Analyse II

Cours de deuxième année

donné à l'Ecole normale supérieure de Lyon

année universitaire 2003-2004

Cédric Villani

Unité de Mathématiques Pures et Appliquées

Ecole normale supérieure de Lyon

46 allée d'Italie

69364 Lyon Cedex 07

`cvillani@umpa.ens-lyon.fr`

Table des matières

CHAPITRE I

Panorama naïf des espaces fonctionnels

[Mettre une référence pour le thm de Riesz]

La majeure partie de l'analyse consiste à étudier les propriétés des fonctions en utilisant des méthodes ayant à voir, de près ou de loin, à des techniques d'approximation (limites, topologie, etc.).

Une manière parfois efficace d'aborder certains problèmes d'analyse consiste à raisonner non pas sur des fonctions isolément, mais sur des ensembles entiers de fonctions vérifiant certaines propriétés de nature géométrique ou topologique, appelés **espaces fonctionnels**. L'étude des propriétés de ces espaces est appelée analyse fonctionnelle. Le plus souvent, on étudie des fonctions à valeurs dans un espace vectoriel ($\mathbb{R}$, $\mathbb{C}$, $\mathbb{C}^n$, etc.), constituant des espaces fonctionnels qui sont des **espaces vectoriels**. Dans ce cours, on se préoccupe presque exclusivement de fonctions à valeurs réelles.

De plus, un espace fonctionnel n'a d'intérêt qu'une fois muni d'une topologie, qui en fait un **espace vectoriel topologique**. En outre, le concept de limite étant omniprésent en analyse, on se restreint d'ordinaire à des espaces **complets**.

On peut sans doute considérer Fourier comme le fondateur de l'analyse fonctionnelle ; cependant l'analyse fonctionnelle moderne commence entre 1900 et 1920 avec Volterra, Fredholm, Hilbert, Fréchet, et surtout F. Riesz et Banach, dont les travaux constituent encore les résultats fondamentaux du domaine. Au début des années 1950, un important effort d'abstraction et de généralisation a été effectué par les membres de Bourbaki (en particulier Dieudonné et Schwartz), qui ont en partie remodelé la discipline. Même si cette tendance à l'abstraction est pour l'essentiel passée de mode, il est bon d'être familier avec certains des concepts développés à cette occasion.

Ce chapitre contient un tour d'horizon sommaire des espaces fonctionnels les plus utilisés et de quelques-unes de leurs propriétés. Dans les deux premières sections, on passera en revue les structures de base et les résultats les plus simples de la théorie classique de l'analyse fonctionnelle. Dans beaucoup de cas, il s'agira d'un simple recensement, illustré d'exemples ; on renverra à d'autres ouvrages pour un développement plus complet de la théorie. Des exemples d'espaces fonctionnels populaires seront utilisés pour illustrer les notions abstraites ; dans un deuxième temps, on reviendra sur ces exemples de manière un peu plus systématique.

On suppose le lecteur déjà familier avec le concept d'espace de Banach (espace vectoriel normé complet) et avec ses principales propriétés. L'expérience montre que, face à une situation pratique qui ne relève pas a priori du cadre des espaces de Banach, on peut presque toujours se ramener à de tels espaces.

Sommaire

I-1. Espaces vectoriels topologiques

1.1. Généralités.

Définition I-1 (espace vectoriel topologique). *On appelle $\mathbb{R}$-espace vectoriel topologique (e.v.t.) un espace topologique E muni d'une structure d'espace vectoriel sur $\mathbb{R}$, telle que les opérations $(x,y) \mapsto x+y$ et $(\lambda,x) \mapsto \lambda x$ soient continues $E \times E \to E$ et $\mathbb{R} \times E \to E$, et tel que $\{0\}$ soit une partie fermée.*

Remarques I-2. (i) Dans la définition d'un espace vectoriel topologique, on n'impose pas toujours que $\{0\}$ soit fermé, mais sans cette hypothèse on ne peut guère dire quoi que ce soit d'intéressant sur un espace vectoriel topologique.

(ii) Dans un espace topologique, les translations $\tau_x : y \mapsto x+y$ sont des bijections bicontinues ; en particulier, imposer la fermeture de $\{0\}$ revient à imposer la fermeture des singletons ; et les propriétés topologiques vraies au voisinage de 0 sont automatiquement vraies au voisinage de n'importe quel point de E.

Exemple I-3. Un espace vectoriel normé est un cas particulier d'espace vectoriel topologique. Rappelons qu'une norme sur un espace vectoriel E est une application $N : E \to \mathbb{R}_+$, telle que (i) $N(x+y) \leq N(x) + N(y)$, (ii) $N(\lambda x) = |\lambda| N(x)$, (iii) $N(x) = 0 \Longleftrightarrow x = 0$. Une norme définit automatiquement une distance d sur E par la formule $d(x,y) = N(y-x)$, et fait donc de E un espace topologique, qui satisfait aux axiomes de la Définition I-1.

Une introduction assez complète à la théorie des espaces vectoriels topologiques se trouve dans [Rudin, chapitre 1], accompagnée de nombreuses références. On distingue en pratique diverses catégories d'espaces vectoriels topologiques. Un rôle important dans cette classification est tenu par les **bases de voisinages**.

Définition I-4 (base de voisinages). *Soit E un espace vectoriel topologique. On appelle voisinage de 0 une partie contenant un ouvert contenant 0. On dit qu'une famille $\mathcal{B}$ est une base de voisinages, ou base de voisinages en 0, si c'est une famille de voisinages de 0 telle que pour tout ouvert O contenant 0 on peut trouver $B \in \mathcal{B}$ avec $B \subset O$.*

Exemple I-5. Dans un espace vectoriel normé, les boules de rayon $1/n$, centrées en 0, constituent une base de voisinages.

Pour simplifier, on distinguera quatre niveaux de structure :

– **Les espaces vectoriels topologiques abstraits**, sans aucune autre hypothèse de structure. A un tel niveau de généralité, on pourrait croire que l'on ne peut pas en dire grand-chose ; pourtant les espaces vectoriels topologiques abstraits vérifient plusieurs propriétés intéressantes, voir [Rudin, pp. 10-14] En particulier, on montre qu'un espace vectoriel topologique est automatiquement **séparé** [Rudin, Théorème 1.12], i.e. deux points distincts admettent des voisinages distincts. L'intérêt de ce résultat est cependant relatif, car la propriété de séparation est d'ordinaire très facile à vérifier directement.

– **Les espaces vectoriels topologiques localement convexes, ou e.v.t.l.c.s.** (le “s” signifiant “séparé” et non “séparable” !). Ce sont les espaces vectoriels topologiques dans lesquels 0 admet une base de voisinages convexes.

- **Les espaces de Fréchet**. Ce sont les e.v.t.l.c.s. munis d'une métrique d complète (toute suite de Cauchy est convergente), invariante par translation ($d(x+z, y+z) = d(x,y)$). Il est clair qu'un espace de Fréchet admet une base dénombrable de voisinages convexes : les boules de rayon $1/n$ centrées en 0. Réciproquement, si un e.v.t.l.c.s. admet une base dénombrable de voisinages, on peut le munir d'une métrique compatible avec sa topologie, invariante par translation, pour lesquelles les boules ouvertes sont convexes [Rudin, Théorème 1.24].
- **Les espaces de Banach**. Ce sont les espaces vectoriels normés complets. Si E est un espace de Banach, c'est en particulier un espace de Fréchet pour la distance associée à sa norme.

Remarque I-6. Si E est un espace de Fréchet pour une distance d, alors l'application $N : x \longmapsto d(0,x)$ vérifie deux des trois axiomes d'une norme : en effet,

$$N(x) = 0 \Longleftrightarrow x = 0;$$

$$N(x+y) = d(0, x+y) = d(-x, y) \leq d(-x, 0) + d(0, y) = N(-x) + N(y);$$

mais l'identité $N(\lambda x) = |\lambda| N(x)$ n'est en général pas vraie pour tout $\lambda \in \mathbb{R}$.

La théorie des espaces de Banach est extrêmement développée, et leurs propriétés (théorème du graphe fermé, théorème de Banach-Steinhaus, caractérisation des applications linéaires continues, dualité, existence de topologies faibles, etc.) sont bien connues et développées dans de nombreux ouvrages populaires [Rudin, Chapitre 3], [Rudin-RCA, Chapitre 5], [Edwards], [Brézis], [Dunford-Schwartz]. Les espaces de Banach sont le cadre fonctionnel naturel dans lequel se fait la majeure partie de l'analyse fonctionnelle. Parmi les exemples les plus simples, citons

- l'espace $L^p(\Omega)$ des fonctions L^p sur un espace mesuré (Ω, λ) (Ω peut être un ouvert de $\mathbb{R}^n$ et λ la mesure de Lebesgue), modulo l'identification des fonctions égales presque partout, muni de la norme

$$\|f\|_{L^p(\Omega)} := \left(\int_\Omega |f(x)|^p \, d\lambda(x) \right)^{1/p};$$

- l'espace $C_b(\Omega)$ des fonctions continues bornées sur un espace topologique Ω, muni de la norme

$$\|f\|_{C_b(\Omega)} = \sup_{x \in \Omega} |f(x)|;$$

etc. etc.

On passera d'autres exemples en revue plus loin. Les remarquables techniques d'interpolation qui seront discutées au Chapitre II sont un bon exemple des méthodes que l'on peut développer dans le contexte des espaces de Banach.

Il arrive cependant assez souvent que l'on rencontre des espaces de Fréchet, par exemple comme limites ou intersections d'espaces de Banach. Deux exemples naturels sont

- l'espace $L^p_{\mathrm{loc}}(\Omega)$ des fonctions qui appartiennent à $L^p(K)$ pour tout compact $K \subset \Omega$: on dit que de telles fonctions sont *localement L^p-intégrables* ;
- l'espace des fonctions continues sur Ω, non nécessairement bornées.

Beaucoup des résultats fondamentaux de la théorie des espaces de Banach sont encore valables dans le cadre des espaces de Fréchet.

Dans la suite de cette section, nous allons résumer brièvement les principales notions utiles en théorie des espaces vectoriels topologiques, puis détailler quelques exemples d'espaces fonctionnels d'usage courant. On renverra souvent à [Rudin] pour les démonstrations des théorèmes principaux. Il est important de noter que ces théorèmes peuvent souvent se démontrer "à la main", sans trop d'efforts, sur les exemples considérés ci-après.

1.2. Espaces de Fréchet et semi-normes. Soit E un espace de Fréchet ; il admet alors une base de voisinages dénombrable. La théorie générale des e.v.t.l.c.s. [Rudin, Théorème 1.24] implique que l'on peut munir E d'une métrique compatible avec sa topologie, invariante par translation, de telle sorte que les boules ouvertes $B_r(0)$ soient symétriques ($x \in B_r \Leftrightarrow -x \in B_r$) et convexes. Comme on l'a déjà remarqué, l'application $d(0, \cdot)$ n'est en général pas une norme. En revanche, on peut utiliser d pour construire une **famille de semi-normes séparante**. On rappelle qu'une semi-norme sur E est une application N à valeurs dans $\mathbb{R}_+$ vérifiant $N(x + y) \leq N(x) + N(y)$, $N(\lambda x) = |\lambda| N(x)$ pour tous $x, y \in E$, $\lambda \in \mathbb{R}$.

Définition I-7 (semi-normes séparantes). *Soit E un espace vectoriel topologique, et $\mathcal{P}$ une famille de semi-normes sur E ; on dit que $\mathcal{P}$ est séparante si, pour tout $x \neq 0$, il existe $p \in \mathcal{P}$ tel que $p(x) \neq 0$.*

L'idée principale est d'utiliser une base de voisinages convexes pour construire une "base" de semi-normes. On sait que la boule unité d'une semi-norme est un convexe symétrique ; la réciproque est l'objet de la proposition importante qui suit :

Proposition I-8 (jauge de Minkowski). *Soit E un espace vectoriel topologique, et C un voisinage convexe symétrique de 0. Alors l'application p définie par*

$$p(x) = \inf\{t > 0;\ x/t \in C\}$$

est une semi-norme continue sur E. En outre,

$$x \in C \Longrightarrow p(x) \leq 1; \qquad p(x) < 1 \Longrightarrow x \in C.$$

Démonstration. Montrons d'abord que p est à valeurs dans $\mathbb{R}_+$, i.e. pour tout $x \in E$ il existe $\lambda > 0$ tel que $\lambda x \in C$. C'est une conséquence immédiate de la continuité en 0 de l'application $\lambda \longmapsto \lambda x$.

Il est clair que $p(\lambda x) = \lambda p(x)$ pour $\lambda \geq 0$, et comme C est symétrique on a $p(-x) = p(x)$. Pour prouver que p est une semi-norme, on utilise la relation

$$\frac{x+y}{s+t} = \frac{t}{s+t}\frac{x}{t} + \frac{s}{s+t}\frac{y}{s},$$

qui implique, par convexité de B, que

$$\frac{x}{t}, \frac{y}{s} \in C \Longrightarrow \frac{x+y}{t+s} \in C.$$

On en déduit facilement $p(x + y) \leq p(x) + p(y)$.

Soient $\varepsilon > 0$ et $x \in E$; l'ensemble $x+\varepsilon B$ est alors un voisinage de x. Si $y \in x+\varepsilon B$, on a $|p(x) - p(y)| \leq p(x - y) = \varepsilon p((x - y)/\varepsilon)$; mais $(x - y)/\varepsilon \in B$, cette dernière quantité est donc majorée par ε. On en déduit que p est continue.

Il est clair que $x \in C \Longrightarrow p(x) \leq 1$. Soit maintenant x tel que $p(x) < 1$, alors il existe $\lambda > 1$ tel que $\lambda x \in C$, et par convexité $x \in C$. $\square$

Proposition I-9 (description des espaces de Fréchet en termes de semi-normes). *(i) Soit E un espace de Fréchet. Alors il existe une famille séparante dénombrable $(p_n)_{n\in\mathbb{N}}$ de semi-normes telle que la topologie de E soit compatible avec la distance*

$$D(x,y) = \max_{n\in\mathbb{N}} \frac{2^{-n}p_n(x-y)}{1+p_n(x-y)}. \tag{1}$$

(ii) Réciproquement, soit E un espace vectoriel, et $(p_n)_{n\in\mathbb{N}}$ une famille dénombrable séparante de semi-normes ; si on définit d par la formule

$$d(x,y) = \max_{n\in\mathbb{N}} \frac{2^{-n}p_n(x-y)}{1+p_n(x-y)},$$

alors E, muni de la topologie induite par d, est un e.v.t.l.c.s. et la métrique d est invariante par translation. **Pour établir que E est un espace de Fréchet, il suffit donc de prouver que d est complète.**

En outre, une base dénombrable de voisinages est donnée par les intersections finies d'ensembles de la forme

$$V(k,n) := \{x \in E;\ p_n(x) < 1/k\}, \qquad (k,n) \in \mathbb{N}\times\mathbb{N}.$$

En d'autres termes, les ouverts sont les **unions arbitraires d'intersections finies de** $V(k,n)$**.** *De plus, la topologie est inchangée si l'on multiplie les p_n par des nombres A_n strictement positifs arbitraires.*

Remarques I-10. (i) Dans la pratique, c'est toujours par une famille de semi-normes que l'on définira un espace de Fréchet. Quitte à remplacer p_k par $\sum_{j\le k} p_k$, on peut toujours les choisir croissantes.

(ii) Le choix du facteur 2^{-n} dans la formule (1) est arbitraire ; on aurait pu la remplacer par n'importe quelle suite de nombres strictement positifs, tendant vers 0, fixée a priori. On aurait aussi pu remplacer la fonction $p_n/(1+p_n)$ par $p_n/(\delta_n+p_n)$, pour n'importe quelle suite δ_n.

(iii) Même si l'une des semi-normes p_n est une norme, il n'y a aucune raison pour que cette norme fasse de E un espace de Banach.

Démonstration. (i) Soit E un espace de Fréchet ; comme on l'a déjà mentionné, on peut choisir une distance d sur E pour laquelle les boules sont symétriques et convexes. La famille $B(0,1/n)$ forme alors une base de voisinages dénombrables convexes. On définit p_n comme la jauge de Minkowski associée à $B(0,1/n)$: c'est une semi-norme.

Montrons que la famille (p_n) est séparante. Soit $x \neq 0$, et n tel que $d(0,x) > 1/n$. La boule $B(0,1/n)$ étant convexe, l'ensemble des $\lambda \geq 0$ tels que $\lambda x \in B(0,1/n)$ est un intervalle de $\mathbb{R}_+$ contenant 0. Comme il ne contient pas 1, il est borné, et la quantité $\inf\{t;\ x/t \in B(0,1/n)\}$ est strictement positive.

La boule $B_D(0,r)$ pour la distance D est donnée par l'intersection des ensembles $\{p_n/(1+p_n) \leq 2^n r\}$. Seuls les indices n tels que $2^n r \leq 1$ comptent dans cette intersection, et la fonction $X \mapsto X/(1+X)$ est croissante ; la boule $B_D(0,r)$ est donc donnée par l'intersection d'un nombre **fini** d'ensembles de la forme $p_n^{-1}([0,r_n))$. Les p_n étant continues, une telle intersection est ouverte. Les boules ouvertes pour D sont donc des ouverts de E.

Réciproquement, montrons que toute boule $B(0,1/m)$ contient une boule $B_D(0,r)$ pour la distance D. Soit x dans $B_D(0,r)$, alors

$$\frac{p_m(x)}{1+p_m(x)} \leq r2^m.$$

En particulier,

$$p_m(x) \leq \frac{2^m r}{1-2^m r},$$

qui est strictement inférieur à 1 si r est choisi assez petit. Un tel x appartient à $B(0,1/m)$ par la Proposition I-8. On conclut que d et D définissent bien la même topologie.

(ii) Pour tous $X, Y \geq 0$, l'inégalité

$$\frac{X+Y}{1+X+Y} \leq \frac{X}{1+X} + \frac{Y}{1+Y},$$

que l'on peut établir simplement en étudiant les variations de ces expressions en fonction de X, implique que la fonction $(x,y) \longmapsto p(x-y)/(1+p(x-y))$ vérifie l'inégalité triangulaire dès que p est une semi-norme. Il s'ensuit que d vérifie aussi cette inégalité. En outre, $d(x,0)=0$ si et seulement si pour toute norme p_n, $p_n(x) = 0$, ce qui n'est possible que si $x=0$ puisque les p_n forment une famille séparante. On conclut que d est bien une distance, clairement invariante par translation.

Chacune des semi-normes p_n est continue pour cette distance, et une boule ouverte pour d est une intersection d'ensembles de la forme $p_n^{-1}([0,r_n))$, qui sont tous convexes puisque les p_n sont des semi-normes. L'espace E est donc bien un e.v.t.l.c.s. muni d'une distance invariante.

On laisse en exercice les dernières assertions de l'énoncé ; un examen attentif de la preuve précédente suffit presque à les démontrer (voir [Rudin, Théorème 1.37] pour la définition de la topologie via les $V(k,n)$). $\square$

Exemple I-11. Soit Ω un ouvert de $\mathbb{R}^n$, et

$$\Omega_k := \{x \in \Omega;\ d(x,\partial\Omega) > k^{-1}.$$

Les Ω_k forment une suite emboîtée d'ouverts qui “emplissent” tout Ω. Pour tout $p \geq 1$, la famille de semi-normes

$$p_k(f) := \|f\|_{L^p(\Omega_k)}$$

définit une structure d'espace de Fréchet sur l'espace $L^p_{\mathrm{loc}}(\Omega)$ des fonctions qui sont p-intégrables sur tout compact. La famille de semi-normes

$$p_k(f) := \sup_{\Omega_k} |f|$$

définit une structure d'espace de Fréchet sur l'espace $C(\Omega)$ (que l'on peut aussi noter $C_{\mathrm{loc}}(\Omega)$ des fonctions continues sur Ω.

Souvent, les espaces de Fréchet apparaissent comme des limites d'espaces de Banach. Voici un exemple simple :

Proposition I-12 (limite décroissante d'espaces de Banach). *Soit $(E_n, \|\cdot\|)_{n\in\mathbb{N}}$ une famille décroissante d'espaces de Banach, i.e. $E_{n+1} \subset E_n$ pour tout n, avec injection continue. Alors, $E := \cap E_n$, muni de la famille des normes $\|\cdot\|_n$, est un espace de Fréchet.*

Démonstration de la Proposition I-12. Les $p_n = \|\cdot\|_n$ sont des normes, elles forment donc une famille séparante. Il suffit de vérifier la complétude. Soit donc $(x_k)_{k\in\mathbb{N}}$ une famille de Cauchy pour la distance

$$d(x,y) = \sup_n 2^{-n} \frac{p_n(x-y)}{1+p_n(x-y)}.$$

L'inégalité

$$p_n(x-y) \leq \frac{2^n d(x,y)}{1-2^n d(x,y)}$$

montre que (x_k) est une suite de Cauchy dans tous les E_n ; elle converge donc dans E_n vers une limite $x^{(n)}$. Par continuité de l'injection, elle converge également dans E_{n-1} vers $x^{(n)}$; par unicité de la limite, on a donc $x^{(n-1)} = x^{(n)}$, et partant, tous les $x^{(n)}$ sont égaux à un élément x qui appartient à tous les E_n. On a alors

$$\frac{p_n(x_k - x)}{1+p_n(x_k-x)} \xrightarrow[k\to\infty]{} 0$$

pour tout n, et il est facile d'en déduire que $d(x_k,x) \to 0$ quand $k\to\infty$. La distance d est donc bien complète. $\square$

1.3. Quotient d'espaces vectoriels topologiques. Un espace vectoriel E étant donné, une opération d'usage très courant pour construire un espace "plus petit" consiste à en prendre le **quotient** par un sous-espace vectoriel F, i.e. quotienter E par la relation d'équivalence

$$x\mathcal{R}y \Longrightarrow x-y\in F.$$

Dans le cas des espaces vectoriels topologiques, on impose presque toujours à F d'être fermé. La proposition suivante [Rudin, Théorème 1.41] assure que les principales propriétés de E sont alors transmises à l'espace quotient E/F.

Proposition I-13 (quotient d'espace vectoriel topologique). *Soient E un espace vectoriel topologique et F un sous-espace vectoriel fermé de E. Soit π l'application linéaire de E dans E/F qui à x associe $x+F$, ou "surjection canonique" de E dans E/F. On munit E/F de la plus petite topologie qui rende π continue : une partie de E/F est ouverte si et seulement si son image réciproque par π est ouverte. Cette topologie est appelée topologie quotient. Alors*

(i) L'espace E/F est un espace vectoriel topologique ;

(ii) Si E est un e.v.t.l.c.s., alors E/F aussi ;

(iii) Si E est un espace de Fréchet, alors E/F aussi ;

(iv) Si E est un espace de Banach, alors E/F aussi.

Démonstration. (i) L'application π est continue et commute avec l'addition : $\pi(x+y) = \pi(x)+\pi(y)$. Soit O un ouvert de E/F, l'ensemble de toutes les classes $X, Y \in E/F$ coïncide avec la réunion des $f+x+y$, où $\pi(x+y)\in O$, $f\in F$. Puisque l'addition sur $E\times E$ et l'application π sont continues, l'ensemble A des couples (x,y) qui vérifient $\pi(x+y)\in O$ est ouvert. Pour tout $f\in F$, l'ensemble $f+A$ est ouvert, car image d'un ouvert par une translation. Il s'ensuit que la réunion des $f+A$ est également ouverte. Autrement dit, l'addition est bien continue de $(E/F)\times(E/F)$ dans E/F.

Par un raisonnement similaire, en utilisant le fait que π commute avec la multiplication scalaire, on montre que la multiplication scalaire est continue de $\mathbb{R}\times(E/F)$ dans E/F.

Enfin, la classe nulle de E/F est l'ensemble F, qui est fermé par hypothèse : la partie $\{0\}$ de E/F est donc fermée.

(ii) Puisque π est surjective, on a $\pi^{-1}(\pi(O)) = F + O$ pour tout ouvert O de E. L'espace $F + O$ est une union de translatés de O, donc d'ouverts, c'est donc un ouvert. Il s'ensuit que π est ouverte (l'image de tout ouvert est un ouvert) et envoie base de voisinages sur base de voisinages. Par linéarité, elle envoie base de voisinages convexes sur base de voisinages convexes.

(iii) Soit d_E une distance pour laquelle E est un espace de Fréchet. On vérifie que la distance

$$d_{E/F}(\pi(x), \pi(y)) := \inf_{f\in F} d_E(x-y, f)$$

est bien définie sur E/F, métrise la topologie quotient, et est invariante par translation ; en particulier, on note que $d_{E/F}(\pi(x), \pi(y)) = 0$ si et seulement la distance de $x-y$ à F est nulle, ce qui signifie $x - y \in F$ puisque F est fermé. Il reste à montrer que E/F est complet. Soit donc $(X_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans E/F ; par récurrence on peut extraire une sous-suite, notée $(X_{n_k})_{k\in\mathbb{N}}$, telle que

$$d_{E/F}(X_{n_k}, X_{n_{k+1}}) \leq 2^{-k}.$$

Par récurrence, on peut alors choisir des représentants x_k de X_{n_k} (i.e. $X_{n_k} = x_{n_k}+F$) tels que

$$d_E(x_k, x_{k+1}) \leq 2 \times 2^{-k}.$$

On vérifie facilement que la suite $(x_k)_{k\in\mathbb{N}}$ est alors de Cauchy pour d, elle converge donc dans E vers x. Par continuité de π, on a $X_{n_k} \to X := x + F$. La suite (X_n) est donc une suite de Cauchy admettant une sous-suite convergente ; elle est donc elle-même convergente.

(iv) Si E est un espace de Banach, la définition de la distance $d_{E/F}$ ci-dessus se transforme en

$$d_{E/F}(\pi(x), \pi(y)) := \inf_{f\in F} \|x - y - f\|$$

et l'application $N(z) := d_{E/F}(0, z)$ définit bien une norme sur E/F, pour laquelle E/F est complet. $\square$

Remarque I-14. On retiendra la définition de la distance quotient :

$$\begin{aligned} d_{E/F}(\pi(x), \pi(y)) &= \inf_{f\in F} d_E(x-y, f) \\ &= \inf_{x'\mathcal{R}x,\, y'\mathcal{R}y} d_E(x', y'). \end{aligned}$$

Dans le cas où E est un espace de Banach, on obtient ainsi la définition de la norme quotient :

$$\|x\|_{E/F} = \inf_{f\in F} \|x - f\|.$$

Remarque I-15. Noter, dans la preuve précédente, l'utilisation du critère pratique suivant : *un espace métrique E est complet si et seulement si, de toute suite (x_n) telle que $d(x_n, x_{n+1}) < 2^{-n}$, on peut extraire une sous-suite convergente.*

1.4. Complétude. Si X est un espace métrique, on sait bien définir les concepts de suite de Cauchy et de complétude : une suite (x_n) est de Cauchy si $d(x_m, x_n) \longrightarrow 0$ quand $(m, n) \to \infty$; et l'espace est dit complet si toute suite de Cauchy converge. Cette notion est si importante qu'il semble désirable, pour étudier des e.v.t.l.c.s. assez généraux, de l'étendre à un cadre non métrique.

Définition I-16 (complétude). *Soit E un espace vectoriel topologique. On dit que $(x_n)_{n\in\mathbb{N}}$, à valeurs dans E, est une suite de Cauchy si, pour tout voisinage V de 0, il existe $N \in \mathbb{N}$ tel que pour $m, n \geq N$ on ait $x_n - x_m \in V$. En d'autres termes, la suite (x_n) est de Cauchy si $x_n - x_m \longrightarrow 0$ pour $(m, n) \to \infty$. On dit que E est complet si toute suite de Cauchy à valeurs dans E est convergente.*

Remarque I-17. Il est clair que toute suite convergente est une suite de Cauchy, comme on le souhaite.

Exemple I-18. Vérifier que cette notion généralise la notion habituelle (définie par une métrique) de suite de Cauchy.

1.5. Bornitude. Un autre outil fondamental de l'étude des espaces de Banach est la caractérisation des applications linéaires continues comme les applications linéaires bornées sur la boule unité. Pour généraliser ce résultat à un cadre non métrique, on utilise la notion d'**ensemble borné**.

Définition I-19 (ensembles bornés). *Soit E un espace vectoriel topologique. On dit que $B \subset E$ est borné si, pour tout voisinage V de 0, il existe $\lambda > 0$ tel que $\lambda B \subset V$.*

Exemple I-20. Tout ensemble compact de E est borné [Rudin, Théorème 1.15].

Remarques I-21. (i) En abrégé, un ensemble borné est un ensemble que l'on peut envoyer dans n'importe quel voisinage de 0 par homothétie. Pour éviter les confusions, on peut dire que c'est la définition d'un ensemble borné au sens des espaces vectoriels topologiques.

(ii) Il est clair que dans un espace vectoriel normé, un ensemble B est borné si et seulement si il est borné au sens de la norme, i.e. $\sup\{\|x\|;\ x \in B\} < +\infty$. En revanche, dans un espace de Fréchet, un ensemble B borné au sens de la distance, i.e. tel que $\sup\{d(0, x);\ x \in B\} < +\infty$, n'est pas nécessairement borné u sens des espaces vectoriels topologiques. Pour s'en convaincre, noter que, quitte à remplacer d par $d/(1 + d)$, on peut toujours supposer que E tout entier est borné au sens de la distance, alors qu'il est bien sûr non borné (sauf cas trivial) au sens des espaces vectoriels topologiques.

(iii) On peut aller plus loin dans la dernière remarque et montrer qu'en général, les boules ne sont jamais bornées dans un espace de Fréchet qui ne soit pas un espace de Banach. En fait, on peut prouver que, dans un espace vectoriel topologique arbitraire E, *si l'origine admet un voisinage convexe borné, alors E est normable* [Rudin, Théorème 1.39].

Dans la pratique, le principal critère de bornitude est fourni par le théorème suivant [Rudin, Théorème 1.37] :

Théorème I-22 (parties bornées des espaces de Fréchet). *Soit E un espace de Fréchet, défini par une famille dénombrable de semi-normes $(p_n)_{n\in\mathbb{N}}$ comme dans le*

Théorème I-9. Alors une partie B de E est bornée si et seulement si il existe une suite de nombres positifs $(M_n)_{n\in\mathbb{N}}$ telle que

$$\forall n, \qquad \sup_{x\in B} p_n(x) \leq M_n.$$

On notera bien que ce résultat n'impose aucune hypothèse d'uniformité sur les nombres M_n.

1.6. Applications linéaires continues. On définit les applications linéaires continues de manière naturelle, comme les applications linéaires qui sont continues au sens de la topologie, ou de manière équivalente continues en 0 ; elles sont alors automatiquement uniformément continues [Rudin, Théorème 1.17]. Si l'espace d'arrivée est $\mathbb{R}$, on parle de forme linéaire continue.

Définition I-23 (Applications linéaires bornées). *Soient E et F deux espaces vectoriels topologiques, et $\Lambda : E \to F$ une application linéaire. On dit que Λ est bornée si elle envoie les parties bornées de E sur des parties bornées de F.*

Les principaux résultats sur les liens entre bornitude et continuité des applications linéaires sont résumés dans les théorème suivants [Rudin, Théorèmes 1.18 et 1.32]

Théorème I-24 (applications linéaires continues). *Soient E et F deux espaces vectoriels topologiques ; alors*

(i) Toute application linéaire continue entre E et F est bornée.

(ii) Si E est métrisable, alors la continuité équivaut à la bornitude, et équivaut également à la **continuité séquentielle** *en 0, i.e. l'image d'une suite qui converge vers 0 dans E est une suite qui converge vers 0 dans F.*

Théorème I-25 (formes linéaires continues). *Soient E un espace vectoriel topologique, et Λ une forme linéaire sur E. Alors les quatre propositions suivantes sont équivalentes : (a) Λ est continue, (b) le noyau de Λ est fermé, (c) le noyau de Λ n'est pas dense, (d) Λ est bornée dans un voisinage de 0.*

Dans le cas des espaces de Fréchet, c'est le critère suivant que l'on utilise en pratique.

Théorème I-26 (applications linéaires continues entre espaces de Fréchet). *Soient E et F deux espaces de Fréchet, dont la topologie est définie par des familles $(p_n)_{n\in\mathbb{N}}$ et $(q_n)_{n\in\mathbb{N}}$ respectives de semi-normes. Alors, une application linéaire $L : E \to F$ est bornée si et seulement si pour tout $k \in \mathbb{N}$ il existe $J = J(k) \in \mathbb{N}$ et $C = C(k) \geq 0$ tels que*

$$\forall x \in E, \qquad q_k(Lx) \leq C \max_{1\leq j\leq J} p_j(x).$$

Démonstration. Si l'inégalité ci-dessus est vérifiée, et B est une partie bornée de E, on sait qu'il existe des nombres A_j tels que $p_j(x) \leq A_j$ pour tout $x \in E$. Alors, pour tout $y \in T(B)$ on aura $q_k(y) \leq C(k)A_{j(k)}$, ce qui prouve que $T(B)$ est bornée.

Réciproquement, si L est continue, elle est en particulier continue en 0, et l'image réciproque du voisinage $\{q_k \leq 1\}$ de 0 dans F est un voisinage de 0 dans E. Il contient donc l'intersection d'un nombre fini de voisinages de la forme $\{p_j \leq r_j\}$. On pose $r = \inf r_j$, $J = \max j$, on voit que si $\max_{j=1}^{J} p_j(x) < r$, alors $q_k(x) \leq 1$. La conclusion en découle par linéarité. □

Particularisons le critère précédent au cas où l'espace d'arrivée est $\mathbb{R}$.

Théorème I-27 (formes linéaires continues sur un espace de Fréchet). *Soit E un espace de Fréchet, dont la topologie est définie par des familles $(p_n)_{n\in\mathbb{N}}$ de semi-normes. Alors, une forme linéaire $\Lambda : E \to \mathbb{R}$ est bornée si et seulement si il existe $J \in \mathbb{N}$ et $C \geq 0$ tels que*

$$\forall x \in E, \qquad |\Lambda x| \leq C \max_{1\leq j\leq J} p_j(x).$$

Exemple I-28. (i) Soit $\varphi : \mathbb{R}^n \to \mathbb{R}$ une fonction mesurable localement bornée. Alors

$$\Lambda : f \longmapsto f\varphi$$

définit une application linéaire bornée (et donc continue) de $L^p_{\mathrm{loc}}(\mathbb{R}^n)$ dans $L^p_{\mathrm{loc}}(\mathbb{R}^n)$.

(ii) Soit $\varphi : \mathbb{R}^n \to \mathbb{R}$ une application continue à support compact. Alors

$$\Lambda : f \longmapsto f * \varphi$$

définit une application linéaire bornée (et donc continue) de $L^1_{\mathrm{loc}}(\mathbb{R}^n)$ dans $C_{\mathrm{loc}}(\mathbb{R}^n)$.

Remarque I-29. Le lecteur pourrait légitimement demander qu'on lui propose maintenant des exemples de formes linéaires discontinues sur des espaces de Fréchet. Malheureusement ou heureusement, c'est sans intérêt : dans le cadre de l'axiomatique mathématique traditionnelle, la construction de formes linéaires discontinues nécessite l'utilisation de l'axiome du choix dans sa version la plus forte, axiome dont l'utilisation prête à controverse, et à juste titre selon l'auteur. En pratique, tant que l'on travaille sur un espace de Fréchet, "toutes les formes linéaires sont continues".

1.7. Propriété de Heine-Borel. Dans un espace vectoriel normé de dimension finie, les compacts sont les fermés bornés. On sait bien que ce résultat est faux dans un espace vectoriel normé de dimension infinie :

Théorème I-30 (Théorème de F. Riesz). *Soit E un espace vectoriel normé ; alors sa boule unité est compacte si et seulement si E est de dimension infinie.*

En particulier, dans un espace de Banach E, il y a équivalence entre les deux énoncés : "les parties compactes sont les parties fermées et bornées" et "E est de dimension finie", de sorte qu'en dimension infinie, la compacité est une propriété beaucoup plus forte que la bornitude.

Il se trouve que cette exclusion n'existe pas dans les espaces de Fréchet ou dans les e.v.t.l.c.s., et que l'on trouve des exemples naturels d'espaces fonctionnels (non normés bien sûr) dans lesquels les ensembles fermés et bornés sont compacts. On donne un nom à cette propriété :

Définition I-31 (propriété de Heine-Borel). *Soit E un espace vectoriel topologique. On dit qu'il satisfait la propriété de Heine-Borel si toute partie fermée et bornée de E est compacte.*

Remarque I-32. On a déjà mentionné qu'un compact est toujours borné ; il est bien sûr fermé.

Exemple I-33. Soit Ω un ouvert de $\mathbb{C}$, et $H(\Omega)$ l'espace des fonctions holomorphes de Ω dans $\mathbb{C}$. Une fonction dans $H(\Omega)$ peut bien sûr diverger près du bord de Ω, mais elle est C^∞ dans Ω. On peut munir $H(\Omega)$ d'une structure d'espace de

Fréchet naturelle, correspondant à la "convergence uniforme sur les compacts" : pour cela, on introduit une suite exhaustive Ω_k de sous-ouverts de Ω, et les semi-normes

$$p_k(f) := \sup_{z\in\Omega_k} |f(z)|.$$

En exercice, le lecteur pourra caractériser les parties bornées de $H(\Omega)$, et vérifier, en utilisant les propriétés des fonctions holomorphes et le théorème d'Ascoli, que de toute partie bornée et fermée on peut extraire une suite qui converge uniformément sur tout compact. Autrement dit, $H(\Omega)$ possède la propriété de Heine-Borel. Nous reviendrons plus tard sur cet exemple...

1.8. Topologies faibles.

Définition I-34 (topologies faible et faible-$*$). *Soient E un espace vectoriel topologique, et E^* son dual topologique, i.e. l'espace vectoriel des formes linéaires continues sur E. Tout $x \in E$ s'identifie naturellement à une forme linéaire sur E^*, l'évaluation en x. On appelle*

- topologie forte sur E : la topologie initiale de E ;

- topologie faible sur E : la topologie la plus grossière qui rende continus tous les éléments de E^* ;

- topologie faible-$*$ sur E^* : la topologie la plus grossière qui rende continus tous les éléments de E.

Remarque I-35. Il est clair que si E est réflexif, i.e. $E^{**} = E$, alors la topologie faible et la topologie faible-$*$ sur E^* coïncident. Même si E n'est pas réflexif, on peut montrer [Rudin p. 68] que toute forme linéaire continue sur E^*, muni de la topologie faible-$*$, est l'opération d'évaluation en un certain élément de E.

Les topologies faibles ne sont en général pas métrisables, et ne peuvent donc se définir par la seule donnée des notions de convergence de suites. Cependant, c'est bien la notion de convergence faible de suites qui est utile en pratique, on va donc la définir explicitement sans craindre la redondance.

Définition I-36 (convergence faible). *Soient E un espace vectoriel topologique et E^* son dual topologique. On dit qu'une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X converge au sens faible vers $x \in E$ si pour tout $\Lambda \in E^*$ on a*

$$\Lambda x_n \xrightarrow[n\to\infty]{} \Lambda x.$$

On dit qu'une suite $(\Lambda_n)_{n\in\mathbb{N}}$ d'éléments de E^* converge au sens faible-$*$ vers $\Lambda \in E^*$ si, pour tout $x \in E$, on a

$$\Lambda_n x \xrightarrow[n\to\infty]{} \Lambda x.$$

Il est clair que l'on pourrait définir des notions plus générales en remplaçant l'espace E^* par l'espace des applications linéaires continues de E dans F, où F est un espace vectoriel topologique.

Exemple I-37. Soit Ω un ouvert de $\mathbb{R}^n$. Pour $1 < p < \infty$, les topologies faible et faible-$*$ sur $L^p(\Omega)$ coïncident, et la notion de convergence associée est

$$\forall g \in L^{p'}(\Omega), \qquad \int_\Omega f_n g \longrightarrow \int_\Omega fg$$

(convergence contre des fonctions test dans L^p). Pour $p = \infty$, la topologie faible-$*$ correspond à la convergence contre des fonctions test dans L^1 ; pour $p = 1$, la topologie faible correspond à la convergence contre des fonctions test dans L^∞. On s'interdira en revanche de considérer la convergence faible-$*$ dans L^1, ou la convergence faible dans L^∞ (la question de savoir si L^1 est le dual de L^∞ touche à de subtiles questions d'axiomatique, et la réponse est négative si l'on admet l'axiome du choix.....)

Quel est l'intérêt d'appauvrir la topologie ? Une des motivations majeures peut se formuler comme suit : *moins il y a d'ouverts, plus il y a de compacts.* Il est beaucoup plus facile d'être compact pour la topologie faible que pour la topologie forte. Ainsi, l'injection de E^* dans $w*-E*$ est compacte, au sens suivant : les boules dans l'espace vectoriel normé E^* sont compactes pour la topologie faible-$*$. Cette propriété est particulièrement utile pour certains théorèmes d'existence faisant appel à des méthodes non constructives.

I-2. Résultats de continuité automatique

Dans cette section, nous allons passer en revue quelques théorèmes très utilisés en analyse fonctionnelle, qui ont tous la forme générale : sous certaines hypothèses de complétude, certaines applications (linéaires ou bilinéaires) sont automatiquement continues. Très puissants, ces résultats rendent de grands services dans les démonstrations théoriques, mais il convient de les considérer avec la plus grande méfiance : du fait de leur côté non constructif, ils mèneront presque toujours, dans des situations concrètes, à des résultats désastreux. Ainsi, ils ne donnent aucun ordre de grandeur des constantes mises en jeu (normes d'applications linéaires continues, etc.), et il est en pratique impossible d'adapter leur preuve pour obtenir de telles estimations constructives. En particulier, il est en pratique inutile de connaître leur démonstration, et nous les admettrons tous.

Nous recenserons trois grands principes : le théorème de Banach-Steinhaus, celui du graphe fermé, et celui de l'application ouverte. Leur démonstration est d'habitude subordonnée au théorème de Baire : *dans un espace métrique complet, une intersection dénombrable d'ouverts denses est dense.* Les preuves qui en résultent ne sont pas très difficiles mais particulièrement opaques à l'intuition (voir [Rudin]). On peut souvent se passer de cet usage du théorème de Baire, mais **la complétude est une hypothèse essentielle**.

2.1. Théorème de Banach-Steinhaus. Le théorème de Banach-Steinhaus est un outil puissant mais principalement théorique, peu constructif, dont il vaut mieux éviter l'usage si possible. Il admet plusieurs versions, dont la plus simple à retenir est peut-être l'énoncé informel suivant : **sous une hypothèse de complétude, un ensemble faiblement borné est uniformément borné**. Voici un énoncé plus précis [Rudin, Théorèmes 2.4, 2.5 et 2.6] :

Théorème I-38 (théorème de Banach-Steinhaus). *Soient E un espace de Fréchet et F un espace vectoriel topologique, et soit $\mathcal{L} \subset L(E, F)$ une famille d'applications linéaires continues de E dans F. On suppose que $\mathcal{L}$ est faiblement borné, au sens où $\mathcal{L}x$ est borné dans F pour tout $x \in E$. Alors $\mathcal{L}$ est uniformément borné, au sens où pour tout ensemble borné $\mathcal{A}$ de E, il existe un ensemble borné $\mathcal{B}$ de F tel que les images $L(\mathcal{A})$, pour $L \in \mathcal{L}$, soient toutes incluses dans $\mathcal{B}$.*

Si l'on particularise ce théorème au cas où F est un espace de Fréchet, et que l'on utilise la caractérisation des ensembles bornés en termes de semi-normes, on obtient l'énoncé suivant.

Théorème I-39 (théorème de Banach-Steinhaus entre espaces de Fréchet). *Soient E et F deux espaces de Fréchet, munis de familles de semi-normes $(p_k)_{k\in\mathbb{N}}$ et $(q_\ell)_{\ell\in\mathbb{N}}$. Soit $\mathcal{L}$ une famille d'applications linéaires continues de E dans F. On suppose que pour tout $x \in E$, et pour tout $\ell \in N$,*

$$\sup_{L\in\mathcal{L}} q_\ell(L(x)) < +\infty.$$

Alors, pour tout $\ell \in \mathbb{N}$ on peut trouver un indice k et une constante C_ℓ telle que

$$\sup_{L\in\mathcal{L}} q_\ell(L(x)) \leq Cp_k(x).$$

Le plus souvent, on utilise le théorème de Banach-Steinhaus à travers les deux conséquences suivants, qui assurent automatiquement la continuité de certains opérateurs [Rudin, Théorèmes 2.8 et 2.17].

Théorème I-40 (Une limite d'applications linéaires continues est continue). *Soient E un espace de Fréchet et F un espace vectoriel topologique. Soit $(L_n)_{n\in\mathbb{N}}$ une famille d'applications linéaires continues de E dans F. On suppose que pour tout $x \in E$, la limite*

$$Lx := \lim_{n\to\infty} L_n x$$

existe dans F. Alors l'application linéaire L ainsi définie est automatiquement continue.

Théorème I-41 (Une forme bilinéaire séparément continue est continue). *Soient E et F deux espaces de Fréchet, et G un espace topologique. Soit b une forme bilinéaire continue de $E \times F$ dans G. Si b est séparément continue par rapport à son premier et son deuxième argument, alors elle est continue sur $E \times F$.*

Remarque I-42. Encore une fois, il est impossible de construire explicitement des formes bilinéaires non continues...

2.2. Théorème de l'application ouverte. On dit qu'une application f est ouverte si l'image par f de tout ouvert est un ouvert. C'est le cas en particulier des réciproques de bijections continues. Le théorème suivant [Rudin, Théorèmes 2.11 et 2.12] traite du lien entre cette notion et la linéarité.

Théorème I-43 (les applications linéaires injectives sont ouvertes). *Soient E et F deux espaces de Fréchet, et $L : E \to F$ une application linéaire continue injective. Alors L est ouverte.*

Le corollaire suivant est particulièrement frappant. Dans le cadre des espaces de Banach, on l'appelle **théorème de Banach**.

Corollaire I-44 (réciproque des applications linéaires continues). *Soient E et F deux espaces de Fréchet, et $L : E \to F$ une application linéaire bijective continue. Alors l'application linéaire L^{-1} est continue.*

2.3. Théorème du graphe fermé. Une application $f : X \to Y$ étant donnée, on appelle graphe de f l'ensemble des couples $(x, f(x))$. Si X et Y sont des espaces vectoriels topologiques, il est clair que le graphe d'une application continue $f : X \to Y$ est fermé dans $X \times Y$, (comme image réciproque de $\{0\}$ par l'application continue $(x, y) \longmapsto y - f(x)$) ; mais la réciproque est en général fausse. Cependant, elle est vraie pour des applications linéaires, sous une hypothèse de complétude [Rudin, Théorème 2.15].

Théorème I-45 (théorème du graphe fermé). *Soient E et F deux espaces de Fréchet, et soit $L : E \to F$ une application linéaire dont le graphe est fermé. Alors L est continue.*

I-3. Résultats d'existence et de compacité

Le célèbre théorème de Hahn-Banach est le point de départ de divers résultats d'existence et de compacité, très utilisés pour démontrer l'existence de certains objets, par exemple dans le domaine des équations aux dérivées partielles. Contrairement aux résultats évoqués dans la section précédente, le théorème de Hahn-Banach ne se sert pas de la complétude. Mais, ce qui est peut-être encore pire, il utilise la version forte de l'axiome du choix ! Le théorème de Hahn-Banach a d'ailleurs été inventé dans le cadre des redoutables paradoxes de Banach-Tarski, que l'on peut considérer comme un argument convaincant pour éradiquer la version forte de l'axiome du choix de toutes les mathématiques appliquées. Cependant, dans la plupart des cas, on peut se contenter d'appliquer des formes faibles de ces théorèmes, basées sur un argument de séparabilité ou de dénombrabilité, qui ne nécessitent pas l'axiome du choix dans sa version forte. Dans la suite, nous énoncerons seulement ces versions faibles, et mentionnons les formes fortes en remarques.

3.1. Théorème de Hahn-Banach. On trouvera le théorème suivant au début de [Brézis] ou dans [Rudin, Théorème 3.2]. Comme on peut le deviner d'après l'énoncé, il n'a rien à voir avec la topologie et relève plutôt de la logique axiomatique.

Théorème I-46 (théorème de prolongement de Hahn-Banach). *Soit E un espace vectoriel de dimension dénombrable, et soit p une application sous-additive, positivement homogène de degré 1 ($p(x + y) \leq p(x) + p(y)$, et $p(tx) = tp(x)$ pour $t \geq 0$). Soit F un sous-espace vectoriel quelconque de E, et soit λ une forme linéaire sur F, majorée par p en chaque point de F. Alors il existe une forme linéaire Λ sur E, qui prolonge λ, telle que*

$$\forall x \in E, \qquad -p(-x) \leq \Lambda x \leq p(x).$$

Corollaire I-47 (théorème de Hahn-Banach pour une semi-norme). **Forme forte :** *Soit E un espace vectoriel de dimension dénombrable, et soit p une semi-norme sur E. Soit F un sous-espace vectoriel quelconque de E, et soit λ une forme linéaire sur F, majorée par p en chaque point de F. Alors il existe une forme linéaire Λ sur E, qui prolonge λ et qui soit majorée par p en chaque point de E. En particulier, si p est une norme, munissant E d'une structure d'espace vectoriel normé, alors la norme de Λ en tant que forme linéaire continue vis-à-vis de p est égale à la norme de λ.*

Remarque I-48. On dit ici que E est de dimension dénombrable si il est engendré par une quantité dénombrable de vecteurs. La forme forte de l'axiome du choix permet de se passer de cette hypothèse.

Remarque I-49. On rappelle que la norme d'une application linéaire continue est définie par $\|L\| := \sup \|Lx\|/\|x\|$, où le supremum est pris sur tous les vecteurs x non nuls.

Le théorème de Hahn-Banach admet d'autres interprétations en termes de formes linéaires séparant des convexes, comme on peut le voir dans [Rudin] ou [Brézis]. On en déduit également le théorème suivant.

Théorème I-50 (existence de forme linéaire normalisante). *Soient E un espace vectoriel normé séparable, et $x_0 \neq 0$ un élément de E. Alors il existe une forme linéaire $\Lambda \in E^*$, de norme 1, telle que $\Lambda x_0 = \|x_0\|$.*

Remarque I-51. La forme forte de l'axiome du choix permet de se passer de l'hypothèse de séparabilité.

Remarque I-52. Encore une fois, ce théorème rend service dans le développement de la théorie générale des espaces vectoriels topologiques, mais ce serait une grave erreur de l'appliquer dans des espaces familiers, tels que $L^p(\mathbb{R}^n)$. En effet, dans ce cas particulier, non seulement on peut démontrer le résultat de manière très simple, mais en plus on peut construire la forme linéaire Λ explicitement, ce qui est infiniment plus intéressant.

3.2. Convexité et topologie faible. La topologie faible est (sauf en dimension finie) beaucoup plus grossière que la topologie forte. En particulier, un ensemble fermé pour la topologie faible ne l'est pas forcément pour la topologie forte. Le résultat suivant, basé sur le théorème de Hahn-Banach, montre que cela est cependant vrai des ensembles convexes [Rudin, Théorème 3.12] :

Dans tout ce paragraphe, les hypothèses de séparabilité peuvent être évitées si l'on admet la forme forte de l'axiome du choix.

Théorème I-53 (convexes faiblement fermés). *Soient E un e.v.t.l.c.s. séparable, et C une partie convexe de E. Alors l'adhérence de C pour la topologie faible coïncide avec l'adhérence de C pour la topologie forte. En particulier, C est fermé pour la topologie faible si et seulement si il est fermé pour la topologie forte.*

La traduction en termes de fonctionnelles est souvent utile [Brézis].

Théorème I-54 (continuité des fonctionnelles convexes s.c.i). *Soit E un e.v.t.l.c.s. séparable, et $J : E \to \mathbb{R}$ une fonction convexe, semi-continue inférieurement pour la topologie forte. Alors J est semi-continue inférieurement pour la topologie faible.*

3.3. Théorème de Banach-Alaoglu. Le théorème suivant [Rudin, Théorème 3.15] joue un rôle important dans de nombreuses branches de l'analyse fonctionnelle et des équations aux dérivées partielles. Nous l'énonçons uniquement pour un espace séparable, une hypothèse que l'on ne rencontre pas d'habitude dans les ouvrages d'analyse fonctionnelle : en effet, comme le théorème de Hahn-Banach, le théorème de Banach-Alaoglu sans hypothèse de séparabilité repose fondamentalement sur la forme forte de l'axiome du choix, cette fois via le théorème de compacité de Tychonov. Et comme le théorème de Hahn-Banach, il ne nécessite aucune hypothèse de complétude.

Théorème I-55 (compacité de la boule unité faible-$*$). *Soient E un espace topologique séparable, et V un voisinage de 0 dans E. Soit*

$$K := \{\Lambda \in E^*;\ \forall x \in V,\ |\Lambda x| \leq 1\}.$$

Alors K est séquentiellement compact pour la topologie faible-$*$.

Remarque I-56. On rappelle que la compacité séquentielle signifie la possibilité d'extraire de toute suite d'éléments de K une sous-suite convergente. La deuxième partie de l'énoncé vient du fait utile suivant : si E est séparable, alors K est métrisable. Si on ne suppose pas E séparable, on peut encore montrer que K est un compact, mais à condition d'admettre la forme forte de l'axiome du choix ; en outre ce n'est pas un compact métrique, de sorte que la compacité séquentielle n'est pas a priori vérifiée.

En considérant le cas particulier où V est la boule unité associée à une semi-norme continue (qui peut être la norme de E si E est un espace vectoriel normé, ou un membre d'une famille de semi-normes définissant la topologie de E si E est un espace de Fréchet), on obtient la version suivante.

Corollaire I-57 (théorème de Banach-Alaoglu et semi-normes). *Soient E un espace vectoriel topologique séparable, p une semi-norme continue sur E et $C > 0$ une constante. Alors l'ensemble des formes linéaires $\Lambda \in E^*$ vérifiant*

$$\forall x \in E, \qquad |\Lambda x| \leq Cp(x)$$

est séquentiellement compact pour la topologie faible-$*$.

I-4. Espaces de Banach célèbres

4.1. Espaces de Lebesgue.

Définition I-58 (espace de Lebesgue). *Soient Ω un espace mesuré, et $p \in [1, +\infty]$. Pour $p < +\infty$, on définit $L^p(\Omega)$ comme l'ensemble des fonctions mesurables $f : \Omega \to \mathbb{R}$ telles que f^p soit intégrable ; on $L^\infty(\Omega)$ comme l'ensemble des fonctions mesurables f telles que f est presque partout bornée par une constante finie. On pose alors*

$$\|f\|_{L^p} := \left(\int |f|^p\right)^{1/p}, \qquad \|f\|_{L^\infty} := \inf\{M;\ |f| \leq M \ \ \textit{presque partout}\}.$$

*Si l'on convient d'***identifier deux fonctions égales presque partout***, l'ensemble des classes d'équivalence est alors un espace de Banach pour la norme $\|\cdot\|_{L^p}$, que l'on note toujours $L^p(\Omega)$.*

Remarque I-59. Sans l'identification des fonctions égales presque partout, les fonctions $\|\cdot\|_{L^p}$ ne sont que des semi-normes. Cette identification est sans conséquence pour la plupart des problèmes concrets, mais parfois désastreuse quand on veut étudier des propriétés fines de fonctions mesurables, par exemple si on veut étudier la dimension de Hausdorff d'un ensemble de points de discontinuité de f... De tels problèmes ne relèvent pas de l'analyse fonctionnelle.

Quelques-unes des principales propriétés des espaces de Lebesgue sont rappelées ci-dessous. On pourra consulter [Brézis, Chapitres 3 et 4] pour un exposé systématique et des démonstrations.

Proposition I-60 (propriétés des espaces de Lebesgue). *Soit Ω un espace mesuré, et L^p l'espace de Lebesgue associé, pour un exposant $p \in [1, +\infty]$.*

(i) Pour $1 \leq p < +\infty$, *on a* $(L^p)^* = L^{p'}$, *où* p' *est l'exposant conjugué de* p : $1/p + 1/p' = 1$. *En revanche* $(L^\infty)^*$ *est strictement plus grand que* L^1, *et d'ailleurs* L^1 *n'est le dual d'aucun espace de Banach.*

(ii) Si Ω *est* σ*-fini, les fonctions simples (fonctions mesurables ne prenant qu'un nombre fini de valeurs) dont le support est de mesure finie forment un sous-espace dense dans* L^p *pour* $p < +\infty$.

(iii) Si Ω *est un ouvert de* $\mathbb{R}^n$, *muni de la mesure de Lebesgue, alors pour* $p < +\infty$ *les fonctions* C^∞ *à support compact dans* Ω *sont denses dans* L^p, *et* L^p *est séparable.*

(iv) Si l'on munit L^p *de la topologie faible* $\sigma(L^p, L^{p'})$ *pour* $p < +\infty$, *et* L^∞ *de la topologie faible-** $\sigma(L^\infty, L^1)$, *on obtient un espace vectoriel topologique, non métrisable, mais dans lequel les boules de* L^p *sont des compacts métrisables, pour tout* $p > 1$.

Les espaces de Lebesgue constituent une très bonne échelle pour quantifier l'intégrabilité des fonctions ; cependant, les spécialistes d'analyse fonctionnelle ont souvent besoin d'espaces plus fins qui viennent s'intercaler entre les espaces de Lebesgue : par exemple, les espaces de Lorentz, d'Orlicz, de Marcinkiewicz, de Hardy, ou l'espace BMO de John-Nirenberg. Nous allons seulement considérer les espaces de Lorentz, qui jouent un rôle important en théorie de l'interpolation ; les espaces de Marcinkiewicz en sont un cas limite.

4.2. Espaces de Lorentz et de Marcinkiewicz. L'idée sous-jacente aux espaces de Lorentz est de pouvoir détecter des corrections logarithmiques à l'intégrabilité L^p. Pour cela, on se ramène au cas où la fonction f est une fonction décroissante sur $\mathbb{R}_+$, via l'examen de la distribution des valeurs de f.

Définition I-61 (réarrangement décroissant). *Soit* $f : (X, \mu) \to \mathbb{R}$ *une fonction mesurable, tendant vers 0 à l'infini, au sens où*

$$\forall a > 0, \qquad \mu\{|f| > a\} < +\infty.$$

On définit son réarrangement décroissant f^* *sur* $\mathbb{R}_+$ *par la formule*

$$f^*(t) = \inf\Big\{s \geq 0;\ \mu\{|f| > s\} \leq t\Big\}.$$

C'est une fonction de $\mathbb{R}_+$ *dans* $\mathbb{R}_+$, *décroissante, continue à droite, telle que pour tout* $t > 0$ *et pour tout* $\varepsilon < t$,

$$\mu\{|f| > t - \varepsilon\} \leq \lambda\{f^* > t\} \leq \mu\{|f| > t\},$$

λ *désignant la mesure de Lebesgue sur* $\mathbb{R}_+$.

Définition I-62 (espaces de Lorentz). *Soient* (X, μ) *un espace mesuré, et* $p, q \in [1, +\infty[$. *On définit l'espace de Lorentz* $L^{p,q}(X)$ *comme l'ensemble des fonctions mesurables* f, *tendant vers 0 à l'infini, pour lesquelles*

$$\|f\|_{L^{p,q}(X)} := \left(\int_0^{+\infty} (t^{1/p} f^*(t))^q \, \frac{dt}{t}\right)^{1/q} < +\infty.$$

Dans le cas où $q = \infty$, la définition est plus simple :

Définition I-63 (espaces de Marcinkiewicz). *Soit $p \in [1, +\infty[$ et (X, μ) un espace mesuré. On note $M^p(X) = L^{p,\infty}(X)$, l'espace des fonctions f telles que*

$$\forall t \geq 0, \quad \mu[\{x;\ |f(x)| \geq t\}] \leq C/t^{1/p}.$$

pour une certaine constante $C \geq 0$.

L'espace M^p est aussi appelé espace "espace L^p faible". L'inégalité de Chebyshev montre que $L^p \subset M^p$, l'inclusion étant en général stricte.

Exemple I-64. Les espaces de Lebesgue ne permettent pas de faire la différence entre les fonctions

$$h_{\alpha,\beta} : x \longmapsto \frac{[\log(1/|x|)]^\beta}{|x|^\alpha}$$

pour des valeurs différentes de β. En effet, si X est la boule unité de $\mathbb{R}^n$, la fonction $h_{\alpha,\beta}$ pour $\beta > 0$ appartient à $L^p(X)$ si et seulement si $p < n/\alpha$.

Les espaces de Lorentz au contraire voient la différence : la fonction $h_{\alpha,\beta}$ pour $\beta > 0$ appartient à $L^{p,q}$ si et seulement si $p < n/\alpha$ ou $p = n/\alpha$ et $q < 1/\beta$. Pour $\beta = 0$, cette fonction appartient à $M^p = L^{p,\infty}$.

4.3. Espaces de fonctions continues.

Définition I-65. *Soit Ω un espace topologique. On définit $C_b(\Omega)$ comme l'espace des fonctions continues bornées de Ω dans $\mathbb{R}$; $C_0(\Omega)$ comme l'espace des fonctions continues sur Ω, tendant vers 0 à l'infini ; $C_c(\Omega)$ comme l'espace des fonctions continues à support compact dans Ω. Ces trois espaces, munis de la norme du sup,*

$$\|f\|_\infty := \sup_\Omega |f|,$$

sont des espaces vectoriels topologiques ; les espaces $C_b(\Omega)$ et $C_0(\Omega)$ sont des espaces de Banach, alors que $C_c(\Omega)$ ne l'est pas en général.

Remarque I-66. Si $\Omega = \mathbb{R}^n$, alors $C_0(\Omega)$ est l'adhérence de $C_c(\Omega)$.

Du point de vue quantitatif, ces espaces ne donnent guère de renseignement supplémentaire par rapport aux espaces de Lebesgue, puisque la norme est la même que la norme L^∞. Mais il y a bien sûr une différence considérable entre ces espaces et L^∞.

Si Ω est un ouvert de $\mathbb{R}^n$, le théorème de représentation de Riesz identifie toutes les formes linéaires positives sur $C_c(\Omega)$ avec l'espace des **mesures** de Borel (mesures positives), finies sur les compacts. On s'intéressera dans la suite aux mesures finies, ou plutôt à l'espace vectoriel engendré par les mesures finies, i.e. les mesures de Radon.

Définition I-67 (mesures de Radon). *Soit Ω un ouvert de $\mathbb{R}^n$, on définit $M(\Omega)$, l'espace des mesures de Radon sur Ω, comme l'espace vectoriel de toutes les mesures qui s'écrivent comme différence de deux mesures de Borel finies. C'est un espace vectoriel normé quand on le munit de la norme de la variation totale,*

$$\|\mu\|_{VT} := \inf\{\mu_+[\Omega] + \mu_-[\Omega];\ \mu = \mu_+ - \mu_-\}.$$

4.4. Espaces de fonctions différentiables. En analyse, on est sans cesse préoccupé par des questions de régularité, typiquement pour des fonctions définies sur $\mathbb{R}^n$ ou sur un ouvert de $\mathbb{R}^n$. Si les espaces de Lebesgue permettent d'estimer la taille d'une fonction, ils ne sont pas adaptés à étudier sa régularité. Une première mesure de régularité est donnée par les espaces C^k de fonctions k fois différentiables :

Définition I-68 (espace C^k). *Soient Ω un ouvert de $\mathbb{R}^n$, et $k \in \mathbb{N}$, $k \geq 1$. On définit l'espace $C^k(\Omega)$ comme l'espace des fonctions k fois dérivables dans Ω, dont toutes les dérivées sont bornées jusqu'à l'ordre k. C'est un espace de Banach quand on le munit de la norme*

$$\|f\|_{C^k} := \|f\|_\infty + \sum_{|\beta|=k} \|\nabla^\beta f\|_\infty.$$

Remarque I-69. On note parfois cet espace $C^k(\overline{\Omega})$ pour insister sur le fait que l'on prend le supremum jusqu'au bord. Une notation plus appropriée serait sans doute $C_b^k(\Omega)$.

Cette échelle est cependant assez grossière, et très vite on est amené à introduire des espaces de régularité intermédiaire.

Définition I-70 (espaces de Hölder). *Soient Ω un ouvert de $\mathbb{R}^n$, $k \in \mathbb{N}$, $\alpha \in]0,1]$. On définit l'espace $C^{k,\alpha}(\Omega)$ comme l'espace des fonctions k fois dérivables de Ω dans $\mathbb{R}$, pour lesquelles*

$$\|f\|_{C^{k,\alpha}} := \|f\|_{L^\infty} + \sum_{|\beta|=k} \sup_{x\neq y} \frac{|\nabla^\beta f(x) - \nabla^\beta f(y)|}{|x-y|^\alpha} < +\infty.$$

La fonction $\|\cdot\|_{C^{k,\alpha}}$ définit alors une norme sur $C^{k,\alpha}(\Omega)$.

Remarque I-71. L'espace $C^{0,1}$ coïncide avec l'espace des fonctions Lipschitziennes ; la différence entre $C^{0,1}$ et C^1 est à peu près aussi subtile que celle qui existe entre L^∞ et C_b. On peut aussi imposer que le quotient $|f(x) - f(y)|/|x-y|^\alpha$ tende vers 0 quand $x \to y$, et obtenir ainsi un espace légèrement plus petit que le précédent.

Définition I-72 (fonctions C^∞ à support compact). *Soit Ω un ouvert de $\mathbb{R}^n$. L'intersection des espaces $C_c(\Omega)$ et $C^k(\Omega)$, pour tout k, est appelée espace des fonctions C^∞ à support compact ; on la note $C_c^\infty(\Omega)$ ou $\mathcal{D}(\Omega)$.*

Puisque $C_c(\Omega)$ n'est déjà pas un espace de Banach, il semble inutile de tenter de normer $\mathcal{D}(\Omega)$.

4.5. Espaces de Sobolev. Les espaces de Sobolev connaissent une popularité immense en analyse fonctionnelle, particulièrement en théorie des équations aux dérivées partielles. Leur définition mélange norme L^p et régularité. En particulier, elle fait intervenir des normes L^p de dérivées, sans supposer a priori que les fonctions en jeu soient dérivables ! Pour contourner cette difficulté, nous allons les définir comme limites de fonctions très régulières.

Définition I-73 (espaces de Sobolev). *(i) Soient Ω un ouvert de $\mathbb{R}^n$, $p \in [1,+\infty]$ et k un entier. On appelle espace de Sobolev $W^{k,p}(\Omega)$ la complétion de $C_c^\infty(\Omega)$ pour la norme*

$$\|f\|_{W^{k,p}} := \|f\|_{L^p} + \sum_{|\beta|=k} \|\nabla^\beta f\|_{L^p}.$$

L'espace ainsi obtenu est un espace de Banach.

(ii) Soient Ω un ouvert de $\mathbb{R}^n$, et k un entier. On appelle espace de Sobolev $H^k(\Omega)$ la complétion de $C_c^\infty(\Omega)$ pour la norme

$$\|f\|_{H^k} := \left(\|f\|_{L^2}^2 + \sum_{|\beta|=k} \|\nabla^\beta f\|_{L^2} \right)^{1/2}.$$

L'espace ainsi obtenu est un espace de Hilbert, qui coïncide avec $W^{k,2}$.

Remarque I-74. Bien sûr, $W^{0,p} = L^p$, et $H^0 = L^2$.

Là encore, il est souvent utile d'introduire des espaces fractionnaires, dont la définition va ressembler à celle des espaces de Hölder. Il y a deux définitions "naturelles", qui malheureusement correspondent à deux familles d'espaces distincts. Ces espaces sont connus sous différents noms dans la littérature. Dans un cas, la définition fait intervenir des transformées de Fourier ; nous supposerons alors que Ω est $\mathbb{R}^n$ tout entier, même si on peut généraliser à des ouverts quelconques par une procédure de localisation.

Définition I-75 (espaces de Sobolev fractionnaires). *(i) Soient Ω un ouvert de $\mathbb{R}^n$, $p \in [1, +\infty[$ et s un nombre réel positif, $s \notin \mathbb{N}$, que l'on décompose en sa partie entière k, et sa partie fractionnaire α. On appelle espace de Sobolev fractionnaire $W^{s,p}(\Omega)$ la complétion de $C_c^\infty(\Omega)$ pour la norme*

$$\|f\|_{W^{s,p}} := \|f\|_{W^{k,p}} + \sum_{|\beta|=k} [D^\beta f]_{W^{\alpha,p}},$$

$$\|f\|_{W^{\alpha,p}} = \left(\int_{\Omega\times\Omega} \frac{|f(x) - f(y)|^p}{|x-y|^{n+\alpha p}}\, dx\, dy \right)^{1/p}.$$

L'espace ainsi obtenu est un espace de Banach.

(ii) Soient $\Omega = \mathbb{R}^n$, $p \in [1, +\infty[$ et s un nombre réel positif, $s \notin \mathbb{N}$, que l'on décompose en sa partie entière k, et sa partie fractionnaire α. On appelle espace de Sobolev fractionnaire $H^{s,p}(\mathbb{R}^n)$ la complétion de $C_c^\infty(\mathbb{R}^n)$ pour la norme

$$\|f\|_{H^{s,p}} := \|f\|_{W^{k,p}} + \sum_{|\beta|=k} [D^\beta f]_{W^{\alpha,p}},$$

$$\|f\|_{W^{\alpha,p}} = \|(-\Delta)^{\alpha/2} f\|_{L^p},$$

où l'opérateur de Laplacien fractionnaire $D^\alpha = (-\Delta)^{\alpha/2}$ est défini sur la transformée de Fourier $\widehat{f}$ de f par $\widehat{D^\alpha f}(\xi) = |\xi|^\alpha f$. L'espace ainsi obtenu est un espace de Banach.

(iii) Soient $\Omega = \mathbb{R}^n$, et s un nombre réel positif. On définit l'espace de Sobolev fractionnaire $H^s(\mathbb{R}^n)$ omme la complétion de $C_c^\infty(\mathbb{R}^n)$ pour la norme

$$\|f\|_{H^s} = \left(\int (1 + |\xi|^2)^{s/2} |\widehat{f}|^(\xi) \right),$$

où $\widehat{f}$ est la transformée de Fourier de f. L'espace ainsi obtenu est un espace de Hilbert, qui coïncide avec $W^{s,2}$ et avec $H^{s,2}$; quand s est un entier, il coïncide avec les espaces $W^{s,2}$ et H^s déjà introduits.

Remarque I-76. Selon la position de p par rapport à 2, l'espace $H^{s,p}$ est inclus dans $W^{s,p}$, ou le contraire.

4.6. Espaces de Besov et autres familles. Pour la plupart des problèmes que l'on rencontre en pratique, les espaces de Sobolev fractionnaires sont largement suffisants. Les spécialistes en revanche ont souvent besoin d'espaces plus fins, dont les espaces de Besov sont un exemple. Leur définition généralise les espaces de Sobolev fractionnaires. En fait, tous les espaces que nous avons vus : Lebesgue, Hölder, Sobolev, font partie de deux grandes familles d'espaces de Banach, les espaces de Besov et les espaces de Triebel-Lizorkin, qui comprennent aussi d'autres espaces célèbres : espaces de Lorentz, Hardy, BMO, ... On pourra consulter le passionnant ouvrage [Frazier-Jawerth-Weiss], ou le livre de référence [Triebel] pour en savoir plus sur ces familles d'espaces et leurs descriptions en termes d'ondelettes.

Définition I-77 (espace de Besov). *Soient $\Omega = \mathbb{R}^n$ $1 \le p, q < +\infty$, $s \ge 0$. On définit l'espace de Besov $B^{s,p}_q(\mathbb{R}^n)$ comme la complétion de $C^\infty_c(\mathbb{R}^n)$ pour la norme*

$$\|f\|_{B^{s,p}_q} := \|f\|_{W^{[s],p}} + \sum_{|\beta|=k} [D^\beta f]_{B^{\alpha,p}_q},$$

où l'on définit, si alpha > 0,

$$\|f\|_{B^{\alpha,p}_q} = \left(\int_{\mathbb{R}^n} \frac{dh}{|h|^{n+\alpha q}} \left(\int_{\mathbb{R}^n} |f(x+h) - f(x)|^p \, dx \right)^{q/p} \right)^{1/q},$$

et si $\alpha = 0$,

$$\|f\|_{B^{\alpha,p}_q} = \left(\int_{\mathbb{R}^n} \frac{dh}{|h|^{n+q}} \left(\int_{\mathbb{R}^n} |f(x+2h) - 2f(x+h) + f(x)|^p \, dx \right)^{q/p} \right)^{1/q}.$$

L'espace ainsi obtenu est un espace de Banach. Si $s \notin \mathbb{N}$ et $p = q$, il coïncide avec $W^{s,p}$.

Remarque I-78. (i) Il y a d'autres façons équivalentes de définir les espaces de Besov, qui font intervenir une généralisation de l'analyse de Fourier, appelée théorie de Littlewood-Paley ; pour en savoir plus on pourra consulter [Frazier-Jawerth-Weiss].

(ii) Il n'y a pas du tout consensus, dans la littérature, sur la place respective des indices p, q et s dans la notation des espaces de Besov !

4.7. Espaces à poids. On peut obtenir des variantes de tous les espaces construits à partir de normes L^p en remplaçant la mesure de Lebesgue par une autre mesure, le plus souvent absolument continue, mais qui précise par exemple le comportement des fonctions à l'infini. Ainsi, on peut introduire les espaces de Lebesgue à poids, $L^p_\kappa(\mathbb{R}^n)$, via la norme

$$\|f\|_{L^p_\kappa} = \left(\int_{\mathbb{R}^n} |f(x)|^p (1+|x|)^{\kappa p} \, dx \right)^{1/p},$$

ou les espaces de Sobolev à poids, $H^k_\kappa(\mathbb{R}^n)$, via la norme

$$\|f\|_{H^k_\kappa} = \|(1+|x|^2)^{\kappa/2} f\|_{H^k},$$

etc.

I-5. Espaces de Fréchet célèbres

Comme on l'a vu, les espaces de Fréchet apparaissent naturellement comme limites d'intersections décroissantes d'espaces de Banach. Deux situations naturelles se présentent : quand on veut imposer aux fonctions des bornes *locales* et non globales ; et quand on veut imposer une régularité infinie.

5.1. Espaces de fonctions continues ou holomorphes sur un ouvert. Pour une fonction définie sur un ouvert, il n'est pas toujours naturel d'imposer des bornes globales ; on voudra souvent que f ait une certaine régularité localement, i.e. sur tout sous-ensemble compact de l'ouvert considéré.

Le lemme topologique suivant sera souvent utilisé.

Lemme I-79 (approximation d'un ouvert par des compacts). *Soit Ω un ouvert de $\mathbb{R}^n$, alors il existe une suite $(K_j)_{j\in\mathbb{N}}$ de compacts de Ω, et une suite $(O_j)_{j\in\mathbb{N}}$ d'ouverts, telles que*

$$K_j \subset O_j \subset K_{j+1}, \qquad \Omega = \bigcup_{j\geq 0} K_j.$$

Une telle suite $(K_j)_{j\in\mathbb{N}}$ est appelée suite exhaustive de compacts de Ω. Si K est un compact arbitraire de Ω, il existe j tel que $K \subset K_j$.

Les suites exhaustives permettent de définir des espaces locaux. Citons deux exemples importants.

Définition I-80 (fonctions continues sur un ouvert). *Soit Ω un ouvert de $\mathbb{R}^n$, on définit $C(\Omega)$ comme l'espace de toutes les fonctions continues de Ω dans $\mathbb{R}$. Soit $(K_j)_{j\in\mathbb{N}}$ une suite exhaustive de compacts de Ω ; pour tout j on définit une semi-norme p_j par la formule $p_j(f) = \sup_{K_j} |f|$. La famille (p_j) est une famille séparante de semi-normes, qui munit $C(\Omega)$ d'une structure d'espace de Fréchet, indépendante du choix de la suite exhaustive.*

Définition I-81 (fonctions holomorphes sur un ouvert). *Soit Ω un ouvert de $\mathbb{C}$, on définit $H(\Omega)$ comme l'espace de toutes les fonctions holomorphes de Ω dans $\mathbb{R}$. Soit $(K_j)_{j\in\mathbb{N}}$ une suite exhaustive de compacts de Ω ; pour tout j on définit une semi-norme p_j par la formule $p_j(f) = \sup_{K_j} |f|$. La famille (p_j) est une famille séparante de semi-normes, qui munit $H(\Omega)$ d'une structure d'espace de Fréchet, indépendante du choix de la suite exhaustive.*

Remarque I-82. Comme les semi-normes p_j sont croissantes, une base de voisinages de 0 dans $C(\Omega)$ est donnée par les ensembles

$$V_j := \{f \in C(\Omega);\ \sup_{K_j} |f| < 1/j\}.$$

L'espace $C_b(\Omega)$ est très différent de l'espace $C(\Omega)$ puisqu'il ne contient que des fonctions continues bornées, alors que $C(\Omega)$ contient toutes les fonctions continues sur Ω. De même, on pourrait munir l'espace de toutes les fonctions k fois dérivables, dont la dérivée est α-Hölderienne, d'une structure d'espace de Fréchet ; l'espace ainsi construit serait beaucoup plus gros que l'espace de Banach que nous avons noté $C^{k,\alpha}$ précédemment. On pourrait noter ce dernier $C^{k,\alpha}_b$ et réserver la notation $C^{k,\alpha}$ à l'espace de Fréchet ; une autre convention consisterait à noter l'espace de Fréchet $C^{k,\alpha}_{\text{loc}}$, où le symbole “loc” est une abréviation de “local”. On peut définir de même les **espaces de Lebesgue locaux**, L^p_{loc}, les **espaces de Sobolev locaux**, $W^{k,p}_{\mathrm{loc}}$, H^s_{loc}, les espaces de Besov locaux, etc.

Faisons quelques commentaires sur ces espaces [Rudin p. 33-34]. Un ensemble $\mathcal{F}$ de fonctions de $C(\Omega)$ est borné si et seulement si il existe une suite $(M_j)_{j\in\mathbb{N}}$ de nombres positifs tels que

$$\sup_{f\in\mathcal{F}} \|f\|_{K_j} \leq M_j.$$

Il est clair que les V_j ne sont pas bornés, ce qui prouve que $C(\Omega)$ n'est pas normable.

En ce qui concerne $H(\Omega)$, qui est un sous-espace fermé de $C(\Omega)$, il n'est pas si évident que les ensembles $V_j \cap H(\Omega)$ ne soient pas bornés. Cependant, un autre argument permet de montrer que $H(\Omega)$ n'est pas normable : par un théorème classique d'analyse complexe, toute famille holomorphe bornée sur K_j admet une sous-suite convergeant uniformément sur K_{j-1} vers une fonction holomorphe. On en déduit que toute partie fermée et bornée de $H(\Omega)$ est compacte, autrement dit $H(\Omega)$ possède la propriété de Heine-Borel. Comme c'est un espace de dimension infinie, il n'est donc pas normable.

5.2. Espaces de fonctions très régulières.

Définition I-83 (Espace des fonctions infiniment différentiables sur un compact). *Soit K un compact de $\mathbb{R}^n$, d'intérieur non vide. On définit $\mathcal{D}_K$ comme l'espace des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$ différentiables à tout ordre, dont le support est contenu dans K. Pour tout j on définit une semi-norme p_j par la formule*

$$p_j(f) = \sup_{|\beta|=j} \sup_{x\in K} |\nabla^\beta f(x)|.$$

La famille (p_j) est une famille séparante de semi-normes, qui munit $\mathcal{D}_K$ d'une structure d'espace de Fréchet, indépendante du choix de la suite exhaustive.

Définition I-84 (Espace des fonctions infiniment différentiables). *Soit Ω un ouvert de $\mathbb{R}^n$, on définit $C^\infty(\Omega)$ comme l'espace des fonctions de Ω dans $\mathbb{R}$ différentiables à tout ordre. Soit $(K_j)_{j\in\mathbb{N}}$ une suite exhaustive de compacts de Ω ; pour tout j on définit une semi-norme p_j par la formule*

$$p_j(f) = \sup_{|\beta|=j} \sup_{x\in K_j} |\nabla^\beta f(x)|.$$

La famille (p_j) est une famille séparante de semi-normes, qui munit $C^\infty(\Omega)$ d'une structure d'espace de Fréchet, indépendante du choix de la suite exhaustive.

En utilisant le théorème d'Ascoli et un argument d'extraction diagonale, on vérifie sans peine que $\mathcal{D}_K$ et $C^\infty(\Omega)$ ont la propriété de Heine-Borel : tout ensemble fermé et borné est compact. En particulier, ces espaces ne sont pas normables.

Citons un dernier exemple où l'on se soucie non seulement de la régularité infinie de f, mais aussi de sa décroissance très rapide à l'infini ; c'est une limite d'espaces de fonctions C^k à poids. Cet espace est appelé espace des fonctions régulières à décroissance rapide, ou classe de Schwartz.

Définition I-85 (Classe de Schwartz). *On définit $\mathcal{S}(\mathbb{R}^n)$ comme l'espace des fonctions C^∞ dont toutes les dérivées décroissent à l'infini plus vite que toute puissance inverse de $|x|$. Pour tout $j \in \mathbb{N}$ on définit une semi-norme (qui se trouve être* *une norme) sur $\mathcal{S}$ par la formule*

$$p_j(f) = \sup_{|\beta|=j} \sup_{x\in\mathbb{R}^n} |(1+|x|)^j \nabla^\beta f(x)|.$$

La famille (p_j) est une famille séparante de semi-normes, qui munit $\mathcal{S}(\mathbb{R}^n)$ d'une structure d'espace de Fréchet.

I-6. E.v.t.l.c.s. célèbres

Les e.v.t.l.c.s. qui ne sont pas des espaces de Fréchet apparaissent le plus souvent comme des espaces munis de topologies faibles. Si E est un espace de Banach, on peut le munir de la topologie faible $\sigma(E, E^*)$, i.e. la topologie la plus grossière qui rende continues toutes les formes linéaires continues sur E. Cette topologie fait de E un e.v.t.l.c.s. dont le dual est toujours E^* [Rudin, Théorème 3.10], et qui n'est **pas** métrisable si E est de dimension infinie.

Pour se convaincre de la non-métrisabilité, le petit exercice qui suit est instructif. Considérons le cas où E est un espace de Banach de dimension infinie, et E^* est muni de la topologie faible ; on sait que les topologies faible et forte sur E^* sont différentes. Si ces deux topologies ont les mêmes suites convergentes, la topologie faible n'est donc pas métrisable, puisque les suites convergentes définissent la topologie dans un espace métrique. Si en revanche il existe une suite (y_k) qui converge faiblement vers 0, sans converger en norme vers 0, alors quitte à normer la famille (y_k) on peut supposer qu'elle est faite de vecteurs unitaires, et alors l'ensemble

$$A := \{y_k + k y_\ell; \quad \ell \geq k\}$$

admet y_k dans son adhérence faible, pour tout k, et donc également 0. Mais on vérifie facilement qu'aucune des suites d'éléments de A ne peut converger vers 0. Cela montre que l'adhérence de A ne coïncide pas avec son adhérence séquentielle, et donc que la topologie n'est pas métrisable.

Remarque I-86. Si E^* est séparable, les boules de E sont métrisables, ce qui en pratique est un substitut tout-à-fait acceptable à la métrisabilité de E entier.

Il est rare, dans des problèmes courants, que l'on rencontre d'autres espaces de fonctions qui ne soient pas des espaces de Fréchet ou des espaces duaux. La principale exception est l'espace $\mathcal{D}(\Omega)$ des **fonctions C^∞ à support compact dans l'ouvert** Ω. On peut le munir d'une topologie qui en fait un e.v.t.l.c.s. complet, dont le dual est l'espace $\mathcal{D}'(\Omega)$ des **distributions sur** Ω. L'espace $\mathcal{D}(\Omega)$ ainsi construit n'est pas métrisable, et en particulier n'est pas un espace de Fréchet.

La théorie des distributions, née dans les années 50, a fourni une motivation majeure à l'étude des e.v.t.l.c.s. abstraits ; jusque là, la quasi-totalité de l'analyse fonctionnelle traitait d'espaces de Banach. L'espace des distributions est extrêmement grand, et contient **tous** les espaces fonctionnels que nous avons mentionnés jusqu'à présent et leur dual, y compris les espaces à poids et les espaces locaux.

Références

Pour un non-spécialiste, les références incontournables sont le livre de **W. Rudin**, *Functional Analysis* (McGraw-Hill, New York, 1991, 2e édition), et le livre de **H. Brézis**, *Analyse fonctionnelle, Théorie et applications* (Masson, Paris, 1992, 3e tirage). Le livre de Brézis ne traite que des espaces de Banach les plus classiques, mais contient de nombreuses références. Un autre ouvrage classique, très complet mais plus ancien, est le livre de **N. Dunford et J.T. Schwartz**, *Linear Operators* (Interscience Publishers Inc., New York, 1958).

Malheureusement, tous ces ouvrages utilisent sans hésitation la forme forte de l'axiome du choix.

Il existe de nombreux ouvrages sur les espaces fonctionnels “sophistiqués” tels que les espaces de Besov, Hardy, BMO, etc. Une référence très complète est le livre de **H. Triebel**, *Theory of function spaces* (Birkhäuser, 1992). Une référence enthousiasmante est l'opuscule de **M. Frazier, B. Jawerth et G. Weiss**, *Littlewood-Paley theory and the study of function spaces* (CBMS Regional Conference Series in Mathematics, 79, American Mathematical Society, Providence, 1991).

CHAPITRE II

Interpolation

Les espaces de Banach constituent le cadre privilégié de la majeure partie des analystes ; en particulier, la norme permet de quantifier la “taille” des éléments, et de procéder à des estimations.

L'interpolation entre espaces de Banach fait partie de la trousse à outils développée pour faciliter les calculs dans ce contexte. Imaginée pour la première fois par M. Riesz en 1926, la théorie de l'interpolation a pris son essor en 1939, quand Thorin d'une part, Marcinkiewicz d'autre part, mirent au point les démonstrations des deux théorèmes emblématiques de l'interpolation ; ces théorèmes allaient ouvrir la voie, l'un à l'interpolation complexe, et l'autre à l'interpolation réelle, méthodes développées principalement dans les années 50 et 60 par Stein, Zygmund, Calderón, Lions, Peetre, et qui font aujourd'hui partie du bagage courant des analystes.

Sommaire

II-1. Introduction

1.1. Motivations. Le but principal de l'interpolation est de construire des recettes permettant des court-circuits dans des estimations fastidieuses, qui font intervenir soit des opérateurs linéaires compliqués (transformée de Fourier, opérateurs solutions de certaines équations aux dérivées partielles...), soit des espaces compliqués (par exemple espaces de Banach à valeurs vectorielles, tels que $W^{1,p}(\mathbb{R}^n; L^q(\mathbb{R}^k))$). L'idée essentielle est que l'on peut obtenir des renseignements sur des espaces ou des opérateurs “intermédiaires” en fonction de renseignements sur des espaces ou opérateurs “extrémaux” ; pour comprendre ce principe par une analogie élémentaire, remarquons que l'on peut estimer les valeurs prises par une fonction croissante sur un segment $[a, b] \subset \mathbb{R}$, par la seule connaissance des valeurs de cette fonction en a et b.

Nous ne parlerons ici que d'interpolation **linéaire**, théorie la plus développée et la plus utilisée. Il existe cependant des techniques d'interpolation multilinéaire, et même des techniques d'interpolation encore plus générale, cependant d'usage peu commode.

L'interpolation ne doit pas être considérée comme une panacée : c'est une technique très commode et assez universelle, mais quelque peu "molle" ; il est rare que l'on arrive à des résultats optimaux (en particulier au niveau des constantes intervenant dans des inégalités de continuité) par cette technique. Nous verrons quelques exemples confirmant cette règle.

1.2. Définitions. Il est intuitif que les espaces $L^p(\Omega)$ de Lebesgue pour $p_0 \leq p \leq p_1$ se situent "entre L^{p_0} et L^{p_1}". Dans le cas où Ω a une mesure finie, c'est clair puisque les espaces L^p forment une famille croissante en p ; mais comment formaliser cette idée dans le cas général ? La même question se pose très souvent de manière naturelle quand on considère des familles d'espaces de Banach dépendant d'un ou de plusieurs paramètres réels.

Il s'avère que l'on peut, de manière très générale, définir des espaces de Banach situés "entre" deux espaces extrêmes arbitraires, même s'ils ne font pas a priori partie d'une famille paramétrée.

Définition II-87 (couple d'interpolation). *On dit que deux espaces de Banach X et Y forment un couple d'interpolation si tous deux s'injectent continûment dans un même espace vectoriel topologique (séparé) V.*

Exemple II-88. Tous les espaces de Banach célèbres que nous avons rencontrés dans le chapitre précédent s'injectent continûment dans l'espace des distributions, que nous étudierons au chapitre suivant.

L'espace V apparaissant dans la Définition II-87 importe peu ; son existence sert uniquement à garantir la possibilité de définir l'espace somme $X + Y$, cadre naturel de la théorie de l'interpolation. Une conséquence de la proposition suivante est que l'on peut toujours, sans perte de généralité, choisir $V = X + Y$.

Proposition II-89 (espaces intersection et somme). *Soient X et Y des espaces de Banach formant un couple d'interpolation. Alors $X \cap Y$ et $X + Y$ sont des espaces de Banach si on les munit des normes respectives*

$$\|v\|_{X\cap Y} = \max(\|v\|_X, \|v\|_Y),$$

$$\|v\|_{X+Y} = \inf_{x+y=v} (\|x\|_X + \|y\|_Y).$$

De plus $X \cap Y$ s'injecte continûment dans X et Y, qui eux-mêmes s'injectent continûment dans $X + Y$.

Démonstration. Il est facile de vérifier que les espaces ainsi définis sont bien des espaces vectoriels normés ; reste à montrer leur complétude. Pour l'espace $X \cap Y$, c'est clair : l'intersection de deux espaces complets est un espace complet. En ce qui concerne l'espace $X + Y$, on peut remarquer qu'il s'identifie à $X \times Y/D$, où

$D = \{(x,-x);\ x \in X \cap Y\}$ est fermé car image réciproque du fermé $\{0\}$ par l'application continue "addition". Enfin, il est clair que

$$\|x\|_{X+Y} \le \|x\|_X \le \|x\|_{X\cap Y}, \qquad \|y\|_{X+Y} \le \|y\|_Y \le \|y\|_{X\cap Y},$$

ce qui établit les injections annoncées. □

Définition II-90 (espace intermédiaire). *Soient X et Y deux espaces de Banach formant un couple d'interpolation. On appelle espace intermédiaire entre X et Y tout espace de Banach E tel que*

$$X \cap Y \subset E \subset X + Y,$$

avec injections continues.

Exemple II-91. On peut montrer (exercice) que tous les espaces L^p $(1 \le p \le \infty)$ sont des espaces intermédiaires entre L^1 et L^∞.

Si E est un espace intermédiaire entre X et Y, alors $x \in E$ dès que $x \in X$ et $x \in Y$, ce qui justifie l'appellation d'"espace intermédiaire". Quand on établit qu'un espace de Banach est intermédiaire entre X et Y, on cherche toujours à obtenir explicitement une estimation de la forme

$$\|v\|_E \le F(\|v\|_X, \|v\|_Y);$$

le cas le plus courant en pratique est celui où une fonction de la forme $F(a,b) = Ca^{1-\theta}b^\theta$ $(C \ge 0, \theta \in (0,1))$ convient.

En théorie de l'interpolation, on cherche à obtenir des renseignements non seulement sur les normes des vecteurs, mais également sur les normes des applications linéaires. Cela justifie la définition suivante.

Définition II-92 (espace d'interpolation). *Soient X et Y deux espaces de Banach formant un couple d'interpolation, et soit E un espace intermédiaire entre X et Y. On dit que E est un espace d'interpolation entre X et Y si toute application linéaire de $X+Y$ dans $X+Y$, continue de X dans X et de Y dans Y, est automatiquement continue de E dans E.*

En d'autres termes, pour toute application linéaire $L : X+Y \to X+Y$,

$$\|L\|_{X\to X} < +\infty, \quad \|L\|_{Y\to Y} < +\infty \quad \Longrightarrow \|L\|_{E\to E} < +\infty.$$

On a utilisé ici la notation

$$\|L\|_{A\to B} := \sup_{\|x\|_A \le 1} \|Lx\|_B$$

pour désigner la norme de l'application linéaire L en tant qu'application linéaire continue entre A et B.

Là encore, quand on établit qu'un espace E est un espace d'interpolation entre X et Y, on cherche à obtenir une estimation explicite de la forme

$$\|L\|_{E\to E} \le F(\|L\|_{X\to X}, \|L\|_{Y\to Y}).$$

Si le concept d'espace d'interpolation est utile pour classifier les espaces de Banach de manière "intrinsèque", un concept beaucoup plus utile en pratique est celui de **méthode d'interpolation**.

Définition II-93 (méthode d'interpolation). *On appelle méthode d'interpolation une application qui à un couple d'interpolation* (X, Y) *associe un espace de Banach* $I(E, F)$*, et possédant la propriété caractéristique suivante : pour tous couples d'interpolations* (X, Y) *et* (X', Y')*, et pour toute application linéaire continue* $L : X + Y \to X' + Y'$*, on a*

$$\|L\|_{X \to X'} < +\infty, \quad \|L\|_{Y \to Y'} < +\infty \quad \Longrightarrow \|L\|_{I(X,Y) \to I(X',Y')} < +\infty.$$

Le principal objet de la théorie de l'interpolation consiste à construire des méthodes d'interpolation aussi commodes et générales que possible, et à établir des estimations de la forme

$$\|L\|_{I(X,Y) \to I(X',Y')} \leq F(\|L\|_{X \to X'}, \|L\|_{Y \to Y'}),$$

aussi explicites que possible.

Il existe deux grandes méthodes populaires d'interpolation : l'**interpolation réelle** et l'**interpolation complexe**. Toutes deux font intervenir des paramètres auxiliaires qui permettent de classifier les espaces d'interpolation qu'elles construisent. Nous commencerons par exposer l'interpolation complexe, dont le principe est plus facile d'accès, même si la formalisation en est plus délicate.

II-2. Interpolation complexe

Cette méthode, aussi appelée méthode de Calderón-Zygmund, a été mise au point dans les années 50 et 60, le formalisme définitif étant établi par Caldéron en 1964. Basée sur le principe du maximum pour les fonctions holomorphes, la méthode complexe prend ses racines dans la preuve que Thorin (1939) donna d'un théorème d'interpolation célèbre dû à Riesz (1926). Ce théorème, appelé **théorème de Riesz-Thorin**, est encore aujourd'hui le chapitre de l'interpolation complexe qu'il est le plus important de connaître.

2.1. Théorème de Riesz-Thorin.

Théorème II-94 (théorème d'interpolation de Riesz-Thorin). *Soient* X *et* Y *deux espaces mesurés* σ*-finis, et* p_0*,* p_1*,* q_0*,* q_1 *des exposants compris entre 1 et* ∞ *au sens large. Soit* T *un opérateur linéaire défini de* $L^{p_0}(X) + L^{q_0}(Y)$ *dans* $L^{p_1}(X) + L^{q_1}(Y)$*, tel que* T *est borné* $L^{p_0}(X) \to L^{p_1}(Y)$*, et* $L^{q_0}(X) \to L^{q_1}(Y)$*. Alors, pour tout* θ*,*

$$T \text{ est borné } L^{p_\theta}(X) \to L^{q_\theta}(Y),$$

où

$$\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \qquad \frac{1}{q_\theta} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

En outre, si on pose $M_\theta = \|T\|_{L^{p_\theta} \to L^{q_\theta}}$*, alors*

$$M_\theta \leq M_0^{1-\theta} M_1^\theta.$$

Corollaire II-95 (les espaces de Lebesgue sont en interpolation). *Soient* X *et* Y *deux espaces mesurés* σ*-finis, et* p, q *des nombres réels tels que* $1 \leq p \leq q \leq \infty$*. Soit* $T : L^p + L^q \to L^p + L^q$ *un opérateur linéaire, borné* $L^p \to L^p$ *et* $L^q \to L^q$*, alors* T *est borné* $L^r \to L^r$ *pour tout* $r \in [p, q]$*.*

La preuve repose sur deux arguments d'analyse réelle et un argument d'analyse complexe :

1/ la **densité des fonctions simples** dans les espaces de Lebesgue. On rappelle que, un espace X étant σ-fini, on peut approcher toute fonction f dans L^p par une suite de fonctions simples supportées par des ensembles de mesure finie. De telles fonctions sont dans L^r pour tout r.

2/ la **dualité entre espaces** L^p pour se ramener à des problèmes scalaires.

3/ le **principe du maximum** pour les fonctions holomorphes : ces fonctions atteignent leur maximum (en module) sur le bord ; pour les estimations obtenues par interpolation ce sera pareil. C'est cette *rigidité des fonctions holomorphes* qui est cruciale dans l'interpolation complexe. Elle se traduit par le lemme suivant :

Lemme II-96 (Lemme des trois lignes). *Soit $S := \{x+iy;\ x \in [0,1];\ y \in \mathbb{R}\} \subset \mathbb{C}$ une bande du plan complexe, et soit $f : S \to \mathbb{C}$ une fonction continue bornée, holomorphe dans l'intérieur de S. Alors,*

(i) $\sup_S |f| = \sup_{\partial S} |f|$;

(ii) soit $M_\theta := \sup_{y\in\mathbb{R}} |f(\theta+iy)|$; *alors*

$$M_\theta \le M_1^\theta M_0^{1-\theta}.$$

Démonstration. 1. Supposons d'abord que f a pour limite 0 à l'infini, et soit $\varepsilon < \|f\|_\infty$; puisque f tend vers 0 à l'infini, il existe $M \in \mathbb{R}$ tel que $|f|$ (vu comme une fonction sur $\mathbb{R}^2$) atteint son maximum sur $[0,1] \times [-M, M]$. On conclut la preuve de (i) en appliquant le principe du maximum pour les fonctions holomorphes définies sur des ouverts bornés.

2. Dans le cas général où f ne converge pas forcément vers 0, on s'y ramène en considérant z_0 tel que $|f(z_0)| \ge (1-\delta)\|f\|_\infty$ et en posant $g(z) = e^{-\lambda(z-z_0)^2} f(z)$, $\lambda > 0$. En appliquant le résultat précédent, on voit que $|g(z)|$ atteint son maximum sur le bord ; or ce maximum est au moins $|g(z_0)| \ge (1-\delta)\|f\|_\infty$. En particulier,

$$\sup_{\partial S} |f| \ge (1-\delta)\|f\|_\infty,$$

et on conclut (i) en faisant tendre δ vers 0.

3. L'énoncé (ii) est obtenu à partir de (i) en posant $h(z) = e^{-\lambda z} f(z)$, $\lambda \in \mathbb{R}$. Alors

$$M_\theta \le e^{\lambda\theta} \sup_S |h| \le e^{\lambda\theta} \sup_{\partial S} |h| \le e^{\lambda\theta} \max(M_0, e^{-\lambda} M_1).$$

On choisit λ de sorte que

$$M_0 = e^{-\lambda} M_1,$$

i.e. $e^\lambda = M_1/M_0$. L'estimation ci-dessus devient alors

$$M_\theta \le M_1^\theta M_0^{1-\theta}.$$

□

Preuve du Théorème de Riesz-Thorin. On note $p = p_\theta$, $q = q_\theta$; et $M_j = \|T\|_{L^{p_j} \to L^{q_j}}$. On va utiliser la formule de dualité

$$\|f\|_{L^q} = \sup_{g \in L^{q'}} \left(\frac{1}{\|g\|_{L^{q'}}} \int fg \right)$$

pour se ramener à une formulation bilinéaire de la continuité de T entre L^p et L^q. Si $q > 1$, cette formule est conséquence de ce que $L^q = (L^{q'})'$; dans le cas $q = 1$ on la démontre directement en posant $g = f/|f|$.

Appelons fonction simple une combinaison linéaire de fonctions indicatrices d'ensembles mesurables de mesure finie ; ces fonctions sont dans tous les espaces L^p. On sait que toute fonction L^p est limite de fonctions simples au sens de la norme L^p si $p < \infty$. En traitant à part le cas $p = \infty$, on voit que

$$\|f\|_{L^q} = \sup_{g \in L^{q'};\ g \text{ simple}} \left(\frac{1}{\|g\|_{L^{q'}}} \int fg \right).$$

Montrer que T est borné $L^p \to L^q$ avec norme au plus $M_1^\theta M_0^{1-\theta}$, revient à prouver que

$$\|Tf\|_{L^q} \le M_1^\theta M_0^{1-\theta} \|f\|_{L^p} \tag{2}$$

pour toute fonction $f \in L^p$. Encore une fois, par densité et en traitant à part le cas $p = \infty$, on voit qu'il suffit d'établir (2) dans le cas où f est une fonction simple. En combinant cela avec le raisonnement de dualité vu précédemment, on conclut que notre but est de prouver

$$\left| \int (Tf) g \right| \le M_1^\theta M_0^{1-\theta} \|f\|_{L^p} \|g\|_{L^{q'}}. \tag{3}$$

C'est ici qu'intervient l'idée-clé du Théorème de Riesz-Thorin : on va introduire un paramètre d'interpolation $z \in S$, et faire varier toutes les quantités ci-dessus en fonction de z. Etant données deux fonctions simples f et g, on pose donc

$$f_z(x) = |f(x)|^{p\left(\frac{1-z}{p_0} + \frac{z}{p_1}\right)} \frac{f(x)}{|f(x)|},$$

$$g_z(y) = |g(y)|^{q'\left(\frac{1-z}{q'_0} + \frac{z}{q'_1}\right)} \frac{g(y)}{|g(y)|},$$

avec la convention $0/0 = 0$. Ces fonctions f_z sont simples, en particulier dans tous les L^r, et il s'ensuit que $Tf_z \in L^{q_0} \cap L^{q_1}$ pour tout z ; la fonction

$$\varphi : z \longmapsto \int (Tf_z) g_z$$

est donc bien définie. En décomposant f_z et g_z en combinaison linéaire de fonctions indicatrices, on voit qu'en fait on peut écrire φ sous la forme

$$\varphi(z) = \sum_{1 \le k \le K} a_k^{\lambda_k z + \mu_k}, \qquad \lambda \in \mathbb{R},\ \mu \in \mathbb{R};$$

en particulier φ est holomorphe et bornée dans S, et on peut appliquer le lemme des trois lignes :

$$|\varphi(\theta)| \le \left(\sup_{t \in \mathbb{R}} |\varphi(it)|^{1-\theta} \right) \left(\sup_{t \in \mathbb{R}} |\varphi(1+it)|^{1-\theta} \right).$$

Mais $\varphi(\theta)$ n'est autre que $\int Tfg$. Par ailleurs,

$$\left| \int Tf_{it}\, g_{it} \right| \le \|Tf_{it}\|_{L^{q_0}} \|g_{it}\|_{L^{q'_0}} \le \|T\|_{L^{p_0} \to L^{q_0}} \|f\|_{L^p}^{p/p_0} \|g\|_{L^{q'}}^{q'/q'_0},$$

et l'on peut faire une majoration similaire pour les $z = 1 + it$. On conclut facilement en remettant le tout en ordre. $\square$

Une variante intéressante, où l'on s'autorise une dépendance de l'opérateur, est la suivante. Convenons qu'une famille (T_z) définit une famille holomorphe d'opérateurs si la fonction $z \longmapsto T_z f$ est holomorphe pour tout f simple. On peut alors changer, dans l'énoncé du Théorème de Riesz-Thorin, l'opérateur T en une famille holomorphes d'opérateurs T_z ; l'hypothèse de bornes $L^{p_0} \to L^{q_0}$ et $L^{p_1} \to L^{q_1}$ sur T est alors remplacée par une hypothèse similaire sur T_0 et T_1 respectivement.

Théorème II-97 (théorème d'interpolation de Stein). *Soient X et Y deux espaces mesurés σ-finis, et p_0, p_1, q_0, q_1 des exposants compris entre 1 et ∞ au sens large. Soit $(T_z)_{z\in D}$ une famille holomorphe d'opérateurs linéaires définis sur une partie D du plan complexe incluant S. On suppose que T_0 est borné de $L^{p_0}(X)+L^{q_0}(Y)$ dans $L^{p_1}(X)+L^{q_1}(Y)$, tel que T_1 est borné $L^{p_0}(X) \to L^{p_1}(Y)$, et $L^{q_0}(X) \to L^{q_1}(Y)$. Alors,*

$$T_\theta \text{ est borné } L^p(X) \to L^q(Y),$$

En outre, si on pose $M_\theta = \|T_z\|_{L^{p_\theta}\to L^{q_\theta}}$, alors

$$M_\theta \leq M_0^{1-\theta} M_1^\theta.$$

Exemple II-98. Soit μ une mesure et w une fonction positive ; la famille d'opérateurs

$$T_z : f \longmapsto w^z f$$

satisfait aux hypothèses du théorème. Le théorème d'interpolation de Stein devient alors un théorème d'interpolation **entre espaces de Lebesgue à poids**. Par exemple, si v est une fonction positive et si l'on définit

$$\|f\|_{L^p_\kappa} = \|f v^\kappa\|_{L^p},$$

alors on a, pour tout opérateur linéaire S,

$$\|S\|_{L^{p_\theta}_{\kappa_\theta} \to L^{q_\theta}_{\lambda_\theta}} \leq \|S\|^{1-\theta}_{L^{p_0}_{\kappa_0} \to L^{q_0}_{\lambda_0}} \|S\|^{\theta}_{L^{p_1}_{\kappa_1} \to L^{q_1}_{\lambda_1}}.$$

2.2. Interpolation complexe abstraite. Dans la preuve du Théorème de Riesz-Thorin, les espaces considérés L^{p_0} et L^{p_1} formaient un couple d'interpolation, puisque tous deux s'injectent continûment dans $L^{p_0}+L^{p_1}$. On aimerait définir une procédure abstraite qui généralise l'énoncé du théorème. C'est exactement ce que fait l'interpolation complexe. Il n'est pas très important en pratique d'en connaître la définition précise, le plus souvent on l'applique à des espaces "bien connus", pour lesquels les espaces d'interpolation se trouvent dans les ouvrages de référence. Notons une difficulté technique importante : dans un cadre abstrait, on n'a pas forcément de fonctions simples à disposition. En revanche, le principe du maximum s'applique sans problème pour des fonctions holomorphes à valeurs dans un espace de Banach. Rappelons-en la définition :

Définition II-99 (fonction holomorphe à valeurs vectorielles). *Soient S un ouvert de $\mathbb{C}$ et E un $\mathbb{C}$-espace de Banach. On dit que f est holomorphe de S dans E si pour toute forme linéaire $\xi : E \to \mathbb{C}$, la fonction $z \longmapsto \langle \xi, f(z) \langle$ est holomorphe.*

En utilisant le théorème de Hahn-Banach et le principe du maximum classique, on montre facilement que la norme d'une fonction holomorphe bornée asdfde S dans E ne peut atteindre son maximum à l'intérieur de S.

Pour $\theta \in]0,1[$, l'espace d'interpolation $[X,Y]_\theta$ sera défini comme l'espace des traces en θ de l'ensemble des fonctions holomorphes à valeurs dans $X+Y$:

Définition II-100 (espace interpolé). *Soient* (X, Y) *un couple d'interpolation, et soit* $\mathcal{F}(X, Y)$ *l'espace de toutes les fonctions* $f : S \to X + Y$*, holomorphes dans* S *et continues sur* $\overline{S}$*, telles que* $t \longmapsto f(it) \in C(\mathbb{R}, X)$ *et* $t \longmapsto f(1+it) \in C(\mathbb{R}, Y)$*, muni de la norme* $\|f\|_{\mathcal{F}} = \sup_{\partial S} |f|$*. On définit* $[X, Y]_\theta$ *comme l'ensemble de tous les* $f(\theta)$*, où* f *décrit* $\mathcal{F}(X, Y)$*. On le munit de la norme*

$$\|a\|_{[X,Y]_\theta} := \inf_{f(\theta)=a} \|f\|_{\mathcal{F}(X,Y)}.$$

En utilisant le principe du maximum, on montre que $\mathcal{F}(X, Y)$ est un espace de Banach, et que $[X, Y]_\theta$, quotient de $\mathcal{F}(X, Y)$ par l'espace fermé des fontions de $\mathcal{F}(X, Y)$ qui s'annulent en θ, est un espace de Banach également. En outre, les fonctions de $\mathcal{F}(X, Y)$ qui tendent vers 0 à l'infini forment une partie dense de $\mathcal{F}(X, Y)$ tout entier (ainsi que l'ensemble des fonctions $\exp(\alpha z^2 + \beta)h$, $h \in X \cap Y$). Le principe du maximum permet de prouver que $[X, Y]_\theta \subset X + Y$, soit

$$X \cap Y \subset [X, Y]_\theta \subset X + Y.$$

Parmi les propriétés élémentaires de l'interpolation, on note $[X, Y]_\theta = [Y, X]_{1-\theta}$; et $[X, X]_\theta = X$ (démonstration : principe du maximum encore).

Le théorème principal dans le sujet est l'analogue du théorème de Riesz-Thorin :

Théorème II-101 (Théorème d'interpolation complexe). *Soient* (X_0, X_1) *et* (Y_0, Y_1) *deux couples d'interpolation. On suppose que* T *est borné de* X_0 *dans* Y_0 *et de* X_1 *dans* Y_1*. Alors, pour tout* $\theta \in [0, 1]$*,* T *est automatiquement borné de* $[X_0, X_1]_\theta$ *dans* $[Y_0, Y_1]_\theta$*, et*

$$\|T\|_{[X_0,X_1]_\theta \to [Y_0,Y_1]_\theta} \leq \|T\|^{1-\theta}_{X_0 \to Y_0} \|T\|^{\theta}_{X_1 \to Y_1}.$$

Corollaire II-102 (estimation de norme intermédiaire). *Soient* Y_0 *et* Y_1 *deux espaces de Banach formant un couple d'interpolation. Alors, pour tout* $z \in Y_0 \cap Y_1$ *on a*

$$\|z\|_{[Y_0,Y_1]_\theta} \leq \|z\|^{1-\theta}_{Y_0} \|z\|^{\theta}_{Y_1}.$$

Ce corollaire est un cas particulier du théorème précédent, dans lequel on a posé $X_0 = X_1 = \mathbb{R}$, fixé z et défini $T(t) = tz$; la norme de cette application est la norme du vecteur z lui-même.

On conclut cette section en énonçant la variante du théorème d'interpolation complexe concernant une famille d'opérateurs dépendant holomorphiquement d'un paramètre complexe.

Définition II-103 (dépendance holomorphe). *Soient* S *un ouvert de* $\mathbb{C}$*,* E *et* F *deux espaces de Banach. On dit qu'une famille* $(T_z)_{z \in S}$ *d'opérateurs linéaires continus de* E *dans* F *dépend de manière holomorphe du paramètre* z *si, pour tous* $x \in E$ *et* $\xi \in F^*$*,*

$$z \longmapsto \langle \xi, T_z x \rangle \quad \text{est holomorphe.}$$

Théorème II-104 (Théorème d'interpolation complexe, variante). *Soit* S *la bande ouverte* $\{x + iy;\ 0 < x < 1\} \subset \mathbb{C}$*, et soit* $z \longmapsto T_z$ *une famille d'opérateurs linéaires, dépendant de* z *de manière holomorphe. Soient* (X_0, X_1) *et* (Y_0, Y_1) *deux couples d'interpolation. On suppose que pour tout* $t \in \mathbb{R}$*,*

$$\|T_{it}\|_{X_0 \to Y_0} \leq M_0; \qquad \|T_{1+it}\|_{X_1 \to Y_1} \leq M_1.$$

Alors, pour tout $\theta \in (0, 1)$*,* $t \in \mathbb{R}$*,*

$$\|T_{\theta+it}\|_{[X_0,X_1]_\theta \to [Y_0,Y_1]_\theta} \leq M_0^{1-\theta} M_1^{\theta}.$$

2.3. Exemples d'espaces interpolés. La définition précédente d'espace interpolé n'est guère explicite, et un travail essentiel qui reste à faire en pratique consiste à représenter l'espace interpolé de manière simple. Voici deux exemples :

- Soit L^p l'espace de Lebesgue d'ordre p sur un espace σ-fini ; alors

$$[L^{p_0}, L^{p_1}]_\theta = L^p \qquad \frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1} \qquad (1 \le p_0, p_1 \le \infty).$$

- Soit $H^{s,p}$ l'espace de Sobolev défini par transformée de Fourier, sur $\mathbb{R}^n$ pour simplifier ; alors

$$[H^{s_0,p_0}, H^{s_1,p_1}]_\theta = H^{s,p}, \qquad s = (1-\theta)s_0 + \theta s_1, \quad \frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1} \qquad (1 < p_0, p_1 < \infty, s_0, s_1 > 0).$$

- Les espaces de Lorentz en revanche ne se déduisent pas les uns des autres par interpolation complexe. Une conséquence notable, l'un des pièges de la théorie de l'interpolation complexe, est que l'on ne peut pas faire d'interpolation à poids (interpolation de Stein) entre espaces de Lorentz.

2.4. Applications. L'application la plus célèbre de la méthode complexe est le théorème de Hausdorff-Young sur la continuité de la transformée de Fourier. C'est d'ailleurs une variante de ce théorème qui constituait la motivation première du théorème de Riesz-Thorin. Nous avons choisi pour définir la transformée de Fourier la convention qui permet d'énoncer le théorème de la manière la plus naturelle.

Théorème II-105 (Hausdorff-Young). *Soit $\mathcal{F}$ l'opérateur de transformée de Fourier :*

$$\mathcal{F}f(\xi) = \widehat{f}(\xi) := \int_{\mathbb{R}^n} e^{-i2\pi x\cdot\xi} f(x)\, dx.$$

Alors, pour tout $p \in [1,2]$, l'opérateur $\mathcal{F}$ est borné de $L^p(\mathbb{R}^n)$ dans $L^{p'}(\mathbb{R}^n)$ ($p' = p/(p-1)$), et

$$\|\mathcal{F}f\|_{L^{p'}} \le \|f\|_{L^p}.$$

Remarques II-106. (i) Cette inégalité n'est **pas** optimale : pour $1 < p < \infty$, on peut améliorer la constante 1. L'**inégalité de Hausdorff-Young optimale** s'écrit

$$\|\widehat{f}\|_{L^{p'}} \le C_p^n \|f\|_{L^p}, \qquad C_p := \frac{p^{1/p}}{p'^{1/p'}},$$

et la constante C_p est strictement inférieure à 1 dès que $p > 1$. Cette inégalité fait partie d'un domaine des mathématiques (l'étude des constantes optimales dans certaines inégalités intégrales sur $\mathbb{R}^n$) qui s'est développé avec vigueur dans les années 70, et qui est toujours en activité.

(ii) En revanche, les espaces obtenus par cette méthode sont optimaux : la transformée de Fourier n'est continue de L^p dans L^q que si p est compris entre 1 et 2, et $p' = p$.

(iii) Les résultats démontrés par Hausdorff et Young concernaient en fait les séries de Fourier, pour lesquelles une inégalité similaire est valide. Pour celle-ci, les constantes obtenues par interpolation sont optimales. Il est bon de remarquer également que ces résultats avaient été établis par Hausdorff en 1912 (dans le cas particulier où $p' \in 2\mathbb{N}$) et par Young en 1923 (dans le cas général) avant l'invention de l'interpolation, ce qui montre bien que l'interpolation est rarement indispensable.

D'autres applications bien connues de l'interpolation complexe concernent les inégalités de convolution de Young, le théorème de multiplicateurs de Calderón-Zygmund, les inégalités de Strichartz... A titre d'exemple, à partir des inégalités très simples

$$\|f * g\|_{L^1} \leq \|f\|_{L^1}\|g\|_{L^1},$$

$$\|f * g\|_{L^\infty} \leq \|f\|_{L^{p'}}\|g\|_{L^p},$$

et de l'identité $\nabla(f * g) = (\nabla f) * g$, on peut prouver des inégalités comme

$$\|f * g\|_{L^r} \leq \|f\|_{L^p}\|g\|_{L^q} \qquad 1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q},$$

$$\|f * g\|_{H^{\theta,r}} \leq \|f\|_{H^{\theta,p}}\|g\|_{L^q},$$

etc...

II-3. Interpolation réelle

L'interpolation réelle a été principalement développée à partir de la fin des années 50 par Lions et Peetre. Elégante et souple, elle rend de grands services, en relation avec la théorie de l'approximation. Comme l'interpolation complexe, elle s'est développée par abstraction d'un théorème célèbre et très utile : le théorème d'interpolation de Marcinkiewicz.

3.1. Théorème de Marcinkiewicz. Le théorème de Riesz-Thorin impliquait qu'un opérateur borné $L^1 \to L^1$ et $L^\infty \to L^\infty$ est automatiquement borné $L^p \to L^p$. Le théorème d'interpolation de Marcinkiewicz, légèrement plus précis, permet de traiter des cas limites où l'opérateur n'est pas borné $L^1 \to L^1$, mais seulement continu de L^1 dans l'espace dit "L^1 faible", $L^{1,\infty}$, qui est un cas limite d'espace de Lorentz. En contrepartie, les constantes que l'on obtient par cette méthode sont moins bonne que celles que l'on obtient par application du théorème de Riesz-Thorin.

Commençons par rappeler la définition de l'espace $L^{1,\infty}$:

Définition II-107 (espace $L^{1,\infty}$). *Soit (Ω, μ) un espace mesuré. On définit $L^{1,\infty}(\Omega)$ comme l'espace des fonctions f telles que*

$$\forall \alpha > 0 \mu[\{f \geq \alpha\}] \leq \frac{C}{\alpha}$$

pour une constante $C \geq 0$. Pour tout $f \in L^{1,\infty}$ on définit

$$\|f\|_{L^{1,\infty}} = \sup_{\alpha>0} \alpha\mu[f \geq \alpha].$$

Remarque II-108. Par l'inégalité de Chebyshev, l'espace L^1 est inclus dans $L^{1,\infty}$; l'inclusion est en général stricte.

Théorème II-109 (Théorème d'interpolation de Marcinkiewicz). *Soient (X, μ) et (Y, ν) des espaces mesurés σ-finis, et soit T un opérateur linéaire continu de $L^1(X)$ dans $L^{1,\infty}(Y)$ et de $L^\infty(X)$ dans $L^\infty(Y)$. Alors, pour tout $p \in]1, +\infty]$, T se prolonge continûment en un opérateur continu de $L^p(X)$ dans $L^p(Y)$. Plus précisément, il* *existe une constante numérique C ($C = e^{1/e} \leq 2$ convient) telle que pour tout $p \in [1, +\infty]$,*

$$\|T\|_{L^p \to L^p} \leq \frac{2p}{p-1} \|T\|_{L^1 \to L^{1,\infty}}^{1/p} \|T\|_{L^\infty \to L^\infty}^{1-1/p}.$$

Remarque II-110. Le théorème reste vrai pour un opérateur sous-linéaire, ce qui illustre déjà une propriété de souplesse qui faisait défaut à l'interpolation complexe.

Démonstration du théorème de Marcinkiewicz. Cette fois nous allons démontrer le théorème directement, sans passer par des fonctions simples. La preuve fait intervenir deux idées principales :

- représenter les normes des fonctions en jeu au moyen de la taille de leurs "ensembles de sur-niveau", i.e. le lieu des points où ces fonctions sont plus grandes qu'un certain paramètre t,

- décomposer la fonction en jeu en la somme de deux fonctions appartenant aux espaces que l'on interpole, où les deux fonctions sont choisies indépendamment *pour chaque valeur du paramètre*.

Ecrivons donc

$$\|Tf\|_{L^\infty} \leq M_1 \|f\|_{L^\infty}, \qquad \|Tf\|_{L^{1,\infty}} \leq M_0 \|f\|_{L^1}.$$

La deuxième inégalité se réécrit

$$\forall t > 0, \quad t\, mes\{|Tf| > t\} \leq M_0 \|f\|_{L^1}.$$

Sans perte de généralité on supposera que $M_0^{1-\theta} M_1^\theta = 1$; on peut toujours se ramener à ce cas en multipliant T par une constante convenable.

Par application du théorème de Fubini dans $\Omega \times \mathbb{R}$, on a

$$\int |f|^p = p \int_0^{+\infty} mes\{|f| > t\} t^{p-1}\, dt.$$

En particulier,

$$\int |Tf|^p = p \int_0^{+\infty} mes\{|Tf| > t\} t^{p-1}\, dt.$$

Pour tout $t \geq 0$ on écrit alors

$$f = f_1^{(t)} + f_2^{(t)}, \qquad f_1^{(t)} = f 1_{|f| \leq At}, \quad f_2^{(t)} = f_2 1_{|f| > At}.$$

La borne $L^\infty \to L^\infty$ entraîne que pour tout $t \geq 0$,

$$|Tf_1^{(t)}| \leq M_1 A t.$$

En particulier,

$$mes\{|Tf| > t\} \leq mes\{|Tf_2^{(t)}| > (1 - M_1 A)t\}.$$

En reportant cette inégalité dans la représentation de $\int |Tf|^p$, on trouve

$$\int |Tf|^p \leq p \int_0^{+\infty} mes\{|Tf_2^{(t)} > (1-M_1A)t\} t^{p-1}\, dt = (1-M_1A)^{-1} p \int_0^{+\infty} \Big((1-M_1A)t\, mes\{|Tf_2^{(t)} > (1$$

$$\leq (1-M_1A)^{-1} p M_0 \int_0^{+\infty} \|f_2^{(t)}\|_{L^1} t^{p-2}\, dt = (1-M_1A)^{-1} p M_0 \int_0^{+\infty} \int |f| 1_{|f|>At}\, t^{p-2}\, dt.$$

On applique alors Fubini et un changement de variable évident pour réécrire le dernier terme sous la forme

$$\begin{aligned}(1-M_1A)^{-1}pM_0 \int |f| \left(\int_0^{|f|/A} t^{p-2}\,dt \right) \\ = \frac{pM_0}{(p-1)(1-M_1A)A^{p-1}} \int |f|^p.\end{aligned}$$

On pose $M_1A = \lambda$, la constante apparaissant en facteur de $\int |f|^p$ est minimale pour $\lambda = 1/p'$, et vaut $c_p M_0 M_1^{p-1}$, avec

$$c_p = \frac{p^{p+1}}{(p-1)^p} = \left(p^{1/p} \frac{p}{p-1} \right)^p,$$

que l'on majore en utilisant $p^{1/p} \le e^{1/e}$. La preuve est complète. $\square$

Le théorème de Marcinkiewicz admet une version plus générale [Zygmund, tome II, chapitre XII, théorème 4.6] qui le rapproche davantage du théorème de Riesz-Thorin :

Théorème II-111 (Théorème d'interpolation de Marcinkiewicz généralisé). *Soient (X,μ) et (Y,ν) des espaces mesurés σ-finis, soient $p_0, q_0, p_1, q_1 \in [1,+\infty]$ avec $q_0 \neq q_1$, $p_0 \le q_0$, $p_1 \le q_1$. Si T est linéaire continu de $L^{p_0}(X)$ dans $L^{q_0,\infty}(Y)$ et de $L^{p_1}(X)$ dans $L^{q_1,\infty}(Y)$, alors pour tout $\theta \in (0,1)$, T se prolonge en un opérateur continu de $L^{p_\theta}(X)$ dans $L^{q_\theta}(Y)$, où*

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \qquad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

En outre, pour tout θ il existe une constante C_θ, ne dépendant que de θ, p_0, p_1, q_0, q_1, telle que

$$\|T\|_{L^{p_\theta} \to L^{q_\theta}} \le C_\theta \|T\|^{1-\theta}_{L^{p_0} \to L^{q_0,\infty}} \|T\|^{\theta}_{L^{p_1} \to L^{q_1,\infty}}.$$

La preuve est assez similaire à celle que nous avons déjà vue et nous l'admettrons.

3.2. Interpolation réelle abstraite. Un point-clé dans la démonstration du théorème de Marcinkiewicz consistait à écrire, pour *chaque* valeur de t, la décomposition

$$f = f_1 + f_2; \qquad \|f_1\|_{L^\infty} = t; \quad \|f_2\|_{L^1}\text{minimal},$$

et à contrôler la valeur de $\|f_2\|_{L^1}$ en fonction de t.

Pour généraliser cette approche à un cadre abstrait, on pourrait imaginer de décomposer un élément v de $X+Y$ en une somme de la forme $x^{(t)} + y^{(t)}$, pour chaque valeur d'un paramètre t, de sorte que (par exemple) $\|x^{(t)}\|_X / \|y^{(t)}\|_Y \simeq t$, et s'intéresser alors au comportement de la fonction $t \longmapsto \|x^{(t)}\|_X$. La solution finalement retenue est assez proche de cette idée.

Définition II-112 (fonction K de Lions-Peetre). *Soient X et Y deux espaces de Banach formant un couple d'interpolation. Pour tout $z \in X+Y$ on définit la K-fonction associée à z par*

$$\forall t > 0, \qquad K(t,z) := \inf_{z=x+y} \big(\|x\|_X + t\|y\|_Y \big),$$

où l'infimum est pris sur toutes les décompositions de z en la somme d'un élément x de X et d'un élément y de Y.

Remarques II-113. (i) Pour tout $t > 0$ fixé, la fonction que l'on minimise pour définir $K(t, z)$ est une norme sur $X+Y$, équivalente à la norme habituelle. Cependant les constantes d'équivalence ne sont bien sûr pas uniformes en t.

(ii) Il est facile de voir que pour tout z, la fonction $K(\cdot, z)$ est croissante et concave de $\mathbb{R}_+$ dans $\mathbb{R}_+$.

(iii) Cette procédure est assez proche de l'idée évoquée précédemment ; en effet, on s'attend à ce que l'infimum dans la définition de la fonction K soit obtenu quand les deux termes de la somme sont du même ordre de grandeur, soit $\|x\|_X/\|y\|_Y \simeq t$.

(iv) La définition de la fonction K fait jouer un rôle asymétrique à X et Y, de sorte que $K(t, z)$ doit implicitement être interprété comme $K_{X,Y}(t, z)$; mais finalement les espaces d'interpolation que l'on en déduira seront symétriques.

C'est le comportement de la fonction K (vue comme une fonction de t) quand $t \to 0$ et $t \to \infty$ qui va servir à définir les espaces d'interpolation ; cette méthode est appelée la K-méthode. L'idée est de comparer la fonction-K à des fonctions de référence du type t^θ, par exemple en imposant qu'une certaine puissance de $K(t, z)/t^\theta$ soit une fonction intégrable de t pour la mesure de référence dt/t sur $\mathbb{R}_+$.

Définition II-114 (espaces d'interpolation réelle). *Soient X et Y deux espaces de Banach formant un couple d'interpolation. On définit sur $\mathbb{R}_+ \times (X+Y)$ la fonction K associée à ces deux espaces. Pour tout $\theta \in (0, 1)$, on définit*

$$(X, Y)_\theta := \left\{ z \in X + Y;\ \lim_{t\to 0} \frac{K(t, z)}{t^\theta} = \lim_{t\to+\infty} \frac{K(t, z)}{t^\theta} = 0 \right\}.$$

Un paramètre auxiliaire $p \in [1, \infty]$ étant donné, on définit également

$$(X, Y)_{\theta,p} := \left\{ z \in X + Y;\ t \longmapsto \frac{K(t, z)}{t^\theta} \in L^p(dt/t) \right\}.$$

Ces espaces sont des espaces de Banach si on les munit des normes

$$\|z\|_{(X,Y)_{\theta,p}} := \left\| \frac{K(t, z)}{t^\theta} \right\|_{L^p(dt/t)};$$

$$\|z\|_{(X,Y)_\theta} := \|z\|_{(X,Y)_{\theta,\infty}} = \sup_{t>0} \frac{K(t, z)}{t^\theta}.$$

Avant de continuer, faisons brièvement le lien avec la théorie de l'approximation. Soit par exemple $z \in (X, Y)_{\theta,\infty}$; on sait alors que pour tout $t > 0$, on peut trouver $x = x^{(t)} \in X$ et $y = y^{(t)} \in Y$ tels que $\|x\|+t\|y\| = O(t^\theta)$. Nécessairement $\|x^{(t)}\| \longrightarrow 0$ quand $t \to 0$, et donc $y^{(t)} \longrightarrow z$ dans $X + Y$. En particulier, tous les éléments de l'espace $(X, Y)_{\theta,\infty}$ peuvent être approchés dans $X + Y$ par des éléments de Y ; en faisant tendre cette fois t vers l'infini, on vérifie qu'ils peuvent aussi être approchés dans $X + Y$ par des éléments de X. Ce qui compte dans la définition des espaces d'interpolation, c'est la vitesse (en fonction de t) à laquelle cette approximation peut se faire.

Proposition II-115 (Propriétés élémentaires de l'interpolation réelle). *Soient X, Y deux espaces de Banach en interpolation. Alors* *(i) Pour tous $\theta \in (0,1), p \in [1,\infty]$, on a*

$$(X,Y)_\theta = (Y,X)_{1-\theta}; \qquad (X,Y)_{\theta,p} = (Y,X)_{1-\theta,p},$$

avec égalité des normes ;

(ii) Pour tout $\theta \in (0,1), p \in [1,\infty]$, on a

$$(X,X)_\theta = (X,X)_{\theta,p} = X,$$

avec équivalence des normes ;

(iii) Dès que $\theta \in (0,1)$, et $1 \leq p_1 \leq p_2 \leq \infty$, on a

$$X \cap Y \subset (X,Y)_{\theta,p_1} \subset (X,Y)_{\theta,p_2} \subset (X,Y)_\theta \subset (X,Y)_{\theta,\infty} \subset \overline{X} \cap \overline{Y} \subset X+Y.$$

(iv) Si $Y \subset X$, alors les espaces $(X,Y)_{\theta,p}$ sont emboîtés selon un ordre lexicographique où θ est prioritaire :

$$0 < \theta_1 \leq \theta_2, \quad 1 \leq p_1, p_2 \leq \infty \quad \Longrightarrow (X,Y)_{\theta_2,p_2} \subset (X,Y)_{\theta_1,p_1}.$$

Les propriétés (i) et (ii) sont des propriétés naturelles que l'on attend de n'importe quelle théorie d'interpolation. Les propriétés (iii) et (iv) reflètent l'intuition que dans la définition des espaces $(X,Y)_{\theta,p}$, le paramètre p doit être vu comme un “raffinement” du paramètre θ.

Remarques II-116. On voit à ce stade quelques différences intéressantes avec l'interpolation complexe. Tout d'abord, dans l'énoncé (ii), la constante d'équivalence des normes n'est pas égale à 1, elle dépend de θ. Ensuite, les valeurs $\theta = 0$ et $\theta = 1$ ne sont pas autorisées, contrairement à ce qui se passe en interpolation complexe. Plus précisément, $(X,Y)_{0,p}$ est réduit à l'espace vectoriel trivial $\{0\}$ si $p < \infty$, tandis que l'espace $(X,Y)_{0,\infty}$ est “légèrement” plus grand que X.

Démonstration. Nous allons démontrer les énoncés (i) à (iii) du théorème ci-dessus, l'énoncé (iv) étant laissé en exercice.

Tout d'abord, la propriété (i) est conséquence de la relation d'homogénéité

$$K_{X,Y}(t,z) = tK_{Y,X}(t^{-1},z)$$

et du changement de variable $t \longmapsto t^{-1}$, qui laisse la mesure de référence dt/t invariante.

Pour établir la propriété (ii), on note que $X+Y = X$, et que $K_{X,X}(t,x) \leq \min(t,1)\|x\|_X$, et en déroulant les définitions on obtient l'équivalence souhaitée, avec une constante

$$\left[\int_0^{+\infty} \left(\frac{\min(t,1)^p}{t^\theta}\right)^p \frac{dt}{t}\right]^{1/p}.$$

Montrons maintenant que $(X,Y)_{\theta,p} \subset (X,Y)_{\theta,\infty}$. En partant de l'identité

$$\frac{1}{\theta^p} = (\theta p)^{1/p} \left(\int_t^{+\infty} \frac{ds}{s^{\theta p+1}}\right)^{1/p},$$

et en utilisant la croissance de la fonction $K(\cdot,z)$, on obtient que pour tout $t>0$,

$$\begin{aligned}\frac{K(t,z)}{t^\theta} &\leq (\theta p)^{1/p}\left(\int_t^{+\infty}\frac{K(t,z)^p}{s^{\theta p+1}}\,ds\right)^{1/p}\\ &\leq (\theta p)^{1/p}\left(\int_0^{+\infty}\frac{K(s,z)^p}{s^{\theta p+1}}\,ds\right)^{1/p},\end{aligned}$$

ce qui prouve le résultat.

Montrons enfin que $(X,Y)_{\theta,p_1}\subset(X,Y)_{\theta,p_2}$ dès que $p_1\leq p_2$. Pour cela on écrit

$$\begin{aligned}\|z\|_{(X,Y)_{\theta,p_2}} &\leq \left(\int_0^{+\infty}\frac{K(t,z)^{p_2}}{t^{\theta p_2+1}}\,dt\right)^{1/p_2} \leq \left(\int_0^{+\infty}\frac{K(t,z)^{p_1}}{t^{\theta p_1+1}}\,dt\right)^{1/p_2}\sup_{t>0}\left[\frac{K(t,z)}{t^\theta}\right]^{1-p_1/p_2}\\ &\leq \|z\|^{p_1/p_2}_{(X,Y)_{\theta,p_1}}\|z\|^{1-p_1/p_2}_{(X,Y)_{\theta,\infty}} \leq \|z\|_{(X,Y)_{\theta,p_1}}.\end{aligned}$$

□

Remarque II-117. En interpolation complexe, les espaces interpolés étaient définis en termes de traces de fonctions holomorphes à valeurs vectorielles. Existe-t-il une caractérisation des espaces d'interpolation réelle en termes de traces ? La réponse est affirmative : d'après un célèbre résultat de J.-L. Lions, les espaces d'interpolation réelle peuvent se définir en termes de traces de fonctions vivant dans des espaces de Sobolev à valeurs vectorielles. Cette remarque justifie encore plus la dénomination d'"interpolation réelle", au sens où la dérivée holomorphe y est remplacée par la dérivée usuelle, les espaces de fonctions holomorphes par des espaces de Sobolev.

Le résultat qui suit est l'analogue abstrait du théorème de Marcinkiewicz.

Théorème II-118 (théorème d'interpolation réelle abstraite). *Soient (X_0,X_1) et (Y_0,Y_1) deux couples d'interpolation. On suppose que T est borné de X_0 dans Y_0 et de X_1 dans Y_1. Alors, pour tous $\theta\in(0,1)$, $p\in[1,\infty]$, T admet une extension continue de $(X_0,X_1)_\theta$ dans $(Y_0,Y_1)_\theta$, et de $(X_0,X_1)_{\theta,p}$ dans $(Y_0,Y_1)_{\theta,p}$. En outre,*

$$\|T\|_{(X_0,X_1)_\theta\to(Y_0,Y_1)_\theta}\leq\|T\|^{1-\theta}_{X_0\to Y_0}\|T\|^\theta_{X_1\to Y_1}.$$
$$\|T\|_{(X_0,X_1)_{\theta,p}\to(Y_0,Y_1)_{\theta,p}}\leq\|T\|^{1-\theta}_{X_0\to Y_0}\|T\|^\theta_{X_1\to Y_1}.$$

Corollaire II-119 (estimation de normes intermédiaires par la méthode réelle). *Soient Y_0 et Y_1 deux espaces de Banach formant un couple d'interpolation. Alors, pour tous $\theta\in(0,1)$, $p\in[1,\infty]$, il existe des constantes C_θ et $C_{\theta,p}$ telles que*

$$\forall z\in Y_0\cap Y_1,\qquad \|z\|_{(Y_0,Y_1)_\theta}\leq\|z\|^{1-\theta}_{Y_0}\|z\|^\theta_{Y_1};$$
$$\forall z\in Y_0\cap Y_1,\qquad \|z\|_{(Y_0,Y_1)_{\theta,p}}\leq\|z\|^{1-\theta}_{Y_0}\|z\|^\theta_{Y_1}.$$

Démonstration. Remarquons d'abord que l'on obtient le corollaire à partir du théorème en définissant $X_0=X_1=\mathbb{R}$, en considérant l'application $T:\lambda\longmapsto\lambda z$, et en utilisant l'identité

$$(\mathbb{R},\mathbb{R})_\theta=(\mathbb{R},\mathbb{R})_{\theta,p}=\mathbb{R},$$

avec *équivalence* (et non égalité) des normes.

Pour démontrer le théorème, on écrit, pour tout $t>0$ et pour toute décomposition de $z\in X+Y$ en une somme $x+y$,

$$\|Tx\|_{Y_0}+t\|Ty\|_{Y_1}\leq\|T\|_{X_0\to Y_0}\left(\|x\|_{X_0}+t\frac{\|T\|_{X_1\to Y_1}}{\|T\|_{X_0\to Y_0}}\|y\|_{X_1}\right).$$

Par linéarité, on a $Tz = Tx + Ty$; en passant d'abord à l'infimum sur le membre de droite et en appliquant la définition de la fonction K, on trouve

$$K_{Y_0,Y_1}(t, Tz) \leq \|T\|_{X_0 \to Y_0} K_{X_0,X_1}\left(t \frac{\|T\|_{X_1 \to Y_1}}{\|T\|_{X_0 \to Y_0}},\ z\right).$$

Il s'ensuit

$$\left(\int_0^{+\infty} \frac{K(t, Tz)^p}{t^{\theta p}} \frac{dt}{t}\right)^{1/p} \leq \|T\|_{X_0 \to X_1} \left(\int_0^{+\infty} \frac{K(\lambda t, z)^p}{t^{\theta p}} \frac{dt}{t}\right)^{1/p},$$

où $\lambda := \|T\|_{X_1 \to Y_1} / \|T\|_{X_0 \to Y_0}$. On conclut en effectuant le changement de variables $t' = \lambda t$ dans l'intégrale. □

3.3. Extrait du catalogue d'interpolation réelle. La démonstration du théorème d'interpolation réelle abstraite était d'une simplicité quasi-enfantine ; cependant, il reste un important travail à accomplir avant de pouvoir appliquer ce théorème en pratique : il s'agit d'identifier les espaces interpolés, dans des situations concrètes. Les ouvrages de référence présentent des catalogues où les interpolations de toutes les familles célèbres d'espaces de Banach sont calculées. En voici quelques exemples.

- $(C(\mathbb{R}^n), C^1(\mathbb{R}^n))_{\theta,\infty} = C^\theta(\mathbb{R}^n)$: l'interpolation de classes de fonctions régulières jusqu'à un certain ordre donne des espaces de fonctions Hölderiennes ; noter la valeur du paramètre $q = \infty$. De même, si Ω est un ouvert assez régulier de $\mathbb{R}^n$, $(C(\overline{\Omega}), C^1(\overline{\Omega}))_{\theta,\infty} = C^\theta(\overline{\Omega})$.
- $(L^p(\mathbb{R}^n), W^{1,p}(\mathbb{R}^n))_{\theta,p} = W^{\theta,p}(\mathbb{R}^n)$: on obtient ainsi une classe d'espaces de Sobolev fractionnaires, différente (sauf dans le cas où $p = 2$) de la classe construite par interpolation complexe. On peut faire de même en remplaçant $\mathbb{R}^n$ par un ouvert régulier de $\mathbb{R}^n$.
- $(L^1(\Omega), L^\infty(\Omega))_{\theta,q} = L^{\frac{1}{1-\theta},q}(\Omega)$, dès que Ω est un espace mesuré σ-fini.
- $(L^p(\mathbb{R}^n), W^{1,p}(\mathbb{R}^n))_{\theta,q} = B^{\theta,p}_q(\mathbb{R}^n)$; de manière plus générale, en interpolant des espaces de Sobolev on obtient des espaces de Besov.
- On peut construire d'autres exemples par **réitération** : si on sait déjà que Z_0 et Z_1 sont des espaces intermédiaires entre X et Y, avec

  $$\|z\|_{Z_0} \leq C_0 \|z\|_X^{1-\theta_0} \|z\|_Y^{\theta_0}, \qquad \|z\|_{Z_1} \leq C_1 \|z\|_X^{1-\theta_1} \|z\|_Y^{\theta_1},$$

  alors on a

  $$(Z_0, Z_1)_{\theta,p} = (X, Y)_{(1-\theta)\theta_0 + \theta\theta_1, p}.$$

  Si Z_0 et Z_1 ont déjà été construits par interpolation, on peut ainsi "réitérer" l'interpolation. Par exemple,

  $$((X, Y)_{\theta_0,q_0}, (X, Y)_{\theta_1,q_1})_{\theta,p} = (X, Y)_{(1-\theta)\theta_0 + \theta\theta_1, p}.$$

  On peut ainsi calculer toutes les interpolations entre espaces de Lorentz.
- En revanche, même pour un ouvert Ω très régulier, $C^1(\Omega)$ n'est ni $(C(\Omega), C^2(\Omega))_\theta$, ni $(C(\Omega), C^2(\Omega))_{\theta,\infty}$. En fait on peut montrer que C^1 n'est pas un espace d'interpolation entre C et C^2. En effet, si l'on introduit l'opérateur T_ε par la formule

  $$T_\varepsilon f(x) := \int_{-1}^1 \frac{x}{x^2 + y^2 + \varepsilon} [f(y) - f(0)]\, dy,$$

alors on peut vérifier que T_ε est borné de C dans C et de C^2 dans C^2, mais pas de C^1 dans C^1, comme le montre la famille $f_\varepsilon(x) = \sqrt{x^2+\varepsilon^2}\eta(x)$, où η est une fonction régulière à support compact.

– Cependant, on a l'inégalité

$$\|f\|_{C^1} \le \sqrt{\|f\|_{C^0}\|f\|_{C^2}},$$

qui quantifie le fait que C^1 est un espace intermédiaire entre C et C^2. De ce fait, il vérifie certaines des propriétés des espaces d'interpolation.

A titre d'exemple, pour retrouver l'énoncé du théorème de Marcinkiewicz, on se donne un opérateur T borné de L^{p_0} dans L^{q_0} et de L^{p_1} dans L^{q_1}, et on note que, avec

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{1-\theta}{p_1}, \qquad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{1-\theta}{q_1},$$

on a, pour autant que $q_0 \neq q_1$, et $q_0, q_1 > 1$, les identités d'interpolation

$$(L^{p_0}, L^{p_1})_{\theta,p} = L^{p,p} = L^p;$$

$$(L^{q_0,\infty}, L^{q_1,\infty})_{\theta,p} = \Big((L^1,L^\infty)_{1-1/q_0,\infty}, (L^1,L^\infty)_{1-1/q_1,\infty}\Big)_{\theta,p} = \ldots = L^{q,p} \subset L^{q,q} = L^q.$$

On remarque en passant que l'on n'a pas retrouvé l'énoncé complet du théorème de Marcinkiewicz, puisque les valeurs $q_0 = 1$ ou $q_1 = 1$ ne sont pas admissibles dans la dernière ligne de calculs.

3.4. Un exemple d'identification d'espace interpolé. Démontrons, à titre d'exemple, que pour tout $p \in [1,\infty)$ et pour tout $\theta \in (0,1)$,

$$(L^p(\mathbb{R}^n), W^{1,p}(\mathbb{R}^n))_{\theta,p} = W^{\theta,p}(\mathbb{R}^n).$$

On définit, pour tout $f \in L^p(\mathbb{R}^n)$,

$$K(t,f) := \inf_{f=a+b} \|f\|_{L^p} + t\|b\|_{W^{1,p}};$$

le but est de montrer que sur L^p les normes

$$\|f\|_{W^{\theta,p}} := \|f\|_{L^p} + \left(\int_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|f(x)-f(y)|^p}{|x-y|^{n+\theta p}}\,dx\,dy\right)^{1/p} = \|f\|_{L^p} + \left(\int_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|f(x)-f(x+h)|^p}{|h|^{n+\theta p}}\,dx\,dh\right)^{1/p}$$

et

$$\|f\|_{\theta,p} := \left(\int_0^{+\infty} \frac{K(t,f)^p}{t^{\theta p}}\,\frac{dt}{t}\right)^{1/p}$$

sont simultanément finies ou infinies, et équivalentes. Comme les espaces L^p et $W^{1,p}$ admettent des sous-espaces denses faits de fonctions régulières, il en est de même pour leur interpolé, et il suffit de vérifier l'équivalence des normes pour les fonctions régulières.

Commençons par la partie simple : dominer la norme L^p. Pour toute décomposition de f en $a+b$, avec $a \in L^p$ et $b \in W^{1,p}$, on peut écrire

$$\|f\|_{L^p} \le \|a\|_{L^p} + \|b\|_{L^p} \le \|a\|_{L^p} + \|b\|_{W^{1,p}}.$$

En particulier, pour tout $t \ge 1$ on a

$$\|f\|_{L^p} \le K(t,f),$$

et il s'ensuit
$$\left(\int_0^{+\infty} \frac{K(t,f)^p}{t^{\theta p}} \frac{dt}{t}\right)^{1/p} \geq \|f\|_{L^p} \left(\int_1^{+\infty} \frac{dt}{t^{\theta p+1}}\right)^{1/p},$$
autrement dit
$$\|f\|_{L^p} \leq (\theta p+1)^{1/p}\|f\|_{\theta,p}.$$

On passe maintenant à l'estimation de la semi-norme définissant $W^{\theta,p}$. Pour cela nous utiliserons l'inégalité suivante : pour tout $h \in \mathbb{R}^n \setminus \{0\}$,
$$\int \left(\frac{|g(x+h)-g(x)|}{|h|}\right)^p dx \leq \int |\nabla g|^p.$$
En effet, cette intégrale vaut
$$\int \left(\frac{|\int_0^1 \nabla g(x+th)\cdot h\, dt|}{|h|}\right)^p dx \leq \int \left(\int_0^1 |\nabla g(x+th)|\, dt\right)^p dx;$$
en appliquant successivement l'inégalité de Jensen, le théorème de Fubini et le changement de variable $x \to x+th$, on majore la dernière quantité par
$$\int_{\mathbb{R}^n}\int_0^1 |\nabla g(x+th)|^p\, dt\, dx = \int_0^1\int_{\mathbb{R}^n} |\nabla g(x+th)|^p\, dx\, dt = \int_0^1\int_{\mathbb{R}^n} |\nabla g|^p\, dx\, dt = \int |\nabla g|^p.$$
Soit n'importe quelle décomposition de f en $a+b$, avec $a \in L^p$, $b \in W^{1,p}$; pour tout $h \neq 0$ on peut écrire
$$\int |f(x+h)-f(x)|^p\, dx \leq 2^{p-1}\left(\int |a(x+h)-a(x)|^p\, dx + \int |b(x+h)-b(x)|^p\, dx\right)$$
$$\leq 4^p\left(\int |a(x)|^p\, dx + |h|^p\int |\nabla b(x)|^p\, dx\right) \leq C_p(\|a\|_{L^p} + |h|\|b\|_{W^{1,p}})^p$$
(aucun effort n'a été fait pour avoir une bonne dépendance des constantes par rapport à p). En passant à l'infimum sur toutes les décompositions possibles, on obtient
$$\int |f(x+h)-f(x)|^p\, dx \leq C_p K(|h|,f)^p.$$
En intégrant par rapport à h et en effectuant un changement de coordonnées polaires, on obtient
$$\int_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|f(x)-f(x+h)|^p}{|h|^{n+\theta p}}\, dx\, dh \leq C_p\int_{\mathbb{R}^n} \frac{K(|h|,f)^p}{|h|^{n+\theta p}}\, dh = C_p|S^{n-1}|\int_0^{+\infty} \frac{t^{n-1}K(t,f)^p}{t^{n+\theta p}}\, dt,$$
et l'intégrale de droite est bien celle que l'on attendait. Cela achève la démonstration de l'inégalité $\|f\|_{W^{\theta,p}} \leq \|f\|_{\theta,p}$.

Pour montrer l'inégalité inverse, nous allons utiliser un **argument d'approximation**. Soit φ une fonction positive, d'intégrale 1, de support inclus dans la boule $B(0,1)$. Pour tout $t>0$ on définit
$$\varphi_t(x) := \frac{1}{t^n}\varphi\left(\frac{x}{t}\right).$$
Pour tout $f \in L^p$, la fonction $f*\varphi$ est très régulière, et bien sûr dans $W^{1,p}$. On écrit
$$f = (f - f*\varphi_t) + f*\varphi_t.$$

Le premier morceau converge vers f dans L^p, plus ou moins vite selon la régularité de f ; le deuxième morceau appartient à $W^{1,p}$ pour tout t fixé, mais sa norme dans cet espace explose a priori quand $t \to 0$, plus ou moins vite selon la régularité de f. L'interpolation réelle est basée sur ce principe...

Par définition de la fonction K, on a, pour tout $t > 0$,

$$K(t, f)^p \leq \|f - f * \varphi_t\|_{L^p} + t\|f * \varphi_t\|_{W^{1,p}}.$$

On va majorer séparément les deux morceaux.

Tout d'abord, de l'identité

$$f - f * \varphi_t = \int_{\mathbb{R}^n} [f(x) - f(y)]\, \varphi_t(x - y)\, dy$$

on déduit, par inégalité de Jensen,

$$|f - f * \varphi_t|^p \leq \int_{\mathbb{R}^n} |f(y) - f(x)|^p\, \varphi_t(x - y)\, dy.$$

En intégrant par rapport à $dx\, dt/t^{\theta p+1}$, on trouve

$$\int_0^{+\infty} \frac{\|f - f * \varphi_t\|_{L^p}^p}{t^{\theta p}} \frac{dt}{t} \leq \int_{\mathbb{R}^n \times \mathbb{R}^n} |f(x) - f(y)|^p \left(\int_0^{+\infty} t^{-\theta p} \varphi_t(x - y) \frac{dt}{t} \right).$$

Or

$$\varphi_t(x - y) \leq \frac{\|\varphi\|_{L^\infty}}{t^n} 1_{|x-y| \leq t},$$

d'où

$$\int_0^{+\infty} t^{-\theta p} \varphi_t(x - y) \frac{dt}{t} \leq \|\varphi\|_{L^\infty} \int_{|x-y|}^{\infty} \frac{dt}{t^{1+n+\theta p}} = \frac{C}{|x - y|^{n+\theta p}}.$$

On conclut que

$$\int_0^{+\infty} \frac{\|f - f * \varphi_t\|_{L^p}^p}{t^{\theta p}} \frac{dt}{t} \leq C \int_{\mathbb{R}^n \times \mathbb{R}^n} \frac{|f(x) - f(y)|^p}{|x - y|^{n+\theta p}},$$

ce qui achève la majoration du terme en $f - f * \varphi_t$.

Il reste à majorer

$$\int_0^{+\infty} \frac{t^p \|f * \varphi_t\|_{W^{1,p}}^p}{t^{\theta p}} \frac{dt}{t}.$$

On laisse de côté la norme L^p de $f * \varphi_t$, qui se majore facilement, et on se concentre sur le terme délicat

$$\int_0^{+\infty} \frac{t^p \|f * \nabla \varphi_t\|_{L^p}^p}{t^{\theta p}} \frac{dt}{t}.$$

On écrit alors, en tirant parti de ce que $\int \nabla \varphi = 0$,

$$f * \varphi_t(x) = \int f(x - y) \varphi_t(y)\, dy = \int [f(x - y) - f(x)]\, \nabla \varphi_t(y)\, dy.$$

En appliquant l'inégalité de Jensen avec la mesure de probabilité $|\nabla \varphi_t(y)|/\|\nabla \varphi_t\|_{L^1}$, on trouve

$$|f * \nabla \varphi_t(x)| \leq \|\nabla \varphi_t\|_{L^1}^{p-1} \int |f(x - y) - f(x)|^p\, |\nabla \varphi_t(y)|\, dy.$$

Par conséquent,

$$\|f * \nabla \varphi_t\|_{L^p}^p \leq \|\nabla \varphi_t\|_{L^1}^{p-1} \|\nabla \varphi_t\|_{L^\infty} \int_{|x-y|| \leq t} |f(x-y) - f(x)|^p\, dy\, dx \leq \frac{C}{t^{n+p}} \int_{|x-y|| \leq t} |f(x-y) - f(x)|^p\, dy\, dx$$

En intégrant par rapport à la mesure $t^p\,dt/t^{\theta p+1}$ et en appliquant le théorème de Fubini, on obtient finalement

$$\int_0^{+\infty} \frac{t^p \|f * \nabla\varphi_t\|_{L^p}^p}{t^{\theta p}} \frac{dt}{t} \leq C \int |f(y) - f(x)|^p \left(\int_0^{|x-y|} \frac{dt}{t^{\theta p+n+1}} \right) dy\,dx,$$

ce qui se réduit bien à l'expression souhaitée.

3.5. Applications. L'application la plus connue du théorème d'interpolation de Marcinkiewicz concerne le **théorème de Hardy-Littlewood** sur la fonction maximale.

Définition II-120 (fonction maximale). *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction localement intégrable sur $\mathbb{R}^n$. On définit sa fonction maximale Mf par la formule*

$$Mf(x) := \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} f(y)\,dy.$$

On remarquera que la définition de la fonction maximale se généralise sans difficulté au cas où l'espace $\mathbb{R}^n$ est remplacé par un espace métrique arbitraire, et la mesure de Lebesgue par une mesure arbitraire, localement finie. Le théorème de Hardy-Littlewood concerne l'opérateur maximal $M : f \longmapsto Mf$, et il est classique de le démontrer par application du théorème de Marcinkiewicz.

Théorème II-121 (inégalité maximale de Hardy-Littlewood). *Il existe une constante $C > 0$ telle que pour tout $p \in (1, +\infty]$,*

$$\|Mf\|_{L^p} \leq \frac{C}{p-1} \|f\|_{L^p}.$$

Le théorème d'interpolation réelle abstraite intervient dans de nombreux résultats de recherche récents, par exemple les lemmes de régularité en moyenne en théorie cinétique.

Références

La “bible” des interpolateurs est l'ouvrage de **Bergh et Löfström**, *Interpolation Spaces. An introduction* (Springer-Verlag, Berlin, 1970). Cet ouvrage peu pédagogique ne doit guère être utilisé que comme traité de référence.

Le chapitre 12 du livre d'**A. Zygmund**, *Trigonometric series* (2e ed., Cambridge University Press, New York, 1959), contient une discussion historique du théorème de Riesz-Thorin et de ses variantes.

Le cours d'**A. Lunardi**, *Interpolation Theory* (Scuola Normale Superiore, 1999) constitue une introduction pédagogique et agréable, à débarrasser toutefois de coquilles assez nombreuses.

CHAPITRE III

Distributions

La théorie des distributions, oeuvre de Laurent Schwartz, fournit aux analystes un cadre général au formalisme agréable pour étudier des espaces fonctionnels et des équations aux dérivées partielles ; elle valut la médaille Fields à son auteur dans les années 50. Depuis lors, les distributions constituent le cadre naturel dans lequel beaucoup d'analystes se placent, reprenant les notations et les idées de Schwartz.

Malgré son élégance et sa commodité, la théorie des distributions n'est indispensable ni dans le domaine des équations aux dérivées partielles ni dans celui de l'analyse fonctionnelle. D'une part, les spécialistes d'équations aux dérivées partielles parviennent toujours à trouver des formulations bien adaptées à leurs problèmes, en s'inspirant des idées sous-jacentes à la théorie des distributions, mais sans y avoir recours explicitement. D'autre part, on peut étudier la plupart des espaces fonctionnels intéressants sans la notion de distribution : par exemple, les espaces de Sobolev peuvent être définis en termes de distributions, ou bien en termes de complétion de l'espace des fonctions C^∞ à support compact pour des normes bien choisies. L'équivalence de ces deux points de vue est d'ailleurs un résultat célèbre dû à Meyers et Serrin.

En utilisant uniquement des espaces de Sobolev et leurs espaces duals, on peut obtenir la grande majorité des distributions que l'on rencontre en pratique : par exemple, la dérivée de la "fonction de Dirac" peut être vue comme un élément de $(C^1)^*$, l'espace dual de l'espace des fonctions C^1. Plus généralement, quand on rencontre une distribution dans un problème concret, c'est presque à coup sûr un élément du dual d'un espace de Sobolev bien choisi, au moins localement. La topologie des espaces de Sobolev étant beaucoup plus simple que celle des distributions, on comprend que l'étude des espaces de Sobolev soit plus populaire et certainement plus utile en pratique.

Pour ces raisons, nous ne donnerons que très peu d'éléments de démonstration des résultats principaux de la théorie des distributions. D'ailleurs, même si certains de ces théorèmes sont d'une puissance redoutable, les résultats que l'on obtient en utilisant des méthodes plus terre-à-terre sont souvent meilleurs (plus constructifs, plus quantitatifs...).

Pour autant, la théorie des distributions fournit un formalisme commode et élégant, qui apporte de l'ordre dans le paysage fonctionnel, et facilite la communication entre mathématiciens d'horizons très divers. En outre, les principes qui la sous-tendent, bien plus que les théorèmes principaux, s'avèrent d'une importance capitale en pratique. Pour toutes ces raisons, une bonne familiarité avec le langage des distributions est presque indispensable à un analyste. Schwartz lui-même avait bien conscience que le principal mérite de son approche ne résidait pas dans l'introduction d'outils nouveaux, mais dans une synthèse claire et accessible de recettes multiples qui étaient, déjà auparavant, employées dans des contextes divers.

Sommaire

III-1. Motivations et contexte

L'introduction des distributions, vers 1950, répondait à plusieurs buts.

1.1. Généralisation de la notion de fonction. Depuis le milieu des années 1920, les théoriciens de la physique quantique, et en particulier Dirac, utilisaient des objets étranges qu'ils manipulaient comme des fonctions. Le plus typique était la “fonction de Dirac”, mystérieuse fonction qui vaut 0 partout, sauf en 0 où elle vaut $+\infty$, et dont l'intégrale est égale à 1 (en violation de toutes les règles de la théorie de l'intégration de Lebesgue). Non seulement Dirac utilisait cette fonction à des fins de calcul formel, mais encore il se permettait de la dériver à volonté, se contentant de remarquer que les dérivées successives étaient “de plus en plus singulières”.

Outre leur rôle dans des problèmes de calcul formel, ces objets avaient souvent un sens physique. Ainsi, la “fonction de Dirac” peut représenter une charge électrique isolée (électron, modélisé comme une masse ponctuelle) ; sa dérivée peut représenter un “dipôle”, association de deux charges électriques ponctuelles de signes opposés, infiniment fortes et infiniment proches ; etc. Le “potentiel de double couche” était également une variante de ces objets.

L'utilité et le caractère intuitif de ces objets rendait presque indispensable leur incorporation dans une théorie mathématique ; c'est ce qu'a réalisé la théorie des distributions. Leur interprétation est d'ailleurs très simple, et cause beaucoup moins de maux de tête que les interrogations que l'on peut avoir sur la nature des “fonctions” considérées par Dirac.

1.2. Réhabilitation de la dérivation. Au début du vingtième siècle, la théorie de l'intégration de Lebesgue a pris un essor rapide, et l'intégration apparaît désormais comme l'opération reine de l'analyse. “La plupart” des fonctions sont intégrables, alors que très peu sont dérivables. En outre, la théorie de Lebesgue montre que l'intégration est souvent une opération continue vis-à-vis de la convergence des fonctions, uniforme ou même simple (ponctuelle) ; alors que la dérivation est une opération grossièrement discontinue. Ainsi, si l'on se donne une fonction u périodique, continûment dérivable, la suite de fonctions $(u_n)_{n\geq 1}$ définie par $u_n(x) = u(nx)/n$ converge uniformément vers la fonction nulle, mais la suite u_n' des dérivées ne converge certainement pas vers 0, ni uniformément ni simplement, ni dans un espace de Lebesgue.

Dans la théorie de Schwartz au contraire, ces problèmes disparaîtront : à toute fonction continue on pourra associer des “fonctions dérivées” à tous les ordres, selon une notion qui prolongera celle de dérivée des fonctions continûment dérivables. De manière plus générale, toute distribution sera dérivable à tout ordre, et on pourra définir une notion de convergence qui prolongera la notion de convergence uniforme, et pour laquelle l'opération de dérivée sera continue. Ainsi, dans l'exemple précédent, la suite de fonctions u_n' convergera au sens des distributions vers 0.

On note qu'il existe une autre théorie de la dérivation d'objets non réguliers : la théorie de “dérivation de la mesure”, qui commence avec le théorème de densité de Lebesgue. Cette théorie généralise la dérivation des fonctions continûment dérivables sur $\mathbb{R}$, mais s'avère insuffisante pour réhabiliter la dérivation : d'une part, son interprétation en dimension plus grande que 1 n'est pas très claire, d'autre part elle est limitée aux mesures.

1.3. Extension des espaces de solutions acceptables. Leray en 1934, Sobolev en 1936, introduisent de nouvelles notions de solutions d'équations aux dérivées partielles, appelées solutions faibles ou solutions généralisées, qui permettent de formuler des équations aux dérivées partielles sans supposer nécessairement l'existence de dérivées au sens classique. Leur approche préfigure très bien la théorie des distributions, qui ne sera mise au point qu'une quinzaine d'années plus tard. Les contributions de Leray et Sobolev ne constituent pas les seuls travaux précurseurs de la théorie des distributions [Schwartz, p. 5] : dans les années 30 et 40, de nombreux chercheurs vont utiliser les concepts de solutions généralisées pour étudier les solutions de diverses équations aux dérivées partielles : Courant et Hilbert, Bochner, Friedrichs, Krylov...

Voici un exemple très simple où l'on est naturellement amené à définir des solutions généralisées. L'équation des ondes en dimension 1,

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0$$

admet pour solutions toutes les fonctions régulières de la forme

$$u(t,x) = f(x-ct) + g(x+ct).$$

Cependant, même si f et g ne sont pas régulières mais, disons, continues, la fonction u définie ci-dessus peut s'interpréter comme la propagation de deux ondes (signaux se propageant sans altération) en sens opposé, et on aimerait bien pouvoir la considérer comme une solution de l'équation des ondes ! En outre, une telle fonction est limite uniforme de solutions de l'équation des ondes, ce qui est un autre argument pour souhaiter les considérer comme solutions elles-mêmes. Mais si les dérivées n'existent pas, en quel sens l'équation aux dérivées partielles est-elle satisfaite ?

Un autre exemple est fourni par l'étude de l'équation de transport linéaire,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \xi) = 0,$$

qui décrit l'évolution d'une densité de matière évoluant selon un champ de vitesses ξ donné. Si l'on veut modéliser le transport d'une seule particule, il est naturel de chercher à admettre parmi les solutions une "fonction de Dirac", dont la position varierait comme une trajectoire intégrale du champ ξ. Un tel objet n'est certainement pas une fonction, et encore moins une fonction dérivable...

Nous arrivons maintenant à l'idée majeure de la théorie des distributions, déjà présente dans les travaux de Sobolev et Leray. Du point de vue physique, on peut la motiver comme suit. Le résultat ou l'interprétation d'une expérience physique faisant intervenir une certaine grandeur, qui varie en fonction du temps ou de l'espace, ne dépend que très rarement des valeurs ponctuelles de cette grandeur, mais plus souvent de sa **valeur moyenne** (ou intégrale) prise contre une fonction du temps ou de l'espace, plus ou moins localisée. En d'autres termes, plutôt qu'une fonction u, c'est plutôt une "moyenne" de la forme $\int u\varphi$ qui sera accessible en pratique. Ainsi, un signal électrique oscillant avec une période très élevée fournira, de manière apparemment paradoxale, un signal plat quand on l'utilisera pour alimenter un oscilloscope : dans le processus de moyenne inhérent à la mesure, les valeurs positives et négatives se compenseront.

Un autre exemple où les moyennes jouent un rôle crucial est fourni par la mécanique statistique, où plutôt que de mesurer la densité de présence d'un gaz, on préfère mesurer des quantités physiques telles que la vitesse moyenne ou la température (écart quadratique moyen). Ce principe est si bien ancré, et depuis si longtemps, dans l'esprit des physiciens, que Maxwell en 1867, écrivant ce qui deviendrait plus tard l'équation de Boltzmann, modélisant l'évolution de la densité d'un gaz raréfié dans le formalisme cinétique, choisit comme inconnue non la densité f elle-même, mais la moyenne $\int f\varphi\,dv$ de f prise contre une fonction φ de la vitesse. C'est d'ailleurs également le principe d'un très grand nombre de mesures statistiques : plutôt que de chercher à accéder à la loi μ d'une variable aléatoire X, on préfère souvent estimer des statistiques portant sur des fonctions φ de cette variable aléatoire, ce qui revient à étudier des quantités de la forme $\int \varphi\,d\mu$ plutôt que la loi μ elle-même.

Dans le cas de l'équation des ondes aussi bien que dans celui des équations de transport avec données singulières, on résout ainsi le dilemme qui nous est proposé, en *reformulant l'équation par son action sur les fonctions-test*.

Explicitons plus en détail la réflexion de Sobolev. Pour des raisons ayant trait à l'analyse (fonctionnelle) de certaines classes d'équations aux dérivées partielles, il cherchait à étudier les fonctions possédant une dérivée carré-intégrable, dans un cadre suffisamment général... et en particulier, sans supposer la fonction dérivable ! Il montra que l'on pouvait donner une définition utile et pratique du concept de "fonction dont la dérivée est de carré intégrable", qui ne présuppose pas l'existence d'une dérivée au sens usuel. Son approche peut se résumer ainsi. Si $f : \mathbb{R} \to \mathbb{R}$ est dérivable, et que $f' \in L^2(\mathbb{R})$, alors pour toute fonction-test φ assez régulière et à support compact, on a

$$\int f\varphi' = \int f'\varphi \leq \|f'\|_{L^2}\|\varphi\|_{L^2}.$$

En particulier, la forme linéaire $\varphi \longmapsto \int f\varphi'$ est bornée sur L^2. Pour une fonction continûment dérivable f, cet énoncé est en fait *équivalent* à l'appartenance de f' à L^2. Comme il conserve un sens même si f n'est pas forcément dérivable, mais seulement L^2 par exemple, Sobolev propose de le considérer comme la définition souhaitée.

L'espace ainsi défini est aujourd'hui appelé espace de Sobolev $W^{1,2}$, ou H^1. Sa définition se généralise facilement à des dimensions arbitraires, à d'autres exposants que 2 et à d'autres ordres de dérivabilité. Cet espace fournit un cadre fonctionnel agréable pour étudier les fonctions dont la dérivée est carré-intégrable ; par exemple, on peut établir qu'en dimension 1, de telles fonctions sont automatiquement continues, ce qui n'est pas le cas en dimension plus grande.

1.4. Simplification de problèmes non linéaires ardus. Dans les exemples de la section précédente, l'utilisation de solutions généralisées reflétait une nécessité de la modélisation. Dans d'autres exemples, leur utilisation sera motivée par notre incompétence à prouver l'existence de solutions classiques. L'exemple archétypal en la matière est celui des équations de la mécanique des fluides, en particulier l'équation de Navier-Stokes incompressible, dont la première étude mathématique moderne remonte à par Leray. Cette équation s'écrit

$$\frac{\partial u}{\partial t} + u \cdot \nabla u + \nabla p = \Delta u; \qquad \nabla \cdot u = 0$$

où l'inconnue $u = u(t, x) \in \mathbb{R}^3$ est un champ de vitesses défini pour $t \geq 0$, $x \in \mathbb{R}^3$, et $p = p(t, x)$ est une fonction scalaire (la pression). On note comme d'habitude $u \cdot \nabla u = (u \cdot \nabla)u$ le champ de vecteurs dont la composante d'ordre i est $u \cdot \nabla u_i$; le laplacien vectoriel Δu est défini composante par composante.

L'écriture même de l'équation de Navier-Stokes incompressible présuppose apparemment la dérivabilité de u, mais rien ne semblait permettre d'affirmer l'existence de solutions dérivables pour des données initiales assez générales. Ce problème est d'ailleurs toujours ouvert, et mis à prix pour une forte récompense... Leray eut l'idée de définir une solution de l'équation de Navier-Stokes incompressible par une **formulation duale** qui ne présupposait pas de régularité : une fonction-test φ régulière étant donnée, on intègre l'équation contre φ et on cherche à reporter toutes les dérivées sur φ grâce à des intégrations par parties systématiques. Ainsi,

$$\int \Delta u \cdot \varphi = \int u \cdot \Delta \varphi.$$

Pour faire de même avec le terme convectif $u \cdot \nabla u$, on note que, par l'hypothèse de divergence nulle de u, on peut écrire $u \cdot \nabla u = \nabla \cdot (u \otimes u)$. Le terme de pression p est un peu déroutant, et il est possible de le faire disparaître en l'intégrant contre un champ de divergence nulle. On obtient donc l'information suivante : pour toute fonction-test $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$, C^∞, à support compact, à divergence nulle,

$$\frac{d}{dt} \int u \cdot \varphi - \int u \otimes u : \nabla \varphi = \int u : \Delta \varphi.$$

Il est également facile de construire des variantes de cette formulation qui ne supposent aucune régularité dans la variable t, et on peut vérifier que cette définition intégrée est équivalente à la définition habituelle quand on considère des solutions régulières – de même que dans l'exemple de Sobolev. L'idée de Leray est de considérer de telles identités, faisant intervenir u mais aucune dérivée de u, comme la définition même des solutions de l'équation de Navier-Stokes. De telles solutions sont aujourd'hui appelées **solutions faibles**.

1.5. Idée de départ. Les approches de Leray et de Sobolev se retrouvent unifiées dans la théorie des distributions. Encore aujourd'hui, les deux problématiques qui sous-tendent leurs travaux constituent d'ailleurs les principales motivations de l'emploi des distributions : l'étude des solutions faibles d'équations aux dérivées partielles d'une part, la mise en place d'un cadre général en analyse fonctionnelle d'autre part.

La théorie des distributions exploite au maximum l'idée de moyenner les solutions contre des fonctions-test qui se comportent bien : on définit au départ une classe agréable de fonctions-test, à la fois très régulières et bien localisées : les **fonctions indéfiniment dérivables, à support compact**. Toutes les propriétés des objets que l'on cherche à définir sont alors définies en fonction des "intégrales" de ces objets contre les fonctions-test. La topologie, et en particulier la notion de convergence, sera alors définie en fonction du comportement des intégrales contre des fonctions-test : pour caricaturer, une suite de distributions Λ_k converge vers une distribution Λ si, pour toute fonction-test φ, la suite de réels $\int \Lambda_k \varphi$ converge vers $\int \Lambda \varphi$.

On exprimera alors systématiquement toutes les opérations sur les distributions en fonctions d'opérations sur les intégrales prises contre des fonctions test. Ainsi, la dérivation sera définie au moyen de la formule d'"intégration par parties" $\int \Lambda'\varphi = -\int \Lambda\varphi$ (noter qu'il n'y a pas de terme de bord puisque φ est à support compact).

Reste à donner un sens précis à la notion d'intégrale. Comme on le rappellera plus loin, on peut traduire la théorie de l'intégration de Lebesgue, dans $\mathbb{R}^n$, en termes de formes linéaires sur l'espace des fonctions continues : c'est le point de vue de Riesz. L'approche utilisée par Schwartz consiste à généraliser cette approche, et à définir les distributions comme **le dual de l'espace des fonctions-test**.

1.6. Théories concurrentes ou parallèles. La théorie des distributions n'est pas l'unique tentative de définir la "bonne" notion de fonction généralisée : d'autres suggestions ont été faites par de nombreux auteurs. On trouve dans [Schwartz, Introduction] une liste non exhaustive de ces extensions et théories parallèles : fonctions généralisées de Gelfand, fonctionnelles analytiques de Martineau et de Fantappié, opérateurs de Mikusinski, distributions généralisées de Roumieu, ultradistributions de Sato. Dans plusieurs de ces théories, on cherche à remplacer l'espace des fonctions C^∞ à support compact par un espace de fonctions encore plus régulières, soit analytiques, soit "quasi-analytiques". On dit qu'un espace vectoriel $\mathcal{C}$ de fonctions forme une classe quasi-analytique si tout élément de $\mathcal{C}$, dont toutes les dérivées s'annulent en un point, s'annule identiquement partout. Les classes quasi-analytiques jouissent donc de certaines propriétés de rigidité que n'ont pas les fonctions C^∞, sans être aussi régulières que les fonctions analytiques.

Plusieurs de ces tentatives n'ont pas connu de succès ; un inconvénient majeur des espaces construits à partir des fonctions analytiques est l'absence de notion de support. En revanche, le point de vue quasi-analytique permet de bâtir une théorie assez proche de celle des distributions. On trouvera dans [Hörmander, chapitre 9] une excellente synthèse de cette théorie, basée entre autres sur les idées de Sato et de Martineau, dite "théorie des hyperfonctions".

Il est cependant incontestable que la théorie des distributions est à la fois plus élémentaire, et d'usage beaucoup plus populaire, que la théorie des hyperfonctions et ses variantes.

III-2. Fonctions

Si l'on garde en tête les principes de base des distributions, on a envie de considérer qu'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ n'a de sens qu'au travers des quantités $\int_K f$, où K décrit par exemple l'ensemble de tous les compacts de $\mathbb{R}^n$. Pour que ces quantités soient bien définies, on se restreindra, en théorie des distributions, à des **fonctions localement intégrables**. Un ouvert Ω étant donné (Ω pouvant être $\mathbb{R}^n$), on souhaite donc inclure $L^1_{\mathrm{loc}}(\Omega)$ parmi les distributions.

Il est clair que si l'on se donne deux fonctions f et g localement intégrables, égales presque partout dans Ω, alors pour tout compact K de Ω on a $\int_K f = \int_K g$, et pour toute fonction continue φ on a $\int \varphi f = \int \varphi g$. Le célèbre **théorème de densité de Lebesgue** implique la réciproque.

Théorème III-1 (théorème de densité de Lebesgue). *Soit f localement intégrable dans Ω. Alors*

(i) pour presque tout $x \in \Omega$,

$$f(x) = \lim_{r\to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} f(y)\, dy.$$

(ii) Cette conclusion reste vraie si l'on remplace les fonctions caractéristiques de boules par des approximations continues : soit (ρ_k) *une famille d'applications continues, positives et d'intégrale 1, de support contenu dans la boule de rayon* $1/k$*, alors pour presque tout* $x \in \Omega$*,*

$$f(x) = \lim_{k\to\infty} \int_\Omega \rho(x-y)f(y)\,dy.$$

Nous admettrons cet énoncé, qui est au reste l'un des théorèmes importants de la théorie de la mesure.

Une conclusion que l'on peut tirer est que le point de vue de l'intégration locale ne permet de distinguer les fonctions que modulo la relation d'équivalence "être égal presque partout à". De manière générale, **en théorie des distributions, on identifie des fonctions qui sont égales presque partout**.

Remarque III-2. Ce point de vue mène beaucoup d'auteurs à remplacer le concept de fonction par celui de "fonction définie presque partout", i.e. une classe d'équivalence pour la relation d'équivalence mentionnée ci-dessus. Très utile en théorie des distributions, ce point de vue est assez malheureux dans d'autres circonstances, en particulier parce qu'il interdit l'étude fine des fonctions : par exemple, un énoncé intéressant tel que "l'ensemble des points de non-dérivabilité d'une fonction convexe en dimension n est de mesure de Hausdorff au plus $n-1$" ne s'exprime pas naturellement dans le langage des distributions.

III-3. Mesures

Les mesures constituent l'exemple le plus simple de fonctions généralisées qui entrent dans le cadre des distributions. Toute la théorie des distributions découle naturellement d'un examen attentif de l'espace des mesures. Commençons par définir précisément de quelles "mesures" il est question.

Définition III-3 (mesures). *Soit* Ω *un ouvert de* $\mathbb{R}^n$*.*

(i) Une mesure de Borel sur Ω*, finie sur les compacts, est une fonction d'ensembles* μ *définie sur la tribu borélienne, i.e. la* σ*-algèbre engendrée par les ouverts de* $\mathbb{R}^n$*, à valeurs dans* $[0,+\infty]$*, prenant des valeurs finies sur les ensembles compacts, satisfaisant à l'axiome de* σ*-additivité*

$$\mu[\cup A_k] = \sum_k \mu[A_k]$$

pour toute famille dénombrable $(A_k)_{k\in\mathbb{N}}$ *de boréliens disjoints de* $\mathbb{R}^n$*. Une telle mesure est automatiquement régulière, i.e.*

$$\mu[A] = \sup\Big\{\mu[O];\ \ O \text{ ouvert, } A \subset O\Big\} = \inf\Big\{\mu[K];\ \ K \text{ compact, } K \subset A\Big\}.$$

(ii) Nous appellerons mesure signée, ou mesure de Radon, la différence de deux mesures de Borel, finies sur les compacts. Une telle mesure μ *définit, par restriction aux parties d'un compact fixé de* Ω*, une fonction* σ*-additive d'ensembles, à valeurs dans* $\mathbb{R}$*. On peut la décomposer de manière unique en la différence de deux mesures positives singulières l'une à l'autre, appelées sa partie positive* μ_+ *et sa partie négative* μ_-*.*

(iii) La variation totale d'une mesure signée μ sur un ensemble $A \subset \Omega$ est définie par $\mu_+[A] + \mu_-[A]$. On notera $M(\Omega)$, et on appellera espace des mesures de Radon finies, l'ensemble des mesures de Radon sur Ω dont la variation totale sur Ω est finie ; et on notera $M_{\mathrm{loc}}(\Omega)$ l'ensemble de toutes les mesures de Radon. Le premier est un espace de Banach, le second est un espace de Fréchet. Si A est une partie mesurable de Ω, incluse dans un compact $K \subset \Omega$, on notera également $M(A)$ l'espace des mesures de Radon dans Ω supportées par (concentrées sur) A.

3.1. Les mesures généralisent les fonctions. Soient Ω un ouvert de $\mathbb{R}^n$, et $f : \Omega \to \mathbb{R}$ une fonction localement intégrable. Alors l'application injective $A \longmapsto \int_A f$ définit une mesure de Radon sur Ω. L'espace $L^1_{\mathrm{loc}}(\Omega)$ s'identifie ainsi à un sous-espace de $M_{\mathrm{loc}}(\Omega)$.

Mais l'espace des mesures de Radon est plus gros que l'espace L^1_{loc} ; l'exemple le plus simple et le plus utile d'une mesure de Radon qui ne soit pas une fonction est la "fonction de Dirac", ou plus proprement **masse de Dirac** ou **mesure de Dirac** en un point x, définie par l'identité

$$\delta_x[A] = 1_{x \in A}.$$

Une telle mesure est bien sûr d'intégrale (ou de masse) 1, et elle est nulle en-dehors de x, au sens où son intégrale sur n'importe quelle partie de Ω ne contenant pas x est nulle.

On peut également construire des mesures portées par des hypersurfaces : par exemple, si $f : \mathbb{R}^k \to \mathbb{R}^n$ est une application continue, et que l'on se donne une mesure absolument continue μ sur $\mathbb{R}^k$, alors la mesure image $f\#\mu$ est une mesure sur $\mathbb{R}^n$, concentrée sur l'image de f, qui peut être un arc de courbe, un morceau de surface, etc.

3.2. Le théorème de Riesz. Le théorème de Riesz permet d'interpréter les mesures de Radon dans une perspective "fonctionnelle".

Théorème III-4 (théorème de Riesz). *Soit Ω un ouvert de $\mathbb{R}^n$. Alors on peut identifier*

(i) les mesures de Radon positives avec les formes linéaires positives sur l'espace $C_c(\Omega)$ des fonctions continues à support compact dans Ω ;

(ii) l'espace $M(K)$ avec le dual (topologique) de l'espace $C(K)$, pour tout compact $K \subset \Omega$;

(iii) l'espace $M(\Omega)$ avec le dual (topologique) de l'espace $C_0(\Omega)$, des fonctions continues qui tendent vers 0 en s'approchant du bord de Ω ;
via le crochet de dualité $(\mu, f) \longmapsto \int f\, d\mu$.

L'inégalité

$$\left| \int_\Omega f\, d\mu \right| \leq \|f\|_\infty \|\mu\|_{VT}$$

implique que toute mesure de variation totale finie induit une forme linéaire continue sur l'espace des fonctions continues bornées. Le point (ii) ci-dessus implique que la réciproque est vraie sur un compact : toute forme linéaire Λ sur $C(K)$ vérifiant

$$|\Lambda f| \leq C\|f\|_\infty \tag{4}$$

pour tout $f \in C(K)$, et pour une constante $C < +\infty$ indépendante de f, peut se représenter par une mesure supportée par K.

Remarque III-5. Attention : si l'on admet la forme forte du théorème de Hahn-Banach, on peut prouver que le dual de l'espace $C_b(\Omega)$ est strictement plus grand que $M(\Omega)$.

Une question naturelle se pose alors : peut-on munir les espaces $C_c(\Omega)$ et $M_{\mathrm{loc}}(\Omega)$ de structures topologiques telles que $C_c^* = M_{\mathrm{loc}}$? La réponse est affirmative, et la solution est similaire à celle que nous étudierons dans le cadre des distributions. Pour l'instant, nous allons nous borner à reformuler le théorème de Riesz de la façon suivante :

Théorème III-6 (reformulation du théorème de Riesz). *Soit Ω un ouvert de $\mathbb{R}^n$. Alors on peut identifier l'espace $M_{\mathrm{loc}}(\Omega)$ à l'espace des formes linéaires sur $C_c(\Omega)$ telles que pour tout compact K il existe une constante C_K telle que pour toute fonction continue $\varphi \in C(K)$, on ait $|\Lambda\varphi| \leq C_K\|f\|_\infty$.*

Démonstration. Si $\mu \in M_{\mathrm{loc}}(\Omega)$, alors sa variation totale sur chaque compact $K \subset \Omega$ est finie, ce qui prouve l'existence d'un C_K comme ci-dessus. Réciproquement, si une forme linéaire vérifie l'hypothèse du théorème, alors elle induit sur chaque C_K une forme linéaire continue, et donc une mesure de Radon finie. □

Remarque III-7. Comme on vient de le voir, il est beaucoup plus facile d'introduire les mesures de Radon à partir du point de vue fonctionnel qu'à partir du point de vue ensembliste. Cette constatation avait poussé Bourbaki à prendre le point de vue fonctionnel comme base de la théorie de la mesure. Malheureusement, l'équivalence entre ces deux points de vue n'est valable que dans des espaces localement compacts, ce qui limite l'intérêt pratique de la définition fonctionnelle au cas où l'on travaille dans une partie de $\mathbb{R}^n$.

III-4. Définition des distributions

Pour définir les distributions, on considère la conclusion du Théorème III-6 comme nouveau point de départ, et on généralise la formule pour admettre une classe de formes linéaires beaucoup plus larges, tirant parti de la régularité de l'espace des fonctions-test.

4.1. Espace $\mathcal{D}$.

Définition III-8 (fonctions-test). *Soit Ω un ouvert de $\mathbb{R}^n$; on note $\mathcal{D}(\Omega)$, ou $C_c^\infty(\Omega)$, l'espace vectoriel des fonctions de classe C^∞ à support compact dans Ω.*

Un compact K de Ω étant donné, on note $\mathcal{D}_K$ l'espace des éléments de $\mathcal{D}(\Omega)$ dont le support est inclus dans K.

On rappelle que $\mathcal{D}(\Omega)$ n'est pas réduit à la fonction nulle, et que c'est un espace de dimension infinie. Il existe plusieurs méthodes pour construire des fonctions-test vérifiant certaines propriétés ; quelques-unes sont présentées de manière très synthétique dans [Hörmander, chapitre 1].

4.2. Espace $\mathcal{D}'$.

Définition III-9 (distributions). *Soit Ω un ouvert de $\mathbb{R}^n$. On appelle distribution une forme linéaire Λ sur $\mathcal{D}(\Omega)$ vérifiant la propriété caractéristique suivante : pour tout compact $K \subset \Omega$, il existe $k \in \mathbb{N}$ et $C_K < +\infty$ tel que pour tout $\varphi \in \mathcal{D}_K$,*

$$|\Lambda\varphi| \leq C_K \sum_{|\alpha|\leq k} \sup_K |\nabla^\alpha \varphi|.$$

On note $\mathcal{D}'(\Omega)$ l'espace des distributions.

Deux notions joueront un rôle important dans l'étude et la classification des distributions : l'ordre et le support. La possibilité de définir une notion simple et efficace de support est l'une des raisons du succès des distributions.

DÉFINITION III-10 (ordre). *Soit Ω un ouvert de $\mathbb{R}^n$, et $\Lambda \in \mathcal{D}'(\Omega)$. On appelle ordre de Λ sur un compact $K \subset \Omega$ le plus petit entier k tel qu'il existe $C_K < +\infty$ tel que pour tout $\varphi \in \mathcal{D}_K$,*

$$|\Lambda\varphi| \leq C_K \sum_{|\alpha|\leq k} \sup_K |\nabla^\alpha \varphi|.$$

On appelle ordre de Λ le plus petit k tel que Λ soit d'ordre k sur tout compact $K \subset \Omega$.

REMARQUE III-11. Une distribution peut être d'ordre infini : on peut construire un exemple à titre d'exercice. Dans des problèmes concrets, il est cependant rare que l'on rencontre de tels objets.

DÉFINITION III-12 (support). *Soit Ω un ouvert de $\mathbb{R}^n$, et $\Lambda \in \mathcal{D}'(\Omega)$. On appelle support de Λ le complémentaire de l'union de tous les ouverts ω dans lesquels Λ s'annule, i.e. tels que $\Lambda\varphi = 0$ pour toute fonction $\varphi \in \mathcal{D}(\omega)$. On le notera* $\operatorname{supp}\Lambda$.

PROPOSITION III-13 (définition équivalente du support). *Soient Ω un ouvert de $\mathbb{R}^n$, et $\Lambda \in \mathcal{D}'(\Omega)$. Alors le support de Λ n'est autre que le complémentaire du plus grand ouvert ω dans lequel Λ s'annule.*

DÉMONSTRATION. Il nous suffit de montrer que si Λ s'annule dans une famille d'ouverts $(\omega_i)_{i\in I}$, alors Λ s'annule aussi dans $\bigcup \omega_i$. Pour cela, soit φ une fonction à support compact K dans $\bigcup \omega_i$. Par compacité, K est recouvert par une sous-famille finie $(\omega_i)_{i\in J}$. Par un théorème de **partition de l'unité**, on peut trouver des fonctions χ_i $(i \in J)$, de support compact respectivement inclus dans ω_i, et dont la somme vaut identiquement 1 sur K. On a alors

$$\Lambda\varphi = \Lambda[(\sum_{i\in J} \chi_i)\varphi] = \sum_{i\in J} \Lambda(\chi_i\varphi) = 0,$$

ce qui conclut la preuve. □

4.3. Exemples.

– Il est clair que $L^1_{\mathrm{loc}}(\Omega)$ est un sous-espace de $\mathcal{D}'(\Omega)$; on montre que f définit une distribution nulle dans un ouvert $O \subset \Omega$ si et seulement si f est nulle presque partout dans O. Le support d'une fonction localement intégrable f est le plus petit fermé en-dehors duquel f définit une distribution nulle, c'est donc également le plus petit fermé en-dehors duquel f est nulle presque partout (noter que l'existence de ce plus petit fermé n'est pas complètement évidente a priori) ; et l'ordre de f en tant que distribution est 0. En utilisant le fait que $f * \rho_k \longrightarrow f$ presque partout, où $(\rho_k)_{k\geq 1}$ est une approximation de l'identité pour la convolution, on montre facilement que deux fonctions L^1_{loc} qui sont égales en tant que distributions, sont égales presque partout.
– L'espace $M_{\mathrm{loc}}(\Omega)$ est aussi un sous-espace de $\mathcal{D}'(\Omega)$, au vu du Théorème III-6. Il est également constitué de distributions d'ordre 0.

– Soit $x \in \mathbb{R}$ un nombre réel arbitraire. La forme linéaire

$$f \longmapsto f'(x)$$

définie sur $\mathcal{D}(\mathbb{R})$, définit une distribution d'ordre 1, de support $\{x\}$. Nous l'identifierons bientôt avec la dérivée de la mesure de Dirac en x.

– Plus généralement, si x est un point de $\mathbb{R}^n$ et α un multi-indice, la forme linéaire

$$f \longmapsto \nabla^\alpha f(x)$$

définit une distribution d'ordre $|\alpha|$ et de support $\{x\}$. Nous la noterons $\nabla^\alpha \delta_x$, ce qui pour l'instant n'est qu'une notation formelle.

– Soit $k \in \mathbb{N}$; alors l'espace $H^{-k}(\mathbb{R}^n) := H^k(\mathbb{R}^n)^*$ est un sous-espace de $\mathcal{D}'(\mathbb{R}^n)$. En effet, soit $\Lambda \in H^{-k}$, alors par définition, il existe $C < +\infty$ tel que pour tout $f \in H^k(\mathbb{R}^n)$ on ait

$$|\Lambda f| \leq C\|f\|_{H^k}.$$

En particulier, si f est une fonction de support compact K,

$$|\Lambda f| \leq C\|f\|_{H^k} \leq C' \sup_{|\alpha|\leq k} \|\nabla^\alpha f\|_{L^2} \leq C'|K|^{1/2} \sup_{|\alpha|\leq K} \|\nabla^\alpha f\|_{L^\infty},$$

ce qui montre bien que Λ est une distribution, d'ordre au plus k. De manière plus générale, l'espace $W^{-k,p}(\mathbb{R}^n)$, défini comme l'espace dual de $W^{k,p'}(\mathbb{R}^n)$, est un espace de distributions d'ordre au plus k.

– La fonction $x \longmapsto 1/x$ en dimension 1 n'est pas localement intégrable. Cependant, pour tout $\varphi \in \mathcal{D}(\mathbb{R})$, la quantité

$$\Lambda\varphi := \lim_{\varepsilon\to 0} \int_{|x|>\varepsilon} \frac{\varphi(x)}{x}\, dx$$

est bien définie. En effet, par la formule de Taylor avec reste intégral,

$$\varphi(x) = \varphi(0) + x\int_0^1 \varphi'(tx)\, dt;$$

si l'on introduit une fonction de troncature $\chi \in \mathcal{D}(\Omega)$, paire et telle que $\varphi\chi = \varphi$, on aura

$$\varphi(0)\int_{|x|>\varepsilon} \frac{\chi(x)}{x}\, dx = 0,$$

et partant

$$\int_{|x|>\varepsilon} \frac{\varphi(x)}{x}\, dx = \int_{|x|>\varepsilon} [\int_0^1 \varphi'(tx)\, dt\chi(x)]\, dx \longrightarrow [\varepsilon \to 0] \int_{\mathbb{R}}\int_0^1 \varphi'(tx)\chi(x)\, dt\, dx.$$

On en déduit que Λ est une distribution d'ordre 1 ; on l'appelle **valeur principale** de la fonction $1/x$ et on la note v.p.$(1/x)$.

– La fonction $x \longmapsto 1/|x|$ en dimension 1 n'est pas localement intégrable. Cependant, pour tout $\varphi \in \mathcal{D}(\mathbb{R})$, on a

$$\int_{|x|>\varepsilon} \frac{\varphi(x)}{|x|}\, dx = \varphi(0)c(\varepsilon) + \int_{\mathbb{R}} \frac{x}{|x|}\int_0^1 \varphi'(tx)\chi(x)\, dt\, dx,$$

où la quantité $c(\varepsilon)$ est indépendant de φ et diverge quand $\varepsilon \to 0$. On convient de définir la **partie finie** (par opposition à partie divergente) de $1/|x|$ comme

$$\left\langle \text{p.f.}\left(\frac{1}{|x|}\right), \varphi \right\rangle = \int_{\mathbb{R}} \frac{x}{|x|} \int_0^1 \varphi'(tx)\chi(x)\, dt\, dx.$$

Cette construction se généralise aisément à d'autres fonctions puissances et d'autres dimensions : on peut ainsi définir p.f.$(1/|x|^\alpha)$ pour tout $\alpha \geq n$ dans $\mathbb{R}^n$ (pour $\alpha < n$ le symbole p.f. serait inutile).

4.4. Distributions à valeurs vectorielles. Les distributions généralisent les fonctions à valeurs rélles... que dire des fonctions à valeurs dans d'autres espaces, par exemple un espace vectoriel topologique E ?

- Il est très facile de définir des distributions à valeurs dans $\mathbb{C}$ ou $\mathbb{R}^n$: il suffit de le faire composante par composante. Les distributions à valeurs complexes jouent un rôle important en analyse complexe.

- Il est plus délicat de définir des distributions à valeurs dans un espace vectoriel topologique de dimension infinie. Il existe des théorèmes généraux...

- Une question particulièrement intéressante consiste à généraliser les formes différentielles ou les champs de tenseurs sur une variété différentiable régulière. Les distributions à valeurs dans un fibré tensoriel sur une variété sont appelées **courants**.

III-5. Topologies

Schwartz découvrit une structure topologique relativement naturelle qui transformait $\mathcal{D}'(\Omega)$ en l'espace dual de $\mathcal{D}(\Omega)$, et induisait une notion naturelle de convergence. Cette contribution était particulièrement enthousiasmante dans le contexte des idées mathématiques de l'époque, où beaucoup de chercheurs espéraient avec Bourbaki que l'analyse serait révolutionnée par la mise en place de structures fonctionnelles adaptées.

5.1. Topologie de $\mathcal{D}_K$. Soit $K \subset \Omega$ un compact de Ω. Alors l'espace $\mathcal{D}_K$ des fonctions C^∞ à support compact dans K est un espace de Fréchet, une famille de semi-normes étant données par les normes C^k, $k \in \mathbb{N}$. En particulier, le critère de continuité des applications linéaires définies entre espaces de Fréchet montre qu'une forme linéaire Λ sur $\mathcal{D}_K$ est continue si et seulement si il existe $k \in \mathbb{N}$ et $C < +\infty$ tel que pour tout $f \in \mathcal{D}_K$,

$$|\Lambda f| \leq C \sup_{|\alpha| \leq k} \sup_K \|\nabla^\alpha f\|.$$

On conclut que les distributions sont exactement les formes linéaires sur $\mathcal{D}(\Omega)$ qui induisent des formes linéaires continues sur tous les espaces $\mathcal{D}_K$. Une conséquence importante est la suivante : si $(\varphi_k)_{k\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{D}_K$, convergeant uniformément vers φ, et que toutes les dérivées partielles, à tous les ordres, de φ_k convergent uniformément vers les dérivées correspondantes de φ, alors

$$\Lambda\varphi_k \xrightarrow[k\to\infty]{} \Lambda\varphi.$$

5.2. Topologie de $\mathcal{D}$. Pour aller plus loin, on souhaite maintenant définir une topologie sur $\mathcal{D}(\Omega)$ tout entier. Une première tentative consiste à introduire une suite exhaustive $(K_p)_{p\in\mathbb{N}}$ de compacts, et à définir $\mathcal{D}(\Omega)$ comme un espace de Fréchet pour les semi-normes $\|\cdot\|_{C^p(K_p)}$. Cependant, l'espace ainsi défini n'est **pas complet** : sa complétion est l'espace $C_0^\infty(\Omega)$, constitué des fonctions C^∞ qui tendent vers 0 au bord de Ω, ainsi que toutes leurs dérivées.

Pour faire de $\mathcal{D}(\Omega)$ un espace complet, on cherche à restreindre l'espace des suites de Cauchy, en rendant la propriété de Cauchy beaucoup plus contraignante, ce qui peut se faire en **enrichissant la topologie**. De là vient l'idée de munir $\mathcal{D}(\Omega)$ de la topologie **limite inductive** des $\mathcal{D}_K$, i.e. la plus fine topologie rendant continues toutes les injections $\mathcal{D}_K \subset \mathcal{D}(\Omega)$. En voici une définition possible de la topologie considérée.

Définition III-14 (topologie de $\mathcal{D}(\Omega)$). *Soit Ω un ouvert de $\mathbb{R}^n$. On définit une base locale de voisinages (en 0) dans $\mathcal{D}(\Omega)$ par la donnée de toutes les parties convexes et symétriques A telles que $A\cap\mathcal{D}_K$ soit un ouvert de $\mathcal{D}_K$ pour tout compact $K\subset\Omega$.*

Exemple III-15. Soit $(x_m)_{m\in\mathbb{N}}$ une suite dans Ω, convergeant dans $\mathbb{R}^n$ vers un point de $\partial\Omega$, et soit $(c_m)_{m\in\mathbb{N}}$ une suite arbitraire de nombres strictement positifs. Alors l'ensemble

$$A := \{\varphi\in\mathcal{D}(\Omega);\ \forall m\in\mathbb{N},\ |\varphi(x_m)| < c_m\}$$

est un ouvert convexe symétrique de $\mathcal{D}(\Omega)$. En effet, si K est un compact de Ω, il ne contient qu'un nombre fini de x_m, et c'est donc l'intersection d'un nombre fini d'ensembles ouverts de la forme $\delta_{x_m}^{-1}(]-c_m, c_m[)$.

La convergence est maintenant une propriété beaucoup plus contraignante, comme le montre le théorème suivant [Rudin, Théorème 6.6]

Théorème III-16 (propriétés topologiques de $\mathcal{D}(\Omega)$). *Soit Ω un ouvert de $\mathbb{R}^n$. Alors*

(i) si A est un ensemble borné de $\mathcal{D}(\Omega)$, alors il existe un compact $K\subset\Omega$ tel que $A\subset\mathcal{D}_K$;

(ii) en particulier, si $(f_j)_{j\in\mathbb{N}}$ est une suite de Cauchy dans $\mathcal{D}(\Omega)$, alors il existe un compact fixe K tel que tous les f_j aient leur support dans K ;

(iii) une suite $(f_j)_{j\in\mathbb{N}}$ converge dans $\mathcal{D}(\Omega)$ si et seulement si tous les f_j ont leur support inclus dans un compact fixe K, et si la suite (f_j) converge uniformément, ainsi que la suite de ses dérivées à tous les ordres.

(iv) $\mathcal{D}(\Omega)$ est un espace topologique complet ;

(v) $\mathcal{D}(\Omega)$ est un espace de Montel ;

(vi) Le dual topologique de $\mathcal{D}(\Omega)$ coïncide avec $\mathcal{D}'(\Omega)$.

(vii) pour tout compact $K\subset\Omega$, $\mathcal{D}_K$ est un fermé d'intérieur vide dans $\mathcal{D}(\Omega)$;

(viii) $\mathcal{D}(\Omega)$ n'est pas métrisable.

Esquisse de démonstration. Nous n'indiquerons que certains points-clé dans la démonstration. La propriété cruciale est le point (i), qui assure à la fin la complétude en forçant les suites de Cauchy à rester dans un espace (de Fréchet) $\mathcal{D}_K$.

Pour établir (i), considérons par l'absurde un ensemble borné A qui n'appartienne à aucun $\mathcal{D}_K$. Il existe donc une suite $(\varphi_m)_{m\in\mathbb{N}}$ de fonctions de A, et une suite $(x_m)_{m\in\mathbb{N}}$

de points de Ω, contenue dans aucun compact de Ω, telle que $\varphi_m(x_m) \neq 0$. On pose alors $c_m := |\varphi_m(x_m)|/m$; par l'Exemple III-15, l'ensemble

$$W := \left\{\varphi \in \mathcal{D}(\Omega);\ \forall m \in \mathbb{N},\ |\varphi(x_m)| < c_m\right\}$$

est un voisinage de 0. L'ensemble A étant borné, il doit donc être possible de trouver $\lambda > 0$ tel que $\lambda A \subset W$. En particulier, $\lambda|\varphi(x_m)| < c_m$, ce qui impose $\lambda < 1/m$ pour tout $m \in \mathbb{N}$: nous obtenons une contradiction.

Les propriétés (ii), (iii) et (iv) sont alors des conséquences assez simples de la propriété (i), du fait que $\mathcal{D}_K$ est un espace de Fréchet, et de considérations topologiques générales. La propriété (v) découle assez simplement de ce que les ensembles bornés de $\mathcal{D}$ sont inclus dans un $\mathcal{D}_K$, et que $\mathcal{D}_K$ est un espace de Montel.

La propriété (v) découle de ce qu'une application linéaire continue doit envoyer ensemble borné sur ensemble borné, qu'un ensemble borné de $\mathcal{D}(\Omega)$ est un ensemble borné d'un certain $\mathcal{D}_K$, et de ce que les distributions sont les formes linéaires dont la restriction à chaque $\mathcal{D}_K$ est continue.

La propriété (vi) est presque évidente. D'une part, $\mathcal{D}_K$ est fermé, comme l'intersection des fermés $\delta_x^{-1}(0)$ pour $x \notin K$. D'autre part, soit $\varphi \in \mathcal{D}_K$, et ψ un élément de $\mathcal{D}(\Omega)$ dont le support ne soit pas inclus dans K ; alors la fonction $\varphi+\varepsilon\psi$ n'appartient pas à $\mathcal{D}_K$, mais converge vers φ dans $\mathcal{D}(\Omega)$ quand $\varepsilon \to 0$.

Pour démontrer le point (vii), il suffit d'invoquer le **théorème de Baire**, sous la forme suivante : *dans un espace métrique complet, une union dénombrable de fermés d'intérieur vide est d'intérieur vide.* En l'occurrence, si l'on se donne une suite exhaustive $(K_j)_{j\in\mathbb{N}}$ de compacts de Ω, alors $\mathcal{D}(\Omega)$ est l'union dénombrable des $\mathcal{D}_{K_j}$, dont chacun est d'intérieur vide, mais $\mathcal{D}(\Omega)$ lui-même n'est pas d'intérieur vide. Puisque $\mathcal{D}(\Omega)$ est complet et ne satisfait pas la conclusion du théorème de Baire, on en conclut qu'il n'est pas métrisable. □

Remarque III-17. L'espace $\mathcal{D}$ n'étant pas métrisable, définir la topologie n'est pas équivalent à définir la convergence des suites. Cependant, dans la pratique c'est presque la seule notion de convergence que l'on utilise.

5.3. Topologie de $\mathcal{D}'(\Omega)$. Après avoir transformé $\mathcal{D}(\Omega)$ en espace topologique complet, nous pouvons faire de même pour l'espace des distributions.

Définition III-18 (topologie de $\mathcal{D}'(\Omega)$). *Soit Ω un ouvert de $\mathbb{R}^n$. On munit $\mathcal{D}'(\Omega)$ de la topologie faible-$*$ induite par $\mathcal{D}(\Omega)$, i.e. la topologie la plus grossière qui rende continues toutes les applications*

$$I_\varphi : \Lambda \longmapsto \Lambda\varphi, \qquad \varphi \in \mathcal{D}(\Omega).$$

Remarque III-19. De manière plus explicite, la topologie faible-$*$ est définie par les unions arbitraires d'intersections finies d'ensembles de la forme $I_\varphi^{-1}(O)$, où O est un ouvert de $\mathbb{R}$. Tout ouvert de cette topologie s'écrit donc sous la forme

$$\bigcup\Big\{\Lambda \in \mathcal{D}'(\Omega);\ \Lambda\varphi_1 < \alpha_1, \dots, \Lambda\varphi_k < \alpha_k\Big\},$$

où l'union est prise sur une famille quelconque de multi-uplets $(\varphi_1, \dots, \varphi_k)$ et $(\alpha_1, \dots, \alpha_k)$, l'indice k pouvant varier.

Théorème III-20 (Propriétés topologiques de $\mathcal{D}'(\Omega)$). *Soit Ω un ouvert de $\mathbb{R}^n$. Alors* *(i) une suite* $(\Lambda_k)_{k\in\mathbb{N}}$ *de distributions converge vers une distribution* Λ *si et seulement si* $\Lambda_k\varphi$ *converge vers* $\Lambda\varphi$ *;*

(ii) $\mathcal{D}'(\Omega)$ *est un espace topologique complet ;*

(iii) $\mathcal{D}'(\Omega)$ *est un espace de Montel ;*

(iv) Le dual topologique de $\mathcal{D}'(\Omega)$ *s'identifie à* $\mathcal{D}(\Omega)$*, qui est donc réflexif.*

(v) la topologie de $\mathcal{D}'(\Omega)$ *n'est pas métrisable.*

Nous admettrons encore ce théorème, qui découle de propriétés générales des topologies faibles (voir [Rudin] et [Schwartz]). Pour apprécier le point (iii), il faut savoir qu'un ensemble borné $\mathcal{B}$ de $\mathcal{D}'$ est un ensemble qui prend des valeurs bornées sur tout ensemble borné de $\mathcal{D}$, autrement dit : pour tout K compact de Ω, pour toute suite $(M_k)_{k\in\mathbb{N}}$ de réels positifs, il existe une constante C telle que pour tout $\Lambda \in \mathcal{B}$ et pour tout $\varphi \in \mathcal{D}_K$, on ait $|\Lambda\varphi| \leq C$.

Exemples III-21. Si x_k converge vers x, alors δ_{x_k} converge au sens des distributions vers δ_x, de même que $\nabla^\alpha \delta_{x_k}$, pour tout α.

La convergence uniforme sur tout compact, la convergence dans L^1_{loc} ou plus généralement dans M_{loc}, la convergence dans le dual de n'importe quel espace de Sobolev, impliquent la convergence au sens des distributions.

Remarque III-22. Si la non-métrisabilité de $\mathcal{D}(\Omega)$ pouvait paraître étonnante, il n'en est pas de même de celle de $\mathcal{D}'(\Omega)$: en général, les topologies faibles ne sont pas métrisables en dimension infinie.

III-6. Calcul des distributions

Dans cette section, on va passer en revue toutes les opérations principales sur les distributions. Certaines seront très faciles par construction des distributions, d'autres demanderont certaines précautions, d'autres encore seront quasiment impossibles.

On énoncera à chaque fois les propriétés de continuité séquentielles, utiles pour l'intuition, tout en insistant encore sur le fait que l'on peut presque toujours les contourner en pratique. Ainsi, nous verrons que l'opération d'évaluation est continue de $\mathcal{D}' \times \mathcal{D}$ dans $\mathbb{R}$: si des distributions Λ_n convergent dans $\mathcal{D}'$ vers Λ et des fonctions test φ_n convergent dans $\mathcal{D}$ vers φ, alors $\Lambda_n\varphi_n \to \Lambda\varphi$. Ce résultat, qui repose sur le théorème de Banach-Steinhaus dans les espaces de Fréchet, est d'un intérêt théorique important, mais en pratique il est bien rare que l'on rencontre une suite convergente de distributions qui ne soit pas convergente (au moins localement) dans un espace de Sobolev dual bien choisi, auquel cas le théorème de Banach-Steinhaus dans les espaces de Banach suffirait à assurer la continuité de l'évaluation.

6.1. Addition, soustraction, multiplication scalaire. Les distributions étant définies comme des formes linéaires, on peut les additionner, soustraire, ou multiplier par un scalaire, sans rencontrer aucune difficulté.

6.2. Localisation et recollement. Il arrive souvent que l'on ait besoin de changer momentanément l'ouvert Ω. Un rôle important dans ce problème est joué par les **fonctions plateaux**, fonctions très régulières, prenant leurs valeurs dans $[0, 1]$, identiquement égales à 1 sur un ensemble donné, et nulles en-dehors d'un ensemble donné. Les deux résultats principaux en la matière sont

- le **lemme d'Urysohn** : étant donnés un compact K inclus dans un ouvert O, on peut trouver une fonction-plateau identiquement égale à 1 sur un voisinage de K, à support compact inclus dans O ;

- le théorème de **partition de l'unité** : étant donné un compact K inclus dans une famille (O_i) d'ouverts (que l'on peut grâce à la compacité supposer finie sans perte de généralité), on peut trouver des fonctions plateaux χ_i telles que le support de χ_i soit inclus dans O_i, pour tout i, et $\sum \chi_i = 1$ sur un voisinage de K.

Voyons maintenant comment appliquer ces résultats en pratique.

Première opération : restriction Soient Ω un ouvert de $\mathbb{R}^n$, et Ω' un ouvert de Ω. Toute distribution $\Lambda \in \mathcal{D}'(\Omega)$ définit une distribution sur Ω', simplement par restriction de l'espace des fonctions test : $\mathcal{D}(\Omega') \subset \mathcal{D}(\Omega)$. On peut donc restreindre une distribution à un ouvert plus petit que celui sur lequel elle était définie a priori.

Deuxième opération : localisation Etant donné un compact K (au voisinage duquel on veut étudier la distribution) un ouvert O contenant K, et une distribution $\Lambda \in \mathcal{D}'(\Omega)$, on peut trouver une autre distribution Λ' dont le support est inclus dans O, et qui coïncide avec Λ dans un voisinage de K. Pour cela il suffit de définir

$$\langle \Lambda', \varphi \rangle = \langle \Lambda, \chi\varphi \rangle,$$

où χ est une fonction plateau identiquement 1 au voisinage de K, à support compact inclus dans O.

Troisième opération : recollement Soient $(\Omega_i)_{i\in J}$ une famille (éventuellement infinie) d'ouverts de $\mathbb{R}^n$, et Ω la réunion des Ω_i. Pour chaque i, on se donne une distribution $\Lambda_i \in \mathcal{D}'(\Omega_i)$; peut-on définir dans Ω une distribution Λ dont la restriction à chaque Ω_i soit Λ_i ? Une condition nécessaire est bien sûr $\Lambda_i = \Lambda_j$ dans $O_i \cap O_j$. Réciproquement, si cette condition est satisfaite, on peut effectivement trouver Λ : pour toute fonction φ à support compact dans Ω, on introduit une partition de l'unité (χ_i) telle que $\sum \chi_i = 1$ au voisinage du support de φ, et le support de χ_i soit inclus dans Ω_i. On a alors

$$\varphi = (\sum_i \chi_i)\varphi,$$

et on peut définir

$$\langle \Lambda, \varphi \rangle := \sum_i \langle \Lambda_i, \chi_i\varphi \rangle.$$

On montre (exercice) que (i) la quantité $\langle \Lambda, \varphi \rangle$ ainsi définie est indépendante du choix de la partition de l'unité, et (ii) cette distribution admet Λ_i comme restriction à Ω_i.

6.3. Dérivation. Les distributions ont été en partie inventées pour pouvoir être dérivées... Le théorème suivant montre qu'elles remplissent leur contrat de ce point de vue.

Théorème III-23 (dérivation des distributions). *Soit Ω un ouvert de $\mathbb{R}^n$, et Λ une distribution dans Ω. Pour tout $k \leq n$, on définit la distribution $\partial\Lambda/\partial x_k$ par l'identité*

$$\forall \varphi \in \mathcal{D}(\Omega), \qquad \left(\frac{\partial \Lambda}{\partial x_k}\right)\varphi = -\Lambda\left(\frac{\partial \varphi}{\partial x_k}\right).$$

On en déduit les définitions de ∇, ∇^α, Δ, $\nabla\cdot$, $\nabla\wedge$ etc. au sens des distributions.

L'application $\Lambda \longmapsto \nabla^\alpha \Lambda$ ainsi définie est continue. En particulier, si $\Lambda_k \longrightarrow \Lambda$ au sens des distributions, alors $\nabla^\alpha \Lambda_k \longrightarrow \nabla^\alpha \Lambda$ au sens des distributions. En outre,

$$\mathrm{supp}\,(\nabla\Lambda) \subset \mathrm{supp}\,(\Lambda);$$

$$\mathrm{ordre}\,(\nabla\Lambda) \leq \mathrm{ordre}\,(\Lambda) + 1.$$

Si Λ coïncide avec une fonction continûment différentiable, alors la distribution $\partial\Lambda/\partial x_k$ coïncide avec la dérivée usuelle $\{\partial\Lambda/\partial x_k\}$.

Remarque III-24. Noter que l'opérateur $\nabla\cdot$ est défini de $\mathcal{D}'(\Omega;\mathbb{R}^n) = \mathcal{D}'(\Omega)^n$ dans $\mathcal{D}(\Omega)$, alors que l'opérateur ∇ est défini de $\mathcal{D}'(\Omega)$ dans $\mathcal{D}'(\Omega;\mathbb{R}^n)$. L'adjoint de la divergence au sens des distributions est l'opposé du gradient des fonctions-test, et l'adjoint du gradient au sens des distributions est l'opposé de la divergence des fonctions-test.

Exemples III-25. (i) Soit $f(x) = |x|$ sur $\mathbb{R}$, alors f' est la fonction "signe", $f'' = 2\delta_0$, f''' vaut 2 fois l'application "évaluation de la dérivée en 0". Attention à ne pas confondre ces dérivées au sens des distributions avec les dérivées presque partout, qui sont nulles à patir du rang 2.

(ii) Plus généralement, une fonction de classe C^1 par morceaux étant donnée sur $\mathbb{R}$, sa dérivée au sens des distributions est donnée par la **formule des sauts** :

$$f' = \{f'\} + \sum_{x\in D} [f_+(x) - f_-(x)]\delta_x,$$

où D désigne l'ensemble des points de discontinuité de f. Cette formule se généralise en dimension supérieure à 1, comme suit : soit, dans un ouvert Ω de $\mathbb{R}^n$, une fonction de classe C^1 en-dehors d'une (hyper)surface Σ, régulière et orientée par une normale n, possédant un élément de surface σ. On suppose que, de part et d'autre de Σ, f se prolonge par continuité en une fonction de classe C^1 jusqu'au bord ; on note f_+ (resp. f_-) la restriction de ce prolongement défini du côté où Σ est orientée (resp. du côté opposé à l'augmentation de Σ). Alors,

$$\nabla f = \{\nabla f\} + [f_+ - f_-]n\,\sigma.$$

(iii) Soit $f(x) = x\log|x| - x$ dans $\mathbb{R}$, alors on peut écrire, au sens des distributions, $f'(x) = \log|x|$, $f''(x) = \text{v.p.}(1/x)$.

(iv) Pour $x \in \mathbb{R}^n$, on note $r = |x|$. Soient $g_1(r) = -|r|$, $g_2(r) = \log(1/r)$, et, pour tout $n \geq 3$, $g_n(r) = 1/r^{n-2}$. Alors, pour tout n il existe une constante $c_n > 0$ telle que, dans $\mathbb{R}^n$, on ait, au sens des distributions,

$$-\Delta[g_n] = c_n\delta_0.$$

Pour $n \geq 3$, on a $c_n = 2(n-2)\pi^{n/2}/\Gamma(n/2)$; on pourra le montrer en exercice, en admettant ou en redémontrant la formule de la surface de la sphère en dimension n.

(v) Soit (u_n) la suite de fonctions définie par $u_n(x) = (1/n)\sin(nx)$. Alors la famille (u_n') converge au sens des distributions vers 0.

Bien sûr, les dérivées successives sont en général des distributions de plus en plus singulières, comme on a pu le voir sur les exemples précédents. Dans le cadre des distributions, cela peut se quantifier par la notion d'ordre. Dans le cadre des espaces de Sobolev, la dérivation envoie $W^{k,p}$ dans $W^{k-1,p}$, et partant $W^{-k,q}$ dans $W^{-k-1,q}$.

Par dualité, beaucoup de propriétés agréables des fonctions-test se transmettent aux distributions. Ainsi,

Proposition III-26 (propriétés de la dérivation). *Soient Ω un ouvert de $\mathbb{R}^n$ $\Lambda \in \mathcal{D}'(\Omega)$. Alors*

(i) $\partial\Lambda/\partial x_k$ est limite de quotients différentiels : au sens des distributions, dans tout ouvert Ω' séparé du bord de Ω par une distance positive, on a

$$\frac{\partial \Lambda}{\partial x_k} = \lim_{t\to 0} \frac{\tau_{te_k}\Lambda - \Lambda}{t},$$

où $\tau_h\Lambda = \Lambda(\cdot + h)$ est la translation de Λ, définie par

$$(\tau_h\Lambda)\varphi = \Lambda[\varphi(\cdot - h)].$$

(ii) pour tous indices j et k,

$$\frac{\partial^2\Lambda}{\partial x_k\, x_j} = \frac{\partial^2\Lambda}{\partial x_j\, x_k}.$$

(iii) si $\Omega' \subset \Omega$ et $h \in \mathbb{R}^n$ est tel que $|h| < d(\Omega', \Omega^c)$, alors

$$\tau_h\Lambda - \Lambda = \int_0^1 (\nabla\Lambda \cdot h)\, dt.$$

En particulier, on peut écrire la formule de Taylor avec reste intégral à un ordre arbitraire.

Remarques III-27. Dans l'énoncé (i), la restriction à Ω' n'est pas vraiment indispensable, au sens où la limite du membre de droite est bien définie pour tout $\varphi \in \mathcal{D}(\Omega)$. Pour apprécier l'énoncé (ii), on peut se souvenir qu'il n'est pas vrai en général pour des dérivées usuelles de fonctions qui sont deux fois différentiables mais pas de classe C^2 ; on en déduit que pour ces fonctions, les dérivées distributions d'ordre 2 ne coïncident pas avec les dérivées usuelles. Enfin, dans l'énoncé (iii), noter l'emploi du paramètre de translation h pour “mimer” l'évaluation d'une fonction en deux points différents, opération qui n'a a priori guère de sens pour des distributions.

6.4. Produit tensoriel et évaluation paramétrée. Commençons par l'évaluation, ou crochet de dualité, dont la définition ne fait pas de mystère.

Théorème III-28 (continuité de l'évaluation). *Soit Ω un ouvert de $\mathbb{R}^n$. Alors l'application d'évaluation définit une opération bilinéaire continue de $\mathcal{D}'(\Omega) \times \mathcal{D}(\Omega)$ dans $\mathbb{R}$. En particulier, si $\Lambda_k \longrightarrow \Lambda$ au sens des distributions, et $\varphi_k \longrightarrow \varphi$ dans $\mathcal{D}(\Omega)$, alors*

$$\langle \Lambda_k, \varphi_k\rangle \longrightarrow \langle \Lambda, \varphi\rangle.$$

Exemple III-29. Si $\Lambda_k = \delta_{x_k}$, où $x_k \to x$, alors $\langle \Lambda_k, \varphi_k\rangle = \varphi_k(x_k) \longrightarrow \varphi(x)$.

On considère maintenant l'évaluation partielle, dans laquelle le crochet de dualité n'agit que sur une partie des variables, de sorte que le résultat est une fonction dépendant des autres variables, similaire dans sa définition et ses propriétés à une intégrale à paramètre. En particulier, on dispose d'analogues des théorèmes de dérivation sous l'intégrale et de Fubini.

Théorème III-30 (produit tensoriel et évaluation paramétrée). *Soient* $\Omega \subset \mathbb{R}^n$ *et* $\Omega' \subset \mathbb{R}^{n'}$ *deux ouverts, et soient* $\Lambda \in \mathcal{D}'(\Omega)$, $\Lambda' \in \mathcal{D}'(\Omega')$ *des distributions. Alors*

(i) il existe une unique distribution T *sur* $\Omega \times \Omega'$ *telle que, pour toutes fonctions* $\varphi \in \mathcal{D}(\Omega)$ *et* $\varphi' \in \mathcal{D}(\Omega')$, *on ait*

$$T(\varphi \otimes \varphi') = (\Lambda\varphi)(\Lambda'\varphi').$$

On appelle T *le produit tensoriel de* Λ *par* Λ' *et on le note* $\Lambda \otimes \Lambda'$. *En outre,*

$$\mathrm{supp}\,(\Lambda \otimes \Lambda') = \mathrm{supp}\,(\Lambda) \times \mathrm{supp}\,(\Lambda'),$$

$$\mathrm{ordre}\,(\Lambda \otimes \Lambda') \leq \mathrm{ordre}\,(\Lambda) + \mathrm{ordre}\,(\Lambda').$$

Enfin, l'application $(\Lambda, \Lambda') \longmapsto \Lambda\otimes\Lambda'$ *est continue de* $\mathcal{D}'(\Omega)\times\mathcal{D}'(\Omega')$ *dans* $\mathcal{D}'(\Omega\times\Omega')$.

(ii) Si Λ *et* Λ' *sont des mesures de Radon, le produit tensoriel de* Λ *et* Λ' *au sens des distributions coïncide avec le produit tensoriel au sens des mesures. En particulier, si* Λ *et* Λ' *sont des fonctions localement intégrables, le produit tensoriel* $\Lambda \otimes \Lambda'$ *coïncide avec le produit tensoriel usuel des fonctions* Λ *et* Λ'.

(iii) Si $\Phi \in \mathcal{D}(\Omega \times \Omega')$ *et* $\Lambda \in \mathcal{D}'(\Omega)$, *on peut définir*

$$\Lambda\Phi : y \longmapsto \langle \Lambda, \Phi(\cdot, y)\rangle \qquad \in \mathcal{D}(\Omega')$$

et

$$\Lambda'\Phi : x \longmapsto \langle \Lambda', \Phi(x, \cdot)\rangle \qquad \in \mathcal{D}(Om).$$

Alors,

$$\Lambda'(\Lambda\Phi) = \Lambda(\Lambda'\Phi) = (\Lambda \otimes \Lambda')\Phi.$$

En particulier, on peut "dériver et intégrer sous le crochet" :

$$\nabla_y(\Lambda\Phi) = \Lambda(\nabla_y\Phi), \qquad \int_y (\Lambda\Phi) = \Lambda(\int_y \Phi).$$

Remarque III-31. Souvent, l'ordre de $\Lambda \otimes \Lambda'$ sera exactement la somme des ordres de Λ et de Λ' ; mais ce n'est pas forcément le cas, comme on le voit avec $\Lambda = 0$.

Esquisse de démonstration. On va démontrer (i), (ii) et (iii) simultanément. Comme on va le voir, la démonstration suit presque l'ordre inverse de l'énoncé.

1. Dans un premier temps, on montre que

$$\nabla_y(\Lambda\Phi) = \Lambda(\nabla_y\Phi);$$

si c'est vrai, on en déduira facilement par récurrence que $\Lambda\Phi$ est de classe C^∞ ; comme le support de Φ peut s'inclure dans un produit de compacts, la fonction $\Lambda\Phi$ est nulle en-dehors d'un compact de $\Omega \times \Omega'$, et on en déduira que $\Lambda\Phi$ est bien un élément de $\mathcal{D}$.

2. La démonstration se ramène facilement au cas où le paramètre y varie dans $\Omega' =]-1, 1[$, et on veut établir que la dérivée en $y = 0$ vaut $\Lambda\partial_y\Phi(\cdot, 0)$. Pour cela on écrit

$$\Lambda\Phi(\cdot, y) - \Lambda\Phi(\cdot, 0) - y\Lambda\partial_y\Phi(\cdot, 0)$$
$$= y^2\Lambda\left[\frac{\Phi(\cdot, y) - \Phi(\cdot, 0) - y\Phi(\cdot, 0)}{y^2}\right].$$

La formule de Taylor avec reste intégral montre que la fonction

$$\psi(x,y) = \frac{\Phi(x,y) - \Phi(x,0) - y\Phi(x,0)}{y^2}$$

est de classe C^∞ ; comme elle est à support compact, elle est dans $\mathcal{D}(\Omega \times \Omega')$. On inclut son support dans un produit de compacts $K \times K'$, et on introduit k l'ordre de Λ restreinte aux fonctions à support dans K. La famille de fonctions $\psi(\cdot, y)$ a toutes ses dérivées jusqu'à l'ordre k bornées, uniformément en y ; la famille des nombres $\Lambda\psi(\cdot, y)$ est donc bornée. On en déduit que

$$\Lambda\Phi(\cdot,y) - \Lambda\Phi(\cdot,0) - y\Lambda\partial_y\Phi(\cdot,0) = O(y^2),$$

ce qui conclut la preuve de la différentiabilité.

3. On peut alors considérer l'opération $\Phi \longmapsto \Lambda\Lambda'\Phi$, qui est une forme linéaire sur $\mathcal{D}(\Omega \times \Omega')$. Si Φ est à support dans $K \times K'$, soit k l'ordre de Λ sur K et k' celui de Λ' sur K'. Le nombre $\Lambda\Lambda'\Phi$ se majore en fonction des dérivées de $\Lambda'\Phi$ d'ordre au plus k, qui sont données par l'évaluation de Λ' aux dérivées de Φ d'ordre au plus k, et se majorent donc par des normes uniformes de dérivées de Φ à un ordre au plus $k + k'$. Quand on fait varier K et K', cette estimation montre à la fois que l'opération ainsi définie est une distribution, et que son ordre est au plus la somme des ordres de Λ et de Λ'.

4. L'objet T ainsi défini est donc bien une distribution, et, grâce à la linéarité de Λ', vérifie bien sûr l'identité

$$T(\varphi \otimes \varphi') = (\Lambda\varphi)(\Lambda'\varphi').$$

Il reste à démontrer qu'il ne peut y avoir qu'une seule telle distribution T. C'est une conséquence du lemme énoncé ci-après, d'intérêt indépendant. La partie (ii) de l'énoncé est également conséquence de ce lemme.

5. La continuité de l'opération de produit tensoriel par rapport à chacun de ses arguments est facile ; la continuité de l'opération bilinéaire s'ensuit alors par des arguments topologiques abstraits (délicats). □

Lemme III-32 (tensorisation des fonctions test). *Soient Ω et Ω' des ouverts de $\mathbb{R}^n$ et $\mathbb{R}^m$ respectivement. Alors*

$$\mathcal{D}(\Omega) \otimes \mathcal{D}(\Omega') = \mathcal{D}(\Omega \times \Omega') :$$

l'ensemble des combinaisons linéaires de produits tensoriels $\varphi \otimes \varphi'$ est dense dans l'ensemble des fonctions-test dans les deux variables.

De manière plus explicite : étant donnés K et K' compacts de Ω et Ω' respectivement, $\ell \in \mathbb{N}$, et $\Phi \in \mathcal{D}_{K\times K'}(\Omega \times \Omega')$, pour tout $\varepsilon > 0$ on peut trouver une fonction Φ_ε de la forme

$$\Phi_\varepsilon(x,y) = \sum_{k=1}^{N} \varphi_k(x)\varphi'_k(y),$$

où les φ_k (resp. φ'_k) sont des éléments de $\mathcal{D}_K(\Omega)$ (resp. $\mathcal{D}_{K'}(\Omega')$), et

$$|\alpha| \leq \ell \implies \|\Phi - \Phi_\varepsilon\|_\infty \leq \varepsilon.$$

En particulier, une distribution sur $\Omega \times \Omega'$ est uniquement déterminée par sa valeur sur les produits de fonctions de la forme $\varphi \otimes \varphi'$, $\varphi \in \mathcal{D}(\Omega)$, $\varphi' \in \mathcal{D}(\Omega')$.

On peut prouver ce lemme de manière élémentaire, en utilisant le théorème d'approximation polynomiale uniforme de Weierstrass, en plusieurs variables (en prenant soin d'approcher d'abord des dérivées d'ordre élevé et en intégrant ensuite autant de fois qu'il le faut). En effet, les polynômes se décomposent comme des sommes de produits tensoriels. Ces fonctions ne sont pas à support compact, mais on peut s'y ramener par troncature.

6.5. Multiplication. La multiplication des distributions pose des problèmes considérables et ne se résout que partiellement.

Définition III-33 (multiplication d'une distribution par une fonction régulière). *Soient* Ω *un ouvert de* $\mathbb{R}^n$, $\psi \in C^\infty_{\text{loc}}(\Omega)$ *et* $\Lambda \in \mathcal{D}'(\Omega)$. *On définit alors la distribution* $\psi\Lambda$ *par la formule*

$$\forall \varphi \in \mathcal{D}(\Omega), \qquad \langle \psi\Lambda, \varphi\rangle = \langle \Lambda, \psi\varphi\rangle.$$

Cette définition est bien sûr licite puisque le produit de deux fonctions C^∞ est C^∞, et à support compact dès que l'une des deux fonctions est à support compact.

Exemples III-34. (i) Dans $\mathbb{R}$, $x\delta = 0$, un exemple qui peut être surprenant à première vue.

(ii) Toujours dans $\mathbb{R}$, $x\text{v.p.}(1/x) = 1$ au sens des distributions, puisque par définition

$$\langle x\text{v.p.}(1/x), \varphi\rangle = \langle \text{v.p.}(1/x), x\varphi\rangle = \lim_{\varepsilon\to 0} \int_{|x|>\varepsilon} \varphi(x)\, dx = \int \varphi.$$

Nous admettrons le théorème suivant, dont le seul point un peu délicat est la continuité bilinéaire du produit.

Proposition III-35 (propriétés de la multiplication). *Soit* Ω *un ouvert de* $\mathbb{R}^n$. *La multiplication, au sens des distributions, d'une fonction* C^∞ *par une fonction localement intégrable coïncide avec l'opération habituelle de multiplication ponctuelle. Plus généralement, la multiplication, au sens des distributions, d'une fonction* ψ *de classe* C^∞ *par une mesure* μ *coïncide avec la mesure* $\psi\mu$.

La multiplication est une opération continue. En outre, pour tous $\psi \in C^\infty_{\text{loc}}(\Omega)$, $\Lambda \in \mathcal{D}'(\Omega)$,

$$\text{supp}\,(\psi\Lambda) \subset \text{supp}\,(\psi) \bigcap \text{supp}\,(\Lambda),$$

$$\text{ordre}\,(\psi\Lambda) \leq \text{ordre}\,(\Lambda).$$

Ces résultats règlent la question de la multiplication par des fonctions très régulières. Pour le reste, on en est réduit à une discussion au cas par cas, basée sur des recettes plus ou moins astucieuses.

Exemples III-36. (i) Soient $f \in L^p_{\text{loc}}(\Omega)$, $g \in L^q_{\text{loc}}(\Omega)$, avec $(1/p)+(1/q) = 1$, $p, q \geq 1$, alors $fg \in L^1_{\text{loc}}(\Omega)$, et on décide que le produit de distributions de f par g coïncide avec le produit usuel.

(ii) Soient $\Lambda \in W^{-k,p}_{\text{loc}}(\Omega)$, $f \in W^{k,q}_{\text{loc}}$, alors on peut définir la distribution $f\Lambda$ par la formule

$$\forall \varphi \in \mathcal{D}(\Omega), \qquad \langle f\Lambda, \varphi\rangle := \langle \Lambda, f\varphi\rangle,$$

ce qui est licite puisque $f\varphi \in W^{k,q}_c$ par la formule de Leibniz...

(iii) Dans $\mathbb{R}$, on considère $f(x) = g(x) = 1/\sqrt{|x|}$, alors $fg(x) = 1/|x| \notin L^1_{\text{loc}}(\mathbb{R})$. Dans ce cas, le produit n'est pas défini a priori, et il y a peu de chances que l'on parvienne à le définir d'une manière "raisonnable".

(iv) L'exemple, ou plutôt le contre-exemple, suivant, a permis à Schwartz en 1954 de démontrer "l'impossibilité de la multiplication des distributions" en général. On se place dans $\mathbb{R}$. Comme on a déjà vu, on a $x\text{v.p.}(1/x) = 1$ et $x\delta = 0$ au sens des distributions. Si l'on parvient à définir un produit raisonnable sur l'espace des distributions, alors

$$\delta(x\text{v.p.}(1/x)) = \delta, \qquad (\delta x)\text{v.p.}(1/x) = 0,$$

ce qui semble écarter tout espoir d'une telle théorie. Notons que la distribution v.p.$(1/x)$ est la dérivée seconde d'une fonction continue : si on l'écarte dans l'espoir de permettre la définition d'un produit général, on perd soit la propriété de stabilité par dérivation, soit l'appartenance des fonctions continues à l'espace des distributions, deux des conditions importantes du cahier des charges de la théorie.

(v) Soit H la fonction de Heaviside : $H(x) = 1_{x\geq 0}$. Il est plus ou moins clair (par localisation...) que le produit de H par δ_x devrait être la distribution nulle pour $x < 0$, et la distribution δ_x pour $x > 0$. On cherche maintenant à définir le produit de H par δ_0, au sens des distributions. Si le produit est une opération continue, on devrait avoir $H\delta_0 = \lim H\delta_{1/n} = \delta_0$, mais aussi $H\delta_0 = \lim H\delta_{-1/n} = 0$, ce qui est bien sûr contradictoire. On en conclut qu'on ne peut probablement pas définir de notion raisonnable de produit de H par δ_0.

(vi) Considérons maintenant, dans le plan (x, y), la multiplication $H(x)\delta_{x=0}$: le problème se pose de manière aussi insoluble qu'auparavant. En revanche, il est facile de définir $H(y)\delta_{x=0}$: c'est la masse de Dirac "uniforme" portée par la demi-ligne verticale $(x = 0, y \geq 0)$. On voit sur cet exemple que le degré de "singularité" des deux distributions en jeu n'est pas seule en cause, et que la géométrie locale de la singularité joue aussi un rôle important. Ici la multiplication est possible car les directions de singularité pour les deux distributions sont orthogonales. Cette remarque est la base de la théorie des fronts d'onde, qui permet de multiplier certaines distributions après examen de la géométrie de leurs singularités.

Signalons pour conclure cette discussion l'existence de certains procédés élaborés de multiplication des distributions, basés sur l'analyse de Littlewood-Paley, en particulier l'**algorithme du paraproduit**, développé par Bony.

6.6. Produit de convolution et dérivation fractionnaire. La définition du produit de convolution se fait sans douleur à partir du produit tensoriel. Cependant, il faut prendre garde aux problèmes considérables que peuvent poser des supports non compacts ! Ces problèmes se posent déjà avec la convolution des fonctions (comment définir la fonction $1 * 1$?), ils se posent pour les distributions avec encore plus d'acuité et de manière plus sournoise. Ainsi, consid'erons, sur $\mathbb{R}$, les trois distributions 1, δ' et H. L'élément δ étant élément neutre pour la convolution (comme on s'en convainc facilement en écrivant la définition), et la convolution devant commuter avec la dérivation, on a

$$(1 * \delta') * H = 1' * H = 0 * H = 0;$$

$$1 * (\delta' * H) = 1 * \delta = 1.$$

Le problème ici vient de ce que les supports de 1 et de H sont tous deux infinis dans une direction. On remédie à cela en imposant une hypothèse de **convolutivité des supports**.

Définition III-37 (supports convolutifs). *Soient $F_1, \ldots, F_k$ des fermés de $\mathbb{R}^n$. On dit qu'ils sont convolutifs si pour tout $R \geq 0$ il existe $\rho(R) \geq 0$ tel que*

$$\forall (x_1, \ldots, x_k) \in F_1 \times \ldots \times F_k, \quad |\sum x_j| \leq R \Longrightarrow \forall j,\ |x_j| \leq \rho(R).$$

Exemples III-38. (i) Deux demi-droites de $\mathbb{R}$ sont convolutives si et seulement si elles s'étendent à l'infini dans des directions opposées.

(ii) Si l'on se donne k fermés dont tous, sauf un, sont compacts, alors ils sont convolutifs.

Théorème III-39 (produit de convolution des distributions). *(i) Soient $\Lambda, \Lambda' \in \mathcal{D}'(\mathbb{R}^n)$. On suppose que Λ est à support compact. On peut alors définir la distribution $\Lambda * \Lambda'$ par la formule*

$$\langle \Lambda * \Lambda', \varphi \rangle = \langle \Lambda, \psi \rangle,$$

$$\psi(x) = \langle \Lambda', \varphi(x - \cdot) \rangle;$$

soit encore, avec des notations peu rigoureuses mais claires,

$$\langle \Lambda * \Lambda', \varphi \rangle = \langle \Lambda(x)\, \Lambda'(y), \varphi(x - y) \rangle.$$

On a alors

$$\operatorname{supp}(\Lambda * \Lambda') \subset \operatorname{supp}(\Lambda) + \operatorname{supp}(\Lambda'),$$

$$\operatorname{ordre}(\Lambda * \Lambda') \leq \operatorname{ordre}(\Lambda) + \operatorname{ordre}(\Lambda').$$

(ii) Si Λ et Λ' sont des fonctions L^1_{loc}, l'une d'entre elles étant à support compact, le produit de convolution $\Lambda * \Lambda'$ coïncide avec le produit de convolution usuel.

(iii) L'opération de convolution commute avec les translations et avec les dérivations à tous ordres ;

(iv) Ces définitions et propriétés s'étendent à un nombre fini arbitraire de distributions dont les supports sont convolutifs.

(v) Si l'on se donne des fermés convolutifs $F_1, \ldots, F_k$, l'opération de convolution $(\Lambda_1, \ldots, \Lambda_k) \longmapsto \Lambda_1 * \ldots * \Lambda_k$ est associative, commutative et continue si on la restreint aux distributions Λ_i dont le support est contenu dans F_i.

On admet encore ce théorème, dont certaines parties sont faciles et d'autres beaucoup plus délicates.

Exemple III-40. Soit $\Lambda \in \mathcal{D}'(\mathbb{R}^n)$ une distribution quelconque, alors $\delta * \Lambda = \Lambda * \delta = \Lambda$; $\nabla \delta * \Lambda = \nabla \Lambda$, etc.

6.7. Régularisation.

Définition III-41 (approximation régulière de δ). *On appelle approximation régulière de δ, ou approximation de l'identité pour la convolution, une famille $(\varphi_t)_{t>0}$ de la forme*

$$\varphi_t(x) = \frac{1}{t^n}\varphi\left(\frac{x}{t}\right),$$

où $\varphi \in \mathcal{D}(\mathbb{R}^n)$, $\operatorname{supp}\varphi \subset B(0,1)$, $\varphi \geq 0$, $\int \varphi = 1$, φ radialement symétrique.

Remarque III-42. La propriété de symétrie radiale ne jouera aucun rôle pour nous, mais elle est parfois précieuse dans des problèmes subtils.

On vérifie facilement qu'une approximation régulière (φ_t) de δ converge vers la masse de Dirac δ_0, au sens de la convergence faible des mesures (avec l'espace C_b comme espace de fonctions-test), donc a fortiori au sens des distributions. La fonction φ_t étant de classe C^∞ et à support compact, on peut la convoler avec n'importe quelle distribution. La continuité de l'opération de convolution implique alors le théorème suivant.

Théorème III-43 (approximation par convolution). *Soit $(\varphi_t)_{t>0}$ une approximation régulière de δ. Alors*

(i) Soit $\Lambda \in \mathcal{D}'(\mathbb{R}^n)$, alors $\Lambda * \varphi_t$ définit une famille de fonctions de classe C^∞ qui converge au sens des distributions vers Λ quand $t \to 0$.

(ii) Soient Ω un ouvert de $\mathbb{R}^n$, et Λ une distribution sur Ω, à support compact K. Pour $t > 0$ assez petit ($t < \operatorname{dist}(K, \partial\Omega)$), le produit de convolution $\Lambda * \varphi_t$ est bien défini comme dans $\mathcal{D}(\Omega)$, et converge au sens des distributions vers Λ quand $t \to 0$.

Corollaire III-44 (densité de $\mathcal{D}$ dans $\mathcal{D}'$). *Soit Ω un ouvert de $\mathbb{R}^n$ et $\Lambda \in \mathcal{D}'(\Omega)$. Soient $(\varphi_t)_{t>0}$ une approximation de l'unité, $(K_p)_{p\in\mathbb{N}}$ une suite exhaustive de compacts dans Ω, et χ_p une famille de fonctions plateaux dans $\mathcal{D}(\Omega)$, χ_p étant identiquement égale à 1 au voisinage de K_p et nulle hors de K_{p+1}. Pour tout $p \geq 1$, on choisit $t(p) \in (0, \operatorname{dist}(K_{p+1}, \partial\Omega))$. Alors*

$$\chi_p \varphi_{t(p)} * \Lambda \xrightarrow[p\to\infty]{} \Lambda, \qquad \text{au sens des distributions.}$$

En particulier, $\mathcal{D}(\Omega)$ est dense dans $\mathcal{D}'(\Omega)$.

Remarque III-45. Ces résultats généraux d'approximation ne sont pas très précis au sens où ils ne fournissent que la convergence au sens des distributions. Dès qu'on a des informations sur la régularité de la distribution limite Λ, on peut renforcer le sens de la convergence. Par exemple, si Λ appartient à L^p_{loc} ($1 \leq p < \infty$) ou C_{loc}, alors la convergence a lieu dans le même espace. En utilisant la commutation entre convolution et dérivation, et un peu d'interpolation, on montre facilement que si Λ appartient à un espace parmi $\{L^p_{\mathrm{loc}}, H^k_{\mathrm{loc}}, W^{s,p}_{\mathrm{loc}}, C^\alpha_{\mathrm{loc}}, C^{k,\alpha}_{\mathrm{loc}}, H^{-k}_{\mathrm{loc}}, W^{-s,p}_{\mathrm{loc}}$, alors la convergence a lieu dans ce même espace. Il n'en va pas de même pour des espaces construits à partir des mesures de Radon ou de L^∞. Par exemple, même si Λ est une mesure à support compact, le produit de convolution $\Lambda * \varphi_t$ ne converge en général vers Λ qu'au sens faible, et non au sens de la variation totale. De même, si Λ est une fonction L^∞, la convergence de $\Lambda * \varphi_t$ vers Λ a lieu au sens faible-$*$, mais certainement pas au sens de la convergence uniforme – sinon Λ serait une fonction *continue*.

6.8. Division. Le problème de la division est facile à comprendre : étant données deux distributions N et D (“numérateur” et “dénominateur”), caractériser toutes les distributions Λ telles que $D\Lambda = N$. Afin de pouvoir définir a priori le produit $D\Lambda$, il est naturel de supposer que D est une fonction C^∞. Dans tout ouvert où D ne s'annule pas, on peut écrire $\Lambda = N/D$, et tous les problèmes proviennent donc des *annulations du dénominateur.*

Exemple III-46. Dans $\mathbb{R}$, considérons la division de 1 par x. Toute distribution Λ vérifiant $x\Lambda = 1$ doit coïncider avec la fonction $1/x$ dans $\mathbb{R}\setminus\{0\}$. Une solution est la distribution v.p.$(1/x)$, mais il reste encore de la place pour ajouter une distribution dont le support serait $\{0\}$. De fait, puisque $x\delta = 0$, n'importe quelle distribution de la forme

$$\Lambda = \text{v.p.}(1/x) + c\delta_0, \qquad c \in \mathbb{R}$$

est solution du problème.

Pour prouver que ce sont les seules, il suffit de démontrer le

Lemme III-47 (division de 0). *Soit Λ une distribution sur $\mathbb{R}$, telle que $x\Lambda = 0$. Alors $\Lambda \in \mathbb{R}\delta_0$. Plus généralement, si $x^\ell\Lambda = 0$ pour un certain entier $\ell \geq 1$, alors Λ est combinaison linéaire de dérivées de δ d'ordre inférieur ou égal à ℓ.*

Ce lemme deviendra évident une fois que nous aurons caractérisé les distributions dont le support est réduit à un point. Avec son aide, on peut montrer le résultat suivant [Schwartz, p. 123 et 125] :

Théorème III-48 (division dans $\mathbb{R}$). *Soient N une distribution sur $\mathbb{R}$ et ℓ un entier supérieur ou égal à 1. Alors il existe une infinité de distributions Λ solutions de*

$$x^\ell\Lambda = N,$$

et deux quelconques d'entre elles diffèrent par une combinaison linéaire de dérivées de δ d'ordre au plus ℓ.

Avec ce théorème, on résout sans difficulté le problème de la division par une fonction dont les zéros sont isolés et d'ordre fini, voir [Schwartz, p. 125].

Le problème de la division dans les dimensions supérieures est beaucoup plus délicat, même s'il admet des cas particuliers simples. On trouvera une discussion et des références dans [Schwartz, pp. 126-128].

6.9. Transformée de Fourier. La transformée de Fourier est une opération linéaire, dont la définition ne devrait donc poser aucun problème par transposition. Une subtilité importante apparaît cependant : pour définir une notion de transformée de Fourier adaptée à l'analyse fonctionnelle, on souhaite avoir accès à l'identité

$$\langle f, \overline{g}\rangle = \langle \widehat{f}, \overline{\widehat{g}}\rangle.$$

On souhaite également que la transformée de Fourier d'une distribution soit définie en tant que distribution. Mais ce n'est pas parce que $g \in \mathcal{D}(\mathbb{R}^n)$ que $\widehat{g} \in \mathcal{D}(\mathbb{R}^n)$! En fait, il est impossible à une fonction régulière non nulle d'avoir une transformée de Fourier à support compact ; le mieux que l'on puisse espérer de la transformée de Fourier d'une fonction C^∞ à support compact, c'est de décroître plus vite que toute puissance inverse de $|x|$, comme on le voit en intégrant par parties.

On va donc agrandir la classe des fonctions-test utilisées pour définir la transformée de Fourier ; ce faisant, on réduira automatiquement la classe des distributions concernées. La nouvelle classe de fonctions-test est appelée **classe de Schwartz**, et la nouvelle classe de distributions est l'espace des **distributions tempérées**.

DÉFINITION III-49 (classe de Schwartz). *On appelle classe de Schwartz de* $\mathbb{R}^n$*, et on note* $\mathcal{S}(\mathbb{R}^n)$ *ou* $\mathcal{S}$*, l'espace des fonctions-test de classe* C^∞*, décroissant à l'infini plus vite que toute puissance inverse de* $|x|$*. Muni des semi-normes*

$$p_k(f) := \sup_{|\alpha|\le k} \sup_{x\in\mathbb{R}^n} (1+|x|)^k |\nabla^\alpha f(x)|,$$

$\mathcal{S}(\mathbb{R}^n)$ *est un espace de Fréchet.*

Les deux théorèmes suivants expriment tout l'intérêt de la classe de Schwartz. Ils sont assez faciles [Rudin, chapitre 7], du moins si on dispose déjà du théorème d'inversion dans L^2...

THÉORÈME III-50 (densité de $\mathcal{S}$). *L'espace* $\mathcal{S}(\mathbb{R}^n)$ *s'injecte continûment et densément dans l'espace* $\mathcal{D}(\mathbb{R}^n)$*.*

THÉORÈME III-51 (stabilité de $\mathcal{S}$). *L'espace* $\mathcal{S}(\mathbb{R}^n)$ *est stable par la transformée de Fourier* $f \longmapsto \widehat{f} = \mathcal{F}f$*, où*

$$\widehat{f}(\xi) := \frac{1}{(2\pi)^{n/2}} \int e^{-i\xi\cdot x} f(x)\,dx.$$

Il est également stable par dérivation, convolution et multiplication par des polynômes de plusieurs variables.

Toutes les formules établies en analyse de Fourier dans l'espace $L^2(\mathbb{R}^n)$ *restent vraies dans l'espace* $\mathcal{S}(\mathbb{R}^n)$*, et en particulier on a les relations d'isométrie*

$$\forall f,g \in \mathcal{S}(\mathbb{R}^n), \qquad \int f\widehat{g} = \int \widehat{f}\widehat{\overline{g}},$$

$$\forall f,g \in \mathcal{S}(\mathbb{R}^n), \qquad \mathcal{F}(f*g) = (2\pi)^{-n}\widehat{f}\widehat{g},$$

$$\forall f \in \mathcal{S}(\mathbb{R}^), \qquad \mathcal{F}(\nabla^\alpha f)(\xi) = (-i)^|\alpha|\xi^\alpha \widehat{f}(\xi).$$

$$\overline{\mathcal{F}}\mathcal{F} = \mathrm{Id},$$

où $\overline{\mathcal{F}}$ *est définie par la même expression que la transformée de Fourier, au remplacement près de* i *par* $-i$*.*

DÉFINITION III-52 (distributions tempérées). *Les distributions tempérées sur* $\mathbb{R}^n$ *sont les distributions qui s'étendent en des formes linéaires continues sur* $\mathcal{S}(\mathbb{R}^n)$*. On note* $\mathcal{S}'(\mathbb{R}^n) = \mathcal{S}'$ *l'espace des distributions tempérées.*

Cette définition n'est pas très explicite : c'est “le plus gros” espace fonctionnel sur lequel on peut transporter à faible prix toutes les propriétés de $\mathcal{S}$, par dualité.

EXEMPLES III-53. Voici quelques distributions tempérées :

- tous les polynômes, et plus généralement toutes les fonctions qui croissent au plus polynomialement à l'infini ;
- toutes les fonctions qui s'écrivent comme le produit d'une fonction de L^p et d'un polynôme ;
- les distributions à support compact ;
- les mesures de Radon de μ telles que $\int (1+|x|^2)^{-k}\,d|\mu|(x) < \infty$ pour un certain $k < \infty$;
- et toutes leurs dérivées.

Remarque III-54. En premire approximation, $\mathcal{S}'$ contient donc des distributions qui ne croissent pas trop vite l'infini (au plus polynomialement, en un certain sens). Mais attention, certaines distributions peuvent appartenir S' grce des **oscillations** l'infini : ainsi, e^x ne dfinit pas une distribution tempre sur $\mathbb{R}$, mais $e^x \cos(e^x)$ si.

Théorème III-55 (transformée de Fourier dans $\mathcal{S}'$). *Soit $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$. On définit la transformée de Fourier $\mathcal{F}\Lambda$ de Λ par la formule*

$$\forall \varphi \in \mathcal{S}(\mathbb{R}^n), \qquad \langle \mathcal{F}\Lambda, \varphi \rangle = \langle \Lambda, \psi \rangle,$$
$$\psi(x) := \frac{1}{(2\pi)^{n/2}} \int e^{ix\cdot\xi} \varphi(-\xi)\, d\xi.$$

*(***à vérifier***) C'est une distribution tempérée.*

La transformée de Fourier ainsi définie est une bijection de $\mathcal{S}'$ dans $\mathcal{S}'$. De plus, si $\varphi \in \mathcal{S}(\mathbb{R}^n)$ et $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$, alors $\varphi * u \in C^\infty_{\mathrm{loc}}(\mathbb{R}^n)$, a une croissance au plus polynomiale, et

$$\widehat{\varphi\Lambda} = (2\pi)^{-n}(\widehat{\varphi} * \widehat{\Lambda}).$$

Si $f \in L^1 + L^2(\mathbb{R}^n)$, la transformée de f au sens des distributions coïncide avec la transformée de Fourier usuelle de f.

Exemples III-56. - la transformée de Fourier d'une masse de Dirac (de masse arbitraire) est une constante, et vice versa ;

- dès que $\alpha \neq 0$, la transformée de Fourier de la distribution p.f.$(|x|^\alpha)$ est un multiple de la distribution p.f.$(|x|^{-(n+\alpha)})$, et vice versa ; le symbole de partie finie étant inutile dans certains cas ;

6.10. Transformée de Laplace. Soit une fonction sur $\mathbb{R}_+$, à croissance au plus exponentielle, au sens où il existe $\lambda_0 > 0$ tel que

$$\forall t \in \mathbb{R}_+, \qquad |f(t)| \le Ce^{\lambda_0 t};$$

alors la transformée de Laplace Lf de f est définie par la formule

$$Lf(\lambda) := \int e^{-\lambda t} f(t)\, dt,$$

pour tout $\lambda \in \mathbb{C}$, de partie réelle strictement supérieure à λ_0. C'est une fonction holomorphe dans ce demi-plan. Le principal intérêt de la transformée de Laplace est de convertir, tout comme la transformée de Fourier, la convolution en multiplication.

Il est facile de généraliser la définition habituelle à une classe assez large de distributions :

Définition III-57 (transformée de Laplace des distributions). *Soit $\Lambda \in \mathcal{D}(\mathbb{R})$, à support inclus dans $\mathbb{R}_+$, telle que $e^{-\lambda_0 \cdot}\Lambda \in \mathcal{S}'(\mathbb{R})$ pour un certain $\lambda_0 > 0$. On définit sa transformée de Laplace $L\Lambda$ sur le demi-plan $(\Re z > \lambda_0)$ par la formule*

$$L\Lambda(z) := \langle e^{-\lambda_0 t}\Lambda,\, e^{(\lambda_0 - \lambda)t}\,\chi \rangle,$$

où χ est une fonction plateau à support dans $[-1, +\infty[$, identiquement égale à 1 au voisinage de $\mathbb{R}_+$.

La transformée ainsi définie vérifie la propriété

$$L(\Lambda_1 * \Lambda_2) = (L\Lambda_1)(L\Lambda_2).$$

Remarque III-58. Soit $\mathcal{A}$ l'ensemble de toutes les distributions de la forme $\Lambda + f$, où Λ est une distribution sur $\mathbb{R}$, à support compact inclus dans $\mathbb{R}_+$, et f est une fonction continue sur $\mathbb{R}_+$, identiquement nulle au voisinage de 0, telle que $e^{-\lambda_0 t} f(t)$ appartient à $\mathcal{S}(\mathbb{R})$ pour un $\lambda_0 > 0$. On peut montrer que $\mathcal{A}$ est stable par convolution, et définir facilement la transformée de Laplace sur $\mathcal{A}$.

6.11. Composition et changement de variables. - permet de définir des distributions sur des variétés, etc.
- deux formalismes : image directe et composition
- formules pour les mesures, et on étend aux distributions

6.12. Dérivation des fonctions composées. - position du problème
- cas des fonctions régulières
- par approximation, non-linéarités C^1, bornées et Lipschitz
- plus général ?

III-7. Théorèmes de structure

Dans cette section, nous allons passer en revue les théorèmes les plus connus concernant la caractérisation des distributions vérifiant certaines propriétés. Beaucoup de ces résultats ont un intérêt purement théorique.

7.1. Théorème de l'inclusion des noyaux. Le premier résultat n'a rien à voir avec la structure de distribution, mais uniquement avec celle de forme linéaire ; cependant on l'applique assez fréquemment dans le contexte des distributions.

Théorème III-59 (théorème de l'inclusion des noyaux). *Soient E un espace vectoriel, et $\Lambda, \Lambda_1, \ldots, \Lambda_n$ des formes linéaires sur E. Alors $\Lambda \in \mathrm{Vect}\,(\Lambda_1, \ldots, \Lambda_n)$ si et seulement si*

$$\bigcap_{1 \leq i \leq n} \ker \Lambda_i \subset \ker \Lambda.$$

Démonstration. Une implication est évidente, et on montre seulement la réciproque : si l'intersection des $\ker \Lambda_i$ est incluse dans $\ker \Lambda$, alors Λ appartient à l'espace vectoriel engendré par les Λ_i.

Soit $T : E \to \mathbb{R}^n$ qui à x associe le vecteur $(\Lambda_1(x), \ldots, \Lambda_n(x))$. L'hypothèse de l'énoncé se traduit par $\ker T \subset \ker \Lambda$. Par le théorème de factorisation des applications linéaires, il existe une forme linéaire ϕ sur $\mathbb{R}^n$ telle que $\phi \circ T = \Lambda$. On peut représenter la forme linéaire ϕ comme le produit scalaire par un vecteur $(\lambda_1, \ldots, \lambda_n)$, ce qui donne $\sum \lambda_i \Lambda_i = \Lambda$. □

7.2. Distributions positives.

Théorème III-60 (les distributions positives sont les mesures). *Soient Ω un ouvert de $\mathbb{R}^n$, et Λ une distribution dans Ω, positive au sens où, pour tout $\varphi \in \mathcal{D}(\Omega)$,*

$$\varphi \geq 0 \Longrightarrow \Lambda\varphi \geq 0.$$

Alors Λ est une mesure positive.

Remarque III-61. La réciproque étant évidente, le théorème montre que les distributions positives sont exactement les mesures positives.

Démonstration. Soit K un compact de Ω, on appelle $C_0(K)$ l'adhérence de l'ensemble $\mathcal{D}_K$ pour la norme uniforme. On fixe une suite exhaustive $(K_j)_{j\in\mathbb{N}}$ de compacts dans Ω, et on note $C_0(K_j)$ l'espace des fonctions continues dans Ω, à support compact inclus dans K_j. Un argument facile de régularisation $\mathcal{D}_{K_{j+1}}$ est dense dans $C_0(K_j)$. Si l'on montre que Λ est une forme linéaire continue sur $\mathcal{D}_{K_{j+1}}(\Omega)$, muni de la norme de la convergence uniforme, alors elle se prolongera de manière unique en une forme linéaire continue sur $C_0(K_j)$; ce prolongement sera nécessairement positif. Tous les prolongements ainsi définis de Λ aux $C_0(K_j)$ constitueront une forme linéaire positive sur l'espace $C_c(\Omega)$ des fonctions continues dans Ω à support compact, et par le théorème de Riesz cette forme linéaire s'identifiera à une mesure positive. Tout le problème consiste donc à établir une inégalité de la forme

$$\forall \varphi \in \mathcal{D}_K, \qquad |\Lambda\varphi| \leq C\|\varphi\|_\infty.$$

Pour cela on introduit une fonction plateau, à support compact dans Ω et identiquement égale à 1 au voisinage de K. On a alors

$$\varphi \leq \chi\|\varphi\|_\infty,$$

d'où par positivité et linéarité

$$\Lambda\varphi \leq (\Lambda\chi)\|\varphi\|_\infty.$$

En changeant φ en $-\varphi$, on obtient

$$|\Lambda\varphi| \leq (\Lambda\chi)\|\varphi\|_\infty,$$

ce qui prouve le résultat. □

Remarque III-62. La mesure peut être de masse infinie, auquel cas les quantités $\Lambda\chi$ tendent vers l'infini quand $j \to \infty$.

7.3. Distributions à support compact. Nous allons voir deux théorèmes concernant les distributions à support compact : d'une part, leur ordre k est toujours fini ; d'autre part, si une fonction-test φ a toutes ses dérivées jusqu'à l'ordre k identiquement nulles sur le support de la distribution Λ, alors $\Lambda\varphi = 0$.

Théorème III-63 (support compact implique ordre fini). *Soient Ω un ouvert de $\mathbb{R}^n$ et $\Lambda \in \mathcal{D}'(\Omega)$ une distribution à support compact K. Alors il existe un compact K' de Ω, un entier $k \in \mathbb{N}$ et une constante $C \geq 0$, tels que pour tout $\varphi \in \mathcal{D}(\Omega)$,*

$$|\Lambda\varphi| \leq C \sup_{|\alpha|\leq k} \sup_{x\in K'} |\nabla^\alpha \varphi(x)|.$$

Démonstration. On introduit une fonction plateau χ, identiquement égale à 1 au voisinage de K, et on pose $K' := \operatorname{supp} \chi$. Pour tout $\varphi \in \mathcal{D}(\Omega)$, on écrit

$$\varphi = \chi\varphi + (1-\chi)\varphi;$$

la deuxième fonction a son support dans l'ouvert $\Omega \setminus K'$ qui est inclus dans le complémentaire du support de Λ, et elle annule donc Λ. Il s'ensuit que

$$\Lambda\varphi = \Lambda(\chi\varphi).$$

On applique la définition des distributions pour borner cette quantité en fonction des dérivées de $\chi\varphi$ jusqu'à un ordre fini ; on applique ensuite la formule de Leibniz pour borner les suprema des dérivées de $\chi\varphi$ en fonction des suprema des dérivées de φ. □

Remarque III-64. On ne peut pas, en général, choisir $K' = K$ dans le théorème.

Théorème III-65 (condition suffisante d'annulation). *Soient* Ω *un ouvert de* $\mathbb{R}^n$, Λ *une distribution dans* Ω, *de support compact* K *et d'ordre* k. *Soit* $\varphi \in \mathcal{D}(\Omega)$ *une fonction-test dont toutes les dérivées partielles jusqu'à l'ordre* k *s'annulent identiquement sur* K. *Alors* $\Lambda\varphi = 0$.

Démonstration. La démonstration consiste à reprendre l'argument précédent et à choisir χ de manière assez fine : on peut montrer que l'on peut choisir χ supportée par l'ensemble $\{y; \mathrm{dist}\,(y, K) \leq \varepsilon\}$, et en même temps

$$|\nabla^\alpha \chi(x)| \leq C\varepsilon^{-|\alpha|}.$$

On écrit soigneusement la formule de Leibnitz, et on fait tendre ε vers 0 ; on parvient ainsi à dominer $|\Lambda\varphi|$ par une expression qui tend vers 0 quand $\varepsilon \to 0$. On trouvera les détails dans [Hörmander p. 46]. □

7.4. Distributions à support ponctuel. On s'intéresse maintenant au cas où Λ a son support réduit à un point. Pour fixer les idées, on supposera que ce point est l'origine 0.

Théorème III-66 (distributions à support ponctuel). *Soit* Ω *un ouvert contenant 0, et* Λ *une distribution dans* Ω, *dont le support est réduit à* $\{0\}$. *Alors* Λ *est une combinaison linéaire finie de dérivées (au sens des distributions) de la masse de Dirac en 0.*

Remarque III-67. Par localisation, on en déduit facilement qu'une distribution dont le support est fait de points isolés s'écrit comme une combinaison linéaire finie de dérivées de masses de Dirac en ces points. Une généralisation plus intéressante consiste à considérer des distributions supportées par des sous-variétés de dimension finie ; voir [Hörmander, Théorème 2.3.5].

Démonstration. D'après le théorème précédent, Λ est d'ordre fini k. Si les formes linéaires $\nabla^\alpha\delta$ s'annulent sur φ pour $|\alpha| \leq k$, i.e. si $\nabla^\alpha\varphi(0) = 0$, alors les dérivées de φ d'ordre au plus k seront toutes identiquement nulles sur le support de Λ, et donc $\Lambda\varphi$ sera nul. On conclut en appliquant le théorème d'inclusion des noyaux. □

7.5. Distributions comme dérivées. L'un des buts de la théorie des distributions était de bâtir un cadre fonctionnel contenant les fonctions continues, dans lequel la dérivation soit toujours possible. Les théorèmes qui suivent montrent que le cadre ainsi construit est le plus économique possible : localement, toute distribution s'écrit comme la dérivée à un ordre suffisant d'une fonction continue.

.....

7.6. Opérateurs de convolution. Le théorème suivant montre que l'opération de convolution est très naturelle, en un certain sens.

Théorème III-68 (caractérisation des opérateurs de convolution). *Soit* L *une application linéaire continue de* $\mathcal{D}(\mathbb{R}^n)$ *dans* $\mathcal{D}(\mathbb{R}^n)$, *commutant avec les translations. Il existe alors une unique distribution* $\Lambda \in \mathcal{D}'(\mathbb{R}^n)$ *à support compact, telle que*

$$\forall\varphi \in \mathcal{D}(\mathbb{R}^n), \qquad L\varphi = \Lambda * \varphi.$$

DÉMONSTRATION. Si la distribution Λ existe, alors

$$\langle \Lambda, \varphi \rangle = \Lambda * \varphi(-\cdot) = [L\varphi(-\cdot)](0).$$

Réciproquement, on vérifie que la distribution Λ ainsi définie a toutes les propriétés souhaitées. $\square$

7.7. Théorème des noyaux de Schwartz.

THÉORÈME III-69 (théorème des noyaux de Schwartz). *Soient Ω_1 et Ω_2 des ouverts de $\mathbb{R}^{n_1}$ et $\mathbb{R}^{n_2}$ respectivement. Alors, les applications linéaires $\mathcal{K}$, séquentiellement continues de $\mathcal{D}(\Omega_1)$ dans $\mathcal{D}'(\Omega_2)$, s'identifient aux distributions K sur $\Omega_1 \times \Omega_2$, via*

$$\langle \mathcal{K}\varphi, \psi \rangle = \langle K, \varphi \otimes \psi \rangle.$$

DÉMONSTRATION. Voir [Hörmander, section 5.2] $\square$

Références

L'ouvrage classique de référence sur les distributions reste, encore aujourd'hui, l'ouvrage fondateur de **L. Schwartz**, *Théorie des distributions*, (Hermann, Paris, 1966). On admire encore aujourd'hui sa pédagogie et sa clarté ; les parties ayant trait à la topologie des distributions apparaissent en revanche un peu désuètes.

Une brève et efficace exposition de la théorie des distributions constitue l'essentiel de la partie II du traité de **W. Rudin**, *Functional Analysis* (McGraw Hill, New York, 1991, 2e édition), on y trouve en particulier une discussion de la topologie des distributions.

Un autre exposé, très accessible et entraînant, qui ne fait aucune référence à la topologie des distributions, se trouve dans l'ouvrage de **E. Lieb** et **M. Loss**, *Analysis* (AMS, Graduate Series in Mathematics 14, 1991, 2e édition). Dans la même catégorie, on recommande également le *Cours d'Analyse* de **J.-M. Bony** (Cours de l'Ecole Polytechnique, Edition 1992), particulièrement clair et pédagogique. Ce cours est rendu auto-contenu par la présence de quelques rappels d'analyse fonctionnelle des espaces de Fréchet en appendice.

Pour un étudiant, l'exposé le plus complet des distributions est sans doute celui que l'on trouve dans l'ouvrage célèbre de **L. Hörmander**, *The analysis of linear partial differential operators I : Distribution theory and Fourier analysis* (Springer, Berlin, 1990, 2e édition ; Springer Study Edition). Outre la théorie classique des distributions, on y trouvera aussi un très bel exposé sur les classes de fonctions quasi-analytiques, et sur les “hyperfonctions”, ou généralisations des distributions construites à partir des classes quasi-analytiques. La grande clarté de la présentation et la richesse des résultats compensent la concision parfois un peu exagérée des preuves.