applications multilinéaires continues assure alors le résultat. □

Le théorème de Hahn-Banach qui suit, et que nous admettrons, est un théorème fondamental de prolongement de forme linéaire continue définie sur un sous-espace vectoriel. Il nous assure de l'existence de suffisamment de formes linéaires continues sur un espace normé. En particulier, le dual d'un espace vectoriel normé n'est pas réduit au vecteur nul (cela n'a rien d'évident *a priori*) car nous pouvons prolonger une forme linéaire non nulle définie, par exemple, sur une droite.

Théorème 3.1.1 *Soient E un espace vectoriel normé, F un sous espace vectoriel de E et ℓ une forme linéaire continue sur F muni de la topologie induite, c'est-à-dire telle que*

$$\exists\, C \geq 0 \text{ telle que } \forall x \in F, |\ell(x)| \leq C\|x\|_E.$$

Il existe alors un élément $\tilde{\ell}$ de E' tel que $\tilde{\ell}_{|F} = \ell$ et tel que $\|\tilde{\ell}\|_{E'} = \|\ell\|_{F'}$.

On utilisera en particulier le corollaire suivant.

Corollaire 3.1.1 *Soit x un élément non nul d'un espace normé E. Il existe une forme linéaire continue ℓ telle que l'on ait*

$$\langle \ell, x \rangle = \|x\|_E \quad \text{et} \quad \|\ell\|_{E'} = 1.$$

Démonstration. On pose $F = \mathbb{K}x$ et pour tout $y = \lambda x \in F$, $\ell(y) = \lambda \|x\|_E$. On a donc $\ell(x) = 1 \times \|x\|_E$ et pour tout $y \in F$, $|\ell(y)| = |\lambda| \|x\|_E = \|\lambda x\|_E = \|y\|_E$, d'où $\|\ell\|_{F'} = 1$. □

L'opération de passage au dual associe un espace de Banach non trivial à tout espace vectoriel normé. On peut naturellement réitérer le processus.

Définition 3.1.2 *Soit E un espace vectoriel normé et E' son dual. On appelle* bidual *de E, et l'on note E'', le dual de E'.*

Tout aussi naturellement le bidual de E est un espace de Banach. Il contient E au sens suivant.

Proposition 3.1.2 *Soit ι l'application de E dans E'' qui à $x \in E$ associe la forme linéaire continue sur E' définie par $\iota(x)(\ell) = \ell(x)$. Alors $\iota \in \mathscr{L}(E, E'')$ avec $\|\iota(x)\|_{E''} = \|x\|_E$. C'est donc une injection isométrique de E dans E''.*

Démonstration. Il est clair que ι est linéaire de E dans E''. De plus, elle est continue car

$$|\iota(x)(\ell)| \leq \|\ell\|_{E'} \|x\|_E,$$

ce qui montre aussi au passage que $\|\iota(x)\|_{E''} \leq \|x\|_E$. Pour démontrer l'égalité, on utilise le corollaire 3.1.1. En effet, pour $x \neq 0$ de E, il existe $\ell \in E'$ telle que $|\iota(x)(\ell)| = \|x\|_E$ et $\|\ell\|_{E'} = 1$. □

On voit déjà qu'en général il n'y a aucune raison pour que l'application ι soit bijective, car même si E n'est pas complet, E'' l'est toujours. La situation est en fait nettement plus subtile que cela car ce n'est pas seulement une question de complétude, nous en verrons des exemples plus loin. Ceci nous amène à poser la définition suivante.

Définition 3.1.3 *On dit qu'un espace vectoriel normé est* réflexif *si l'application ι est surjective.*

On dit souvent de ce cas, avec un léger abus de langage sans grand danger en général, que $E'' = E$ à la place de $\iota(E) = E''$. Un espace vectoriel normé est donc réflexif s'il est canoniquement (*i.e.* via ι) isométrique à son bidual. Il faut faire attention à ce dernier point. Aussi surprenant que cela paraisse, il existe en effet un exemple d'un espace de Banach qui est isométrique à son bidual, mais qui n'est pas réflexif ! L'isométrie en question n'est pas l'injection canonique ι.

Proposition 3.1.3 *Si E est un espace vectoriel normé réflexif, alors c'est un espace de Banach.*

Démonstration. D'après la remarque ci-dessus, un espace réflexif est isométrique à son bidual qui est un espace de Banach, puisque c'est le dual de l'espace normé E', d'où la proposition. □

Dans la même veine, mentionnons,

Proposition 3.1.4 *Si E est un espace de Banach, alors $\imath(E)$ est un sous-espace fermé de E''.*

En d'autres termes un espace de Banach s'identifie naturellement à un sous-espace fermé de son bidual.

Clairement, tout espace normé de dimension finie est réflexif. Il existe des espaces de Banach, et non des moindres, qui ne sont pas réflexifs. Notons enfin qu'il n'existe en général pas d'application linéaire continue naturelle entre E et E' qui permette d'identifier l'un à une partie de l'autre.

3.2 Identification d'un dual à l'aide d'un espace de Banach

Cette section est très importante pour la suite du cours, notamment en vue de la théorie des distributions. Il arrive fréquemment que l'on dise que tel espace « est le dual » de tel autre. Il s'agit de comprendre précisément ce que cet abus de langage signifie.

Proposition 3.2.1 *Soient E et F deux espaces de Banach et B une forme bilinéaire continue sur l'espace $F \times E$. L'application δ_B qui à $y \in F$ associe la forme linéaire $\delta_B(y)$ sur E définie par*

$$\begin{aligned} \delta_B(y) \colon E &\to \mathbb{K} \\ x &\mapsto \langle \delta_B(y), x\rangle = B(y,x) \end{aligned} \tag{3.1}$$

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

Démonstration. C'est très simple. Tout d'abord, il est clair que tout le monde est bien linéaire comme annoncé. Comme la forme bilinéaire B est continue, il existe une constante C telle que

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

En particulier, $\delta_B(y) \in E'$ et 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' à F *si et seulement si l'application δ_B définie par* (3.1) *est un isomorphisme de F sur E'.*

En clair, cela signifie que B identifie E' à F s'il existe une constante $C > 0$ telle que, pour toute forme linéaire continue ℓ sur E, il existe un élément y de F et un seul tel que,

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

Remarques. D'après le théorème de Banach 2.2.5, il suffit de démontrer que l'application δ_B est bijective. Un exemple déjà connu : si E est réflexif, alors la forme bilinéaire sur $E \times E'$, $B(x,\ell) = \langle \ell, x\rangle$, identifie E'' à E. L'identification est encore plus agréable si c'est une isométrie, c'est-à-dire si $\|\delta_B(y)\|_{E'} = \|y\|_F$, ce qui est le cas dans l'exemple précédent.

La définition 3.2.1 est utilisée pour identifier, si possible, le dual d'un espace normé, qui est un espace défini de façon « abstraite », avec un autre espace normé « concret ». Donnons-en un premier exemple assez typique tout en restant élémentaire.

Théorème 3.2.1 *Soient p un réel supérieur ou égal à 1, $E = \ell^p(\mathbb{N})$ et $F = \ell^{p'}(\mathbb{N})$ avec $p' = p/(p-1)$ si $p > 1$, $p' = +\infty$ si $p = 1$. On considère la forme bilinéaire B suivante*

$$\begin{aligned} B\colon F \times E &\to \mathbb{K} \\ (y,x) &\mapsto B(y,x) \stackrel{déf}{=} \sum_{n\in\mathbb{N}} x_n y_n. \end{aligned}$$

Alors, la forme bilinéaire B identifie isométriquement $(\ell^p(\mathbb{N}))'$ à $\ell^{p'}(\mathbb{N})$.

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 et $\|\delta_B(y)\|_{(\ell^p(\mathbb{N}))'} = \|y\|_{\ell^{p'}(\mathbb{N})}$ (en prenant, si p' est fini, la suite $x_n = \bar{y}_n|y_n|^{p'-2}$ si $y_n \neq 0$, 0 sinon, et si $p' = +\infty$, on choisit un entier n tel que $|y_n| \geq \|y\|_{\ell^\infty(\mathbb{N})} - \varepsilon$ et l'on considère la suite de $\ell^1(\mathbb{N})$ qui vaut toujours 0, sauf $x_n = \bar{y}_n/|y_n|$). L'application δ_B est isométrique, donc injective.

Démontrons qu'elle est surjective. Soit φ une forme linéaire continue sur $\ell^p(\mathbb{N})$. Cherchons un élément y de $\ell^{p'}(\mathbb{N})$ telle que

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

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 l'élément $y = (y_n)_{n\in\mathbb{N}}$ satisfait

$$y_n = \langle \varphi, e_n \rangle \text{ pour tout } n \in \mathbb{N}. \tag{3.4}$$

Il s'agit maintenant de démontrer dans un premier temps que la suite $y = (y_n)_{n\in\mathbb{N}}$ définie par (3.4) est bien un élément de $\ell^{p'}(\mathbb{N})$, puis que $\delta_B(y) = \varphi$.

Traitons d'abord le cas $p = 1$. Pour tout entier n, on a

$$\begin{aligned} |y_n| &\le \|\varphi\|_{(\ell^1(\mathbb{N}))'}\|e_n\|_{\ell^1(\mathbb{N})} \\ &= \|\varphi\|_{(\ell^1(\mathbb{N}))'}. \end{aligned}$$

Donc y appartient bien à $\ell^\infty(\mathbb{N})$ avec $\|y\|_{\ell^\infty(\mathbb{N})} \le \|\varphi\|_{(\ell^1(\mathbb{N}))'}$.

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

$$\forall\, r \in \mathbb{N}, \quad x_{r,n} = \frac{\overline{y}_n}{|y_n|}|y_n|^{\frac{1}{p-1}} \quad \text{si } n \le r \text{ et } y_n \neq 0, \quad x_{r,n} = 0 \quad \text{sinon.}$$

Ils sont bien dans $\ell^p(\mathbb{N})$, puisqu'ils n'ont qu'un nombre fini de termes non nuls (on ne sait encore rien sur y !). Remarquons que par construction

$$|x_{r,n}|^p = |y_n|^{\frac{p}{p-1}} = |y_n|^{p'} \text{ et } x_{r,n}y_n = |y_n|^{\frac{1}{p-1}+1} = |y_n|^{p'} \text{ pour tout } n \le r.$$

Par définition de y_n et linéarité de φ, on voit que

$$\varphi(x_r) = \sum_{n=0}^{r} x_{r,n}\varphi(e_n) = \sum_{n=0}^{r} x_{r,n}y_n = \sum_{n=0}^{r} |y_n|^{p'}.$$

D'un autre côté, la forme linéaire φ étant continue sur $\ell^p(\mathbb{N})$, on a également

$$\begin{aligned} \sum_{n=0}^{r} |y_n|^{p'} = |\varphi(x_r)| &\le \|\varphi\|_{(\ell^p(\mathbb{N}))'}\|x_r\|_{\ell^p(\mathbb{N})} \\ &= \|\varphi\|_{(\ell^p(\mathbb{N}))'}\Big(\sum_{n=0}^{r} |y_n|^{p'}\Big)^{\frac{1}{p}}. \end{aligned} \tag{3.5}$$

Divisant (3.5) par le dernier terme du membre de droite (s'il est non nul, sinon il n'y a rien à démontrer), on en déduit que

$$\Big(\sum_{n=0}^{r} |y_n|^{p'}\Big)^{1-\frac{1}{p}} \le \|\varphi\|_{(\ell^p(\mathbb{N}))'}.$$

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

Soit enfin $x \in \ell^p(\mathbb{N})$ quelconque et notons x_k la suite tronquée au rang k (*i.e.* $x_{k,n} = x_n$ si $n \leq k$, $x_{k,n} = 0$ si $n > k$). Il est clair que $x_k \to x$ dans $\ell^p(\mathbb{N})$ quand $k \to +\infty$. De plus, comme plus haut, $\delta_B(y)(x_k) = \langle \varphi, x_k \rangle$. La continuité de φ et de $\delta_B(y)$ implique que $\delta_B(y)(x) = \langle \varphi, x \rangle = \varphi(x)$. □

Commentaire. L'identification du dual de ℓ^p avec $\ell^{p'}$ par l'intermédiaire de la forme bilinéaire $B(y,x) = \sum_{n\in\mathbb{N}} x_n y_n$ étant une isométrie, on a souvent tendance à dire par abus de langage que le dual de ℓ^p est purement et simplement $\ell^{p'}$ et que le crochet de dualité est simplement la forme bilinéaire B elle-même. Cet abus de langage ne présente aucun danger, autant donc s'y habituer tout de suite.

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

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

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 linéaire B identifie $c_0(\mathbb{N})'$ à l'espace $\ell^1(\mathbb{N})$.

Remarques. Les espaces $\ell^p(\mathbb{N})$ pour $1 < p < +\infty$ sont donc des espaces réflexifs, puisque $(p')' = p$ (plus précisément, il faudrait vérifier que l'injection canonique ι est réalisée à l'aide de l'application δ_B, ce qui n'est pas difficile).

Dans l'introduction 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 l'espace $\ell^\infty(\mathbb{N})$ n'est pas séparable. Or on sait que l'espace $\ell^1(\mathbb{N})$ l'est, donc $\ell^1(\mathbb{N})$ et $(\ell^1(\mathbb{N}))'$ ne sont pas isomorphes. Notons enfin que nous n'avons rien dit sur le dual de $\ell^\infty(\mathbb{N})$, c'est-à-dire le bidual de $\ell^1(\mathbb{N})$ et que nous n'en dirons prudemment rien de plus pour le moment.

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

$$\ell^{2,s}(\mathbb{N}) \stackrel{déf}{=} \{u = (u_n)_{n\in\mathbb{N}}; ((1+n^2)^{\frac{s}{2}} u_n)_{n\in\mathbb{N}} \in \ell^2(\mathbb{N})\}.$$

et on le munit de la norme

$$\|u\|_{\ell^{2,s}(\mathbb{N})} \stackrel{déf}{=} \Big(\sum_{n\in\mathbb{N}} (1+n^2)^s |u_n|^2\Big)^{\frac{1}{2}}.$$

Soient B_1 et B_2 les deux formes bilinéaires définies ci-après par

$$\begin{aligned} B_1 \colon \ell^{2,s}(\mathbb{N}) \times \ell^{2,s}(\mathbb{N}) &\to \mathbb{K} \\ (y,x) &\mapsto B_1(y,x) \stackrel{\text{déf}}{=} \sum_{n\in\mathbb{N}} (1+n^2)^s x_n y_n. \end{aligned}$$

$$\begin{aligned} B_2 \colon \ell^{2,-s}(\mathbb{N}) \times \ell^{2,s}(\mathbb{N}) &\to \mathbb{K} \\ (y,x) &\mapsto B_2(y,x) \stackrel{\text{déf}}{=} \sum_{n\in\mathbb{N}} x_n y_n. \end{aligned}$$

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 $\ell^{2,-s}(\mathbb{N})$.

Nous admettons momentanément le théorème suivant, qui sera démontré en détail au chapitre 5.

Théorème 3.2.2 *Soient p un réel supérieur à 1 et B la forme bilinéaire définie par*

$$\begin{aligned} B \colon L^{p'} \times L^p &\to \mathbb{K} \\ (g,f) &\mapsto \int_\Omega f(x)g(x)d\mu(x). \end{aligned}$$

Alors la forme bilinéaire B identifie le dual de $L^p(\Omega,d\mu)$ à $L^{p'}(\Omega,d\mu)$.

3.3 Une notion affaiblie de 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 faiblement-étoile (faiblement-$*$) vers ℓ, et l'on note $\ell_n \stackrel{*}{\rightharpoonup} \ell$, si*

$$\forall x \in E, \quad \langle \ell_n, x\rangle \to \langle \ell, x\rangle.$$

Mentionnons qu'il existe une topologie sur E' dont les suites convergentes sont exactement celles qui convergent faiblement-$*$ au sens précédent. On l'appelle la topologie faible-$*$ sur E'. C'est une topologie qui n'est pas métrisable en dimension infinie. Comme son étude en tant que topologie n'est pas cruciale pour la suite du cours, nous ne l'introduirons pas ici. La notion pratique qui va nous servir est uniquement celle de convergence faible-$*$.

La convergence faible-$*$ apparaît comme la convergence simple des formes linéaires vues comme applications de E dans K. L'autre notion que nous connaissons déjà sur E', la convergence au sens de la norme duale, est la convergence uniforme sur les boules de E. Il est évident que si $\ell_n \to \ell$ dans E', au sens de la norme (on dit aussi au sens fort), alors $\ell_n \overset{*}{\rightharpoonup} \ell$. La réciproque n'est pas vraie en général.

Nous allons maintenant étudier quelques propriétés de la convergence faible-$*$ qui sont importantes pour les applications.

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, pour tout x de E, la suite $(\langle \ell_n, x\rangle)_{n\in\mathbb{N}}$ converge et désignons par $\langle \ell, x\rangle$ sa limite. 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*

$$\|\ell\|_{E'} \leq \liminf_{n\to\infty} \|\ell_n\|_{E'}.$$

Ceci n'est qu'une traduction dans le cadre des formes linéaires du théorème de Banach-Steinhaus 2.2.1. Remarquons que l'on ne suppose pas $\ell \in E'$, ce qui explique que l'hypothèse n'est pas $\ell_n \overset{*}{\rightharpoonup} \ell$, même si l'on se retrouve *in fine* dans ce cadre. L'inégalité peut naturellement être stricte (chercher un exemple).

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. Si $\ell_n \overset{*}{\rightharpoonup} \ell$ et $\|x_n - x\|_E \to 0$ alors $\langle \ell_n, x_n\rangle \to \langle \ell, x\rangle$.*

Pour démontrer cela, commençons par décomposer le crochet de dualité comme suit

$$\langle \ell_n, x_n\rangle = \langle \ell_n, x_n - x\rangle + \langle \ell_n, x\rangle.$$

Par la convergence faible-$*$, on a $\langle \ell_n, x\rangle \to \langle \ell, x\rangle$. D'après le théorème 3.3.1, la suite $(\ell_n)_{n\in\mathbb{N}}$ est bornée dans E'. Donc, il existe un réel M tel que l'on ait $|\langle \ell_n, x_n - x\rangle| \leq \|\ell_n\|_{E'}\|x_n - x\|_E \leq M\|x_n - x\|_E$, ce qui implique que $\langle \ell_n, x_n - x\rangle \to 0$, d'où la proposition. □

Le théorème de compacité faible-$*$ de la boule qui suit est fondamental car il permet de trouver une valeur d'adhérence à une suite à partir d'une simple borne, même en dimension infinie. Rappelons que, dans ce cas, la boule unité de E' n'est pas compacte pour la topologie métrique définie par la norme. Pour retrouver cette compacité, il faut affaiblir la topologie (c'est-à-dire autoriser moins d'ouverts). La convergence obtenue est naturellement plus faible que la convergence en norme.

Théorème 3.3.2 *Soit E un espace séparable, on considère une suite de formes linéaires continues $(\ell_n)_{n\in\mathbb{N}}$ bornée dans E'. Il existe un élément ℓ de E' et une fonction strictement croissante ψ de $\mathbb{N}$ dans $\mathbb{N}$ tels que l'on ait*

$$\ell_{\psi(n)} \overset{*}{\rightharpoonup} \ell.$$

Démonstration. Soit $(x_p)_{p\in\mathbb{N}}$ un ensemble dénombrable dense dans E. Nous allons utiliser le procédé diagonal déjà vu lors de la démonstration du théorème d'Ascoli.

On procède par récurrence. 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 φ_0 de $\mathbb{N}$ dans $\mathbb{N}$ telle que l'on ait

$$\lim_{n\to\infty}\langle \ell_{\varphi_0(n)}, x_0\rangle = \lambda_0.$$

Supposons construits des fonctions $(\varphi_j)_{0\leq j\leq m}$ strictement croissantes de $\mathbb{N}$ dans $\mathbb{N}$ et des scalaires $(\lambda_j)_{1\leq j\leq m}$ tels que, pour tout $j\leq m$, on ait

$$\lim_{n\to\infty}\langle \ell_{\varphi_0\circ\cdots\circ\varphi_j(n)}, x_j\rangle = \lambda_j.$$

La suite $(\langle \ell_{\varphi_0\circ\cdots\circ\varphi_m(n)}, x_{m+1}\rangle)_{n\in\mathbb{N}}$ 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

$$\lim_{n\to\infty}\langle \ell_{\varphi_0\circ\cdots\circ\varphi_{m+1}(n)}, x_{m+1}\rangle = \lambda_{m+1}.$$

Posons $\psi(n) = \varphi_0\circ\cdots\circ\varphi_n(n)$. Comme nous l'avons déjà vu page 43 lors de la démonstration du théorème d'Ascoli, l'application ψ est strictement croissante. De plus, on a $\langle \ell_{\psi(n)}, x_j\rangle \to \lambda_j$ quand $n\to+\infty$ pour tout j, puisqu'il s'agit pour tout j d'une suite extraite (à partir du rang $n=j$) d'une suite convergente.

Considérons l'application ℓ définie de $\{x_p,\ p\in\mathbb{N}\}$ dans $\mathbb{K}$ par

$$\ell(x_p) = \lambda_p.$$

Comme $|\langle \ell_{\psi(n)}, x_p\rangle - \langle \ell_{\psi(n)}, x_q\rangle| \leq M\|x_p - x_q\|_E$ avec $M \overset{\text{déf}}{=} \sup_{n\in\mathbb{N}}\|\ell_n\|_{E'}$, passant à la limite on obtient

$$|\ell(x_p) - \ell(x_q)| \leq M\|x_p - x_q\|_E.$$

D'après le théorème 1.3.1 de prolongement des applications uniformément continues sur une partie dense, ℓ se prolonge par continuité de façon unique à E. De plus, par construction, ℓ est M-lipschitzienne.

On a pour tout x de E,

$$\ell(x) - \ell_{\psi(n)}(x) = \ell(x) - \ell(x_p) + \ell(x_p) - \langle \ell_{\psi(n)}, x_p\rangle + \langle \ell_{\psi(n)}, x_p - x\rangle,$$

(on ne sait pas encore que ℓ est linéaire). D'où, comme $\ell(x_p) = \lambda_p$, en passant aux valeurs absolues ou modules,

$$|\ell(x) - \ell_{\psi(n)}(x)| \leq 2M\|x_p - x\|_E + |\lambda_p - \langle \ell_{\psi(n)}, x_p\rangle|.$$

Soit ε un réel strictement positif arbitraire. La suite $(x_p)_{p\in\mathbb{N}}$ étant dense, il existe un entier p tel que $\|x_p - x\|_E < \frac{\varepsilon}{4M}$. De plus, une fois p choisi, comme

$\langle \ell_{\psi(n)}, x_p \rangle \to \lambda_p$, il existe un entier n_0 tel que pour tout $n \geq n_0$, $|\lambda_p - \langle \ell_{\psi(n)}, x_p \rangle| < \frac{\varepsilon}{2}$.

Par conséquent, pour tout $n \geq n_0$,

$$|\ell(x) - \ell_{\psi(n)}(x)| < \varepsilon,$$

c'est-à-dire

$$\forall x \in E, \quad \langle \ell_{\psi(n)}, x \rangle \to \ell(x),$$

et le théorème est démontré. □

Remarques. Il est toujours vrai que les boules fermées du dual d'un espace de Banach E sont compactes pour la topologie faible-$*$, sans hypothèse de séparabilité sur E, c'est le théorème de Banach-Alaoglu. Par contre, si on ne fait pas cette hypothèse de séparabilité, on ne peut pas dire que toute suite bornée de E' admet une sous-suite faiblement-$*$ convergente. En effet, ces compacts ne sont pas en général métrisables. On peut montrer que la restriction de la topologie faible-$*$ à toute boule du dual d'un espace E séparable est une topologie métrisable, ce qui montre bien que le théorème 3.3.2 est un cas particulier du théorème de compacité plus général que l'on vient de mentionner. Tout ceci n'empêche bien sûr pas la topologie faible-$*$ de ne pas être métrisable sur tout le dual en dimension infinie.

3.4 La notion d'application linéaire transposée

Cette notion est basée sur la proposition suivante.

Proposition 3.4.1 *Soient E et F deux espaces vectoriels normés et f un élément de l'espace $\mathscr{L}(E,F)$. L'application ${}^t f$ de F' dans E' définie par*

$$\forall \ell \in F', \forall x \in E, \quad {}^t f(\ell)(x) = \langle \ell, f(x) \rangle$$

est une application linéaire continue.

Définition 3.4.1 *L'application linéaire définie ci-dessus s'appelle l'application linéaire* transposée.

Démonstration. Il est évident que ${}^t f(\ell)$ est une forme linéaire sur E. Pour les questions de continuité, on écrit

$$|\langle \ell, f(x) \rangle| \leq \|\ell\|_{F'} \|f(x)\|_F \leq \|\ell\|_{F'} \|f\|_{\mathscr{L}(E,F)} \|x\|_E,$$

d'où ${}^t f(\ell) \in E'$. Il est également évident que l'application ${}^t f$ est linéaire de F' dans E'. De plus,

$$\|{}^t f(\ell)\|_{E'} = \sup_{\|x\|_E \leq 1} |\langle \ell, f(x) \rangle| \leq \|f\|_{\mathscr{L}(E,F)} \|\ell\|_{F'},$$

d'après l'inégalité précédente, d'où la proposition. □

Il découle de ce qui précède que l'application $f \mapsto {}^t f$ est linéaire continue de $\mathscr{L}(E,F)$ dans $\mathscr{L}(F',E')$ avec $\|{}^t f\|_{\mathscr{L}(F',E')} \leq \|f\|_{\mathscr{L}(E,F)}$.

On peut bien entendu itérer l'opération, mais il faut faire attention à ce que ${}^t({}^t f) \in \mathscr{L}(E'',F'')$ et donc on n'a pas en général ${}^t({}^t f) = f$.

Proposition 3.4.2 *Soient E, F et G trois espaces vectoriels normés, $f \in \mathscr{L}(E,F)$ et $g \in \mathscr{L}(F,G)$. On a*

$${}^t(g \circ f) = {}^t f \circ {}^t g$$

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

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

Démonstration. Ces relations sont purement algébriques. En effet, pour toute forme linéaire ℓ sur G et pour tout x de E, on a

$$\begin{aligned} \langle {}^t f \circ {}^t g(\ell), x\rangle &= \langle {}^t g(\ell), f(x)\rangle \\ &= \langle \ell, g \circ f(x)\rangle \\ &= \langle {}^t(g \circ f)(\ell), x\rangle. \end{aligned}$$

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

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

Donc ${}^t u$ est un élément inversible de $\mathscr{L}(E')$ et $({}^t u)^{-1} = {}^t(u^{-1})$. □

La proposition suivante donne un avant-goût de ce que nous verrons souvent au sujet des distributions.

Proposition 3.4.3 *Soient E et F des espaces vectoriels normés, on considère un élément f de $\mathscr{L}(E,F)$, une suite $(\ell_n)_{n\in\mathbb{N}}$ de F' et un élément ℓ de F'. On a*

$$\ell_n \overset{*}{\rightharpoonup} \ell \textit{ dans } F' \Longrightarrow {}^t f(\ell_n) \overset{*}{\rightharpoonup} {}^t f(\ell) \textit{ dans } E'.$$

Démonstration. En effet, on a pour tout x appartenant à E,

$$\langle {}^t f(\ell_n), x\rangle = \langle \ell_n, f(x)\rangle.$$

Par hypothèse, on a

$$\langle \ell_n, f(x)\rangle \to \langle \ell, f(x)\rangle.$$

Mais, par définition de la transposée, $\langle \ell, f(x)\rangle = \langle {}^t f(\ell), x\rangle$. Ainsi donc

$$\forall x \in E, \quad \langle {}^t f(\ell_n), x\rangle \to \langle {}^t f(\ell), x\rangle,$$

c'est-à-dire exactement ${}^t f(\ell_n) \overset{*}{\rightharpoonup} {}^t f(\ell)$. □

Chapitre 4

Espaces de Hilbert réels et complexes

4.1 Produit scalaire et orthogonalité

Dans toute la suite, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, $\bar{\lambda}$ est le complexe conjugué de λ. Rappelons tout d'abord la définition d'un produit scalaire.

Définition 4.1.1 *On distingue les cas réel et complexe, bien qu'ils soient très semblables.*

i) Cas réel. Soit E un $\mathbb{R}$-espace vectoriel. Une forme bilinéaire f sur E est appelée produit scalaire *sur E si elle est symétrique, définie, positive, c'est-à-dire vérifie*

$$\begin{aligned} \forall (x,y) \in E^2,\ f(x,y) = f(y,x) \quad &(\textit{symétrique}),\\ \forall x \in E,\ f(x,x) \in \mathbb{R}_+ \quad &(\textit{positive}),\\ f(x,x) = 0 \Leftrightarrow x = 0 \quad &(\textit{définie}). \end{aligned}$$

ii) Cas complexe. Soient E et F deux espaces vectoriels sur $\mathbb{C}$. 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 $\mathbb{C}$, on ait*

$$\ell(\lambda x + y) = \bar{\lambda}\ell(x) + \ell(y).$$

Une application f de $E \times E$ dans $\mathbb{C}$ est une forme sesquilinéaire sur *E si elle est linéaire par rapport à la première variable et antilinéaire par rapport à la seconde. Autrement dit*

$$\forall \lambda \in \mathbb{C},\ \forall (x,y,z) \in E^3 \quad \begin{aligned} f(\lambda x + y, z) &= \lambda f(x,y) + f(y,z)\\ f(z, \lambda x + y) &= \bar{\lambda} f(z,x) + f(z,y). \end{aligned}$$

Une forme sesquilinéaire sur E est appelée produit scalaire *si elle est hermitienne, définie, positive, c'est-à-dire vérifie*

$$\forall (x,y) \in E^2,\ f(x,y) = \overline{f(y,x)} \quad (\textit{hermitienne}),$$
$$\forall x \in E,\ f(x,x) \in \mathbb{R}_+ \quad (\textit{positive}),$$
$$f(x,x) = 0 \Leftrightarrow x = 0 \quad (\textit{définie}).$$

Dans les deux cas, un espace vectoriel sur $\mathbb{K}$ *muni d'un produit scalaire est appelé un* espace préhilbertien.

Remarques. On note très souvent un produit scalaire $(\cdot|\cdot)$ ou $(\cdot,\cdot)$ et $\|x\|^2 = (x|x)$. Un espace préhilbertien réel de dimension finie est appelé *espace euclidien* alors qu'un espace préhilbertien complexe de dimension finie est appelé *espace hermitien.* Notons que l'on rencontre aussi dans la littérature la convention inverse pour la définition de la sesquilinéarité, à savoir antilinéarité par rapport à la première variable et linéarité par rapport à la deuxième variable. Les deux conventions ont leurs avantages et leurs inconvénients, la théorie sous-jacente étant fondamentalement la même.
Convention. Par abus de langage et pour ne pas dédoubler les démonstrations qui sont identiques dans le cas réel et le cas complexe, on conviendra qu'une application antilinéaire entre deux espaces vectoriels sur $\mathbb{K}$ est une application linéaire ordinaire si $\mathbb{K} = \mathbb{R}$ et antilinéaire au sens ci-dessus si $\mathbb{K} = \mathbb{C}$. On prendra la convention analogue pour les formes sesquilinéaires/bilinéaires. Ces conventions sont cohérentes avec le fait que $\bar{\lambda} = \lambda$ si $\lambda \in \mathbb{R}$. On notera $\mathfrak{Re}\,\lambda$ et $\mathfrak{Im}\,\lambda$ les parties réelle et imaginaire de λ.

Proposition 4.1.1 *Soient E un espace vectoriel sur* $\mathbb{K}$ *et* $(\cdot|\cdot)$ *un produit scalaire sur E. L'application* $x \mapsto \|x\| = (x|x)^{\frac{1}{2}}$ *est une norme sur E et l'on a pour tout* $(x,y) \in E^2$

$$|(x|y)| \leq \|x\|\,\|y\| \quad (\textit{inégalité de Cauchy-Schwarz}),$$

et

$$\frac{1}{2}(\|x\|^2 + \|y\|^2) = \left\|\frac{x+y}{2}\right\|^2 + \left\|\frac{x-y}{2}\right\|^2 \quad (\textit{identité du parallélogramme}).$$

Démonstration. Ceci est très classique. Commençons par l'inégalité de Cauchy-Schwarz. Si $x = 0$, il n'y a rien à démontrer. Supposons $x \neq 0$. Pour tout $t \in \mathbb{K}$,

$$\|tx + y\|^2 \geq 0.$$

En développant le premier membre, on obtient

$$\forall t \in \mathbb{K}, \quad |t|^2\|x\|^2 + 2\,\mathfrak{Re}(t(x|y)) + \|y\|^2 \geq 0.$$

L'inégalité de Cauchy-Schwarz découle immédiatement du choix $t = -\frac{\overline{(x|y)}}{\|x\|^2} \in \mathbb{K}$.

Pour démontrer que $x \mapsto (x|x)^{\frac{1}{2}}$ est une norme, il suffit d'écrire à l'aide de l'inégalité de Cauchy-Schwarz que

$$\begin{aligned}\|x+y\|^2 = (x+y|x+y) &= (x|x) + 2\,\mathfrak{Re}(x|y) + (y|y)\\ &\leq \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2\\ &= (\|x\| + \|y\|)^2\end{aligned}$$

d'où l'inégalité triangulaire, et que

$$\|\lambda x\|^2 = (\lambda x|\lambda x) = \lambda\overline{\lambda}\|x\|^2 = |\lambda|^2\|x\|^2,$$

d'où la positivité homogène. Pour démontrer l'identité du parallélogramme, il suffit d'observer que

$$\begin{aligned}2\left\|\frac{x+y}{2}\right\|^2 + 2\left\|\frac{x-y}{2}\right\|^2 &= \frac{1}{2}\|x+y\|^2 + \frac{1}{2}\|x-y\|^2\\ &= \frac{1}{2}(\|x\|^2 + 2\,\mathfrak{Re}(x|y) + \|y\|^2 + \|x\|^2 - 2\,\mathfrak{Re}(x|y) + \|y\|^2)\\ &= \|x\|^2 + \|y\|^2.\end{aligned}$$

où l'on a utilisé de façon répétée que $(x|y) + (y|x) = 2\,\mathfrak{Re}(x|y)$. □

Exercice 4.1.1 *Soit $(E, \|\cdot\|)$ un espace vectoriel normé sur $\mathbb{K}$. Supposons que*

$$\forall (x,y) \in E^2,\ \frac{1}{2}(\|x\|^2 + \|y\|^2) = \left\|\frac{x+y}{2}\right\|^2 + \left\|\frac{x-y}{2}\right\|^2.$$

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 révèle extrêmement fécond dans les espaces de dimension infinie, notamment dans les espaces de fonctions.

Définition 4.1.2 *Soient E un espace préhilbertien et x et y deux éléments de E. On dit que x et y sont* orthogonaux (*et l'on note $x \perp y$) si* $(x|y) = 0$.

Rappelons pour mémoire le vénérable théorème de Pythagore :

$$x \perp y \Rightarrow \|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

Dans le cas d'un espace préhilbertien réel, l'implication a également lieu en sens inverse.

Proposition 4.1.2 *Soit E un espace préhilbertien et A une partie quelconque de E. On pose*

$$A^{\perp} \stackrel{\text{déf}}{=} \{x \in E; \forall a \in A,\ x \perp a\}.$$

La partie $A^{\perp}$ est un sous-espace vectoriel fermé de E appelé orthogonal de A.

Démonstration. L'application linéaire $\mathscr{L}_a$ définie par $x \mapsto (x|a)$ est une application linéaire continue de E dans $\mathbb{K}$ par Cauchy-Schwarz. Le noyau de $\mathscr{L}_a$ est donc un sous-espace vectoriel fermé de E. Il est clair que

$$A^{\perp} = \bigcap_{a \in A} \ker \mathscr{L}_a.$$

Donc $A^{\perp}$ est un sous-espace vectoriel fermé en tant qu'intersection de sous-espaces vectoriels fermés. □

Exercice 4.1.2 *Soit E un espace préhilbertien et A une partie de E. Démontrez que $\bar{A}^{\perp} = A^{\perp} = \mathrm{vect}(A)^{\perp}$.*

La définition suivante est parfaitement naturelle au vu de tout ce qui précède.

Définition 4.1.3 *Soit H un espace préhilbertien. On dit que H est un* espace de Hilbert *s'il est complet pour la norme associée à son produit scalaire.*

Les espaces de Hilbert (ou hilbertiens) sont donc des cas particuliers d'espaces de Banach. La 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. Notons à ce propos que les espaces euclidiens et hermitiens sont automatiquement hilbertiens.

Exemples fondamentaux. Les espaces $\ell^2(\mathbb{N})$, muni du produit scalaire $(x|y) = \sum_{n\in\mathbb{N}} x_n \bar{y}_n$, et $L^2(X, d\mu)$, muni du produit scalaire $(f|g) = \int_X f(x)\bar{g}(x)\,d\mu$, sont des espaces de Hilbert (en fait le premier est un cas particulier du second).

4.2 Propriétés des espaces de Hilbert

Le théorème de Pythagore se généralise à une infinité de vecteurs de la façon 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 H. Alors la série $\sum_n x_n$ converge si et seulement si la série $\sum_n \|x_n\|^2$ converge et l'on a*

$$\Big\|\sum_{n\in\mathbb{N}} x_n\Big\|^2 = \sum_{n\in\mathbb{N}} \|x_n\|^2.$$

Démonstration. Introduisons les sommes partielles $S_q = \sum_{p=0}^{q} x_p$. Supposons d'abord que la série $\sum_n x_n$ soit convergente dans H et notons $S = \lim S_q$ sa somme. Par orthogonalité deux à deux des vecteurs, le théorème de Pythagore assure que

$$\sum_{p=0}^{q} \|x_p\|^2 = \|S_q\|^2. \tag{4.1}$$

Le membre de droite de l'inégalité ci-dessus converge vers $\|S\|^2$, donc est majoré indépendemment de q. Par conséquent, la série à termes positifs $\sum_n \|x_n\|^2$ est majorée donc converge.

Réciproquement, si la série $\sum_n \|x_n\|^2$ converge, alors on a

$$\begin{aligned} \|S_{q+q'} - S_q\|^2 &= \sum_{q+1}^{q+q'} \|x_p\|^2 \\ &\leq \sum_{q+1}^{\infty} \|x_p\|^2 \to 0 \text{ quand } q \to +\infty. \end{aligned}$$

La suite des sommes partielles est donc une suite de Cauchy et comme H est complet, elle est convergente.

Enfin, en passant à la limite dans l'égalité (4.1) ci-dessus, 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 H telle que*

$$\sum_{n=0}^{\infty} \|x_n\|^2 < \infty.$$

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 reposent sur le théorème de projection orthogonale sur les convexes fermés.

Théorème 4.2.2 *Soit C une partie convexe fermée d'un espace de Hilbert H. Pour tout point x de H, il existe un unique point de C, noté $p_C(x)$ et appelé projection de x sur C, tel que*

$$\|x - p_C(x)\| = \inf_{y\in C} \|x - y\|.$$

Démonstration. Démontrons tout d'abord l'unicité. Soient deux points y_1 et y_2 de C réalisant le minimum. D'après l'identité du parallélogramme, on a

$$\frac{1}{2}\|y_1-y_2\|^2 = \|x-y_1\|^2+\|x-y_2\|^2-2\left\|x-\frac{y_1+y_2}{2}\right\|^2.$$

Or, C est convexe, donc $(y_1+y_2)/2$ appartient à C. Par définition de y_1 et de y_2, on en déduit que $\|y_1-y_2\|^2\leq 0$, donc que $y_1=y_2$, ce qui assure l'unicité.

Démontrons maintenant l'existence. Par définition de la borne inférieure, il existe une suite $(y_n)_{n\in\mathbb{N}}$ d'éléments de C telle

$$\lim_{n\to\infty}\|x-y_n\| = d \stackrel{\text{déf}}{=} \inf_{y\in C}\|x-y\|.$$

On utilise à nouveau l'identité du parallélogramme pour dire que, pour tout couple d'entiers (n,m), on a

$$\|y_n-y_m\|^2 = 2\|x-y_n\|^2+2\|x-y_m\|^2-4\left\|x-\frac{y_n+y_m}{2}\right\|^2.$$

La partie C étant convexe, le milieu de y_n et de y_m appartient aussi à C. Ainsi, l'on a

$$\|y_n-y_m\|^2 \leq 2\|x-y_n\|^2+2\|x-y_m\|^2-4d^2.$$

Vu la définition de la suite $(y_n)_{n\in\mathbb{N}}$, ceci implique que cette suite est de Cauchy. L'espace H étant de Hilbert, il est complet. Comme la partie C est fermée, elle est complète et donc la suite $(y_n)_{n\in\mathbb{N}}$ converge dans C, ce qui conclut la démonstration de ce théorème. □

Exercice 4.2.2 *Soient C une partie convexe fermée et x un point d'un espace de Hilbert H. Démontrez que $p_C(x)$ est l'unique point de C tel que, pour tout point y de C, on ait*

$$\mathfrak{Re}(x-p_C(x)|p_C(x)-y)\geq 0.$$

Exercice 4.2.3 *Soit C une partie convexe fermée d'un espace de Hilbert H. Démontrez que l'application p_C est lipschitzienne de rapport* 1.

Presque toutes les propriétés importantes 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 H. L'application p_F est linéaire continue de H dans F et l'on a*

$$H=F\oplus F^\perp \quad \textit{et} \quad x=p_F(x)+p_{F^\perp}(x).$$

Démonstration. Considérons un élément x quelconque de H. Pour tout point y de F, et pour tout réel λ, on a

$$\|x - p_F(x) + \lambda y\|^2 = \lambda^2 \|y\|^2 + 2\lambda\, \mathfrak{Re}(x - p_F(x)|y) + \|x - p_F(x)\|^2.$$

Par définition de la projection, on a

$$\|x - p_F(x) + \lambda y\|^2 \geq \|x - p_F(x)\|^2.$$

Ceci impose donc que, pour tout y appartenant à F, on ait

$$\mathfrak{Re}(x - p_F(x)|y) = 0.$$

Si $\mathbb{K} = \mathbb{R}$, il n'y a rien de plus à dire. Si $\mathbb{K} = \mathbb{C}$, il suffit de changer y en iy pour obtenir que

$$x - p_F(x) \in F^\perp.$$

Nous venons donc de démontrer que $H = F + F^\perp$, ainsi que le fait que p_F est linéaire (par unicité de la projection). Mais, si x appartient à $F \cap F^\perp$, alors $(x|x) = 0$, donc $x = 0$. Le corollaire est ainsi démontré. $\square$

Dans un espace de Hilbert, on dispose d'un critère commode pour caractériser les parties totales.

Corollaire 4.2.2 *Soit A une partie quelconque d'un espace de Hilbert H. Cette partie A est totale si et seulement si $A^\perp = \{0\}$.*

Démonstration. On utilise l'exercice 4.1.2 qui affirme 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 H, ce qui signifie exactement que la partie A est totale.

Réciproquement, si la partie A est totale, alors par définition $\overline{\mathrm{vect}(A)}$ est égal à H, d'où $A^\perp = \{0\}$. $\square$

Dans le cas où A est lui-même un sous-espace vectoriel, on voit donc qu'une caractérisation de sa densité est $A^\perp = \{0\}$.

Exercice 4.2.4 *Soit A une partie quelconque d'un espace de Hilbert H. Démontrez que*

$$(A^\perp)^\perp = \overline{\mathrm{vect}(A)}.$$

Définition 4.2.1 *Soit H un espace de Hilbert séparable de dimension infinie. On appelle* base hilbertienne *ou* base orthonormale *de H toute suite $(e_n)_{n\in\mathbb{N}}$ d'éléments de H qui est totale et telle que*

$$(e_n|e_m) = \delta_{nm}. \tag{4.2}$$

Théorème 4.2.3 *Dans un espace de Hilbert séparable de dimension infinie H, il existe des bases hilbertiennes.*

Démonstration. Considérons une partie dénombrable totale libre $(a_n)_{n\in\mathbb{N}}$. Nous allons utiliser le procécé d'orthogonalisation de Schmidt. Rappelons ce procédé. On pose
$$e_0 = \frac{a_0}{\|a_0\|}\cdot$$
Supposons définis des vecteurs $(e_j)_{0\leq j\leq n}$ vérifiant la relation (4.2) et tels que
$$\mathrm{vect}\{a_0,\cdots,a_n\} = \mathrm{vect}\{e_0,\cdots,e_n\}.$$
Posons
$$e_{n+1} = \frac{f_{n+1}}{\|f_{n+1}\|} \quad \text{avec} \quad f_{n+1} = a_{n+1} - \sum_{j=0}^{n}(a_{n+1}|e_j)e_j.$$
La vérification de l'hypothèse de récurrence est immédiate. Le théorème est ainsi démontré. □

Théorème 4.2.4 *Soit H un espace de Hilbert séparable et $(e_n)_{n\in\mathbb{N}}$ une base hilbertienne de H. L'application $\mathscr{I}$ définie par*
$$\begin{aligned}\mathscr{I} : H &\to \ell^2(\mathbb{N})\\ x &\mapsto ((x|e_n))_{n\in\mathbb{N}}\end{aligned}$$
est une bijection linéaire isométrique. En particulier, on a les identités de Bessel *et* de 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. Montrons d'abord que l'application $\mathscr{I}$ envoie bien H dans $\ell^2(\mathbb{N})$. Pour cela, posons
$$x_N \stackrel{\text{déf}}{=} \sum_{n=0}^{N}(x|e_n)e_n.$$
On a
$$(x|x_N) = \sum_{n=0}^{N}\overline{(x|e_n)}(x|e_n) = \sum_{n=0}^{N}|(x|e_n)|^2 = \|x_N\|^2.$$
Appliquant l'inégalité de Cauchy-Schwarz au membre de gauche, on en déduit que $\|x_N\|^2 \leq \|x\|\,\|x_N\|$. Par conséquent, soit $\|x_N\| = 0$ et il n'y a rien à démontrer, soit on divise par $\|x_N\|$ puis on élève au carré, d'où
$$\sum_{n=0}^{N}|(x|e_n)|^2 \leq \|x\|^2.$$

La série à termes positifs $\sum_{n\in\mathbb{N}} |(x|e_n)|^2$ est majorée, donc convergente, c'est-à-dire que $\mathscr{I}(x)$ appartient à $\ell^2(\mathbb{N})$.

Par le théorème 4.2.1, la série $\sum_{n\in\mathbb{N}}(x|e_n)e_n$ converge dans H. Soit $\tilde{x}$ sa somme. D'après le calcul précédent, $(x-x_N|e_n)=0$ pour tout $n\leq N$. Passant à la limite quand $N\to+\infty$, on en déduit que $x-\tilde{x}$ appartient à l'orthogonal de la base hilbertienne, lequel est réduit au vecteur nul puisqu'il s'agit d'une partie totale. Ceci nous donne l'identité de Bessel, puis celle de Parseval en réappliquant le théorème 4.2.1.

L'injectivité de l'application $\mathscr{I}$ et le fait qu'il s'agit d'une isométrie résultent simplement de l'identité de Parseval. Sa surjectivité découle du théorème 4.2.1. Le théorème 4.2.3 est ainsi démontré. $\square$

Remarques. Une conséquence facile du théorème précédent est que

$$\forall (x,y)\in H^2, \quad (x|y)=\sum_{n\in\mathbb{N}}(x|e_n)\overline{(y|e_n)}.$$

Le théorème 4.2.3 peut être interprété comme signifiant que, d'un certain point de vue, il n'y a qu'un seul espace de Hilbert séparable de dimension infinie, c'est $\ell^2(\mathbb{N})$. Par contre, cet espace a plusieurs réalisations concrètes comme $L^2(\Omega,dx)$ par exemple. On peut également trouver des espaces de Hilbert non séparables, mais leur utilité dans les applications est pour le moins limitée. Le concept de base hilbertienne s'adapte également au cas non séparable.

Il est important de ne pas confondre base hilbertienne et base algébrique. La décomposition d'un vecteur sur une base algébrique ne fait intervenir qu'un nombre fini de vecteurs de cette dernière. L'existence de bases algébriques pour les espaces de dimension infinie dépendant de l'axiome du choix, il est en général impossible d'exhiber explicitement une base algébrique. Par contre, dans le cas d'une base hilbertienne, la décomposition d'un vecteur utilise une infinité de termes, et il s'agit d'une égalité avec une série convergente. On exhibe également très explicitement des bases hilbertiennes dans les espaces hilbertiens courants.

On peut donner comme application de l'existence de base hilbertienne le théorème suivant.

Théorème 4.2.5 *Soit H un espace de Hilbert. Un opérateur linéaire continu est compact si et seulement si il est limite dans $\mathscr{L}(H)$ d'une suite d'opérateurs de rang fini.*

Démonstration. D'après le théorème 2.4.2, on sait que toute limite de suite d'opérateurs compacts, donc en particulier de rang fini, est un opérateur compact.

Réciproquement, soit A un opérateur compact. L'image $A(B)$ par A de la boule unité B de H est une partie totale de $\operatorname{im}A$. Comme l'adhérence de $A(B)$ est compacte, l'espace $H_0 \stackrel{\text{déf}}{=} \overline{\text{im}A}$ est un espace de Hilbert séparable. Il admet donc une base hilbertienne. Désignons par p_n la projection orthogonale sur l'espace engendré par les n premiers vecteurs de cette base. Nous allons démontrer que la suite des opérateurs $A_n \stackrel{\text{déf}}{=} p_n \circ A$ (qui sont évidemment de rang fini) converge vers A.

Soit ε un réel strictement positif. L'opérateur A étant compact, il existe une famille finie $(x_j)_{1\leq j\leq N_\varepsilon}$ telle que

$$\forall x \in B\,,\ \exists\, j \in \{1,\cdots,N_\varepsilon\} \,/\, \|Ax - Ax_j\| < \frac{\varepsilon}{3}\cdot$$

Donc, pour tout x appartenant à la boule unité, on a

$$\|A_n x - Ax\| \leq \sup_{1\leq j\leq N_\varepsilon} \|A_n x_j - Ax_j\| + \frac{2\varepsilon}{3}\cdot$$

Mais, pour tout indice j, $\lim_{n\to\infty} p_n(Ax_j) = Ax_j$. Donc, il existe un entier n_0 tel que

$$n \geq n_0 \Rightarrow \sup_{1\leq j\leq N_\varepsilon} \|A_n x_j - Ax_j\| < \frac{\varepsilon}{3}\cdot$$

Ainsi, il existe un entier n_0 tel que

$$n \geq n_0 \Rightarrow \sup_{\|x\|\leq 1} \|A_n x - Ax\| < \varepsilon.$$

Le théorème est ainsi démontré. □

En référence au théorème 4.2.5, on dit qu'un espace de Hilbert possède la *propriété d'approximation*. Attention, cette propriété d'apparence naturelle n'est pas partagée par tous les espaces de Banach.

Exercice 4.2.5 *Soient H un espace de Hilbert complexe, séparable, de dimension infinie et K un compact de $\mathbb{C}$. Démontrez qu'il existe un élément A de $\mathscr{L}(H)$ tel que* $\mathrm{Sp}(A) = K$.

4.3 Dualité des espaces de Hilbert

La dualité des espaces euclidiens et hermitiens en dimension finie est familière. Soit H un espace de Hilbert de dimension finie, on sait que l'application δ définie par

$$\begin{aligned} \delta\colon H &\to H' \\ x &\mapsto \delta(x)\,:\ y \mapsto (y|x) \end{aligned}$$

est une bijection antilinéaire et isométrique de H sur H' qui permet d'identifier H et son dual par l'intermédiaire du produit scalaire. Il est remarquable que pour un espace de Hilbert de dimension infinie, la situation ne diffère pas. C'est le fameux théorème de représentation de Riesz (plutôt un des fameux théorèmes d'un des Riesz).

Théorème 4.3.1 *Soit H un espace de Hilbert, on considère l'application δ définie par*

$$\begin{aligned} \delta : H &\to H' \\ x &\mapsto \delta(x) : y \mapsto (y|x) \end{aligned}$$

C'est une bijection antilinéaire isométrique.

Démonstration. Le fait que $\|\delta(x)\|_{H'} \leq \|x\|$, donc en particulier que $\delta(x)$ est une forme linéaire continue, résulte de l'inégalité de Cauchy-Schwarz. Le fait que δ est antilinéaire est évident. De plus,

$$\left\langle \delta(x), \frac{x}{\|x\|} \right\rangle = \|x\|,$$

donc δ est une isométrie. L'injectivité de δ en résulte. Pour démontrer la surjectivité de δ, considérons un élément ℓ de $H' \setminus \{0\}$. Son noyau est un hyperplan (un espace de codimension 1, comme tout noyau de forme linéaire non nulle) fermé de H. Son orthogonal est donc de dimension 1. Soit e un vecteur unitaire de $\ker \ell^{\perp}$, on pose $x = \overline{\ell(e)}e$. Pour tout $y \in H$, écrivons $y = (y|e)e + z$ avec $z \in \ker \ell$. Il vient

$$\ell(y) = (y|e)\ell(e) = (y|x) = \langle \delta(x), y \rangle.$$

Le théorème est ainsi démontré. $\square$

Remarques. Dans la pratique, le théorème de Riesz est utilisé sous la forme suivante : pour toute forme linéaire continue ℓ sur H, il existe un unique $x \in H$ qui représente cette forme linéaire grâce au produit scalaire au sens où

$$\forall y \in H, \quad \ell(y) = (y|x).$$

Il faut également noter que le théorème de Riesz implique que la norme duale sur H' est une norme hilbertienne, ce qui n'est aucunement évident a priori. En particulier, on a un produit scalaire naturel sur H' défini par

$$(\ell_1|\ell_2)_{H'} = (\delta^{-1}(\ell_2)|\delta^{-1}(\ell_1))_H$$

en remarquant pour l'inversion de l'ordre des termes, que δ^{-1} est antilinéaire.

Corollaire 4.3.1 *Tout espace de Hilbert est réflexif.*

Démonstration. Notons δ' l'application de H' dans H'' définie par le théorème de Riesz, avec H' muni du produit scalaire naturel ci-dessus. Pour tout $\psi \in H''$, il existe donc $\ell \in H'$ tel que $\psi = \delta'(\ell)$. Mais il existe aussi $x \in H$ tel que $\ell = \delta(x)$. En d'autres termes, $\delta'(\delta(x)) = \psi$. Par définition de δ', ceci signifie que pour tout $\ell' \in H'$,

$$\begin{aligned}\langle \psi, \ell' \rangle_{H'',H'} &= (\ell' | (\delta')^{-1}(\psi))_{H'} = (\ell' | \ell)_{H'} \\ &= (\ell' | \delta(x))_{H'} = (x | \delta^{-1}(\ell'))_H = \langle \ell', x \rangle_{H',H} = \langle \imath(x), \ell' \rangle_{H'',H'}.\end{aligned}$$

L'injection canonique $\imath$ est donc surjective. □

On peut déduire du théorème de Riesz les deux corollaires ci-dessous qui montrent d'autres identifications du dual d'un espace de Hilbert donné à l'aide de formes bilinéaires (et non sequilinéaires). Ceci pour souligner qu'il n'y a pas qu'une unique façon d'identifier le dual d'un espace de Hilbert.

Corollaire 4.3.2 *La forme bilinéaire définie par*

$$\begin{aligned} B\colon L^2(X,d\mu) \times L^2(X,d\mu) &\to \mathbb{K} \\ (f,g) &\mapsto \int_X f(x)g(x)\,d\mu(x) \end{aligned}$$

identifie le dual de $L^2(X,d\mu)$ à $L^2(X,d\mu)$.

Corollaire 4.3.3 *La forme bilinéaire $\widetilde{B}$ définie par*

$$\begin{aligned} \widetilde{B}\colon L^2(\mathbb{R}^d,(1+|\xi|^2)^{-s}d\xi) \times L^2(\mathbb{R}^d,(1+|\xi|^2)^{s}d\xi) &\to \mathbb{K} \\ (f,g) &\mapsto \int_{\mathbb{R}^d} f(\xi)g(-\xi)d\xi \end{aligned}$$

identifie $L^2(\mathbb{R}^d,(1+|\xi|^2)^{-s}d\xi)$ au dual de $L^2(\mathbb{R}^d,(1+|\xi|^2)^{s}d\xi)$.

Démonstration. Pour démonter le premier corollaire, il suffit d'observer que $\delta_B(f) = \delta(\bar{f})$ ce qui assure immédiatement le résultat (qui n'est autre que le théorème de Riesz lui-même dans le cas $\mathbb{K} = \mathbb{R}$).

La démonstration du second corollaire est très légèrement plus délicate. Posons pour tout $s \in \mathbb{R}$,

$$M_s(f)(\xi) \stackrel{\text{déf}}{=} (1+|\xi|^2)^s f(\xi) \quad \text{et} \quad \check{M}_s(f)(\xi) \stackrel{\text{déf}}{=} (1+|\xi|^2)^s f(-\xi).$$

Nous allons démontrer que $\delta_{\widetilde{B}} = {}^t\check{M}_s \circ \delta_B \circ M_{-s}$, ce qui assurera le résultat d'après le corollaire précédent et la proposition 3.4.2.

$$
\begin{aligned}
\langle \delta_{\widetilde{B}}(f), g\rangle &= \int_{\mathbb{R}^d} f(\xi)g(-\xi)d\xi \\
&= \int_{\mathbb{R}^d} (1+|\xi|^2)^{-s} f(\xi)(1+|\xi|^2)^s g(-\xi)d\xi \\
&= B(M_{-s}(f), \check{M}_s(g)) \\
&= \langle \delta_B(M_{-s}(f)), \check{M}_s(g)\rangle \\
&= \langle {}^t\check{M}_s \circ \delta_B \circ M_{-s}(f), g\rangle,
\end{aligned}
$$

d'où le corollaire. □

Dans un espace de Hilbert, on dispose d'une nouvelle notion de convergence affaiblie.

Définition 4.3.1 *Soient $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments d'un espace de Hilbert H et x un élément de H. On dit que la suite $(x_n)_{n\in\mathbb{N}}$ converge faiblement vers x et l'on note $x_n \rightharpoonup x$ si et seulement si*

$$\forall y \in H, \quad (y|x_n) \to (y|x) \text{ quand } n \to +\infty.$$

Comme pour la convergence faible-$*$ sur un dual, il existe une topologie sur H, dite topologie faible, dont les suites convergentes sont exactement celles définies ci-dessus. Pour la même raison, nous n'étudierons pas cette topologie faible.

La définition ci-dessus peut être reformulée par

$$x_n \rightharpoonup x \Longleftrightarrow \delta(x_n) \overset{*}{\rightharpoonup} \delta(x),$$

en utilisant l'identification du dual d'un espace de Hilbert avec lui-même du théorème de Riesz.

Le théorème suivant donne quelques propriétés simples de la convergence faible.

Théorème 4.3.2 *Soient $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments d'un espace de Hilbert H et x un élément de H. On a alors*

$$\text{si } \lim_{n\to\infty} \|x_n - x\| = 0, \text{ alors } x_n \rightharpoonup x, \tag{4.3}$$

$$\text{si } x_n \rightharpoonup x \text{ et } \|x_n\| \to \|x\|, \text{ alors } \lim_{n\to\infty} \|x_n - x\| = 0. \tag{4.4}$$

De plus, si $x_n \rightharpoonup x$, alors la suite $(x_n)_{n\in\mathbb{N}}$ est bornée.

Démonstration. Le premier point du théorème résulte simplement du fait que

$$|(y|x_n) - (y|x)| \leq \|y\| \, \|x_n - x\|.$$

Quant au second point, il suffit d'écrire que

$$\|x_n - x\|^2 = \|x_n\|^2 - 2\,\mathfrak{Re}(x|x_n) + \|x_n\|^2;$$

Comme la suite $(x_n)_{n\in\mathbb{N}}$ tend faiblement vers x, on a

$$\mathfrak{Re}(x|x_n) \to \mathfrak{Re}(x|x) = \|x\|^2.$$

Le deuxième point est ainsi démontré. Le troisième point est encore un corollaire du théorème de Banach-Steinhaus. En effet, nous avons noté que $x_n \rightharpoonup x$ est équivalent à $\delta(x_n) \overset{*}{\rightharpoonup} \delta(x)$, ce qui implique que $\delta(x_n)$ est bornée dans H'. Or, le théorème de représentation de Riesz dit que δ est un isométrie. □

Exemple. Soit H un espace de Hilbert de dimension infinie et séparable. Considérons une base hilbertienne $(e_n)_{n\in\mathbb{N}}$ de H. D'après le théorème 4.2.4, pour tout x de H, la suite $((x|e_n))_{n\in\mathbb{N}}$ appartient à $\ell^2(\mathbb{N})$. Donc elle tend vers 0. Voilà un exemple de suite faiblement convergente vers 0 qui ne converge pas au sens de la norme vers 0 puisque $\|e_n\| = 1$.

Proposition 4.3.1 *Soient $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ deux suites d'éléments de H telles que*

$$x_n \to x \text{ et } y_n \rightharpoonup y \text{ quand } n \to +\infty.$$

Alors, on a $(x_n|y_n) \to (x|y)$.

Ce n'est qu'une réécriture dans le cadre hilbertien de la proposition 3.3.1.

Exercice 4.3.1 *Trouvez deux suites $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ d'éléments de H telles que*

$$x_n \rightharpoonup x, y_n \rightharpoonup y \text{ et } (x_n|y_n) \not\to (x|y).$$

Exercice 4.3.2 *Démontrez que, dans un espace de Hilbert 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 un 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 2.1.2 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 3.3.2.

Théorème 4.3.3 *Soit* $(x_n)_{n\in\mathbb{N}}$ *une suite bornée d'un espace de Hilbert séparable H, on peut en extraire une sous-suite faiblement convergente.*

Remarque. Le théorème reste vrai même si H n'est pas séparable.

Le théorème de Lax-Milgram suivant est très souvent utilisé pour résoudre des problèmes d'équations aux dérivées partielles elliptiques.

Théorème 4.3.4 *Soit* H *un espace de Hilbert,* ℓ *une forme linéaire continue sur* H *et* a *une forme bilinéaire continue sur* H *telle qu'il existe* $\alpha > 0$ *tel que*

$$\forall v \in H, \quad \mathfrak{Re}\, a(v,v) \geq \alpha \|v\|^2 \text{ (on dit que } a \text{ est } H\text{-elliptique)}.$$

Alors le problème : trouver $u \in H$ *tel que*

$$\forall v \in H, \quad a(u,v) = \ell(v),$$

admet une solution unique. De plus, l'application qui à ℓ *associe* u *est linéaire continue de* H' *dans* H.

Démonstration. La forme bilinéaire a étant continue, il existe une constante M telle que

$$\forall v,w \in H, \quad |a(v,w)| \leq M\|v\|\|w\|.$$

On en déduit immédiatement que l'application A de H dans H' définie par $\langle A(v), w\rangle = a(v,w)$ est linéaire continue, avec $\|A\|_{\mathscr{L}(H,H')} \leq M$.

L'application A est injective. En effet, si $A(v) = 0$, alors en particulier $0 = \langle A(v), v\rangle = a(v,v) = \mathfrak{Re}\, a(v,v) + i\,\mathfrak{Im}\, a(v,v)$, d'où $0 = \mathfrak{Re}\, a(v,v) \geq \alpha\|v\|^2$ et par conséquent $v = 0$.

L'application A est surjective. En effet, soit $v_n \in H$ une suite telle que $A(v_n)$ converge dans H' vers un certain ℓ. Cette suite est donc telle que $(A(v_n))_{n\in\mathbb{N}}$ est de Cauchy dans H'. Par conséquent, comme

$$\begin{aligned}
\alpha\|v_n - v_m\|_H^2 &\leq \mathfrak{Re}\, a(v_n - v_m, v_n - v_m)\\
&\leq |a(v_n - v_m, v_n - v_m)|\\
&= |\langle A(v_n - v_m), v_n - v_m\rangle|\\
&\leq \|A(v_n) - A(v_m)\|_{H'}\|v_n - v_m\|_H,
\end{aligned}$$

on en déduit que

$$\|v_n - v_m\|_H \leq \alpha^{-1}\|A(v_n) - A(v_m)\|_{H'},$$

donc $(v_n)_{n\in\mathbb{N}}$ est de Cauchy dans H. Il existe donc $v \in H$ tel que $v_n \to v$ dans H, ce qui implique que $A(v_n) \to A(v) = \ell$ dans H'. En d'autres termes, on vient de montrer que l'image de A est fermée dans H'. On va maintenant montrer qu'elle est dense. Pour cela, il suffit de remarquer que si $w \in H'$ est tel que $(A(v)|w)_{H'} = 0$ pour tout $v \in H$, alors prenant $v = \delta^{-1}w$, on en déduit encore une fois que $w = 0$. Ceci implique que l'orthogonal de l'image de A dans H' est réduit au vecteur nul, donc que cette image est dense. L'image de A étant fermée et dense, elle est égale à H', c'est-à-dire que A est surjective.

Enfin, on remarque que $u = A^{-1}\ell$ et que $\|u\|_H \leq \alpha^{-1}\|\ell\|_{H'}$, ce qui montre le théorème. $\square$

Remarque. Dans le cas d'un espace de Hilbert réel, la condition de H-ellipticité s'écrit simplement

$$\forall v \in H, \quad a(v,v) \geq \alpha\|v\|^2.$$

On remarque également en multipliant par $-i$, qu'il suffit que la partie imaginaire de la forme bilinéaire satisfasse la condition de H-ellipticité pour que la conclusion du théorème de Lax-Milgram soit assurée.

4.4 La notion d'adjoint

La notion d'adjoint est cousine dans le cadre hilbertien de celle de transposition. Elle est bien connue en dimension finie. Étant donné un produit scalaire sur un espace H de dimension finie d, l'adjointe A^* d'une application linéaire est défini par la relation

$$\forall (x,y), \qquad (Ax|y) = (x|A^*y).$$

On sait que si $(A_{ij})_{1\leq i,j\leq d}$ est la matrice de A dans une base orthonormée, alors la matrice de A^* dans cette même base est donnée par $A^*_{ij} = \bar{A}_{ji}$. De plus, c'est un théorème bien classique d'algèbre linéaire que si A est autoadjointe (i.e. $A = A^*$), alors 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 H. Il existe un unique opérateur linéaire A^* continu sur H tel que*

$$\forall (x,y) \in H \times H, \ (Ax|y) = (x|A^*y).$$

De plus, l'application qui à A associe A^* est une application antilinéaire isométrique de $\mathscr{L}(H)$ dans lui-même.

Démonstration. L'unicité de l'opérateur A^* résulte immédiatement du fait que $H^\perp = \{0\}$. Soit $\mathscr{L}_A : H \to H'$ l'application définie par

$$\langle \mathscr{L}_A(y), x\rangle = (Ax|y).$$

Il est visible que l'application $\mathscr{L}_A$ est une application antilinéaire continue de H dans H'. Posons $A^* \stackrel{\text{déf}}{=} \delta^{-1} \circ \mathscr{L}_A$. Par définition de A^*, on a

$$(x|A^*y) = \langle \mathscr{L}_A(y), x \rangle = (Ax|y).$$

De plus, on a

$$\|A\| = \sup_{\|x\| \leq 1, \|y\| \leq 1} |(Ax|y)| = \sup_{\|x\| \leq 1, \|y\| \leq 1} |(A^*x|y)| = \|A^*\|.$$

Le théorème 4.4.1 est donc démontré. □

Définition 4.4.1 *On appelle l'opérateur A^* ci-dessus l'*opérateur adjoint de A. *On dit qu'un élément A de $\mathscr{L}(H)$ est* auto-adjoint *si $A = A^*$.*

On a défini l'adjoint pour un opérateur continu. Notons que la notion prend encore plus d'intérêt, mais se complique notablement, pour les opérateurs non continus. Nous n'aborderons pas ces questions ici. Donnons quelques propriétés de l'adjonction.

Proposition 4.4.1 *L'opération de passage à l'adjoint est une involution, c'est-à-dire que $A = A^{**}$. On a aussi*

$$(\ker A)^{\perp} = \overline{\operatorname{im} A^*} \quad et \quad \ker A = (\operatorname{im} A^*)^{\perp}. \tag{4.5}$$

pour tout $A \in \mathscr{L}(H)$. De plus, si la suite $(x_n)_{n \in \mathbb{N}}$ converge faiblement vers x, alors la suite $(A^*x_n)_{n \in \mathbb{N}}$ converge faiblement vers A^*x. Enfin, si A est compact, alors A^* l'est aussi.

Démonstration. Le premier point résulte du fait que si $(Ax|y) = (x|A^*y)$, alors $(A^*y|x) = (y|Ax)$. Pour démontrer les relations (4.5), écrivons que

$$\forall y \in H,\ (Ax|y) = 0 \Leftrightarrow \forall y \in H,\ (x|A^*y) = 0.$$

La traduction ensembliste de l'équivalence ci-dessus est exactement l'égalité entre les deux ensembles $\ker A$ et $(\operatorname{im} A^*)^{\perp}$. 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 H, on a $(y|A^*x_n) = (Ay|x_n)$. La convergence faible de la suite $(x_n)_{n \in \mathbb{N}}$ entraîne donc immédiatement celle de la suite $(A^*x_n)_{n \in \mathbb{N}}$.

Supposons maintenant que A soit compact. Le théorème 4.2.5 affirme l'existence d'une suite $(A_n)_{n \in \mathbb{N}}$ d'opérateurs de rang fini convergeant vers A. Comme l'opération * est continue de $\mathscr{L}(H)$ dans lui-même, il suffit de démontrer que l'adjoint d'un opérateur de rang fini l'est aussi. Soit C un opérateur de rang fini. D'après les relations (4.5), le noyau de C^* est l'orthogonal de $\operatorname{im} C$, qui est de dimension finie. Donc, l'image de C^* est égale à l'image de C^* restreint au sous-espace $\operatorname{im} C$, qui est de dimension finie. Donc C^* est de rang fini.

La démonstration de la proposition est ainsi achevée. □

Remarque. Pour la propriété de convergence faible, l'adjoint n'a rien magique, puisque $A = (A^*)^*$. Notons également que si A est de rang fini, alors $\operatorname{rang} A^* = \operatorname{rang} A$. En effet, on a un isomorphisme algébrique entre $H/\ker A^*$ et $\operatorname{im} A^*$, et un isomorphisme topologique entre $H/(\operatorname{im} A)^\perp$ et $\operatorname{im} A$ car $\operatorname{im} A$ est fermée.

Mentionnons une dernière propriété importante.

Proposition 4.4.2 *Soit A un isomorphisme sur un espace de Hilbert H. Alors A^* est un isomorphisme avec $(A^*)^{-1} = (A^{-1})^*$.*

Démonstration. Posons $B = (A^{-1})^*$. Pour tout $(x,y) \in H^2$, on a

$$(Ax|By) = (A^{-1}(Ax)|y) = (x|y) \text{ d'une part, et d'autre part } (Ax|By) = (x|A^*By),$$

d'où $A^*B = \mathrm{Id}$. De même

$$(Ax|BA^*y) = (x|A^*y) = (Ax|y),$$

d'où $BA^* = \mathrm{Id}$ puisque A est surjective. □

4.5 Spectre d'un opérateur auto-adjoint compact

On donne ici l'analogue infini dimensionnel du théorème de diagonalisabilité orthogonale des matrices symétriques ou hermitiennes en dimension finie. C'est un résultat d'une grande importance en théorie des équations aux dérivées partielles.

Proposition 4.5.1 *Tout opérateur compact de H dans H transforme les suites faiblement convergentes en suites fortement convergentes.*

Démonstration. Soit A un opérateur compact et $(x_n)_{n\in\mathbb{N}}$ une suite qui tend faiblement vers x. D'après la proposition 4.4.1 appliquée à A^*, on a $Ax_n \rightharpoonup Ax$. D'un autre côté, la suite $(x_n)_{n\in\mathbb{N}}$ est bornée, donc la suite $(Ax_n)_{n\in\mathbb{N}}$ est relativement compacte. Soit $(n_p)_{p\in\mathbb{N}}$ une sous-suite telle que $A_{x_{n_p}} \to y$. Clairement, comme *a fortiori* $A_{x_{n_p}} \rightharpoonup y$, on a $y = Ax$, ce qui montre que la suite $(Ax_n)_{n\in\mathbb{N}}$ n'a qu'une seule valeur d'adhérence. Comme elle est dans un compact, elle converge vers cette valeur d'adhérence. □

Théorème 4.5.1 *Soit A un opérateur compact auto-adjoint d'un espace de Hilbert H séparable de dimension infinie.*

i) Le spectre de A est la réunion de $\{0\}$ et soit d'une suite $(\lambda_j)_{j\in\mathbb{N}}$ de valeurs propres réelles non nulles tendant vers 0, soit d'un nombre fini de valeurs propres réelles non nulles.

ii) L'espace vectoriel $\ker(A-\lambda_j \operatorname{Id})$ *est de dimension finie si* λ_j *est non nul.*

iii) Il existe une base hilbertienne $(e_j)_{j\in\mathbb{N}}$ *de* H *formée de vecteurs propres et l'on a*

$$\forall x \in H, \quad Ax = \sum_{j\in\mathbb{N}} \tilde{\lambda}_j (x|e_j)e_j \text{ et } \|Ax\|^2 = \sum_{j\in\mathbb{N}} \tilde{\lambda}_j^2 |(x|e_j)|^2,$$

où la famille $(\tilde{\lambda}_j)_{j\in\mathbb{N}}$ *est formée de la suite ou de la famille finie des valeurs propres non nulles, éventuellement complétée par la valeur propre nulle. De plus* $\|A\|_{\mathscr{L}(H)} = \max_{j\in\mathbb{N}} |\tilde{\lambda}_j|$.

Démonstration. Démontrons le premier point en commençant par montrer que 0 appartient au spectre. En effet, si 0 n'appartenait pas au spectre, A serait un isomorphisme et d'après le théorème de l'application ouverte, l'image de la boule unité contiendrait une boule, donc ne serait pas relativement compacte.

Montrons ensuite que tout élément non nul du spectre est une valeur propre. Notons d'abord que si λ appartient au spectre de A, alors $\bar{\lambda}$ aussi. En effet, $(A-\lambda \operatorname{Id})^* = A - \bar{\lambda} \operatorname{Id}$ et si l'un est un isomorphisme, l'autre aussi. Soit $\lambda \neq 0$ appartenant au spectre et tel que $A-\lambda \operatorname{Id}$ soit injective. On va montrer qu'alors $A-\lambda \operatorname{Id}$ est surjective, ce qui sera une contradiction.

Montrons d'abord que l'image de $A-\lambda \operatorname{Id}$ est fermée. Pour cela, on note qu'il existe une constante $C>0$ telle que l'on ait l'inégalité

$$\forall x \in H, \quad \|x\| \leq C\|Ax-\lambda x\|. \tag{4.6}$$

En effet, s'il n'existait aucune telle constante C, on construirait une suite $(x_n)_{n\in\mathbb{N}}$ telle que

$$\|x_n\| = 1 \text{ et } \|Ax_n - \lambda x_n\| \leq \frac{1}{n}.$$

Extrayant si besoin est une sous-suite, on peut aussi bien supposer qu'il existe $x \in \bar{B}(0,1)$ tel que $x_n \rightharpoonup x$. Comme A est compact, on en déduit que $Ax_n \to Ax$ au sens de la norme. Par conséquent

$$\|x_n - \lambda^{-1}Ax\| \leq |\lambda|^{-1}\|Ax_n - Ax\| + |\lambda|^{-1}\|\lambda x_n - Ax_n\| \to 0 \text{ quand } n \to +\infty.$$

On en déduit deux choses : d'une part la suite x_n converge fortement, donc sa norme passe à la limite, c'est-à-dire $\|x\| = 1$; d'autre part $x = \lambda^{-1}Ax$, c'est-à-dire $x \in \ker(A-\lambda \operatorname{Id})$. Or on a supposé que $A-\lambda \operatorname{Id}$ est injective, donc c'est une contradiction. Par conséquent, l'inégalité (4.6) a bien lieu. Cette inégalité montre immédiatement que $\operatorname{im}(A-\lambda \operatorname{Id})$ est fermée.

Notons maintenant que toutes les valeurs propres sont réelles. En effet, soit λ une valeur propre associée à un vecteur propre unitaire x. Comme A est auto-adjoint, il vient

$$\lambda = (\lambda x|x) = (Ax|x) = (x|Ax) = (x|\lambda x) = \bar{\lambda}.$$

Soit alors λ un élément du spectre non nul qui n'est pas une valeur propre. D'après ce qui précède, $\bar{\lambda}$ appartient aussi au spectre et n'est pas une valeur propre. Par conséquent,

$$H = (\ker(A - \bar{\lambda}\,\mathrm{Id}))^{\perp} = \overline{\mathrm{im}(A - \lambda\,\mathrm{Id})},$$

c'est-à-dire que l'image de $A - \lambda\,\mathrm{Id}$ est dense. Comme elle est fermée, on en déduit que $A - \lambda\,\mathrm{Id}$ est surjective, donc que λ n'appartient pas au spectre, contradiction.

Les éléments non nuls du spectre de A sont donc tous des valeurs propres.

Notons enfin que si $\lambda_1 \neq \lambda_2$ sont deux valeurs propres distinctes, les vecteurs propres x_1, x_2 qui leur sont associés sont orthogonaux. En effet

$$\lambda_1(x_1|x_2) = (Ax_1|x_2) = (x_1|Ax_2) = \lambda_2(x_1|x_2).$$

Résumons ce que nous avons montré jusqu'à présent. Le spectre de A est formé de 0 et de valeurs propres réelles non nulles, associées à des vecteurs propres deux-à-deux orthogonaux. Comme H est séparable, ceci implique que les valeurs propres sont en nombre au plus dénombrable. Soit il n'en existe qu'un nombre fini non nulles, et on a terminé le premier point, soit on en a une suite $(\lambda_j)_{j\in\mathbb{N}}$. Donnons nous $(e_j)_{j\in\mathbb{N}}$ une suite de vecteurs propres orthonormés correspondants. La suite $(e_j)_{j\in\mathbb{N}}$ tend faiblement vers 0. Comme A est compact, la suite $(Ae_j)_{j\in\mathbb{N}}$ tend vers 0 en norme. Mais comme $Ae_j = \lambda_j e_j$, on en déduit que

$$|\lambda_j|\,\|e_j\| \to 0.$$

Les vecteurs e_j étant de norme 1, la suite $(\lambda_j)_{j\in\mathbb{N}}$ tend vers 0.

Démontrons ensuite le deuxième point du théorème. Posons $E_j = \ker(A - \lambda_j\,\mathrm{Id})$. 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 $(x_{n_p})_{p\in\mathbb{N}}$ telle que la suite $(Ax_{n_p})_{p\in\mathbb{N}}$ converge. Mais $Ax_{n_p} = \lambda_j x_{n_p}$ et λ_j est non nulle. La boule unité de E_j est donc compacte. D'après le théorème 2.1.2 de Riesz, la dimension de E_j est finie.

Pour le troisième point, on va d'abord montrer que A admet au moins une valeur propre. Pour cela, on pose

$$\lambda_0 \stackrel{\text{déf}}{=} \sup_{\|x\|=1} |(Ax|x)|.$$

Si $\lambda_0 = 0$, on a alors $(Ax|x) = 0$ pour tout élément x de H. Or, l'opérateur A étant auto-adjoint, on a

$$(Ax|y) = \frac{1}{2}\Big((A(x+y)|x+y) - (Ax|x) - (Ay|y)\Big).$$

Donc, si $\lambda_0 = 0$, on a $(Ax|y) = 0$ pour tout couple (x,y) d'éléments de H, ce qui implique évidemment que $A = 0$. Comme $A^* = A$, $(Ax|x)$ est réel, donc quitte à changer A en $-A$, on peut supposer que

$$\lambda_0 = \sup_{\|x\|=1} (Ax|x).$$

Par définition, il existe une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de la sphère unité telle que

$$\lim_{n\to\infty} (Ax_n|x_n) = \lambda_0$$

On peut trouver un x dans $\bar{B}(0,1)$ et une sous-suite tels que $x_{n_p} \rightharpoonup x$. Comme l'opérateur A est compact, on en déduit que $Ax_{n_p} \to Ax$. D'après la proposition 4.3.1, il vient alors

$$(Ax_{n_p}|x_{n_p}) \to (Ax|x) = \lambda_0.$$

Bien sûr, $x \neq 0$ puisque λ_0 est strictement positif. De plus, si l'on avait $\|x\| < 1$, alors on aurait

$$\Big(A\frac{x}{\|x\|}\Big|\frac{x}{\|x\|}\Big) = \frac{\lambda_0}{\|x\|^2} > \lambda_0,$$

ce qui contredit la maximalité de λ_0 car λ_0 est strictement positif. Donc $\|x\| = 1$, et un simple argument d'homogénéité permet alors d'affirmer que

$$\forall z \in H, \quad \lambda_0\|z\|^2 - (Az|z) \geq 0.$$

Prenant $z = x + th$ pour $t \in \mathbb{R}_+^*$ et faisant tendre t vers 0^+, on obtient

$$\forall h \in H,\ 2\lambda_0 \mathfrak{Re}(x|h) - 2\mathfrak{Re}(Ax|h) = 0,$$

ce qui entraîne que $Ax = \lambda_0 x$. Nous avons vu que x était non nul, donc λ_0 est bien valeur propre de A.

Soient λ_j, $j \in \mathbb{N}$ ou $1 \leq j \leq N$, les valeurs propres non nulles de A. Dans chaque E_j et dans $\ker A$ choisissons une famille orthonormée. La réunion de ces familles orthonormées est orthonormée et totale. En effet, son orthogonal M est stable par A et le spectre de la restriction de A à M ne contient pas de valeur propre, ce qui contredit ce qui précède, sauf si $M = \{0\}$. Nous pouvons alors renuméroter cette base hilbertienne de telle sorte que la plus grande valeur propre en valeur absolue soit $\tilde{\lambda}_0 = \lambda_0$, en renumérotant également les valeurs propres non nulles et nulles.

Nous avons ainsi construit une base hilbertienne $(e_j)_{j\in\mathbb{N}}$ de H formée de vecteurs propres. Pour tout $x \in H$, les identités de Bessel et de Parseval impliquent que $\sum_{j\in\mathbb{N}} \tilde{\lambda}_j^2 |(x|e_j)|^2 \leq \lambda_0^2\|x\|^2$ est une série convergente, d'où l'on déduit immédiatement que

$$Ax = \sum_{j\in\mathbb{N}} \tilde{\lambda}_j (x|e_j)e_j \text{ et } \|Ax\|^2 = \sum_{j\in\mathbb{N}} \tilde{\lambda}_j^2 |(x|e_j)|^2.$$

On en déduit que

$$\|Ax\|^2 \leq \tilde{\lambda}_0^2 \|x\|^2.$$

Comme par ailleurs, on sait déjà que $|\tilde{\lambda}_0| \leq \|A\|_{\mathscr{L}(H)}$, le théorème est entièrement démontré. □

Remarque. Mis à part 0, le spectre d'un opérateur ne contient donc que des valeurs propres. La valeur 0 peut ou peut ne pas être valeur propre. Si elle ne l'est pas, alors A est injective. Si elle l'est, sa multiplicité peut prendre toutes les valeurs possibles : n'importe quel nombre entier strictement positif si A n'est pas injective, mais $\ker A$ est de dimension finie, et $+\infty$ si $\ker A$ est de dimension infinie.

Chapitre 5

Les espaces L^p

5.1 Rappels sur la théorie de la mesure et l'intégration

Dans cette section, nous allons rappeler — et parfois redémontrer — les principaux énoncés de base de la théorie de la mesure et de l'intégrale de Lebesgue.

Définition 5.1.1 *Soit X un ensemble. On dit qu'une partie $\mathscr{M} \subset \mathscr{P}(X)$ est une* σ-algèbre (*ou* tribu) sur X *si elle satisfait les trois propriétés*

- *L'ensemble X appartient à $\mathscr{M}$,*
- *Si $A \in \mathscr{M}$ alors $X \setminus A \in \mathscr{M}$.*
- *Si $(A_j)_{j\in\mathbb{N}}$ est une famille dénombrable d'éléments de $\mathscr{M}$, alors*

$$\bigcup_{j=1}^{+\infty} A_j \in \mathscr{M}.$$

Le couple $(X, \mathscr{M})$ est appelé espace mesurable. *Les éléments de $\mathscr{M}$ sont appelés les* ensembles mesurables *de X.*

Noter les analogies et les différences entre cette notion et celle de topologie introduite au chapitre 1. Les propriétés suivantes des ensembles mesurables sont élémentaires : $\emptyset \in \mathscr{M}$, $\cup_{j=1}^{n} A_j \in \mathscr{M}$, $\cap_{j=1}^{+\infty} A_j \in \mathscr{M}$ et $\cap_{j=1}^{n} A_j \in \mathscr{M}$.

Étant donnée une partie $\mathscr{A}$ de $\mathscr{P}(X)$, il existe une plus petite σ-algèbre qui contient $\mathscr{A}$. On l'appelle la σ-algèbre *engendrée* par $\mathscr{A}$. Elle est construite simplement, mais de façon pas très explicite, en prenant l'intersection de toutes les σ-algèbres qui contiennent $\mathscr{A}$. Dans le cas où X est aussi un espace topologique, muni d'une topologie $\mathscr{O}$, la σ-algèbre $\mathscr{B}$ engendrée par les ouverts de $\mathscr{O}$ est appelée *tribu borélienne.* Ses éléments sont des *boréliens.* Il est facile de voir que les fermés, les réunions dénombrables de fermés (ou F_σ) et les intersections dénombrables d'ouverts (ou G_δ) sont des boréliens. Mais ce sont loin d'être les seuls boréliens !

Définition 5.1.2 *Soit f une application de X espace mesurable dans Y espace topologique. On dit que f est* mesurable *si pour tout ouvert V de Y, $f^{-1}(V)$ est mesurable.*

Clairement, une application f est mesurable si et seulement si l'image réciproque de tout fermé est mesurable (le vérifier), et plus généralement si et seulement si l'image réciproque de tout borélien est mesurable. Dans le cas où X est un espace topologique, il est clair que toute fonction continue est mesurable pour la tribu borélienne. Les applications mesurables à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$ sont dans ce cas appelées *fonctions boréliennes.*

Dans le cas des fonctions à valeurs réelles ou infinies, on a la caractérisation suivante. On munit $\overline{\mathbb{R}} = [-\infty, +\infty]$ de la topologie de l'ordre, dont une base d'ouverts est donnée par les intervalles ouverts de $\mathbb{R}$ et les intervalles de la forme $[-\infty, a[$ et $]a, +\infty]$.

Proposition 5.1.1 *Une fonction f de X dans $\overline{\mathbb{R}}$ est mesurable si et seulement si pour tout $a \in \overline{\mathbb{R}}$, l'ensemble $f^{-1}([a, +\infty])$ est mesurable.*

Démonstration. Si f est mesurable, comme $[a, +\infty]$ est fermé, $f^{-1}([a, +\infty])$ est bien mesurable. Réciproquement, considérons un intervalle ouvert $]a, b[$ de $\mathbb{R}$. Montrons que $f^{-1}(]a, b[)$ est mesurable. On a

$$]a, b[=]a, +\infty] \cap [-\infty, b[=]a, +\infty] \cap (\overline{\mathbb{R}} \setminus [b, +\infty]).$$

Par conséquent,

$$f^{-1}(]a, b[) = f^{-1}(]a, +\infty]) \cap (X \setminus f^{-1}([b, +\infty])),$$

et l'on a gagné si l'on montre que $f^{-1}(]a, +\infty])$ est mesurable. Or,

$$]a, +\infty] = \bigcup_{n=1}^{\infty} [a + 1/n, +\infty] \text{ d'où } f^{-1}(]a, +\infty]) = \bigcup_{n=1}^{\infty} f^{-1}([a + 1/n, +\infty])$$

qui est une union dénombrable d'ensemble mesurables. On note alors que tout ouvert V de $\mathbb{R}$ est réunion dénombrable d'intervalles ouverts, $V = \cup_{n=1}^{\infty}]a_n, b_n[$, donc $f^{-1}(V) = \cup_{n=1}^{\infty} f^{-1}(]a_n, b_n[)$ est mesurable. Pour conclure, on remarque que les ouverts de $\overline{\mathbb{R}}$ ne sont rien d'autre que les ouverts de $\mathbb{R}$, éventuellement complétés par $\{-\infty\}$ ou $\{+\infty\}$. Or, prenant $a = +\infty$, on sait déjà que $f^{-1}(\{+\infty\})$ est mesurable, et comme $f^{-1}(\{-\infty\}) = f^{-1}(\overline{\mathbb{R}}) \setminus \cup_{n \in \mathbb{N}} f^{-1}([-n, +\infty])$, l'ensemble $f^{-1}(\{-\infty\})$ est aussi mesurable. □

Remarque. Pour les fonctions à valeurs dans $\mathbb{R}$, il suffit de montrer que pour tout $a \in \mathbb{R}$, l'ensemble $f^{-1}([a,+\infty[)$ (ou $f^{-1}(]a,+\infty[)$) est mesurable.

Les fonctions mesurables ont de nombreuses propriétés de stabilité.

Proposition 5.1.2 i) *Si f et g sont mesurables de X dans $\mathbb{R}$ alors (f,g) est mesurable de X dans $\mathbb{R}^2$.*

ii) *Si h est continue de Y dans Z et f est mesurable de X dans Y, alors $h \circ f$ est mesurable de X dans Z.*

iii) *Si f_1 et f_2 sont mesurables de X dans $\mathbb{R}$ alors $f_1 + f_2$, $f_1 f_2$, $f_1 + i f_2$ sont mesurables de X dans $\mathbb{C}$. L'ensemble des fonctions mesurables à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$ est un espace vectoriel.*

Démonstration. ii) est évident. iii) découle de i) et de ii). Pour montrer i), il suffit de remarquer que tout ouvert de $\mathbb{R}^2$ est réunion dénombrable de rectangles ouverts et que $(f,g)^{-1}(]a,b[\times]c,d[) = f^{-1}(]a,b[) \cap g^{-1}(]c,d[)$. $\square$

Une conséquence utile de ii) est que si f est mesurable à valeurs dans $\overline{\mathbb{R}}$, alors $|f|$, f_+ et f_- sont mesurables à valeurs dans $\overline{\mathbb{R}}_+ = [0,+\infty]$. De même, si f est mesurable à valeurs dans $\mathbb{C}$, alors $|f|$, $\mathfrak{Re}\, f_+$, $\mathfrak{Re}\, f_-$, $\mathfrak{Im}\, f_+$ et $\mathfrak{Im}\, f_-$ sont mesurables à valeurs dans $\mathbb{R}_+$. Remarquons que l'ensemble des fonctions mesurables à valeurs dans un espace vectoriel topologique est un espace vectoriel.

Une autre caractéristique de la mesurabilité des fonctions est sa très grande stabilité par passage à la limite, à la grande différence de la continuité.

Proposition 5.1.3 *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables de X dans $\overline{\mathbb{R}}$, alors les fonctions $\inf_{n\in\mathbb{N}} f_n$, $\sup_{n\in\mathbb{N}} f_n$, $\liminf_{n\to+\infty} f_n$ et $\limsup_{n\to+\infty} f_n$ sont mesurables de X dans $\overline{\mathbb{R}}$. De plus, si l'on définit une fonction f de X dans $\overline{\mathbb{R}}$ par $f(x) = 0$ si la suite $(f_n(x))_{n\in\mathbb{N}}$ ne converge pas et $f(x) = \lim_{n\to+\infty} f_n(x)$ sinon, alors f est mesurable de X dans $\overline{\mathbb{R}}$.*

Démonstration. Montrons d'abord la mesurabilité de la borne inférieure. Soit $g = \inf_{n\in\mathbb{N}} f_n$. Pour tout $a \in \overline{\mathbb{R}}$, dire que $x \in g^{-1}([a,+\infty])$ est équivalent à dire que pour tout $n \in \mathbb{N}$, $f_n(x) \geq a$. En d'autres termes, $g^{-1}([a,+\infty]) = \cap_{n\in\mathbb{N}} f_n^{-1}([a,+\infty])$ qui est donc mesurable. Comme $\sup_{n\in\mathbb{N}} f_n = -\inf_{n\in\mathbb{N}}(-f_n)$, $\liminf_{n\to+\infty} f_n = \sup_{n\in\mathbb{N}}\big(\inf_{k\geq n} f_k\big)$ et $\limsup_{n\to+\infty} f_n = -\liminf_{n\to+\infty}(-f_n)$, ces trois autres fonctions sont aussi mesurables.

Considérons l'ensemble $E = \{x \in X; f_n(x) \text{ converge}\} = h^{-1}(\{0\})$ où $h = \limsup_{n\to+\infty} f_n - \liminf_{n\to+\infty} f_n$. L'ensemble E est donc mesurable. Par conséquent, posant $l = \limsup_{n\to+\infty} f_n$, on voit que si $a > 0$, alors $f^{-1}([a,+\infty]) = l^{-1}([a,+\infty]) \setminus E$ est mesurable, et si $a \leq 0$, alors $f^{-1}([a,+\infty]) = l^{-1}([a,+\infty]) \cup E$ est aussi mesurable. $\square$

Remarque. En particulier, si une suite de fonctions mesurables converge simplement, alors sa limite est une fonction mesurable.

Parmi les fonctions mesurables à valeurs dans $\mathbb{R}$ (ou dans un espace vectoriel), il en est de plus simples que les autres, ce sont les fonctions simples ou fonctions étagées.

Définition 5.1.3 *Soit $(X,\mathscr{M})$ un espace mesurable. On appelle* fonction étagée *toute application mesurable de X à valeurs dans $\overline{\mathbb{R}}$ dont l'image est finie.*

En d'autres termes, une fonction étagée est une fonction mesurable qui ne prend qu'un nombre fini de valeurs. Il est facile de voir que toute fonction étagée s'écrit sous la forme $f = \sum_{i=1}^m \lambda_i \mathbf{1}_{E_i}$ où les λ_i sont les valeurs prises par f, les ensembles E_i forment une partition mesurable de X et $\mathbf{1}_{E_i}$ désigne la fonction caractéristique de E_i, qui vaut 1 en tout point de E_i et 0 ailleurs. Réciproquement, toute fonction qui peut s'écrire sous cette forme (sans que les E_i ne forment nécessairement une partition de X) est une fonction étagée. Les fonctions étagées permettent d'approcher les fonctions mesurables générales.

Lemme 5.1.1 *Soit f une fonction mesurable de X à valeurs dans $\overline{\mathbb{R}}_+$. Il existe une suite croissante de fonctions étagées à valeurs dans $\mathbb{R}_+$ qui converge simplement vers f.*

Démonstration. Définissons les fonctions f_n pour $n \geq 1$ par

$$f_n \stackrel{\text{déf}}{=} \sum_{j=0}^{(2^n)^2-1} \frac{j}{2^n} \mathbf{1}_{f^{-1}\left(\left[\frac{j}{2^n}, \frac{j+1}{2^n}\right[\right)} + 2^n \mathbf{1}_{f^{-1}([2^n,+\infty])}.$$

Il est clair qu'il s'agit de fonctions étagées. Les valeurs que peut prendre f_n vont de 0 à 2^n par incréments de $1/2^n$. Si x est tel que $f(x)$ est fini, alors il existe n_0 tel que pour tout $n \geq n_0$, on peut encadrer $f(x)$ par de telles valeurs, c'est-à-dire qu'il existe un unique $0 \leq j \leq (2^n)^2 - 1$ tel que $j/2^n \leq f(x) < (j+1)/2^n$. Par conséquent, on a pour tout $n \geq n_0$, $0 \leq f(x) - f_n(x) \leq 1/2^n$. Donc $f_n(x) \leq f(x)$ et $f_n(x) \to f(x)$. Si $f(x) = +\infty$, alors trivialement $f_n(x) \leq f(x)$, et comme $x \in f^{-1}([n,+\infty])$ pour tout n, $f_n(x) = 2^n \to f(x)$.

Il reste à montrer qu'il s'agit d'une suite croissante. Si $f(x) = +\infty$, c'est clair d'après ce qui précède. Si $f(x)$ est fini, alors il existe n_0 tel que pour tout $n \leq n_0$, $f(x) \geq 2^n$ auquel cas $f_n(x) = 2^n$ est croissante. Puis pour $n \geq n_0 + 1$ il existe $0 \leq j \leq (2^n)^2 - 1$ tel que $j/2^n \leq f(x) < (j+1)/2^n$. On a alors $f_n(x) = j/2^n$. Si $2j/2^{n+1} \leq f(x) < (2j+1)/2^{n+1}$, alors $f_{n+1}(x) = 2j/2^{n+1} = f_n(x)$, par contre si $(2j+1)/2^{n+1} \leq f(x) < (2j+2)/2^{n+1}$, alors $f_{n+1}(x) = (2j+1)/2^{n+1} > f_n(x)$. Dans les deux cas, on a bien $f_{n+1}(x) \geq f_n(x)$. □

Remarque. Si f est de plus bornée, alors il est clair que la suite f_n converge uniformément vers f.

Définition 5.1.4 *Soit $(X, \mathscr{M})$ un espace mesurable. On appelle* mesure positive *sur $\mathscr{M}$ toute application μ de $\mathscr{M}$ à valeurs dans $[0, +\infty]$, dénombrablement additive, c'est-à-dire telle que pour toute famille dénombrable $(A_j)_{j\in\mathbb{N}}$ d'éléments disjoints de $\mathscr{M}$, on ait*

$$\mu\Big(\bigcup_{j=1}^{+\infty} A_j\Big) = \sum_{j=1}^{+\infty} \mu(A_j),$$

et telle qu'il existe $A \in \mathscr{M}$ tel que $\mu(A) < +\infty$. Le triplet $(X, \mathscr{M}, \mu)$ est appelé espace mesuré.

Remarques. La série ci-dessus est à termes positifs, sa convergence dans $[0, +\infty]$ est donc toujours assurée. On dit souvent qu'un couple (X, μ) est un espace mesuré, la σ-algèbre étant sous-entendue. Une mesure définie sur les boréliens est appelée *mesure de Borel* ou *mesure borélienne.*

Les propriétés suivantes des mesures positives découlent de leur définition de façon élémentaire : $\mu(\emptyset) = 0$; pour toute famille finie d'ensemble mesurables disjoints $\mu\big(\cup_{j=1}^n A_j\big) = \sum_{j=1}^n \mu(A_j)$; si $A \subset B$, $\mu(A) \leq \mu(B)$; si $A_j \subset A_{j+1}$, $\mu(A_n) \to \mu\big(\cup_{j=1}^\infty A_j\big)$ quand $n \to +\infty$ (cette propriété, qui découle immédiatement de l'additivité dénombrable, est fondamentale pour le théorème de convergence monotone) ; si $A_j \supset A_{j+1}$ et s'il existe i tel que $\mu(A_i) < +\infty$, $\mu(A_n) \to \mu\big(\cap_{j=1}^\infty A_j\big)$ quand $n \to +\infty$.

La mesure que nous utiliserons le plus souvent en analyse est la mesure de Lebesgue sur $\mathbb{R}^d$. Si $P = \prod_{i=1}^d]a_i, b_i[$ est un pavé de $\mathbb{R}^d$, on sait définir son volume d-dimensionnel $\operatorname{vol} P = \prod_{i=1}^d (b_i - a_i)$.

Théorème 5.1.1 *Il existe une mesure positive sur une σ-algèbre $\mathscr{M}$ de $\mathbb{R}^d$ notée $\mathscr{L}_d$ telle que*

i) *$\mathscr{M}$ contient les boréliens.*

ii) *Pour tout pavé P, $\mathscr{L}_d(P) = \operatorname{vol} P$.*

iii) *La mesure $\mathscr{L}_d$ est invariante par translation, i.e., pour tout A dans $\mathscr{M}$ et tout x dans $\mathbb{R}^d$, $A + x \in \mathscr{M}$ et $\mathscr{L}_d(A + x) = \mathscr{L}_d(A)$.*

iv) *Si μ est une mesure de Borel positive invariante par translation et finie sur tout compact, alors il existe $c \geq 0$ telle que $\mu(A) = c\mathscr{L}_d(A)$ pour tout borélien A.*

La démonstration de l'existence (et de l'unicité au sens de iv)) de la mesure de Lebesgue est un peu délicate. Nous ne la donnons pas ici.

Une remarque importante concernant la mesurabilité des fonctions. Dans la pratique, quand on travaille sur $\mathbb{R}^d$ avec la mesure de Lebesgue, il est inutile de s'obnubiler sur la question de la mesurabilité des fonctions auxquelles on a affaire. En effet, étant donnée la richesse de la tribu borélienne sur $\mathbb{R}^d$ (complétée par les ensembles de mesure de Lebesgue nulle) et l'extraordinaire stabilité de la classe des fonctions mesurables par rapport aux opérations que l'on peut appliquer à ces fonctions, on peut dire que, grosso modo, toute fonction qui s'écrit un tant soit peu explicitement est automatiquement mesurable (car par exemple continue, ou limite presque partout d'une suite de fonctions continues, ou composée de fonctions boréliennes, etc.).

En fait, pour être plus précis et pour ceux qui s'intéressent à la théorie des ensembles, on peut montrer que l'existence d'une fonction sur $\mathbb{R}$ non mesurable pour la mesure de Lebesgue est équivalente à l'axiome du choix, ou plus exactement que l'on peut sans contradiction remplacer l'axiome du choix général par une version plus faible de ce même axiome, complété par l'axiome de Solovay qui prescrit que toutes les parties de $\mathbb{R}$ sont Lebesgue-mesurables. Attention toutefois, les boréliens n'épuisent jamais tous les ensembles Lebesgue-mesurables, quels que soient les axiomes de théorie des ensembles que l'on choisisse.

Il convient quand même d'infléchir légèrement les remarques qui précèdent : dans certains contextes et notamment en probabilités, on ne travaille pas forcément sur $\mathbb{R}^d$ ni avec la tribu borélienne, et dans ce cas, les questions de mesurabilité des fonctions deviennent cruciales.

Soit (X,μ) un espace mesuré. L'intégrale de Lebesgue permet d'associer simplement à toute fonction f à valeurs dans $\overline{\mathbb{R}}_+$ et mesurable sur X et à tout ensemble mesurable A, un élément de $\overline{\mathbb{R}}_+$, appelé *intégrale* de f sur A par rapport à la mesure μ et noté $\int_A f\,d\mu$. On commence par le cas des fonctions étagées.

Définition 5.1.5 *Soit $f = \sum_{i=1}^m \lambda_i \mathbf{1}_{E_i}$ une fonction étagée de X dans $\mathbb{R}_+$. On appelle intégrale de f sur A la quantité*

$$\int_A f\,d\mu = \sum_{i=1}^m \lambda_i \mu(E_i \cap A) \in \overline{\mathbb{R}}_+,$$

avec la convention $0\cdot(+\infty) = 0$.

Il est important de noter que dans cette définition, on prend f à valeurs *finies*. Par contre, on peut avoir $\mu(E_i \cap A) = +\infty$ d'où la nécessité de la convention si $\lambda_i = 0$.

Il n'est pas très difficile de montrer que cette quantité ne dépend que de f et non pas de la représentation de f en combinaison linéaire de fonctions caractéristiques choisie (en effet, il y a une infinité de telles représentations pour une même fonction simple). On montre également sans grande difficulté les propriétés fondamentales suivantes :

i) $\int_A f\,d\mu = \int_X f\mathbf{1}_A\,d\mu$;

ii) si $f_1 \leq f_2$ alors $\int_A f_1\, d\mu \leq \int_A f_2\, d\mu$;
iii) pour tout $\lambda \geq 0$, $\int_A \lambda f\, d\mu = \lambda \int_A f\, d\mu$ et
iv) $\int_A (f_1 + f_2)\, d\mu = \int_A f_1\, d\mu + \int_A f_2\, d\mu$.

De plus, à f fixée, on a évidemment que $\varphi_f(A) = \int_A f\, d\mu$ est une mesure positive sur X. Nous pouvons maintenant définir l'intégrale de Lebesgue d'une fonction mesurable à valeurs dans $\overline{\mathbb{R}}_+$.

Définition 5.1.6 *Soit f une fonction mesurable de X dans $\overline{\mathbb{R}}_+$. On appelle intégrale de f sur A la quantité*

$$\int_A f\, d\mu = \sup\left\{ \int_A s\, d\mu; s \textit{ fonction étagée à valeurs dans } \mathbb{R}_+ , s \leq f \right\}.$$

Par définition, cette intégrale est un élément de $\overline{\mathbb{R}}_+$. Elle satisfait visiblement les propriétés i) à iii) et la nouvelle définition coïncide avec la définition précédente quand f est une fonction étagée à valeurs dans $\mathbb{R}_+$.

Dans le cas où $X = \mathbb{R}^d$ et $\mu = \mathscr{L}_d$ est la mesure de Lebesgue, on note l'intégrale simplement $\int_A f\, dx$.

Toute la puissance de l'intégrale de Lebesgue tient au théorème absolument remarquable suivant, le théorème de convergence monotone de Lebesgue.

Théorème 5.1.2 *Soit $(f_n)_{n\in\mathbb{N}}$ une suite croissante de fonctions mesurables sur X à valeurs dans $\bar{\mathbb{R}}_+$. Cette suite converge simplement vers une fonction mesurable f et l'on a*

$$\int_X f_n\, d\mu \to \int_X f\, d\mu \textit{ quand } n \to +\infty.$$

Démonstration. Il est clair qu'une suite croissante de fonctions à valeurs dans $\overline{\mathbb{R}}_+$ converge simplement vers une limite, puisque toute suite croissante dans $\overline{\mathbb{R}}_+$ est convergente. De plus cette limite f est mesurable, comme limite simple de fonctions mesurables. On peut donc définir son intégrale.

De même, la suite $\int_X f_n\, d\mu$ est croissante à valeurs dans $\overline{\mathbb{R}}_+$, donc convergente. Comme $f_n \leq f$, par ii) on a évidemment

$$\lim_{n\to\infty} \int_X f_n\, d\mu \leq \int_X f\, d\mu.$$

Toute la difficulté tient dans l'inégalité inverse.

Pour cela on se donne une fonction étagée s à valeurs dans $\mathbb{R}_+$ telle que $s \leq f$ et un nombre positif $\alpha < 1$. Définissons des ensembles mesurables

$$E_n = \{x \in X; f_n(x) \geq \alpha s(x)\}.$$

Comme la suite f_n est croissante, on a $E_n \subset E_{n+1}$. Soit $x \in X$. Si $s(x) = 0$, alors $x \in E_n$ pour tout n. Si $s(x) > 0$, alors comme $f_n(x) \to f(x) \geq s(x) > \alpha s(x)$, il existe n_0 tel que $x \in E_n$ pour tout $n \geq n_0$. Par conséquent, on vient de montrer que $X = \cup_{n\in\mathbb{N}} E_n$. Or $f_n \geq f_n \mathbf{1}_{E_n}$, donc

$$\int_X f_n \, d\mu \geq \int_{E_n} f_n \, d\mu \geq \int_{E_n} \alpha s \, d\mu = \alpha \varphi_s(E_n),$$

où φ_s est la mesure positive associée à l'intégrale de s. Comme la famille dénombrable E_n est croissante et d'union égale à X, par les propriétés élémentaires des mesures positives, on en déduit que

$$\varphi_s(E_n) \to \varphi_s(X) \text{ quand } n \to +\infty,$$

soit

$$\lim_{n\to\infty} \int_X f_n \, d\mu \geq \alpha \varphi_s(X) = \alpha \int_X s \, d\mu.$$

Prenant d'abord la borne supérieure du membre de droite sur tous les $\alpha < 1$, puis sur toutes les fonctions s, on obtient le théorème. □

Remarque. En particulier, on peut appliquer le théorème de convergence monotone à la suite de fonctions étagées construite au Lemme 5.1.1. Ceci implique d'ailleurs que $\int_X f_1 \, d\mu + \int_X f_2 \, d\mu = \int_X (f_1 + f_2) \, d\mu$, i.e., la propriété iv).

Lemme 5.1.2 (de Fatou) *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables de X dans $\overline{\mathbb{R}}_+$. On a*

$$\int_X (\liminf_{n\to\infty} f_n) \, d\mu \leq \liminf_{n\to\infty} \Big(\int_X f_n \, d\mu \Big).$$

Démonstration. C'est une conséquence immédiate du théorème de convergence monotone. En effet, $\liminf_{n\in\mathbb{N}} f_n = \sup_n \inf_{k\geq n} f_k$ par définition. Or, pour tout n, $g_n = \inf_{k\geq n} f_k \leq f_n$, d'où

$$\int_X g_n \, d\mu \leq \int_X f_n \, d\mu.$$

Mais la suite g_n est croissante et tend simplement vers $\sup_n g_n = \liminf_{n\in\mathbb{N}} f_n$. Donc, par le théorème de convergence monotone, $\int_X g_n \, d\mu \to \int_X (\liminf_{n\to\infty} f_n) \, d\mu$, d'où le résultat en passant à la limite inférieure dans l'inégalité ci-dessus. □

Il est facile de retenir le lemme de Fatou, car celui-ci exprime simplement que l'intégrale est une fonction semi-continue inférieure pour la convergence simple. Une fois que l'on sait définir l'intégrale d'une fonction à valeurs dans $\overline{\mathbb{R}}_+$, on peut définir l'intégrale d'une fonction à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$.

Définition 5.1.7 *On dit qu'une fonction f mesurable de X dans $\mathbb{R}$ ou $\mathbb{C}$ est* intégrable *si*

$$\int_X |f| \, d\mu < +\infty.$$

On pose alors

$$\int_X f\,d\mu = \int_X (\mathfrak{Re}\, f)_+\,d\mu - \int_X (\mathfrak{Re}\, f)_-\,d\mu + i\int_X (\mathfrak{Im}\, f)_+\,d\mu - i\int_X (\mathfrak{Im}\, f)_-\,d\mu.$$

Comme $|f|$ majore chacune des quatre intégrandes du membre de droite, il est bien clair que ces quatre intégrales sont des réels, *i.e.*, ne sont pas infinies. La combinaison linéaire définissant l'intégrale de f a donc bien un sens. Elle n'en aurait pas si on autorisait $\int_X |f|\,d\mu = +\infty$, car on pourrait voir apparaître des formes indéterminées $\infty - \infty$. L'intégrale satisfait les propriétés de linéarité et d'additivité usuelles. De plus, on a l'inégalité d'usage extrêmement fréquent suivante :

$$\left|\int_X f\,d\mu\right| \leq \int_X |f|\,d\mu,$$

(faire bien attention que le membre de gauche n'a de sens que pour les fonctions intégrables, alors que le membre de droite est défini pour toute fonction mesurable). En effet, soit $\theta \in \mathbb{R}$ tel que $\left|\int_X f\,d\mu\right| = e^{i\theta}\int_X f\,d\mu$. Il vient $\left|\int_X f\,d\mu\right| = \int_X e^{i\theta} f\,d\mu = \int_X \mathfrak{Re}(e^{i\theta} f)\,d\mu \leq \int_X |e^{i\theta} f|\,d\mu = \int_X |f|\,d\mu$.

Le théorème de convergence dominée de Lebesgue qui suit est l'un des plus importants en analyse.

Théorème 5.1.3 *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables sur X à valeurs dans $\mathbb{C}$ qui converge simplement*

$$f_n(x) \to f(x) \text{ quand } n \to +\infty,$$

et telle qu'il existe une fonction g sur X à valeurs dans $\overline{\mathbb{R}}_+$ intégrable telle que l'on ait

$$\forall x \in X, \quad |f_n(x)| \leq g(x),$$

alors f est intégrable et l'on a

$$\int_X |f_n - f|\,d\mu \to 0 \text{ quand } n \to +\infty.$$

En particulier,

$$\int_X f_n\,d\mu \to \int_X f\,d\mu.$$

Démonstration. C'est une conséquence immédiate du lemme de Fatou. On sait que f est mesurable et la majoration uniforme des f_n par g montre qu'elle est intégrable. Comme $|f_n - f| \leq 2g$, on voit que $2g - |f_n - f| \geq 0$ et $\lim_{n\to+\infty}(2g - |f_n - f|) = 2g$. Par le lemme de Fatou, il vient

$$\int_X 2g\,d\mu \leq \liminf_{n\to+\infty}\Big(\int_X (2g - |f_n - f|)\,d\mu\Big) = \int_X 2g\,d\mu - \limsup_{n\to+\infty}\Big(\int_X |f_n - f|\,d\mu\Big).$$

Comme $\int_X 2g\,d\mu < +\infty$, on en déduit que

$$\limsup_{n\to+\infty}\Big(\int_X |f_n - f|\,d\mu\Big) \le 0,$$

et une suite positive dont la limite supérieure est négative tend bien sûr vers 0. On remarque ensuite que

$$\Big|\int_X f_n\,d\mu - \int_X f\,d\mu\Big| = \Big|\int_X (f_n - f)\,d\mu\Big| \le \int_X |f_n - f|\,d\mu \to 0.$$

□

Définition 5.1.8 *On dit qu'une propriété a lieu* μ-presque partout *ou* pour μ-presque tout *x, si elle a lieu sauf sur un ensemble de μ-mesure nulle.*

Les ensembles de μ-mesure nulle sont appelés ensembles μ-*négligeables*. La terminologie vient du fait que les ensembles μ-négligeables « ne sont pas vus » par la mesure μ. Attention, cela ne signifie pas nécessairement qu'ils soient « petits », tout dépend de la mesure μ. Par exemple, si $\mu = \delta_0$, la masse de Dirac en 0 sur $\mathbb{R}$ définie par $\mu(A) = 1$ si $0 \in A$ et $\mu(A) = 0$ si $0 \notin A$, alors $\mathbb{R}^*$ est μ-négligeable. Si μ est la mesure de Lebesgue sur $\mathbb{R}$, alors $\mathbb{Q}$ est μ-négligeable. Si μ est la mesure de comptage, $\mu(A) = \operatorname{Card} A$, alors le seul ensemble μ-négligeable est l'ensemble vide.

Proposition 5.1.4 *La relation d'égalité μ-presque partout est une relation d'équivalence sur l'ensemble des fonctions mesurables.*

Démonstration. C'est presque évident. Notons $f \sim_\mu g$ si $f = g$ μ-presque partout, c'est-à-dire s'il existe un ensemble N tel que $\mu(N) = 0$ et tel que $f(x) = g(x)$ pour tout $x \notin N$. Clairement, $f \sim_\mu f$ en prenant $N = \emptyset$; si $f \sim_\mu g$, alors $g \sim_\mu f$; si $f \sim_\mu g$ et $g \sim_\mu h$, notant N_1 le premier ensemble négligeable et N_2 le deuxième ensemble négligeable, on voit que $f(x) = g(x) = h(x)$ pour tout $x \notin N_1 \cup N_2$ et $\mu(N_1 \cup N_2) \le \mu(N_1) + \mu(N_2) = 0$, donc $f \sim_\mu h$. □

Attention, dire que deux fonctions sont égales μ-presque partout ne signifie pas nécessairement qu'elles se ressemblent beaucoup, tout dépend encore une fois de la mesure μ. Par exemple, si $\mu = \delta_0$, alors $f \sim_\mu g$ si et seulement si $f(0) = g(0)$. Si μ est la mesure de Lebesgue, alors $\mathbf{1}_{\mathbb{Q}} \sim_\mu 0$. Si μ est la mesure de comptage, alors $f \sim_\mu g$ si et seulement si $f = g$.

Cette relation d'équivalence définit une partition de l'ensemble des fonctions mesurables en classes d'équivalence. Du point de vue de l'intégrale, les éléments d'une même classe d'équivalence sont indiscernables au sens suivant : si F est une fonction continue de $\mathbb{C}$ dans $\overline{\mathbb{R}}$ et si $f \sim_\mu g$, alors

$$\int_X F \circ f \, d\mu = \int_X F \circ g \, d\mu.$$

En effet, $\left|\int_X (F \circ f - F \circ g)\, d\mu\right| = \left|\int_N (F \circ f - F \circ g)\, d\mu\right| \leq \mu(N) \sup_N |F \circ f - F \circ g| = 0$ avec la convention $0 \cdot \infty = 0$. De telles quantités intégrales ne dépendent donc que de la classe d'équivalence, et non pas d'un représentant particulier de cette classe. Elles passent au quotient par la relation d'équivalence, c'est-à-dire sont en fait définies sur l'ensemble des classes d'équivalence modulo l'égalité μ-presque partout.

Convention importante. On peut en général sans grand danger confondre une fonction mesurable sur X et sa classe d'équivalence modulo l'égalité μ-presque partout, et l'on parle couramment de « fonction », alors que l'objet dont on parle est en réalité une classe d'équivalence. Cet abus de langage permet d'éviter une certaine lourdeur, il est pratiqué de façon quasi-systématique mais il convient de garder en arrière-plan de son esprit que c'est bien un abus de langage. En effet, dans certaines situations très particulières, il peut advenir que l'on soit contraint de considérer non pas une classe d'équivalence, mais un de ses représentants. Ce ne sera pas le cas dans ce cours. Une autre facette de cette convention est que l'on identifie systématiquement les fonctions continues à leur classe d'équivalence modulo l'égalité μ-presque partout. C'est en ce sens que l'on pourra éventuellement parler d'injection de l'espace des fonctions continues dans un espace de type L^p (au moins pour certaines mesures).

Remarque tout aussi importante. Le lemme de Fatou et les théorèmes de convergence monotone et dominée restent valables si l'on remplace les diverses limites et majorations ponctuelles par des limites et des majorations μ-presque partout.

L'inégalité de Hölder joue un rôle crucial dans la théorie des espaces L^p. Dans toute la suite du cours, nous prendrons la convention suivante :

$$\text{Si } p \in \,]1, \infty[, \; p' \stackrel{\text{déf}}{=} \frac{p}{p-1}, \text{ si } p = 1, \; p' \stackrel{\text{déf}}{=} +\infty, \text{ et si } p = +\infty, \; p' \stackrel{\text{déf}}{=} 1.$$

Les exposants p et p' sont dits conjugués et, en convenant que $1/+\infty = 0$, on a dans tous les cas

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

Lemme 5.1.3 *Soit f une fonction mesurable sur X à valeurs dans $\overline{\mathbb{R}}_+$ telle que $\int_X f \, d\mu = 0$. Alors $f = 0$ μ-presque partout.*

Démonstration. Soient $E = \{x \in X; f(x) > 0\}$ et $E_n = \{x \in X; f(x) \geq 1/n\}$. Clairement, $E_n \subset E_{n+1}$ et $E = \cup_{n\in\mathbb{N}^*} E_n$. Par les propriétés élémentaires des mesures, on en déduit que $\mu(E) = \lim_{n\to+\infty} \mu(E_n)$. Or

$$\mu(E_n) = \int_{E_n} d\mu \leq n \int_{E_n} f\,d\mu \leq n \int_X f\,d\mu = 0.$$

Donc $\mu(E) = 0$, ce qui signifie bien que $f \sim_\mu 0$. □

Voici l'énoncé de l'inégalité de Hölder.

Théorème 5.1.4 *Soient (X,μ) un espace mesuré, f et g deux fonctions mesurables sur X à valeurs dans $\overline{\mathbb{R}}_+$. Alors pour tout $1 < p < +\infty$, on a*

$$\int_X fg\,d\mu \leq \Big(\int_X f^p\,d\mu\Big)^{1/p} \Big(\int_X g^{p'}\,d\mu\Big)^{1/p'}.$$

Démonstration. Soit $A = \big(\int_X f^p\,d\mu\big)^{1/p}$ et $B = \big(\int_X g^{p'}\,d\mu\big)^{1/p'}$. Si $A = 0$, alors $f = 0$ μ-presque partout, donc fg aussi et l'inégalité est vérifiée. Si $A > 0$ et $B = +\infty$, il n'y a rien à démontrer. Plaçons-nous donc dans le cas où $0 < A, B < +\infty$ et posons $F = f/A$ et $G = g/B$. Il vient immédiatement

$$\int_X F^p\,d\mu = \int_X G^{p'}\,d\mu = 1.$$

Pour tout $x \in X$ tel que $0 < F(x) < +\infty$ et $0 < G(x) < +\infty$, on pose

$$s = p \ln F(x) \text{ et } t = p' \ln G(x).$$

Comme $1/p + 1/p' = 1$, la convexité de la fonction exponentielle implique que

$$e^{\frac{s}{p}+\frac{t}{p'}} \leq \frac{1}{p}e^s + \frac{1}{p'}e^t,$$

c'est-à-dire

$$F(x)G(x) \leq \frac{1}{p}F(x)^p + \frac{1}{p'}G(x)^{p'}.$$

Cette inégalité reste trivialement vraie si au moins un de ses termes vaut 0 ou $+\infty$. L'intégrant sur X, on obtient

$$\int_X FG\,d\mu \leq \frac{1}{p}\int_X F^p\,d\mu + \frac{1}{p'}\int_X G^{p'}\,d\mu = 1$$

qui n'est autre que l'inégalité de Hölder. □

Commentaires. Remarquons que chacun des termes de cette inégalité peut prendre la valeur $+\infty$.

Une conséquence importante de l'inégalité de Hölder est l'inégalité de Minkowski.

Corollaire 5.1.1 *Soient (X,μ) un espace mesuré, f et g deux fonctions mesurables sur X à valeurs dans $\overline{\mathbb{R}}_+$. Alors pour tout $1 < p < +\infty$, on a*

$$\Big(\int_X (f+g)^p \, d\mu\Big)^{1/p} \le \Big(\int_X f^p \, d\mu\Big)^{1/p} + \Big(\int_X g^p \, d\mu\Big)^{1/p}.$$

Démonstration. Il suffit d'appliquer l'inégalité de Hölder à la fonction $(f+g)^p = f(f+g)^{p-1} + g(f+g)^{p-1}$. □

Nous pouvons maintenant définir les espaces L^p.

Définition 5.1.9 *Pour tout $p \in [1,+\infty[$ et toute fonction f mesurable sur X à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$, on pose*

$$\|f\|_{L^p(X,d\mu)} = \Big(\int_X |f|^p \, d\mu\Big)^{1/p}.$$

Soit $S_f = \{\alpha \in \mathbb{R}; \mu(|f|^{-1}(]\alpha,+\infty])=0\}$. On pose

$$\|f\|_{L^\infty(X,d\mu)} = \inf S_f \text{ si } S_f \neq \emptyset \text{ et } \|f\|_{L^\infty(X,d\mu)} = +\infty \text{ si } S_f = \emptyset.$$

La signification de la quantité $\|f\|_{L^\infty(X,d\mu)}$ quand elle est finie est que $|f|$ ne peut dépasser strictement cette valeur que sur un ensemble de mesure nulle, mais que $|f|$ dépasse toute valeur qui lui est strictement inférieure sur un ensemble de mesure strictement positive. Quand elle est infinie, $|f|$ prend des valeurs aussi grandes que l'on veut sur des ensembles de mesure strictement positive. C'est pourquoi on l'appelle aussi la borne supérieure essentielle de $|f|$, $\sup\mathrm{ess}\,|f|$. Il est facile de voir que quand S_f n'est pas vide, $\|f\|_{L^\infty(X,d\mu)} \in S_f$ et que $|f(x)| \le \lambda$ μ-presque partout si et seulement si $\lambda \ge \|f\|_{L^\infty(X,d\mu)}$.

Définition 5.1.10 *Pour tout $p \in [1,+\infty]$ on appelle $L^p(X,d\mu)$ l'ensemble des (classes de) fonctions mesurables sur X à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$ telles que*

$$\|f\|_{L^p(X,d\mu)} < +\infty.$$

L'ensemble $L^p(X,d\mu)$ est un espace vectoriel, il est normé par $\|f\|_{L^p(X,d\mu)}$ et complet pour cette norme.

Démonstration. Le fait qu'il s'agisse d'un espace vectoriel découle directement de la définition dans les cas $p = 1$ et $p = +\infty$ et de l'inégalité de Minkowski pour $1 < p < +\infty$. Le fait que l'on ait affaire à une norme découle du lemme 5.1.3 et de l'inégalité de Minkowski et de l'inégalité triangulaire dans $\mathbb{R}$ ou $\mathbb{C}$.

Montrons rapidement que $L^p(X,d\mu)$ est complet. Supposons d'abord $1 \leq p < +\infty$. On se donne une suite de Cauchy f_n dans $L^p(X,d\mu)$. Pour tout $\varepsilon > 0$, il existe n_0 tel que pour tout $n,m \geq n_0$, $\|f_n - f_m\|_{L^p(X,d\mu)} \leq \varepsilon$. Prenant $\varepsilon = 2^{-k}$ pour $k \in \mathbb{N}$, on extrait donc une sous-suite f_{n_k} telle que $\|f_{n_k} - f_{n_{k+1}}\|_{L^p(X,d\mu)} \leq 2^{-k}$. On considère la fonction

$$g(x) = \sum_{k=0}^{\infty} |f_{n_{k+1}} - f_{n_k}|(x).$$

Cette fonction est bien définie μ-presque partout comme série à termes positifs et prend ses valeurs dans $[0,+\infty]$. De plus, c'est une limite ponctuelle de fonctions mesurables, elle est donc mesurable. Comme $\|g\|_{L^p(X,d\mu)} \leq \sum_{k=0}^{\infty} 2^{-k} \leq 2$ (Minkowski + Fatou), g appartient en fait à $L^p(X,d\mu)$. En particulier g est finie μ-presque partout.

Par conséquent, il existe un ensemble μ-négligeable N tel que $g(x) < +\infty$ si $x \notin N$ et pour x en dehors de cet ensemble la série

$$u(x) = \sum_{k=0}^{\infty} (f_{n_{k+1}} - f_{n_k})(x)$$

est absolument convergente. Elle définit donc une fonction mesurable et, comme $|u(x)| \leq g(x)$, appartenant à $L^p(X,d\mu)$. Pour conclure, il suffit de montrer que $f_{n_k} \to (u + f_{n_0})$ dans $L^p(X,d\mu)$ quand $k \to +\infty$. Or $f_{n_k} - f_{n_0} = \sum_{l=0}^{k-1}(f_{n_{l+1}} - f_{n_l})(x)$, donc $f_{n_k} - (u + f_{n_0}) \to 0$ μ-presque partout. De plus, $|f_{n_k} - (u + f_{n_0})|^p \leq (|u| + |g|)^p$ avec u et g dans $L^p(X,d\mu)$. Le théorème de convergence dominée de Lebesgue montre donc que $\|f_{n_k} - (u + f_{n_0})\|_{L^p(X,d\mu)} \to 0$ quand $k \to +\infty$.

Le cas $p = +\infty$ est plus simple. Soit A_k et $B_{m,n}$ les ensembles sur lesquels $|f_k(x)| > \|f_k\|_{L^\infty(X,d\mu)}$ et $|f_m(x) - f_n(x)| > \|f_m - f_n\|_{L^\infty(X,d\mu)}$ respectivement. Ces ensembles sont de mesure nulle et sont en nombre dénombrable. Par conséquent, leur réunion E est de mesure nulle. En dehors de E, les suites $f_n(x)$ sont de Cauchy dans $\mathbb{R}$ ou $\mathbb{C}$ et bornées par $\sup \|f_k\|_{L^\infty(X,d\mu)}$. Elles y convergent donc vers une fonction f bornée par $\sup \|f_k\|_{L^\infty(X,d\mu)}$. Si on définit $f(x) = 0$ pour $x \in E$, alors $f \in L^\infty(X,d\mu)$ et $\|f_n - f\|_{L^\infty(X,d\mu)} \to 0$ par construction. $\square$

Remarque. Quand $X = \mathbb{N}$ et μ est la mesure de comptage, alors $L^p(X,d\mu) = \ell^p$ que l'on retrouve donc comme cas particulier.

On déduit facilement de l'inégalité de Hölder le

Corollaire 5.1.2 *Soient* (X,μ) *un espace mesuré,* $p \in [1,+\infty]$, 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*

$$\int_X |fg|d\mu \leq \|f\|_{L^p}\|g\|_{L^{p'}}.$$

On en déduit aussi facilement le corollaire suivant.

Corollaire 5.1.3 *Soient* p *et* q *deux éléments de* $[1,+\infty]$ *tels que*

$$\frac{1}{p}+\frac{1}{q} \leq 1$$

et définissons $r \in [1,+\infty]$ *par*

$$\frac{1}{r}=\frac{1}{p}+\frac{1}{q}\cdot$$

Alors l'application

$$\begin{cases} L^p \times L^q & \rightarrow L^r \\ (f,g) & \mapsto fg \end{cases}$$

est une application bilinéaire continue.

Démonstration. Le résultat est évident si $p=q=r=+\infty$. Sinon on peut supposer que $p<+\infty$ et l'on applique l'inégalité de Hölder avec $s=p/r$. Il vient

$$\int_X |fg|^r d\mu \leq \|f\|_{L^p}^r\|g\|_{L^q}^r,$$

ce qui signifie que

$$\|fg\|_{L^r} \leq \|f\|_{L^p}\|g\|_{L^q},$$

d'où le corollaire. □

On peut réécrire le théorème de convergence dominée de Lebesgue dans le contexte des espaces L^p. La démonstration est immédiate.

Théorème 5.1.5 *Soient* $1 \leq p < +\infty$ *et* $(f_n)_{n\in\mathbb{N}}$ *une suite de fonctions de* $L^p(X,d\mu)$*. Si, pour* μ*-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* μ*-presque tout* x *de* X*, on ait*

$$|f_n(x)| \leq g(x),$$

alors, on a

$$f \in L^p \quad \textit{et} \quad \lim_{n\to\infty}\|f_n-f\|_{L^p}=0.$$

Corollaire 5.1.4 *L'espace $L^\infty(X,d\mu)\cap L^p(X,d\mu)$ est dense dans $L^p(X,d\mu)$.*

Démonstration. En effet, posons

$$f_n = \mathbf{1}_{|f|\leq n} f.$$

Comme une fonction de L^p est finie μ-presque partout, 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 convergence dominée assure alors le résultat. □

Le théorème de convergence dominée admet une réciproque partielle au sens du théorème suivant.

Théorème 5.1.6 *Soit $p\in[1,+\infty]$, on considère une suite $(f_n)_{n\in N}$ de $L^p(X,d\mu)$ qui converge dans $L^p(X,d\mu)$ vers une fonction f. Il existe une sous-suite n_p et une fonction $g\in L^p(X,d\mu)$ telles que*

$$\left.\begin{array}{r}\lim_{p\to\infty} f_{n_p}(x) = f(x),\\ |f_{n_p}(x)| \leq g(x),\end{array}\right\} \ \mu\text{-presque partout par rapport à } x.$$

Démonstration. Ceci est un sous-produit de la démonstration du fait que $L^p(X,d\mu)$ est complet. □

On supposera dans la suite, comme il est classique, que la mesure μ est σ-*finie*, c'est-à-dire que X est une réunion dénombrable d'ensembles de mesure finie. C'est le cas par exemple de la mesure de Lebesgue, puisque $\mathbb{R}^d$ est σ-compact, *i.e.*, réunion dénombrable de compacts, et que chaque compact est de mesure de Lebesgue finie.

L'inégalité de Hölder est fondamentale. Nous allons voir qu'elle est optimale au sens suivant.

Lemme 5.1.4 *Soit (X,μ) un espace mesuré. Soit f une fonction mesurable et positive et p un élément de $[1,+\infty]$. On a alors*

$$\left(\int_X f^p\,d\mu\right)^{1/p} = \sup_{g\geq 0, \|g\|_{L^{p'}}\leq 1} \int_X fg\,d\mu$$

pour $p<+\infty$ et

$$\sup\operatorname{ess} f = \sup_{g\geq 0, \|g\|_{L^1}\leq 1} \int_X fg\,d\mu$$

pour $p = +\infty$. Si de plus la fonction f appartient à L^p, alors

$$\|f\|_{L^p} = \sup_{\|g\|_{L^{p'}} \leq 1} \Big| \int_X fg\, d\mu \Big|.$$

Démonstration. Soit f mesurable sur X à valeurs dans $[0, +\infty]$. Si $f = 0$, il n'y a rien à démontrer. Supposons donc $f \neq 0$. D'après l'inégalité de Hölder, on a toujours

$$\Big(\int_X f^p\, d\mu\Big)^{1/p} \geq \sup_{\|g\|_{L^{p'}} \leq 1} \int_X fg\, d\mu \quad \text{et} \quad \sup\text{ess} f \geq \sup_{\|g\|_{L^1} \leq 1} \int_X fg\, d\mu$$

Commençons par démontrer le lemme dans le cas où $p = +\infty$. Pour tout $\lambda > 0$, on pose $E_\lambda \stackrel{\text{déf}}{=} \{x \in X; f(x) \geq \lambda\}$. Soit λ tel que

$$\mu(E_\lambda) > 0.$$

Prenons un ensemble mesurable $F_\lambda \subset E_\lambda$ tel que $0 < \mu(F_\lambda) < +\infty$ (il en existe un puisque μ est supposée être σ-finie) et soit $g_\lambda = \mu(F_\lambda)^{-1}\mathbf{1}_{F_\lambda}$. C'est une fonction de L^1, positive, supportée dans E_λ, et d'intégrale 1. On a alors

$$\int_X fg_\lambda\, d\mu = \int_{E_\lambda} fg_\lambda\, d\mu \geq \lambda \int_{E_\lambda} g_\lambda\, d\mu = \lambda,$$

d'où le lemme dans ce cas.

Supposons maintenant que p soit réel. Remarquons tout d'abord que si $f = +\infty$ sur un ensemble de mesure strictement positive, alors l'égalité est trivialement satisfaite. On peut donc supposer que f est finie presque partout. Considérons alors une suite $(E_n)_{n\in\mathbb{N}}$ croissante d'ensembles de mesure finie dont la réunion est X (toujours car μ est σ-finie) et posons d'abord

$$f_n = \mathbf{1}_{E_n \cap \{f \leq n\}} f.$$

Il est clair que la fonction f_n appartient à $L^1 \cap L^\infty$ donc à L^p pour tout $p \in [1, +\infty[$. On peut donc définir alors

$$g_n = \mathbf{1}_{E_n \cap \{f \leq n\}} \|f_n\|_{L^p}^{-\frac{p}{p'}} f_n^{p-1}$$

(le premier facteur dans l'expression de g_n étant inutile pour $p > 1$). Il est également clair que $g_n^{p'} = \|f_n\|_{L^p}^{-p} f_n^p$ pour $p > 1$ et $g_n = \mathbf{1}_{E_n \cap \{f \leq n\}}$ pour $p = 1$, d'où $\|g_n\|_{L^{p'}} = 1$ (au moins pour n assez grand dans le deuxième cas).

La définition des fonctions f_n et g_n montre que $fg_n = f_n g_n = \|f_n\|_{L^p}^{-\frac{p}{p'}} f_n^p$, donc

$$\int_X fg_n\,d\mu = \Big(\int_X f_n^p\,d\mu\Big)\|f_n\|_{L^p}^{-\frac{p}{p'}} = \|f_n\|_{L^p}^{p-\frac{p}{p'}} = \|f_n\|_{L^p}.$$

On en déduit que

$$\int_X f_n^p\,d\mu \leq \Big(\sup_{g\geq 0, \|g\|_{L^{p'}}\leq 1}\int_X fg\,d\mu\Big)^p.$$

Comme f est finie presque partout, le théorème de convergence monotone de Lebesgue appliqué à la suite croissante $(f_n^p)_{n\in\mathbb{N}}$ implique alors immédiatement que

$$\int_X f^p\,d\mu \leq \Big(\sup_{g\geq 0, \|g\|_{L^{p'}}\leq 1}\int_X fg\,d\mu\Big)^p,$$

d'où le lemme pour $f \geq 0$.

Soit maintenant f à valeurs complexes et appartenant à L^p. On pose $\mathrm{sgn} f(x) = \bar{f}(x)/|f(x)|$ si $f(x) \neq 0$ et $\mathrm{sgn} f(x) = 1$ sinon. Pour tout $\tilde{g} \in L^{p'}$ à valeurs positives, on pose également $g = \mathrm{sgn} f \tilde{g}$. Il vient $|g| = \tilde{g}$ et $fg = |f|\tilde{g}$. Par conséquent,

$$\sup_{\|g\|_{L^{p'}}\leq 1}\Big|\int_X fg\,d\mu\Big| \geq \sup_{\tilde{g}\geq 0, \|\tilde{g}\|_{L^{p'}}\leq 1}\int_X |f|\tilde{g}\,d\mu,$$

ce qui implique d'après l'étude du cas positif que

$$\sup_{\|g\|_{L^{p'}}\leq 1}\Big|\int_X fg\,d\mu\Big| \geq \Big(\int_X |f|^p\,d\mu\Big)^{1/p} = \|f\|_{L^p},$$

d'où le lemme. □

Noter les précautions prises dans le cas positif pour prendre en compte la possibilité que les intégrales prennent la valeur $+\infty$.

Nous rappelons maintenant sans démonstration quelques théorèmes utiles.

Théorème 5.1.7 (de dérivation sous le signe somme) *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*

$$z \mapsto f(z,x)$$

est différentiable sur Ω, si pour tout z de Ω,

$$x \mapsto f(z,x) \quad \textit{et} \quad x \mapsto D_z f(z,x)$$

sont intégrables sur X et si enfin, pour presque tout x de X et pour tout z de Ω, on a

$$|D_z f(z,x)| \leq g(x) \quad avec \quad g \in L^1(X, d\mu),$$

alors l'application

$$z \mapsto \int_X f(z,x)\, d\mu(x)$$

est différentiable sur Ω et l'on a

$$D\Big(\int_X f(z,x)\, d\mu(x)\Big) = \int_X D_z f(z,x)\, d\mu(x).$$

C'est une conséquence du théorème de convergence dominée.

Théorème 5.1.8 (de Fubini positif) *Soient (X_1, μ_1) et (X_2, μ_2) deux espaces mesurés σ-finis et F une fonction mesurable positive de $X_1 \times X_2$ dans $[0, +\infty]$. On a*

$$\begin{aligned}\int_{X_1 \times X_2} F(x_1,x_2)\, d\mu_1(x_1) \otimes d\mu_2(x_2) &= \int_{X_1} \Big(\int_{X_2} F(x_1,x_2)\, d\mu_2(x_2)\Big) d\mu_1(x_1) \\ &= \int_{X_2} \Big(\int_{X_1} F(x_1,x_2)\, d\mu_1(x_1)\Big) d\mu_2(x_2).\end{aligned}$$

Notons qu'une bonne partie de la difficulté que nous éludons ici réside dans la définition de la mesure produit $\mu_1 \otimes \mu_2$ et d'une σ-algèbre adéquate sur $X_1 \times X_2$. 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*

$$\|f\|_{L^p}^p = p \int_0^{+\infty} \lambda^{p-1} \mu(|f| > \lambda) d\lambda.$$

Théorème 5.1.9 (de Fubini) *Soient (X_1, μ_1) et (X_2, μ_2) deux espaces mesurés σ-finis et F une fonction mesurable sur $X_1 \times X_2$. Si la fonction F appartient à $L^1(X_1 \times X_2, d\mu_1 \otimes d\mu_2)$, alors*

i) *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 $\int_{X_2} F(x_1,x_2)\, d\mu_2(x_2)$ appartient à $L^1(X_1, d\mu_1)$ et*

$$\Big\| \int_{X_2} F(x_1,x_2)\, d\mu_2(x_2) \Big\|_{L^1(X_1, d\mu_1)} \leq \|F\|_{L^1(X_1 \times X_2, d\mu_1 \otimes d\mu_2)}.$$

ii) *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 $\int_{X_1} F(x_1,x_2)\, d\mu_1(x_1)$ appartient à $L^1(X_2, d\mu_2)$ et*

$$\Big\| \int_{X_1} F(x_1,x_2)\, d\mu_1(x_1) \Big\|_{L^1(X_2, d\mu_2)} \leq \|F\|_{L^1(X_1 \times X_2, d\mu_1 \otimes d\mu_2)}.$$

iii) *Enfin, dans ce cas on a*

$$\begin{aligned}\int_{X_1\times X_2} F(x_1,x_2)\,d\mu_1(x_1)\otimes d\mu_2(x_2) &= \int_{X_1}\Big(\int_{X_2} F(x_1,x_2)\,d\mu_2(x_2)\Big)d\mu_1(x_1)\\ &= \int_{X_2}\Big(\int_{X_1} F(x_1,x_2)\,d\mu_1(x_1)\Big)d\mu_2(x_2).\end{aligned}$$

Mentionnons enfin un dernier théorème classique, très important en théorie de la mesure.

Théorème 5.1.10 (de Radon-Nykodym) *Soit $(X,\mathscr{M})$ un espace mesurable et μ et ν deux mesures positives sur $\mathscr{M}$ telles que $\nu(X)<+\infty$ et μ est σ-finie. Si pour tout $E\in\mathscr{M}$ tel que $\mu(E)=0$ on a $\nu(E)=0$, alors il existe une unique fonction $h\in L^1(X,d\mu)$ positive telle que*

$$\forall A\in\mathscr{M},\quad \nu(A)=\int_A h\,d\mu.$$

La fonction h est appelée dérivée de Radon-Nykodym *de la mesure ν par rapport à la mesure μ.*

Démonstration. L'unicité de h est évidente. Démontrons en l'existence en commençant par le cas où $\mu(X)<+\infty$. On pose $\lambda=\mu+\nu$, c'est une mesure positive finie sur $\mathscr{M}$. Considérons la forme linéaire sur $L^2(X,d\lambda)$

$$\ell(f)=\int_X f\,d\nu.$$

Cette forme linéaire est continue sur $L^2(X,d\lambda)$, car par l'inégalité de Cauchy-Schwarz,

$$\begin{aligned}|\ell(f)| &\le \Big(\int_X f^2\,d\nu\Big)^{1/2}\Big(\int_X d\nu\Big)^{1/2}\\ &\le \Big(\int_X f^2\,d\lambda\Big)^{1/2}(\nu(X))^{1/2}=\|f\|_{L^2(X,d\lambda)}(\nu(X))^{1/2}.\end{aligned}$$

Par le théorème de représentation de Riesz 4.3.1, il existe donc une fonction $g\in L^2(X,d\lambda)$ telle que

$$\forall f\in L^2(X,d\lambda),\quad \ell(f)=\int_X gf\,d\lambda.$$

En utilisant la définition de λ, cette égalité se réécrit sous la forme suivante :

$$\forall f\in L^2(X,d\lambda),\quad \int_X(1-g)f\,d\nu=\int_X gf\,d\mu. \tag{5.1}$$

En particulier, pour tout $A \in \mathscr{M}$, on peut prendre $f = \mathbf{1}_A$, d'où

$$\int_A (1-g)\, d\nu = \int_A g\, d\mu.$$

On déduit immédiatement de cette égalité que $0 \leq g \leq 1$ μ- et ν-presque partout. On peut donc sans modifier aucune des intégrales supposer que $0 \leq g \leq 1$ partout. Ceci nous amène à introduire les ensembles

$$R = \{x \in X; g(x) < 1\} \quad \text{et} \quad S = \{x \in X; g(x) = 1\},$$

de telle sorte que $X = R \cup S$, $R \cap S = \emptyset$ et $\mu(S) = 0$. Posant alors

$$\nu_r(A) = \nu(A \cap R) \quad \text{et} \quad \nu_s(A) = \nu(A \cap S),$$

on obtient une décomposition de la mesure ν

$$\nu = \nu_r + \nu_s. \tag{5.2}$$

Comme $g \in L^\infty(X, d\lambda)$, on peut prendre $f = \left(\sum_{i=0}^k g^i\right)\mathbf{1}_A$ dans l'égalité (5.1), ce qui donne

$$\int_{A\cap R} (1-g^{k+1})\, d\nu = \int_A \left(\sum_{i=1}^{k+1} g^i\right) d\mu.$$

Sur l'ensemble R, $1 - g^{k+1}(x) \to 1$ quand $k \to +\infty$, donc par le théorème de convergence dominée, on voit que

$$\int_{A\cap R} (1-g^{k+1})\, d\nu \to \nu_r(A) \text{ quand } k \to +\infty.$$

Comme g est positive, la suite de fonctions $h_k = \sum_{i=1}^{k+1} g^i$ est croissante. Elle converge donc vers une fonction h mesurable à valeurs dans $[0, +\infty]$ et par le théorème de convergence monotone, on voit que

$$\int_A h_k\, d\mu \to \int_A h\, d\mu \text{ quand } k \to +\infty.$$

On a donc montré que pour tout $A \in \mathscr{M}$, on a

$$\nu_r(A) = \int_A h\, d\mu.$$

Prenant $A = X$, on en déduit que $h \in L^1(X, d\mu)$ et par conséquent

$$\forall A \in \mathscr{M}, \quad \nu(A) = \int_A h\, d\mu + \nu_s(A).$$

Comme $\mu(A\cap S)=0$, on a $\nu_s(A)=\nu(A\cap S)=0$, d'où le théorème quand μ est finie.

Soit pour finir $(E_n)_{n\in\mathbb{N}}$ une suite croissante d'ensembles de μ-mesure finie dont la réunion est X. Pour chaque n, il existe une unique fonction $h_n\in L^1(E_n,d\mu)$ telle que

$$\forall A\in\mathscr{M},\quad \nu(A\cap E_n)=\int_{A\cap E_n} h_n\,d\mu.$$

L'unicité de la dérivée de Radon-Nykodym implique immédiatement que si $n\geq m$, alors $(h_n)_{|E_m}=h_m$. Il existe donc une fonction mesurable positive h sur X telle que $h_n=\mathbf{1}_{E_n}h$, si bien que l'égalité ci-dessus devient

$$\forall A\in\mathscr{M},\quad \nu(A\cap E_n)=\int_A \mathbf{1}_{E_n}h\,d\mu.$$

Le théorème de convergence monotone permet de passer à la limite dans les deux membres quand $n\to+\infty$,

$$\forall A\in\mathscr{M},\quad \nu(A)=\int_A h\,d\mu,$$

et le choix $A=X$ montre enfin que $h\in L^1(X,d\mu)$. □

Remarques. Quand l'hypothèse $\mu(E)=0\Rightarrow\nu(E)=0$ a lieu, on dit que ν est *absolument continue* par rapport à μ. En l'absence de cette hypothèse, on a démontré en fait un petit peu plus que ce qui est annoncé dans l'énoncé du théorème, à savoir l'existence et l'unicité de la *décomposition de Lebesgue* (5.2) d'une mesure positive finie par rapport à une mesure positive σ-finie. La mesure ν_s est appelée *partie singulière* de ν par rapport à μ, la mesure ν_r est sa partie absolument continue. Remarquons que si ν est également σ-finie, on obtient encore une dérivée de Radon-Nykodym h, mais celle-ci n'est plus dans $L^1(X,d\mu)$. Enfin, si une des deux mesures n'est pas σ-finie, le théorème de Radon-Nykodym peut être mis en défaut.

On a jusqu'à présent seulement considéré des mesures positives. On peut tout aussi bien considérer des mesures réelles ou complexes, c'est-à-dire des fonctions dénombrablement additives sur $\mathscr{M}$ à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$, les séries qui interviennent dans cette définition étant absolument convergentes. Si ν est une telle mesure, on définit sa *variation totale* $|\nu|$ par $|\nu(A)|=\sup\bigl(\sum_{i=0}^{\infty}|\nu(A_i)|\bigr)$, le sup étant pris sur toutes les partitions mesurables dénombrables de A. On démontre alors que $|\nu|$ est une mesure positive finie. Ceci conduit à la *décomposition de Jordan* d'une mesure réelle, en posant

$$\nu_+=\frac{1}{2}(|\nu|+\nu)\text{ et }\nu_-=\frac{1}{2}(|\nu|-\nu)\text{ si bien que }\nu=\nu_+-\nu_-\text{ et }|\nu|=\nu_++\nu_-.$$

Les mesures ν_+ et ν_- sont deux mesures positives finies appelées *variations positive et négative* de ν. Une mesure complexe se décompose en parties réelle et imaginaire qui toutes deux admettent une décomposition de Jordan, soit au total quatre mesures positives finies. Vu les considérations précédentes, le théorème de Radon-Nykodym reste vrai pour de telles mesures.

Corollaire 5.1.5 *Soit $(X,\mathscr{M})$ un espace mesurable et μ une mesure positive σ-finie et ν une mesure réelle ou complexe sur $\mathscr{M}$. Si pour tout $E \in \mathscr{M}$ tel que $\mu(E)=0$ on a $|\nu|(E)=0$, alors il existe une unique fonction $h \in L^1(X,d\mu)$ réelle ou complexe telle que*

$$\forall A \in \mathscr{M}, \quad \nu(A) = \int_A h\,d\mu.$$

Pour clore cette section, nous allons définir quelques espaces fonctionnels, les espaces L^p locaux. Ici, X sera un espace métrique σ-compact muni de la tribu borélienne et μ une mesure de Borel positive finie sur tous les compacts de X (une telle mesure est appelée *mesure de Radon*), donc σ-finie.

Définition 5.1.11 *L'espace noté $L^p_{\mathrm{loc}}(X,d\mu)$ des* fonctions localement L^p *sur X est l'ensemble des fonctions mesurables f sur X telles que, pour tout compact K de X, on ait*

$$f \in L^p(K,d\mu).$$

Remarque. Attention, $L^p_{\mathrm{loc}}(X,d\mu)$ n'est pas un espace normé de façon naturelle, au sens où sa topologie naturelle, que nous ne précisons pas ici, n'est pas en général engendrée par une norme, sauf si X est lui-même compact, auquel cas $L^p_{\mathrm{loc}}(X,d\mu) = L^p(X,d\mu)$.

Proposition 5.1.5 *Si $p \le q$, alors $L^q_{\mathrm{loc}}(X,d\mu)$ est inclus dans $L^p_{\mathrm{loc}}(X,d\mu)$.*

Démonstration. Pour démontrer cela, il suffit d'utiliser l'inégalité de Hölder qui implique en particulier que $q \ge p \Rightarrow L^q(K,d\mu) \subset L^p(K,d\mu)$. □

Définition 5.1.12 *Soit f une fonction de $L^1_{\mathrm{loc}}(X,d\mu)$, le* support (par rapport à la mesure μ) *de f noté* $\mathrm{Supp}\, f$ *est le complémentaire du plus grand ouvert U tel que $f_{|U}$ soit nulle μ-presque partout.*

Remarque. Il est bien clair que cette notion ne dépend que de la classe d'équivalence de f modulo l'égalité μ-presque partout, et non pas du représentant f choisi pour tester l'annulation sur un ouvert. Le support d'une fonction localement intégrable est donc toujours un fermé, le plus grand ouvert en question existant bien, c'est la réunion de tous les ouverts où f s'annule μ-presque partout.

On peut caractériser le support d'une fonction f à l'aide d'intégrales. Pour cela, on suppose que X est *localement compact*, c'est-à-dire que chaque point est le centre d'une boule fermée qui est compacte. C'est naturellement le cas dans $\mathbb{R}^d$.

Proposition 5.1.6 *Soit f une fonction de $L^1_{\text{loc}}(X,d\mu)$. Un point x de X n'appartient pas au support de f si et seulement il existe un réel $r>0$ tel que, pour toute fonction mesurable bornée φ nulle en dehors de la boule $\bar{B}(x,r)$, on ait*

$$\int_X \varphi f \, d\mu = 0.$$

Démonstration. Si $x \notin \operatorname{Supp} f$, il existe $r>0$ tel que $f_{|B(x,r)} = 0$ μ-presque partout.

Réciproquement, soit x ayant la propriété de l'énoncé. On peut choisir r assez petit pour que $K = \bar{B}(x,r)$ soit compact. Pour toute fonction φ bornée et nulle en dehors de K, φf est donc intégrable sur X et

$$\int_X \varphi f \, d\mu = \int_K \varphi f \, d\mu.$$

On applique alors le lemme 5.1.4 sur K. Il vient

$$\|f\|_{L^1(K,d\mu)} = \sup_{\|\varphi\|_{L^\infty} \leq 1} \int_K \varphi f \, d\mu = 0,$$

donc f est nulle μ-presque partout sur $B(x,r)$. □

Remarque. La notion de support pour une fonction $L^1_{\text{loc}}(X,d\mu)$ dépend fondamentalement de la mesure μ. C'est pourquoi elle peut jouer des tours par rapport à l'intuition. Par exemple, si $X = \mathbb{R}$ et $\mu = \delta_0$ la masse de Dirac en 0, alors le support de la fonction $f(x) = 1$ pour tout $x \in \mathbb{R}$ est le fermé $\{0\}$! En effet, elle est bien nulle μ-presque partout sur $\mathbb{R}^*$, vu que $\mu(\mathbb{R}^*) = 0$, et elle n'est pas nulle μ-presque partout sur $\mathbb{R}$, vu que $\mu(\mathbb{R}) = 1$. Quand μ est la mesure de Lebesgue sur un ouvert Ω de $\mathbb{R}^d$, la notion de support d'une fonction $L^1_{\text{loc}}(\Omega)$ est plus conforme à l'intuition, puisqu'elle coïncide avec la notion classique (complémentaire du plus grand ouvert où la fonction s'annule) pour les fonctions continues sur Ω.

5.2 Densité des fonctions continues dans L^p

Dans ce chapitre, Ω désignera un ouvert de $\mathbb{R}^d$ muni de la distance euclidienne usuelle, μ une mesure de Borel positive finie sur tous les compacts de Ω (donc σ-finie) et nous ne considérerons que des fonctions à valeurs réelles, le cas complexe étant tout-à-fait analogue. Ces hypothèses ne sont pas essentielles, et l'on peut aisément généraliser ce qui suit avec quelques aménagements, mais c'est le cadre dans lequel on appliquera ces résultats ultérieurement. On admettra tout d'abord un résultat de théorie de la mesure.

Théorème 5.2.1 *Pour tout ensemble A appartenant à la tribu borélienne complétée et de mesure finie et pour tout ε strictement positif, il existe un ouvert U et un compact K tel que*

$$K \subset A \subset U \quad et \quad \mu(U \setminus K) < \varepsilon.$$

On dit que μ est une mesure régulière. Rappelons que la tribu borélienne est la plus petite σ-algèbre qui contient tous les ouverts. On la complète en lui ajoutant tous les ensembles qui sont compris au sens de l'inclusion entre deux boréliens dont la mesure de leur différence ensembliste est nulle. Du théorème 5.2.1, on déduit l'énoncé fondamental suivant.

Théorème 5.2.2 *Soit $p \in [1,+\infty[$, l'espace $C_c(\Omega)$ des fonctions continues à support compact dans Ω est dense dans $L^p(\Omega, d\mu)$.*

Démonstration. La démonstration de ce théorème se fait en plusieurs étapes. Soit d'abord A un borélien borné. On va approcher sa fonction caractéristique par une suite de fonctions continues à support compact, c'est-à-dire que

$$\forall \varepsilon > 0,\ \exists\, f \in C_c(\Omega) \text{ telle que } \|\mathbf{1}_A - f\|_{L^p} < \varepsilon.$$

D'après le théorème 5.2.1, il existe un ouvert U et un compact K tels que

$$K \subset A \subset U \quad \text{et} \quad \mu(U \setminus K) < \varepsilon^p.$$

Comme on a supposé que A est borné, et on peut supposer que l'adhérence de U est compacte. Il suffit pour cela de prendre l'intersection de U avec une boule contenant A. D'après la proposition 1.2.7, il existe une fonction continue f de Ω à valeurs dans $[0,1]$ telle que

$$f_{|\Omega \setminus U} = 0 \quad \text{et} \quad f_{|K} = 1.$$

Comme $\bar{U}$ est compact, f est à support compact dans Ω. De plus, on voit facilement que

$$|\mathbf{1}_A - f| \leq \mathbf{1}_{U\setminus K} \quad \text{et donc que} \quad |\mathbf{1}_A - f|^p \leq \mathbf{1}_{U\setminus K}.$$

D'où il vient

$$\|\mathbf{1}_A - f\|_{L^p} < \varepsilon.$$

Soit maintenant A un borélien de mesure finie et approchons sa fonction caractéristique. Soit $(K_n)_{n\in\mathbb{N}}$ une suite exhaustive de compacts. D'après le théorème de convergence dominée, il est clair que

$$\lim_{n\to\infty} \mathbf{1}_{K_n \cap A} = \mathbf{1}_A \quad \text{dans} \quad L^p.$$

Comme $K_n \cap A$ est borné, on sait l'approcher par une suite de fonctions continues à support compact par l'étape précédente. Un argument de double limite permet donc d'approcher $\mathbf{1}_A$ dans L^p.

Une fois que l'on sait approcher une fonction caractéristique d'un borélien de mesure finie, on sait approcher une fonction étagée non nulle sur un ensemble de mesure finie. En effet, une telle fonction est une combinaison linéaire de fonctions caractéristiques de boréliens de mesure finie, et l'inégalité triangulaire et la positivité homogène de la norme permettent de conclure dans ce cas.

Soit alors $f \in L^p(\Omega, d\mu)$ positive. On considère les fonctions étagées f_n définies par la formule du lemme 5.1.1 pour $n \geq 1$. Celles-ci sont telles que $\lim_{n\to\infty} f_n(x) = f(x)$ et $f_n(x) \leq f(x)$ pour tout x dans Ω. Le théorème de convergence dominée assure que $f_n \to f$ dans L^p quand $n \to +\infty$. Les fonctions f_n étant dans L^p, les ensembles $A_{n,j} = f^{-1}\left(\left[\frac{j}{2^n}, \frac{j+1}{2^n}\right[\right)$ et $A_n = f^{-1}([2^n, +\infty])$ correspondant à des valeurs strictement positives sont de mesure finie (attention, $A_{n,0}$ peut ne pas être de mesure finie, mais $0 \times \mathbf{1}_{A_{n,0}}$ est excellemment approché par la fonction continue à support compact 0). On est donc dans la situation précédente et l'on sait, pour chaque n fixé approcher f_n dans $L^p(\Omega, d\mu)$ par une suite de fonctions continues à support compact. Un argument de double limite permet donc d'approcher f dans L^p.

Pour conclure, il suffit de remarquer que $f = f_+ - f_-$. □

Remarque. Les fonctions continues à support compact ne sont clairement pas denses dans $L^\infty(X, d\mu)$. Quelle est leur adhérence ?

Corollaire 5.2.1 *Pour tout réel $p \geq 1$, l'espace $L^p(\Omega, d\mu)$ est séparable.*

Démonstration. Soit $(K_n)_{n\in\mathbb{N}}$ une suite exhaustive de compacts dans Ω. D'après le théorème précédent, l'ensemble $\cup_{n\in\mathbb{N}} C(K_n)$ est dense dans $L^p(\Omega, d\mu)$. Or pour chaque n, $C(K_n)$ est séparable pour la topologie de la norme uniforme, donc a fortiori pour la topologie L^p, en raison du théorème de Weierstrass et en approchant chaque polynôme à coefficients réels par une suite de polynômes à coefficients rationnels. Un argument de suite diagonale permet donc de conclure. □

Remarque. Rappelons que $L^\infty(\Omega, d\mathscr{L}_d)$ est un exemple d'espace de Banach concret non séparable.

Exercice 5.2.1 *Soient $x \in \mathbb{R}^d$ et τ_x l'application définie par*

$$\begin{aligned} \tau_x \colon L^p(\mathbb{R}^d, d\mathscr{L}_d) &\to L^p(\mathbb{R}^d, d\mathscr{L}_d) \\ f &\mapsto \tau_x(f) : y \mapsto f(x-y). \end{aligned}$$

L'application τ_x est une isométrie de $L^p(\mathbb{R}^d, d\mathscr{L}_d)$ dans $L^p(\mathbb{R}^d, d\mathscr{L}_d)$ et, si p est réel, alors, on a

$$\forall f \in L^p,\ \lim_{x\to 0} \|\tau_x f - \check{f}\|_{L^p} = 0 \quad \textit{avec} \quad \check{f}(y) \stackrel{\textit{déf}}{=} f(-y).$$

5.3 Dualité entre L^p et $L^{p'}$

Le théorème principal et d'ailleurs unique de cette section est le théorème suivant, qui dit comment l'on peut représenter une forme linéaire continue sur les espaces L^p lorsque p est réel. Nous traiterons le cas des fonctions à valeurs réelles, le cas complexe étant tout-à-fait analogue. Pour simplifier, nous nous plaçons toujours dans le cas où Ω est un ouvert de $\mathbb{R}^d$.

Théorème 5.3.1 *Soient $p \in [1, +\infty[$ et B la forme bilinéaire définie par*

$$\begin{array}{rcl} B\colon L^{p'}(\Omega, d\mu) \times L^p(\Omega, d\mu) & \to & \mathbb{R} \\ (g, f) & \mapsto & \displaystyle\int_\Omega f(x)g(x)\,d\mu. \end{array}$$

Alors la forme bilinéaire B identifie le dual de $L^p(\Omega, d\mu)$ à $L^{p'}(\Omega, d\mu)$.

Démonstration. On va donner une preuve de ce théorème fondamental qui repose sur le théorème de Radon-Nykodym. On a déjà noté que, comme μ est finie sur les compacts, elle est σ-finie. Par l'inégalité de Hölder, l'application δ_B (voir section 3.2) est isométrique, donc injective. On va montrer qu'elle est surjective.

Commençons par le cas où $\mu(\Omega) < +\infty$. Soit Φ une forme linéaire continue sur $L^p(\Omega, d\mu)$. Pour tout $A \subset \Omega$ mesurable, comme $\mu(A) < +\infty$, on a $\mathbf{1}_A \in L^p(\Omega, d\mu)$ et l'on est en droit de poser

$$\nu(A) = \Phi(\mathbf{1}_A).$$

Montrons que ν est une mesure réelle. Si $A \cap B = \emptyset$, alors $\mathbf{1}_{A\cup B} = \mathbf{1}_A + \mathbf{1}_B$, d'où par linéarité de Φ,

$$\nu(A \cup B) = \nu(A) + \nu(B).$$

On en déduit que ν est additive. Montrons qu'elle est aussi dénombrablement additive. Soit $(A_n)_{n\in\mathbb{N}}$ une famille dénombrable disjointe d'ensembles mesurables, et notons $B_k = \bigcup_{i=0}^k A_i$ et $B = \bigcup_{i=0}^{+\infty} A_i$. Il vient

$$\|\mathbf{1}_B - \mathbf{1}_{B_k}\|_{L^p} = \big(\mu(B \setminus B_k)\big)^{1/p} \to 0 \quad \text{quand} \quad k \to +\infty,$$

puisque μ est une mesure finie (c'est ici que l'on a utilisé ici le fait que $p < +\infty$). La continuité de Φ implique donc que

$$\sum_{i=0}^k \nu(A_i) = \nu(B_k) \to \nu(B) \quad \text{quand} \quad k \to +\infty,$$

c'est-à-dire que ν est bien dénombrablement additive.

Soit E tel que $\mu(E)=0$ et soit $F\subset E$. La fonction caractéristique de F est donc nulle μ-presque partout, sa classe d'équivalence dans L^p est la classe nulle, donc $\nu(F)=0$. Ceci implique que $|\nu|(E)=0$ par la définition de la variation totale. On peut donc appliquer le théorème de Radon-Nykodym au couple (ν,μ) pour en déduire qu'il existe une fonction $g\in L^1(\Omega,d\mu)$ telle que

$$\forall A\in\mathscr{M},\quad \Phi(\mathbf{1}_A)=\nu(A)=\int_A g(x)\,d\mu=\int_\Omega g(x)\mathbf{1}_A\,d\mu.$$

Par linéarité, pour toute fonction f étagée, de la forme $f(x)=\sum_{i=1}^k a_i\mathbf{1}_{A_i}(x)$, on obtient

$$\Phi(f)=\int_\Omega g(x)f(x)\,d\mu. \tag{5.3}$$

Par une approximation supplémentaire, cette égalité reste vraie pout tout $f\in L^\infty(\Omega,d\mu)$, *cf.* la démonstration du lemme 5.1.1. La suite f_n de fonctions étagées qui y est construite tend vers f dans $L^p(\Omega,d\mu)$, d'où la convergence du membre de gauche, et uniformément, d'où la convergence de l'intégrale du membre de droite car $g\in L^1(\Omega,d\mu)$.

Il faut maintenant montrer que $g\in L^{p'}(\Omega,d\mu)$ et que l'égalité (5.3) reste vraie pour tout $f\in L^p(\Omega,d\mu)$. On pose $\operatorname{signe} t=-1$ si $t<0$ et $\operatorname{signe} t=1$ si $t\geq 0$. On distingue deux cas suivant les valeurs de p.

Dans le cas où $p=1$, si l'on prend $f=\operatorname{signe} g\mathbf{1}_A$, on déduit de (5.3) que

$$\int_A |g(x)|\,d\mu\leq \|\Phi\|_{(L^1)'}\|\mathbf{1}_A\|_{L^1}=\|\Phi\|_{(L^1)'}\mu(A),$$

d'où $|g(x)|\leq\|\Phi\|_{(L^1)'}$ μ-presque partout. En effet, si l'on pose $B_n=\{x\in\Omega;|g(x)|\geq \|\Phi\|_{(L^1)'}+1/n\}$ et $B=\{x\in\Omega;|g(x)|>\|\Phi\|_{(L^1)'}\}$, l'inégalité précédente implique que $\mu(B_n)=0$, d'où $\mu(B)=0$ puisque $B_n\subset B_{n+1}$ et $B=\cup B_n$. En résumé, on a $g\in L^\infty(\Omega,d\mu)$ et $\|g\|_{L^\infty}\leq\|\Phi\|_{(L^1)'}$.

Soit maintenant $1<p<+\infty$. Introduisons les fonctions

$$g_n(x)=\mathbf{1}_{\{|g|\leq n\}}(x)g(x) \text{ et } f_n(x)=\operatorname{signe}(g(x))\mathbf{1}_{\{|g|\leq n\}}(x)|g(x)|^{\frac{1}{p-1}}.$$

On a $f_n\in L^\infty(\Omega,d\mu)$, $gf_n=|g_n|^{p'}$ et $|f_n|^p=|g_n|^{p'}$. Appliquant (5.3), il vient

$$\int_\Omega |g_n(x)|^{p'}\,d\mu=\Phi(f_n)\leq\|\Phi\|_{(L^p)'}\|f_n\|_{L^p}=\|\Phi\|_{(L^p)'}\|g_n\|_{L^{p'}}^{\frac{p'}{p}},$$

d'où immédiatement

$$\|g_n\|_{L^{p'}}\leq\|\Phi\|_{(L^p)'}.$$

La suite $|g_n|^{p'}$ converge en croissant vers $|g|^{p'}$, donc par le théorème de convergence monotone, on obtient

$$g\in L^{p'}(\Omega,d\mu)\quad\text{ et }\quad\|g\|_{L^{p'}}\leq\|\Phi\|_{(L^p)'}.$$

Il est maintenant trivial de vérifier que (5.3) reste vraie pour tout $f \in L^p(\Omega, d\mu)$ puisque $L^\infty(\Omega, d\mu)$ est dense dans $L^p(\Omega, d\mu)$ et que les deux membres sont de formes linéaires continues sur $L^p(\Omega, d\mu)$. L'inégalité de Hölder implique alors $\|g\|_{L^{p'}} = \|\Phi\|_{(L^p)'}$, ce qui termine la démonstration dans le cas $\mu(\Omega) < +\infty$.

On passe au cas général comme dans la démonstration du théorème de Radon-Nykodym. □

Remarques. On n'a pas vraiment utilisé le fait que Ω est un ouvert de $\mathbb{R}^d$. Le théorème reste vrai plus généralement au sens où le dual de $L^p(X, d\mu)$ est bien $L^{p'}(X, d\mu)$ pour $1 \leq p < +\infty$ si μ est σ-finie. Notons aussi que $L^1(\Omega, d\mu)$ n'est pas en général (sauf pour certaines mesures μ très particulières) le dual de $L^\infty(\Omega, d\mu)$: il existe des formes linéaires continues sur $L^\infty(\Omega, d\mu)$ qui ne sont pas représentées par une fonction de $L^1(\Omega, d\mu)$. Donnons en un exemple. Supposons que $\Omega =]-1, 1[$ et définissons $\Phi(f) = f(0)$ pour tout $f \in \mathscr{C}([-1, 1])$. C'est une forme linéaire continue sur le sous-espace vectoriel $\mathscr{C}([-1, 1])$, donc par le théorème de Hahn-Banach, elle admet un prolongement continu à $L^\infty(\Omega, dx)$ tout entier. Il est facile de voir que cette forme linéaire n'est pas représentée par une fonction g de $L^1(\Omega, dx)$.

Développons en quelques mots une description possible du dual de $L^\infty(\Omega, dx)$. Nous avons vu que tout élément Φ de $L^\infty(\Omega, dx)'$ donne naissance à une mesure finiment additive ν et que cette mesure est bornée puisque pour tout A mesurable, $|\nu(A)| \leq \|\Phi\|_{L^\infty(\Omega, dx)'}$. Par ailleurs, si $\mu(A) = 0$, alors $\nu(A) = 0$ également. On peut montrer que de telles mesures ont une variation totale finie et définir à partir de là une notion d'intégrale de fonctions mesurables bornées en prolongeant par continuité la formule évidente pour les fonctions étagées. Le dual de L^∞ est identifié de la sorte à l'espace des mesures finiment additives, bornées, normé par la variation totale de l'espace entier, à travers cette notion d'intégrale. Le crochet de dualité est donc $\langle \nu, f \rangle = \int_\Omega f \, d\nu$.

Corollaire 5.3.1 *Sous les hypothèses du théorème précédent, $L^p(X, d\mu)$ est réflexif pour tout $1 < p < +\infty$.*

Démonstration. On omet la démonstration, qui ne se réduit pas, on le rappelle, à juste constater que $(p')' = p$! □

5.4 Convolution et régularisation

Dans toute cette section, nous travaillerons dans l'espace $\mathbb{R}^d$ muni de la mesure de Lebesgue. La convolution est un outil crucial pour l'étude des fonctions sur $\mathbb{R}^d$. L'idée de la convolution repose sur le théorème suivant.

Théorème 5.4.1 *Soient f et g deux fonctions de L^1. Alors, pour presque tout x de $\mathbb{R}^d$, la fonction*

$$y \mapsto f(x-y)g(y)$$

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 de* convolution $\star$ *ainsi définie est une application bilinéaire continue de $L^1\times L^1$ dans L^1. De plus, on a $f\star g = g\star f$.*

Démonstration. Comme les fonctions f et g sont supposées être dans L^1, on a

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} |f(x-y)|\,|g(y)|\,dxdy = \int_{\mathbb{R}^d} |g(y)|\Big(\int_{\mathbb{R}^d} |f(x-y)|\,dx\Big)dy,$$

par le théorème de Fubini positif 5.1.8. Or le changement de variable $z = x-y$ montre que $\int_{\mathbb{R}^d} |f(x-y)|\,dx = \int_{\mathbb{R}^d} |f(z)|\,dz$. Par conséquent,

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} |f(x-y)|\,|g(y)|\,dxdy = \|f\|_{L^1}\|g\|_{L^1} < +\infty$$

d'où

$$|f(x-y)|\,|g(y)| \in L^1(\mathbb{R}^d\times\mathbb{R}^d).$$

Les conclusions du théorème 5.1.8 entraînent immédiatement la première partie du théorème. □

Le même changement de variable montre que

$$f\star g(x) = \int_{\mathbb{R}^d} f(x-y)g(y)\,dy = \int_{\mathbb{R}^d} f(z)g(x-z)\,dz = g\star f(x).$$

Remarque. La convolution est commutative. Elle est aussi associative. En effet, on vérifie très facilement que $(f\star g)\star h = f\star(g\star h)$. Ces deux propriétés ne sont en fait que le reflet des propriétés analogues de l'addition dans $\mathbb{R}^d$.

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. Choisissons un représentant quelconque de la classe d'équivalence modulo l'égalité presque partout f, que nous notons également f. D'après l'inégalité de Hölder, pour tout x dans $\mathbb{R}^d$, l'application

$$y \mapsto f(x-y)g(y)$$

est intégrable et pour tout $x \in \mathbb{R}^d$,

$$|(f \star g)(x)| \leq \|f\|_{L^p} \|g\|_{L^{p'}},$$

d'où le théorème. □

Remarque. Le choix d'un représentant de f fournit un représentant de $f \star g$. En effet, si $f = f'$ presque partout, alors clairement, $f \star g = f' \star g$ presque partout. Par contre, les représentants de la convoluée ainsi obtenus sont tous bornés partout, et pas seulement presque partout. En fait, il est facile de voir que $f \star g$ est continue bornée par densité de $C_c(\mathbb{R}^d)$ dans L^p et $L^{p'}$.

Le théorème ci-dessus se généralise encore. On peut définir la convolution de deux fonctions si elles appartiennent à des espaces L^p convenables. Plus précisément :

Théorème 5.4.3 *Soit (p,q,r) un triplet de réels tel que*

$$1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q} \cdot \tag{5.4}$$

Alors, pour presque tout x de $\mathbb{R}^d$, la fonction

$$y \mapsto f(x-y)g(y)$$

est intégrable et la fonction $f \star g$ définie sur $\mathbb{R}^d$ en l'intégrant par rapport à y appartient à L^r. L'application $(f,g) \mapsto f \star g$ est bilinéaire continue de $L^p \times L^q$ dans L^r avec

$$\|f \star g\|_{L^r} \leq \|f\|_{L^p} \|g\|_{L^q}.$$

Démonstration. Dans le cas général, cette preuve est un exemple d'application du lemme 5.1.4. Nous avons déjà vu les cas $p = q = r = 1$ et $q = p'$, $r = +\infty$. On peut donc supposer que $1 < r, r' < +\infty$ et $p, q < +\infty$.

Soit φ une fonction de $L^{r'}$, nous allons majorer

$$I_\varphi(f,g) \stackrel{\text{déf}}{=} \int_{\mathbb{R}^d \times \mathbb{R}^d} |f(x-y)g(y)\varphi(x)| \, dxdy.$$

Soient α et β deux réels de l'intervalle $]0,1[$ à choisir intelligemment plus tard. On écrit

$$I_\varphi(f,g) = \int_{\mathbb{R}^d \times \mathbb{R}^d} |f(x-y)|^{1-\alpha} |g(y)|^{1-\beta} |\varphi(x)| \, |f(x-y)|^{\alpha} |g(y)|^{\beta} \, dxdy.$$

Appliquons l'inégalité de Hölder avec la mesure

$$d\mu = |f(x-y)|^{\alpha}|g(y)|^{\beta}\,dxdy$$

et les exposants r et r'. Il en résulte que

$$\begin{aligned}
I_\varphi(f,g) &\le I_{\varphi,1}(f,g)^{\frac{1}{r'}} I_{\varphi,2}(f,g)^{\frac{1}{r}} \quad \text{avec}\\
I_{\varphi,1}(f,g) &= \int_{\mathbb{R}^d\times\mathbb{R}^d} |\varphi(x)|^{r'}|f(x-y)|^{\alpha}|g(y)|^{\beta}\,dxdy \quad \text{et}\\
I_{\varphi,2}(f,g) &= \int_{\mathbb{R}^d\times\mathbb{R}^d} |f(x-y)|^{(1-\alpha)r+\alpha}|g(y)|^{(1-\beta)r+\beta}\,dxdy.
\end{aligned}$$

Majorons $I_{\varphi,1}(f,g)$. Pour ce faire, on choisit les deux paramètres α et β pour que les divers exposants qui apparaissent soient agréables. Plus précisément, on prend

$$\alpha = \frac{r-p}{r-1} \quad \text{et} \quad \beta = \frac{r-q}{r-1},$$

qui sont bien compris entre 0 et 1 car $r > p$ et $r > q$. Il résulte de ce choix et de la relation (5.4) que

$$\frac{\alpha}{p}+\frac{\beta}{q} = 1.$$

L'inégalité de Hölder entraîne alors que

$$\begin{aligned}
\int_{\mathbb{R}^d} |f(x-y)|^{\alpha}|g(y)|^{\beta}\,dy &\le \Big(\int_{\mathbb{R}^d} |f(x-y)|^{\alpha\times\frac{p}{\alpha}}\,dy\Big)^{\frac{\alpha}{p}} \Big(\int_{\mathbb{R}^d} |g(y)|^{\beta\times\frac{q}{\beta}}\,dy\Big)^{\frac{\beta}{q}}\\
&= \|f\|_{L^p}^{\alpha}\|g\|_{L^q}^{\beta}.
\end{aligned}$$

D'où il résulte par Fubini (en intégrant d'abord en y, puis en x) que

$$I_{\varphi,1}(f,g)^{\frac{1}{r'}} \le \|\varphi\|_{L^{r'}}\|f\|_{L^p}^{\frac{\alpha}{r'}}\|g\|_{L^q}^{\frac{\beta}{r'}}.$$

Pour majorer $I_{\varphi,2}(f,g)$, remarquons que

$$(1-\alpha)r+\alpha = p \quad \text{et que} \quad (1-\beta)r+\beta = q.$$

Ainsi on a, encore par Hölder et Fubini,

$$\begin{aligned}
I_{\varphi,2}(f,g)^{\frac{1}{r}} &\le \Big(\int_{\mathbb{R}^d} |f(x)|^p\,dx\Big)^{\frac{1}{r}} \Big(\int_{\mathbb{R}^d} |g(x)|^q\,dx\Big)^{\frac{1}{r}}\\
&\le \|f\|_{L^p}^{\frac{p}{r}}\|g\|_{L^q}^{\frac{q}{r}}.
\end{aligned}$$

On trouve donc finalement que

$$I_\varphi(f,g) \leq \|\varphi\|_{L^{r'}} \|f\|_{L^p}^{\frac{\alpha}{r'}+\frac{p}{r}} \|g\|_{L^q}^{\frac{\beta}{r'}+\frac{q}{r}}.$$

Mais, par définition de α et β, on a

$$\frac{\alpha}{r'} + \frac{p}{r} = \frac{\beta}{r'} + \frac{q}{r} = 1,$$

d'où

$$I_\varphi(f,g) \leq \|\varphi\|_{L^{r'}} \|f\|_{L^p} \|g\|_{L^q}.$$

On déduit en premier lieu de l'inégalité qui précède que la fonction $(x,y) \mapsto |f(x-y)g(y)\varphi(x)|$ est intégrable sur $\mathbb{R}^d \times \mathbb{R}^d$. Par le théorème de Fubini, on en déduit que pour presque tout x, la fonction $y \mapsto |\varphi(x)|\,|f(x-y)g(y)|$ est intégrable, d'où en choisissant une famille dénombrable appropriée de fonctions φ_n, que pour presque tout x, la fonction $y \mapsto |f(x-y)g(y)|$ est intégrable. On peut donc définir $f \star g$ par la formule intégrale usuelle et l'on a par Fubini positif

$$\begin{aligned}\int_{\mathbb{R}^d} |f\star g(x)\varphi(x)|\,dx &= \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} |f(x-y)g(y)|\,dy\Big)|\varphi(x)|\,dx \\ &= I_\varphi(f,g) \leq \|\varphi\|_{L^{r'}} \|f\|_{L^p} \|g\|_{L^q}.\end{aligned}$$

On en déduit que $f\star g \in L^r$ avec $\|f\star g\|_{L^r} \leq \|f\|_{L^p}\|g\|_{L^q}$ par le lemme 5.1.4. □

Remarque. Les deux cas extrêmes $p=q=1$ et $q=p'$ faisant intervenir uniquement l'un le théorème de Fubini, l'autre l'inégalité de Hölder, il est normal que les cas intermédiaires fassent intervenir simultanément ces deux résultats.

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

$$\mathrm{Supp}(f\star g) \subset \mathrm{Adh}\,(\mathrm{Supp}\, f + \mathrm{Supp}\, g).$$

Démonstration. L'hypothèse faite sur les exposants nous permet de définir la convolution d'après le théorème précédent. Soit x un point de $\mathbb{R}^d$ et ρ un réel strictement positif tels que

$$B(x,\rho) \cap (\mathrm{Supp}\, f + \mathrm{Supp}\, g) = \emptyset.$$

Pour toute fonction φ bornée et nulle en dehors de $B(x,\rho)$, on a

$$\begin{aligned}\int_{\mathbb{R}^d\times\mathbb{R}^d} \varphi(x)f(x-y)g(y)\,dxdy &= \int_{\mathbb{R}^d\times\mathbb{R}^d} \varphi(z+y)f(z)g(y)\,dzdy \\ &= 0,\end{aligned}$$

car l'intégrande est identiquement nulle. On en déduit le théorème en appliquant le lemme 5.1.4. □

La convolution est une opération cruciale. Elle permet, parmi bien d'autres choses, 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. On se ramène d'abord à une fonction à support compact en posant

$$\varphi_R(x) = \left(\int_{B(0,R)} \varphi(y)\,dy\right)^{-1}\mathbf{1}_{B(0,R)}(x)\varphi(x).$$

Cette fonction est bien définie pour R assez grand, elle est d'intégrale 1, à support dans la boule de centre 0 et de rayon R. De plus,

$$\begin{aligned}\|\varphi_\varepsilon - \varphi_{R,\varepsilon}\|_{L^1} &= \varepsilon^{-d}\int_{|z|\leq\varepsilon R}\left|1-\left(\int_{B(0,R)}\varphi(y)\,dy\right)^{-1}\right||\varphi(\varepsilon^{-1}z)|\,dz\\ &\qquad\qquad +\varepsilon^{-d}\int_{|z|>\varepsilon R}|\varphi(\varepsilon^{-1}z)|\,dz\\ &= \left|1-\left(\int_{B(0,R)}\varphi(y)\,dy\right)^{-1}\right|\int_{|x|\leq R}|\varphi(x)|\,dx\\ &\qquad\qquad +\int_{|x|>R}|\varphi(x)|\,dx \to 0 \text{ quand } R\to+\infty\end{aligned}$$

uniformément par rapport à ε. Pour toute fonction g continue à support compact, on peut donc écrire

$$\varphi_\varepsilon\star f - f = (\varphi_\varepsilon - \varphi_{R,\varepsilon})\star f + \varphi_{R,\varepsilon}\star(f-g) + (\varphi_{R,\varepsilon}\star g - g) + (g-f).$$

Majorons chaque terme en norme L^p. On a

$$\|(\varphi_\varepsilon - \varphi_{R,\varepsilon})\star f\|_{L^p} \leq \|\varphi_\varepsilon - \varphi_{R,\varepsilon}\|_{L^1}\|f\|_{L^p},$$

puis

$$\|\varphi_{R,\varepsilon}\star(f-g)\|_{L^p} \leq \|\varphi_{R,\varepsilon}\|_{L^1}\|f-g\|_{L^p} = \|\varphi_R\|_{L^1}\|f-g\|_{L^p} \leq 2\|\varphi\|_{L^1}\|f-g\|_{L^p}$$

si R est assez grand, car $\int_{B(0,R)}\varphi(y)\,dy \to 1$. En résumé, il vient

$$\|\varphi_\varepsilon\star f - f\|_{L^p} \leq \|\varphi_\varepsilon - \varphi_{R,\varepsilon}\|_{L^1}\|f\|_{L^p} + (2\|\varphi\|_{L^1}+1)\|f-g\|_{L^p} + \|\varphi_{R,\varepsilon}\star g - g\|_{L^p}.$$

Par densité des fonctions continues à support compact dans L^p, il suffit de donc considérer le dernier terme à R fixé. Comme φ_R est d'intégrale 1, on obtient par un calcul simple que pour tout $x \in \mathbb{R}^d$,

$$\varphi_{R,\varepsilon} \star g(x) - g(x) = \int_{B(0,R)} \varphi_R(z)\big(g(x-\varepsilon z) - g(x)\big)\,dz.$$

La fonction g est continue à support compact, donc uniformément continue. Le support de $\varphi_{R,\varepsilon}$ est inclus dans $\bar{B}(0,R\varepsilon)$, ce qui implique que pour $\varepsilon \leq 1$, le support de $\varphi_{R,\varepsilon} \star g$ est inclus dans $K = \mathrm{Adh}\,(\bar{B}(0,R) + \mathrm{Supp}\,g)$ qui est un compact donc de mesure finie. Par conséquent, pour tout $\eta > 0$, il existe $\alpha > 0$ tel que $|y-y'| \leq \alpha R$ implique que $|g(y) - g(y')| \leq \eta \|\varphi_R\|_{L^1}^{-1}(\mathscr{L}_d(K))^{-1/p}$. On voit donc que dès que $\varepsilon \leq \alpha$,

$$|\varphi_{R,\varepsilon} \star g(x) - g(x)| \leq \eta(\mathscr{L}_d(K))^{-1/p},$$

d'où

$$\|\varphi_{R,\varepsilon} \star g - g\|_{L^p} \leq \eta.$$

Le théorème est ainsi démontré. □

Remarque. Si l'on examine la dernière étape de la démonstration, on s'aperçoit que si f est uniformément continue sur $\mathbb{R}^d$ et φ est à support compact, alors $\varphi_\varepsilon \star f$ est bien définie et converge uniformément vers f sur $\mathbb{R}^d$.

Remarque importante. Ce théorème ne s'applique pas pour $p = +\infty$.

Exercice 5.4.1 *Considérer pour f la fonction de Heaviside H (la fonction caractéristique des réels positifs) et pour φ une fonction continue à support compact, paire et d'intégrale* 1. *Démontrer alors que*

$$\lim_{\varepsilon \to 0} \|\varphi_\varepsilon \star f - f\|_{L^\infty} \neq 0.$$

Comme nous allons le voir, le théorème 5.4.5 est extrêmement important pour la régularisation des fonctions localement intégrables. En effet, le théorème de dérivation sous l'intégrale assure que si φ est une fonction indéfiniment différentiable à support compact, alors la fonction $\varphi_\varepsilon \star f$ est aussi indéfiniment différentiable. Ainsi l'on aura approché au sens de L^p toute fonction de L^p, pour p réel, par une suite de fonctions C^∞.

Théorème 5.4.6 *Soit $k \in \mathbb{N}$. Si φ appartient à $C^k(\mathbb{R}^d)$ et est à support compact, alors pour tout $f \in L^p(\mathbb{R}^d)$, $\varphi \star f \in C^k(\mathbb{R}^d)$. Si φ appartient à $C^\infty(\mathbb{R}^d)$ et est à support compact, alors $\varphi \star f \in C^\infty(\mathbb{R}^d)$.*

Démonstration. Commençons par le cas $k = 0$. Si φ est continue à support compact, alors elle est trivialement dans $L^{p'}(\mathbb{R}^d)$ et la convolution $\varphi \star f$ est bien définie et appartient à $L^\infty(\mathbb{R}^d)$ (φ est aussi trivialement dans $L^1(\mathbb{R}^d)$ donc $\varphi \star f$ appartient aussi à $L^p(\mathbb{R}^d)$). Dans la suite, on choisit un représentant de f.

Soit $x \in \mathbb{R}^d$ et x_n une suite de points de $\mathbb{R}^d$ qui tend vers x. Il est facile de voir que l'ensemble

$$K = \overline{\bigcup_n (x_n - \operatorname{Supp} \varphi)}$$

est fermé borné, donc compact. On a

$$\begin{aligned}\varphi \star f(x_n) - \varphi \star f(x) &= \int_{\mathbb{R}^d} (\varphi(x_n - y) - \varphi(x - y)) f(y)\, dy \\ &= \int_K (\varphi(x_n - y) - \varphi(x - y)) f(y)\, dy.\end{aligned}$$

Par conséquent,

$$|\varphi \star f(x_n) - \varphi \star f(x)| \leq \int_K |\varphi(x_n - y) - \varphi(x - y)||f(y)|\, dy.$$

Soit $M = \max_{\operatorname{Supp} \varphi} |\varphi|$. Comme f appartient à $L^p(\mathbb{R}^d)$, elle est dans $L^1_{\text{loc}}(\mathbb{R}^d)$, donc intégrable sur K. Comme φ est continue, on a donc

$$\begin{aligned}|\varphi(x_n - y) - \varphi(x - y)||f(y)| &\to 0 \text{ quand } n \to +\infty \text{ p.p. en } y \\ |\varphi(x_n - y) - \varphi(x - y)||f(y)| &\leq 2M|f(y)|,\end{aligned}$$

et le théorème de convergence dominée nous donne la continuité de $\varphi \star f$ en x.

Regardons maintenant le cas $k = 1$. Soit $x_0 \in \mathbb{R}^d$ un point fixé et $\Omega = B(x_0, 1)$ une boule ouverte qui contient ce point. Comme plus haut, on note que

$$K' = \overline{\cup_{x \in \Omega} (x - \operatorname{Supp} \varphi)}$$

est un compact, et l'on a

$$\varphi \star f(x) = \int_{K'} \varphi(x - y) f(y)\, dy,$$

pour tout $x \in \Omega$. On note aussi que la fonction $h \colon \Omega \times K'$, $h(x,y) = \varphi(x-y)f(y)$ est différentiable par rapport à x, intégrable sur K' ainsi que sa différentielle par rapport à x et que

$$|D_x h(x,y)| \leq \max_{\operatorname{Supp} \varphi} |D\varphi||f(y)|,$$

avec f intégrable sur K'. Le théorème de dérivation sous le signe somme 5.1.7 assure donc que $\varphi \star f$ est de classe C^1 sur Ω avec $\frac{\partial(\varphi \star f)}{\partial x_j}(x) = \frac{\partial \varphi}{\partial x_j} \star f(x)$.

Le cas général se fait par récurrence sur k. □

Remarque. La seule propriété de f que l'on a utilisé est que f est localement intégrable. Dans ce cas, $\varphi \star f$ est également bien défini car φ est à support compact, et appartient à $C^k(\mathbb{R}^d)$. L'existence de fonctions indéfiniment différentiables à support compact doit être démontrée, car elle est loin d'aller de soi. Commençons par en fabriquer une à la main.

Lemme 5.4.1 *Soit f la fonction de $\mathbb{R}$ dans $\mathbb{R}$ définie par*

$$f(t) = e^{\frac{1}{t-1}} \text{ si } t < 1 \quad \text{et} \quad f(t) = 0 \text{ sinon.}$$

Cette fonction est indéfiniment différentiable sur $\mathbb{R}$.

Démonstration. Il suffit d'observer que dans l'intervalle $]-\infty, 1[$,

$$f^{(k)}(t) = \frac{P_k(t)}{(t-1)^{2k}} e^{\frac{1}{t-1}},$$

où P_k est un polynôme. Les dérivées à tout ordre se raccordent donc à 0 en $t = 1$. Les détails sont laissés en exercice. $\square$

Corollaire 5.4.1 *Soit Ω un ouvert de $\mathbb{R}^d$, on désigne par $\mathscr{D}(\Omega)$ l'ensemble des fonctions indéfiniment différentiables et à support compact inclus dans Ω. L'espace $\mathscr{D}(\Omega)$ est différent de $\{0\}$.*

Démonstration. On note $|x|^2$ le carré de la norme euclidienne de x. C'est une fonction de classe C^∞ sur $\mathbb{R}^d$. Soit $x_0 \in \Omega$ et $\alpha > 0$ tel que $\bar{B}(x_0, \alpha) \subset \Omega$. La fonction

$$\varphi(x) = f\Big(\frac{|x - x_0|^2}{\alpha^2}\Big)$$

où f est la fonction définie dans le lemme 5.4.1 est C^∞ à support dans $\bar{B}(x_0, \alpha)$. $\square$

Une fois que l'on a construit cette fonction particulière, on en a en fait énormément à notre disposition.

Corollaire 5.4.2 *Pour tout réel $p \geq 1$, l'espace $\mathscr{D}(\mathbb{R}^d)$ 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 (on prend la fonction précédente, elle est positive, donc d'intégrale strictement positive, et on la divise par cette intégrale). Considérons la famille (φ_e)

$$\varphi_\varepsilon(x) = \varepsilon^{-d} \varphi\left(\frac{x}{\varepsilon}\right) \cdot$$

Considérons alors une fonction f de L^p. On peut trouver une fonction g de L^p 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 est indéfiniment différentiable, il en est de même pour la fonction g_ε par le théorème 5.4.6. Comme on a vu que $g_\varepsilon \to g$ dans $L^p(\mathbb{R}^d)$ quand $\varepsilon \to 0$, on a bien le corollaire. □

Remarque. Une famille (φ_ε) qui a les propriétés ci-dessus est appelée *suite régularisante* ou bien *approximation de l'identité.*

Corollaire 5.4.3 *Pour tout réel $p \geq 1$, l'espace $\mathscr{D}(\Omega)$ est dense dans $L^p(\Omega)$.*

Démonstration. Considérons une suite de exhaustive de compacts $(K_n)_{n\in\mathbb{N}}$ de Ω (voir la définition 1.6.1 page 54). On pose alors $f_n \stackrel{\text{déf}}{=} \mathbf{1}_{K_n} f$, que l'on étend par 0 en dehors de Ω. D'après le théorème 5.1.5, 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 sur $\mathbb{R}^d$. De plus, son support est inclus dans le compact $K_n + \bar{B}(0,\varepsilon)$, c'est-à-dire l'ensemble des points à distance au plus ε de K_n. D'après la proposition 1.6.1 page 53, il existe $\varepsilon_n > 0$ tel que pour $\varepsilon \leq \varepsilon_n$, alors le compact $K_n + \bar{B}(0,\varepsilon)$ est inclus dans Ω, d'où le corollaire. □

Remarque. Une autre application de la régularisation est la construction de partitions de l'unité de classe C^∞.

Chapitre 6

Distributions sur un ouvert de $\mathbb{R}^d$

6.1 Présentation des idées

Les deux idées de base qui fondent la théorie des distributions sont la dualité et la transposition. Le but principal est de dégager une généralisation de la notion de fonction la plus vaste possible qui permette de garder un sens aux opérateurs différentiels linéaires.

Tout repose sur l'observation suivante. Soit une fonction f fixée sur un ouvert Ω de $\mathbb{R}^d$. Il est possible de lire ses propriétés à travers les quantités scalaires suivantes :

$$\langle f, \varphi \rangle = \int_\Omega f(x)\varphi(x)\,dx,$$

à condition de prendre suffisamment de fonctions φ, dites fonctions-test, par exemple les fonctions indéfiniment différentiables à support compact, pour pouvoir faire la différence entre deux fonctions distinctes. L'idée est de ne pas regarder les fonctions en tant que telles, mais de les considérer comme des formes linéaires sur certains espaces de fonctions. Ce faisant, on sera amené à généraliser considérablement le concept même de fonction.

En particulier, on regardera la convergence d'une suite de fonctions $(f_n)_{n\in\mathbb{N}}$, non plus au travers de la convergence ponctuelle, mais au travers de la convergence des quantités $\int_\Omega f_n(x)\varphi(x)\,dx$. Cette idée est étroitement liée à celles de convergence faible-$*$ des chapitres 3 et 5, mais dans un cadre encore plus général. Vérifions d'abord que les fonctions indéfiniment différentiables à support compact sont assez nombreuses pour retrouver ces notions de convergence faible-$*$.

Proposition 6.1.1 *Soient Ω un ouvert de $\mathbb{R}^d$, p un élément de $]1,+\infty[$, $(f_n)_{n\in\mathbb{N}}$ une suite bornée d'éléments de $L^p(\Omega)$ et f une fonction de $L^p(\Omega)$. On a alors*

$$f_n \overset{*}{\rightharpoonup} f \text{ quand } n\to\infty \Leftrightarrow \forall\,\varphi\in\mathscr{D}(\Omega), \int_\Omega f_n(x)\varphi(x)\,dx \to \int_\Omega f(x)\varphi(x)\,dx,$$

où $\mathscr{D}(\Omega)$ désigne l'espace des fonctions indéfiniment différentiables à support compact dans Ω.

Démonstration. Il est clair que seule l'implication de droite à gauche est à démontrer puisque $\mathscr{D}(\Omega) \subset L^{p'}(\Omega)$. Soit ε un réel strictement positif et g une fonction de $L^{p'}$. D'après le corollaire 5.4.2, il existe une fonction φ appartenant à $\mathscr{D}(\Omega)$ telle que

$$\|g-\varphi\|_{L^{p'}} \leq \frac{\varepsilon}{2(\sup\|f_n\|_{L^p}+\|f\|_{L^p})}.$$

On a alors

$$\Big|\int_\Omega f_n(x)g(x)\,dx - \int_\Omega f(x)g(x)\,dx\Big| \leq \Big|\int_\Omega f_n(x)\varphi(x)\,dx - \int_\Omega f(x)\varphi(x)\,dx\Big| + \frac{\varepsilon}{2},$$

d'où la proposition. □

Tout ceci est bel et bien, mais nous sommes toujours dans le cadre classique des fonctions. Voyons un autre exemple qui montre que ces considérations amènent naturellement à sortir de ce cadre. Soient f une fonction positive de $L^1(\mathbb{R}^d)$ d'intégrale 1, nulle en dehors d'une boule et $(\varepsilon_n)_{n\in\mathbb{N}}$ une suite de nombre réels tendant vers 0, et considérons la suite définie par

$$f_{\varepsilon_n}(x) \stackrel{\text{déf}}{=} \varepsilon_n^{-d} f\Big(\frac{x}{\varepsilon_n}\Big),$$

cf. le théorème 5.4.5. Cette suite est bornée dans $L^1(\mathbb{R}^d)$, néanmoins, pour n'importe quelle sous-suite $(\varepsilon_{n_p})_{p\in\mathbb{N}}$, il n'existe aucune fonction $\widetilde{f}$ appartenant à $L^1_{\text{loc}}(\mathbb{R}^d)$ telle que, pour toute fonction $\varphi \in \mathscr{D}(\mathbb{R}^d)$, on ait

$$\lim_{p\to\infty} \int_{\mathbb{R}^d} f_{\varepsilon_{n_p}}(x)\varphi(x)\,dx = \int_{\mathbb{R}^d} \widetilde{f}(x)\varphi(x)\,dx.$$

En effet, pour toute fonction bornée g nulle sur un voisinage de l'origine, on a

$$\int_{\mathbb{R}^d} f_{\varepsilon_n}(x)g(x)\,dx = 0$$

à partir d'un certain rang n puisqu'alors les supports de f_{ε_n} et de g sont disjoints. Donc, d'après le lemme 5.1.4, si la fonction $\widetilde{f}$ existe, alors elle est nulle en dehors de l'origine. Donc elle est nulle dans L^1. Mais ceci est impossible puisque si une fonction g vaut identiquement 1 près de l'origine, alors

$$\int_{\mathbb{R}^d} f_{\varepsilon_n}(x)g(x)\,dx = 1$$

à partir d'un certain rang, ce qui est contradictoire.

En tant que suite de fonctions, la suite construite ci-dessus converge uniformément vers 0 en dehors de tout voisinage de 0 et vers $+\infty$ ou 0 en 0 suivant les valeurs de $f(0)$. Ce qui précède montre que cette convergence ponctuelle perd certaines informations. Vers quoi donc converge alors la suite $(f_{\varepsilon_n})_{n\in\mathbb{N}}$ quand on la regarde comme suite de formes linéaires ? La proposition suivante répond à la question.

Proposition 6.1.2 *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de $L^1(\mathbb{R}^d)$ positives et d'intégrale* 1 *telle que, pour tout entier n, le support de f_n soit inclus dans une boule de rayon α_n, la suite de réels strictement positifs $(\alpha_n)_{n\in\mathbb{N}}$ vérifiant $\lim\limits_{n\to\infty} \alpha_n = 0$. On a alors*

$$\forall\, \varphi \in C_0(\mathbb{R}^d),\ \lim_{n\to\infty} \int_{\mathbb{R}^d} f_n(x)\varphi(x)\, dx = \varphi(0).$$

Démonstration. En effet, on peut écrire

$$\begin{aligned} |\langle f_n, \varphi\rangle - \varphi(0)| &= \left| \int_{\mathbb{R}^d} f_n(x)\varphi(x)\, dx - \varphi(0) \int_{\mathbb{R}^d} f_n(x)\, dx \right| \\ &\le \int_{\mathbb{R}^d} f_n(x)|\varphi(x) - \varphi(0)|\, dx \\ &= \int_{B(0,\alpha_n)} f_n(x)|\varphi(x) - \varphi(0)|\, dx \\ &\le \sup_{x\in B(0,\alpha_n)} |\varphi(x) - \varphi(0)|. \end{aligned}$$

La continuité de la fonction φ assure la conclusion de la proposition. □

On voit donc que la suite f_n converge de ce point de vue vers la forme linéaire $\varphi \mapsto \varphi(0)$. Cette proposition admet une interprétation physique. Si les fonctions f_n sont vues comme des densités de masse ou de charge électrique qui se concentrent autour de l'origine en gardant une masse totale égale à 1, la « limite » doit être vue comme une masse ou une charge ponctuelle portée par l'origine.

Du point de vue mathématique, on voit ici un exemple d'une suite de fonctions bornée dans L^1 qui souhaite sortir à la limite de l'espace L^1 (ce qui ne se produit pas dans les L^p, $p > 1$) pour converger vers un objet d'une toute autre nature.

Considérons maintenant la suite $(g_n)_{n\in\mathbb{N}}$ de fonctions de la variable réelle définie par

$$g_n(x) = -n^2 \text{ si } x \in \left]-\frac{1}{n}, 0\right[,\ g_n(x) = n^2 \text{ si } x \in \left]0, \frac{1}{n}\right[\text{ et } g_n(x) = 0 \text{ sinon.}$$

Cette suite n'est même pas bornée dans $L^1(\mathbb{R})$. Néanmoins, soit φ une fonction de $\mathscr{D}(\mathbb{R})$, alors on a

$$\langle g_n, \varphi\rangle = \varphi(0)\int_{\mathbb{R}} g_n(x)\,dx + \varphi'(0)\int_{\mathbb{R}} x g_n(x)\,dx \\ + \int_{\mathbb{R}} (\varphi(x) - \varphi(0) - x\varphi'(0)) g_n(x)\,dx.$$

Le premier terme est nul à cause de la parité des fonctions g_n. Le second terme vaut $\varphi'(0)$. Quant au troisième, il est majoré par

$$\frac{\|\varphi''\|_{L^\infty}}{2}\int_{\mathbb{R}} x^2 g_n(x)\,dx = \frac{\|\varphi''\|_{L^\infty}}{3n}.$$

La convergence mise en évidence ici

$$\langle g_n, \varphi\rangle \to \varphi'(0)$$

est la convergence vers une forme linéaire définie sur les fonctions de classe C^1. Comme dans le cas précédent, la notion de convergence ponctuelle ne donne aucun renseignement pertinent. Par contre, la limite peut s'interpréter en disant que la limite d'une charge totale n positive située entre 0 et $1/n$ et d'une charge totale n négative située entre $-1/n$ et 0 est un dipôle de moment 1.

Une des raisons majeures pour introduire les distributions est que celles-ci permettent de formuler et de résoudre bon nombre de problèmes d'équations aux dérivées partielles. Nous allons voir que l'on peut dériver indéfiniment toute distribution, même s'il s'agit d'une fonction qui n'est pas dérivable au sens classique.

Donnons un exemple tiré de la physique. Soit Ω un ouvert borné de $\mathbb{R}^3$. On suppose que la frontière de Ω est couverte d'une fine couche de métal totalement conductrice. On place en un point x_0 de Ω, par ailleurs supposé vide, une charge électrique ponctuelle, disons un électron, et l'on veut calculer le potentiel électrique qui s'établit dans Ω. La physique nous dit qu'il s'agit d'une fonction u définie sur $\bar{\Omega}$ telle que $u = 0$ sur $\partial\Omega$, qui est une surface équipotentielle, et telle que $\Delta u = q\delta_{x_0}$ où le Laplacien est l'opérateur différentiel $\Delta = \sum_{i=1}^{3} \frac{\partial^2}{\partial x_i^2}$, q est la charge de l'électron et δ_{x_0} est la masse de Dirac en x_0, objet que nous avons déjà rencontré en tant que mesure, et qui revient ici en tant que distribution comme limite de densités de charges qui se concentrent en x_0 et dont la définition comme forme linéaire est $\langle \delta_{x_0}, \varphi\rangle = \varphi(x_0)$. Or cette masse de Dirac n'est absolument pas une fonction. On voit donc mal comment u pourrait être une fonction classiquement deux fois dérivable. Donner un sens à l'équation de Laplace précédente dans un contexte classique n'est pas aisé. Par contre, celle-ci se formule très naturellement dans le cadre des distributions où la dérivation est un jeu d'enfant.

Prenons un second exemple, encore issu de la physique. Considérons une corde élastique infiniment longue et infiniment mince et intéressons-nous aux déplacements transverses de faible amplitude de cette corde. On cherche donc une fonction

$u(x,t)$ qui représente le déplacement transverse du point de la corde d'abscisse x à l'instant t. Ce déplacement doit satisfaire l'équation des ondes $\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0$ (cette équation modélise également la propagation d'une onde électromagnétique plane dans un milieu homogène infini). Parmi toutes les solutions se trouvent les ondes progressives qui sont des fonctions de la forme $u(x,t) = f(x \pm ct)$ avec f de classe C^2, des profils qui se déplacent vers la droite ou vers la gauche à la vitesse c sans changer de forme, dont on vérifie aisément qu'elles satisfont l'équation des ondes. Or la restriction f de classe C^2 n'est pas du tout physique et l'on souhaiterait pouvoir considérer des solutions moins régulières, par exemple f Lipschitzienne, disons affine par morceaux pour fixer les idées. Une telle fonction n'étant pas dérivable ne peut en aucun cas être considérée comme une solution de l'équation des ondes au sens classique. C'est pourtant une solution physiquement acceptable et en fait on peut voir qu'elle satisfait bien l'équation des ondes en tant que distribution.

Il y a même des situations plus compliquées (dynamique des gaz, écoulement de l'air autour d'un avion) dans lesquelles la physique s'intéresse explicitement à des solutions carrément discontinues d'équations aux dérivées partielles, ce sont les ondes de choc. Tout ceci s'exprime aisément en termes de distributions. On pourrait continuer longtemps cette liste.

6.2 Définition des distributions et premières propriétés

Au début de cette section, nous allons introduire des notations que nous utiliserons dans toute la suite du cours. Ces notations permettent de manipuler de façon pratique les dérivées partielles à tous ordres. Soit α un multi-entier, c'est-à-dire un élément de $\mathbb{N}^d$, on appelle longueur de α et l'on note $|\alpha|$ l'entier $\alpha_1 + \cdots + \alpha_d$. La notation $\beta \leq \alpha$ signifie que, pour tout j, $\beta_j \leq \alpha_j$; la notation $\beta < \alpha$ signifie que $\beta \leq \alpha$ et que $\beta \neq \alpha$. Enfin, si f est une fonction $|\alpha|$ fois différentiable, sur un ouvert de $\mathbb{R}^d$, alors, on pose

$$\partial^\alpha f \stackrel{\text{déf}}{=} \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}.$$

Ainsi, pour $\alpha = (0,0,\ldots,0)$, on a $\partial^\alpha f = f$, pour $\alpha = (1,0,\ldots,0)$, on a $\partial^\alpha f = \frac{\partial f}{\partial x_1}$, pour $\alpha = (1,1,\ldots,0)$, on a $\partial^\alpha f = \frac{\partial^2 f}{\partial x_1 \partial x_2}$, etc.

Définition 6.2.1 *On appelle* fonction-test *sur un ouvert Ω de $\mathbb{R}^d$ une fonction de $\mathscr{D}(\Omega;\mathbb{K}) = \mathscr{D}(\Omega)$. Pour tout compact K de l'ouvert Ω, on désigne par $\mathscr{D}_K$ l'ensemble des fonctions-test à support dans K.*

Définition 6.2.2 *Une forme linéaire u sur $\mathscr{D}(\Omega)$ est appelée* distribution *sur Ω si et seulement si, pour tout compact K de Ω, il existe un entier N et une constante C (qui dépendent a priori de K) tels que*

$$\forall\, \varphi \in \mathscr{D}_K,\ |\langle u, \varphi\rangle| \le C \sup_{|\alpha|\le N} \|\partial^\alpha \varphi\|_{L^\infty}. \tag{6.1}$$

L'ensemble des distributions sur Ω est un espace vectoriel noté $\mathscr{D}'(\Omega)$.

Remarque. Si $\mathbb{K} = \mathbb{R}$, on parle de distributions à valeurs réelles, si $\mathbb{K} = \mathbb{C}$ de distributions à valeurs complexes. Dans les deux cas, il s'agit visiblement d'imposer une condition de continuité sur les formes linéaires u considérées. Comme la topologie sur $\mathscr{D}(\Omega)$ qui est cachée derrière est un peu compliquée (elle n'est pas normée, ni même métrisable), et que l'on en a rarement, sinon jamais, besoin dans les applications, nous la passerons allègrement sous silence. Il est néanmoins utile de savoir dire quand une suite converge dans $\mathscr{D}(\Omega)$.

Définition 6.2.3 *On dit qu'une suite φ_n d'éléments de $\mathscr{D}(\Omega)$ converge vers $\varphi \in \mathscr{D}(\Omega)$ au sens de $\mathscr{D}(\Omega)$ si les deux conditions suivantes sont satisfaites :*
i) Il existe un compact $K \subset \Omega$ qui contient le support de tous les φ_n,
ii) Pour tout multi-entier α, $\partial^\alpha \varphi_n$ converge uniformément vers $\partial^\alpha \varphi$.

Cette définition de convergence permet en fait de caractériser en pratique les distributions.

Proposition 6.2.1 *Soit u une distribution sur Ω et $(\varphi_n)_{n\in\mathbb{N}}$ une suite de fonctions-test qui converge vers φ dans $\mathscr{D}(\Omega)$. Alors on a*

$$\langle u, \varphi_n\rangle \to \langle u, \varphi\rangle. \tag{6.2}$$

Réciproquement, si une forme linéaire u sur $\mathscr{D}(\Omega)$ satisfait la propriété (6.2) pour toute suite $(\varphi_n)_{n\in\mathbb{N}}$ qui converge dans $\mathscr{D}(\Omega)$, alors c'est une distribution.

Démonstration. Soit u une distribution et $(\varphi_n)_{n\in\mathbb{N}}$ une suite qui converge vers φ dans $\mathscr{D}(\Omega)$. Par la définition 6.2.3, il existe un compact K qui contient tous les supports des φ_n ainsi que celui de φ. On peut donc trouver un entier N et une constante C tels que l'inégalité (6.1) a lieu. Or $\varphi_n - \varphi$ est aussi à support dans K, donc

$$|\langle u, \varphi_n\rangle - \langle u, \varphi\rangle| = |\langle u, \varphi_n - \varphi\rangle| \le C \sup_{|\alpha|\le N} \|\partial^\alpha(\varphi_n - \varphi)\|_{L^\infty},$$

d'où la conclusion puisque toutes les dérivées partielles de φ_n convergent uniformément vers celles de φ sur K, et que l'on prend le sup sur un nombre fini d'entre elles.

La réciproque est nettement plus délicate et nous l'admettrons. □

Remarque de première importance. Bien que nous ne l'ayons pas démontrée, la réciproque de la proposition 6.2.1 fournit un moyen pratique de montrer qu'une forme linéaire sur $\mathscr{D}(\Omega)$ donnée est une distribution : il suffit de la tester sur les suites convergentes dans $\mathscr{D}(\Omega)$ (même si la topologie de $\mathscr{D}(\Omega)$ n'est pas du tout une topologie métrisable). Il est bien souvent plus facile de vérifier cette continuité séquentielle que de montrer l'existence pour tout compact K d'un entier N et d'une contante C tels que (6.1) ait lieu.

Donnons maintenant la définition de l'ordre d'une distribution.

Définition 6.2.4 *Soit u une distribution sur Ω. Elle est dite* d'ordre fini *si et seulement si il existe un entier N_0 tel que l'inégalité* (6.1) *soit vraie avec $N = N_0$ pour tout compact K de Ω. L'*ordre *de la distribution u est alors défini comme étant le plus petit entier positif vérifiant cette propriété.*

Voyons tout de suite quelques exemples. Soit δ_0 la forme linéaire définie par

$$\begin{aligned} \delta_0 : \mathscr{D}(\mathbb{R}^d) &\to \mathbb{K} \\ \varphi &\mapsto \varphi(0). \end{aligned}$$

De toute évidence, cette forme linéaire est une distribution sur $\mathbb{R}^d$ d'ordre 0. On l'appelle la *masse de Dirac* en 0. De même, la forme linéaire δ_0' définie par

$$\begin{aligned} \delta_0' : \mathscr{D}(\mathbb{R}) &\to \mathbb{K} \\ \varphi &\mapsto -\varphi'(0). \end{aligned}$$

est une distribution sur $\mathbb{R}$ d'ordre 1.

Il est maintenant essentiel de se demander ce que deviennent les fonctions usuelles dans ce cadre. Il faut se restreindre aux fonctions localement intégrables. En effet, la façon la plus simple de définir une forme linéaire sur $\mathscr{D}(\Omega)$ à partir d'une fonction f est de poser

$$\langle f, \varphi \rangle = \int_\Omega f(x)\varphi(x)\, dx$$

pour toute fonction-test φ. Or pour que l'intégrale ait un sens, il faut bien que la fonction f soit dans $L^1(K)$ pour tout compact K de Ω puisque les fonctions-test φ sont à support compact. C'est le minimum que l'on puisse exiger. Nous considérons toujours les fonctions $L^1_{\mathrm{loc}}(\Omega)$ comme des distributions grâce au théorème suivant.

Théorème 6.2.1 *Soit ι l'application linéaire définie par*

$$\begin{aligned} \iota : L^1_{\mathrm{loc}}(\Omega) &\to \mathscr{D}'(\Omega) \\ f &\mapsto \iota(f) : \varphi \mapsto \int_\Omega f(x)\varphi(x)\, dx \end{aligned}$$

L'application $\imath$ est une injection. De plus, elle est continue au sens où, pour tout compact K de Ω, on a

$$\forall \varphi \in \mathscr{D}_K,\ |\langle \imath(f), \varphi\rangle| \leq \|f\|_{L^1(K)}\|\varphi\|_{L^\infty}.$$

Démonstration. Vérifions d'abord que $\imath(f)$ est bien une distribution. C'est clair car si $\varphi \in \mathscr{D}_K$, alors

$$|\langle \imath(f), \varphi\rangle| \leq \int_\Omega |f(x)\varphi(x)|\,dx \leq \|f\|_{L^1(K)}\|\varphi\|_{L^\infty},$$

c'est-à-dire que l'inégalité de définition est satisfaite avec $C = \|f\|_{L^1(K)}$ et $N = 0$, d'où aussi la continuité de l'application $\imath$. Montrons qu'il s'agit d'une injection. Pour cela, considérons $f \in L^1_{\text{loc}}(\Omega)$ qui est telle que $\imath(f) = 0$, c'est-à-dire $\langle \imath(f), \varphi\rangle = 0$ pour tout $\varphi \in \mathscr{D}(\Omega)$. Par densité des fonctions de $\mathscr{D}(\Omega)$ dans les fonctions continues à support compact, on en déduit que $\int_\Omega f(x)v(x)\,dx = 0$ pour toute fonction v continue à support compact.

Soit $B(x_0, r) \subset \Omega$. Comme on peut approcher presque partout la fonction caractéristique de tout ensemble mesurable de A par une suite de fonctions continues à support compact de façon dominée, on en déduit que $\int_A f(x)\,dx = 0$. Par linéarité de l'intégrale, il s'ensuit que $\int_{B(x_0,r)} f(x)g(x)\,dx = 0$ pour toute fonction étagée g à support dans $B(x_0, r)$. Comme toute fonction mesurable bornée est limite uniforme croissante de telles fonctions étagées, on applique la proposition 5.1.6 pour conclure que $f = 0$ p.p. sur $B(x_0, r)$, donc finalement, $f = 0$ sur Ω. $\square$

On voit que cette identification, que nous ferons systématiquement et sans jamais plus utiliser la notation $\imath(f)$, nous amène à considérer les fonctions localement intégrables comme des distributions d'ordre 0.

Avertissement sans frais. Si les fonctions localement intégrables s'identifient à des distributions à l'aide de l'intégrale, il faut bien se garder d'écrire le crochet d'une distribution générale avec une fonction-test comme une intégrale ! Une écriture du style $\int_\Omega u(x)\varphi(x)\,dx$ où u est seulement une distribution n'a *absolument aucun sens* et doit être évitée à tout prix, bien qu'elle soit courante dans la littérature physique, au moins tant que l'on ne s'est pas assuré le cas échéant que u se trouve être en fait une fonction localement intégrable.

Nous allons maintenant donner une définition de la convergence d'une suite de distributions.

Définition 6.2.5 *Soient $(u_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathscr{D}'(\Omega)$ et u une distribution sur Ω. On dit que la suite $(u_n)_{n\in\mathbb{N}}$ converge vers u (au sens des distributions) si et seulement si*

$$\forall \varphi \in \mathscr{D}(\Omega),\ \langle u_n, \varphi\rangle \to \langle u, \varphi\rangle \text{ quand } n \to +\infty.$$

Bien que l'on puisse définir sur l'espace des distributions une topologie pour laquelle les suites convergentes sont précisément les suites satisfaisant à la définition ci-dessus, nous nous limiterons à ce concept de suite convergente qui est suffisant dans la plupart des applications. Cette définition de la convergence est exactement analogue à la notion de convergence faible-$*$ dans le dual d'un espace vectoriel normé que nous avons déjà rencontrée. Ceci n'est bien sûr pas fortuit.

Proposition 6.2.2 *Soit p un élément de $]1,\infty]$, on considère une suite bornée $(f_n)_{n\in\mathbb{N}}$ de fonctions de $L^p(\Omega)$ et f une distribution sur Ω. On a alors*

$$f_n \to f \text{ dans } \mathscr{D}'(\Omega) \quad \Longleftrightarrow \quad f \in L^p(\Omega) \text{ et } f_n \overset{*}{\rightharpoonup} f.$$

Démonstration. L'implication de droite à gauche est immédiate. Le fait que la suite $(f_n)_{n\in\mathbb{N}}$ converge dans $\mathscr{D}'(\Omega)$ signifie que

$$\forall \varphi \in \mathscr{D}(\Omega),\ \lim_{n\to\infty} \langle f_n, \varphi\rangle = \langle f, \varphi\rangle.$$

Mais, les distributions f_n étant des fonctions localement intégrables, on a

$$\langle f_n, \varphi\rangle = \int_\Omega f_n(x)\varphi(x)\,dx.$$

La suite $(f_n)_{n\in\mathbb{N}}$ étant bornée dans L^p, il existe une constante C telle que

$$\forall \varphi \in \mathscr{D}(\Omega),\ \left|\int_\Omega f_n(x)\varphi(x)\,dx\right| \le C\|\varphi\|_{L^{p'}}.$$

Par passage à la limite, on en déduit que

$$\forall \varphi \in \mathscr{D}(\Omega),\ |\langle f, \varphi\rangle| \le C\|\varphi\|_{L^{p'}}.$$

La distribution f est donc une forme linéaire sur $\mathscr{D}(\Omega)$ qui est continue pour la topologie de la norme $L^{p'}$. Comme par hypothèse, $p > 1$, on a $p' < +\infty$, donc $\mathscr{D}(\Omega)$ est dense dans $L^{p'}(\Omega)$. D'après le théorème de prolongement 2.2.2, cette forme linéaire admet un unique prolongement en une forme linéaire continue sur $L^{p'}(\Omega)$, notée $\widetilde{f}$. D'après le théorème 5.3.1, il existe alors une fonction f^* appartenant à $L^p(\Omega)$ telle que

$$\forall g \in L^{p'}(\Omega),\ \int_\Omega f^*(x)g(x)\,dx = \langle \widetilde{f}, g\rangle.$$

Cette égalité est *a fortiori* vraie pour toute fonction-test φ, d'où il vient

$$\forall \varphi \in \mathscr{D}(\Omega),\ \int_\Omega f^*(x)\varphi(x)\,dx = \langle f, \varphi\rangle.$$

D'après le théorème 6.2.1, ceci signifie que l'on peut identifier la distribution f à la fonction f^*, ce que l'on abrège usuellement en déclarant que f appartient à l'espace $L^p(\Omega)$. On peut donc écrire que

$$\forall\, \varphi \in \mathscr{D}(\Omega),\ \lim_{n\to\infty} \int_\Omega f_n(x)\varphi(x)\,dx = \int_\Omega f(x)\varphi(x)\,dx.$$

On conclut alors en appliquant la proposition 6.1.1. □

Le lien avec la convergence faible-$*$ apparaît également au travers du théorème suivant, analogue du théorème 3.3.1.

Théorème 6.2.2 *Soit $(u_n)_{n\in\mathbb{N}}$ une suite de distributions sur un ouvert Ω de $\mathbb{R}^d$ telle qu'il existe une forme linéaire u sur $\mathscr{D}(\Omega)$ telle que*

$$\forall\, \varphi \in \mathscr{D}(\Omega)\,,\ \lim_{n\to+\infty} \langle u_n, \varphi\rangle = \langle u, \varphi\rangle.$$

Alors u est une distribution, c'est-à-dire que

$$\forall\, K \text{ compact de } \Omega, \exists\, N, C, \forall\, \varphi \in \mathscr{D}_K,\ |\langle u, \varphi\rangle| \leq C \sup_{|\alpha|\leq N} \|\partial^\alpha \varphi\|_{L^\infty}.$$

Ce théorème (que nous admettons) est en fait une conséquence de l'extension du théorème de Banach-Steinhaus aux espaces de Fréchet (nous n'avons pas rencontré les espaces de Fréchet, disons rapidement qu'il s'agit d'espaces vectoriels munis d'une topologie rendant continues les opérations d'espace vectoriel, métrisable et complète, mais en général pas normée, en omettant une partie de l'histoire...).

Remarque. Ce théorème a une portée non négligeable. Il permet, lorsqu'une forme linéaire sur $\mathscr{D}(\Omega)$ est mise en évidence comme limite d'une suite de distributions, de se dispenser de vérifier que la forme linéaire limite vérifie les inégalités de la définition d'une distribution. Cela dit, en pratique, le plus souvent on doit établir des inégalités de ce type uniformément par rapport à n pour montrer la convergence ! Le gain est donc plus apparent que réel.

6.3 Opérations sur les distributions

Nous allons dans cette section, étendre aux distributions quelques opérations qui nous sont familières pour des fonctions régulières. Commençons par un exemple très simple, mais qui donne la tonalité générale, la conjugaison complexe (dans le cas des distributions à valeurs complexes).

Définition 6.3.1 *Soit u une distribution sur un ouvert Ω de $\mathbb{R}^d$. On définit la distribution $\bar{u}$ par*

$$\forall\, \varphi \in \mathscr{D}(\Omega),\ \langle \overline{u}, \varphi\rangle \stackrel{déf}{=} \overline{\langle u, \overline{\varphi}\rangle}.$$

Démonstration. Tout d'abord, $\overline{u}$ est bien une forme linéaire sur $\mathscr{D}(\Omega)$. Utilisons la proposition 6.2.1. Soit $(\varphi_n)_{n\in\mathbb{N}}$ une suite qui converge vers φ dans $\mathscr{D}(\Omega)$. Il est à peu près évident que $\overline{\varphi}_n \to \overline{\varphi}$ également dans $\mathscr{D}(\Omega)$. Comme u est une distribution, on en déduit que $\langle u, \overline{\varphi}_n\rangle \to \langle u, \overline{\varphi}\rangle$. Comme la conjugaison est continue de $\mathbb{C}$ dans $\mathbb{C}$, on voit que $\langle \overline{u}, \varphi_n\rangle \to \langle \overline{u}, \varphi\rangle$ et $\overline{u}$ est une distribution. □

On voit que l'idée sous-jacente est de travailler par transposition, en passant toute la difficulté (éventuelle) sur la fonction-test.

Il convient de vérifier que, lorsque f est une fonction $L^1_{\text{loc}}(\Omega)$, la conjugaison au sens des distributions coïncide avec la conjugaison ordinaire pour les fonctions. En effet, pour toute fonction-test φ,

$$\langle \overline{\imath(f)}, \varphi\rangle = \overline{\langle \imath(f), \overline{\varphi}\rangle} = \overline{\int_\Omega f\,\overline{\varphi}\,dx} = \int_\Omega \overline{f}\varphi\,dx = \langle \imath(\overline{f}), \varphi\rangle,$$

donc les deux formes linéaires $\overline{\imath(f)}$ et $\imath(\overline{f})$ sont égales.

De plus, la conjugaison définie ci-dessus est séquentiellement continue, c'est-à-dire que si l'on a une suite $(u_n)_{n\in\mathbb{N}}$ de distributions convergeant vers une distribution u, alors $\overline{u_n} \to \overline{u}$ au sens de $\mathscr{D}'(\Omega)$. En effet, pour toute fonction-test φ, on a

$$\lim_{n\to\infty} \langle u_n, \overline{\varphi}\rangle = \langle u, \overline{\varphi}\rangle.$$

Il en résulte que

$$\begin{aligned}\lim_{n\to\infty} \overline{\langle u_n, \overline{\varphi}\rangle} &= \overline{\langle u, \overline{\varphi}\rangle}\\ &= \langle \overline{u}, \varphi\rangle.\end{aligned}$$

Nous allons définir sur le même modèle, c'est-à-dire par transposition, diverses opérations sur les distributions. La plus importante, et de très loin, est la dérivation. C'est elle qui rend les distributions aussi essentielles en analyse.

Définition 6.3.2 *Soit D_j la forme linéaire sur $\mathscr{D}(\Omega)$ définie par*

$$\big\langle D_j u, \varphi\big\rangle = -\Big\langle u, \frac{\partial \varphi}{\partial x_j}\Big\rangle.$$

Alors, $D_j u$ est une distribution que l'on appelle dérivée partielle par rapport à la j-ème variable d'une distribution *u et que l'on note* $\frac{\partial u}{\partial x_j}$.

Démonstration. C'est clair car si $\varphi_n \to \varphi$ dans $\mathscr{D}(\Omega)$, on a visiblement $\partial\varphi_n/\partial x_j \to \partial\varphi/\partial x_j$ dans $\mathscr{D}(\Omega)$ (le vérifier !), donc

$$\big\langle D_j u, \varphi_n\big\rangle = -\Big\langle u, \frac{\partial \varphi_n}{\partial x_j}\Big\rangle \to -\Big\langle u, \frac{\partial \varphi}{\partial x_j}\Big\rangle = \big\langle D_j u, \varphi\big\rangle.$$

On conclut à l'aide de la proposition 6.2.1. □

La dérivée d'une distribution est donc toujours définie et c'est une distribution. Il convient de vérifier si cette définition — ainsi que la notation afférente — est cohérente avec la notion de dérivation usuelle quand celle-ci s'applique. C'est l'objet de la proposition suivante.

Proposition 6.3.1 *Soit $f \in C^1(\Omega)$. On a*

$$D_j f = \frac{\partial f}{\partial x_j}.$$

Démonstration. Si f est de classe C^1 sur Ω, $\partial f/\partial x_j$ est continue sur Ω, donc ces deux fonctions sont localement intégrables et sont par conséquent des distributions. Par définition de D_j, on a

$$\forall \varphi \in \mathscr{D}(\Omega), \quad \langle D_j f, \varphi\rangle = -\Big\langle f, \frac{\partial \varphi}{\partial x_j}\Big\rangle = -\int_\Omega f(x)\frac{\partial \varphi}{\partial x_j}(x)\,dx$$

car f est L^1_{loc}. La fonction φ est à support compact, donc il existe une bande $B = \{x \in \mathbb{R}^d; |x_j| < M\}$ telle que $\operatorname{supp}\varphi \subset B$. Comme le produit $f\partial\varphi/\partial x_j$ est identiquement nul en dehors de $\operatorname{supp}\varphi$, on peut étendre f de façon quelconque, mais C^1, à B tout entier sans modifier l'intégrale. Par le théorème de Fubini, il vient donc

$$\int_\Omega f(x)\frac{\partial \varphi}{\partial x_j}(x)\,dx = \int_B f(x)\frac{\partial \varphi}{\partial x_j}(x)\,dx = \int_{\mathbb{R}^{d-1}}\Big(\int_{-M}^{M} f(x)\frac{\partial \varphi}{\partial x_j}(x)\,dx_j\Big)dx'$$

où x' désigne le point $(x_1,\dots,x_{j-1},x_{j+1},\dots,x_d)$. Une intégration par parties usuelle dans la variable x_j donne

$$\int_\Omega f(x)\frac{\partial \varphi}{\partial x_j}(x)\,dx = \int_{\mathbb{R}^{d-1}}[f(x)\varphi(x)]_{x_j=-M}^{x_j=M} - \int_{\mathbb{R}^{d-1}}\Big(\int_{-M}^{M} \frac{\partial f}{\partial x_j}(x)\varphi(x)\,dx_j\Big)dx'.$$

Or, par définition de la bande B, si $|x_j| = M$, alors $\varphi(x) = 0$. Finalement, en réappliquant Fubini, on obtient

$$\forall \varphi \in \mathscr{D}(\Omega), \quad \langle D_j f, \varphi\rangle = \int_\Omega \frac{\partial f}{\partial x_j}(x)\varphi(x)\,dx = \Big\langle \frac{\partial f}{\partial x_j}, \varphi\Big\rangle,$$

ce qui montre le résultat à l'aide de l'identification des fonctions localement intégrables avec des distributions. □

Remarques importantes. On est donc parfaitement fondé à utiliser la notation des dérivées partielles usuelles pour la dérivation des distributions. Bien sûr, on peut dériver ainsi toutes les fonctions localement intégrables, même celles qui ne sont absolument pas dérivables au sens classique ! Évidemment, dans ce cas, la dérivée ne sera pas une fonction, mais une distribution. De plus, la dérivation est une application trivialement séquentiellement continue dans $\mathscr{D}'$, au sens où

$$u_n \to u \text{ dans } \mathscr{D}'(\Omega) \Longrightarrow \frac{\partial u_n}{\partial x_j} \to \frac{\partial u}{\partial x_j} \text{ dans } \mathscr{D}'(\Omega).$$

En effet, $\langle \partial u_n/\partial x_j, \varphi\rangle = -\langle u_n, \partial\varphi/\partial x_j\rangle \to -\langle u, \partial\varphi/\partial x_j\rangle = \langle \partial u/\partial x_j, \varphi\rangle$. Cette propriété frappante n'a pas d'analogue pour la dérivation au sens classique. Il faut naturellement bien garder à l'esprit que la convergence en question est très faible. La formule *qu'il faut retenir* est

$$\Big\langle \frac{\partial u}{\partial x_j}, \varphi \Big\rangle = -\Big\langle u, \frac{\partial \varphi}{\partial x_j}\Big\rangle.$$

L'opération de dérivation peut bien sûr être itérée. On peut donc définir les dérivées partielles successives à tout ordre par la formule

$$\langle \partial^\alpha u, \varphi\rangle = (-1)^{|\alpha|} \langle u, \partial^\alpha \varphi\rangle .$$

Encore une remarque. On a démontré en cours de route que si f est C^1 sur Ω et g est C^1 *à support compact* sur Ω, alors

$$\int_\Omega \frac{\partial f}{\partial x_j}(x) g(x)\, dx = -\int_\Omega f(x) \frac{\partial g}{\partial x_j}(x)\, dx.$$

C'est un cas particulier de la *formule d'intégration par partie en dimension d*, une formule de la plus grande importance qu'il convient de bien connaître. Quand on ne suppose pas g à support compact, la formule d'intégration par partie devient beaucoup plus compliquée : l'analogue du terme tout intégré en dimension 1 est en effet assez délicat à définir.

Exemple. Il est facile de vérifier que la distribution δ_0' définie plus haut est la dérivée de la masse de Dirac en 0 sur $\mathbb{R}$.

Voyons maintenant quelques opérations supplémentaires de moindre importance, comme la dilatation.

Définition 6.3.3 *Soit λ un réel strictement positif. On définit l'application A_λ par*

$$A_\lambda \varphi(x) = \lambda^{-d} \varphi\left(\frac{x}{\lambda}\right).$$

Pour tout $u \in \mathscr{D}'(\mathbb{R}^d)$, on définit la forme linéaire u_λ par $\langle u_\lambda, \varphi\rangle = \langle u, A_\lambda \varphi\rangle$. Alors u_λ est une distribution que l'on appelle la dilatée *de u de rapport λ.*

Démonstration. C'est tout aussi clair que précédemment. Par ailleurs, on étend ainsi aux distributions la dilatation usuelle. En effet, pour tout fonction f localement intégrable, on a

$$\langle f_\lambda, \varphi\rangle = \langle f, A_\lambda \varphi\rangle = \int_{\mathbb{R}^d} f(x)\lambda^{-d}\varphi\left(\frac{x}{\lambda}\right) dx = \int_{\mathbb{R}^d} f(\lambda x)\varphi(x)\, dx,$$

d'où $f_\lambda(x) = f(\lambda x)$. Notons enfin que si $u_n \to u$ dans $\mathscr{D}'(\mathbb{R}^d)$, alors $u_{n,\lambda} \to u_\lambda$ dans $\mathscr{D}'(\mathbb{R}^d)$, tout aussi trivialement. □

Dans la même veine, on a la translation.

Définition 6.3.4 *Soient a un vecteur de $\mathbb{R}^d$ et A_a l'application définie par*

$$A_a\varphi(x) = \varphi(x+a).$$

Pour tout $u \in \mathscr{D}'(\mathbb{R}^d)$, on définit la forme linéaire τ_a par $\langle \tau_a u, \varphi\rangle = \langle u, A_a\varphi\rangle$. Alors $\tau_a u$ est une distribution que l'on appelle translatée *de u par a.*

Démonstration. Remarquons que cela étend bien la notion de translation des fonctions car si f est une fonction localement intégrable, on a,

$$\int_{\mathbb{R}^d} f(x-a)\varphi(x)\, dx = \int_{\mathbb{R}^d} f(x)\varphi(x+a)\, dx.$$

□

La dilatation et la translation sont des cas particuliers de changements de variables pour les distributions.

Définition 6.3.5 *Soit Θ un C^∞-difféomorphisme d'un ouvert Ω de $\mathbb{R}^d$ dans lui-même. On définit l'application A_Θ par*

$$A_\Theta\varphi(x) \stackrel{\text{déf}}{=} J(\Theta^{-1}(x))^{-1}\varphi(\Theta^{-1}(x))$$

où $J(y) = |\det \nabla\Theta(y)|$ est le jacobien de Θ, $\nabla\Theta$ désignant ici la matrice $d\times d$ des dérivées partielles premières des composantes de Θ. Pour tout $u \in \mathscr{D}'(\Omega)$, on définit la forme linéaire $u\circ\Theta$ par $\langle u\circ\Theta, \varphi\rangle = \langle u, A_\Theta\varphi\rangle$. Alors $u\circ\Theta$ est une distribution sur Ω que l'on appelle la composée *de u avec Θ.*

Démonstration. Il suffit de remarquer que $A_\Theta\varphi \in \mathscr{D}(\Omega)$. De plus, d'après la formule de Leibniz, on a

$$\partial^\alpha(A_\Theta\varphi) = \sum_{\beta\le\alpha} C_\beta^\alpha \partial^{\alpha-\beta}(J(\Theta^{-1}(x))^{-1})\partial^\beta(\varphi\circ\Theta^{-1}).$$

La formule de dérivation composée d'ordre quelconque s'écrit

$$\partial^\beta(\varphi\circ\Theta^{-1}) = \sum_{|\alpha|=1}^{|\beta|} \partial^\alpha\varphi(\Theta^{-1})P_\alpha(\partial^\gamma(\Theta^{-1})_i),$$

où P_α est un polynôme en les dérivées partielles $\partial^\gamma(\Theta^{-1})_i$ des composantes de Θ^{-1} avec $1 \le |\gamma| \le |\beta| - |\alpha| + 1$. Il en résulte que si $\varphi_n \to \varphi$ dans $\mathscr{D}(\Omega)$, alors $A_\Theta\varphi_n \to A_\Theta\varphi$ dans $\mathscr{D}(\Omega)$, d'où le résultat.

Il convient de vérifier que l'on a ainsi étendu aux distributions la composition usuelle. En effet, pour tout fonction f localement intégrable, on a

$$\langle f\circ\Theta,\varphi\rangle = \langle f, A_\Theta\varphi\rangle = \int_\Omega f(x)J(\Theta^{-1}(x))^{-1}\varphi(\Theta^{-1}(x))\,dx = \int_\Omega f(\Theta(y))\varphi(y)\,dy,$$

en posant $x = \Theta(y)$. □

On peut multiplier toute distribution par une fonction de classe C^∞.

Définition 6.3.6 *Soit f une fonction indéfiniment différentiable sur un ouvert Ω de $\mathbb{R}^d$. Pour tout $u \in \mathscr{D}'(\Omega)$, on définit la forme linéaire $A_f u$ par $\langle A_f u,\varphi\rangle = \langle u, f\varphi\rangle$. Alors $A_f u$ est une distribution sur Ω que l'on note fu.*

Démonstration. Il suffit de remarquer que $f\varphi \in \mathscr{D}(\Omega)$. De plus, d'après la formule de Leibniz, on a

$$\partial^\alpha(f\varphi) = \sum_{\beta\le\alpha} C_\beta^\alpha \partial^{\alpha-\beta} f \partial^\beta\varphi.$$

Donc, si $\varphi_n \to \varphi$ dans $\mathscr{D}(\Omega)$, visiblement $f\varphi_n \to f\varphi$ dans $\mathscr{D}(\Omega)$, d'où le résultat.

De plus, si g est localement intégrable, alors

$$\langle A_f g,\varphi\rangle = \langle g, f\varphi\rangle = \int_\Omega g(x)f(x)\varphi(x)\,dx,$$

d'où $A_f g = gf$ et l'explication du terme de multiplication d'une distribution par une fonction C^∞. □

Remarque. On ne peut pas en général multiplier une distribution par une fonction qui soit moins que C^∞, encore moins par une autre distribution. En fait, on peut montrer qu'il n'existe aucune multiplication raisonnable sur l'espace $\mathscr{D}'(\Omega)$.

On va maintenant définir et étudier la convolution d'une distribution sur $\mathbb{R}^d$ par une fonction indéfiniment différentiable à support compact. Pour toute fonction ψ définie sur $\mathbb{R}^d$, on note $\check\psi$ la fonction définie sur $\mathbb{R}^d$ par

$$\check\psi(x) \overset{\text{déf}}{=} \psi(-x).$$

Définition 6.3.7 *Soit $u \in \mathscr{D}'(\mathbb{R}^d)$ et $\psi \in \mathscr{D}(\mathbb{R}^d)$. On définit la convoluée de u avec ψ par*
$$\forall x \in \mathbb{R}^d, \quad u \star \psi(x) = \langle u, \check{\psi}(-x+\cdot)\rangle.$$

La convolution d'une distribution par une fonction indéfiniment différentiable à support compact est donc une fonction. Il est clair que si $u \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, on retrouve la convolution usuelle.

Proposition 6.3.2 *La fonction $u \star \psi$ est indéfiniment différentiable et l'on a*
$$\partial^\alpha(u\star\psi) = (\partial^\alpha u)\star\psi = u\star(\partial^\alpha\psi).$$

Démonstration. Soit e_i le i-ème vecteur de base. On considère le quotient différentiel
$$\frac{u\star\psi(x+he_i) - u\star\psi(x)}{h} = \Big\langle u, \frac{\check{\psi}(-x-he_i+\cdot) - \check{\psi}(-x+\cdot)}{h}\Big\rangle.$$
Considérons une suite $h_n \to 0$ avec $|h_n| \leq 1$. Les supports des fonctions $h_n^{-1}(\check{\psi}(-x-h_ne_i+\cdot) - \check{\psi}(-x+\cdot))$ sont contenus dans un même compact. De plus, par le théorème des accroissements finis, cette suite converge uniformément vers $-\partial_i\check{\psi}(-x+\cdot)$. De même, pour tout multi-entier α, on voit que $\partial^\alpha(h_n^{-1}(\check{\psi}(-x-h_ne_i+\cdot) - \check{\psi}(-x+\cdot)))$ converge uniformément vers $-\partial^\alpha(\partial_i\check{\psi}(-x+\cdot))$. On en déduit que cette suite converge dans $\mathscr{D}(\mathbb{R}^d)$, d'où
$$\frac{u\star\psi(x+h_ne_i) - u\star\psi(x)}{h_n} \to -\langle u, \partial_i\check{\psi}(-x+\cdot)\rangle = \langle u, \partial_i\psi(\check{x+\cdot})\rangle = \langle \partial_i u, \check{\psi}(-x+\cdot)\rangle,$$
d'où le résultat pour les dérivées d'ordre 1. Le cas général s'en déduit par récurrence. $\square$

Remarque. Ce résultat implique en particulier que $u\star\psi$ est une fonction localement intégrable, donc aussi une distribution. L'action de cette distribution sur les fonctions-test s'exprime également par transposition.

Proposition 6.3.3 *Pour toute distribution u et tout couple de fonctions indéfiniment différentiables à support compact (φ,ψ), on a*
$$\langle u\star\psi, \varphi\rangle = \langle u, \check{\psi}\star\varphi\rangle.$$

Démonstration. Par définition de la convolution, on a
$$\begin{aligned}\langle u\star\psi,\varphi\rangle &= \int_{\mathbb{R}^d}(u\star\psi)(x)\varphi(x)\,dx\\ &= \int_{\mathbb{R}^d}\langle u, \check{\psi}(-x+\cdot)\rangle\varphi(x)\,dx = \int_{\mathbb{R}^d}\langle u, \check{\psi}(-x+\cdot)\varphi(x)\rangle\,dx.\end{aligned}$$

Il suffit donc de montrer que pour toute fonction $\Phi \in \mathscr{D}(\mathbb{R}^d \times \mathbb{R}^d)$, on a

$$\int_{\mathbb{R}^d} \langle u, \Phi(x,\cdot)\rangle \, dx = \Big\langle u, \int_{\mathbb{R}^d} \Phi(x,\cdot)\, dx \Big\rangle.$$

Ce résultat est trivialement vrai si Φ est de la forme

$$\Phi(x,y) = \sum_{i=1}^{k} \phi_i(x)\psi_i(y).$$

On conclut par un argument de densité de l'ensemble de ces fonctions, le produit tensoriel $\mathscr{D}(\mathbb{R}^d) \otimes \mathscr{D}(\mathbb{R}^d)$, dans $\mathscr{D}(\mathbb{R}^d \times \mathbb{R}^d)$, densité que nous montrons un peu plus loin. □

On en déduit un résultat de densité des fonctions indéfiniment différentiables à support compact dans les distributions.

Proposition 6.3.4 *Pour tout $u \in \mathscr{D}'(\mathbb{R}^d)$, il existe une suite $(u_n)_{n\in\mathbb{N}}$ de fonctions de $\mathscr{D}(\mathbb{R}^d)$ telle que*

$$u_n \to u \text{ dans } \mathscr{D}'(\mathbb{R}^d).$$

Démonstration. Il suffit de prendre une suite régularisante ψ_n, une fonction θ de $\mathscr{D}(\mathbb{R}^d)$ identiquement égale à 1 sur la boule unité et de poser $u_n(x) = \theta(x/n)u \star \psi_n(x)$. La suite $\check{\psi}_n$ est aussi une suite régularisante, donc

$$\check{\psi}_n \star \varphi \to \varphi \text{ dans } \mathscr{D}(\mathbb{R}^d).$$

Comme il est clair que $\theta(x/n)\check{\psi}_n \star \varphi = \check{\psi}_n \star \varphi$ pour n assez grand, on en déduit le résultat grâce à la proposition 6.3.3. □

Montrons rapidement la densité du produit tensoriel évoquée plus haut. On rappelle que le produit tensoriel $\mathscr{D}(\mathbb{R}^d) \otimes \mathscr{D}(\mathbb{R}^d)$ est l'espace des fonctions définies sur $\mathbb{R}^d \times \mathbb{R}^d$ qui s'écrivent sous la forme $\Phi(x,y) = \sum_{i=1}^k \phi_i(x)\psi_i(y)$ pour un certain k avec ϕ_i et ψ_i éléments de $\mathscr{D}(\mathbb{R}^d)$. Il n'y a bien sûr pas unicité d'une telle décomposition.

Proposition 6.3.5 *Pour tout $\Phi \in \mathscr{D}(\mathbb{R}^d \times \mathbb{R}^d)$, il existe une suite $(\Phi_n)_{n\in\mathbb{N}}$ de fonctions de $\mathscr{D}(\mathbb{R}^d) \otimes \mathscr{D}(\mathbb{R}^d)$ telle que*

$$\Phi_n \to \Phi \text{ dans } \mathscr{D}(\mathbb{R}^d \times \mathbb{R}^d).$$

Démonstration. Soit $\Phi \in \mathscr{D}(\mathbb{R}^d \times \mathbb{R}^d)$ et K un compact de $\mathbb{R}^d \times \mathbb{R}^d$ dont l'intérieur contient le support de Φ, de la forme $K = [-\frac{L}{2}, \frac{L}{2}]^d \times [-\frac{L}{2}, \frac{L}{2}]^d$. On étend la restriction de Φ à K en une fonction $\widetilde{\Phi}$, K-périodique sur $\mathbb{R}^d \times \mathbb{R}^d$. Cette fonction admet donc un développement en série de Fourier

$$\widetilde{\Phi}(x,y) = \sum_{(\alpha,\beta)\in\mathbb{N}^d\times\mathbb{N}^d} c_{\alpha\beta} e^{i\frac{2\pi}{L}((x|\alpha)+(y|\beta))}$$

où $(\cdot|\cdot)$ désigne le produit scalaire usuel sur $\mathbb{R}^d$. Les coefficients de Fourier sont donnés par

$$\begin{aligned} c_{\alpha\beta} &= L^{-2d} \int_K e^{-i\frac{2\pi}{L}((x|\alpha)+(y|\beta))} \widetilde{\Phi}(x,y)\, dxdy \\ &= L^{-2d} \int_{\mathbb{R}^d\times\mathbb{R}^d} e^{-i\frac{2\pi}{L}((x|\alpha)+(y|\beta))} \Phi(x,y)\, dxdy. \end{aligned}$$

En intégrant cette formule k fois par parties, on voit que pour tout k, il existe une constante C_k telle que pour tous multi-indices α et β, on ait

$$|c_{\alpha\beta}| \leq C_k(1+|\alpha|+|\beta|)^{-k}.$$

Choisissons maintenant une fonction $\theta \in \mathscr{D}(\mathbb{R}^d)$ telle que $\theta\otimes\theta = 1$ sur le support de Φ et $\theta\otimes\theta$ est à support dans K. Il vient donc

$$\Phi(x,y) = \widetilde{\Phi}(x,y)\theta(x)\theta(y) = \sum_{(\alpha,\beta)\in\mathbb{N}^d\times\mathbb{N}^d} c_{\alpha\beta} e^{i\frac{2\pi}{L}(x|\alpha)}\theta(x) e^{i\frac{2\pi}{L}(y|\beta)}\theta(y).$$

Posant

$$\Phi_n(x,y) = \sum_{\max(|\alpha|,|\beta|)\leq n} c_{\alpha\beta} e^{i\frac{2\pi}{L}(x|\alpha)}\theta(x) e^{i\frac{2\pi}{L}(y|\beta)}\theta(y),$$

on voit que $\Phi_n \in \mathscr{D}(\mathbb{R}^d)\otimes\mathscr{D}(\mathbb{R}^d)$ et que $\Phi_n \to \Phi$ dans $\mathscr{D}(\mathbb{R}^d\times\mathbb{R}^d)$, puisque la série de chaque dérivée partielle est normalement convergente. Il suffit en effet de prendre k assez grand dans l'estimation des coefficients de Fourier ci-dessus. Par ailleurs, tous les termes de cette série sont à support dans K. □

Remarque. Le résultat de densité ci-dessus, légèrement généralisé à un produit de deux ouverts de deux espaces de dimension différente — mais c'est bien exactement la même preuve — permet de définir le produit tensoriel de deux distributions $u_1 \in \mathscr{D}'(\Omega_1)$ et $u_2 \in \mathscr{D}'(\Omega_2)$, avec Ω_i ouvert de $\mathbb{R}^{d_i}$. On définit en effet

$$\langle u_1\otimes u_2, \varphi_1\otimes\varphi_2\rangle = \langle u_1,\varphi_1\rangle\langle u_2,\varphi_2\rangle$$

puis $u_1\otimes u_2$ sur $\mathscr{D}(\Omega_1)\otimes\mathscr{D}(\Omega_2)$ par linéarité (le résultat ne dépendant pas de la décomposition), puis par densité sur $\mathscr{D}(\Omega_1\times\Omega_2)$. Ceci permet d'ailleurs d'établir que

$$\langle u_1\otimes u_2, \Phi\rangle = \langle u_1,\psi_1\rangle = \langle u_2,\psi_2\rangle$$

où
$$\psi_1(x) = \langle u_2, \Phi(x,\cdot)\rangle, \quad \psi_2(y) = \langle u_1, \Phi(\cdot,y)\rangle$$
formule que l'on a utilisé dans le cas particulier de $u \otimes 1$ plus haut.

À partir du produit tensoriel de deux distributions, il est possible de définir la convolution de deux distributions avec une plus grande généralité que ce que l'on a exposé plus haut. Néanmoins, il n'est pas possible de convoluer deux distributions quelconques. Pour que leur convolution soit définie, les deux distributions doivent satisfaire des conditions de support, notion que nous rencontrerons un peu plus loin.

6.4 Exemples de distributions et d'équations dans $\mathscr{D}'$

Nous avons déjà rencontré la masse de Dirac et sa dérivée. Considérons maintenant la fonction de Heaviside. La fonction Heaviside H est la fonction caractéristique des réels positifs. C'est une fonction de $L^\infty(\mathbb{R})$, elle est donc localement intégrable et c'est une distribution sur $\mathbb{R}$. Calculons sa dérivée au sens des distributions. Par définition, c'est la distribution H' définie par
$$\forall \varphi \in \mathscr{D}(\mathbb{R}), \quad \langle H', \varphi\rangle = -\langle H, \varphi'\rangle.$$
Par définition de H, on a
$$\begin{aligned}\langle H', \varphi\rangle &= -\int_0^\infty \varphi'(x)dx\\ &= \varphi(0).\end{aligned}$$
La dérivée de la fonction de Heaviside est donc la masse de Dirac. On a ainsi trouvé une primitive au sens des distributions à la masse de Dirac (à titre d'exercice, montrer qu'une primitive de H au sens des distributions est x_+, et calculer des primitives successives).

Remarquons que la fonction de Heaviside est dérivable partout au sens usuel sauf en 0, et que cette dérivée au sens usuel est nulle là où elle existe. La notion usuelle de dérivée est donc inopérante au point où justement se produit la variation de H. De la dérivée au sens des distributions de H, on déduit aisément la propriété suivante, dite formule des sauts.

Théorème 6.4.1 *Soit f une fonction de classe C^1 par morceaux sur $\mathbb{R}$. On désigne par $(a_n)_{1\le n\le N}$ ses points de discontinuité. Sa dérivée au sens des distributions est la somme de la dérivée usuelle sur $\mathbb{R}\setminus\{a_1,\cdots,a_N\}$ et de la distribution*
$$\sum_{n=1}^N (f(a_n^+) - f(a_n^-))\delta_{a_n},$$

où δ_{a_n} désigne la masse de Dirac au point a_n, i.e. la distribution définie par

$$\langle \delta_{a_n}, \varphi \rangle = \varphi(a_n).$$

On pourrait penser que la classe des fonctions localement intégrables est suffisamment large pour contenir toutes les fonctions que l'on aimerait pouvoir manipuler au sens des distributions. Il n'en est rien. Ainsi, la fonction $x \mapsto 1/x$ est localement intégrable sur $\mathbb{R}_+^*$ et y définit donc une distribution. Elle n'est par contre pas localement intégrable sur $\mathbb{R}$ et ne s'identifie pas directement à une distribution sur $\mathbb{R}$, en raison de sa singularité non intégrable en 0.

Pour contourner cette difficulté, on utilise l'artifice suivant. On définit une forme linéaire u sur $\mathscr{D}(\mathbb{R})$ par

$$\langle u, \varphi \rangle = \int_{-R}^{R} \frac{\varphi(x) - \varphi(0)}{x}\, dx,$$

où R est un réel strictement positif tel que le support de φ soit inclus dans l'intervalle $]-R, R[$. Vérifions d'abord que cette définition ne dépend pas du choix de R. En effet, une fonction-test φ étant donnée, choisissons un autre nombre réel R' ayant la même propriété vis-à-vis du support de φ. On a, en supposant que R est le plus grand des deux (sans perte de généralité),

$$\begin{aligned}\int_{-R}^{R} \frac{\varphi(x) - \varphi(0)}{x} dx - \int_{-R'}^{R'} \frac{\varphi(x) - \varphi(0)}{x}\, dx &= -\varphi(0)\Big(\int_{-R}^{-R'} \frac{dx}{x} + \int_{R'}^{R} \frac{dx}{x}\Big) \\ &= 0.\end{aligned}$$

Nous allons ensuite montrer que l'on a bien affaire à une distribution sur $\mathbb{R}$. C'est clair, car on a

$$\Big| \int_{-R}^{R} \frac{\varphi(x) - \varphi(0)}{x} dx \Big| \leq 2R \|\varphi'\|_{L^\infty},$$

par le théorème des accroissements finis. Donc, nous avons bien défini une distribution sur $\mathbb{R}$. Cette distribution u s'appelle la valeur principale de $1/x$ et elle est souvent notée $\mathrm{vp}\,\frac{1}{x}$. Pour comprendre cette terminologie, nous allons étudier son action sur des fonctions-test de $\mathscr{D}(\mathbb{R} \setminus \{0\})$. Soit φ une telle fonction-test et R un réel associé. Par définition, on a

$$\begin{aligned}\Big\langle \mathrm{vp}\,\frac{1}{x}, \varphi \Big\rangle &= \int_{-R}^{R} \frac{\varphi(x) - \varphi(0)}{x} dx \\ &= \int_{-R}^{R} \frac{\varphi(x)}{x} dx \\ &= \int_{\mathbb{R}} \frac{\varphi(x)}{x} dx \\ &= \Big\langle \Big(\frac{1}{x}\Big)_{|\mathbb{R}\setminus\{0\}}, \varphi \Big\rangle.\end{aligned}$$

On voit que la distribution valeur principale de $1/x$ considérée comme distribution sur $\mathbb{R}\setminus\{0\}$ est égale à la fonction $1/x \in L^1_{\text{loc}}(\mathbb{R}\setminus\{0\})$ (c'est assez subtil, il faut bien comprendre ce qu'est un compact de $\mathbb{R}\setminus\{0\}$). On pourra s'amuser à montrer que, bien qu'elle représente d'une certaine façon une fonction, cette distribution valeur principale est d'ordre 1 en tant que distribution sur $\mathbb{R}$, et non pas d'ordre 0 (alors qu'elle est bien sûr d'ordre 0 en tant que distribution sur $\mathbb{R}\setminus\{0\}$).

On peut également voir la valeur principale de $1/x$ comme la limite, au sens des distributions, d'une suite de fonctions classiques. Soit $(f_n)_{n\in\mathbb{N}}$ la suite de fonctions définie par

$$f_n(x) = \frac{1}{x} \quad \text{si} \quad |x| \geq \frac{1}{n} \quad \text{et} \quad f_n(x) = 0 \quad \text{sinon.}$$

Il n'est pas trop difficile de montrer que

$$\lim_{n\to\infty} f_n = u \text{ dans } \mathscr{D}'(\mathbb{R}).$$

Étudions maintenant un exemple sur lequel la continuité de la dérivation mène à des résultats qui peuvent paraître paradoxaux, sauf si l'on ne perd pas de vue qu'il s'agit de distributions. Soit f la fonction définie sur $\mathbb{R}$ par

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{n^2} \sin(nx).$$

Cette série (de Fourier) étant normalement convergente, elle définit une fonction continue sur $\mathbb{R}$. Par contre, la série des dérivées n'est pas normalement convergente, et il n'est pas clair que f soit dérivable au sens classique. Les choses s'aggravent au fur et à mesure que l'on dérive la série terme à terme. Par contre, nous pouvons dériver autant de fois que nous le souhaitons la fonction f à condition de le faire au sens des distributions. En effet, pour toute fonction-test φ, on a

$$\lim_{N\to\infty} \int_{\mathbb{R}} \left(\sum_{n=1}^{N} \frac{1}{n^2} \sin(nx) \right) \varphi(x)\, dx = \int_{\mathbb{R}} f(x)\varphi(x)\, dx,$$

puisque la série converge uniformément et que φ est à support compact. Donc on a

$$\sum_{n=1}^{N} \frac{1}{n^2} \sin(n\cdot) \to f \quad \text{dans} \quad \mathscr{D}'(\mathbb{R}).$$

D'après la continuité de la dérivation dans $\mathscr{D}'(\mathbb{R})$, qui est une banalité, rappelons-le, on en déduit que, pour tout entier ℓ, on a

$$\sum_{n=1}^{N} \frac{1}{n^2} \frac{d^l}{dx^l} \sin(n\,\cdot) \to \frac{d^l f}{dx^l} \quad \text{dans} \quad \mathscr{D}'(\mathbb{R}).$$

Donc, en particulier, la suite $\left(\sum_{n=1}^N \sin(n\cdot)\right)_{N\in\mathbb{N}}$ converge dans $\mathscr{D}'(\mathbb{R})$, mais aussi les suites $\left(\sum_{n=1}^N n\cos(n\cdot)\right)_{N\in\mathbb{N}}$, $\left(\sum_{n=1}^N n^2\sin(n\cdot)\right)_{N\in\mathbb{N}}$, etc. Démontrer cette propriété directement est un exercice très formateur.

Donnons pour finir un exemple de distribution sur $\mathbb{R}$ qui n'est pas d'ordre fini

$$u = \sum_{n=0}^{+\infty} \delta_n^{(n)}.$$

Nous allons maintenant exposer quelques exemples de résolution d'équations dont l'inconnue est une distribution. La première propriété montre que les distributions se prêtent à un calcul des primitives semblables à celui des fonctions.

Proposition 6.4.1 *Soit u une distribution de $\mathscr{D}'(\mathbb{R})$. Si $u' = 0$, alors la distribution u est une fonction constante (et réciproquement).*

Démonstration. Fixons une fonction θ appartenant à $\mathscr{D}(\mathbb{R})$ et d'intégrale 1. Soit φ une fonction indéfiniment différentiable à support compact quelconque sur $\mathbb{R}$. La fonction

$$\varphi - \Big(\int_{-\infty}^{+\infty} \varphi(t)\,dt\Big)\theta$$

est d'intégrale nulle. Donc la fonction

$$\psi(x) \stackrel{\text{déf}}{=} \int_{-\infty}^{x} \Big(\varphi(s) - \Big(\int_{-\infty}^{+\infty} \varphi(t)\,dt\Big)\theta(s)\Big)\,ds$$

est indéfiniment différentiable à support compact et

$$\psi' = \varphi - \Big(\int_{-\infty}^{+\infty} \varphi(x)\,dx\Big)\theta$$

D'où il résulte que

$$\varphi = \psi' + \Big(\int_{-\infty}^{+\infty} \varphi(x)\,dx\Big)\theta.$$

En appliquant la distribution u à l'égalité ci-dessus entre fonctions indéfiniment différentiables à support compact, on trouve que

$$\begin{aligned}
\langle u, \varphi\rangle &= \langle u, \psi'\rangle + \Big(\int_{-\infty}^{+\infty} \varphi(x)\,dx\Big)\langle u, \theta\rangle \\
&= -\langle u', \psi\rangle + \Big(\int_{-\infty}^{+\infty} \varphi(x)\,dx\Big)\langle u, \theta\rangle \\
&= \Big(\int_{-\infty}^{+\infty} \varphi(x)\,dx\Big)\langle u, \theta\rangle \\
&= \int_{-\infty}^{+\infty} \langle u, \theta\rangle \varphi(x)\,dx,
\end{aligned}$$

et la distribution u s'identifie à la fonction constante $x \mapsto \langle u, \theta\rangle$. La réciproque est évidente. $\square$

Remarque. Le résultat reste vrai sur un intervalle ouvert de $\mathbb{R}$ et plus généralement en dimension quelconque : si Ω est un ouvert *connexe* de $\mathbb{R}^d$ et u est une distribution sur Ω telle que $\partial_i u = 0$ pour $i = 1, \dots, d$, alors u est une constante.

Corollaire 6.4.1 *Soit u une distribution sur $\mathbb{R}$. Il existe $U \in \mathscr{D}'(\mathbb{R})$ telle que $U' = u$ et pour toute distribution V telle que $V' = u$, on a $V = U + c$ où c est une constante.*

Démonstration. On définit un opérateur sur $\mathscr{D}(\mathbb{R})$ comme à la proposition précédente par

$$T(\varphi)(x) = \int_{-\infty}^{x} \Big(\varphi(s) - \Big(\int_{-\infty}^{+\infty} \varphi(t)\,dt\Big)\theta(s)\Big)\,ds,$$

où θ est choisie une fois pour toutes. On vérifie aisément que T est séquentiellement continu, c'est-à-dire que si $\varphi_n \to \varphi$ dans $\mathscr{D}(\mathbb{R})$, alors $T(\varphi_n) \to T(\varphi)$ dans $\mathscr{D}(\mathbb{R})$.

On définit alors, pour tout $\varphi \in \mathscr{D}(\mathbb{R})$

$$\langle U, \varphi\rangle = -\langle u, T(\varphi)\rangle.$$

D'après la continuité de T, on voit que U est une distribution. De plus,

$$\langle U', \varphi\rangle = -\langle U, \varphi'\rangle = \langle u, T(\varphi')\rangle = \langle u, \varphi\rangle,$$

donc U est bien une primitive de u au sens de $\mathscr{D}'(\mathbb{R})$. Enfin, si $V' = u$, alors $(V-U)' = 0$ et on conclut grâce à la proposition 6.4.1. $\square$

Remarque. On peut donc prendre indéfiniment des primitives de distributions, et l'on aura que si $U^{(n)} = V^{(n)}$, alors $V = U + P$ où P est un polynôme de degré inférieur à $n-1$, sur $\mathbb{R}$ ou sur un intervalle ouvert de $\mathbb{R}$. L'analogue en dimension supérieure est par contre nettement plus délicat.

Proposition 6.4.2 *Soit u une distribution sur $\mathbb{R}^d$. Alors*

$$\forall\, j \in \{1, \cdots, d\},\ x_j u = 0 \Leftrightarrow \exists\, \lambda \in \mathbb{K}, u = \lambda\delta_0.$$

Démonstration. On utilise la formule de Taylor avec reste intégral. Soit φ une fonction-test. On a

$$\varphi(x) = \varphi(0) + \sum_{j=1}^{d} x_j \int_0^1 \frac{\partial\varphi}{\partial x_j}(tx)\,dt.$$

Soit B_n la boule centrée en 0 et de rayon $n \in \mathbb{N}$ et χ_n une fonction de $\mathscr{D}(\mathbb{R}^d)$ valant identiquement 1 sur B_n. Pour toute fonction-test φ dont le support est inclus dans B_n, on a bien sûr $\chi_n\varphi = \varphi$, d'où il vient que

$$\varphi(x) = \varphi(0)\chi_n(x) + \sum_{j=1}^{d} x_j\varphi_j(x) \quad \text{avec} \quad \varphi_j(x) \stackrel{\text{déf}}{=} \chi_n(x)\int_0^1 \frac{\partial\varphi}{\partial x_j}(tx)\,dt.$$

Par le théorème de dérivation sous le signe somme, et du fait que χ_n est à support compact, on voit que $\varphi_j \in \mathscr{D}(\mathbb{R}^d)$, d'où

$$\langle u, \varphi \rangle = \varphi(0)\langle u, \chi_n \rangle + \sum_{j=1}^{d} \langle u, x_j \varphi_j \rangle.$$

Le fait que $x_j u = 0$ implique que $\langle u, x_j \varphi_j \rangle = 0$. Donc, on a

$$\langle u, \varphi \rangle = \lambda_n \varphi(0),$$

pour toute fonction-test φ à support dans B_n, avec $\lambda_n = \langle u, \chi_n \rangle$. Maintenant, si φ est à support dans B_n, elle est a fortiori à support dans B_{n+1}. On en déduit donc que $\lambda_{n+1} = \lambda_n = \lambda$ est indépendant de n. Comme le support de toute fonction-test est inclus dans une boule B_n, finalement, $u = \lambda \delta_0$.

Réciproquement, il est facile de vérifier que $x_j(\lambda \delta_0) = 0$. □

Voyons maintenant un exemple d'équation algébrique dont la solution au sens des distributions réserve une petite surprise.

Proposition 6.4.3 *Soit u une distribution sur $\mathbb{R}$. Alors*

$$xu = 1 \Leftrightarrow \exists\, \lambda \in \mathbb{K}, u = \mathrm{vp}\,\frac{1}{x} + \lambda \delta_0,$$

où $\mathrm{vp}\,\frac{1}{x}$ *est la distribution définie page 173.*

Démonstration. Vérifions d'abord que $x\,\mathrm{vp}\,\frac{1}{x} = 1$. Soit φ une fonction-test et R un réel tel que le support de φ soit inclus dans l'intervalle ouvert $]-R, R[$. Par définition de la valeur principale de $1/x$ et de la multiplication par une fonction C^∞, on a

$$\left\langle x\,\mathrm{vp}\,\frac{1}{x}, \varphi \right\rangle = \left\langle \mathrm{vp}\,\frac{1}{x}, x\varphi \right\rangle = \int_{-R}^{R} \frac{x\varphi(x) - 0\varphi(0)}{x} dx = \int_{-R}^{R} \varphi(x) dx = \langle 1, \varphi \rangle.$$

On a donc bien $x\,\mathrm{vp}\left(\frac{1}{x}\right) = 1$.

D'autre part, si u_1 et u_2 sont deux distributions telles que $xu_1 = xu_2 = 1$, alors $x(u_1 - u_2) = 0$ et la proposition 6.4.2 conclut la démonstration. □

6.5 Support d'une distribution, distributions à support compact

Les distributions sont a priori définies de façon globale sur Ω. Dire qu'une distribution vaut ceci ou cela en tel ou tel point de Ω n'a absolument aucun sens.

Les distributions conservent néanmoins un caractère de localité qui se traduit notamment dans le concept de support d'une distribution. Celui-ci est fondé sur la notion de restriction d'une distribution à un ouvert. Soient ω et Ω deux ouverts de $\mathbb{R}^d$ tels que le premier soit inclus dans le second. Pour toute fonction $\varphi \in \mathscr{D}(\omega)$, son prolongement par 0 à Ω, noté $\tilde{\varphi}$, est visiblement un élément de $\mathscr{D}(\Omega)$. Cette propriété nous mène naturellement à la définition suivante.

Définition 6.5.1 *Soient $\omega \subset \Omega$ deux ouverts de $\mathbb{R}^d$. Considérons une distribution u sur Ω. On définit alors la restriction de u à ω, notée $u_{|\omega}$, par*

$$\forall\, \varphi \in \mathscr{D}(\omega),\ \langle u_{|\omega}, \varphi\rangle \stackrel{\text{déf}}{=} \langle u, \tilde{\varphi}\rangle.$$

Le fait que la forme linéaire $u_{|\omega}$ est bien une distribution est élémentaire.

Définition 6.5.2 *Soit u une distribution sur Ω, on appelle* support *de u et l'on note* $\operatorname{Supp} u$ *le complémentaire du plus grand ouvert ω inclus dans Ω tel que $u_{|\omega} = 0$.*

Remarques. Le support d'une distribution est par définition un fermé de Ω pour la topologie induite. De plus, cette notion de support d'une distribution prolonge celle du support d'une fonction. En effet, soit f une fonction de $L^1_{\text{loc}}(\Omega)$ et ω un ouvert tel que $f_{|\omega} = 0$ au sens des distributions. D'après la définition précédente, ceci signifie que

$$\forall\, \varphi \in \mathscr{D}(\omega),\ \langle f, \varphi\rangle = \int_\omega f(x)\varphi(x)\, dx = 0.$$

D'après le théorème 6.2.1, ceci signifie exactement que, pour presque tout x de ω, on a que $f(x) = 0$. Ceci montre bien que la définition de support d'une distribution coïncide bien avec la définition usuelle dans le cas des fonctions localement intégrables.

Définition 6.5.3 *Une distribution sur Ω dont le support est un compact de Ω est assez logiquement appelée* distribution à support compact. *L'espace vectoriel des distributions à support compact est désigné par $\mathscr{E}'(\Omega)$.*

Nous allons maintenant voir que l'espace des distributions à support compact peut s'identifier au dual de l'espace des fonctions indéfiniment différentiables sur Ω. Pour cela, commençons par démontrer le lemme suivant.

Lemme 6.5.1 *Soient u une distribution à support compact sur Ω et χ une fonction de $\mathscr{D}(\Omega)$ valant identiquement 1 sur le support de u. Alors, on a*

$$\forall\, \varphi \in \mathscr{D}(\Omega),\ \langle u, \chi\varphi\rangle = \langle u, \varphi\rangle.$$

Démonstration. Notons tout d'abord qu'une telle fonction χ existe car le support de u est compact. Il suffit d'écrire que, pour toute fonction-test φ, on a

$$\langle u, \chi\varphi\rangle = \langle u, \varphi\rangle + \langle u, \chi\varphi - \varphi\rangle.$$

Par définition de χ, le support de $(\chi - 1)\varphi$ est un compact inclus dans le complémentaire du support de u, donc par définition du support de u, on a

$$\langle u, \chi\varphi - \varphi\rangle = 0,$$

d'où le lemme. □

Proposition 6.5.1 *Soit u une distribution à support compact. Il existe une unique forme linéaire $\widetilde{u}$ définie sur $C^\infty(\Omega)$ (l'ensemble des fonctions indéfiniment différentiables sur Ω) prolongeant u et telle qu'il existe un compact $K' \subset \Omega$, une constante C et un entier N tels que*

$$\forall\, \phi \in C^\infty(\Omega),\ |\langle \widetilde{u}, \phi\rangle| \leq C \sup_{\substack{x\in K' \\ |\alpha|\leq N}} |\partial^\alpha \phi(x)|.$$

Démonstration. Soit χ une fonction-test de $\mathscr{D}(\Omega)$ valant 1 sur le support de u, on définit $\widetilde{u}$ sur $C^\infty(\Omega)^*$ par

$$\langle \widetilde{u}, \phi\rangle \stackrel{\text{déf}}{=} \langle u, \chi\phi\rangle.$$

Par un argument analogue à celui qui vient d'être exposé, on vérifie aisément que cette définition ne dépend pas du choix de χ. De plus, le lemme 6.5.1 ci-dessus montre que que $\widetilde{u}$ prolonge bien u. Comme les fonctions $\chi\phi$ sont toutes à support dans le même compact $K' = \operatorname{Supp}\chi$ et u étant une distribution, il existe une constante C et un entier N tels que

$$\forall\, \phi \in C^\infty(\Omega),\ |\langle \widetilde{u}, \phi\rangle| \leq C \sup_{|\alpha|\leq N} \|\partial^\alpha(\chi\phi)\|_{L^\infty}.$$

La formule de Leibniz assure que

$$\forall\, \phi \in C^\infty(\Omega),\ |\langle \widetilde{u}, \phi\rangle| \leq C \sup_{\substack{x\in K' \\ |\alpha|\leq N}} |\partial^\alpha \phi(x)|. \tag{6.3}$$

Ceci montre l'existence du prolongement $\tilde{u}$. On admettra qu'il y a également unicité (voir proposition suivante). □

Corollaire 6.5.1 *Toute distribution à support compact est d'ordre fini.*

Démonstration. Ceci n'est qu'une traduction de l'inégalité (6.3). □

Proposition 6.5.2 *Soit $\widetilde{u}$ une forme linéaire sur $C^\infty(\Omega)$ continue au sens de* (6.3). *Alor $\widetilde{u}$ définit par restriction une distribution à support compact.*

Démonstration. Il est tout à fait clair que la restriction de $\widetilde{u}$ à $\mathscr{D}(\Omega)$ définit un élément u de $\mathscr{D}'(\Omega)$. De plus, si le support de φ est contenu dans le complémentaire de K', alors φ est nulle sur K' et l'estimation (6.3) implique trivialement que $\langle u, \varphi\rangle = 0$. Par conséquent, u est à support dans K'. □

Remarque L'espace $C^\infty(\Omega)$ étant parfois noté $\mathscr{E}(\Omega)$, la notation $\mathscr{E}'(\Omega)$ utilisée pour les distributions à support compact insiste sur le fait que l'on identifie les formes linéaires sur $C^\infty(\Omega)$ « continues » – c'est-à-dire vérifiant (6.3) – et les distributions sur Ω à support compact. Naturellement, tout ceci peut être mis dans un cadre topologique rigoureux.

Mentionnons dans le même ordre d'idées la propriété de *faisceau* des distributions : on peut « recoller » deux (ou plus) distributions définies sur deux (ou plus) ouverts qui coïncident au sens de la restriction sur leur intersection, en une distribution sur la réunion des ouverts. Plus précisément,

Proposition 6.5.3 *Soient Ω_1 et Ω_2 deux ouverts, $u_1 \in \mathscr{D}'(\Omega_1)$ et $u_2 \in \mathscr{D}'(\Omega_2)$ deux distributions telles que $u_{1|\Omega_1\cap\Omega_2} = u_{2|\Omega_1\cap\Omega_2}$. Alors il existe une unique distribution $u \in \mathscr{D}'(\Omega_1 \cup \Omega_2)$ telle que $u_{|\Omega_1} = u_1$ et $u_{|\Omega_2} = u_2$.*

Démonstration. Soit (θ_1, θ_2) une partition de l'unité C^∞ associée aux ouverts Ω_1 et Ω_2, c'est-à-dire que $\operatorname{Supp}\theta_i \subset \Omega_i$, $0 \le \theta_i \le 1$ et $\theta_1 + \theta_2 = 1$. Notons $\Omega = \Omega_1 \cup \Omega_2$.

Pour tout $\varphi \in \mathscr{D}(\Omega)$, on définit

$$\langle u, \varphi\rangle = \langle u_1, \varphi\theta_1\rangle + \langle u_2, \varphi\theta_2\rangle.$$

On vérifie facilement que $u \in \mathscr{D}'(\Omega)$. Calculons sa restriction aux ouverts Ω_i. Si φ appartient à $\mathscr{D}(\Omega_1)$, alors $\varphi\theta_2$ appartient à $\mathscr{D}(\Omega_1 \cap \Omega_2)$. On voit donc dans ce cas que

$$\langle u_2, \varphi\theta_2\rangle = \langle u_{2|\Omega_1\cap\Omega_2}, \varphi\theta_2\rangle = \langle u_{1|\Omega_1\cap\Omega_2}, \varphi\theta_2\rangle = \langle u_1, \varphi\theta_2\rangle,$$

d'où

$$\langle u, \varphi\rangle = \langle u_1, \varphi\theta_1\rangle + \langle u_1, \varphi\theta_2\rangle = \langle u_1, \varphi(\theta_1 + \theta_2)\rangle = \langle u_1, \varphi\rangle,$$

et donc $u_{|\Omega_1} = u_1$, et de même sur Ω_2.

Il faut maintenant s'assurer de l'unicité de u. Soit $v \in \mathscr{D}'(\Omega)$ une distribution qui réponde à la question. Pour tout $\varphi \in \mathscr{D}(\Omega)$, on a $\varphi = \varphi(\theta_1 + \theta_2) = \varphi\theta_1 + \varphi\theta_2$ avec $\varphi\theta_i \in \mathscr{D}(\Omega_i)$. On a donc

$$\begin{aligned}\langle v, \varphi\rangle = \langle v, \varphi\theta_1\rangle + \langle v, \varphi\theta_2\rangle &= \langle v_{|\Omega_1}, \varphi\theta_1\rangle + \langle v_{|\Omega_2}, \varphi\theta_2\rangle \\ &= \langle u_1, \varphi\theta_1\rangle + \langle u_2, \varphi\theta_2\rangle = \langle u, \varphi\rangle,\end{aligned}$$

donc $v = u$. □

Remarquons que si $\Omega_1 \cap \Omega_2 = \emptyset$, le résultat de recollement est vrai sans aucune condition de compatibilité entre u_1 et u_2 ! Comme indiqué plus haut, on généralise sans difficulté à un nombre quelconque d'ouverts.

Chapitre 7

Transformation de Fourier

La transformation de Fourier est un des outils de base de l'analyse.

7.1 Définition sur l'espace L^1.

Définition 7.1.1 *Soit f une fonction de $L^1(\mathbb{R}^d;\mathbb{C})$. On appelle* transformée de Fourier *de f et l'on note $\widehat{f}$ ou bien $\mathscr{F}f$ la fonction définie par*

$$\forall \xi \in \mathbb{R}^d, \quad \widehat{f}(\xi) = \mathscr{F}f(\xi) = \int_{\mathbb{R}^d} e^{-i(x|\xi)} f(x)\, dx \in \mathbb{C}. \tag{7.1}$$

On utilisera indifféremment les deux notations pour la transformée de Fourier d'une fonction.

Il est clair que la transformée de Fourier d'une fonction de L^1 est une fonction bornée, mais nous avons le théorème plus précis suivant.

Théorème 7.1.1 *L'application $\mathscr{F}$ est linéaire continue de $L^1(\mathbb{R}^d;\mathbb{C})$ à valeurs dans l'espace $C_0(\mathbb{R}^d;\mathbb{C})$ des fonctions continues nulles à l'infini muni de la norme $L^\infty(\mathbb{R}^d;\mathbb{C})$.*

Démonstration. Le corollaire 5.4.2 affirme en particulier que les fonctions indéfiniment différentiables à support compact sur $\mathbb{R}^d$ sont denses dans l'espace L^1. Comme il est évident que

$$\|\widehat{f}\|_{L^\infty} \leq \|f\|_{L^1}, \tag{7.2}$$

on en déduit que la transformation de Fourier est à valeurs dans les fonctions continues bornées. En effet, si $\varphi \in \mathscr{D}(\mathbb{R}^d)$, alors $\mathscr{F}\varphi$ est une fonction continue et si $\varphi_n \to f$ dans L^1, alors $\mathscr{F}\varphi_n \to \mathscr{F}f$ uniformément par l'estimation (7.2).

Il suffit dès lors de démontrer que

$$\forall \varphi \in \mathscr{D}(\mathbb{R}^d), \lim_{|\xi|\to\infty} \widehat{\varphi}(\xi) = 0 \tag{7.3}$$

pour assurer le théorème par le même argument de convergence uniforme.

Pour démontrer (7.3), nous allons utiliser une idée que nous reverrons fréquemment : la régularité d'une fonction et la décroissance à l'infini de sa transformée de Fourier sont indissolublement liées. En effet, remarquons que

$$(1+|\xi|^2)e^{-i(x|\xi)} = (\mathrm{Id}-\Delta)e^{-i(x|\xi)} \quad \text{avec} \quad \Delta = \sum_{j=1}^{d} \frac{\partial^2}{\partial x_j^2}.$$

Ainsi donc, il vient

$$(1+|\xi|^2)\widehat{\varphi}(\xi) = \int_{\mathbb{R}^d} \left((\mathrm{Id}-\Delta)e^{-i(\cdot|\xi)}\right)(x)\varphi(x)\,dx.$$

La fonction φ est indéfiniment différentiable à support compact ; on peut donc intégrer deux fois par parties dans l'intégrale ci-dessus. On trouve alors que

$$\begin{aligned}(1+|\xi|^2)\widehat{\varphi}(\xi) &= \int_{\mathbb{R}^d} e^{-i(x|\xi)})(\mathrm{Id}-\Delta)\varphi(x)\,dx \\ &= \mathscr{F}\left((\mathrm{Id}-\Delta)\varphi\right).\end{aligned}$$

Mais la fonction φ étant indéfiniment différentiable à support compact, $(\mathrm{Id}-\Delta)\varphi$ également et est donc intégrale. Ainsi donc, on a

$$|(1+|\xi|^2)\widehat{\varphi}(\xi)| \leq \|(\mathrm{Id}-\Delta)\varphi\|_{L^1}.$$

Divisant par $1+|\xi|^2$, ceci implique l'assertion (7.3). □

Exercice 7.1.1 *Démontrez par une méthode analogue que, si φ est une fonction indéfiniment différentiable à support compact sur $\mathbb{R}^d$, alors, pour tout entier N, il existe une constante C telle que*

$$|\widehat{\varphi}(\xi)| \leq C(1+|\xi|^2)^{-N}.$$

La proposition suivante décrit partiellement les rapports entre transformation de Fourier et convolution.

Proposition 7.1.1 *Soient f et g deux fonctions de L^1. On a*

$$\mathscr{F}(f\star g) = \widehat{f}\widehat{g}.$$

Démonstration. La fonction F_ξ définie de $\mathbb{R}^d \times \mathbb{R}^d$ dans $\mathbb{C}$ par

$$F_\xi(x,y) \stackrel{\text{déf}}{=} e^{-i(x|\xi)}f(x-y)g(y) = e^{-i(x-y|\xi)}f(x-y)e^{-i(y|\xi)}g(y)$$

est, pour tout ξ appartenant à $\mathbb{R}^d$, une fonction de $L^1(\mathbb{R}^d \times \mathbb{R}^d)$. Le théorème de Fubini assure que

$$\begin{aligned}
\mathscr{F}(f\star g)(\xi) &= \int_{\mathbb{R}^d} e^{-i(x|\xi)} \left(\int_{\mathbb{R}^d} f(x-y)g(y)\,dy\right) dx\\
&= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} e^{-i(x-y|\xi)} f(x-y)\,dx\right) e^{-i(y|\xi)} g(y)\,dy\\
&= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} e^{-i(z|\xi)} f(z)\,dz\right) e^{-i(y|\xi)} g(y)\,dy\\
&= \int_{\mathbb{R}^d} e^{-i(z|\xi)} f(z)\,dz \int_{\mathbb{R}^d} e^{-i(y|\xi)} g(y)\,dy\\
&= \widehat{f}(\xi)\widehat{g}(\xi),
\end{aligned}$$

par le changement de variable $z = x - y$. □

Remarquons l'harmonie de ce résultat : L^1 est une algèbre pour la convolution et C_0 est une algèbre pour le produit ordinaire.

Nous allons maintenant examiner l'influence des dilatations sur la transformation de Fourier.

Proposition 7.1.2 *Soient f une fonction de $L^1(\mathbb{R}^d)$ et λ un réel strictement positif. On a alors*

$$\mathscr{F}(f(\lambda\cdot))(\xi) = \lambda^{-d}\widehat{f}\Big(\frac{\xi}{\lambda}\Big).$$

Démonstration. Il suffit de faire le changement de variables $y = \lambda x$ dans l'intégrale définissant la transformée de Fourier, ce qui donne

$$\mathscr{F}(f(\lambda\cdot))(\xi) = \int_{\mathbb{R}^d} e^{-i(x|\xi)} f(\lambda x)\,dx = \lambda^{-d}\int_{\mathbb{R}^d} e^{-i(\frac{y}{\lambda}|\xi)} f(y)\,dy = \lambda^{-d}\widehat{f}\Big(\frac{\xi}{\lambda}\Big),$$

d'où la proposition. □

Une transformée de Fourier importante est celle des gaussiennes, $G_a(x) = e^{-\frac{a}{2}|x|^2}$ pour $a > 0$.

Proposition 7.1.3 *Soit a un réel strictement positif. On a alors*

$$\mathscr{F}G_a(\xi) = \left(\frac{2\pi}{a}\right)^{\frac{d}{2}} e^{-\frac{|\xi|^2}{2a}} = \left(\frac{2\pi}{a}\right)^{\frac{d}{2}} G_{\frac{1}{a}}(\xi). \tag{7.4}$$

Démonstration. Commençons par la dimension $d = 1$ et posons

$$g(\xi) \stackrel{\text{déf}}{=} \int_{\mathbb{R}} e^{-ix\xi} e^{-x^2}\,dx.$$

La fonction g est bien évidemment dérivable et l'on peut écrire en justifiant aisément la dérivation sous le signe somme et l'intégration par parties

$$\begin{aligned} g'(\xi) &= \int_{\mathbb{R}} -ixe^{-ix\xi}e^{-x^2}dx \\ &= \int_{\mathbb{R}} ie^{-ix\xi}\frac{1}{2}\frac{d}{dx}(e^{-x^2})\,dx \\ &= \frac{1}{2}\int_{\mathbb{R}} -i\frac{d}{dx}\left(e^{-ix\xi}\right)e^{-x^2}dx \\ &= \frac{1}{2}\int_{\mathbb{R}} i^2\xi e^{-ix\xi}e^{-x^2}dx \\ &= -\frac{\xi}{2}g(\xi). \end{aligned}$$

Cette équation différentielle est facile à résoudre. On a

$$g(\xi) = g(0)e^{-\frac{\xi^2}{4}}.$$

Il est classique que

$$g(0) = \int_{\mathbb{R}} e^{-x^2}dx = \sqrt{\pi}.$$

La proposition 7.1.2 assure alors le résultat en dimension 1. En dimension d quelconque, il suffit d'observer que

$$e^{-ix\xi}e^{-\frac{a}{2}|x|^2} = \prod_{j=1}^{d} e^{-ix_j\xi_j}e^{-\frac{a}{2}x_j^2}$$

et que, si les f_j sont des fonctions de $L^1(\mathbb{R})$, posant $(f_1\otimes\cdots\otimes f_d)(x) = \prod_{j=1}^{d} f_j(x_j)$, on a

$$\mathscr{F}(f_1\otimes\cdots\otimes f_d) = \widehat{f_1}\otimes\cdots\otimes\widehat{f_d},$$

par le théorème de Fubini. On applique alors d fois la formule (7.4) établie dans le cas unidimensionnel. □

En itérant le résultat, on a immédiatement le corollaire suivant.

Corollaire 7.1.1 *On a*

$$\mathscr{F}(\mathscr{F}G_a) = (2\pi)^d G_a.$$

Le corollaire se réécrit sous la forme suivante, cas particulier important du l'inversion de Fourier étudiée au prochain paragraphe.

Corollaire 7.1.2 *On a*

$$G_a(x) = (2\pi)^{-d}\int_{\mathbb{R}^d} e^{i(x|\xi)}\widehat{G_a}(\xi)\,d\xi,$$

Démonstration. Il suffit d'utiliser la parité de la gaussienne. □

7.2 Formule d'inversion de Fourier et extension à L^2

Le théorème d'inversion de Fourier qui suit est absolument fondamental. Il permet, sous certaines conditions, de reconstituer une fonction à partir de sa transformée de Fourier.

Théorème 7.2.1 *Soit f une fonction de $L^1(\mathbb{R}^d)$ telle que $\widehat{f}$ soit aussi dans L^1. On a alors*

$$f(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{i(x|\xi)} \widehat{f}(\xi)\, d\xi. \tag{7.5}$$

Démonstration. Posons

$$G(x) = (2\pi)^{-\frac{d}{2}} e^{-\frac{|x|^2}{2}} \quad \text{et} \quad G_\varepsilon(x) = \frac{1}{\varepsilon^d} G\left(\frac{x}{\varepsilon}\right) \cdot$$

La fonction G est une fonction positive d'intégrale 1. D'après le théorème 5.4.5, on a pour tout f dans $L^1(\mathbb{R}^d)$

$$\lim_{\varepsilon \to 0} \|G_\varepsilon \star f - f\|_{L^1} = 0. \tag{7.6}$$

De plus, la proposition 7.1.3 nous dit que $\widehat{G}_\varepsilon(\xi) = e^{-\frac{\varepsilon^2 |\xi|^2}{2}}$, d'où

$$\forall \xi \in \mathbb{R}^d, \quad |\widehat{G}_\varepsilon(\xi)| \leq 1 \quad \text{et} \quad \lim_{\varepsilon \to 0} \widehat{G}_\varepsilon(\xi) = 1.$$

On voit donc que

$$\forall \xi \in \mathbb{R}^d, \quad |\widehat{G}_\varepsilon(\xi) \hat{f}(\xi)| \leq |\hat{f}(\xi)| \quad \text{et que} \quad \lim_{\varepsilon \to 0} \widehat{G}_\varepsilon(\xi) \hat{f}(\xi) = \hat{f}(\xi),$$

Comme par hypothèse $\hat{f}$ est dans $L^1(\mathbb{R}^d)$, on peut donc appliquer le théorème de convergence dominée de Lebesgue pour obtenir

$$(2\pi)^{-d} \int_{\mathbb{R}^d} e^{i(x|\xi)} \widehat{f}(\xi)\, d\xi = (2\pi)^{-d} \lim_{\varepsilon \to 0} \int_{\mathbb{R}^d} e^{i(x|\xi)} \widehat{G}_\varepsilon(\xi) \widehat{f}(\xi)\, d\xi. \tag{7.7}$$

D'un autre côté, la fonction f appartient à $L^1(\mathbb{R}^d)$, par conséquent $\widehat{G}_\varepsilon \otimes f$ est une fonction de $L^1(\mathbb{R}^d \times \mathbb{R}^d)$ et d'après le théorème de Fubini, on a

$$\begin{aligned}
(2\pi)^{-d} \int_{\mathbb{R}^d} e^{i(x|\xi)} \widehat{G}_\varepsilon(\xi) \widehat{f}(\xi)\, d\xi &= (2\pi)^{-d} \int_{\mathbb{R}^d \times \mathbb{R}^d} e^{i(x|\xi)} \widehat{G}_\varepsilon(\xi) e^{-i(y|\xi)} f(y)\, d\xi dy \\
&= \int_{\mathbb{R}^d} \left((2\pi)^{-d} \int_{\mathbb{R}^d} e^{i(x-y|\xi)} \widehat{G}_\varepsilon(\xi)\, d\xi \right) f(y)\, dy \\
&= \int_{\mathbb{R}^d} G_\varepsilon(x-y) f(y)\, dy = (G_\varepsilon \star f)(x), \qquad (7.8)
\end{aligned}$$

où l'on a utilisé le corollaire 7.1.2. Le théorème découle alors des relations (7.6), (7.7) et (7.8). □

Remarques. La transformation décrite par (7.5) s'appelle *transformation de Fourier inverse*. On peut écrire ce résultat sous la forme

$$\mathscr{F}^2 f = (2\pi)^d \check{f} \text{ ou } \mathscr{F}(\check{\mathscr{F}} f) = (2\pi)^d f \quad \text{avec} \quad \check{f}(x) \stackrel{\text{déf}}{=} f(-x) \tag{7.9}$$

pour une fonction f telle que f et $\widehat{f}$ soient dans L^1. Ce théorème signifie qu'une fonction f dans L^1 telle que $\widehat{f}$ soit aussi dans L^1 (par exemple une fonction indéfiniment différentiable à support compact comme le montre l'exercice 7.1.1) s'écrit comme une superposition d'oscillations — les fonctions $e^{i(x|\xi)}$ — la transformée de Fourier de f apparaissant alors comme la densité d'oscillation à la fréquence ξ. Remarquons pour finir qu'une fonction de L^1 dont la transformée de Fourier est dans L^1 est nécessairement continue et nulle à l'infini.

Le théorème d'inversion de Fourier a une conséquence très importante, le théorème de Fourier-Plancherel.

Théorème 7.2.2 *L'application $(2\pi)^{-\frac{d}{2}}\mathscr{F}$ se prolonge de $L^1 \cap L^2$ à L^2 en une isométrie bijective de L^2 dans L^2.*

La démonstration repose sur le lemme suivant, très simple mais crucial.

Lemme 7.2.1 *Soient f et g deux fonctions de L^1 ; on a*

$$\int_{\mathbb{R}^d} \widehat{f}(x) g(x)\, dx = \int_{\mathbb{R}^d} f(x) \widehat{g}(x)\, dx.$$

Démonstration. Observons que la fonction de $\mathbb{R}^d \times \mathbb{R}^d$ dans $\mathbb{C}$ définie par

$$(x,y) \mapsto e^{-i(x|y)} f(x) g(y)$$

est intégrable sur $\mathbb{R}^d \times \mathbb{R}^d$. Le théorème de Fubini assure alors que

$$\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} e^{-i(x|y)} f(x) dx \right) g(y)\, dy = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} e^{-i(x|y)} g(y) dy \right) f(x)\, dx,$$

ce qui est exactement la conclusion du lemme. □

Démonstration du théorème de Fourier-Plancherel. Soit φ une fonction de $\mathscr{D}(\mathbb{R}^d)$. Nous allons appliquer le lemme 7.2.1 au couple $(\varphi, \overline{\mathscr{F}(\varphi)})$. Pour cela, notons d'abord que

$$\begin{aligned} \overline{\mathscr{F}\varphi(\xi)} = \overline{\int_{\mathbb{R}^d} e^{-i(x|\xi)} \varphi(x)\, dx} &= \int_{\mathbb{R}^d} e^{i(x|\xi)} \overline{\varphi}(x)\, dx \\ &= \int_{\mathbb{R}^d} e^{-i(x|\xi)} \check{\overline{\varphi}}(x)\, dx = \mathscr{F}\left(\check{\overline{\varphi}}\right)(\xi), \end{aligned}$$

en changeant x en $-x$. On trouve donc que

$$\begin{aligned}\|\widehat{\varphi}\|_{L^2}^2 = \int_{\mathbb{R}^d} \mathscr{F}\varphi(\xi)\overline{\mathscr{F}\varphi(\xi)}\,d\xi &= \int_{\mathbb{R}^d} \mathscr{F}\varphi(\xi)\mathscr{F}\left(\check{\overline{\varphi}}\right)(\xi)\,d\xi \\ &= \int_{\mathbb{R}^d} \varphi(x)\mathscr{F}\left(\mathscr{F}\left(\check{\overline{\varphi}}\right)\right)(x)\,dx = (2\pi)^d \int_{\mathbb{R}^d} \varphi(x)\check{\check{\overline{\varphi}}}(x)\,dx.\end{aligned}$$

Comme clairement $\check{\check{\overline{\varphi}}}(x) = \overline{\varphi}(x)$, on obtenu que

$$\|(2\pi)^{-d/2}\widehat{\varphi}\|_{L^2}^2 = \|\varphi\|_{L^2}^2$$

pour toute fonction indéfiniment différentiable à support compact. L'ensemble de ces fonctions étant dense dans L^2, la transformation de Fourier (normalisée par le facteur $(2\pi)^{-d/2}$) se prolonge donc en une isométrie de L^2 sur L^2. Comme c'est une isométrie, elle est injective. Il reste à démontrer qu'elle est surjective. Son image est bien sûr fermée puisque c'est une isométrie. De plus, d'après la formule d'inversion de Fourier et l'exercice 7.1.1, l'image de L^2 par $\mathscr{F}$ contient l'ensemble des fonctions indéfiniment différentiables à support compact qui est dense dans L^2, d'où le théorème. □

Remarque. Attention, pour une fonction de L^2, la formule intégrale de définition de la transformée de Fourier (7.1) n'est *a priori* pas valable telle quelle (sauf si la fonction appartient aussi à L^1, bien sûr). Pour calculer la transformée de Fourier d'une fonction de L^2, on n'a à ce stade d'autre choix que de passer par une approximation et un passage à la limite.

7.3 Espace de Schwartz et transformation de Fourier

Pour tout multi-entier α et $x \in \mathbb{R}^d$, on pose $x^\alpha \stackrel{\text{déf}}{=} \prod_j x_j^{\alpha_j}$ et pour toute fonction f qui est $|\alpha|$ fois différentiable, on pose $D^\alpha f = \left(\frac{\partial}{i}\right)^\alpha f$.

Définition 7.3.1 *L'espace $\mathscr{S}(\mathbb{R}^d)$, noté simplement $\mathscr{S}$ en l'absence d'ambiguité, est l'espace des fonctions u indéfiniment différentiables sur $\mathbb{R}^d$ telles que, pour tout entier k, on ait*

$$\|u\|_{k,\mathscr{S}} \stackrel{\text{déf}}{=} \sup_{\substack{x\in\mathbb{R}^d \\ |\alpha|,|\beta|\le k}} |x^\beta \partial^\alpha u(x)| < +\infty.$$

Remarques. On appelle $\mathscr{S}$ l'*espace de Schwartz* ou bien l'espace des *fonctions à décroissance rapide*, au sens où elles tendent vers 0 à l'infini, ainsi que toutes leurs dérivées partielles, plus rapidement que l'inverse de n'importe quel polynôme. Il est clair que $\mathscr{S}$ est inclus dans $L^1 \cap L^\infty$ ainsi que dans tous les L^p, que l'application linéaire ∂^α envoie $\mathscr{S}$ dans $\mathscr{S}$ avec

$$\|\partial^\alpha u\|_{k,\mathscr{S}} \leq \|u\|_{k+|\alpha|,\mathscr{S}}.$$

On peut voir l'espace $\mathscr{S}$ comme l'intersection pour tous les entiers k des espaces $\mathscr{S}_k(\mathbb{R}^d)$ formés des fonctions u, k fois continûment différentiables et telles que $\|u\|_{k,\mathscr{S}}$ soit finie. Chacun de ces espaces, une fois muni de la norme correspondante, est un espace de Banach. La structure d'espace vectoriel topologique comme intersection décroissance d'une famille dénombrable d'espaces de Banach est un cas particulier de ce que l'on appelle structure d'espace de Fréchet (laquelle ne sera pas abordée en tant que telle dans ce cours). Elle permet de munir l'espace vectoriel obtenu par intersection d'une structure d'espace métrique complet. La distance dans $\mathscr{S}$ est

$$d(u,v) = \sum_{k\in\mathbb{N}} 2^{-k}\min(1, \|u-v\|_{k,\mathscr{S}}).$$

Il faut noter que la topologie de $(\mathscr{S}, d)$ n'est pas une topologie d'espace vectoriel normé.

Donnons maintenant quelques exemples d'éléments de $\mathscr{S}$. Les fonctions indéfiniment différentiables à support compact sont bien sûr des éléments de $\mathscr{S}$. De plus, les gaussiennes, c'est-à-dire les fonctions du type $e^{-a|x|^2}$ avec a strictement positif, sont aussi des éléments de $\mathscr{S}$.

Pour comprendre les calculs d'estimations que l'on effectue couramment sur l'espace $\mathscr{S}$, on peut s'appuyer sur la proposition suivante. On note provisoirement $|x| = (|x_1|, |x_2|, \ldots, |x_d|)$.

Proposition 7.3.1 *Soit q une fonction de $\mathbb{R}^d$ dans $\mathbb{R}$ telle qu'il existe un polynôme P sur $\mathbb{R}^d$ à coefficients positifs avec $|q(x)| \leq P(|x|)$. Alors il existe une constante C et un entier k tels que, pour tout $u \in \mathscr{S}$,*

$$\forall x \in \mathbb{R}^d, \quad |q(x)u(x)| \leq C\|u\|_{k,\mathscr{S}}.$$

Démonstration. Le polynôme P s'écrit comme une somme de monômes $P(|x|) = \sum_{j=1}^m \lambda_j |x^{\beta_j}|$ où β_j est un multi-entier. On note $k = \max_{1\leq j\leq m} |\beta_j|$ le degré total de P. Il vient donc

$$|q(x)u(x)| \leq P(|x|)|u(x)| = \sum_{j=1}^m \lambda_j |x^{\beta_j} u(x)| \leq m(\max_{1\leq j\leq m} \lambda_j)(\max_{1\leq j\leq m} |x^{\beta_j} u(x)|),$$

d'où le résultat avec $C = m\max_{1\leq j\leq m} \lambda_j$. □

Remarque. Le même résultat s'applique clairement avec $\partial^\alpha u$ au lieu de u. En particulier, on en déduit que si $u \in \mathscr{S}$, alors pour tout N entier, il existe une constante $C(N,\alpha)$ telle que

$$\forall x \in \mathbb{R}^d, \quad |\partial^\alpha u(x)| \leq \frac{C(N,\alpha)}{(1+|x|^2)^N},$$

d'où le terme de fonctions à décroissance rapide utilisé pour qualifier les éléments de $\mathscr{S}$.

Proposition 7.3.2 *Pour toute fonction f de $\mathscr{S}$, on a*

$$\mathscr{F}(D^\alpha f)(\xi) = \xi^\alpha \widehat{f}(\xi) \quad et \quad \mathscr{F}(x^\alpha f)(\xi) = (-D_\xi)^\alpha \widehat{f}(\xi).$$

Démonstration. On procède par intégrations par parties successives, lesquelles sont licites en raison de la décroissance rapide de toutes les fonctions considérées (on pourra les justifier en détail).

$$\begin{aligned}\mathscr{F}(D^\alpha f)(\xi) &= \int_{\mathbb{R}^d} e^{-i(x|\xi)} \left(\frac{\partial}{i}\right)^\alpha f(x)\,dx \\ &= \int_{\mathbb{R}^d} \left(-\frac{\partial}{i}\right)^\alpha e^{-i(x|\xi)} f(x)\,dx = \xi^\alpha \int_{\mathbb{R}^d} e^{-i(x|\xi)} f(x)\,dx = \xi^\alpha \widehat{f}(\xi).\end{aligned}$$

De même, on a

$$\begin{aligned}\mathscr{F}(x^\alpha f)(x) = \int_{\mathbb{R}^d} e^{-i(x|\xi)} x^\alpha f(x)\,dx &= \int_{\mathbb{R}^d} \left(-\frac{\partial_\xi}{i}\right)^\alpha e^{-i(x|\xi)} f(x)\,dx \\ &= (-D_\xi)^\alpha \int_{\mathbb{R}^d} e^{-i(x|\xi)} f(x)\,dx = (-D_\xi)^\alpha \widehat{f}(\xi),\end{aligned}$$

en dérivant sous le signe somme. □

L'espace $\mathscr{S}$ est remarquablement adapté à la transformation de Fourier : c'est l'un des rares espaces explicitement définis qui soient stables par transformation de Fourier. Nous avons déjà vu que l'espace L^2 possédait cette propriété de stabilité.

Théorème 7.3.1 *La transformation de Fourier est un isomorphisme de $\mathscr{S}$ dans $\mathscr{S}$.*

Démonstration. Pour montrer que $\mathscr{F}(\mathscr{S}) \subset \mathscr{S}$, on doit majorer les quantités $\xi^\beta \partial^\alpha \widehat{u}(\xi)$, ou ce qui revient au même $\xi^\beta (-D)^\alpha \widehat{u}(\xi)$. D'après la proposition 7.3.2, on a

$$\xi^\beta (-D)^\alpha \widehat{u}(\xi) = \int_{\mathbb{R}^d} e^{-i(x|\xi)} D^\beta (x^\alpha u)(x)\,dx$$

D'après la formule de Leibniz, il vient

$$\xi^\beta(-D)^\alpha\widehat{u}(\xi) = \sum_{|\gamma|,|\delta|\leq k} A_{\gamma,\delta}\int_{\mathbb{R}^d} e^{-i(x|\xi)}x^\gamma\partial^\delta u(x)\,dx,$$

où $A_{\gamma,\delta}$ sont des constantes qu'il n'est pas utile d'expliciter. Or, on a

$$\left|\int_{\mathbb{R}^d} e^{-i(x|\xi)}x^\gamma\partial^\delta u(x)dx\right| \leq \int_{\mathbb{R}^d}(1+|x|^2)^{-\frac{d+1}{2}}(1+|x|^2)^{\frac{d+1}{2}}|x^\gamma\partial^\delta u(x)|\,dx.$$

En majorant le terme $(1+|x|^2)^{\frac{d+1}{2}}|x^\gamma\partial^\delta u(x)|$, il vient

$$\begin{aligned}|\xi^\beta\partial^\alpha\widehat{u}(\xi)| &\leq C_k\|u\|_{k+d+1,\mathscr{S}}\int_{\mathbb{R}^d}(1+|x|^2)^{-\frac{d+1}{2}}\,dx\\ &\leq C_k\|u\|_{k+d+1,\mathscr{S}}\end{aligned}$$

où C_k est une constante qui dépend de k, et ce pour tout couple (α,β) de multi-entiers de longueur inférieure ou égale à k. Ceci montre que

$$\|\hat{u}\|_{k,\mathscr{S}} \leq C_k\|u\|_{k+d+1,\mathscr{S}} < +\infty,$$

d'où la stabilité de $\mathscr{S}$ par transformation de Fourier.

Pour la continuité, on commence par noter que d'après la structure de la distance sur $\mathscr{S}$, il suffit de montrer la continuité en 0. De plus, il est assez facile de voir que (exercice)

$$d(u_n,0)\to 0 \Longleftrightarrow \forall k\in\mathbb{N}, \|u_n\|_{k,\mathscr{S}}\to 0.$$

Par conséquent, si $u_n\to 0$ dans $\mathscr{S}$, alors $\|u_n\|_{k+d+1,\mathscr{S}}\to 0$ pour tout k, ce qui implique d'après l'inégalité précédente que $\|\hat{u}_n\|_{k,\mathscr{S}}\to 0$, c'est-à-dire $\hat{u}\to 0$ dans $\mathscr{S}$.

Comme $\mathscr{S}\subset L^2$ et que la transformation de Fourier est injective sur L^2, elle le reste *a fortiori* sur $\mathscr{S}$. Enfin, comme $\mathscr{F}(\mathscr{S})\subset\mathscr{S}\subset L^1$, on peut appliquer la formule d'inversion de Fourier à des fonctions de $\mathscr{S}$, ce qui montre que tout $u\in\mathscr{S}$ peut s'écrire $u=(2\pi)^{-d}\mathscr{F}(\check{\mathscr{F}}u)$ avec $\check{\mathscr{F}}u\in\mathscr{S}$. Donc, la transformation de Fourier est surjective sur $\mathscr{S}$ et son inverse y est clairement continue. □

Proposition 7.3.3 *L'espace $\mathscr{D}(\mathbb{R}^d)$ est dense dans $\mathscr{S}$, au sens où*

$$\forall u\in\mathscr{S},\ \forall\varepsilon>0,\ \forall k\in\mathbb{N},\ \exists\,\varphi\in\mathscr{D} \text{ telle que } \|u-\varphi\|_{k,\mathscr{S}}<\varepsilon.$$

Démonstration. Soit θ un élément de $C^\infty(\mathbb{R})$ égal à 0 sur $[-1/2,1/2]$, à 1 en dehors de $[-1,1]$ et compris entre 0 et 1 (il en existe : prendre la primitive nulle en

0 d'une fonction C^∞ à support dans $[1/2,1]$, positive et d'intégrale 1, complétée par parité). Clairement,

$$\max_{t\in\mathbb{R}}\left|\frac{\theta(t/n)}{t}\right| \leq \frac{2}{n}.$$

On pose

$$\chi(x) = \prod_{i=1}^{d}(1-\theta(x_i)).$$

C'est un élément de $\mathscr{D}(\mathbb{R}^d)$. De plus,

$$1-\chi(x) = \sum_{i=1}^{d} \theta(x_i)\psi_i(x)$$

où les ψ_i sont des fonctions bornées. Posons pour $n \geq 1$

$$\varphi_n(x) \stackrel{\text{déf}}{=} \chi\Big(\frac{x}{n}\Big)u(x).$$

Il est clair que pour tout couple de multi-entiers de longueur inférieure à k, on a

$$x^\beta \partial^\alpha(u-\varphi_n) = x^\beta\Big(1-\chi\Big(\frac{x}{n}\Big)\Big)\partial^\alpha u - \sum_{0\leq\gamma<\alpha} C_\alpha^\gamma x^\beta \partial^\gamma u \frac{1}{n^{|\alpha-\gamma|}}\partial^{\alpha-\gamma}\chi\Big(\frac{x}{n}\Big),$$

avec $C_\alpha^\gamma = \prod_{i=1}^d C_{\alpha_i}^{\gamma_i}$. Par construction de χ,

$$\Big|x^\beta\Big(1-\chi\Big(\frac{x}{n}\Big)\Big)\partial^\alpha u\Big| \leq \frac{2\max_{i,x}|\psi_i|}{n}\sum_{i=1}^{d}|x^{\beta^i}\partial^\alpha u(x)| \leq \frac{2d\max_{i,x}|\psi_i|}{n}\|u\|_{k+1,\mathscr{S}},$$

où le multi-indice β^i est défini par $\beta_i^i = \beta_i + 1$, $\beta_j^i = \beta_j$ pour $j \neq i$. De plus,

$$\Big|x^\beta \partial^\gamma u \frac{1}{n^{|\alpha-\gamma|}}\partial^{\alpha-\gamma}\chi\Big(\frac{x}{n}\Big)\Big| \leq \frac{C(k,\chi)}{n}\|u\|_{k,\mathscr{S}},$$

où la constante $C(k,\chi)$ ne dépend que de k et de χ. Finalement, on obtient

$$|x^\beta \partial^\alpha(u-\varphi_n)(x)| \leq \frac{C'(k,\chi)}{n}\|u\|_{k+1,\mathscr{S}}$$

où la constante $C'(k,\chi)$ ne dépend que de k et de χ, et ce pour tout couple (α,β) de multi-entiers de longueur plus petite que k. D'où

$$\|u-\varphi_n\|_{k,\mathscr{S}} \leq \frac{C'(k,\chi)}{n}\|u\|_{k+1,\mathscr{S}},$$

ce qui clôt la preuve de la proposition. □

Remarque. On peut vérifier que cette propriété de densité est exactement celle correspondant à la distance sur $\mathscr{S}$.

On peut démontrer maintenant la proposition suivante qui généralise la proposition 7.1.1 à d'autres contextes.

Proposition 7.3.4 *Si (f,g) appartient à $L^2 \times L^2$ ou bien à $\mathscr{F}(L^1) \times L^1$, alors, on a*

$$\mathscr{F}(fg) = (2\pi)^{-d}\widehat{f} \star \widehat{g}.$$

Démonstration. Commençons par montrer la formule pour f et g dans $\mathscr{S}$. Comme $\widehat{f}$ et $\widehat{g}$ sont aussi dans $\mathscr{S} \subset L^1$, on a dans ce cas

$$\mathscr{F}(\widehat{f} \star \widehat{g}) = \widehat{\widehat{f}}\,\widehat{\widehat{g}} = (2\pi)^{2d}\check{f}\check{g} = (2\pi)^{2d}(\check{fg}) = (2\pi)^d \mathscr{F}(\mathscr{F}(fg)).$$

Appliquant $\mathscr{F}^{-1}$ aux deux membres, on obtient le résultat dans ce cas.

Il suffit maintenant de démontrer que l'application bilinéaire

$$(f,g) \mapsto \mathscr{F}(fg) - (2\pi)^{-d}\widehat{f} \star \widehat{g}$$

est continue de $\mathscr{S} \times \mathscr{S}$ dans L^∞ quand on munit $\mathscr{S} \times \mathscr{S}$ de la norme $L^2 \times L^2$ ainsi que de la norme $\mathscr{F}(L^1) \times L^1$. L'espace $\mathscr{F}(L^1) \subset L^\infty$ est naturellement muni de la norme L^1 de l'image réciproque, *i.e.* si $f \in \mathscr{F}(L^1)$, $\|f\|_{\mathscr{F}(L^1)} = \|\mathscr{F}^{-1}(f)\|_{L^1}$. Comme $\mathscr{D} \times \mathscr{D}$ est dense dans $\mathscr{S} \times \mathscr{S}$ et dans $L^2 \times L^2$ et que $\mathscr{F}(\mathscr{D}) \times \mathscr{D}$ est dense dans $\mathscr{F}(\mathscr{S}) \times \mathscr{S} = \mathscr{S} \times \mathscr{S}$ et dans $\mathscr{F}(L^1) \times L^1$, on concluera dans les deux cas que l'application bilinéaire est identiquement nulle.

Dans le premier cas, on a

$$\|\mathscr{F}(fg)\|_{L^\infty} \leq \|fg\|_{L^1} \leq \|f\|_{L^2}\|g\|_{L^2},$$

ainsi que

$$\|\widehat{f} \star \widehat{g}\|_{L^\infty} \leq \|\widehat{f}\|_{L^2}\|\widehat{g}\|_{L^2} \leq (2\pi)^d\|f\|_{L^2}\|g\|_{L^2}.$$

De manière analogue, dans le deuxième cas, on a

$$\|\mathscr{F}(fg)\|_{L^\infty} \leq \|fg\|_{L^1} \leq \|f\|_{L^\infty}\|g\|_{L^1} \leq \|\mathscr{F}^{-1}(f)\|_{L^1}\|g\|_{L^1},$$

ainsi que

$$\|\widehat{f} \star \widehat{g})\|_{L^\infty} \leq \|\widehat{f}\|_{L^1}\|\widehat{g}\|_{L^\infty} \leq \|\widehat{f}\|_{L^1}\|g\|_{L^1} = (2\pi)^d\|\mathscr{F}^{-1}(f)\|_{L^1}\|g\|_{L^1}.$$

D'où la proposition. $\square$

Remarque. La transformée de Fourier d'un élément de $\mathscr{F}(L^1)$ est trivialement définie par $\mathscr{F}(f) = (2\pi)^d (\mathscr{F}^{-1}(f))\check{} \in L^1$.

Le lemme suivant montre que l'on ne peut espérer d'une fonction et sa transformée de Fourier soient toutes deux dans $\mathscr{D}$. Autrement dit, l'espace $\mathscr{D}$ n'est absolument pas stable par transformation de Fourier.

Lemme 7.3.1 *Soit φ une fonction de $\mathscr{D}(\mathbb{R})$. Si $\widehat{\varphi}$ appartient à $\mathscr{D}(\mathbb{R})$, alors $\varphi = 0$.*

Démonstration. En effet, on a, par définition de la transformation de Fourier

$$\widehat{\varphi}(\xi) = \int_{\mathbb{R}} e^{-ix\xi} \varphi(x)\, dx.$$

Soit ζ un nombre complexe quelconque. La fonction définie par $x \mapsto e^{-ix\zeta}\varphi(x)$ est une fonction C^∞ à support compact. On peut donc prolonger la fonction $\widehat{\varphi}$ à tout le plan complexe $\mathbb{C}$ par la formule

$$\widehat{\varphi}(\zeta) = \int_{\mathbb{R}} e^{-ix\zeta} \varphi(x)\, dx.$$

De plus, le théorème de dérivation sous l'intégrale nous assure que cette fonction est holomorphe. La seule fonction holomorphe dont les zéros ne sont pas isolés est la fonction nulle, d'où le lemme. □

Le résultat précédent est en rapport avec le principe d'incertitude : on ne peut pas localiser une fonction simultanément en espace et en fréquence.

Exercice 7.3.1 *Démontrez plus précisément que si $\varphi \in \mathscr{D}(\mathbb{R}) \setminus \{0\}$, alors* $\operatorname{Supp} \widehat{\varphi} = \mathbb{R}$.

7.4 Distributions tempérées

Définition 7.4.1 *On appelle* distribution tempérée *toute forme linéaire u sur l'espace $\mathscr{S}(\mathbb{R}^d)$ ayant la propriété de continuité suivante : il existe un entier k et une constante C telle que*

$$|\langle u, \phi \rangle| \leq C \|\phi\|_{k,\mathscr{S}},$$

pour toute fonction ϕ dans l'espace $\mathscr{S}(\mathbb{R}^d)$, où les normes $\|\cdot\|_{k,\mathscr{S}}$ sont définies page 187. On note $\mathscr{S}'(\mathbb{R}^d)$ ou simplement $\mathscr{S}'$ l'espace des distributions tempérées. Enfin, on dit que u_n converge vers u dans $\mathscr{S}'$ si et seulement si $\langle u_n - u, \phi \rangle \to 0$ pour toute fonction ϕ dans $\mathscr{S}$.

L'espace $\mathscr{S}'$ est exactement le dual topologique de $\mathscr{S}$. En effet, il est clair que si u est une distribution tempérée, alors c'est une forme linéaire continue sur $\mathscr{S}$. Réciproquement, on peut voir (exercice, en raisonnant par contraposition) que toute forme linéaire continue sur $\mathscr{S}$ satisfait la définition d'une distribution tempérée.

On n'augmente pas la généralité en considérant les distributions tempérées.

Proposition 7.4.1 *Toute distribution tempérée est une distribution sur* $\mathbb{R}^d$.

Démonstration. En effet

$$\forall \varphi \in \mathscr{D}_K, |\langle u, \varphi\rangle| \leq C\|\varphi\|_{k,\mathscr{S}} = C \sup_{\substack{x\in\mathbb{R}^d \\ |\alpha|,|\beta|\leq k}} |x^\beta \partial^\alpha \varphi(x)| \leq C_K \sup_{|\alpha|\leq k} \|\partial^\alpha \varphi\|_{L^\infty}.$$

□

La réciproque n'est pas vraie, c'est-à-dire qu'il existe des distributions sur $\mathbb{R}^d$ qui ne sont pas des distributions tempérées.

La théorie des distributions visant à généraliser la notion de fonction, il est légitime de se demander où apparaissent les fonctions dans ce cadre. Pour répondre à cette question, on définit l'espace suivant

Définition 7.4.2 *Soit p un élément de $[1,+\infty]$, on désigne par $L^p_M(\mathbb{R}^d)$ l'ensemble des fonctions localement intégrables f telles qu'il existe un entier N tel que $(1+|x|)^{-N} f(x)$ soit dans $L^p(\mathbb{R}^d)$.*

Exercice 7.4.1 *Démontrer que les espaces $L^p_M(\mathbb{R}^d)$ forment une famille décroissante d'espaces, c'est-à-dire que si $p \leq q$, alors $L^q_M(\mathbb{R}^d) \subset L^p_M(\mathbb{R}^d)$.*

Exercice 7.4.2 *Démontrer que les fonctions polynomiales et les polynômes trigonométriques appartiennent à $L^1_M(\mathbb{R}^d)$ mais pas les fonctions à croissance exponentielle comme $x \mapsto \exp(|x|)$.*

Remarque L'exercice ci-dessus et le théorème qui suit montrent que la notion de distribution tempérée contient une information sur le comportement à l'infini.

L'identification des fonctions avec les distributions tempérées se fait grâce au théorème que l'on prouvera en exercice.

Théorème 7.4.1 *Soit ι l'application définie par*

$$\begin{array}{rcl} \iota : L^1_M(\mathbb{R}^d) & \to & \mathscr{S}'(\mathbb{R}^d) \\ f & \mapsto & \iota(f) : \phi \mapsto \displaystyle\int_{\mathbb{R}^d} f(x)\phi(x)dx \end{array}$$

est une injection linéaire. De plus, on a la propriété de « continuité » suivante. Soit $N \in \mathbb{N}$ tel que $(1+|x|)^{-N} f(x) \in L^1(\mathbb{R}^d)$, alors

$$|\langle \iota(f), \phi\rangle| \leq \|(1+|\cdot|)^{-N} f\|_{L^1} \|\phi\|_{N,\mathscr{S}}.$$

Remarque. Tous les espaces $L^p(\mathbb{R}^d)$ sont inclus dans $L^1_M(\mathbb{R}^d)$. On peut donc considérer toute fonction de $L^p(\mathbb{R}^d)$ comme une distribution tempérée.

7.5 Opérations et transformation de Fourier sur les distributions tempérées

Nous allons définir un certain nombre d'opérations linéaires continues sur les fonctions de $\mathscr{S}$, puis nous étendrons ces opérations par dualité à l'espace $\mathscr{S}'$.

Tout d'abord, définissons quelques espaces de fonctions.

Définition 7.5.1 *On appelle* espace des fonctions à croissance modérée (*ou lente*) *et l'on désigne par* $\mathscr{O}_M$ *l'espace des fonctions* f *indéfiniment différentiables sur* $\mathbb{R}^d$ *telles que*

$$\forall k \in \mathbb{N}, \exists N \in \mathbb{N}, \exists C > 0 \text{ tel que } \sup_{|\alpha|=k} |\partial^\alpha f(x)| \leq C(1+|x|)^N.$$

On désigne par $L^1_{\mathscr{S}}$ *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 $\mathscr{O}_M$. Les fonctions localement intégrables décroissant à l'infini plus vite que toute puissance négative de $|x|$ sont d'excellents exemples de fonctions de $L^1_{\mathscr{S}}$.

Proposition 7.5.1 *Les applications linéaires suivantes sont continues de* $\mathscr{S}$ *dans* $\mathscr{S}$ *:*

i) L'application ∂^α*,*
ii) pour tout $f \in \mathscr{O}_M$*, la multiplication par* f*,*
iii) pour tout $f \in L^1_{\mathscr{S}}$*, la convolution par* f*,*
iv) la transformation de Fourier.

Démonstration. On a déjà vu que la continuité sur $\mathscr{S}$ est impliquée par la continuité en 0, et que celle-ci est équivalente à la continuité pour chaque norme $\|\cdot\|_{k,\mathscr{S}}$.

i) Il est clair que, pour tout entier k, on a

$$\|\partial^\alpha \phi\|_{k,\mathscr{S}} \leq \|\phi\|_{k+|\alpha|,\mathscr{S}}.$$

ii) Appliquant la formule de Leibniz, il vient

$$\partial^\alpha(f\phi) = \sum_{\beta \leq \alpha} C_\alpha^\beta \partial^{\alpha-\beta} f \partial^\beta \phi.$$

Comme la fonction f appartient à $\mathscr{O}_M$, pour tout $k \in \mathbb{N}$, il existe un entier N tel que, pour tout multi-entier γ de longueur plus petite que k, on ait

$$\forall x \in \mathbb{R}^d,\ |\partial^\gamma f(x)| \leq C(1+|x|)^N.$$

On en déduit alors que

$$\|f\phi\|_{k,\mathscr{S}} \le C_k \|\phi\|_{k+N,\mathscr{S}}.$$

iii) Soit k un entier et α et β deux multi-entiers de longueur inférieure à k. Par dérivation sous le signe somme, on a

$$\partial^\alpha (f \star \phi)(x) = \int_{\mathbb{R}^d} f(y) \partial^\alpha \phi(x-y) dy.$$

Les fonctions $t \mapsto t^{\beta_i}$ sont convexes, croissantes sur $\mathbb{R}_+$. Par conséquent, on a

$$|x^\beta| \le \prod_{i=1}^d (|y_i| + |x_i - y_i|)^{\beta_i} \le 2^{|\beta|-d} \prod_{i=1}^d (|y_i|^{\beta_i} + |x_i - y_i|^{\beta_i}) \le 2^{k-d} \sum_{i,j=1}^d |y_i|^{\beta_i} |x_j - y_j|^{\beta_j}.$$

Ainsi donc, on peut écrire que

$$|x^\beta| |\partial^\alpha (f \star \phi)(x)| \le 2^{k-d} \sum_{i,j=1}^d \int_{\mathbb{R}^d} |y_i|^{\beta_i} |f(y)| |x_j - y_j|^{\beta_j} |\partial^\alpha \phi(x-y)| \, dy.$$

Or, $|x_j - y_j|^{\beta_j} |\partial^\alpha \phi(x-y)| \le \|\phi\|_{k,\mathscr{S}}$ et $|y_i|^{\beta_i} \le 1 + |y|^k$. On en déduit alors que

$$\|f \star \phi\|_{k,\mathscr{S}} \le d^2 2^{k-d} \|(1+|\cdot|)^k f\|_{L^1} \|\phi\|_{k,\mathscr{S}}.$$

iv) Il s'agit du théorème 7.3.1, démontré page 189. □

Nous allons maintenant étendre ces opérations à l'espace $\mathscr{S}'$.

Définition 7.5.2 *Pour tout $u \in \mathscr{S}'$, on définit*
i) la dérivation par

$$\langle \partial^\alpha u, \phi \rangle = (-1)^{|\alpha|} \langle u, \partial^\alpha \phi \rangle,$$

ii) la multiplication par une fonction f de $\mathscr{O}_M$ par

$$\langle fu, \phi \rangle = \langle u, f\phi \rangle,$$

iii) la convolution par une fonction f de $L^1_{\mathscr{S}}$ par

$$\langle f \star u, \phi \rangle = \langle u, \check{f} \star \phi \rangle,$$

iv) et enfin la transformée de Fourier par

$$\langle \mathscr{F}(u), \phi \rangle = \langle u, \mathscr{F}(\phi) \rangle.$$

Toutes ces opérations sont continues sur $\mathscr{S}'$.

Démonstration. La continuité est évidente grâce aux définitions par transposition et à la définition de la convergence dans $\mathscr{S}'$. La vérification du fait que l'on étend bien ainsi à $\mathscr{S}'$ les notions usuelles est sans surprise. Regardons quand même le cas de la transformation de Fourier d'un peu plus près. Considérons une fonction f de l'espace $L^1(\mathbb{R}^d) \subset L^1_M(\mathbb{R}^d)$. En notant exceptionnellement u_f la distribution tempérée associée à la fonction f, on a par définition

$$\forall \phi \in \mathscr{S}, \quad \langle \widehat{u_f}, \phi \rangle = \int_{\mathbb{R}^d} f(x)\widehat{\phi}(x)\,dx.$$

Le lemme 7.2.1 affirme en particulier que

$$\forall \phi \in \mathscr{S}, \quad \int_{\mathbb{R}^d} \widehat{f}(x)\phi(x)\,dx = \int_{\mathbb{R}^d} f(x)\widehat{\phi}(x)\,dx.$$

On en déduit que $\widehat{u_f} = \widehat{f}$. □

Nous avons aussi défini la transformation de Fourier sur L^2 par prolongement continu de $L^1 \cap L^2$ à L^2. Comme $L^2 \subset L^1_M \subset \mathscr{S}'$, il convient aussi de vérifier que notre définition au sens des distributions tempérées, restreinte à l'espace L^2, coïncide avec la transformation de Fourier L^2. On sait déjà qu'elle coïncide avec la transformation de Fourier usuelle sur $L^1 \cap L^2$, lequel est dense dans L^2. Soit $u \in L^2$ et $u^n \in L^1 \cap L^2$ une suite telle que $u^n \to u$ dans L^2. Notons que si $u^n \to u$ dans L^2, alors trivialement, $u^n \to u$ dans $\mathscr{S}'$. Par continuité de la transformation de Fourier sur ces deux espaces, on en déduit que $\hat{u}^n \to \hat{u}$ dans L^2 et $\mathscr{F}(u^n) \to \mathscr{F}(u)$ dans $\mathscr{S}'$ (en distinguant très provisoirement les deux définitions de la transformée de Fourier). Comme $\hat{u}^n = \mathscr{F}(u^n)$, on voit que $\hat{u} = \mathscr{F}(u)$.

Nous allons maintenant étendre aux distributions tempérées les diverses formules sur la transformation de Fourier démontrées dans le cadre de la transformation de Fourier sur $\mathscr{S}$. Commençons par le théorème d'inversion de Fourier en version distributions tempérées.

Théorème 7.5.1 *La transformation de Fourier est inversible sur $\mathscr{S}'$ et l'on a*

$$\forall u \in \mathscr{S}', \quad \mathscr{F}^{-1}u = (2\pi)^{-d}(\check{\mathscr{F}}u).$$

Démonstration. Définissons d'abord pour toute distribution tempérée v la distribution tempérée $\check{v}$ par transposition, comme d'habitude,

$$\forall \phi \in \mathscr{S}, \quad \langle \check{v}, \phi \rangle = \langle v, \check{\phi} \rangle$$

Calculons ensuite $\mathscr{F}(\check{\mathscr{F}}u)$. Il vient

$$\forall \phi \in \mathscr{S}, \quad \langle \mathscr{F}(\check{\mathscr{F}}u), \phi \rangle = \langle u, \mathscr{F}(\check{\mathscr{F}}\phi) \rangle.$$

Or, par le théorème 7.2.1 d'inversion de Fourier dans L^1, on a

$$\mathscr{F}(\check{\mathscr{F}}\phi) = (2\pi)^d \phi,$$

d'où le théorème. □

Proposition 7.5.2 *Soit u une distribution tempérée. On a alors*

$$\mathscr{F}(D^\alpha u) = \xi^\alpha \widehat{u} \quad et \quad \mathscr{F}(x^\alpha u) = (-D)^\alpha \widehat{u}.$$

Démonstration. On utilise la proposition 7.3.2. Par définition,

$$\begin{aligned}\langle \mathscr{F}(D^\alpha u), \phi\rangle &= \langle D^\alpha u, \widehat{\phi}\rangle \\ &= \langle u, (-D)^\alpha \widehat{\phi}\rangle.\end{aligned}$$

La proposition 7.3.2 nous dit que

$$(-D)^\alpha \widehat{\phi} = \mathscr{F}(x^\alpha \phi).$$

D'où il résulte que

$$\begin{aligned}\langle \mathscr{F}(D^\alpha u), \phi\rangle &= \langle u, \mathscr{F}(x^\alpha \phi)\rangle \\ &= \langle \widehat{u}, (\xi^\alpha \phi)\rangle \\ &= \langle \xi^\alpha \widehat{u}, \phi\rangle.\end{aligned}$$

Comme la démonstration de la seconde relation est en tout point analogue, elle est laissée au lecteur. □

Ces relations sont extrêmement importantes dans les calculs pratiques de transformée de Fourier. Par exemple, calculons la transformée de Fourier de la masse de Dirac δ_0. Par définition de la transformation de Fourier, on a

$$\langle \widehat{\delta_0}, \phi\rangle = \langle \delta_0, \widehat{\phi}\rangle = \widehat{\phi}(0) = \int_{\mathbb{R}^d} \phi(x)dx = \langle 1, \phi\rangle$$

pour tout $\phi \in \mathscr{S}$. En d'autres termes,

$$\widehat{\delta_0} = 1. \tag{7.10}$$

De ceci et du théorème d'inversion de Fourier 7.5.1, on déduit que

$$\mathscr{F}(1) = (2\pi)^d \delta_0. \tag{7.11}$$

L'utilisation pratique des formules ci-dessus est très souvent couplée avec des résolutions d'équations simples sur les distributions.

Parmi toutes les applications de la transformation de Fourier, citons en exemple celle-ci.

Théorème 7.5.2 *Soit $f \in \mathscr{S}'(\mathbb{R}^d)$, il existe une unique distribution u de $\mathscr{S}'(\mathbb{R}^d)$ solution de*

$$(-\Delta + 1)u = f \quad où \quad \Delta = \sum_{j=1}^d \frac{\partial^2}{\partial x_j^2}.$$

Démonstration. Utilisons la proposition 7.5.2 pour affirmer que

$$\mathscr{F}((-\Delta+1)u) = (1+|\xi|^2)\widehat{u}.$$

Le théorème 7.5.1 d'inversion de Fourier permet alors d'écrire que

$$(-\Delta+1)u = f \Longleftrightarrow (1+|\xi|^2)\widehat{u} = \widehat{f}.$$

La fonction $\xi \mapsto (1+|\xi|^2)^{-1}$ est une fonction indéfiniment différentiable à croissance lente. On peut donc multiplier l'égalité précédente par cette fonction et l'on obtient

$$\widehat{u} = \frac{\widehat{f}}{1+|\xi|^2}.$$

Par conséquent, l'équation aux dérivées partielles considérée a une et une seule solution dans $\mathscr{S}'$, qui s'écrit

$$u = \mathscr{F}^{-1}\Big(\frac{\widehat{f}}{1+|\xi|^2}\Big),$$

en appliquant la transformation de Fourier inverse. □

Un cas particulier intéressant est celui où $f = \delta_0$. La solution correspondante

$$u_0 = \mathscr{F}^{-1}\Big(\frac{1}{1+|\xi|^2}\Big),$$

est appelée *solution élémentaire* de l'opérateur $-\Delta+1$. On vérifie que, sous réserve de pouvoir effectuer la convolution, la solution avec second membre f est alors donnée par $u = u_0 \star f$.

L'utilité de la transformation de Fourier dans les problèmes d'équations aux dérivées partielles est bien illustrée par le théorème précédent : la transformation de Fourier transforme les opérateurs différentiels en opérateurs algébriques, plus faciles à manipuler.

Chapitre 8

Espaces de Sobolev

Dans ce cours, nous nous limiterons aux espaces de Sobolev modelés sur L^2.

8.1 Définition des espaces de Sobolev sur $\mathbb{R}^d$

Définition 8.1.1 *Soit s un réel, on dit qu'une distribution tempérée u appartient à* l'espace de Sobolev d'indice *s, noté $H^s(\mathbb{R}^d)$, ou simplement H^s en l'absence d'ambiguïté, si et seulement si*

$$\widehat{u} \in L^2(\mathbb{R}^d; (1+|\xi|^2)^s \, d\xi).$$

et l'on note

$$\|u\|_{H^s}^2 = \int_{\mathbb{R}^d} (1+|\xi|^2)^s |\widehat{u}(\xi)|^2 \, d\xi.$$

Proposition 8.1.1 *Pour tout s réel, l'espace H^s, muni de la norme $\|\cdot\|_{H^s}$, est un espace de Hilbert.*

Démonstration. Le fait que la norme $\|\cdot\|_{H^s}$ provienne du produit scalaire

$$(u|v)_{H^s} \stackrel{\text{déf}}{=} \int_{\mathbb{R}^d} (1+|\xi|^2)^s \widehat{u}(\xi) \overline{\widehat{v}(\xi)} \, d\xi$$

est évident. Démontrons que H^s est un espace complet. Soit $(u_n)_{n\in\mathbb{N}}$ une suite de Cauchy de H^s. Par définition de la norme, la suite $(\widehat{u}_n)_{n\in\mathbb{N}}$ est une suite de Cauchy de l'espace $L^2(\mathbb{R}^d; (1+|\xi|^2)^s \, d\xi)$, lequel est complet. Donc, il existe une fonction $\widetilde{u}$ appartenant à l'espace $L^2(\mathbb{R}^d; (1+|\xi|^2)^s \, d\xi)$ telle que

$$\lim_{n\to\infty} \|\widehat{u}_n - \widetilde{u}\|_{L^2(\mathbb{R}^d; (1+|\xi|^2)^s d\xi)} = 0. \tag{8.1}$$

Comme $L^2(\mathbb{R}^d; (1+|\xi|^2)^s \, d\xi)$ est inclus dans l'espace des distributions tempérées $\mathscr{S}'$, on peut poser $u = \mathscr{F}^{-1}\widetilde{u}$. Par construction, c'est un élément de H^s et d'après (8.1), on a bien $\|u_n - u\|_{H^s} \to 0$ quand $n \to +\infty$. $\square$

Pour dire brièvement les choses, la transformation de Fourier est un isomorphisme isométrique de H^s sur $L^2(\mathbb{R}^d;(1+|\xi|^2)^s d\xi)$. En particulier, $H^0 = L^2$ joue le rôle d'espace-pivot dans l'échelle des espaces H^s, $s \in \mathbb{R}$. Il est clair que $H^{s_2} \hookrightarrow H^{s_1}$, où la flèche $\hookrightarrow$ désigne une injection continue, dès que $s_1 \leq s_2$.

Le cas où s est un entier naturel est particulier.

Proposition 8.1.2 *Soit m est un entier positif. L'espace $H^m(\mathbb{R}^d)$ est l'espace des fonctions u de $L^2(\mathbb{R}^d)$ dont toutes les dérivées d'ordre inférieur ou égal à m sont des distributions appartenant à $L^2(\mathbb{R}^d)$. De plus, la norme*

$$\widetilde{\|}u\|_{H^m}^2 \stackrel{déf}{=} \sum_{|\alpha|\leq m} \|\partial^\alpha u\|_{L^2}^2$$

munit l'espace H^m d'une structure d'espace de Hilbert et cette norme est équivalente à la norme $\|\cdot\|_{H^m}$ définie à l'aide de la transformation de Fourier.

Démonstration. Le fait que

$$\widetilde{\|}u\|_{H^m}^2 = \widetilde{(}u|u)_{H^m} \quad \text{avec} \quad \widetilde{(}u|v)_{H^m} \stackrel{\text{déf}}{=} \sum_{|\alpha|\leq m} \int_{\mathbb{R}^d} \partial^\alpha u(x)\overline{\partial^\alpha v(x)}dx.$$

assure que la norme $\widetilde{\|}\cdot\|_{H^m}$ provient d'un produit scalaire. De plus, il existe une constante C telle que

$$\forall\, \xi \in \mathbb{R}^d,\ C^{-1}\Big(1+\sum_{0<|\alpha|\leq m} |\xi|^{2|\alpha|}\Big) \leq (1+|\xi|^2)^m \leq C\Big(1+\sum_{0<|\alpha|\leq m} |\xi|^{2|\alpha|}\Big). \tag{8.2}$$

Comme la transformation de Fourier est, à une constante près, une isométrie de L^2 sur L^2, alors on a

$$\partial^\alpha u \in L^2 \Leftrightarrow \xi^\alpha \widehat{u} \in L^2.$$

Donc, on en déduit que

$$u \in H^m \Leftrightarrow \forall\, \alpha \text{ tel que } |\alpha| \leq m,\ \partial^\alpha u \in L^2.$$

Enfin, l'inégalité (8.2) assure l'équivalence des normes grâce au fait que la transformation de Fourier est une isométrie à une constante près. □

Remarques. Cette proposition nous indique comment on peut définir des espaces de Sobolev d'indice entier positif modelés sur les espaces L^p pour $p \in [1,+\infty]$. Il suffit de définir l'espace $W^{m,p}(\mathbb{R}^d)$ comme l'espace des fonctions u de $L^p(\mathbb{R}^d)$ dont toutes les dérivées d'ordre inférieur ou égal à m sont des distributions appartenant à $L^p(\mathbb{R}^d)$, muni de la norme idoine.

On voit aussi que l'indice s de l'espace de Sobolev s'interprète comme un ordre de dérivées appartenant à L^2 dans le cas entier, donc une mesure de régularité. Dans le cas s réel, on a aussi une interprétation en termes de dérivées d'ordre fractionnaire.

Exercice 8.1.1 *Démontrez que l'espace $\mathscr{S}$ est inclus dans l'espace H^s, et ce pour tout réel s.*

Exercice 8.1.2 *Démontrer que la masse de Dirac δ_0 appartient à l'espace $H^{-\frac{d}{2}-\varepsilon}$ pour tout réel strictement positif ε, mais que δ_0 n'appartient pas à l'espace $H^{-\frac{d}{2}}$.*

Exercice 8.1.3 *Démontrez que, pour toute distribution à support compact u, il existe un réel s telle que u appartienne à l'espace de Sobolev H^s.*

Exercice 8.1.4 *Démontrer que la constante 1 n'appartient à H^s pour aucun réel s.*

Exercice 8.1.5 *Soit s un réel de l'intervalle $]0,1[$. Démontrer que l'espace H^s est l'espace des fonctions u de L^2 telles que*

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{|u(x)-u(y)|^2}{|x-y|^{d+2s}}\,dxdy < +\infty.$$

Démontrer de plus qu'il existe une constante C telle que, pour toute fonction u de H^s, on ait

$$C^{-1}\|u\|_{H^s}^2 \leq \|u\|_{L^2}^2 + \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{|u(x)-u(y)|^2}{|x-y|^{d+2s}}\,dxdy \leq C\|u\|_{H^s}^2.$$

Comment pourrait-on définir un espace $W^{s,p}$ de façon analogue ?

Théorème 8.1.1 *Soit s un réel quelconque. L'espace $\mathscr{D}(\mathbb{R}^d)$ est dense dans $H^s(\mathbb{R}^d)$ et la multiplication par une fonction de $\mathscr{S}$ est une application continue de H^s dans lui-même.*

Démonstration. Pour prouver le premier point de ce théorème, il suffit de démontrer que l'orthogonal de $\mathscr{D}(\mathbb{R}^d)$ est réduit au vecteur nul, voir corollaire 4.2.2. Considérons une distribution u de H^s telle que, pour toute fonction-test φ, on ait $(\varphi|u)_{H^s} = 0$. Ceci signifie que, pour toute fonction-test φ, on a

$$\int_{\mathbb{R}^d} \widehat{\varphi}(\xi)(1+|\xi|^2)^s\overline{\widehat{u}(\xi)}\,d\xi = 0.$$

En utilisant que $\mathscr{S}$ est inclus dans H^s, que, d'après la proposition 7.3.3, $\mathscr{D}$ est dense dans $\mathscr{S}$, et que la transformation de Fourier est un isomorphisme de $\mathscr{S}$ dans lui-même, on a, pour toute fonction ϕ de $\mathscr{S}$,

$$\int_{\mathbb{R}^d} \phi(\xi)(1+|\xi|^2)^s \overline{\widehat{u}(\xi)}\, d\xi = 0.$$

Comme $\mathscr{S}$ est dense dans L^2, ceci entraîne que l'on a $(1+|\xi|^2)^s \widehat{u}(\xi) = 0$, donc $\widehat{u} = 0$, donc $u = 0$.

Démontrons maintenant le second point du théorème. Cette démonstration est présentée ici à titre culturel. Dans les calculs qui suivent, la constante C est générique et sa valeur peut changer d'une occurrence à l'autre. Soit ϕ une fonction de $\mathscr{S}$.

D'après la proposition 7.3.4, on sait que

$$\widehat{\phi u} = (2\pi)^{-d} \widehat{\phi} \star \widehat{u}.$$

On souhaite montrer qu'il existe une constante C telle que $\|\phi u\|_{H^s} \leq C\|u\|_{H^s}$. Il s'agit donc de majorer la norme L^2 de la fonction définie par

$$U(\xi) = (1+|\xi^2|)^{\frac{s}{2}} \int_{\mathbb{R}^d} |\widehat{\phi}(\xi-\eta)| \times |\widehat{u}(\eta)|\, d\eta$$

par une constante fois la norme H^s de u. Posons $I_1(\xi) = \{\eta \,/\, 2|\xi-\eta| \leq |\eta|\}$ et $I_2(\xi) = \{\eta \,/\, 2|\xi-\eta| \geq |\eta|\}$. Il est clair que l'on a

$$\begin{aligned} U(\xi) &= U_1(\xi) + U_2(\xi) \quad \text{avec} \\ U_j(\xi) &= (1+|\xi|^2)^{\frac{s}{2}} \int_{I_j(\xi)} |\widehat{\phi}(\xi-\eta)| \times |\widehat{u}(\eta)|\, d\eta. \end{aligned}$$

Tout d'abord, observons que si $\eta \in I_1(\xi)$,alors on a

$$\frac{1}{2}|\eta| \leq |\xi| \leq \frac{3}{2}|\eta|.$$

On en déduit que, pour tout réel s, il existe une constante C telle que, pour tout couple (ξ,η) tel que η appartienne à $I_1(\xi)$,

$$(1+|\xi|^2)^s \leq C(1+|\eta|^2)^s.$$

Il vient que

$$U_1(\xi) \leq C \int_{\mathbb{R}^d} |\widehat{\phi}(\xi-\eta)|(1+|\eta|^2)^{\frac{s}{2}} |\widehat{u}(\eta)|\, d\eta = C\widehat{\phi} \star w(\xi),$$

où $w(\eta) = (1+|\eta|^2)^{\frac{s}{2}}|\widehat{u}(\eta)|$. Comme $\widehat{\phi}$ appartient à $\mathscr{S}$, en particulier $\widehat{\phi}$ appartient à L^1, donc

$$\|U_1\|_{L^2} \leq C\|\widehat{\phi}\|_{L^1}\|w\|_{L^2} = C\|\widehat{\phi}\|_{L^1}\|u\|_{H^s},$$

par le théorème 5.4.3.

Par ailleurs, si η appartient à $I_2(\xi)$, on a

$$\frac{1}{2}|\xi-\eta| \leq |\xi| \leq \frac{3}{2}|\xi-\eta| \quad \text{et} \quad |\eta| \leq 2|\xi-\eta|.$$

Distinguons deux cas. Si $s \geq 0$, alors $(1+|\xi|^2)^{\frac{s}{2}} \leq C(1+|\xi-\eta|^2)^{\frac{s}{2}}$, d'où

$$U_2(\xi) \leq C\int_{\mathbb{R}^d} |\widehat{\phi}(\xi-\eta)|(1+|\xi-\eta|^2)^{\frac{s}{2}}|\widehat{u}|\,d\eta = Cw\star\widehat{u}(\xi),$$

où $w(\xi) = |\widehat{\phi}(\xi)|(1+|\xi|^2)^{\frac{s}{2}} \in L^1$ puisque $\phi \in \mathscr{S}$. On en déduit comme précédemment que $U_2 \in L^2$ avec

$$\|U_2\|_{L^2} \leq C\|w\|_{L^1}\|\widehat{u}\|_{L^2} \leq C\|w\|_{L^1}\|u\|_{H^s},$$

d'où le résultat dans ce cas.

Si maintenant $s < 0$, on note que l'on a toujours $(1+|\xi|^2)^{\frac{s}{2}} \leq C(1+|\xi-\eta|^2)^{\frac{s}{2}}$, mais aussi que $(1+|\eta|^2)^{-\frac{s}{2}} \leq C(1+|\xi-\eta|^2)^{-\frac{s}{2}}$. Par conséquent, on peut écrire

$$U_2(\xi) \leq C\int_{\mathbb{R}^d} |\widehat{\phi}(\xi-\eta)|(1+|\eta|^2)^{\frac{s}{2}}|\widehat{u}(\eta)|\,d\eta,$$

et l'on obtient une nouvelle fois

$$\|U_2\|_{L^2} \leq C\|u\|_{H^s}$$

par le même argument d'estimation de convolution que plus haut. □

Exercice 8.1.6 *Soit $\mathscr{F}L^1 = \{u \in \mathscr{S}' \,/\, \widehat{u} \in L^1\}$. Démontrez que, pour tout réel positif s, le produit est une application bilinéaire continue de $\mathscr{F}L^1 \cap H^s \times \mathscr{F}L^1 \cap H^s$. Qu'en déduire lorsque s est strictement supérieur à $d/2$?*

Exercice 8.1.7 *Soit s un réel strictement supérieur à $1/2$. Démontrez que l'application restriction γ définie par*

$$\begin{array}{rcl} \gamma\colon \mathscr{D}(\mathbb{R}^d) & \longrightarrow & \mathscr{D}(\mathbb{R}^{d-1}) \\ \varphi & \longmapsto & \gamma(\varphi) : (x_2,\cdots,x_d) \mapsto \varphi(0,x_2,\cdots,x_d) \end{array}$$

se prolonge en une application continue de $H^s(\mathbb{R}^d)$ dans $H^{s-\frac{1}{2}}(\mathbb{R}^{d-1})$.
Indication : *On écrira que*

$$\mathscr{F}_{\mathbb{R}^{d-1}}\varphi(0,\xi_2,\cdots,\xi_d) = (2\pi)^{-1}\int_{\mathbb{R}} \widehat{\varphi}(\xi_1,\xi_2,\cdots,\xi_d)\,d\xi_1.$$

Nous allons maintenant démontrer un théorème qui décrit le dual de l'espace H^s.

Théorème 8.1.2 *La forme bilinéaire B définie par*

$$
\begin{aligned}
B\colon \mathscr{S} \times \mathscr{S} &\to \mathbb{K} \\
(u, \varphi) &\mapsto \int_{\mathbb{R}^d} u(x)\varphi(x)\,dx
\end{aligned}
$$

se prolonge en une forme bilinéaire continue de $H^{-s} \times H^s$ dans $\mathbb{K}$. De plus, l'application $\delta_B\colon H^{-s} \to (H^s)'$ qui lui est associée est un isomorphisme linéaire isométrique (à une constante près), c'est-à-dire que la forme bilinéaire B identifie l'espace H^{-s} au dual de H^s.

Démonstration. Le point important de la démonstration de ce théorème est lié à la formule d'inversion de Fourier. Le lemme 7.2.1 page 186 et la formule d'inversion de Fourier assure que l'on a, pour tout couple (u, φ) de fonctions de $\mathscr{S}$,

$$
\begin{aligned}
B(u, \varphi) &= \int_{\mathbb{R}^d} u(x)\varphi(x)\,dx \\
&= \int_{\mathbb{R}^d} u(x)\mathscr{F}(\mathscr{F}^{-1}\varphi)(x)\,dx \\
&= \int_{\mathbb{R}^d} \widehat{u}(\xi)(\mathscr{F}^{-1}\varphi)(\xi)\,d\xi \\
&= (2\pi)^{-d} \int_{\mathbb{R}^d} \widehat{u}(\xi)\widehat{\varphi}(-\xi)\,d\xi. \qquad (8.3)
\end{aligned}
$$

On en déduit donc immédiatement en multipliant et en divisant par $(1+|\xi|^2)^{\frac{s}{2}}$ et en utilisant l'inégalité de Cauchy-Schwarz que

$$
|B(u, \varphi)| \leq (2\pi)^{-d} \|u\|_{H^s} \|\varphi\|_{H^{-s}},
$$

d'où le premier point du théorème. Le fait que l'application δ_B soit injective résulte du fait que, si, pour toute fonction $\varphi \in \mathscr{S}$, on a $B(u, \varphi) = 0$, alors u est l'élément nul de $\mathscr{S}'$. Reste à démontrer qu'elle est surjective. En fait, nous allons démontrer directement que δ_B est un isomorphisme.

Pour tout réel σ, la transformation de Fourier est par construction un isomorphisme isométrique de H^σ sur $L^2_\sigma = L^2(\mathbb{R}^d, (1+|\xi|^2)^\sigma d\xi)$. Considérons maintenant la forme bilinéaire $\check{B}$ définie par

$$
\begin{aligned}
\check{B}\colon L^2_{-s} \times L^2_s &\to \mathbb{K} \\
(\varphi, u) &\mapsto (2\pi)^{-d} \textstyle\int_{\mathbb{R}^d} \widehat{u}(\xi)\widehat{\varphi}(-\xi)\,d\xi.
\end{aligned}
$$

Si nous démontrons que

$$
\delta_B = {}^t\mathscr{F}\,\delta_{\check{B}}\mathscr{F}, \qquad (8.4)
$$

alors le théorème 8.1.2 est démontré. En effet, comme on sait que $\mathscr{F}$ est un isomorphisme de H^s sur L^2_s, d'après la proposition 3.4.2 page 92, l'application linéaire ${}^t\mathscr{F}$ est un isomorphisme de $(L^2_s)'$ sur $(H^s)'$. D'après le corollaire 4.3.3 page 104, on sait déjà que $\delta_{\check{B}}$ est un isomorphisme de $(L^2_s)'$ sur L^2_{-s}.

Pour démontrer la formule (8.4), écrivons que

$$\begin{aligned} \langle {}^t\mathscr{F}\delta_{\check{B}}\mathscr{F}u, \varphi\rangle &= \langle \mathscr{F}\delta_{\check{B}}\mathscr{F}u, \mathscr{F}\varphi\rangle \\ &= \delta_{\check{B}}(\mathscr{F}u, \mathscr{F}\varphi) \end{aligned}$$

D'après la formule (8.3), on a

$$\langle {}^t\mathscr{F}\delta_{\check{B}}\mathscr{F}u, \varphi\rangle = \langle \delta_B(u), \varphi\rangle,$$

d'où le théorème. □

8.2 Injections de Sobolev

Le but de cette section est d'étudier les propriétés d'inclusion des espaces $H^s(\mathbb{R}^d)$ dans les espaces $L^p(\mathbb{R}^d)$. Nous allons démontrer le théorème suivant.

Théorème 8.2.1 *Si $s > d/2$, alors l'espace H^s est s'injecte continûment dans l'espace des fonctions continues nulles à l'infini. Si $0 \leq s < d/2$, alors l'espace H^s s'injecte continûment dans L^{s^*} où $s^* = \frac{2d}{d-2s}$.*

Démonstration. Rappelons que l'on dit qu'un espace E s'injecte continûment dans un espace F si $E \subset F$ (algébriquement, ou parfois à un isomorphisme près) et que l'injection de E dans F est continue, c'est-à-dire qu'il existe une constante C telle que pour tout $u \in E$, $\|u\|_F \leq C\|u\|_E$.

Le premier point du théorème est très facile à démontrer. Écrivons pour tout u dans H^s,

$$|\widehat{u}(\xi)| = (1+|\xi|^2)^{-s/2}(1+|\xi|^2)^{s/2}|\widehat{u}(\xi)|. \tag{8.5}$$

Le fait que $s > d/2$ implique que la fonction

$$\xi \mapsto (1+|\xi|^2)^{-s/2}$$

appartient à L^2. Donc, par l'inégalité de Cauchy-Schwarz, on en déduit que

$$\|\widehat{u}\|_{L^1} \leq \Big(\int_{\mathbb{R}^d}(1+|\xi|^2)^{-s}d\xi\Big)^{\frac{1}{2}}\|u\|_{H^s}.$$

Or on sait que si $\widehat{u}$ est dans L^1, alors u est une fonction continue qui tend vers zéro à l'infini avec

$$\|u\|_{L^\infty} \leq (2\pi)^{-d}\|\widehat{u}\|_{L^1}, \tag{8.6}$$

d'où le premier point.

La démonstration du second point est plus délicate et est présentée ici à titre culturel. Les constantes C qui apparaissent sont génériques, et peuvent varier d'une ligne à l'autre.

On utilise alors le résultat de l'exercice 5.1.1 page 133 que nous rappelons. Pour tout p dans l'intervalle $]1,+\infty[$, on a, pour toute fonction mesurable f,

$$\|f\|_{L^p}^p = p\int_0^\infty \lambda^{p-1} m(|f| > \lambda)\, d\lambda,$$

égalité valable que les deux membres soient finis ou infinis.

Nous allons décomposer u de manière à faire apparaître des normes de type L^2. Pour ce faire, découpons u en « basses » et « hautes » fréquences en posant

$$u = u_{1,A} + u_{2,A} \quad \text{avec} \quad u_{1,A} = \mathscr{F}^{-1}(\mathbf{1}_{B(0,A)}\widehat{u}) \quad \text{et} \quad u_{2,A} = \mathscr{F}^{-1}(\mathbf{1}_{B^c(0,A)}\widehat{u}). \tag{8.7}$$

Comme la fonction $\mathbf{1}_{B(0,A)}\widehat{u}$ est L^2 et à support compact, elle est dans L^1, donc $u_{1,A}$ est bornée par transformation de Fourier inverse. Plus précisément,

$$\begin{aligned}
\|u_{1,A}\|_{L^\infty} &\leq (2\pi)^{-d}\|\widehat{u_{1,A}}\|_{L^1} \\
&= (2\pi)^{-d}\int_{B(0,A)} |\xi|^{-s}|\xi|^s|\widehat{u}(\xi)|\, d\xi \\
&\leq (2\pi)^{-d}\Big(\int_{B(0,A)} |\xi|^{-2s}\, d\xi\Big)^{\frac{1}{2}}\|u\|_{H^s} \\
&= CA^{\frac{d}{2}-s}\|u\|_{H^s}.
\end{aligned} \tag{8.8}$$

Or, d'après l'inégalité triangulaire, on a

$$\{|u| > \lambda\} \subset \{2|u_{1,A}| > \lambda\} \cup \{2|u_{2,A}| > \lambda\},$$

donc $m(|u| > \lambda) \leq m(2|u_{1,A}| > \lambda) + m(2|u_{2,A}| > \lambda)$. D'après l'inégalité (8.8) ci-dessus, si l'on pose

$$A_\lambda = \Big(\frac{\lambda}{2C\|u\|_{H^s}}\Big)^{\frac{2}{d-2s}} = \Big(\frac{\lambda}{2C\|u\|_{H^s}}\Big)^{\frac{s^*}{d}}$$

alors

$$m\Big(|u_{1,A_\lambda}| > \frac{\lambda}{2}\Big) = 0.$$

On en déduit donc que

$$\|u\|_{L^{s^*}}^{s^*} \leq s^* \int_0^\infty \lambda^{s^*-1} m(2|u_{2,A_\lambda}| > \lambda)\, d\lambda.$$

Il est bien connu (c'est l'inégalité de Bienaimé-Tchebychev) que

$$\begin{aligned} m\big(2|u_{2,A_\lambda}| > \lambda\big) &= \int_{\{2|u_{2,A_\lambda}|>\lambda\}} dx \\ &\leq \int_{\{2|u_{2,A_\lambda}|>\lambda\}} \frac{4|u_{2,A_\lambda}(x)|^2}{\lambda^2}\, dx \\ &= 4\frac{\|u_{2,A_\lambda}\|_{L^2}^2}{\lambda^2}. \end{aligned}$$

Il en résulte donc que l'on a

$$\|u\|_{L^{s^*}}^{s^*} \leq C \int_0^\infty \lambda^{s^*-3} \|u_{2,A_\lambda}\|_{L^2}^2\, d\lambda. \tag{8.9}$$

La transformation de Fourier est (à une constante près) une isométrie de L^2. On a donc

$$\|u_{2,A_\lambda}\|_{L^2}^2 = (2\pi)^{-d} \int_{\{|\xi|\geq A_\lambda\}} |\widehat{u}(\xi)|^2\, d\xi.$$

D'après l'inégalité (8.9), il vient

$$\|u\|_{L^{s^*}}^{s^*} \leq C \int_{\mathbb{R}_+\times\mathbb{R}^d} \lambda^{s^*-3} \mathbf{1}_{\{(\lambda,\xi)/\ |\xi|\geq A_\lambda\}}(\lambda,\xi) |\widehat{u}(\xi)|^2\, d\xi d\lambda.$$

Or, par définition de A_λ, on a

$$|\xi| \geq A_\lambda \Leftrightarrow \lambda \leq C(\xi) \stackrel{\text{déf}}{=} 2C\|u\|_{H^s}|\xi|^{\frac{d}{s^*}}.$$

D'après le théorème de Fubini, on a donc

$$\begin{aligned} \|u\|_{L^{s^*}}^{s^*} &\leq C \int_{\mathbb{R}^d} \left(\int_0^{C(\xi)} \lambda^{s^*-3}\, d\lambda \right) |\widehat{u}(\xi)|^2\, d\xi \\ &\leq C\|u\|_{H^s}^{s^*-2} \int_{\mathbb{R}^d} |\xi|^{\frac{d(s^*-2)}{s^*}} |\widehat{u}(\xi)|^2\, d\xi \leq C\|u\|_{H^s}^{s^*}, \end{aligned}$$

car $\frac{d(s^*-2)}{s^*} = 2s$. □

Remarques. Une façon de deviner (ou de se rappeler) l'exposant de Sobolev

$$s^* = \frac{2d}{d-2s}$$

est l'utilisation d'un argument d'homogénéité. En effet, soit v une fonction sur $\mathbb{R}^d$, désignons par v_λ la fonction $v_\lambda(x) = v(\lambda x)$. On a

$$\|v_\lambda\|_{L^{s^*}} = \lambda^{-\frac{d}{s^*}} \|v\|_{L^{s^*}}.$$

De plus, d'après la proposition 7.1.2, on a

$$\begin{aligned}\int_{\mathbb{R}^d} |\xi|^{2s} |\widehat{v_\lambda}(\xi)|^2 d\xi &= \lambda^{-2d} \int_{\mathbb{R}^d} |\xi|^{2s} |\widehat{v}(\lambda^{-1}\xi)|^2 d\xi \\ &= \lambda^{-d+2s} \|v\|^2_{\dot{H}^s},\end{aligned}$$

en posant

$$\|v\|^2_{\dot{H}^s} \stackrel{\text{déf}}{=} \int_{\mathbb{R}^d} |\xi|^{2s} |\widehat{v}(\xi)|^2 d\xi.$$

Si l'on a une inégalité d'injection continue, les deux quantités $\|\cdot\|^2_{L^p}$ et $\|\cdot\|^2_{\dot{H}^s}$ doivent être comparables pour tout λ. On en déduit que l'on doit avoir $-\frac{2d}{s^*} = -d+2s$. L'exposant s^* est donc le seul pour lequel une telle injection peut avoir lieu. Ceci ne montre évidemment pas qu'elle a bien lieu.

Notons que l'on a $2 \leq \frac{2d}{d-2s} \leq +\infty$, la borne inférieure étant atteinte pour $s = 0$. L'injection de Sobolev permet donc de gagner de l'intégrabilité dès que l'indice de dérivation s est strictement positif.

8.3 Les espaces $H^m(\Omega)$, $H^m_0(\Omega)$ et $H^{-m}(\Omega)$

Dans cette section, on étudie quelques espaces de Sobolev sur des ouverts quelconques de $\mathbb{R}^d$. Dans ce contexte, la transformation de Fourier est moins opérante. On se restreint aux espaces de Sobolev d'indice entier, étant entendu qu'il est possible de définir aussi dans ce cas des espaces de Sobolev d'indice réel quelconque.

Définition 8.3.1 *Soient m un entier naturel et Ω un ouvert de $\mathbb{R}^d$. L'espace $H^m(\Omega)$ est l'espace des fonctions de $L^2(\Omega)$ dont toutes les dérivées jusqu'à l'ordre m sont dans $L^2(\Omega)$, muni de la norme*

$$\|u\|_{H^m(\Omega)} = \Big(\sum_{|\alpha| \leq m} \|\partial^\alpha u\|^2_{L^2(\Omega)} \Big)^{1/2}.$$

L'espace $H_0^m(\Omega)$ est l'adhérence au sens de la norme $H^m(\Omega)$ de $\mathscr{D}(\Omega)$, muni de cette norme. L'espace $H^{-m}(\Omega)$ est l'ensemble des distributions u sur Ω telles que

$$\|u\|_{H^{-m}(\Omega)} \stackrel{déf}{=} \sup_{\substack{\|\varphi\|_{H^m(\Omega)}\leq 1 \\ \varphi\in\mathscr{D}(\Omega)}} |\langle u,\varphi\rangle| < \infty.$$

Proposition 8.3.1 *L'espace $H^m(\Omega)$ muni de sa norme est un espace de Hilbert.*

La démonstration est un exercice facile laissé au lecteur. L'espace $H^{-m}(\Omega)$ s'identifie naturellement au dual de $H_0^m(\Omega)$ grâce au théorème suivant.

Théorème 8.3.1 *L'application bilinéaire définie par*

$$\begin{array}{rcl} B\colon H^{-m}(\Omega)\times\mathscr{D}(\Omega) & \to & \mathbb{K} \\ (u,\varphi) & \mapsto & \langle u,\varphi\rangle \end{array}$$

se prolonge en une application bilinéaire continue de $H^{-m}(\Omega)\times H_0^m(\Omega)$ dans $\mathbb{K}$, toujours notée B. De plus, l'application $\delta_B\colon H^{-m}(\Omega)\to(H_0^m(\Omega))'$ qui lui est associée est un isomorphisme linéaire isométrique, c'est-à-dire que la forme bilinéaire B identifie l'espace $H^{-m}(\Omega)$ au dual de $H_0^m(\Omega)$.

Démonstration. Le fait que l'application bilinéaire B se prolonge n'est rien d'autre que la traduction dans la présente situation du théorème de prolongement 2.2.2. Soit ℓ une forme linéaire continue sur $H_0^m(\Omega)$. Sa restriction à $\mathscr{D}(\Omega)$ est une distribution u sur Ω telle que

$$\forall\,\varphi\in\mathscr{D}(\Omega),\ \langle u,\varphi\rangle = \langle \ell,\varphi\rangle$$

d'où le théorème par définition de la norme sur $(H_0^m(\Omega))'$. □

Donnons quelques propriétés des espaces d'indice négatif. On introduit d'abord trois applications :

– Le prolongement par 0 en dehors de Ω, qui réalise une injection isométrique de $H_0^m(\Omega)$ dans $H^m(\mathbb{R}^d)$. En effet, cette propriété est trivialement vraie sur $\mathscr{D}(\Omega)$, donc elle le reste par densité. L'espace $H_0^m(\Omega)$ peut donc être considéré comme un sous-espace vectoriel fermé de $H^m(\mathbb{R}^d)$. Notons $H_0^m(\Omega)^\perp$ son orthogonal.

– L'application linéaire pr de $H^{-m}(\Omega)$ dans $H^{-m}(\mathbb{R}^d) = H^m(\mathbb{R}^d)'$ par

$$\langle pr(u), f\rangle = \langle u, f_0\rangle, \tag{8.10}$$

où l'on a décomposé $H^m(\mathbb{R}^d)$ en somme orthogonale : $f = f_0 + f_\perp$ avec $f\in H_0^m(\Omega)$ et $f_\perp\in H_0^m(\Omega)^\perp$

– Enfin, on note r est la restriction des distributions de $\mathbb{R}^d$ sur Ω.

Nous allons démontrer le petit lemme suivant.

Lemme 8.3.1 *L'application pr est une application linéaire isométrique de $H^{-m}(\Omega)$ dans $H^{-m}(\mathbb{R}^d)$ et l'application r est une application linéaire continue de $H^{-m}(\mathbb{R}^d)$ dans $H^{-m}(\Omega)$. De plus, on a*

$$r \circ pr = \mathrm{Id}_{H^{-m}(\Omega)} .$$

Démonstration. D'après le théorème 8.1.2, on sait que

$$\|pr(u)\|_{H^{-m}(\mathbb{R}^d)} = \sup_{\|\varphi\|_{H^m(\mathbb{R}^d)} \leq 1} |\langle pr(u), \varphi \rangle|.$$

La projection orthogonale sur un sous-espace fermé d'un espace de Hilbert étant une application linéaire de norme 1, on peut écrire, vu la définition de pr,

$$\begin{aligned} \|pr(u)\|_{H^{-m}(\mathbb{R}^d)} &= \sup_{\|\varphi\|_{H^m(\mathbb{R}^d)} \leq 1} |\langle pr(u), \varphi \rangle| \\ &= \sup_{\|\varphi\|_{H^m(\Omega)} \leq 1} |\langle u, \varphi \rangle| \\ &= \|u\|_{H^{-m}(\Omega)}. \end{aligned} \tag{8.11}$$

Par ailleurs, on a, pour toute distribution v de $H^{-m}(\mathbb{R}^d)$,

$$\begin{aligned} \sup_{\substack{\|\varphi\|_{H^m(\Omega)} \leq 1 \\ \varphi \in \mathscr{D}(\Omega)}} |\langle r(v), \varphi \rangle| &= \sup_{\substack{\|\varphi\|_{H^m(\Omega)} \leq 1 \\ \varphi \in \mathscr{D}(\Omega)}} |\langle v, \varphi \rangle| \\ &\leq \sup_{\substack{\|\varphi\|_{H^m(\mathbb{R}^d)} \leq 1 \\ \varphi \in \mathscr{D}(\mathbb{R}^d)}} |\langle v, \varphi \rangle| \\ &= \|v\|_{H^{-m}(\mathbb{R}^d)}. \end{aligned}$$

On voit donc que $r(v) \in H^{-m}(\Omega)$ et que l'application r est continue de $H^{-m}(\mathbb{R}^d)$ dans $H^{-m}(\Omega)$.

Enfin, il est clair que l'on a, pour toute fonction $f \in H_0^m(\Omega)$,

$$\begin{aligned} \langle r \circ pr(u), f \rangle &= \langle pr(u), f \rangle \\ &= \langle u, f \rangle, \end{aligned}$$

d'où le lemme. □

Théorème 8.3.2 i) *L'espace $H^{-m}(\Omega)$ est l'ensemble des restrictions à Ω des éléments de $H^{-m}(\mathbb{R}^d)$,*

ii) *La norme $\|\cdot\|_{H^{-m}(\Omega)}$ définie ci-dessus est une norme hilbertienne.*

iii) *L'espace $\mathscr{D}(\Omega)$ est dense dans l'espace $H^{-m}(\Omega)$.*

Démonstration. Le premier point est clairement assuré par le lemme. Quant au deuxième point, il suffit de remarquer que $H^{-m}(\Omega)$ est isométrique au dual de l'espace de Hilbert $H_0^m(\Omega)$.

Démontrons maintenant que $\mathscr{D}(\Omega)$ est dense dans $H^{-m}(\Omega)$. Pour ce faire, considérons une distribution u dans $H^{-m}(\Omega)$. Le théorème 8.1.1 dit en particulier que $\mathscr{D}(\mathbb{R}^d)$ est dense dans $H^{-m}(\mathbb{R}^d)$. Il existe donc une suite $(\psi_n)_{n\in\mathbb{N}}$ d'éléments de $\mathscr{D}(\mathbb{R}^d)$ telle que

$$\lim_{n\to\infty} \psi_n = pr(u) \quad \text{dans} \quad H^{-m}(\mathbb{R}^d).$$

D'après le lemme précédent, ceci implique que

$$\lim_{n\to\infty} r(\psi_n) = u \quad \text{dans} \quad H^{-m}(\Omega).$$

Mais, la restriction à Ω d'une fonction de $L^2(\mathbb{R}^d)$ étant une fonction de $L^2(\Omega)$, et $\mathscr{D}(\Omega)$ étant dense dans $L^2(\Omega)$, il existe une suite de fonctions $(\varphi_n)_{n\in\mathbb{N}}$ de $\mathscr{D}(\Omega)$ telle que, pour tout entier n, on ait

$$\|r(\psi_n) - \varphi_n\|_{L^2(\Omega)} \leq \frac{1}{n+1}\cdot$$

Comme la norme $L^2(\Omega)$ est plus grande que la norme $H^{-m}(\Omega)$, il est clair que $\varphi_n \to u$ au sens de la norme de $H^{-m}(\Omega)$. □

Montrons maintenant l'inégalité de Poincaré, qui est un résultat très important.

Théorème 8.3.3 *Soit Ω un ouvert de $\mathbb{R}^d$ borné dans une direction. Il existe une constante C telle que*

$$\forall u \in H_0^1(\Omega),\ \|u\|_{L^2} \leq C\Big(\sum_{j=1}^d \|\partial_j u\|_{L^2}^2\Big)^{\frac{1}{2}}.$$

Démonstration. Supposons d'abord que Ω soit borné dans la direction x_1. Soit R un réel strictement positif tel que Ω soit inclus dans $]-R,R[\times\mathbb{R}^{d-1}$. On a alors, pour toute fonction-test φ,

$$\varphi(x_1,\cdots,x_d) = \int_{-R}^{x_1} \frac{\partial\varphi}{\partial x_1}(s,x_2,\cdots,x_d)\,ds.$$

L'inégalité de Cauchy-Schwarz implique que pour tout $x\in\Omega$,

$$|\varphi(x_1,\cdots,x_d)|^2 \leq 2R\int_{-R}^{R}\Big|\frac{\partial\varphi}{\partial x_1}(s,x_2,\cdots,x_d)\Big|^2\,ds.$$

En intégrant cette inégalité sur Ω, il vient

$$\int_\Omega |\varphi(x_1,\cdots,x_d)|^2\,dx \leq 2R\int_\Omega\int_{-R}^{R}\Big|\frac{\partial\varphi}{\partial x_1}(s,x_2,\cdots,x_d)\Big|^2\,dsdx = 4R^2\Big\|\frac{\partial\varphi}{\partial x_1}\Big\|^2_{L^2(\Omega)},$$

par Fubini. Comme $\mathscr{D}(\Omega)$ est dense dans $H_0^1(\Omega)$, le théorème est démontré dans le cas où l'ouvert est borné dans la direction x_1. Le cas général s'en déduit par rotation. □

L'inégalité de Poincaré implique de manière évidente le corollaire suivant.

Corollaire 8.3.1 *Soit Ω un ouvert de $\mathbb{R}^d$ borné dans une direction. L'application*

$$u \mapsto \Big(\sum_{|\alpha|=m}\|\partial^\alpha u\|^2_{L^2}\Big)^{\frac{1}{2}}$$

est une norme hilbertienne sur $H_0^m(\Omega)$ équivalente à la norme de $H^m(\Omega)$.

Dans le cas d'un ouvert borné dans une direction, on a donc le choix d'utiliser la norme H^m entière ou bien la semi-norme ci-dessus, et il en va de même pour calculer les normes duales sur H^{-m}.

Pour conclure ce chapitre sur les espaces de Sobolev, nous allons énoncer et admettre le théorème de Rellich, un résultat de compacité très important.

Théorème 8.3.4 *Soit Ω un ouvert borné de $\mathbb{R}^d$ avec $d \geq 2$. Si p est un réel strictement inférieur à*

$$2^* = \frac{2d}{d-2},$$

alors l'injection de $H_0^1(\Omega)$ dans $L^p(\Omega)$ est compacte.

Dans la pratique, on utilise souvent le corollaire suivant.

Corollaire 8.3.2 *Soit $(u_n)_{n\in\mathbb{N}}$ une suite bornée de fonctions de $H_0^1(\Omega)$, il existe une sous-suite n_l et une fonction u de $H_0^1(\Omega)$ telles que*

$$\forall\, 1 \leq p < 2^*,\ \lim_{l\to\infty}\|u_{n_l} - u\|_{L^p(\Omega)} = 0.$$

Chapitre 9

Deux exemples de résolution d'EDP

Le but de ce chapitre conclusif est de donner deux exemples d'applications de la notion de dualité à la résolution d'équations aux dérivées partielles. Le choix de ces applications est arbitraire, à ceci près qu'elles sont d'énoncé simple.

9.1 Résolution du problème de Dirichlet

Dans toute cette section, Ω désigne un ouvert borné de $\mathbb{R}^d$ et les distributions considérées seront toutes supposées à valeurs réelles. On cherche à résoudre le problème de Dirichlet pour l'équation de Poisson, c'est-à-dire qu'étant donné une fonction f dans $L^2(\Omega)$, ou mieux une distribution dans $H^{-1}(\Omega)$, on cherche une fonction u telle que

$$-\Delta u = f$$

au sens des distributions. Il est clair qu'il faut mettre une condition supplémentaire si l'on veut avoir l'unicité. En effet, toute fonction constante est solution de $\Delta u = 0$. Énonçons le théorème précis.

Théorème 9.1.1 *Pour toute distribution f appartenant à $H^{-1}(\Omega)$, il existe une unique fonction u dans l'espace $H_0^1(\Omega)$ telle que*

$$-\Delta u = f.$$

Démonstration. Commençons par démontrer l'unicité. Soit u une fonction de $H_0^1(\Omega)$ solution de

$$-\Delta u = 0.$$

On en déduit que, pour toute fonction $\varphi \in \mathscr{D}(\Omega)$, on a

$$\langle -\Delta u, \varphi \rangle = 0.$$

Par définition de la dérivée des distributions, ceci s'écrit

$$\sum_{j=1}^{d} \langle \partial_j u, \partial_j \varphi \rangle = 0.$$

Comme u appartient à $H_0^1(\Omega)$, ses dérivées appartiennent à $L^2(\Omega)$. Il en résulte que

$$\int_\Omega \nabla u(x) \cdot \nabla \varphi(x)\, dx = 0. \tag{9.1}$$

L'application

$$h \mapsto \int_\Omega \nabla u(x) \cdot \nabla h(x)\, dx$$

est une application linéaire continue sur $H_0^1(\Omega)$, donc l'identité (9.1) ci-dessus est vraie pour toute fonction h de l'espace $H_0^1(\Omega)$, donc en particulier pour $h = u$. On en déduit que $\nabla u = 0$, et donc d'après l'inégalité de Poincaré, que $u = 0$, ce qui établit l'unicité.

L'existence utilise de manière cruciale le concept de dualité. Soit f un élément de $H^{-1}(\Omega)$. Le théorème 8.3.1 affirme que la forme linéaire définie par

$$\varphi \to \langle f, \varphi \rangle$$

est une forme linéaire continue sur $H_0^1(\Omega)$. D'autre part, l'inégalité de Poincaré implique que l'application définie de $H_0^1(\Omega)$ dans $\mathbb{R}^+$ par

$$u \to \int_\Omega |\nabla u|^2\, dx$$

est une norme hilbertienne sur $H_0^1(\Omega)$. Le théorème 4.3.1 de représentation de Riesz assure qu'il existe une fonction u de $H_0^1(\Omega)$ telle que, pour toute fonction v de $H_0^1(\Omega)$, on ait

$$\langle f, v \rangle = \int_\Omega \nabla u(x) \cdot \nabla v(x)\, dx.$$

Cette relation est en particulier vraie pour toute fonction $v = \varphi$ appartenant à $\mathscr{D}(\Omega)$. Par définition de la dérivée des distributions, on a donc

$$\langle f, v \rangle = \int_\Omega \nabla u(x) \cdot \nabla \varphi(x) dx = \sum_{j=1}^{d} \langle \frac{\partial u}{\partial x_j}, \frac{\partial \varphi}{\partial x_j} \rangle = \sum_{j=1}^{d} -\langle \frac{\partial^2 u}{\partial x_j^2}, \varphi \rangle = -\langle \Delta u, \varphi \rangle,$$

c'est-à-dire précisément $-\Delta u = f$ au sens des distributions. □

Remarquons que ce théorème donne une description des éléments de $H^{-1}(\Omega)$. Soit f une distribution de $H^{-1}(\Omega)$, alors f est somme de dérivées secondes (au sens des distributions) de fonctions de $H_0^1(\Omega)$ et aussi une somme de dérivées premières de fonctions de $H^0(\Omega) = L^2(\Omega)$.

On aurait aussi pu utiliser le théorème de Lax-Milgram au lieu du théorème de représentation de Riesz, mais ce n'était guère utile dans ce cas particulier.

9.2 Résolubilité locale des opérateurs elliptiques

Définition 9.2.1 *On appelle opérateur différentiel (linéaire) d'ordre m sur un ouvert Ω de $\mathbb{R}^d$ tout opérateur P de $\mathcal{D}'(\Omega)$ dans lui-même de la forme*

$$Pu = \sum_{|\alpha|\leq m} a_\alpha(x)\partial^\alpha u$$

où $(a_\alpha)_{|\alpha|\leq m}$ est une famille de fonctions indéfiniment différentiables sur Ω.

A tout opérateur différentiel P, on associe un polynôme sur $\mathbb{C}^d$ par

$$P(\xi) = \sum_{|\alpha|\leq m} a_\alpha(x)\xi^\alpha$$

à x fixé. On ne distinguera pas l'opérateur et le polynôme dans la notation.

Nous allons maintenant définir la notion d'opérateur elliptique.

Définition 9.2.2 *Soit P un opérateur différentiel d'ordre m sur Ω ; on dit que P est un opérateur elliptique si et seulement si, pour tout point x_0 de Ω, il existe une constante c strictement positive et un voisinage V de x_0 tel que*

$$\forall x \in V, \forall \xi \in \mathbb{R}^d, \left| \sum_{|\alpha|=m} a_\alpha(x)\xi^\alpha \right| \geq c|\xi|^m.$$

Exemples Soit $A(x) = (a_{ij}(x))_{1\leq i,j\leq d}$ une fonction indéfiniment différentiable à valeurs dans les matrices réelles symétriques définies positives. L'opérateur définie par

$$P = \sum_{ij} a_{ij}(x)\partial_i\partial_j$$

est elliptique d'ordre 2. En particulier, le laplacien est un opérateur elliptique.

L'objet de cette section est la démonstration du théorème suivant.

Théorème 9.2.1 *Soient Ω un ouvert de $\mathbb{R}^d$ et P un opérateur elliptique d'ordre m sur Ω. Pour tout point x_0 de Ω, il existe un réel R strictement positif tel que, pour toute distribution f dans $H^{-m}(B(x_0,R))$, il existe une fonction u de $L^2(B(x_0,R))$ telle que $Pu = f$.*

Ce théorème dit bien que l'opérateur P est localement résoluble, c'est-à-dire qu'en chaque point, on peut trouver une boule centrée en ce point sur laquelle on peut résoudre l'équation aux dérivées partielles $Pu = f$ pour une distribution f pas trop irrégulière. La démonstration de ce théorème se décompose en deux parties.

Dans la première partie, on démontre une inégalité relative aux fonctions de $\mathscr{D}(\Omega)$, dite « estimation *a priori* ». Dans la seconde partie, on utilise le concept de dualité pour déduire de l'inégalité précédemment démontrée l'existence de la solution.

Commençons par l'estimation *a priori*.

Lemme 9.2.1 *Soient P un opérateur elliptique d'ordre m sur Ω et x_0 un point de Ω. Il existe deux réels C et R strictement positifs tels que*

$$\forall \varphi \in \mathscr{D}(B(x_0,R)), \ \|\varphi\|_{H^m} \leq C\|P\varphi\|_{L^2}.$$

Démonstration. La preuve de ce lemme est assez longue. On démontre tout d'abord l'inégalité pour l'opérateur à coefficients constants

$$P_0 = \sum_{|\alpha|\leq m} a_\alpha(x_0)\partial^\alpha;$$

puis on utilise une méthode dite de pertubation qui permet de conclure.

Il s'agit donc de majorer, pour φ appartenant à $\mathscr{D}(\Omega)$, la quantité

$$\|\varphi\|^2_{H^m} = \int_{\mathbb{R}^d}(1+|\xi|^2)^m|\widehat{\varphi}(\xi)|^2\,d\xi.$$

L'opérateur P étant supposé elliptique, il existe un réel c tel que l'on puisse écrire pour tout $\xi \in \mathbb{R}^d$

$$\begin{aligned}
|\xi|^{2m} &\leq c^2\Big|\sum_{|\alpha|=m} a_\alpha(x_0)\xi^\alpha\Big|^2\\
&= c^2\Big|i^m\sum_{|\alpha|=m} a_\alpha(x_0)\xi^\alpha\Big|^2\\
&= c^2\Big|\sum_{|\alpha|\leq m} a_\alpha(x_0)(i\xi)^\alpha - \sum_{|\alpha|\leq m-1} a_\alpha(x_0)(i\xi)^\alpha\Big|^2\\
&\leq 2c^2\Big|\sum_{|\alpha|\leq m} a_\alpha(x_0)(i\xi)^\alpha\Big|^2 + 2c^2\Big|\sum_{|\alpha|\leq m-1} a_\alpha(x_0)(i\xi)^\alpha\Big|^2\\
&\leq 2c^2|P_0(i\xi)|^2 + C_0(1+|\xi|^{2(m-1)}),
\end{aligned}$$

en utilisant l'inégalité injustement méprisée bien que cruciale, $(A+B)^2 \leq 2A^2 + 2B^2$. Comme $2m > 2(m-1)$, il existe une constante M telle que si $|\xi| \geq M$, on ait alors

$$C_0(1+|\xi|^{2(m-1)}) \leq \frac{1}{2}|\xi|^{2m}.$$

Il en résulte que

$$|\xi| \geq M \Rightarrow |\xi|^{2m} \leq 4c^2|P_0(i\xi)|^2.$$

On note alors que

$$\mathscr{F}(P_0\varphi)(\xi) = P_0(i\xi)\widehat{\varphi}(\xi).$$

Ainsi, on a

$$\begin{aligned}\int_{\mathbb{R}^d}(1+|\xi|^2)^m|\widehat{\varphi}(\xi)|^2\,d\xi \le C_0'\int_{|\xi|\ge M}(|\mathscr{F}(P_0\varphi)(\xi)|^2+|\widehat{\varphi}(\xi)|^2)\,d\xi&\\+\int_{|\xi|\le M}(1+|\xi|^2)^m|\widehat{\varphi}(\xi)|^2\,d\xi&,\end{aligned}$$

d'où

$$\|\varphi\|_{H^m}^2 \le C_1\|P_0\varphi\|_{L^2}^2 + C_2\|\varphi\|_{L^2}^2.$$

Or, l'inégalité de Poincaré démontrée dans le théorème 8.3.3 affirme que, pour toute fonction φ appartenant à $\mathscr{D}(B(x_0,R))$,

$$\|\varphi\|_{L^2}^2 \le CR\|\nabla\varphi\|_{L^2}^2.$$

Comme $\|\nabla\varphi\|_{L^2}^2 \le \|\varphi\|_{H^m}^2$, on a, pour toute fonction φ appartenant à $\mathscr{D}(B(x_0,R))$,

$$\|\varphi\|_{H^m}^2 \le C_1\|P_0\varphi\|_{L^2}^2 + C_2R\|\varphi\|_{H^m}^2$$

En prenant $R \le (2C_2)^{-1}$, on assure l'existence d'une constante C telle que

$$\forall\,\varphi \in \mathscr{D}(B(x_0,R))\,,\ \|\varphi\|_{H^m} \le C_3\|P_0\varphi\|_{L^2}.$$

Nous allons maintenant mettre en œuvre un argument de perturbation. D'après l'inégalité ci-dessus, on peut écrire

$$\begin{aligned}\|\varphi\|_{H^m} &\le C_3\|P_0\varphi\|_{L^2}\\ &\le C_3\|P\varphi\|_{L^2} + C_3\|(P_0-P)\varphi\|_{L^2}.\end{aligned}$$

Or, par définition de P_0, on a

$$(P_0-P)\varphi(x) = \sum_{|\alpha|\le m}(a_\alpha(x_0)-a_\alpha(x))\partial^\alpha\varphi(x).$$

Ainsi donc

$$\begin{aligned}\|(P_0-P)\varphi\|_{L^2} &\le C_5\Big(\sup_{\substack{x\in B(x_0,R)\\|\alpha|\le m}}\|\partial^\alpha\varphi\|_{L^2}|a_\alpha(x_0)-a_\alpha(x)|\Big)\\ &\le C_6\|\varphi\|_{H^m}\sup_{\substack{x\in B(x_0,R)\\|\alpha|\le m}}|a_\alpha(x_0)-a_\alpha(x)|.\end{aligned}$$

Les fonctions a_α étant continues, on peut imposer à R d'être suffisamment petit pour que

$$C_3C_6 \sup_{\substack{x\in B(x_0,R)\\ |\alpha|\leq m}} |a_\alpha(x_0)-a_\alpha(x)| \leq \frac{1}{2}\cdot$$

Il vient alors

$$\|\varphi\|_{H^m} \leq C_3\|P\varphi\|_{L^2} + \frac{1}{2}\|\varphi\|_{H^m}.$$

Le lemme est démontré. □

On introduit l'adjoint formel de l'opérateur P

$$P^*u = \sum_{|\alpha|\leq m} (-\partial)^\alpha(a_\alpha u).$$

Corollaire 9.2.1 *Soient P un opérateur elliptique d'ordre m sur Ω et x_0 un point de Ω. Il existe deux réels C et R strictement positifs tels que*

$$\forall\, v \in H^m_0(B(x_0,R)), \|v\|_{H^m} \leq C\|P^*v\|_{L^2}.$$

Démonstration. Il suffit d'observer que si P est elliptique, alors P^* l'est aussi car, d'après la formule de Leibniz,

$$P^*u = \sum_{|\alpha|=m} a_\alpha(-\partial)^\alpha u + \sum_{|\beta|\leq m-1} b_\beta \partial^\beta u,$$

où les fonctions b_β sont indéfiniment différentiables, et l'on conclut par densité de $\mathscr{D}(B(x_0,R))$ dans $H^m_0(B(x_0,R))$. □

La seconde partie de la démonstration utilise de manière cruciale le concept de dualité, cœur de la théorie des distributions. Voici comment. Considérons l'espace E défini par

$$E = \{P^*v, v \in H^m_0(B(x_0,R))\}.$$

Lemme 9.2.2 *L'espace E est fermé dans $L^2(B(x_0,R))$*

Démonstration. Soit $(g_n)_{n\in\mathbb{N}}$ une suite de fonctions de E telle que, pour un certain $g \in L^2(B(x_0,R))$, on ait

$$\|g_n - g\|_{L^2(B(x_0,R))} \to 0.$$

Par définition de l'espace E, il existe une suite $(v_n)_{n\in\mathbb{N}}$ de fonctions de $H^m_0(B(x_0,R))$ telles que

$$g_n = P^*v_n.$$

D'après le corollaire 9.2.1, on a

$$\|v_p - v_q\|_{H^m(B(x_0,R))} \leq C\|P^*(v_p-v_q)\|_{L^2(B(x_0,R))} = C\|g_p - g_q\|_{L^2(B(x_0,R))}.$$

Comme la suite $(g_n)_{n\in\mathbb{N}}$ est convergente dans $L^2(B(x_0,R))$, elle est de Cauchy. Il résulte de l'inégalité ci-dessus que la suite $(v_n)_{n\in\mathbb{N}}$ est également une suite de Cauchy dans l'espace $H^m(B(x_0,R))$. Donc, il existe une fonction v appartenant à $H_0^m(B(x_0,R))$ telle que l'on ait

$$\lim_{n\to\infty} v_n = v \quad \text{dans} \quad H^m(B(x_0,R)).$$

L'opérateur P^* est un opérateur continu de $H^m(B(x_0,R))$ dans $L^2(B(x_0,R))$. Donc on a

$$g = P^* v,$$

d'où $g \in E$. □

Démonstration du théorème 9.2.1. Nous allons maintenant construire la solution u par le biais d'une forme linéaire continue sur E. Soit f une distribution de $H^{-m}(B(x_0,R))$; on considère la forme linéaire L définie par

$$\begin{aligned} L\colon E &\longrightarrow \mathbb{R} \\ g = P^* v &\longmapsto \langle f, v\rangle. \end{aligned}$$

D'après le corollaire 9.2.1, la fonction v de $H_0^m(B(x_0,R))$ telle que $P^*v = g$ est unique et la forme linéaire L bien définie.

De plus, on a, pour tout $v \in H_0^m(B(x_0,R))$,

$$\begin{aligned} |Lg| = |\langle f, v\rangle| &\leq \|f\|_{H^{-m}(B(x_0,R))}\|v\|_{H^m(B(x_0,R))} \\ &\leq C\|f\|_{H^{-m}(B(x_0,R))}\|P^*v\|_{L^2(B(x_0,R))} \\ &= C\|f\|_{H^{-m}(B(x_0,R))}\|g\|_{L^2(B(x_0,R))}. \end{aligned}$$

Donc la forme linéaire L est continue sur E. Soit $\bar{L}$ la forme linéaire définie sur l'espace $L^2(B(x_0,R))$ par

$$\bar{L}_{|E} = L \quad \text{et} \quad \bar{L}_{|E^\perp} = 0.$$

La forme linéaire $\bar{L}$ est continue sur $L^2(B(x_0,R))$. D'après le théorème de représentation de Riesz, il existe une fonction u de $L^2(B(x_0,R))$ telle que, pour toute fonction g de $L^2(B(x_0,R))$, on ait

$$\bar{L}(g) = \int_{B(x_0,R)} g(x)u(x)\,dx.$$

Cette relation est en particulier vraie pour les fonctions g appartenant à E. Donc, pour toute fonction φ appartenant à $\mathscr{D}(B(x_0,R))$, on a

$$\int_{B(x_0,R)} (P^*\varphi)(x)u(x)\,dx = \langle f, \varphi\rangle.$$

Or, par définition de P^*, on a

$$\int_{B(x_0,R)} (P^*\varphi)(x)u(x)\,dx = \langle Pu, \varphi \rangle$$

et donc

$$\forall\, \varphi \in \mathscr{D}(B(x_0,R)), \quad \langle Pu, \varphi \rangle = \langle f, \varphi \rangle,$$

c'est-à-dire que $Pu = f$ au sens des distributions. □

Remarque. On note que E est en général un sous-espace strict de $L^2(B(x_0,R))$. En effet, dans le cas $m = 2$, $P = P^* = \Delta$, on peut montrer que P est un isomorphisme entre $H^2(B(x_0,R)) \cap H^1_0(B(x_0,R))$ et $L^2(B(x_0,R))$ alors que $E = P(H^2_0(B(x_0,R)))$.