

Jean-Michel Bony

Cours d'analyse

Théorie des distributions
et analyse de Fourier



INTRODUCTION

Ce cours est principalement consacré à deux grands outils de l'Analyse, l'analyse de Fourier et la théorie des distributions, ainsi qu'à diverses applications à des équations de la physique mathématique. Il s'agit de domaines ayant des racines très anciennes, mais qui font toujours l'objet de recherches actives, et qui ont connu dans une période récente des développements très importants dont nous essaierons de donner une idée.

L'analyse de Fourier. — Dite encore analyse harmonique, cette branche est dominée par les concepts de série et surtout de transformation de Fourier, ainsi que par l'opération de convolution qui leur est étroitement reliée.

On peut, sous des hypothèses très générales, représenter une fonction f définie dans $\mathbb{R}^n$ sous la forme suivante

$$f(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{ix \cdot \xi} \widehat{f}(\xi) d\xi,$$

où la fonction $\widehat{f}$ définie dans $\mathbb{R}^n$ et qui se déduit de f par une formule analogue s'appelle la transformée de Fourier de f .

Il faut voir cette formule de la manière suivante. Elle permet d'écrire une fonction quelconque f comme une "superposition" de fonctions oscillantes simples : les applications $x \mapsto e^{ix \cdot \xi}$, chacune d'elles ayant une amplitude $|\widehat{f}(\xi)|$, et un déphasage $\arg \widehat{f}(\xi)$.

Lorsqu'on veut analyser qualitativement ou quantitativement une fonction, l'idée la plus simple est l'analyse "en amplitude" fondée sur les valeurs ponctuelles : en quels points x la fonction est-elle nulle, "petite", "grande"... ? L'analyse de Fourier permet d'y superposer une analyse "en fréquence" : quelles sont les fréquences ξ qui contribuent à l'écriture de f ci-dessus, y contribuent-elles peu ou beaucoup... ? Ce point de vue revient à faire l'analyse "en amplitude" de $\widehat{f}$.

Peut-être d'apparence moins naturelle, l'analyse en fréquence s'est révélée d'une importance aussi grande que l'analyse en amplitude (leurs rôles sont symétriques en mécanique quantique par exemple). Elle prend même un caractère prédominant dans l'étude d'un certain nombre de questions, notamment :

— Les phénomènes oscillants, cela va (presque) de soi.

— Les phénomènes régis par des équations aux dérivées partielles linéaires à coefficients constants. Pour une équation d'évolution par exemple, il existe des solutions particulièrement simples, de la forme $e^{ix \cdot \xi} e^{\alpha t}$, où le nombre complexe α se détermine à partir de ξ par un simple calcul algébrique. L'analyse de Fourier permet d'écrire toute évolution, aussi compliquée soit-elle, comme une superposition de ces évolutions simples.

— Les problèmes de régularité. Comme nous le verrons, un “principe” correspondant à de nombreux théorèmes assure qu'une fonction est d'autant plus régulière que sa transformée de Fourier est petite à l'infini. Cela permet d'utiliser des méthodes éprouvées (majorations, développements asymptotiques...) sur $\hat{f}$ pour étudier la régularité de la fonction f .

La transformation de Fourier a néanmoins un inconvénient, son caractère global. Pour le traitement du signal, ou pour l'étude des équations aux dérivées partielles à coefficients variables, l'analyse en fréquence reste indispensable, mais il faut pouvoir la mener localement en les variables de temps ou d'espace. C'est l'objet de l'analyse microlocale, branche née vers 1970 dont nous ne donnerons qu'un très bref aperçu, dans la section 9.7.

La théorie des distributions. — C'est une extension de la notion de fonction, qui a joué un rôle très important dans le développement de l'Analyse. Bien que son introduction par L. Schwartz soit encore relativement récente, elle a permis de tels progrès en théorie des équations aux dérivées partielles et en analyse harmonique que l'on ne saurait plus parler de ces deux branches sans y avoir recours.

L'ensemble du cours, et notamment l'introduction du chapitre 4, montrera l'intérêt de cette généralisation de la notion de fonction. Cela dit, l'aspect déductif de l'exposition risque de donner une idée fausse du développement historique des mathématiques, qui est tout sauf déductif. L'introduction des distributions est aussi l'aboutissement d'un processus s'étalant sur plus d'un demi-siècle, en mathématiques et en physique⁽¹⁾.

⁽¹⁾Le lecteur intéressé trouvera au chapitre 6 de l'ouvrage de Laurent Schwartz “Un Mathématicien aux prises avec le siècle” (Odile Jacob, 1997) l'histoire de l'invention des distributions assortie d'une analyse du contexte scientifique.

Le calcul symbolique de Heaviside (1893), et surtout le formalisme introduit par P. Dirac (1926) pour les besoins de la mécanique quantique, la célèbre “fonction δ ” notamment, posaient un problème intéressant. La définition de ces concepts comme des fonctions au sens mathématique du terme était parfaitement contradictoire. Néanmoins, utilisés par Dirac lui-même ou d’autres bons physiciens, ils se révélaient efficaces et féconds. Des situations de ce type ne sont pas rares, il en existe actuellement, et elles signifient en général que des progrès mathématiques sont à l’ordre du jour.

En mathématiques, une multitude de concepts et de résultats, parfaitement rigoureux mais un peu épars, développés pendant la première partie de ce siècle n’ont trouvé leur unification que dans le cadre des distributions.

Ainsi, la notion de dérivée au sens des distributions permet de définir des solutions d’équations aux dérivées partielles qui ne sont pas suffisamment dérivables pour être des solutions au sens usuel du terme. En fait, pour presque chaque type d’équation aux dérivées partielles, on avait été amené à définir une ou des notions de solutions généralisées (ou solutions “faibles”). Ces définitions rentrent maintenant dans un cadre commun et beaucoup plus général

De même, comme nous le verrons, beaucoup d’opérateurs permettant de résoudre ces équations s’exprimeront pour nous en termes de convolution par une distribution, alors que l’introduction de ces opérateurs est souvent bien antérieure. Pour l’équation des ondes par exemple, Hadamard les avait exprimés par une formule ressemblant à celle d’une convolution, mais où des intégrales divergentes devaient être remplacées par leur “partie finie”.

Enfin, un pas très important avait été franchi par Sobolev en introduisant les espaces qui portent son nom, nous aurons l’occasion d’en voir l’intérêt, et en contribuant à clarifier le concept de solution faible.

La théorie des équations aux dérivées partielles. — C’est un thème qui ne sera qu’effleuré dans ce cours. Aucun développement systématique ne lui sera consacré, il est clair qu’un sujet d’une telle ampleur nécessite un enseignement spécifique.

Par contre, nous nous sommes efforcés de montrer l’efficacité et la puissance des outils théoriques introduits en les appliquant systématiquement à des équations de la physique mathématique, principalement aux équations de Laplace, de Schrödinger, et aux équations de la propagation des ondes et de la chaleur.

Les trois premiers chapitres doivent être considérés comme des préliminaires, présentant des outils d’utilisation constante dans la suite. Le premier est consacré à *l’intégrale de Lebesgue*. Un exposé complet de la théorie, avec

toutes les démonstrations, aurait été trop long pour le cadre horaire limité de ce cours. Par contre, nous avons pu en présenter les résultats les plus utiles, qui sont relativement peu nombreux, faciles à mémoriser, et considérablement plus efficaces que leurs homologues en théorie de l'intégrale de Riemann.

L'*analyse fonctionnelle* aura une place relativement réduite dans ce cours. Le chapitre 2 est destiné à consolider, et à développer sur quelques points, les connaissances du lecteur sur les espaces métriques, de Banach et de Hilbert. Cela nous suffira pour la suite, où nous n'introduirons pas, sauf pour les espaces normés, de topologie sur les espaces fonctionnels que nous aurons à utiliser.

En ce qui concerne l'espace des distributions, il s'agit d'un choix délibéré : la notion de suite convergente est tout à fait suffisante dans la pratique. En ce qui concerne les espaces de fonctions dérivables, il aurait par contre été raisonnable d'introduire leur structure d'espace métrisable. Si nous avons préféré, sans même parler de semi-norme, écrire explicitement chaque condition de continuité, c'est uniquement pour réaliser, au prix d'un petit nombre de répétitions, une économie de temps et de pensée.

Cela dit, pour ceux de nos lecteurs qui seraient suffisamment pourvus de ces deux denrées, nous avons consacré l'appendice C à l'étude des espaces de Fréchet. Nous y avons aussi donné les démonstrations d'un certain nombre de résultats admis dans le cours. Ils reposent sur le théorème de Banach-Steinhaus, qui repose lui-même sur la théorie des espaces de Baire développée dans l'appendice B.

Enfin nous aurons besoin de quelques compléments de *calcul différentiel*, principalement à plusieurs variables. Il s'agit d'un outil indispensable, les équations de la physique mathématique étant posées en dimension 3 ou 4. Le chapitre 3 rappelle les propriétés fondamentales des fonctions de classe C^k dans un ouvert de $\mathbb{R}^n$, cadre qui sera suffisant pour la suite, et y ajoute des résultats importants sur l'approximation par des fonctions différentiables.

L'appendice A développe un point de vue plus intrinsèque, et donne quelques compléments géométriques ainsi qu'une présentation complète de l'intégrale de surface. Il n'aborde toutefois que quelques aspects de la géométrie différentielle, et ne saurait se substituer à un enseignement consacré à ce sujet.

Le cours proprement dit se compose des chapitres 1 à 10, à l'exclusion des appendices et des textes en petits caractères. Ceux-ci sont destinés à apporter des compléments et selon les cas à satisfaire, ou à piquer, la curiosité du lecteur.

Les chapitres sont divisés en sections au sein desquelles tous les énoncés, des théorèmes aux exercices (qui ne sont pas moins importants) sont numérotés linéairement. Les formules sont numérotées au sein de chaque chapitre.

On trouvera enfin une petite bibliographie à la page 249.

*

* *

Cet ouvrage reprend, presque sans changement, le texte d'un cours trimestriel enseigné à l'École Polytechnique de 1986 à 1996. Il doit beaucoup aux cours d'Analyse de mes prédécesseurs : Laurent Schwartz, le regretté Charles Goulaouic et Yves Meyer. Le cadre trimestriel, certainement trop restreint, a eu sans doute un avantage : nous contraindre à aller droit à l'essentiel. Néanmoins, le sujet mériterait d'être traité à un rythme moins soutenu et d'être précédé d'un enseignement de calcul différentiel et intégral.

CHAPITRE 1

L'INTÉGRALE DE LEBESGUE

La théorie de l'intégrale de Riemann est à bien des égards insuffisante pour les besoins de l'Analyse, et notamment pour étudier les notions (analyse de Fourier, théorie des distributions) qui joueront un rôle central dans ce cours. Ce n'est que dans le cadre de la théorie de l'intégration introduite en 1900 par H. Lebesgue que l'on dispose des “bons” théorèmes de passage à la limite dans les intégrales, et que les espaces de fonctions sommables jouissent de bonnes propriétés.

Pour des raisons de temps, il ne serait pas possible de donner dans le cadre de ce cours un exposé complet, avec toutes les démonstrations, de la théorie de l'intégration. La solution adoptée consiste à admettre d'emblée un petit nombre de résultats, dont la démonstration est longue mais dont l'énoncé est simple, et d'en déduire les théorèmes fondamentaux de la théorie. Nous admettrons notamment dès le début l'existence et les principales propriétés de l'intégrale des fonctions positives — ce qui dans un exposé déductif nécessiterait toute la construction de la mesure de Lebesgue — ainsi que les théorèmes 1.4.1 et 1.5.2.

Pour les mêmes raisons, nous ferons délibérément “comme si” toutes les fonctions étaient mesurables ; nous nous expliquons sur ce point dans la section 1.6.

1.1. Intégrale des fonctions positives

Nous considérerons en fait des fonctions à valeurs dans $\overline{\mathbb{R}_+} = [0, \infty]$. Dans cet ensemble, auquel on étend de façon évidente l'addition, la multiplication par les éléments de $]0, \infty[$ et la relation d'ordre, toute suite croissante est convergente.

Nous admettrons qu'il existe une application

$$f \mapsto \int_{\mathbb{R}^n} f(x) dx$$

(que l'on notera parfois $\int f(x) dx$ voire $\int f$) qui, à toute fonction f , définie sur $\mathbb{R}^n$ et à valeurs dans $\overline{\mathbb{R}_+}$ fait correspondre un élément de $\overline{\mathbb{R}_+}$, et qui vérifie les quatre propriétés suivantes.

(a) *Linéarité* : pour f et g comme ci-dessus, et pour $\lambda, \mu \in]0, \infty[$, on a $\int(\lambda f + \mu g) = \lambda \int f + \mu \int g$.

(b) *Croissance* : si $f(x) \leq g(x)$ pour tout x , on a $\int f \leq \int g$. Cette propriété est en fait conséquence de la précédente.

(c) *Normalisation* : Pour un pavé $P = [a_1, b_1] \times \dots \times [a_n, b_n]$, si on appelle $\mathbf{1}_P$ la fonction égale à 1 sur P et à 0 ailleurs, on a

$$\int \mathbf{1}_P(x) dx = \prod_{i=1}^n (b_i - a_i).$$

(d) *Le théorème fondamental suivant* :

Théorème 1.1.1 (de Beppo Levi ou de la convergence monotone)

Si f_j est une suite croissante de fonctions définies sur $\mathbb{R}^n$ et à valeurs dans $\overline{\mathbb{R}_+}$, on a

$$\int \lim_{j \rightarrow \infty} f_j(x) dx = \lim_{j \rightarrow \infty} \int f_j(x) dx. \quad (1.1)$$

En particulier, pour des fonctions u_j positives, prenant éventuellement la valeur $+\infty$, on a toujours

$$\int \sum_{j=1}^{\infty} u_j(x) dx = \sum_{j=1}^{\infty} \int u_j(x) dx. \quad (1.2)$$

Le second point résulte facilement du premier en considérant la suite croissante constituée des sommes partielles $S_p(x) = \sum_{j=1}^p u_j(x)$.

1.1.2. Mesure des ensembles. — Si A est un sous ensemble de $\mathbb{R}^n$, on note $\mathbf{1}_A$ la fonction égale à 1 sur A et à 0 ailleurs, et on pose

$$\mu(A) = \int \mathbf{1}_A(x) dx.$$

L'application μ s'appelle la mesure de Lebesgue. C'est une application de l'ensemble des parties de $\mathbb{R}^n$ dans $\overline{\mathbb{R}_+}$ qui vérifie les deux propriétés suivantes.

Croissance. — si $A \subset B$, on a $\mu(A) \leq \mu(B)$. Cela résulte de la croissance de l'intégrale (ou si l'on veut de $\mu(B) = \mu(A) + \mu(B \setminus A)$).

Additivité dénombrable. — si des sous-ensembles A_j , $j \in \mathbb{N}$ sont deux à deux disjoints, on a $\mu(\bigcup A_j) = \sum \mu(A_j)$. En effet, en posant $A = \bigcup_j A_j$, on a alors $\mathbf{1}_A = \sum_j \mathbf{1}_{A_j}$, et le résultat découle de (1.2).

Si on ne suppose pas les A_j disjoints deux à deux, on a seulement la sous-additivité : $\mu(\bigcup A_j) \leq \sum \mu(A_j)$. On a également la relation $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$.

Les ensembles de mesure nulle, on dit aussi *ensembles négligeables*, jouent un rôle important dans la théorie. D'après la sous-additivité, *la réunion d'une infinité dénombrable d'ensembles de mesure nulle est de mesure nulle*. Un point (et plus généralement un pavé plat) étant de mesure nulle, il en résulte qu'un ensemble dénombrable est de mesure nulle.

Exercice 1.1.3. — Démontrer que le graphe d'une application continue de $\mathbb{R}$ dans $\mathbb{R}$ est un sous-ensemble de mesure nulle dans $\mathbb{R}^2$ (quels que soient ε et N , on montrera que la portion du graphe comprise entre les abscisses $-N$ et N est contenue dans une réunion de rectangles dont la somme des mesures est $\leq \varepsilon$).

Exercice 1.1.4. — Soit A un ensemble de mesure nulle dans $\mathbb{R}^n$. Démontrer que son complémentaire est partout dense.

On dit qu'une propriété $P(x)$ dépendant d'un point x est vérifiée *presque partout* (en abrégé p.p.) si l'ensemble $\{x \mid \text{non } P(x)\}$ est de mesure nulle. Si chacune des propriétés $P_j(x)$, $j \in \mathbb{N}$, est vraie presque partout, il en est donc de même de $(\forall j, P_j(x))$.

Proposition 1.1.5. — Soit f définie dans $\mathbb{R}^n$ et à valeurs dans $\overline{\mathbb{R}_+}$. On a

$$\int f(x) dx = 0 \iff f(x) = 0 \text{ p.p.}$$

Posons $A = \{x \mid f(x) \neq 0\}$. On a

$$f(x) \leq \lim_{j \rightarrow \infty} j \mathbf{1}_A(x).$$

Si $\mu(A) = 0$, on obtient donc en utilisant (1.1)

$$\int f(x) dx \leq \lim_{j \rightarrow \infty} j \int \mathbf{1}_A(x) dx = 0.$$

Réciproquement, si l'intégrale de f est nulle et en remarquant que $\mathbf{1}_A(x) \leq \lim_{j \rightarrow \infty} j f(x)$, on obtient

$$\mu(A) = \int \mathbf{1}_A(x) dx \leq \lim_{j \rightarrow \infty} j \int f(x) dx = 0.$$

Théorème 1.1.6 (Fatou). — Soit f_j une suite (non nécessairement monotone) de fonctions positives telle que pour chaque $x \in \mathbb{R}^n$, la suite $f_j(x)$ soit convergente, et que la suite $\int f_j(x) dx$ soit également convergente. On a

$$\int \lim_{j \rightarrow \infty} f_j(x) dx \leq \lim_{j \rightarrow \infty} \int f_j(x) dx.$$

Remarquons d'abord que, si une suite u_j d'éléments de $\overline{\mathbb{R}_+}$ converge vers $u \in \overline{\mathbb{R}_+}$, la suite $v_j = \inf_{k \geq j} u_k$ converge également vers u . En effet, dans le cas $u < \infty$, pour tout $\varepsilon > 0$ donné, tous les u_j appartiennent à $[u - \varepsilon, u + \varepsilon]$ à partir d'un certain rang, et tous les v_j appartiennent au même intervalle à partir du même rang. Le raisonnement est analogue si $u = +\infty$, avec des intervalles du type $[A, +\infty]$.

Soit g_j la fonction définie par $g_j(x) = \inf_{k \geq j} f_k(x)$. En posant $f(x) = \lim_{j \rightarrow \infty} f_j(x)$, on a d'après l'argument ci-dessus $f(x) = \lim_{j \rightarrow \infty} g_j(x)$ et de même $\lim_{j \rightarrow \infty} \int f_j(x) dx = \lim_{j \rightarrow \infty} (\inf_{k \geq j} \int f_k(x) dx)$.

D'autre part, l'intégrale de g_j est inférieure à l'intégrale de f_k pour chaque $k \geq j$, et on a donc $\int g_j(x) dx \leq \inf_{k \geq j} \int f_k(x) dx$. En appliquant le théorème de Beppo Levi à la suite croissante des g_j qui converge vers f , on obtient

$$\int f(x) dx = \lim_{j \rightarrow \infty} \int g_j(x) dx \leq \lim_{j \rightarrow \infty} \left(\inf_{k \geq j} \int f_k(x) dx \right) = \lim_{j \rightarrow \infty} \int f_j(x) dx,$$

ce qui est le résultat cherché.

Remarque 1.1.7. — Dans le cas particulier où la limite des intégrales est nulle, on peut en déduire que l'intégrale de la limite $\int \lim_{j \rightarrow \infty} f_j(x) dx$ est nulle.

En général, ce théorème ne fournit qu'une inégalité, et il est important de bien comprendre les deux raisons, illustrées par les deux exemples suivants, qui peuvent empêcher l'intégrale de passer à la limite. Pour en avoir une représentation plus imagée, on pourra penser que les f_j sont des densités de masses positives.

— La fonction f_j est égale à 1 sur l'intervalle $[j, j+1]$ et à 0 dans le complémentaire. On a $f_j(x) \rightarrow 0$ en tout point x alors que $\lim \int f_j(x) dx = 1$. On voit que "la masse disparaît à l'infini".

— La fonction f_j est égale à j sur l'intervalle $]0, 1/j[$ et à 0 dans le complémentaire. On a encore $f_j(x) \rightarrow 0$ en tout point x et $\lim \int f_j(x) dx = 1$. On

peut dire que “la masse se concentre sur un ensemble de mesure nulle”, en l'occurrence l'ensemble réduit à l'origine.

Le lecteur pourra construire beaucoup d'exemples analogues. Nous verrons plus loin que le théorème de Lebesgue, dont l'hypothèse signifie grosso modo que l'on se prémunit contre ces deux types d'accident, permettra de passer à la limite.

Exercice 1.1.8. — Rappelons d'abord la définition de la *limite inférieure d'une suite de nombres réels* bornée inférieurement

$$\liminf_{j \rightarrow \infty} x_j = \lim_{j \rightarrow \infty} \left(\inf_{k \geq j} x_k \right) \in]-\infty, +\infty],$$

les bornes inférieures du membre de droite étant bien définies, et formant une suite qui croît avec j .

Démontrer la forme plus générale suivante du théorème de Fatou : si f_j est une suite quelconque de fonctions positives, on a

$$\int \liminf_{j \rightarrow \infty} f_j(x) \, dx \leq \liminf_{j \rightarrow \infty} \int f_j(x) \, dx.$$

1.2. Fonctions sommables

Définition 1.2.1. — Soit f une fonction définie dans $\mathbb{R}^n$ à valeurs dans $\mathbb{C}$. On dit que f est *sommable* (ou *intégrable au sens de Lebesgue*) si on a

$$\int_{\mathbb{R}^n} |f(x)| \, dx < \infty.$$

L'espace des fonctions sommables est noté $\mathcal{L}^1(\mathbb{R}^n)$.

Dans la pratique, pour montrer qu'une fonction est sommable, il suffit de montrer qu'elle est majorée en module par une fonction positive d'intégrale finie.

Dans le cas où une fonction sommable f est à valeurs réelles, on peut l'écrire sous la forme $f = g - h$ où g et h sont positives et d'intégrales finies. On peut par exemple prendre $f = f_+ - f_-$, en posant $f_{\pm}(x) = \max\{\pm f(x), 0\}$, et les fonctions $f_{\pm}$ sont majorées par la fonction $|f|$ qui est d'intégrale finie.

On définit alors l'intégrale de f par

$$\int_{\mathbb{R}^n} f(x) \, dx = \int_{\mathbb{R}^n} g(x) \, dx - \int_{\mathbb{R}^n} h(x) \, dx,$$

le résultat ne dépendant pas de la décomposition choisie. En posant en effet $r = g - f_+$, on a également $h - f_- = r$, la fonction r étant positive et d'intégrale

finie (elle est majorée par g). On a alors $\int g = \int f_+ + \int r$, $\int h = \int f_- + \int r$ et donc $\int g - \int h = \int f_+ - \int f_-$.

Si $f = f_1 + if_2$ est sommable et à valeurs complexes, les fonctions f_j sont alors sommables, et on pose $\int f = \int f_1 + i \int f_2$.

À partir des propriétés correspondantes pour les fonctions positives, on montre facilement les propriétés élémentaires suivantes.

Linéarité. — L'espace $\mathcal{L}^1(\mathbb{R}^n)$ est un espace vectoriel, et l'application $f \mapsto \int f$ est linéaire de $\mathcal{L}^1(\mathbb{R}^n)$ dans $\mathbb{C}$.

Croissance. — Pour f et g sommables à valeurs réelles vérifiant $f(x) \leq g(x)$ pour tout x , on a $\int f \leq \int g$.

Pour f et g sommables à valeurs complexes, on a

$$\left| \int f(x) dx \right| \leq \int |f(x)| dx, \quad (1.3)$$

et

$$\int |f(x) + g(x)| dx \leq \int |f(x)| dx + \int |g(x)| dx. \quad (1.4)$$

Démontrons par exemple l'additivité pour des fonctions réelles. Si $f = g - h$ et $f' = g' - h'$ sont des décompositions de f et f' en différences de fonctions positives d'intégrale finie, on a la décomposition $f + f' = (g + g') - (h + h')$ et donc $\int(f + f') = \int(g + g') - \int(h + h')$, ce qui entraîne le résultat voulu.

Montrons également (1.3), qui entraîne facilement (1.4). Soit θ un nombre complexe de module 1 tel que $\theta \int f$ soit réel positif. On a alors

$$\left| \int f(x) dx \right| = \int \operatorname{Re}(\theta f(x)) dx \leq \int (\operatorname{Re}(\theta f(x)))_+ dx \leq \int |f(x)| dx.$$

1.2.2. Intégration sur un sous-ensemble. — Considérons un sous-ensemble A de $\mathbb{R}^n$, et f une fonction définie sur A . On notera f_A la fonction égale à $f(x)$ pour $x \in A$ et à 0 sinon.

On dit que la fonction f est sommable sur A si la fonction f_A est sommable dans $\mathbb{R}^n$. On note $\mathcal{L}^1(A)$ l'espace des fonctions sommables sur A , et on pose alors

$$\int_A f(x) dx = \int_{\mathbb{R}^n} f_A(x) dx.$$

Le passage de f à f_A permet d'associer à tout résultat sur l'espace $\mathcal{L}^1(\mathbb{R}^n)$ un résultat correspondant sur $\mathcal{L}^1(A)$, que nous laisserons au lecteur le soin d'énoncer.

Remarque 1.2.3. — Si deux fonctions f et g sont égales presque partout, et si l'une d'elles est sommable, l'autre l'est aussi, et elles ont la même intégrale. On a en effet $|\int f - \int g| \leq \int |f - g|$, quantité qui est nulle d'après la proposition 1.1.5. En particulier, l'intégrale d'une fonction sur un ensemble de mesure nulle est toujours nulle. De même, en dimension 1, il est équivalent d'intégrer une fonction f sur $[a, b]$ ou sur $]a, b[$, le résultat étant habituellement noté $\int_a^b f(x) dx$.

Intégrale d'une fonction définie presque partout. — Soit f une fonction définie seulement presque partout dans $\mathbb{R}^n$. Si l'un de ses prolongements $\tilde{f}$ à l'espace entier est sommable, il résulte de la remarque précédente que tous les autres prolongements le sont aussi et qu'ils ont tous la même intégrale. Par abus de langage et de notation, on dit encore dans ce cas que la fonction f est sommable dans $\mathbb{R}^n$ et on note $\int_{\mathbb{R}^n} f(x) dx$ l'intégrale de $\tilde{f}$.

Le théorème suivant est d'une importance capitale. La supériorité de l'intégrale de Lebesgue sur l'intégrale de Riemann est due en partie au fait que l'on peut intégrer plus de fonctions, mais elle est surtout due au fait que l'on dispose de théorèmes beaucoup plus efficaces. On comparera l'énoncé suivant, et le théorème de dérivation sous le signe somme qui en découle, aux résultats analogues fondés sur la convergence uniforme.

Théorème 1.2.4 (de Lebesgue ou de la convergence dominée)

Soit f_j une suite de fonctions qui converge presque partout vers une fonction f . On suppose qu'il existe une fonction positive sommable fixe h telle que l'on ait $|f_j(x)| \leq h(x)$ p.p. pour tout j . On a alors

$$\int |f(x) - f_j(x)| dx \rightarrow 0 \quad (1.5)$$

et

$$\int f_j(x) dx \rightarrow \int f(x) dx. \quad (1.6)$$

Considérons les ensembles $A_j = \{x \mid |f_j(x)| > h(x)\}$ et l'ensemble B constitué des points x où la suite $f_j(x)$ ne converge pas vers $f(x)$. Ces ensembles sont de mesure nulle par hypothèse, et il en est donc de même de $N = B \cup \left(\bigcup_{j=0}^{\infty} A_j\right)$.

En posant $\tilde{f}_j(x) = f_j(x)$ pour $x \notin N$ et $\tilde{f}_j(x) = 0$ pour $x \in N$, et en définissant de même $\tilde{f}$, il suffit de démontrer le théorème en remplaçant les fonctions f_j par les $\tilde{f}_j$ (les intégrales sont inchangées), et nous nous sommes ainsi ramenés au cas où les majorations et la convergence ont lieu partout.

En posant $g_j(x) = |\tilde{f}_j(x) - \tilde{f}(x)|$, puis $h_j(x) = \sup_{k \geq j} g_k(x)$, on voit facilement que h_j est une suite décroissante de fonctions positives qui tend vers 0 en tout point, et que l'on a $h_j(x) \leq 2h(x)$ en tout point. Il nous reste à prouver que l'intégrale de h_j tend vers 0, ce qui entraîne immédiatement (1.5) et (1.6).

En appliquant le théorème de la convergence monotone à la suite croissante de fonctions positives $2h - h_j$, qui tend vers $2h$ en tout point, on obtient

$$\lim_{j \rightarrow \infty} \int (2h(x) - h_j(x)) dx = 2 \int h(x) dx.$$

Comme la fonction h est sommable, l'intégrale figurant au premier membre est la différence des intégrales de $2h$ et de h_j . Cela prouve que $\int h_j(x) dx \rightarrow 0$ et achève la démonstration.

Exercice 1.2.5. — Soit A_j une suite croissante de sous-ensembles de $\mathbb{R}^n$ et posons $A = \bigcup_j A_j$. Démontrer que, pour toute fonction $f \in \mathcal{L}^1(A)$, on a $\int_A f(x) dx = \lim_{j \rightarrow \infty} \int_{A_j} f(x) dx$. Démontrer qu'une fonction f définie sur A appartient à $\mathcal{L}^1(A)$ si et seulement si $\sup_j \int_{A_j} |f(x)| dx < \infty$.

Théorème 1.2.6 (de dérivation sous le signe somme)

Soient I un intervalle de $\mathbb{R}$ et A un sous-ensemble de $\mathbb{R}^n$. On se donne une fonction f définie sur $A \times I$ vérifiant les trois hypothèses suivantes.

- (a) Pour tout $\lambda \in I$, la fonction $x \mapsto f(x, \lambda)$ est sommable sur A
- (b) La dérivée partielle $\partial f / \partial \lambda(x, \lambda)$ existe en tout point de $A \times I$
- (c) Il existe une fonction h positive et sommable sur A telle que l'on ait $|\partial f / \partial \lambda(x, \lambda)| \leq h(x)$ quels que soient x et λ .

Alors la fonction F définie par

$$F(\lambda) = \int_A f(x, \lambda) dx \tag{1.7}$$

est dérivable dans I , et on a

$$F'(\lambda) = \int_A \frac{\partial f}{\partial \lambda}(x, \lambda) dx. \tag{1.8}$$

1.2.7. Mode d'emploi. — Ce résultat ne prend toute sa force que si on l'accompagne des deux remarques suivantes.

(a) — Avant d'appliquer le théorème, on peut retirer du domaine d'intégration un ensemble de mesure nulle, ce qui ne change pas les intégrales dans (1.7) et (1.8), et donc ne vérifier les hypothèses que dans l'ensemble A' ainsi obtenu.

Par contre, il ne suffirait pas que, pour chaque λ , les hypothèses soient satisfaites sauf sur un sous-ensemble de A , fût-il réduit à un point, qui dépend de λ (voir l'exercice 1.2.10).

(b) — La dérivabilité est une propriété locale. Pour prouver que F est dérivable dans I , il suffit de montrer que F est dérivable dans tout intervalle compact $[c, d] \subset I$. Il suffira donc de trouver des fonctions positives sommables h_{cd} qui majorent $\partial f / \partial \lambda$ en module lorsque λ parcourt $[c, d]$.

Remarque 1.2.8. — Il existe un théorème de *continuité sous le signe somme* pour $F(\lambda) = \int_A f(x, \lambda) dx$, lorsque λ parcourt un ouvert de $\mathbb{R}^p$ ou même un espace métrique quelconque, et le lecteur pourra l'énoncer s'il le désire. Il faut prouver que pour toute suite λ_j tendant vers un point λ_0 , on a $F(\lambda_j) \rightarrow F(\lambda_0)$, et le théorème de Lebesgue fournit immédiatement les conditions voulues.

Lorsque λ parcourt un ouvert $\omega \subset \mathbb{R}^p$, pour prouver que F est de classe C^1 on procède généralement en deux temps. On prouve d'abord l'existence des $\partial F / \partial \lambda_i$ en gelant les autres variables λ_j et en appliquant le théorème 1.2.6 à la fonction de x et λ_i ainsi obtenue. Les $\partial F / \partial \lambda_i$ étant exprimées par une intégrale, on prouve ensuite leur continuité dans ω à l'aide de l'argument ci-dessus.

1.2.9. Démonstration du théorème 1.2.6. — On a, λ étant fixé,

$$\frac{1}{l} (F(\lambda + l) - F(\lambda)) = \int_A g_l(x) dx,$$

où on a posé

$$g_l(x) = \frac{1}{l} (f(x, \lambda + l) - f(x, \lambda)).$$

La fonction g_l converge en tout point vers $\partial f / \partial \lambda$. D'autre part, on a $|g_l(x)| \leq \sup_{0 \leq \theta \leq 1} |\partial f / \partial \lambda(x, \lambda + \theta l)| \leq h(x)$ d'après le théorème des accroissements finis, et le résultat est conséquence immédiate du théorème de Lebesgue.

Exercice 1.2.10. — Soit φ une fonction continue sur $[0, 1]$. On considère, dans $[0, 1] \times [0, 1]$ la fonction f définie par $f(x, \lambda) = \varphi(x)$ si $x \leq \lambda$ et $f(x, \lambda) = 0$ sinon. On pose $F(\lambda) = \int_0^1 f(x, \lambda) dx$. Pour chaque λ la dérivée partielle $\partial f / \partial \lambda$ existe sauf en un point, et elle est majorée par une fonction sommable fixe : la fonction 0. Déterminer la dérivée de F .

Exercice 1.2.11. — (a) Soit f une fonction sommable sur $]0, \infty[$. Pour $\lambda \in [0, \infty[$, on définit sa transformée de Laplace F par $F(\lambda) = \int_0^\infty e^{-\lambda t} f(t) dt$. Démontrer que F est continue sur $[0, \infty[$.

(b) On suppose maintenant que f , toujours définie sur $]0, \infty[$, est telle que $(1+t)^{-N} f(t)$ est sommable pour un N convenable. Démontrer que $F(\lambda)$ est bien définie par la formule ci-dessus pour $\lambda \in]0, \infty[$, et que la fonction F

est de classe C^∞ sur cet intervalle. Déterminer la transformée de Laplace de $t \mapsto t^k f(t)$.

(c) Sous les hypothèses de (b), pour $z = x + iy$ complexe vérifiant $x > 0$, on pose $F(z) = \int_0^\infty e^{-zt} f(t) dt$. Démontrer que F est une fonction de classe C^1 (et même C^∞) des variables x, y dans le demi-plan $x > 0$. Démontrer que l'on a $\partial F / \partial x + i \partial F / \partial y = 0$. En déduire que, pour tout z dans ce demi-plan, il existe $A_z \in \mathbb{C}$ tel que l'on ait, pour h complexe, $F(z+h) = F(z) + A_z h + hr(h)$ avec $\lim_{h \rightarrow 0} r(h) = 0$ (de telles fonctions sont dites holomorphes).

Exercice 1.2.12. — (a) Pour $x > 0$, on pose $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt$ (fonction Γ d'Euler). Démontrer que la fonction Γ est de classe C^∞ sur $]0, \infty[$.

(b) Démontrer que $\Gamma(x) \rightarrow +\infty$ pour $x \rightarrow 0$. (Pourquoi utiliser des ε ? La convergence monotone sur $]0, 1]$ est tellement plus simple.)

(c) Pour $\operatorname{Re} z > 0$, on pose $\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt$. Démontrer que Γ est holomorphe (voir l'exercice précédent) dans le demi-plan $\operatorname{Re} z > 0$.

Exercice 1.2.13. — Calculer la dérivée à droite en 0 de la fonction

$$\lambda \mapsto \int_0^1 (g(x) + \lambda^2)^{1/2} dx$$

où g est une fonction bornée, positive ou nulle sur $]0, 1[$.

1.3. Cas de la dimension 1

Théorème 1.3.1. — Soit f une fonction continue sur un intervalle $[a, b]$ et soit F une primitive de f . On a alors

$$\int_{[a,b]} f(x) dx = F(b) - F(a).$$

Considérons la fonction $G(x) = \int_{[a,x]} f(t) dt$. On a, pour $l > 0$ (le raisonnement serait le même pour $l < 0$)

$$\begin{aligned} \frac{1}{l} (G(x+l) - G(x)) &= \frac{1}{l} \int_{[x,x+l]} f(t) dt \\ &= \frac{1}{l} \int_{[x,x+l]} f(x) dt + \frac{1}{l} \int_{[x,x+l]} (f(t) - f(x)) dt. \end{aligned} \quad (1.9)$$

D'après la propriété (c) de la section 1.1, l'intégrale d'une constante M sur un intervalle $[a, b]$ est égale à $M(b-a)$. Le membre de droite de (1.9) est la somme de $f(x)$ et d'une quantité majorée en module par $\max_{[x,x+l]} |f(t) - f(x)|$ qui tend vers 0 avec l . Il en résulte que G est dérivable, et que $G'(x) = f(x)$. La fonction $F - G$ est donc constante, et on a $G(b) - G(a) = F(b) - F(a)$ ce qui est le résultat voulu.

On voit donc que l'intégrale de Lebesgue est un prolongement de l'intégrale de Riemann des fonctions continues. Cela dit, l'étude préalable de l'intégrale de Riemann n'est nullement nécessaire. Pour le calcul des intégrales de Lebesgue, on peut à la fois utiliser le calcul des primitives dans les cas simples et disposer de théorèmes puissants sur les passages à la limite.

Exercice 1.3.2. — Soit f une fonction intégrable au sens de Riemann sur $[a, b]$. Démontrer que f est sommable et que les deux définitions de l'intégrale de f coïncident. On pourra, pour chaque entier m , poser $h = (b - a)/2^m$ et introduire les fonctions φ_m [resp. ψ_m] constantes sur chaque intervalle $[kh, (k+1)h[$ et y prenant comme valeur le sup [resp. inf] de f sur le même intervalle. On montrera ensuite que les suites (φ_m) et (ψ_m) sont monotones et, en utilisant la proposition 1.1.5, que l'on a $\inf \varphi_m = \sup \psi_m$ p.p.

Bien entendu, on ne se privera pas de la notation classique $\int_a^b f(x) dx$ pour désigner, selon les cas, l'intégrale sur $]a, b[$ ou l'opposé de l'intégrale sur $]b, a[$.

1.3.3. Qu'est-il advenu des intégrales absolument convergentes ... — À la différence de celle de Riemann, l'intégrale de Lebesgue est directement définie de manière globale sur $\mathbb{R}$. Cela rend beaucoup plus nette la distinction entre deux concepts de nature très différente mais qui sont souvent présentés ensemble (sous le nom plus que discutable d'« intégrales impropres ») en théorie de l'intégrale de Riemann.

Quoi qu'il en soit, il est bien sûr intéressant d'avoir des relations entre les intégrales sur $\mathbb{R}$ entier, et celles sur les intervalles finis. Pour $a < b$, on notera f_{ab} la fonction égale à f pour $x \in [a, b]$ et à 0 ailleurs.

Si f est une fonction positive, et si a_n et b_n sont des suites respectivement décroissant vers $-\infty$ et croissant vers $+\infty$, la suite des fonctions $f_{a_n b_n}$ converge en croissant vers la fonction f , et on a convergence des intégrales d'après le théorème de Beppo Levi. On a donc

$$\forall f \geq 0, \quad \int_{\mathbb{R}} f(x) dx = \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow +\infty}} \int_a^b f(x) dx \leq +\infty$$

formule dans laquelle on peut remplacer $\lim$ par $\sup$.

En particulier, en appliquant ce qui précède à la fonction $|f|$, on obtient le résultat suivant : *pour qu'une fonction f à valeurs complexes soit sommable, il faut et il suffit que $\sup \int_a^b |f(x)| dx < +\infty$.*

Si la fonction f est sommable, la suite des fonctions $f_{a_n b_n}$ définies comme ci-dessus converge vers f en chaque point, et est majorée en module par la

fonction sommable fixe $|f(\cdot)|$. On a donc, en vertu du théorème de Lebesgue,

$$\forall f \in \mathcal{L}^1(\mathbb{R}), \quad \int_{\mathbb{R}} f(x) dx = \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow +\infty}} \int_a^b f(x) dx.$$

On voit que le concept d'intégrale d'une fonction sommable généralise celui d'"intégrale de Riemann absolument convergente".

Bien entendu ce qui précède s'applique également aux rapports entre $\int_a^b f(x) dx$ et les $\int_{a_n}^{b_n} f(x) dx$ lorsque a_n et b_n tendent respectivement vers a et b .

1.3.4. ... et des intégrales semi-convergentes? — Il arrive qu'une fonction f , sommable sur tout compact, ne soit pas sommable sur $\mathbb{R}$ mais que les intégrales $\int_a^b f(x) dx$ tendent vers une limite pour $a \rightarrow -\infty$ et $b \rightarrow +\infty$. Un exemple classique est la fonction $\sin x/x$. Il est traditionnel de noter encore $\int_{-\infty}^{+\infty} f(x) dx$ la limite et de parler d'intégrale semi-convergente, mais *de telles expressions ne sont pas vraiment des intégrales — aucun théorème de la théorie de l'intégration ne s'applique à elles — ce sont des limites d'intégrales et il n'y a guère plus à en dire*. Pour les étudier, on applique la théorie de l'intégration aux vraies intégrales $\int_a^b f(x) dx$, et on passe à la limite avec les moyens du bord.

1.4. Intégrales multiples

Théorème 1.4.1 (Fubini). — Soit $f(x, y)$ une fonction définie dans le produit $\mathbb{R}^p \times \mathbb{R}^q$.

(a) Si f est à valeurs dans $\overline{\mathbb{R}}_+$, on a l'égalité suivante, où les trois membres définissent un élément de $\overline{\mathbb{R}}_+$

$$\iint_{\mathbb{R}^{p+q}} f(x, y) dx dy = \int_{\mathbb{R}^p} \left\{ \int_{\mathbb{R}^q} f(x, y) dy \right\} dx = \int_{\mathbb{R}^q} \left\{ \int_{\mathbb{R}^p} f(x, y) dx \right\} dy \quad (1.10)$$

(b) Si f est sommable dans $\mathbb{R}^{p+q}$, les trois membres de (1.10) ont un sens dans $\mathbb{C}$ et sont égaux. Pour être précis, dire que le troisième membre a un sens signifie les deux choses suivantes :

- pour presque chaque y , la fonction $x \mapsto f(x, y)$ est sommable dans $\mathbb{R}^p$,
- la fonction $\varphi(y) = \int f(x, y) dx$ qui est ainsi définie presque partout est sommable dans $\mathbb{R}^q$.

1.4.2. Mode d'emploi. — On désire souvent intervertir les signes d'intégration, et prouver l'égalité des second et troisième membres de (1.10) pour une fonction f à valeurs complexes. On procède en deux temps.

(a) — On prouve d'abord que f est sommable dans $\mathbb{R}^{p+q}$. Pour cela, on utilise la forme (a) du théorème de Fubini pour calculer ou majorer l'intégrale $\iint |f(x, y)| \, dx \, dy$ par intégration successive dans l'ordre que l'on désire.

(b) — Une fois établi le fait que f est sommable "du couple", on utilise à nouveau le théorème de Fubini sous la forme (b).

On prendra garde au fait suivant : il peut très bien arriver (pour une fonction non sommable dans $\mathbb{R}^{p+q}$) que les second et troisième membres de (1.10) aient un sens dans $\mathbb{C}$, et qu'ils soient différents.

Remarque 1.4.3. — Pour calculer l'intégrale double d'une fonction f sur un sous-ensemble A de $\mathbb{R}^{p+q}$, on se ramène au cas de l'intégration de la fonction f_A . Le troisième membre, par exemple, de (1.10) devient

$$\int_{\pi(A)} \left\{ \int_{A_y} f(x, y) \, dx \right\} dy$$

où A_y est la "tranche" $\{x \in \mathbb{R}^p \mid (x, y) \in A\}$, et où $\pi(A)$ est la projection de A sur $\mathbb{R}^q$.

Exercice 1.4.4. — Soit A le graphe d'une application de $\mathbb{R}^{n-1}$ dans $\mathbb{R}$. Démontrer que A est de mesure nulle dans $\mathbb{R}^n$.

Théorème 1.4.5 (de changement de variable). — Soient Ω_1 et Ω_2 deux ouverts de $\mathbb{R}^n$, et Φ un difféomorphisme de Ω_1 sur Ω_2 . On notera $J_\Phi(x)$ le déterminant de la matrice jacobienne de Φ au point x

$$J_\Phi = \begin{vmatrix} \partial\Phi_1/\partial x_1 & \cdots & \partial\Phi_1/\partial x_n \\ \cdots & \cdots & \cdots \\ \partial\Phi_n/\partial x_1 & \cdots & \partial\Phi_n/\partial x_n \end{vmatrix}.$$

Soit f une fonction définie sur Ω_2 .

(a) Si la fonction f est à valeurs dans $\overline{\mathbb{R}_+}$, on a l'égalité suivante, où les deux membres ont un sens dans $\overline{\mathbb{R}_+}$

$$\int_{\Omega_2} f(y) \, dy = \int_{\Omega_1} f(\Phi(x)) |J_\Phi(x)| \, dx. \quad (1.11)$$

(b) Si f est à valeurs complexes, elle est sommable dans Ω_2 si et seulement si la fonction $f(\Phi(x)) |J_\Phi(x)|$ est sommable dans Ω_1 et les deux membres de (1.11) sont alors égaux.

Remarque 1.4.6. — Il est souvent utile, pour se trouver dans les conditions d'application du théorème, de retirer des domaines d'intégration des ensembles de mesure nulle, ce qui ne change aucune des intégrales. Par exemple, les coordonnées polaires ne fournissent pas de difféomorphisme du plan sur un ouvert. Par contre, si on retire du plan l'origine et le demi-axe des x négatif, on obtient un difféomorphisme de $]0, \infty[\times]-\pi, \pi[$ sur l'ouvert ainsi obtenu, et donc la formule classique (le jacobien étant égal à r)

$$\iint_{\mathbb{R}^2} f(x, y) dx dy = \iint_{]0, \infty[\times]-\pi, \pi[} f(r \cos \theta, r \sin \theta) r dr d\theta,$$

valable pour f positive ou sommable.

On utilisera parfois la formule analogue dans $\mathbb{R}^n$. Ce n'est pas exactement une application directe du théorème, mais on s'y ramène facilement en prenant des équations locales de portions de la sphère unité $\mathbb{S}^{n-1}$ de $\mathbb{R}^n$. On notera θ le point courant de celle-ci, et $d\sigma_\theta$ la mesure de surface (voir la section A.3).

$$\int_{\mathbb{R}^n} f(x) dx = \int_0^\infty \int_{\mathbb{S}^{n-1}} f(r\theta) r^{n-1} dr d\sigma_\theta.$$

1.5. Espaces $\mathcal{L}^1$, $\mathcal{L}^2$, $\mathcal{L}^\infty$

1.5.1. L'espace $\mathcal{L}^1(\mathbb{R}^n)$. — Nous l'avons défini comme l'espace des fonctions sommables. Si on pose $\|f\| = \int |f(x)| dx$, on voit facilement que $\|\lambda f\| = |\lambda| \|f\|$, et (1.4) exprime que $\|f + g\| \leq \|f\| + \|g\|$. Ce n'est néanmoins pas une norme sur $\mathcal{L}^1(\mathbb{R}^n)$, car la relation $\|f\| = 0$ équivaut d'après la proposition 1.1.5 à $f = 0$ p.p. et non pas à $f = 0$. Pour la même raison, l'application $(f, g) \mapsto \int |f(x) - g(x)| dx$ vérifie l'inégalité du triangle mais n'est pas une distance.

Cela n'empêche pas de définir la *convergence en moyenne*. On dit qu'une suite f_j d'éléments de $\mathcal{L}^1(\mathbb{R}^n)$ converge en moyenne vers f si $\int |f_j(x) - f(x)| dx$ tend vers 0 pour $j \rightarrow \infty$. On n'a pas unicité de la limite f , mais on a le résultat suivant : pour que f_j converge également en moyenne vers une fonction $\tilde{f}$, il faut et il suffit que $f = \tilde{f}$ p.p.. Il est clair que la condition est suffisante, les intégrales de $|f_j(x) - f(x)|$ et de $|f_j(x) - \tilde{f}(x)|$ étant les mêmes. Réciproquement, si f_j converge en moyenne vers f et $\tilde{f}$, on a $\int |f - \tilde{f}| \leq \int |f - f_j| + \int |f_j - \tilde{f}|$ pour tout j . Le premier membre est donc nul, ce qui entraîne $f = \tilde{f}$ p.p. d'après la proposition 1.1.5.

Tous ces résultats suggèrent de passer au quotient par la relation d'équivalence $f = g$ p.p., ce que nous ferons dans la section 2.6. Le résultat suivant, que nous admettrons, est important.

Théorème 1.5.2. — Soit f un élément de $\mathcal{L}^1(\mathbb{R}^n)$. Il existe une suite f_j de fonctions continues à support compact (c'est-à-dire nulles en dehors d'un ensemble borné) qui converge vers f en moyenne.

Exercice 1.5.3. — Soit $f \in \mathcal{L}^1(\mathbb{R}^n)$. On définit la transformée de Fourier $\widehat{f}$ de f par $\widehat{f}(\xi) = \int e^{-ix \cdot \xi} f(x) dx$ pour tout $\xi \in \mathbb{R}^n$.

(a) Montrer que la fonction $\widehat{f}$ est continue et bornée et que l'on a $\sup_{\xi \in \mathbb{R}^n} |\widehat{f}(\xi)| \leq \int |f(x)| dx$.

(b) Démontrer que, si φ est continue à support compact, la fonction $\widehat{\varphi}$ tend vers 0 à l'infini (on montrera, en posant $\theta = \pi/\xi$, que $\widehat{f}(\xi) = - \int e^{-ix \cdot \xi} f(x+\theta) dx$).

(c) En déduire que, pour $f \in \mathcal{L}^1(\mathbb{R}^n)$, la fonction $\widehat{f}$ tend vers 0 à l'infini.

1.5.4. L'espace $\mathcal{L}^2(\mathbb{R}^n)$. — C'est l'espace des fonctions de carré sommable, c'est-à-dire telles que $\int |f(x)|^2 dx < \infty$. À partir des relations $|f(x) + g(x)|^2 \leq 2(|f(x)|^2 + |g(x)|^2)$ et $|f(x)g(x)| \leq 1/2(|f(x)|^2 + |g(x)|^2)$, on montre facilement que $\mathcal{L}^2(\mathbb{R}^n)$ est un espace vectoriel et que, pour deux éléments de cet espace, la quantité

$$(f | g) = \int_{\mathbb{R}^n} f(x) \overline{g(x)} dx \quad (1.12)$$

définit un élément de $\mathbb{C}$, la fonction à intégrer étant sommable.

Il est facile de voir que cette expression possède toutes les propriétés d'un produit scalaire, à l'exception de la suivante : la nullité de $(f | f)$ équivaut à $f = 0$ p.p. et non à $f = 0$. Cela nous suggère encore de passer au quotient par la relation d'équivalence $f = g$ p.p. Nous le ferons dans la section 2.6 où l'on trouvera d'autres propriétés des fonctions de carré sommable.

La quantité $(f - g | f - g)^{1/2} = \left(\int |f(x) - g(x)|^2 dx \right)^{1/2}$ est appelée *écart quadratique moyen* de f et g et on dit que f_j converge vers f en moyenne quadratique si l'écart quadratique moyen de f et f_j tend vers 0.

1.5.5. L'espace $\mathcal{L}^\infty(\mathbb{R}^n)$. — Il s'agit de l'espace des fonctions qui sont presque partout égales à une fonction bornée.

Théorème et Définition 1.5.6. — Soit f une fonction à valeurs réelles définie sur $\mathbb{R}^n$. On dit que $M \in \mathbb{R}$ est un presque majorant de f si on a $f(x) \leq M$ p.p. Si f est presque majorée, l'ensemble de ses presque majorants possède un plus petit élément, que l'on appelle la borne supérieure essentielle de f et que l'on note $\sup_{\text{ess}} f(x)$.

Soit μ la borne inférieure de l'ensemble des presque majorants de f . Il suffit de montrer que μ est encore un presque majorant de f . Pour chaque entier

j , le nombre $\mu + 1/j$ est un presque majorant et il existe donc un ensemble négligeable A_j en dehors duquel on a $f(x) \leq \mu + 1/j$. L'ensemble $A = \bigcup A_j$ est de mesure nulle, et on a $f(x) \leq \mu$ dans le complémentaire de A ce qui achève la démonstration.

On appelle $\mathcal{L}^\infty(\mathbb{R}^n)$ l'espace des fonctions essentiellement bornées, c'est-à-dire telles que la fonction $|f(x)|$ possède un presque majorant. On vérifie facilement que l'application $f \mapsto \sup \text{ess} |f(x)|$ possède toutes les propriétés d'une norme, à l'exception du fait que la nullité de $\sup \text{ess} |f(x)|$ implique seulement $f = 0$ p.p.

1.6. Sur la construction de l'intégrale

Le but de cette section est de donner un aperçu très superficiel de la construction de l'intégrale de Lebesgue et de dire quelques mots des questions de mesurabilité. Le lecteur intéressé par la théorie pourra se reporter par exemple à l'ouvrage de W. Rudin cité dans la bibliographie.

On définit d'abord la mesure des ensembles pavables, c'est à dire des ensembles qui sont réunion finie de pavés d'intérieur disjoint, comme étant égale à la somme des volumes de ces pavés — après avoir vérifié que le résultat ne dépend pas de la décomposition choisie.

Un premier passage à la limite définit la mesure d'un ouvert Ω de $\mathbb{R}^n$: c'est la borne supérieure des mesures des ensembles pavables contenus dans Ω . Un second passage à la limite définit la mesure d'un compact : c'est la borne inférieure des mesures des ouverts qui le contiennent. On dit alors qu'un ensemble A est *mesurable (pour la mesure de Lebesgue)* si, pour tout $\varepsilon > 0$, on peut trouver une suite de compacts K_j et une suite d'ouverts Ω_j tels que l'on ait $K_j \subset \Omega_j$,

$$\bigcup_{j=0}^{\infty} K_j \subset A \subset \bigcup_{j=0}^{\infty} \Omega_j \quad \text{et} \quad \sum_0^{\infty} \mu(\Omega_j \setminus K_j) \leq \varepsilon.$$

On définit alors la mesure de A par

$$\mu(A) = \inf_{\substack{\Omega \supset A \\ \Omega \text{ ouvert}}} \mu(\Omega) = \sup_{\substack{K \subset A \\ K \text{ compact}}} \mu(K) \leq +\infty.$$

Les résultats les plus importants que l'on obtient au terme de cette construction sont les suivants.

— Les ensembles mesurables forment une *tribu* c'est-à-dire une famille de parties de $\mathbb{R}^n$ qui est stable par passage au complémentaire et par réunion

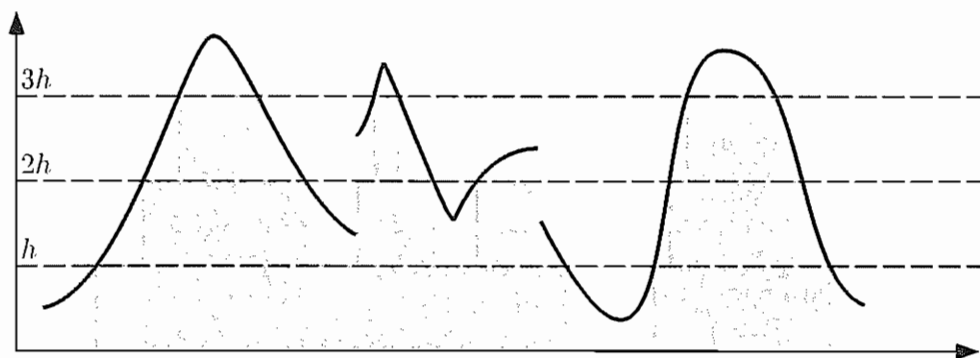
dénombrable (et donc par intersection dénombrable). Cette tribu contient notamment les ouverts et les fermés.

— La mesure μ est une application croissante et dénombrablement additive (voir le n°1.1.2) de la tribu des ensembles mesurables dans $\overline{\mathbb{R}_+}$.

On dit ensuite qu'une application f de $\mathbb{R}^n$ dans $\overline{\mathbb{R}_+}$ est mesurable si l'image réciproque de tout intervalle est un ensemble mesurable, et on définit enfin l'intégrale de f par la formule

$$\int f(x) dx = \sup_{h>0} \sum_{j=0}^{\infty} \mu(f^{-1}([jh, (j+1)h])) jh \quad (1.13)$$

(sauf dans le cas où f prend la valeur $+\infty$ sur un ensemble de mesure > 0 , auquel cas on pose bien entendu $\int f = +\infty$). On peut d'ailleurs remplacer $\sup_{h>0}$ par $\lim_{h \rightarrow 0}$ dans la formule ci-dessus. On interprète souvent cette



La fonction f_h et son intégrale

formule en disant que l'intégrale de Lebesgue découpe f selon l'axe des y , alors que l'intégrale de Riemann la découpe selon l'axe des x . L'expression (1.13) exprime effectivement $\int f(x) dx$ comme limite d'intégrales $\int f_h(x) dx$, où la fonction f_h est égale à kh lorsque $kh \leq f(x) < (k+1)h$.

On voit les avantages de cette méthode : les fonctions f_h approchent f à h près en norme uniforme, même pour des fonctions f très irrégulières, alors que l'approximation uniforme par des fonctions constantes sur des intervalles ne fonctionne que pour des fonctions très particulières (continues ou plus généralement réglées). En outre, l'approximation est "presque" croissante : on voit facilement que l'on a $f_{h/2} \geq f_h$, et la suite des $f_{2^{-n}}$ par exemple est croissante. Cela rend plus naturel le théorème de Beppo Levi, dont la démonstration n'est effectivement plus très difficile à ce stade.

Le prix à payer est bien sûr le fait que, pour définir l'intégrale de f_h , il faut disposer de la mesure des ensembles $f^{-1}([kh, (k+1)h])$, ensembles qui peuvent être très compliqués, ce qui nécessite toute la construction précédente. Au contraire, l'intégrale de Riemann n'utilisait que le concept élémentaire de mesure d'un intervalle.

1.6.1. Existe-t-il des ensembles et des fonctions non mesurables?

L'expérience suggère la réponse *non*. En effet, les ensembles mesurables forment une tribu contenant les ouverts et il en résulte que l'espace des fonctions mesurables contient les fonctions continues et est stable par toutes les opérations dénombrables usuelles : limite d'une suite (ou somme d'une série) de fonctions qui converge en chaque point, sup ou inf dénombrable, ... À titre d'exemple, le lecteur pourra voir dans l'exercice B.2.4 que la fonction égale à 1 en tout point rationnel et à 0 en tout point irrationnel — le type même de la fonction non intégrable au sens de Riemann, alors que c'est une excellente fonction sommable d'intégrale nulle — est limite d'une suite de fonctions dont chacune est limite d'une suite de fonctions continues. On peut bien sûr faire beaucoup plus compliqué, mais on n'arrive jamais à construire une fonction non mesurable sans faire appel à l'axiome du choix.

La véritable réponse à la question posée est : *cela dépend des axiomes mis à la base des mathématiques*. On a en effet les deux résultats suivants.

— Si on adjoint l'axiome du choix aux axiomes usuels de la théorie des ensembles, on peut prouver effectivement qu'il existe des ensembles non mesurables (voir l'exercice 1.6.2).

— Par contre, un résultat relativement récent de logique mathématique (Solovay, 1966) assure que l'on peut adjoindre à ces mêmes axiomes, sans introduire de contradiction, les formes dénombrables de l'axiome du choix et l'axiome "tout sous-ensemble de $\mathbb{R}^n$ est mesurable".

Dans la pratique cela signifie que, à moins de le faire exprès à l'aide de l'axiome du choix, il est exclu que l'on ait à considérer des fonctions non mesurables. C'est pourquoi ce cours a été écrit comme si toutes les fonctions étaient mesurables. La véritable raison est bien sûr une question de temps, il y a mieux à faire que de démontrer, par des méthodes répétitives, des résultats dont on sait d'avance qu'ils sont toujours vrais. Le lecteur n'aura qu'à ajouter mentalement l'adjectif "mesurable" chaque fois qu'il rencontrera le mot "ensemble" ou "fonction".

Cela dit, le lecteur excessivement scrupuleux qui serait choqué par cette façon de faire pourra se placer dans le système d'axiomes autorisé par Solovay. C'est un cadre dans lequel on peut développer toute l'analyse classique, et où tous les énoncés de ce chapitre sont effectivement des théorèmes.

Ce qui précède s'applique à la mesure de Lebesgue, et il ne faudrait pas en conclure que toutes les questions de mesurabilité sont sans intérêt. En théorie des probabilités, on introduit fréquemment plusieurs tribus (dépendant par exemple du temps), la mesurabilité d'une variable aléatoire X par rapport à telle ou telle tribu ayant un contenu probabiliste précis. Dans un tel contexte, la démonstration de la mesurabilité d'une variable aléatoire peut être un résultat important, et éventuellement difficile.

Exercice 1.6.2. — Soit Q_1 l'ensemble des nombres rationnels contenus dans $[-1, 1]$, dont on numérottera les éléments : $Q_1 = \{r_1, r_2, \dots\}$. On considère, dans l'ensemble $[0, 1]$, la relation d'équivalence $(x - y) \in Q_1$. On forme un ensemble A en choisissant un point et un seul dans chaque classe d'équivalence (on notera que l'ensemble des classes d'équivalences n'est pas dénombrable). Démontrer que A n'est pas mesurable. (On introduira les ensembles translatés $A_n = A + r_n$, on montrera que ces ensembles sont disjoints et que l'on a $[0, 1] \subset \bigcup_n A_n \subset [-1, 2]$. On montrera ensuite, en supposant A mesurable, que les hypothèses $\mu(A) = 0$ et $\mu(A) > 0$ conduisent toutes deux à une contradiction.)

1.7. Les quatre opérations

On dit souvent que l'intégration est l'opération inverse de la dérivation et il semble, nous y reviendrons, qu'il y ait quelque apparence de vérité dans cette assertion en dimension 1. Qu'en est-il en dimension supérieure?

Dans $\mathbb{R}^n$, il faut remplacer la dérivation par la différentiation, opération qui à une fonction f (de classe C^1 pour fixer les idées) associe sa différentielle $df = \sum (\partial f / \partial x_i) dx_i$. On connaît bien l'opération inverse, qui est la résolution des équations dites "aux différentielles totales" : les fonctions α_i étant données (disons de classe C^1 dans un ouvert Ω), déterminer f telle que $df = \sum \alpha_i dx_i$. Le lecteur n'ignore pas que les conditions $\partial \alpha_i / \partial x_j = \partial \alpha_j / \partial x_i$ sont nécessaires pour avoir l'existence d'une solution, et qu'elles sont suffisantes dans un "bon" ouvert (convexe, étoilé ou plus généralement "simplement connexe" conviennent).

Quoi qu'il en soit, l'opération inverse de la différentiation n'a rien à voir avec l'intégration dans $\mathbb{R}^n$, et on peut se demander quelle est l'opération inverse de celle-ci.

Avant de poursuivre, nous allons examiner un type de raisonnement courant en Physique. Par exemple, dans un ouvrage d'électrostatique où on demande de calculer le potentiel U créé à l'origine par une répartition de charges de densité $\rho(\mathbf{r})$, ($\mathbf{r} = (x, y, z)$), on pourra trouver le raisonnement suivant.

Considérons autour du point $\mathbf{r}$ un parallélépipède infinitésimal de côtés dx, dy, dz . Il porte une charge $\rho(\mathbf{r}) dx dy dz$ qui crée à l'origine un potentiel

$$dU = \frac{1}{4\pi\epsilon_0 r} \rho(\mathbf{r}) dx dy dz, \quad r = |\mathbf{r}|. \quad (1.14)$$

On a donc

$$U = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\mathbf{r})}{r} dx dy dz. \quad (1.15)$$

Ce texte pose quelques questions : d'où vient le *donc* ci-dessus ; s'il y a une différentielle dU existerait-il une fonction U , et si oui une fonction de quoi?

Pour répondre sur le plan mathématique aux interrogations qui précèdent, il faut en fait introduire deux concepts et énoncer deux théorèmes.

On appelle *mesure* (nous continuons à faire comme si tous les ensembles étaient mesurables) une application $A \mapsto m(A)$ qui à tout sous-ensemble A de $\mathbb{R}^n$ associe un élément de $\overline{\mathbb{R}_+}$, qui vérifie $m(\emptyset) = 0$ et la propriété d'additivité dénombrable : si des A_j , $j \in \mathbb{N}$, sont deux à deux disjoints, on a $m(\cup A_j) = \sum m(A_j)$.

On appelle, lorsqu'elle existe, *densité au point x de la mesure m par rapport à la mesure de Lebesgue μ* , la quantité $f(x)$ définie comme suit

$$f(x) = \frac{dm}{d\mu}(x) = \lim_{r \rightarrow 0} \frac{m(B(x, r))}{\mu(B(x, r))},$$

où $B(x, r)$ désigne la boule (que l'on pourrait remplacer par un cube) de centre x et de rayon r .

Cette quatrième opération, qui à une mesure associe sa densité, va être l'opération inverse de l'intégration. C'est ce que montrent les deux résultats suivants, que nous ne ferons qu'énoncer.

Théorème 1.7.1 (de dérivation de Lebesgue). — *Soit f une fonction positive et localement sommable dans $\mathbb{R}^n$. Alors la fonction d'ensemble m définie par $m(A) = \int_A f(x) dx$ est une mesure qui possède presque partout une densité et on a $\frac{dm}{d\mu}(x) = f(x)$ p.p.*

Théorème 1.7.2 (Radon-Nikodym). — *Soit m une mesure telle que l'on ait $m(K) < \infty$ pour tout compact K et $m(N) = 0$ pour tout ensemble N de mesure de Lebesgue nulle. Il existe alors une fonction f , sommable sur tout compact, telle que l'on ait $m(A) = \int_A f(x) dx$.*

Les mathématiques sous-jacentes au raisonnement électrostatique ci-dessus sont alors claires. On considère la fonction d'ensemble U qui à chaque $A \subset \mathbb{R}^n$ associe le potentiel $U(A)$ créé à l'origine par les charges contenues dans A . Il est raisonnable de penser que c'est une mesure, l'additivité (finie) étant explicitement énoncée dans les ouvrages d'électrostatique sous le nom de principe de superposition. Il reste à en calculer la densité $dU/d\mu$, ce qu'évoque assez bien l'argument infinitésimal et la formule (1.14), et à appliquer les deux théorèmes ci-dessus pour calculer $U = U(\mathbb{R}^3)$, ce qui est résumé dans le *donc*.

Ce type de raisonnement est très courant en physique et chaque fois que, pour une quantité Q , on écrit que dQ est proportionnel à l'élément de volume, on fait appel implicitement à la notion de mesure. Il s'agit d'un concept dont l'importance est comparable à celle du concept de fonction, et l'absence de référence explicite tient sans doute au fait qu'il n'est apparu historiquement qu'au début du XX^e siècle.

L'hypothèse du théorème de Radon-Nikodym ($\mu(N) = 0 \Rightarrow m(N) = 0$) est importante. Elle n'est pas respectée dans le cas de charges portées par des surfaces, des courbes ou des points, et ces cas sont étudiés séparément dans les ouvrages d'électrostatique (la formule (1.15) donnerait toujours $U = 0$).

Les quatre opérations existent en dimension 1, mais on ne les perçoit pas toujours comme différentes. Étant donné une mesure m (de masse finie pour simplifier) sur $\mathbb{R}$, on peut lui associer la fonction croissante F définie par $F(x) = m(]-\infty, x])$. Réciproquement, à une telle fonction F on associe facilement la fonction additive d'intervalles définie par $m([a, b]) = F(b) - F(a)$ (on peut en fait la prolonger en une mesure) et il est clair que, en un

point où F est dérivable, la mesure m admet une densité égale à $F'(x)$. L'opération inverse de la dérivation est le calcul des primitives, tandis que le calcul des "intégrales définies" est plutôt l'inverse du calcul de la densité d'une mesure.

CHAPITRE 2

TOPOLOGIE GÉNÉRALE ET ESPACES FONCTIONNELS

Les sections 2.1 à 2.4 sont probablement en grande partie connues du lecteur. Nous avons notamment rappelé, essentiellement sans démonstration, les propriétés fondamentales des espaces métriques. L'attention du lecteur est toutefois attirée sur les arguments de compacité, qui sont d'un usage constant dans la suite du cours, ainsi que sur les propriétés de base des espaces de Banach et des applications linéaires continues. Des exercices et des compléments en petits caractères montreront au lecteur curieux que ces résultats élémentaires permettent néanmoins d'obtenir assez rapidement des énoncés intéressants et non triviaux.

Les sections 2.5 (espaces de Hilbert) et 2.6 (notamment l'étude des espaces L^1, L^2, L^∞) sont sans doute nouvelles pour le lecteur, et sont importantes.

2.1. Espaces métriques (propriétés topologiques)

Un *espace métrique* (X, d) est un ensemble X muni d'une application d , appelée *distance*, de $X \times X$ dans $[0, +\infty[$ vérifiant les propriétés suivantes : d est symétrique (c'est-à-dire que $d(x, y) = d(y, x)$) ; la distance $d(x, y)$ est nulle si et seulement si on a $x = y$; et enfin d vérifie l'inégalité triangulaire

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

On appelle *boule ouverte* de centre $x \in X$ et de rayon $r > 0$ l'ensemble $B(x, r) = \{y \mid d(x, y) < r\}$. On dit qu'un sous-ensemble U de X est *ouvert* si pour tout point $x \in U$ il existe une boule ouverte centrée en x et contenue dans U . Un sous-ensemble de X est dit *fermé* si son complémentaire est ouvert.

La famille des ouverts est stable par réunion quelconque et par intersection finie. La famille des fermés est stable par intersection quelconque et par réunion finie.

Pour $x \in X$, on appelle *voisinage de x* un ensemble contenant un ouvert contenant x . Pour que V soit un voisinage de x , il faut et il suffit qu'il existe $r > 0$ avec $B(x, r) \subset V$. Une intersection finie de voisinages de x est encore un voisinage de x . Pour qu'un ensemble U soit ouvert, il faut et il suffit que U soit voisinage de chacun de ses points.

Si A est un sous-ensemble de X , la réunion des ouverts contenus dans A , qui est donc le plus grand ouvert contenu dans A , est appelée *intérieur* de A et est notée $\dot{A}$. Pour que $x \in \dot{A}$, il faut et il suffit que A soit un voisinage de x . L'intersection de tous les fermés contenant A , qui est donc le plus petit fermé contenant A , est appelée *adhérence* (ou *fermeture*) de A et est notée $\overline{A}$. Pour que $x \in \overline{A}$, il faut et il suffit que tout voisinage de x rencontre A . Le complémentaire de l'adhérence de A est l'intérieur du complémentaire de A et vice versa. L'ensemble $\overline{A} \setminus \dot{A}$ est appelé *frontière* de A .

On dit qu'un sous-ensemble A est *dense dans X* (ou encore est *partout dense*), si $\overline{A} = X$. Il est équivalent de dire que tout ouvert non vide de X rencontre A .

2.1.1. Limites et continuité. — Soient (E, d) et (E', d') deux espaces métriques, f une application de E dans E' et A un sous-ensemble de E . Soient $x_0 \in \overline{A}$ et $y_0 \in E'$. On dit que $\lim_{x \rightarrow x_0; x \in A} f(x) = y_0$ si, pour tout voisinage W de y_0 , il existe un voisinage V de x_0 tel que $f(V \cap A) \subset W$.

La limite, si elle existe, est unique.

On dit que f est *continue* en x_0 si $\lim_{x \rightarrow x_0; x \in E} f(x) = f(x_0)$ (on sous-entend habituellement le $x \in E$ dans la notation). Il est équivalent de dire que l'image réciproque de tout voisinage de $f(x_0)$ est un voisinage de x_0 (ou encore que pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que $d(x, x_0) \leq \eta$ implique $d'(f(x), f(x_0)) \leq \varepsilon$).

On dit que f est *continue (sur E)* si elle est continue en tout point de E . Il est équivalent de dire que l'image réciproque par f d'un ensemble ouvert [*resp.* fermé] est un ouvert [*resp.* fermé].

On appelle *homéomorphisme* de (E, d) sur (E', d') une bijection de E sur E' qui est continue ainsi que son inverse.

On dit qu'une suite x_j d'éléments de E converge vers x si pour tout voisinage V de x , il existe un indice j_0 à partir duquel tous les x_j appartiennent à V . Il est équivalent de dire que la suite numérique $j \mapsto d(x, x_j)$ converge vers 0.

Composition des limites : la composée de deux applications continues est continue ; si $\lim_{j \rightarrow \infty} x_j = x$ et si f est continue en x , alors $\lim_{j \rightarrow \infty} f(x_j) = f(x)$, etc.

Proposition 2.1.2. — *Pour qu'un sous-ensemble A de E soit fermé, il faut et il suffit que, pour toute suite d'éléments de A qui converge dans E , la limite appartienne à A .*

Proposition 2.1.3. — *Pour qu'une application f de (E, d) dans (E', d') soit continue, il faut et il suffit que, pour toute suite convergente x_j d'éléments de E , on ait $f(\lim_{j \rightarrow \infty} x_j) = \lim_{j \rightarrow \infty} f(x_j)$.*

Proposition 2.1.4. — *Soient f et g continues de (E, d) dans (E', d') . L'ensemble $\{x \in E \mid f(x) = g(x)\}$ est alors fermé. En particulier, si f et g coïncident sur un sous-ensemble partout dense, elles coïncident partout.*

2.1.5. Sous-espace métrique. — Soit (E, d) un espace métrique et soit A un sous-ensemble de E . L'ensemble A , muni de la restriction à $A \times A$ de la fonction d est lui-même un espace métrique. On parle alors du sous-espace (et non plus sous-ensemble) A de E .

2.1.6. Espace métrique produit. — Soient (E_1, d_1) et (E_2, d_2) deux espaces métriques. Leur produit est l'ensemble produit $E_1 \times E_2$ muni de la distance

$$d((x_1, x_2), (y_1, y_2)) = (d_1(x_1, y_1)^2 + d_2(x_2, y_2)^2)^{1/2}$$

Une application f à valeur dans un espace produit admet deux composantes f_1 à valeur dans E_1 et f_2 à valeur dans E_2 . Pour que f soit continue, il faut et il suffit que f_1 et f_2 soient continues.

Si (E, d) est un espace métrique, l'application $(x, y) \mapsto d(x, y)$ est continue de $E \times E$ dans $\mathbb{R}$. Pour tout point x_0 de E , l'application $x \mapsto d(x_0, x)$ est continue de E dans $\mathbb{R}$. Il en est de même, pour tout sous-ensemble $A \neq \emptyset$ de E , de l'application $x \mapsto d(x, A)$, où la distance de x à l'ensemble A est définie par $d(x, A) = \inf_{y \in A} d(x, y)$.

2.1.7. Distances topologiquement équivalentes. — Soient E un ensemble et d, d' deux distances sur E . On dit qu'elles sont topologiquement équivalentes si on a

$$\forall x \in E, \forall \varepsilon > 0, \exists \eta > 0, \forall y \in E, \quad d(x, y) \leq \eta \implies d'(x, y) \leq \varepsilon, \quad (2.1)$$

ainsi que la propriété analogue obtenue en échangeant d et d' . Il est équivalent de dire que l'application identique de E est un homéomorphisme de (E, d) sur (E, d') . Toutes les définitions et propriétés qui précèdent sont inchangées quand on remplace la distance par une distance topologiquement équivalente.

2.1.8. Espaces topologiques. — On appelle espace topologique un ensemble E muni d'une famille de sous-ensembles appelés ouverts satisfaisant aux conditions suivantes : l'ensemble vide et E lui-même sont des ouverts ; la famille des ouverts est stable par réunion quelconque et par intersection finie.

Les notions de fermés, voisinages, limite, continuité, ... se définissent comme précédemment. On dit que l'espace est *séparé* si pour tout couple (x, y) de points distincts, il existe un voisinage V de x et un voisinage W de y avec $V \cap W = \emptyset$. Cela assure l'unicité de la limite.

D'une manière générale, tous les énoncés qui précèdent (ne se référant pas explicitement à la distance) sont valables dans un espace topologique séparé à l'exception des propositions 2.1.2 et 2.1.3. La notion de suite convergente ne caractérise pas en général la topologie (on peut trouver des topologies différentes pour qui les suites convergentes sont les mêmes), et il faut faire appel à la notion de limite d'un filtre pour avoir des énoncés généralisant ces deux propositions.

La donnée d'une distance sur un ensemble lui confère une structure d'espace topologique séparé. Deux distances sont topologiquement équivalentes si et seulement si elles définissent le même espace topologique.

2.2. Espaces métriques (propriétés uniformes)

2.2.1. Distances uniformément équivalentes. — Soient E un ensemble et d, d' deux distances sur E . On dit qu'elles sont uniformément équivalentes si on a

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in E, \forall y \in E, \quad d(x, y) \leq \eta \implies d'(x, y) \leq \varepsilon, \quad (2.2)$$

ainsi que la propriété analogue obtenue en échangeant d et d' .

Le lecteur vérifiera facilement que les notions qui vont suivre restent inchangées lorsque l'on remplace la distance par une distance uniformément équivalente mais que, à la différence des notions du paragraphe précédent, elles peuvent changer si on la remplace par une distance topologiquement équivalente.

Exercice 2.2.2. — Sur $]0, \infty[$, on définit les deux distances $d(x, y) = |x - y|$ et $d'(x, y) = |\log x - \log y|$. Démontrer qu'elles sont topologiquement équivalentes mais pas uniformément équivalentes. La suite $(1/j)$ est-elle de Cauchy pour d , pour d' ? Pour quelle métrique l'espace $]0, \infty[$ est-il complet?

Définition 2.2.3. — Soit (E, d) un espace métrique.

(a) On dit qu'une suite x_j est une suite de Cauchy si on a

$$\lim_{j, k \rightarrow \infty} d(x_j, x_k) = 0.$$

(b) On dit que l'espace est complet si toute suite de Cauchy est convergente.

Bien entendu, dans tout espace métrique, une suite convergente est toujours de Cauchy.

Proposition 2.2.4. — *Soit (E, d) un espace métrique et $A \subset E$.*

(a) *Si E est complet et si A est fermé dans E , le sous-espace A est complet.*

(b) *Si le sous-espace A est complet, A est fermé dans E .*

Soit (x_j) une suite de Cauchy de A . La suite (x_j) est en particulier une suite de Cauchy dans E et converge donc vers un élément $x \in E$. Si A est fermé, d'après la proposition 2.1.2, le point x qui est limite d'une suite d'éléments de A appartient à A . Nous avons donc prouvé que, dans le sous-espace A , toute suite de Cauchy est convergente.

Pour démontrer la partie (b), il suffit de montrer d'après la proposition 2.1.2 que si une suite x_j d'éléments de A converge dans E , sa limite x appartient à A . La suite convergente x_j est une suite de Cauchy dans E , donc une suite de Cauchy dans le sous-espace A . Celui-ci étant complet, elle doit converger vers un élément de A , qui ne peut être que x par unicité de la limite.

Définition 2.2.5. — *Soient (E, d) et (E', d') deux espaces métriques et f une application de E dans E' . On dit que f est uniformément continue si, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que $d(x, y) \leq \eta$ implique $d'(f(x), f(y)) \leq \varepsilon$.*

On remarquera que deux distances d et d' sont uniformément équivalentes si et seulement si l'application identique est uniformément continue de (E, d) dans (E, d') et de (E, d') dans (E, d) .

On vérifie facilement que la fonction distance d elle-même, définie sur $E \times E$, ainsi que les fonctions $x \mapsto d(x, x_0)$ et $x \mapsto d(x, A)$ définies sur E pour $x_0 \in E$ et $A \subset E$, sont uniformément continues.

Proposition 2.2.6 (prolongement des applications)

Considérons deux espaces métriques (E, d) et (E', d') dont le second est complet. Soient F un sous-ensemble dense dans E et g une application uniformément continue du sous-espace F dans E' . Il existe alors une unique application uniformément continue $\tilde{g}$ de E dans E' qui coïncide avec g sur F .

Rappelons au lecteur le principe de la démonstration. Pour tout $x \in E$, il existe une suite x_j d'éléments de F qui converge vers x . Cette suite est de Cauchy dans le sous-espace F , et est transformée par l'application uniformément continue g en une suite de Cauchy de E' . La suite $g(x_j)$ converge donc vers un élément $y \in E'$. On montre facilement que si une autre suite x'_j converge vers x , la suite $g(x'_j)$ converge aussi vers y . On pose alors $\tilde{g}(x) = y$ et on vérifie que $\tilde{g}$ est uniformément continue. L'unicité d'un prolongement continu est déjà connue (proposition 2.1.4).

Un cas particulier important d'application uniformément continue est celui des *applications lipschitziennes*, c'est-à-dire telles que, pour une constante $C > 0$ convenable, on ait $d'(f(x), f(y)) \leq Cd(x, y)$ (on dit alors que f est lipschitzienne de rapport C). Il faut noter que ce concept (et en particulier celui d'application contractante ci-dessous) n'est pas invariant lorsque l'on remplace une distance par une distance uniformément équivalente. Les exercices 2.2.8 et 2.2.9 montreront qu'un choix judicieux de la distance peut accroître notablement la portée du théorème suivant.

Théorème 2.2.7 (du point fixe de Picard). — Soit (E, d) un espace métrique complet, et f une application contractante de E dans lui-même (c'est-à-dire lipschitzienne de rapport $k < 1$). Alors il existe un unique $x \in E$ tel que $f(x) = x$.

La démonstration est simple et fournit de plus un algorithme pour déterminer x . On choisit un point quelconque $x_0 \in E$, et on définit par récurrence la suite x_j par $x_{j+1} = f(x_j)$ pour $j \geq 0$. On montre facilement que $d(x_{j+1}, x_j) \leq k^j d(x_1, x_0)$ et, par inégalité triangulaire, que la suite x_j est de Cauchy. Elle converge donc vers un point x et, par continuité de f , on a $f(x) = x$. L'unicité est facile : si x et y sont deux points fixes, on a $d(x, y) \leq kd(x, y)$, ce qui implique $d(x, y) = 0$.

Exercice 2.2.8 (Équations intégrales de seconde espèce)

On se donne une fonction $K(x, y)$ à valeurs réelles, définie et continue sur $\{(x, y) \in \mathbb{R}^2 \mid 0 \leq y \leq x\}$. Soit $M(T) = \sup \{|K(x, y)| \mid 0 \leq y \leq x \leq T\}$ pour $T \geq 0$. Étant donné une fonction g définie et continue sur $[0, \infty[$, on cherche une fonction continue u vérifiant

$$u(x) - \int_0^x K(x, y)u(y) dy = g(x). \quad (2.3)$$

Le lecteur se rappellera (ou redémontrera) que l'espace $C([0, T])$ des fonctions continues sur $[0, T]$ est complet, si on le munit de la distance de la convergence uniforme : $d(u, v) = \sup_{x \in [0, T]} |u(x) - v(x)|$.

(a) Soit T assez petit pour que $TM(T) < 1$. Démontrer qu'il existe une unique solution définie sur $[0, T]$ de (2.3) (on appliquera le théorème précédent à l'application $u \mapsto g + \int_0^x K(x, y)u(y) dy$).

(b) Montrer que, sur $C([0, T])$, la distance d_γ définie pour $\gamma > 0$ par $d_\gamma(u, v) = \sup_{x \in [0, T]} e^{-\gamma x} |u(x) - v(x)|$, est uniformément équivalente à d .

(c) Démontrer que, quel que soit T , il existe une unique solution définie sur $[0, T]$ de (2.3) (on choisira γ assez grand pour que $M(T)/\gamma < 1$).

(d) En déduire qu'il existe une et une seule solution définie sur $[0, \infty[$ de l'équation (2.3).

Exercice 2.2.9 (Une version précisée du théorème de Cauchy-Lipschitz)

On se donne une fonction $F(x, z)$ à valeurs réelles, définie et continue sur $[0, \infty[\times \mathbb{R}$ et un nombre $u_0 \in \mathbb{R}$. On suppose que F est uniformément lipschitzienne par rapport à z , c'est-à-dire qu'il existe $C > 0$ telle que l'on ait $|F(x, z) - F(x, z')| \leq C |z - z'|$ quels que soient x, z et z' .

(a) Démontrer qu'une fonction u définie sur $[0, T]$ est une solution continûment dérivable de l'équation différentielle

$$u'(x) = F(x, u(x)), \quad u(0) = u_0, \quad (2.4)$$

si et seulement si u est une solution continue de l'équation intégrale

$$u(x) = u_0 + \int_0^x F(t, u(t)) dt. \quad (2.5)$$

(b) En utilisant la distance d de l'exercice précédent, démontrer l'existence et l'unicité d'une solution continue de (2.5) sur $[0, T]$ pour tout $T < 1/C$.

(c) En utilisant la distance d_γ , et en choisissant $\gamma > CT$, démontrer l'existence et l'unicité d'une solution continue de (2.5) dans $[0, T]$ quel que soit T .

(d) En déduire qu'il existe une et une seule solution de (2.4) continûment dérivable sur $[0, \infty[$.

(e) Étendre le résultat au cas des systèmes différentiels ($z \in \mathbb{R}^n$), et au cas où le caractère lipschitzien est local en x et global en z (c'est-à-dire que, pour $x \leq T$, on a $|F(x, z) - F(x, z')| \leq C(T) |z - z'|$ quels que soient z et z'). Dans l'énoncé usuel du théorème de Cauchy-Lipschitz, le caractère lipschitzien est local en x et en z , mais on n'obtient que l'existence *locale* d'une solution.

Remarque 2.2.10. — Dans un espace topologique, les concepts d'application uniformément continue et de suite de Cauchy n'ont aucun sens. Il existe une structure d'espace *uniforme*, plus générale que celle d'espace métrique, pour laquelle ces notions peuvent être définies. Les groupes (et en particulier les espaces vectoriels), munis d'une topologie compatible avec leur structure de groupe, sont naturellement munis de structures d'espace uniforme.

2.3. Espaces métriques compacts

A. Généralités

Définition 2.3.1 (Borel-Lebesgue). — On dit qu'un espace métrique (E, d) est compact si, étant donné un recouvrement ouvert quelconque de E (c'est-à-dire une famille $(U_i)_{i \in I}$ d'ouverts dont la réunion est E), on peut toujours en extraire un recouvrement fini (c'est-à-dire que l'on peut trouver un nombre fini d'indices $(i_1, \dots, i_n)$ avec $E = U_{i_1} \cup \dots \cup U_{i_n}$).

Cette définition a l'avantage de se généraliser aux espaces topologiques quelconques. Dans les espaces métriques, les propriétés suivantes sont parfois plus maniables.

Théorème 2.3.2. — Soit (E, d) un espace métrique. Les trois propriétés suivantes sont équivalentes.

- (a) E est compact.
- (b) Toute suite décroissante de fermés non vides de E a une intersection non vide.
- (c) De toute suite d'éléments de E , on peut extraire une sous-suite convergente.

Pour montrer que (a) implique (b), supposons qu'il existe une suite F_j décroissante de fermés non vides dont l'intersection est vide. Leurs complémentaires constituent alors un recouvrement ouvert de E , dont on peut extraire un recouvrement fini $\mathcal{C}F_{j_1}, \dots, \mathcal{C}F_{j_n}$. Si j est le plus grand de ces indices, on a donc $\mathcal{C}F_j = E$ et $F_j = \emptyset$ ce qui est la contradiction cherchée.

Montrons que (b) entraîne (c). Soit donc x_j une suite d'éléments de E et notons F_j l'adhérence de $A_j = \{x_k \mid k \geq j\}$. Les F_j forment une suite décroissante de fermés et d'après (b) il existe un point y appartenant à tous les F_j . Il est alors facile de construire par récurrence une suite croissante $j(k)$ telle que $d(y, x_{j(k)}) \leq 1/k$: $j(k)$ étant construit, et y appartenant à l'adhérence de $A_{j(k)+1}$, on peut trouver un élément de cet ensemble (dont l'indice sera noté $j(k+1)$) dont la distance à y est inférieure à $1/(k+1)$. La suite $x_{j(k)}$ est donc convergente.

Il est un peu plus délicat de montrer que (c) implique (a). Montrons d'abord que pour tout $\varepsilon > 0$, il existe un recouvrement fini de E par des boules de rayon ε . Dans le cas contraire, en choisissant x_1 arbitraire, il existerait x_2 n'appartenant pas à $B(x_1, \varepsilon)$. Il existerait de même x_3 n'appartenant pas à $B(x_1, \varepsilon) \cup B(x_2, \varepsilon)$ et ainsi de suite. On construirait ainsi une suite x_j de points dont les distances mutuelles seraient $\geq \varepsilon$, cette propriété serait encore vraie pour toute suite extraite, qui ne pourrait donc converger.

Montrons ensuite, toujours sous l'hypothèse (c), que si $(U_i)_{i \in I}$ est un recouvrement ouvert de E , il existe $\varepsilon > 0$ tel que toute boule de rayon ε soit contenue dans au moins l'un des U_i . Dans le cas contraire, on pourrait pour tout n trouver une boule $B(x_n, 1/n)$ qui ne serait contenue dans aucun des U_i . Une sous-suite x_{n_k} convergerait vers un point y . L'un au moins des U_i doit contenir y , et doit donc contenir une boule $B(y, \rho)$ pour un certain $\rho > 0$. Cela conduit à une contradiction dès que k est assez grand pour avoir $d(y, x_{n_k}) < \rho/2$ et $1/n_k \leq \rho/2$.

Il est maintenant facile de conclure. Soit $(U_i)_{i \in I}$ un recouvrement ouvert de E , et soit $\varepsilon > 0$ déterminé comme ci-dessus. On peut recouvrir E par un nombre fini de boules de rayon ε . Choisissons, pour chacune d'elles, l'un des U_i qui la contienne. On obtient ainsi un recouvrement de E par un nombre fini d'ouverts de la famille, ce qui achève la démonstration.

Corollaire 2.3.3. — Soient X et Y deux espaces métriques compacts. L'espace produit $X \times Y$ est alors compact.

Soit (x_j, y_j) une suite d'éléments de $X \times Y$ et montrons que l'on peut en extraire une sous-suite convergente. On peut en extraire une première sous-suite notée (x'_j, y'_j) de telle sorte que x'_j converge vers un point x , puis extraire de celle-ci une seconde sous-suite (x''_j, y''_j) telle que y''_j converge vers un point y . La suite (x''_j, y''_j) converge alors vers (x, y) , ce qui achève la démonstration.

Proposition 2.3.4. — Soit (E, d) un espace métrique et soit $A \subset E$.

(a) Si le sous-espace A est compact, il est fermé dans E .

(b) Si E est compact et si A est fermé dans E , le sous-espace A est compact.

Pour montrer le premier point, il suffit de montrer que si une suite d'éléments de A converge dans E , la limite y appartient à A . Or la compacité assure qu'une sous-suite converge vers un point de A qui ne peut être que y .

Pour montrer le second point, considérons une suite quelconque d'éléments de A . On peut en extraire une sous-suite qui converge dans E . Comme A est fermé, la limite appartient à A ce qui, d'après 2.3.2(c) achève la démonstration.

Proposition 2.3.5. — Tout espace métrique compact est complet

Soit x_i une suite de Cauchy. Il existe alors une sous-suite $x_{j(k)}$ qui converge vers $y \in E$. Il est facile de voir que la suite x_i elle-même converge vers y : pour i assez grand, en choisissant k assez grand, on a $d(x_i, x_{j(k)}) \leq \varepsilon/2$ et $d(y, x_{j(k)}) \leq \varepsilon/2$.

Le lecteur pourrait s'étonner de voir une propriété de nature topologique (la compacité) entraîner une propriété de nature uniforme. L'exercice suivant en fournit une explication.

Exercice 2.3.6. — Soit (E, d) un espace métrique compact, et d' une distance topologiquement équivalente à d . Montrer que d' est uniformément équivalente à d .

Théorème 2.3.7. — Soient E un espace métrique compact, E' un espace métrique et f une application continue de E dans E' . Alors l'image $f(E)$ est un sous-espace compact.

Soit $y_j = f(x_j)$ une suite quelconque d'éléments de $f(E)$. Il suffit d'extraire de la suite x_j une sous-suite qui converge vers un point x , et la suite des images, qui est une sous-suite de la suite (y_j) , converge vers $f(x)$.

Un cas particulier et important est celui où $E' = \mathbb{R}$. L'image $f(E)$ est alors un compact de $\mathbb{R}$, et on sait qu'un tel ensemble est borné et que $\sup f(E)$ et

$\inf f(E)$ appartiennent à $f(E)$. On retrouve le fait qu'une fonction numérique continue sur un compact est bornée et atteint ses bornes.

Théorème 2.3.8. — Soient (E, d) et (E', d') deux espaces métriques, (E, d) étant compact. Toute application continue de E dans E' est alors uniformément continue.

Soit $\varepsilon > 0$. Pour chaque $x \in E$ il existe $\eta_x > 0$ tel que $y \in B(x, \eta_x)$ entraîne $d'(f(y), f(x)) \leq \varepsilon/2$. Les $B(x, \eta_x/2)$ forment un recouvrement ouvert de E , dont on peut extraire un recouvrement fini $(B(x_i, \eta_{x_i}/2))_{i \in I}$. Posons $\eta = \min \{\eta_{x_i}/2 \mid i \in I\}$. Si $d(x, y) < \eta$, les points x et y appartiennent tous deux à l'une des boules $B(x_i, \eta_{x_i})$, et on a donc $d'(f(x), f(y)) \leq \varepsilon$ ce qui achève la démonstration.

Exercice 2.3.9 (Théorème de Dini). — Soit X un espace métrique compact et (f_n) une suite décroissante d'applications continues de X dans $\mathbb{R}$ telle que l'on ait $\lim f_n(x) = 0$ en tout point $x \in X$. Démontrer que la suite (f_n) converge uniformément vers 0.

B. Exemples et applications

Le lecteur sait déjà que les sous-espaces compacts de $\mathbb{R}^n$ (ou d'un espace vectoriel normé de dimension finie sur $\mathbb{R}$) sont exactement les parties fermées et bornées de $\mathbb{R}^n$. Nous verrons plus loin que cette propriété caractérise les espaces de dimension finie.

Théorème et Définition 2.3.10 (Suite exhaustive de compacts)

Soit Ω un ouvert de $\mathbb{R}^n$. On dit qu'une suite $(K_j)_{j \in \mathbb{N}}$ de compacts contenus dans Ω est une suite exhaustive de compacts si on a $K_j \subset \overset{\circ}{K}_{j+1}$ pour tout j et si $\Omega = \bigcup_j K_j$.

Dans tout ouvert Ω de $\mathbb{R}^n$, il existe des suites exhaustives de compacts.

Si (K_j) est une suite exhaustive de compacts, pour tout compact K de Ω , il existe un indice j tel que $K \subset K_j$.

Pour montrer l'existence, il suffit de définir K_j comme étant l'ensemble des x dont la distance à $\mathbb{C}\Omega$ est $\geq 1/j$ et qui vérifient $|x| \leq j$. Il est clair que K_j est un fermé borné de $\mathbb{R}^n$, que la réunion des K_j est Ω , et que chaque point de K_j est centre d'une boule contenue dans K_{j+1} .

Si K est un compact de Ω , les $\overset{\circ}{K}_j$ constituent un recouvrement ouvert de K dont on peut extraire un recouvrement fini. Si j_0 est le plus grand des indices de ce recouvrement, on a $K \subset K_{j_0}$.

Exercice 2.3.11. — Donner un exemple dans $]0, 1[$ d'une suite croissante $\overset{\circ}{K}_j$ de compacts dont la réunion est $]0, 1[$ et d'un compact K qui n'est inclus dans aucun K_j .

Exercice 2.3.12. — On dit qu'un espace métrique X est *localement compact* si tout point $x \in X$ possède un voisinage compact. Montrer que tout espace métrique localement compact qui est réunion dénombrable de compacts possède une suite exhaustive de compacts.

Exercice 2.3.13. — Étant donnés deux sous-ensembles F et G de $\mathbb{R}^n$, on définit leur somme algébrique $F + G = \{x \mid \exists y \in F, \exists z \in G, x = y + z\}$.

- (a) Démontrer que, si F est compact et G fermé, l'ensemble $F + G$ est fermé.
- (b) Donner des exemples de couples d'ensembles fermés dont la somme n'est pas fermée. On pourra considérer dans le plan les ensembles F et G définis respectivement par $xy \geq 1; y > 0$ et par $xy \leq -1; y > 0$. On pourra aussi considérer sur $\mathbb{R}$ les ensembles $F = \{1, 2, 3, \dots\}$ et $G = \{-2+1/2, -3+1/3, \dots\}$.

Exercice 2.3.14. — On dit que deux ensembles fermés F et G de $\mathbb{R}^n$ sont convolutifs si on a la propriété suivante : pour tout $R > 0$, il existe $\rho(R)$ tel que les relations $y \in F, z \in G$ et $|y + z| \leq R$ impliquent $|y| \leq \rho(R)$ et $|z| \leq \rho(R)$.

- (a) Démontrer que si F et G sont des fermés convolutifs, leur somme est fermée.
- (b) Démontrer que si F est compact et G fermé, F et G sont convolutifs.
- (c) On suppose que F et G sont des fermés coniques, c'est-à-dire que $x \in F$ et $\lambda \in [0, \infty[$ entraîne $\lambda x \in F$. Démontrer que F et G sont convolutifs si et seulement si $F \cap (-G) = \{0\}$.

L'exercice suivant montre un exemple important de compacité en dimension infinie

Exercice 2.3.15 (Théorème d'Ascoli). — Soit (X, δ) un espace métrique compact, et $C(X)$ l'espace des fonctions continues sur X à valeurs réelles, muni de la distance de la convergence uniforme $d(f, g) = \sup_{x \in X} |f(x) - g(x)|$.

On dit qu'un ensemble $\mathcal{A} \subset C(X)$ est *uniformément équicontinu* si pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que, quels que soient $x \in X, y \in X$ et $f \in \mathcal{A}$, on ait $\delta(x, y) \leq \eta \Rightarrow |f(x) - f(y)| \leq \varepsilon$. Le théorème d'Ascoli assure que $\mathcal{A}$ est relativement compact dans $C(X)$ (c'est-à-dire contenu dans un compact de $C(X)$) si et seulement si

- (i) $\mathcal{A}$ est équicontinu
- (ii) Pour tout $x \in X$, l'ensemble $\{f(x) \mid f \in \mathcal{A}\}$ est borné.

- (a) Soit $\mathcal{A}$ une partie compacte de $C(X)$. Démontrer que (i) et (ii) sont vérifiés (on montrera que, dans le cas contraire, il existerait $\varepsilon > 0$ et des suites f_n, x_n et y_n telles que $\delta(x_n, y_n) \leq 1/n$ et $|f_n(x_n) - f_n(y_n)| \geq \varepsilon$, et que la suite f_n ne pourrait avoir de sous-suite uniformément convergente).

(b) On se donne réciproquement une partie $\mathcal{A}$ de $C(X)$ qui vérifie les propriétés (i) et (ii). Montrer que son adhérence $\overline{\mathcal{A}}$ dans $C(X)$ est compact. On remarque d'abord que $\overline{\mathcal{A}}$ vérifie aussi (i) et (ii), et il faut montrer que l'on peut extraire d'une suite quelconque f_n d'éléments de $\overline{\mathcal{A}}$ une sous-suite convergente.

(b1) Montrer qu'il existe un sous-ensemble dénombrable $x_1, x_2, \dots$ dense dans X (pour chaque n , on recouvrira X par un nombre fini de boules de rayon $1/n$ et on prendra l'ensemble des centres de ces boules).

(b2) Extraire de la suite f_n une sous-suite (que l'on notera $f_n^{(1)}$) telle que la suite numérique $f_n^{(1)}(x_1)$ soit convergente, puis extraire de la suite $f_n^{(1)}$ une sous-suite (notée $f_n^{(2)}$) qui converge aussi au point $x_2, \dots$. Les suites $f_n^{(k)}$ étant définies par récurrence, on définira une suite g_n (procédé diagonal) par $g_n = f_n^{(n)}$. Démontrer que g_n est une suite extraite de la suite f_n et que $g_n(x_j)$ converge pour tout j vers une limite que l'on notera $g(x_j)$.

(b3) Montrer que la fonction g ainsi définie sur un sous-ensemble dense est uniformément continue, qu'elle possède donc un prolongement continu $\tilde{g}$ sur X , et que la suite g_n converge uniformément vers $\tilde{g}$.

(c) En déduire par exemple que l'ensemble des fonctions f lipschitziennes de rapport C et vérifiant $|f(0)| \leq M$, où C et M sont donnés, est compact dans $C([0, 1])$.

Exercice 2.3.16. — Montrer, toujours sur un espace métrique compact X , qu'une partie $\mathcal{A}$ est uniformément équicontinue si et seulement si elle est équicontinue en chaque point, c'est-à-dire que pour chaque $x_0 \in X$ et chaque $\varepsilon > 0$, on peut trouver $\eta > 0$ tel que, pour tout $f \in \mathcal{A}$, on ait $\delta(x_0, y) \leq \eta \Rightarrow |f(x_0) - f(y)| \leq \varepsilon$.

Théorème 2.3.17 (Stone-Weierstrass). — (a) Soit X un espace métrique compact et notons $C(X)$ l'espace des fonctions continues sur X à valeurs réelles, muni de la distance de la convergence uniforme. Soit $\mathcal{F}$ une sous-algèbre de $C(X)$ (c'est-à-dire un sous-espace vectoriel stable par produit) qui contient les constantes et qui sépare les points de X (c'est-à-dire que pour tout couple (x, y) de points distincts, il existe $f \in \mathcal{F}$ avec $f(x) \neq f(y)$). L'espace $\mathcal{F}$ est alors dense dans $C(X)$.

(b) On se donne maintenant une sous-algèbre $\mathcal{G}$ de l'espace $C(X; \mathbb{C})$ des fonctions continues à valeurs complexes, qui contient les constantes, qui sépare les points de X et qui est stable par conjugaison complexe ($g \in \mathcal{G}$ implique $\bar{g} \in \mathcal{G}$). L'espace $\mathcal{G}$ est alors dense dans $C(X; \mathbb{C})$.

Il est facile de voir que, sous les hypothèses de (b), l'espace $\mathcal{F}$ des fonctions de $\mathcal{G}$ qui sont à valeurs réelles vérifie les hypothèses de (a) : si f sépare les points x et y , l'une des fonctions $(f + \bar{f})$ ou $(f - \bar{f})/i$ les sépare aussi. Pour toute fonction continue à valeurs complexes, il suffit alors d'approcher uniformément ses parties réelle et imaginaire en utilisant (a).

Nous noterons $\overline{\mathcal{F}}$ l'adhérence de $\mathcal{F}$ dans $C(X)$. Nous aurons besoin du lemme élémentaire suivant.

Lemme 2.3.18. — Il existe une suite P_n de polynômes d'une variable telle que les fonctions $P_n(t)$ convergent uniformément vers $|t|$ sur l'intervalle $[-1, 1]$.

Pour n donné, il est facile de trouver $\alpha \in]0, 1[$ tel que la distance (de la convergence uniforme sur $[-1, 1]$) des fonctions $|t|$ et $\sqrt{t^2 + \alpha^2}$ soit inférieure à $1/2n$. D'autre part, la fonction $\sqrt{x}$ est développable en une série entière (de la variable $x - 1$) dans le disque de centre 1 et de rayon 1. Les sommes partielles de cette série convergent uniformément sur tout compact de ce disque. On peut donc trouver une des sommes partielles, qui est un polynôme $Q_n(x)$, telle que l'on ait

$$\sup_{x \in [\alpha, \sqrt{1+\alpha^2}]} |\sqrt{x} - Q_n(x)| \leq 1/2n.$$

Il suffit de poser $P_n(t) = Q_n(t^2 + \alpha^2)$, et on a $\sup_{|t| \leq 1} |P_n(t) - |t|| \leq 1/n$.

Lemme 2.3.19. — *Sous les hypothèses du théorème 2.3.17(a), l'algèbre $\overline{\mathcal{F}}$ est stable pour les opérations $(f, g) \mapsto \sup(f, g)$ et $(f, g) \mapsto \inf(f, g)$.*

Il est facile de montrer que $\overline{\mathcal{F}}$ est une algèbre. Compte tenu des relations $\sup(f, g) = (f+g + |f-g|)/2$ et $\inf(f, g) = (f+g - |f-g|)/2$, il suffit de prouver que $f \in \overline{\mathcal{F}}$ entraîne $|f| \in \overline{\mathcal{F}}$ et, quitte à multiplier f par une constante convenable, on peut supposer $|f| \leq 1$. D'après le lemme 2.3.18, les fonctions $x \mapsto P_n(f(x))$, qui sont des éléments de l'algèbre $\overline{\mathcal{F}}$ convergent uniformément vers $|f|$, ce qui achève la démonstration du lemme.

2.3.20. Démonstration du théorème 2.3.17. — Soient $f \in C(X)$ et $\varepsilon > 0$ fixés. Pour tout couple de points x, y de X , il est possible de trouver une fonction $g_{xy} \in \mathcal{F}$ telle que $g_{xy}(x) = f(x)$ et $g_{xy}(y) = f(y)$ (prendre une combinaison linéaire de la constante 1 et d'une fonction séparant x et y). Fixons provisoirement le point x , et notons U_y le voisinage ouvert de y ainsi défini : $\{z \mid g_{xy}(z) > f(z) - \varepsilon\}$. Les U_y forment un recouvrement ouvert dont on peut extraire un recouvrement fini $U_{y_1}, \dots, U_{y_n}$. En posant $h_x = \sup \{g_{xy_1}, \dots, g_{xy_n}\}$, nous avons construit pour chaque x un élément $h_x \in \overline{\mathcal{F}}$ qui vérifie $h_x(z) \geq f(z) - \varepsilon$ pour tout z , et $h_x(x) = f(x)$.

Pour chaque x , l'ensemble $V_x = \{z \mid h_x(z) < f(z) + \varepsilon\}$ est un voisinage ouvert de x . On peut extraire de la famille des V_x un recouvrement fini $V_{x_1}, \dots, V_{x_p}$ et poser $\varphi = \inf \{h_{x_1}, \dots, h_{x_p}\}$. Nous avons ainsi construit, pour chaque $\varepsilon > 0$, un élément φ qui appartient à $\overline{\mathcal{F}}$ et qui vérifie $f(z) - \varepsilon \leq \varphi(z) \leq f(z) + \varepsilon$ pour tout z , ce qui achève la démonstration du théorème.

Exemple 2.3.21. — Soit K un compact de $\mathbb{R}^n$. Le sous-espace de $C(K)$ constitué des polynômes de n variables vérifie les hypothèses du théorème de Stone-Weierstrass, et on a donc le résultat suivant : pour toute fonction f continue sur K à valeurs réelles [resp. complexes], il existe une suite P_n de polynômes à coefficients réels [resp. complexes] qui converge vers f uniformément sur K .

On peut également écrire f comme somme de la série "téléscopique" $f = P_0 + (P_1 - P_0) + (P_2 - P_1) + \dots$, série de polynômes qui converge uniformément. Cela n'implique nullement que f soit somme d'une série uniformément convergente de monômes (ce qui signifierait que f est développable en série entière).

Exemple 2.3.22. — Soit $\Gamma = \{z \mid |z| = 1\}$ le cercle unité du plan complexe. Le sous-ensemble de $C(\Gamma)$ constitué des fonctions de la forme $\sum_{-p}^q a_n z^n$, $a_n \in \mathbb{C}$ vérifie les hypothèses du théorème 2.3.17(b) (on a $\bar{z} = z^{-1}$ sur Γ).

Les fonctions F continues sur Γ s'identifiant aux fonctions f continues 2π -périodiques sur $\mathbb{R}$ (en posant $f(t) = F(e^{it})$), il en résulte que pour toute fonction continue 2π -périodique f , il existe une suite P_n de polynômes trigonométriques (c'est-à-dire de fonctions de la forme $\sum_{-p}^q a_n e^{int}$) qui converge uniformément vers f .

Là encore, on ne confondra pas cette propriété avec le fait que la série de Fourier de f converge uniformément (ce qui n'est pas vrai pour toute fonction continue périodique).

C. Partitions de l'unité

Il s'agit d'une première approche de concepts qui nous seront utiles ultérieurement, mais dans $\mathbb{R}^n$ et pour des fonctions différentiables. Une partie de la théorie est toutefois de nature purement topologique, et nous la développons ici dans le cadre des fonctions continues sur un espace métrique.

Rappelons que le *support* d'une fonction continue à valeurs réelles ou complexes est le fermé $\text{Supp}(\varphi) = \overline{\{x \mid \varphi(x) \neq 0\}}$.

Proposition 2.3.23. — *Soient F et G deux fermés disjoints d'un espace métrique X . Il existe une application f continue de X dans $[0, 1]$, égale à 0 au voisinage de F et à 1 au voisinage de G .*

La fonction $d(x, F)$ est continue sur X et, F étant fermé, elle ne s'annule que sur F . La fonction g définie par

$$g(x) = \frac{d(x, F)}{d(x, F) + d(x, G)}$$

est continue, le dénominateur ne s'annulant jamais, est à valeurs dans $[0, 1]$ et vaut 0 et 1 respectivement sur F et G . Elle n'est pas nulle au voisinage de F , mais elle permet de définir le fermé $F_1 = g^{-1}([0, 1/3])$ qui contient l'ouvert $g^{-1}([0, 1/3[)$ et est donc un voisinage de F . On définit de même le fermé $G_1 = g^{-1}([2/3, 1])$. Il suffit d'appliquer la construction précédente au couple (F_1, G_1) pour obtenir une fonction f répondant à la question.

Proposition 2.3.24. — *Soient X un espace métrique, K un compact de X et (U_j) , $j = 1, \dots, N$ une famille finie d'ouverts. On suppose $K \subset \bigcup_{j=1}^N U_j$. Il existe alors des compacts $K_j \subset U_j$ tels que $K \subset \bigcup_{j=1}^N K_j$.*

Chaque $x \in K$ appartient à l'un des U_j et est donc centre d'une boule ouverte B_x telle que la boule fermée $\tilde{B}_x$ de même rayon⁽¹⁾ soit contenue dans U_j . Les B_x recouvrent K , et on peut donc en extraire un recouvrement fini $B_{x_1}, \dots, B_{x_p}$. Pour chaque $j \in \{1, \dots, N\}$, soit A_j l'ensemble des l appartenant à $\{1, \dots, p\}$ tels que $\tilde{B}_{x_l}$ soit contenu dans U_j . Les compacts $K_j = K \cap (\bigcup_{l \in A_j} \tilde{B}_{x_l})$ répondent à la question.

⁽¹⁾ Elle contient l'adhérence $\overline{B_x}$ de B_x mais ne lui est pas nécessairement égale : considérer $B(0, 1)$ dans l'espace $X =]-\infty, 0] \cup [1, \infty[$.

Théorème 2.3.25. — *Sous les hypothèses de la proposition précédente, il existe des fonctions φ_j , $j = 1, \dots, N$ continues de X dans $[0, 1]$ telles que l'on ait $\text{Supp}(\varphi_j) \subset U_j$ et*

$$\sum_{j=1}^N \varphi_j = 1 \quad \text{au voisinage de } K.$$

Soient K_j les compacts fournis par la proposition précédente. D'après la proposition 2.3.23, il existe des fonctions continues positives ψ_j égales à 1 au voisinage de K_j et à 0 au voisinage de $\complement U_j$ (c'est-à-dire à support dans U_j). En notant V l'ouvert où $\sum_{j=1}^N \psi_j$ est non nul, il existe de même une fonction ψ_0 égale à 0 au voisinage de K et à 1 au voisinage du complémentaire de V . La fonction $\sum_{j=0}^N \psi_j$ est donc strictement positive partout, et les fonctions φ_j définies par

$$\varphi_j(x) = \frac{\psi_j(x)}{\sum_{j=0}^N \psi_j(x)} \quad j = 1, \dots, N$$

répondent à la question.

Les partitions de l'unité permettent (en posant $f = \sum \varphi_j f$ au voisinage de K) de décomposer une fonction arbitraire en somme de fonctions ayant un "petit" support, la petitesse étant imposée à l'avance. L'exercice suivant montre un exemple typique d'utilisation des partitions de l'unité : le passage du local au global. Nous aurons à utiliser des arguments analogues en théorie des distributions.

Exercice 2.3.26. — Soit X un espace métrique compact (ou plus généralement un espace localement compact, voir l'exercice 2.3.12). Pour tout ouvert U de X , on note $C_0(U)$ l'espace vectoriel des fonctions continues sur U dont le support est compact, fonctions qui peuvent s'identifier, en les prolongeant par 0, aux fonctions définies sur X entier dont le support est un compact contenu dans U . On a ainsi $C_0(U) \subset C_0(V)$ pour $U \subset V$.

On appelle *mesure de Radon* (positive) μ sur U une forme linéaire sur $C_0(U)$ (notée le plus souvent $f \mapsto \int f d\mu$ ou $f \mapsto \int f(x) d\mu(x)$) qui est *positive*, c'est-à-dire que pour toute fonction positive f , on a $\int f d\mu \geq 0$.

Si $W \subset V$ sont des ouverts de X et si μ est une mesure de Radon sur V , sa restriction μ_W est la mesure de Radon sur W définie par $\int f d\mu_W = \int f d\mu$ pour $f \in C_0(W)$.

On se donne un ouvert U de X et $(U_i)_{i \in I}$ un recouvrement ouvert (fini ou non) de U .

(a) Soit μ et ν deux mesures de Radon sur U telles que pour tout i on ait $\mu_{U_i} = \nu_{U_i}$. Montrer que $\mu = \nu$ (il s'agit de démontrer que pour tout $f \in C_0(U)$ on a $\int f d\mu = \int f d\nu$, et on décomposera d'abord f , à l'aide d'une partition de l'unité, en une somme finie de fonctions à support dans l'un des U_i).

(b) Soient maintenant μ_i des mesures de Radon sur les U_i qui "se recollent", c'est-à-dire telles que l'on ait $\mu_i|_{U_i \cap U_j} = \mu_j|_{U_i \cap U_j}$ chaque fois que l'intersection $U_i \cap U_j$ est non vide. Montrer qu'il existe une et une seule mesure de Radon μ sur U telle que l'on ait $\mu|_{U_i} = \mu_i$ pour tout i (on décomposera une fonction quelconque $f \in C_0(U)$ en une somme finie de $f_i \in C_0(U_i)$).

et on posera $\int f d\mu = \sum \int f_i d\mu_i$ en vérifiant que le membre de droite ne dépend pas de la décomposition choisie).

Exercice 2.3.27 (théorème de Tietze-Urysohn). — Soit F un fermé d'un espace métrique X , et f une application continue de F dans $\mathbb{R}$. L'objet de l'exercice est de démontrer qu'il existe une application continue $\tilde{f}$ de X dans $\mathbb{R}$ qui prolonge f , c'est-à-dire telle que f et $\tilde{f}$ coïncident sur F .

(a) Montrer qu'il suffit de démontrer le théorème en supposant f à valeurs dans $] -1, 1[$ et en exigeant $\tilde{f}$ à valeurs dans $] -1, 1[$ (on prolongera d'abord $2/\pi \arctg f$).

(b) Montrer qu'il suffit de démontrer le théorème en supposant f à valeurs dans $] -1, 1[$ (ce que l'on supposera dans les questions suivantes) et en demandant seulement que $\tilde{f}$ soit à valeurs réelles (on multipliera un premier prolongement $\tilde{f}_1$ par une fonction égale à 1 sur F et à 0 sur $\{x \mid |g(x)| \geq 1\}$).

(c) Montrer qu'il existe une application continue g_0 de X dans $[-1/3, 1/3]$ telle que l'on ait $|f(x) - g_0(x)| \leq 2/3$ pour tout $x \in F$ (on posera $F_0 = \{x \in F \mid f(x) \leq -1/3\}$, $G_0 = \{x \in F \mid f(x) \geq 1/3\}$ et on utilisera la proposition 2.3.23)

(d) Montrer qu'il existe une application continue g_1 de X dans $[-\frac{1}{3}\frac{2}{3}, \frac{1}{3}\frac{2}{3}]$ telle que l'on ait $|f(x) - g_0(x) - g_1(x)| \leq (2/3)^2$ pour tout $x \in F$.

(e) Poursuivre la récurrence et conclure en posant $\tilde{f}(x) = \sum_{n=0}^{\infty} g_n(x)$.

2.4. Espaces vectoriels normés

Le corps des scalaires des espaces vectoriels considérés sera toujours égal à $\mathbb{R}$ ou $\mathbb{C}$.

Une *norme* sur un espace vectoriel F est une application $f \mapsto \|f\|$ à valeurs dans $[0, \infty[$ qui est *homogène* ($\|\lambda f\| = |\lambda| \|f\|$), *sous-linéaire* ($\|f + g\| \leq \|f\| + \|g\|$) et telle que l'on ait l'équivalence ($\|f\| = 0$) $\Leftrightarrow (f = 0)$. Un espace vectoriel normé est naturellement muni d'une structure d'espace métrique en posant $d(f, g) = \|f - g\|$.

Il est facile de vérifier que les applications qui définissent la structure algébrique de F : l'application $(f, g) \mapsto f + g$ de $F \times F$ dans F et l'application $(\lambda, f) \mapsto \lambda f$ de $\mathbb{C} \times F$ dans F sont continues. Bien entendu, comme dans tout espace métrique, la distance (et donc la norme) sont continues.

Deux normes $\|\cdot\|_1$ et $\|\cdot\|_2$ sur F sont *équivalentes* s'il existe une constante $C > 0$ telle que l'on ait $C^{-1} \|f\|_1 \leq \|f\|_2 \leq C \|f\|_1$ pour tout f de F . Dans ce cas, les distances associées sont uniformément équivalentes. Le lecteur n'ignore

pas que, sur un espace vectoriel F de dimension finie, toutes les normes sont équivalentes.

Si G est un sous-espace vectoriel de l'espace vectoriel normé F , il est naturellement muni, en restreignant à G l'application norme, d'une structure d'espace vectoriel normé. Si F_1 et F_2 sont deux espaces vectoriels normés, il en est de même de leur produit $F_1 \times F_2$ en posant $\|(f_1, f_2)\| = (\|f_1\|_{F_1}^2 + \|f_2\|_{F_2}^2)^{1/2}$.

Le théorème suivant (que l'on évitera d'utiliser pour des applications non linéaires!) a l'intérêt de ramener les problèmes de continuité à des majorations sur les normes. Une application linéaire sera souvent notée $f \mapsto Lf$ au lieu de $f \mapsto L(f)$.

Théorème 2.4.1. — Soient F et G des espaces vectoriels normés et L une application linéaire de F dans G . Les trois propriétés suivantes sont équivalentes.

- (a) L'application L est continue à l'origine.
- (b) L'application L est uniformément continue.
- (c) Il existe une constante $C \geq 0$ telle que l'on ait

$$\forall f \in F \quad \|Lf\|_G \leq C \|f\|_F.$$

Montrons que (a) implique (c). La continuité à l'origine, pour $\varepsilon = 1$ par exemple, fournit $\eta > 0$ tel que $\|f\|_F \leq \eta$ entraîne $\|Lf\|_G \leq 1$. Par homogénéité, on a donc $\|Lf\|_G \leq (1/\eta) \|f\|_F$ pour tout f .

L'implication (b) $\Rightarrow$ (a) étant triviale, il reste à montrer que (c) entraîne (b). Pour tout $\varepsilon > 0$, il suffit de choisir $\eta = \varepsilon/C$ et on a $(\|f - g\|_F \leq \eta) \Rightarrow (\|Lf - Lg\|_G \leq \varepsilon)$.

On dit qu'une partie $A \subset F$ est *bornée* si elle est contenue dans une boule. La propriété (c) ci-dessus peut être reformulée ainsi : l'image par L d'une partie bornée de F est une partie bornée de G .

Exercice 2.4.2. — Soient E, F, G trois espaces vectoriels normés et B une application bilinéaire de $E \times F$ dans G . Démontrer que B est continue si et seulement s'il existe $C > 0$ telle que $\|B(e, f)\|_G \leq C \|e\|_E \|f\|_F$.

Démontrer qu'une application bilinéaire (non nulle) n'est jamais uniformément continue.

L'espace vectoriel $\mathcal{L}(F, G)$ des applications linéaires continues de F dans G est muni d'une structure d'espace vectoriel normé en posant

$$\|L\|_{\mathcal{L}(F, G)} = \sup_{f \neq 0} \frac{\|Lf\|_G}{\|f\|_F} = \sup_{\|f\|_F \leq 1} \|Lf\|_G.$$

Le troisième membre montre qu'il s'agit de la norme de la convergence uniforme sur la boule unité de F .

En particulier, le *dual topologique* F' de F , l'espace des formes linéaires continues sur F qui n'est autre que $\mathcal{L}(F, \mathbb{C})$, est muni de la norme $\|L\|_{F'} = \sup_{\|f\|_F \leq 1} |L(f)|$.

Définition 2.4.3. — (a) On dit qu'un sous-ensemble A de F est total dans F si le sous-espace vectoriel formé des combinaisons linéaires d'éléments de A est partout dense.

(b) On dit qu'un espace vectoriel normé est séparable s'il existe un sous-ensemble dénombrable dense dans F . Il est équivalent de dire qu'il existe un sous-ensemble total dénombrable.

Établissons ce dernier point en montrant, si A est total et dénombrable, que l'ensemble B des combinaisons linéaires à coefficients rationnels d'éléments de A est dénombrable et partout dense. Pour tout N , l'application $(\lambda_1, \dots, \lambda_N, a_1, \dots, a_N) \mapsto \sum \lambda_j a_j$ de $\mathbb{Q}^N \times A^N$ dans F , dont l'ensemble source est dénombrable, a une image B_N qui est un ensemble dénombrable. L'ensemble B est réunion dénombrable d'ensembles dénombrables et est donc dénombrable.

Soit maintenant $f \in F$ et $\varepsilon > 0$. L'ensemble A étant total, il existe des $a_j \in A$ et des $\lambda_j \in \mathbb{R}$, $j = 1, \dots, N$ tels que $\|f - \sum_{j=1}^N \lambda_j a_j\| \leq \varepsilon/2$. En posant $M = \sup \|a_j\|$ et en choisissant pour chaque j un rationnel μ_j tel que $|\lambda_j - \mu_j| \leq \varepsilon/(2MN)$, on obtient bien $\|f - \sum_{j=1}^N \mu_j a_j\| \leq \varepsilon$.

Définition 2.4.4. — On appelle *espace de Banach* un espace vectoriel normé complet (pour la distance associée à la norme).

L'espace $\mathbb{R}^n$, et plus généralement tout espace vectoriel normé de dimension finie, est complet.

On peut appliquer aux espaces de Banach toutes les propriétés vues dans la section 2.2. En particulier, une application linéaire continue à valeur dans un espace de Banach G , définie sur un sous-espace vectoriel F_1 dense dans l'espace vectoriel normé F est uniformément continue d'après le théorème 2.4.1, et se prolonge donc en une application linéaire continue de F dans G .

Théorème 2.4.5. — (a) Soit F un espace de Banach et considérons une série d'éléments de F , de terme général u_n , normalement convergente, c'est-à-dire telle que la série numérique $\sum \|u_n\|_F$ soit convergente. Alors la série $\sum_0^\infty u_n$ est convergente, c'est-à-dire que ses sommes partielles $S_N = \sum_0^N u_n$ convergent, pour $N \rightarrow \infty$, vers un élément de F .

(b) Réciproquement, soit F un espace vectoriel normé dans lequel toute série normalement convergente est convergente. Alors, F est un espace de Banach.

Si u_n est normalement convergente, il est facile en effet de voir que la suite des sommes partielles est de Cauchy : par inégalité triangulaire, on a $\|S_{N+P} - S_N\| \leq \sum_{n=1}^{\infty} \|u_n\|$ et le membre de droite tend vers 0 pour $N \rightarrow \infty$.

Sous l'hypothèse (b), donnons nous une suite de Cauchy (f_n) et montrons qu'elle est convergente. Pour chaque entier p , il existe $n(p)$ tel que, si n_1 et n_2 sont $\geq n(p)$, on a $\|f_{n_1} - f_{n_2}\| \leq 2^{-p}$. Considérons la suite extraite $g_p = f_{n(p)}$ et posons $u_p = g_{p+1} - g_p$. On a donc $\|u_p\| \leq 2^{-p}$ et la série de terme général u_p est convergente. D'autre part, on a

$$g_p = g_0 + \sum_{k=0}^{p-1} u_k$$

et la suite g_p converge vers une limite $g \in F$.

La fin de la démonstration est valable dans un espace métrique quelconque : si une suite de Cauchy possède une sous-suite convergente, elle est convergente. Pour $\varepsilon > 0$, il existe N tel que, si n_1 et n_2 sont $\geq N$, on a $\|f_{n_1} - f_{n_2}\| \leq \varepsilon$. Pour $n \geq N$ et pour p assez grand, on a $\|f_n - g_p\| \leq \varepsilon$ et donc, en faisant tendre p vers l'infini, $\|f_n - g\| \leq \varepsilon$. Cela montre que la suite (f_n) converge vers g .

Exercice 2.4.6. — (a) Démontrer qu'une série normalement convergente possède la propriété suivante (on dit que la famille (u_n) est sommable) : il existe $S \in F$ tel que, pour tout $\varepsilon > 0$, on peut trouver un sous-ensemble fini $J \subset \mathbb{N}$ tel que, pour tout ensemble fini K contenant J , on ait $\|S - \sum_{n \in K} u_n\| \leq \varepsilon$. Cette propriété est indépendante de l'ordre dans lequel sont rangés les u_n et en particulier, si b est une bijection de $\mathbb{N}$ sur lui-même, la série $u_{b(n)}$ est également convergente et de même somme.

(b) Lorsque F est de dimension finie, montrer que (u_n) est sommable si et seulement si la série est normalement convergente (le lecteur utilisera ou redémontrera le fait que le résultat est vrai pour les séries d'éléments de $\mathbb{R}$). Il n'en est pas de même en dimension infinie comme le montrera l'exemple des séries orthogonales dans un espace de Hilbert.

Dans un espace normé de dimension finie, les ensembles fermés et bornés, et en particulier la boule unité fermée, sont compacts. Un théorème de F. Riesz, que nous proposons en exercice, montre que cette propriété caractérise les espaces de dimension finie.

Exercice 2.4.7. — Soit F un espace vectoriel normé dont la boule unité fermée $\overline{B(0,1)}$ est compacte.

(a) Montrer qu'il existe un ensemble fini f_j , $j = 1, \dots, N$ tel que les boules $B(f_j, 1/2)$ recouvrent $\overline{B(0,1)}$.

(b) Soit f un élément quelconque de F . Montrer qu'il peut se décomposer sous la forme

$$f = \sum_j \lambda_j^{(1)} f_j + r_1$$

avec $|\lambda_j^{(1)}| \leq \|f\|$ et $\|r_1\| \leq \|f\|/2$ (un seul des λ_j est en fait non nul).

(c) Construire par récurrence des $\lambda_j^{(n)}$ et r_n tels que

$$r_{n-1} = \sum_j \lambda_j^{(n)} f_j + r_n$$

avec $|\lambda_j^{(n)}| \leq \|f\|/2^{n-1}$ et $\|r_n\| \leq \|f\|/2^n$.

(d) En déduire que les f_j forment un système de générateurs de F .

Remarque 2.4.8. — Signalons ici sans démonstration quelques théorèmes de la théorie des espaces de Banach que nous n'aurons pas à utiliser dans la suite du cours, mais qui n'en sont pas moins importants.

Le lecteur trouvera dans l'appendice C le théorème de Banach-Steinhaus démontré dans le cadre plus général des espaces de Fréchet. Un autre théorème qui repose également sur la théorie des espaces de Baire, et qui est également valable pour les espaces de Fréchet, est le *théorème des isomorphismes de Banach* dont voici l'énoncé. Soient F et G des espaces de Banach et L une application linéaire continue *bijective* de F sur G . Alors l'application réciproque L^{-1} est continue de G dans F .

Un autre théorème important est le théorème de Hahn-Banach. Soit F un espace vectoriel normé, F_1 un sous-espace vectoriel de F , et L une forme linéaire continue sur F_1 vérifiant donc $|L(f)| \leq C\|f\|$ pour $f \in F_1$. Il existe alors une forme linéaire continue $\tilde{L}$ sur F lui-même, qui prolonge L et qui vérifie de plus $|\tilde{L}(f)| \leq C\|f\|$ pour $f \in F$ avec la même constante C . Nous n'aurons à l'utiliser que pour les espaces de Hilbert, la démonstration étant nettement plus facile dans ce cas.

2.5. Espaces de Hilbert

Un *produit scalaire* sur un espace vectoriel F sur $\mathbb{C}$ est une application $(f, g) \mapsto (f|g)$ de $F \times F$ dans $\mathbb{C}$ qui est *sesquilinéaire* c'est-à-dire linéaire par rapport à la première variable et antilinéaire par rapport à la seconde ($(f|\lambda g) = \overline{\lambda}(f|g)$), qui possède la *symétrie hermitienne* ($(g|f) = \overline{(f|g)}$) et qui est *définie positive* c'est-à-dire que $(f|f) > 0$ pour $f \neq 0$. Lorsque E est un espace vectoriel sur $\mathbb{R}$ les deux premières propriétés doivent être remplacées par la bilinéarité et la symétrie.

Un espace vectoriel muni d'un produit scalaire est appelé *espace préhilbertien*. Le lecteur n'ignore pas qu'il est canoniquement muni d'une structure

d'espace vectoriel normé en posant

$$\|f\| = (f | f)^{1/2},$$

et que l'on a l'inégalité de Cauchy-Schwarz

$$|(f | g)| \leq \|f\| \|g\|.$$

On dit que deux éléments de F sont orthogonaux si leur produit scalaire est nul. Si f et g sont orthogonaux, on a la relation de Pythagore : $\|f + g\|^2 = \|f\|^2 + \|g\|^2$, qui s'étend sans difficulté aux sommes finies d'éléments deux à deux orthogonaux.

Définition 2.5.1. — On appelle *espace de Hilbert* un espace vectoriel muni d'un produit scalaire et complet pour la norme associée.

Un espace de Hilbert est en particulier un espace de Banach et une série normalement convergente y est donc convergente. On a également le résultat plus spécifique suivant, où la condition portant sur les normes est moins forte.

Théorème 2.5.2. — Soit H un espace de Hilbert, et (u_n) une suite d'éléments de H deux à deux orthogonaux. Pour que la série $\sum_0^\infty u_n$ soit convergente, il faut et il suffit que la série $\sum_0^\infty \|u_n\|^2$ soit convergente. On a alors $\|\sum u_n\|^2 = \sum \|u_n\|^2$.

La condition est nécessaire. En effet, en posant $S_N = \sum_0^N u_n$, on a d'après le théorème de Pythagore

$$\sum_0^N \|u_n\|^2 = \|S_N\|^2.$$

En appelant S la somme de la série, le membre de droite converge vers $\|S\|^2$ par continuité de la norme, ce qui entraîne que la série numérique $\sum_0^\infty \|u_n\|^2$ converge.

Réciproquement, si cette série converge, on a

$$\|S_{N+P} - S_N\|^2 = \sum_{N+1}^{N+P} \|u_n\|^2 \leq \sum_{N+1}^\infty \|u_n\|^2, \quad (2.6)$$

Le membre de droite de (2.6) tend vers 0 avec N , ce qui assure que la suite S_N est de Cauchy dans H et est donc convergente.

Exercice 2.5.3. — (a) Montrer plus généralement, pour une suite u_n non nécessairement orthogonale, que la série $\sum_0^\infty u_n$ est convergente si la série double

$$\sum_{m=0}^\infty \sum_{n=0}^\infty |(u_m | u_n)|$$

est convergente. La réciproque est-elle vraie?

(b) Montrer que, sous l'hypothèse ci-dessus, la famille (u_n) est sommable (voir l'exercice 2.4.6).

Exercice 2.5.4. — Soit (u_n) , $n = 1, \dots, N$ une suite finie d'éléments d'un espace de Hilbert réel H . On suppose que pour tout choix de $\varepsilon_n \in \{-1, 1\}$, $n = 1, \dots, N$, on a $\|\sum_1^N \varepsilon_n u_n\| \leq C$. Démontrer que l'on a $\sum_1^N \|u_n\|^2 \leq C^2$ (On calculera la somme des 2^N quantités $\|\sum_1^N \varepsilon_n u_n\|^2$). Étendre ce résultat au cas des séries.

Le théorème suivant joue un rôle très important : tout comme le théorème du point fixe, il affirme l'existence d'un élément vérifiant une (in)égalité. Il sert notamment, directement ou par l'intermédiaire du théorème 2.5.10 ci-dessous, à démontrer l'existence de solutions pour des équations ou inéquations fonctionnelles.

Théorème 2.5.5 (projection sur un convexe fermé)

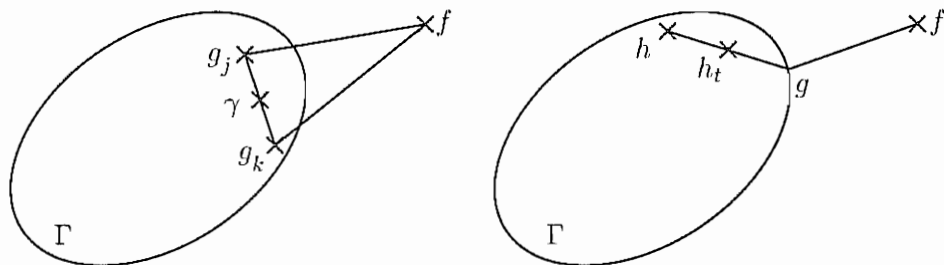
(a) Soient H un espace de Hilbert et Γ une partie convexe non vide et fermée de H . Pour tout $f \in H$ il existe un unique point de Γ (appelé projection de f sur Γ) dont la distance à f soit minimum.

(b) La projection de f sur Γ est l'unique point $g \in \Gamma$ tel que l'on ait $\operatorname{Re} \langle f - g | h - g \rangle \leq 0$ pour tout $h \in \Gamma$.

Le lecteur démontrera facilement, en développant les carrés scalaire du membre de droite, l'identité de la médiane

$$\|u\|^2 + \|v\|^2 = 2 \left\| \frac{u+v}{2} \right\|^2 + \frac{1}{2} \|u-v\|^2 \quad (2.7)$$

qui remonte (au moins) à Euclide.



Posons $d = \inf \{\|f - g\| \mid g \in \Gamma\}$. S'il existait g_1 et g_2 distincts réalisant cette borne inférieure, leur milieu γ appartiendrait aussi à Γ et d'après (2.7) on aurait $2\|f - \gamma\|^2 = 2d^2 - \|g_1 - g_2\|^2/2 < 2d^2$ ce qui est impossible. Cela établit l'unicité de la projection.

Par définition de la borne inférieure, il existe une suite g_j d'éléments de Γ telle que $\|f - g_j\|$ tende vers d . En introduisant le milieu γ de g_j et g_k , on déduit de (2.7)

$$\frac{1}{2} \|g_j - g_k\|^2 = \|f - g_j\|^2 + \|f - g_k\|^2 - 2 \|\gamma - f\|^2 \leq \|f - g_j\|^2 + \|f - g_k\|^2 - 2d^2.$$

Le membre de droite tend vers 0 lorsque j et k tendent vers l'infini. La suite g_j est donc de Cauchy. L'ensemble Γ étant fermé et donc complet, elle converge vers un point $g \in \Gamma$. Par continuité, on a $\|f - g\| = \lim \|f - g_j\| = d$ ce qui achève la démonstration du point (a).

Soit maintenant $h \in \Gamma$. Pour $0 < t < 1$, le point $h_t = g + t(h - g)$ appartient aussi à Γ . On a donc

$$\|g - f\|^2 \leq \|h_t - f\|^2 = \|g - f\|^2 + t^2 \|h - g\|^2 + 2t \operatorname{Re}(g - f | h - g). \quad (2.8)$$

En faisant tendre t vers 0, on voit que le coefficient de t doit donc être positif ou nul, ce qui établit l'inégalité $\operatorname{Re}(f - g | h - g) \leq 0$. Réciproquement, si un point $g \in \Gamma$ vérifie cette inégalité pour tout $h \in \Gamma$, l'égalité de droite dans (2.8), pour $t = 1$, montre que la distance de g à f est inférieure à celle de h à f , ce qui caractérise la projection d'après la partie (a).

Exercice 2.5.6. — On note $f \mapsto p(f)$ la projection sur le convexe fermé Γ . Montrer que p est lipschitzienne de rapport 1 (on pourra considérer $\operatorname{Re}(p(f) - p(g) | f - g)$).

Corollaire 2.5.7 (supplémentaire orthogonal). — Soient H un espace de Hilbert et F un sous-espace vectoriel fermé de H . Posons

$$F^\perp = \{g \in H \mid \forall f \in F, (f | g) = 0\}.$$

L'ensemble $F^\perp$ est un sous-espace vectoriel fermé de H appelé *supplémentaire orthogonal* de F . Tout élément $h \in H$ se décompose de manière unique sous la forme

$$h = f + g \quad f \in F \quad g \in F^\perp. \quad (2.9)$$

En outre, les éléments f et g de la décomposition sont les projections de h sur F et $F^\perp$ respectivement.

Soit en effet $h \in H$ et notons f sa projection sur F . Pour tout élément $f' \in F$ et tout scalaire λ , le point $f + \lambda f'$ appartient à F et on a donc d'après le théorème 2.5.5(b) $\operatorname{Re}(h - f | \lambda f') \leq 0$. Il en résulte que l'on a $(h - f | f') = 0$ pour tout $f' \in F$, et que $g = h - f \in F^\perp$. Cela prouve l'existence de la décomposition.

Si'il existait une autre décomposition $h = f_1 + g_1$ du même type, le vecteur $f - f_1 = g_1 - g$ appartiendrait à F et à $F^\perp$, il serait orthogonal à lui-même et donc nul.

Corollaire 2.5.8 (critère de totalité). — *Soit H un espace de Hilbert et $A \subset H$. Pour que A soit total, il faut et il suffit que le seul vecteur orthogonal à tous les éléments de A soit le vecteur nul.*

Soit F le sous-espace vectoriel constitué des combinaisons linéaires des éléments de A et soit $\overline{F}$ son adhérence. Par linéarité et passage à la limite, il est clair qu'un élément $g \in H$ appartient à $\overline{F}^\perp$ si et seulement si il est orthogonal à tous les éléments de A .

Si A est total, ce qui signifie $\overline{F} = H$, l'unicité de la décomposition (2.9) impose $\overline{F}^\perp = \{0\}$. Réciproquement, si $\overline{F}^\perp = \{0\}$, l'existence de la décomposition (2.9) implique $\overline{F} = H$, ce qui achève la démonstration.

Exercice 2.5.9. — Soient H et K deux espaces de Hilbert, F un sous-espace vectoriel de H , et L une application linéaire continue de F dans K . Montrer qu'il existe un prolongement de L en une application linéaire continue $\tilde{L}$ de H dans K qui vérifie $\|\tilde{L}\|_{\mathcal{L}(H,K)} = \|L\|_{\mathcal{L}(F,K)}$ (on prolongera d'abord L en L_1 par continuité à $\overline{F}$, puis on posera $\tilde{L} = L_1 \circ P$ où P est le projecteur orthogonal sur $\overline{F}$).

L'analogie de ce résultat est vrai, mais plus difficile à démontrer, lorsque H est un espace de Banach et K le corps des scalaires. C'est le théorème de Hahn-Banach (voir la remarque 2.4.8).

Théorème 2.5.10 (F. Riesz). — *Soit H un espace de Hilbert. À tout élément $h \in H$, on peut faire correspondre la forme linéaire continue L_h définie par $L_h(f) = (f|h)$. Réciproquement, étant donnée une forme linéaire continue L sur H , il existe un et un seul $h \in H$ tel que l'on ait $L = L_h$.*

Pour $f \in H$, on a

$$|L_h(f)| = |(f|h)| \leq \|h\|_H \|f\|_H.$$

Cela prouve que L_h est continue et que $\|L_h\|_{H'} \leq \|h\|_H$ (en considérant $L_h(h)$, on voit facilement que l'on a en fait égalité).

Soit maintenant L une forme linéaire continue non identiquement nulle (c'est-à-dire différente de L_0). Le sous-espace vectoriel $F = L^{-1}(0)$ est fermé et distinct de H et $F^\perp$ n'est donc pas réduit à $\{0\}$. Soit g un élément non nul de $F^\perp$. Il n'appartient donc pas à F et le scalaire $\lambda = L(g)$ est non nul. Pour tout $f \in H$, on peut poser

$$f = \frac{L(f)}{L(g)} g + \left(f - \frac{L(f)}{L(g)} g \right) = f_1 + f_2$$

et on remarque que le second terme, qui vérifie $L(f_2) = 0$ appartient à F tandis que le premier appartient à $F^\perp$. En faisant le produit scalaire avec g , on obtient

$$(f | g) = \frac{L(f)}{L(g)} \|g\|^2 + 0.$$

Il suffit maintenant de poser $h = \frac{\overline{L(g)}}{\|g\|^2} g$ pour obtenir $L(f) = (f | h)$ quel que soit f . Cela montre donc que L est égale à L_h .

On obtient facilement l'unicité : si $L_{h_1} = L_{h_2}$, le vecteur $h_1 - h_2$ est orthogonal à tout élément de H et est donc nul.

Remarque 2.5.11. — L'application $h \mapsto L_h$ est une bijection isométrique de H sur H' . Cette application est linéaire dans le cas d'un espace de Hilbert réel. Par contre, cette application est *antilinéaire* dans le cas complexe ($L_{\lambda h} = \overline{\lambda} L_h$). Dans les notations de Dirac de la mécanique quantique, on dit que L_h est le "bra" associé au "ket" h (en mécanique quantique, on convient le plus souvent que le produit scalaire est linéaire par rapport à la seconde variable et antilinéaire par rapport à la première ; quelle que soit la convention, l'application $h \mapsto L_h$ est antilinéaire).

L'application $h \mapsto L_h$ dépend explicitement du produit scalaire de H . Si on le remplace par un autre produit scalaire fournissant une norme équivalente, l'espace reste un espace de Hilbert, mais le nouvel isomorphisme de H sur H' est différent de l'ancien. Ce n'est que dans la mesure où le produit scalaire est bien fixé que l'on peut considérer que les espaces H et H' sont canoniquement (anti-) isomorphes.

Définition 2.5.12. — Soit H un espace de Hilbert séparable. On appelle base hilbertienne (ou base orthonormale) de H une suite finie ou infinie (e_j) , $j = 1, 2, \dots$ qui constitue un système total dans H et qui vérifie les relations d'orthonormalité : $(e_j | e_k) = \delta_{jk}$.

On a utilisé ci-dessus le symbole de Kronecker δ_{jk} égal à 1 si $j = k$ et à 0 sinon. Cette définition n'implique nullement que les e_j constituent une base de H au sens algébrique du terme : on demande que l'espace des combinaisons linéaires des e_j soit partout dense au lieu de demander qu'il coïncide avec H . En fait, un espace de Hilbert (ou de Banach ou même de Fréchet) de dimension infinie n'admet jamais de base algébrique dénombrable, ce que le lecteur pourra démontrer à titre d'exercice après avoir lu l'appendice B.

Théorème 2.5.13. — Dans tout espace de Hilbert séparable, il existe des bases hilbertiennes.

A partir d'un ensemble total $\{a_1, a_2, \dots\}$ fini ou dénombrable, nous allons construire une base hilbertienne par le procédé dit d'orthonormalisation de

Schmidt. On commence par supprimer de la suite a_j tout vecteur qui est combinaison linéaire des précédents. On obtient ainsi une suite b_j qui est telle que l'espace vectoriel E_n engendré par $b_1, \dots, b_n$ soit exactement de dimension n . L'espace des combinaisons linéaires des a_j est exactement la réunion des E_n et l'hypothèse assure qu'il est partout dense.

Posons $e_1 = b_1 / \|b_1\|$, et montrons par récurrence que l'on peut construire une suite orthonormale e_j telle que pour tout n , les $e_1, \dots, e_n$ forment une base de E_n . L'espace des combinaisons linéaires des e_j sera alors également la réunion des E_n et le système des e_j sera bien total.

Les e_j étant supposés construits jusqu'au rang n , posons

$$f_{n+1} = b_{n+1} - \sum_{j=1}^n (b_{n+1} | e_j) e_j.$$

Il est immédiat que ce vecteur est orthogonal aux e_j pour $j = 1, \dots, n$. D'autre part, il appartient à E_{n+1} mais pas à E_n et il en résulte que $(e_1, \dots, e_n, f_{n+1})$ est une base algébrique de E_{n+1} . Il suffit de poser $e_{n+1} = f_{n+1} / \|f_{n+1}\|$ pour en obtenir une base orthonormale, ce qui achève la récurrence et la démonstration.

Théorème 2.5.14. — *Soit H un espace de Hilbert séparable, et (e_j) , $j = 1, 2, \dots$ une base hilbertienne de H .*

(a) *Tout élément $f \in H$ peut se décomposer de façon unique sous forme d'une série convergente dans H*

$$f = \sum_j c_j(f) e_j \quad c_j(f) \in \mathbb{C}.$$

Les composantes $c_j(f)$ sont données par

$$c_j(f) = (f | e_j), \quad (2.10)$$

et vérifient

$$\|f\|^2 = \sum_j |c_j(f)|^2 \quad (\text{Bessel-Parseval}) \quad (2.11)$$

(b) *Réciproquement, étant donnés des scalaires γ_j vérifiant $\sum_j |\gamma_j|^2 < \infty$, la série $\sum_j \gamma_j e_j$ converge dans H et sa somme f vérifie $c_j(f) = \gamma_j$ pour tout j .*

La convergence de la série figurant dans (b) est un cas particulier du théorème 2.5.2, les $\gamma_j e_j$ étant orthogonaux et de norme $|\gamma_j|$.

On obtient facilement le reste du point (b) et du même coup l'unicité de la décomposition et la nécessité des formules (2.10) : si la série $\sum \gamma_j e_j$ converge vers f , par continuité du produit scalaire, on a $(\sum_{j=1}^N \gamma_j e_j | e_k) \rightarrow (f | e_k)$ et le membre de gauche est égal à γ_k dès que N est supérieur à k .

Soit enfin $f \in H$, posons $c_j(f) = (f | e_j)$ et $f_N = \sum_{j=1}^N c_j(f) e_j$. On a

$$(f | f_N) = \sum_{j=1}^N \overline{c_j(f)} (f | e_j) = \sum_{j=1}^N |c_j(f)|^2.$$

Le membre de droite est égal, d'après le théorème de Pythagore, à $\|f_N\|^2$ et on a donc $\|f_N\|^2 = (f | f_N) \leq \|f\| \|f_N\|$. Cela prouve que $\|f_N\|^2$ est majoré pour tout N par $\|f\|^2$, et donc que la série $\sum |c_j(f)|^2$ est convergente. Une nouvelle application du théorème 2.5.2 assure que la série $\sum c_j(f) e_j$ converge vers un élément $g \in H$. Pour tout j , on a $(g - f | e_j) = c_j(f) - c_j(f) = 0$, et l'élément $g - f$ qui est orthogonal à un système total est nécessairement nul. Cela achève la démonstration.

2.5.15. Convergence faible. — Dans un espace de Hilbert H , on dit qu'une suite u_j converge faiblement vers u (ce que l'on note $u_j \rightarrow u$) si pour tout $v \in H$, on a $(u_j | v) \rightarrow (u | v)$.

Cette notion de convergence correspond en fait à une topologie dite faible, mais celle-ci ne peut pas être définie par une métrique, et on peut démontrer que la boule unité fermée de H est compacte pour la topologie faible. Nous nous contenterons de l'énoncé suivant (compacité séquentielle).

Théorème 2.5.16. — Soit u_j une suite d'éléments d'un espace de Hilbert séparable H vérifiant $\|u_j\| \leq M$ pour tout j . On peut alors en extraire une sous-suite v_k qui converge faiblement vers un élément v vérifiant $\|v\| \leq M$.

Soit e_n une base hilbertienne de H . La suite numérique $(u_j | e_1)$ étant bornée, on peut extraire de (u_j) une première sous-suite, notée $u_j^{(1)}$, telle que $(u_j^{(1)} | e_1)$ converge vers une limite $\alpha_1 \in \mathbb{C}$. On peut de même extraire de $u_j^{(1)}$ une sous-suite $u_j^{(2)}$ telle que $(u_j^{(2)} | e_2)$ converge vers une limite α_2 . On construit ainsi par récurrence une suite de suites $u_j^{(n)}$, et on définit enfin la suite v_k par $v_k = u_k^{(k)}$. C'est une suite extraite de la suite u_j et pour tout vecteur de la base hilbertienne, on a $(v_k | e_n) \rightarrow \alpha_n$.

Pour N fixé et pour $k \rightarrow \infty$, on a $(v_k | \sum_1^N \alpha_n e_n) \rightarrow \sum_1^N |\alpha_n|^2$. Le membre de gauche étant majoré d'après Cauchy-Schwarz par $M (\sum_1^N |\alpha_n|^2)^{1/2}$, il en résulte que $\sum_1^N |\alpha_n|^2 \leq M^2$ pour tout N , et donc que $\sum_1^\infty |\alpha_n|^2 \leq M^2$.

Nous pouvons maintenant poser $v = \sum \alpha_n e_n$, on a $\|v\| \leq M$ et il reste à prouver que $v - v_k \rightarrow 0$. Donnons nous un élément quelconque $w = \sum w_n e_n$ de H et $\varepsilon > 0$. On a pour tout N

$$(v - v_k | w) = \left(v - v_k \left| \sum_1^N w_n e_n \right. \right) + \left(v - v_k \left| \sum_{N+1}^\infty w_n e_n \right. \right).$$

Le second terme est majoré en module par $2M \left\| \sum_{N+1}^\infty w_n e_n \right\|$ que l'on peut rendre inférieur à $\varepsilon/2$ en choisissant N assez grand. Cet entier N étant fixé, le premier terme du membre de droite tend vers 0 avec k et est $\leq \varepsilon/2$ pour k assez grand. On a donc montré que $(v - v_k | w) \rightarrow 0$ pour tout $w \in H$, ce qui est le résultat voulu.

On prendra garde au fait que la norme et le produit scalaire ne passent pas à la limite faible. Par exemple, la suite e_n elle-même tend faiblement vers 0 alors que $\|e_n\|$ reste égal

à 1. On a toutefois le résultat suivant, que le lecteur pourra démontrer après avoir lu le théorème C.3.1 : si $u_j \rightarrow u$ et $v_j \rightarrow v$, on a $(u_j | v_j) \rightarrow (u | v)$.

2.6. Espaces fonctionnels classiques

2.6.1. Espaces de fonctions bornées. — Si X est un ensemble, l'espace $\mathcal{F}_b(X, \mathbb{C})$ des applications bornées de X dans $\mathbb{C}$ est un espace de Banach lorsqu'on le munit de la norme de la convergence uniforme : $\|f\| = \sup_{x \in X} |f(x)|$.

En effet, étant donnée une suite de Cauchy f_j , on a pour chaque point x l'inégalité $|f_j(x) - f_k(x)| \leq \|f_j - f_k\|$. La suite $f_j(x)$ est donc de Cauchy, et converge vers un élément de $\mathbb{C}$ que l'on notera $f(x)$. Il reste à montrer que f_j converge vers f uniformément. Quels que soient j et x , on a $|f(x) - f_j(x)| = \lim_{k \rightarrow \infty} |f_k(x) - f_j(x)| \leq \sup_{k \geq j} \|f_j - f_k\|$. On a donc $\|f - f_j\| \leq \sup_{k \geq j} \|f_j - f_k\|$, et le membre de droite tend vers 0 lorsque j tend vers l'infini.

Si X est un espace métrique, l'espace $C_b(X)$ des fonctions continues bornées sur X est un sous-espace du précédent que l'on munit de la même norme. Une limite uniforme de fonctions continues étant continue, cela signifie que $C_b(X)$ est un sous-espace fermé de $\mathcal{F}_b(X, \mathbb{C})$. L'espace $C_b(X)$ est donc complet et est un espace de Banach. Dans le cas où X est compact, l'espace $C_b(X)$ coïncide avec l'espace $C(X)$ de toutes les fonctions continues sur X .

2.6.2. Espaces de suites. — Ces espaces sont analogues à ceux que nous étudierons sur un ouvert de $\mathbb{R}^n$, la sommation des séries remplaçant l'intégration des fonctions.

Exercice 2.6.3. — On appelle l^1 l'espaces des suites $u = (u_n)_{n \in \mathbb{N}}$ telles que $\sum_0^\infty |u_n|$ soit fini. Montrer que l^1 est un espace de Banach si on le munit de la norme $\|u\|_1 = \sum_0^\infty |u_n|$. Démontrer que l'ensemble des suites qui n'ont qu'un nombre fini de composantes non nulles est dense dans l^1 .

Exercice 2.6.4. — On appelle l^2 l'espaces des suites $u = (u_n)_{n \in \mathbb{N}}$ telles que $\sum_0^\infty |u_n|^2$ soit fini. Montrer que l^2 est un espace de Hilbert si on le munit du produit scalaire $(u | v) = \sum_{n=0}^\infty u_n \overline{v_n}$ et donc de la norme $\|u\|_2 = \{\sum_{n=0}^\infty |u_n|^2\}^{1/2}$. Démontrer que l'ensemble des suites qui n'ont qu'un nombre fini de composantes non nulles est dense dans l^2 .

Démontrer que pour tout espace de Hilbert séparable de dimension infinie H , il existe une application linéaire de H dans l^2 qui est une isométrie bijective (cela montre que tous les espaces de Hilbert séparables de dimension infinie sont isomorphes).

Exercice 2.6.5. — On appelle l^∞ l'espace des suites $u = (u_n)_{n \in \mathbb{I}}$ qui sont bornées (c'est-à-dire l'espace $\mathcal{F}_b(\mathbb{N}, \mathbb{C})$). Montrer que l^∞ est un espace de Banach si on le munit de la norme $\|u\|_\infty = \sup_{n \in \mathbb{I}} |u_n|$. Démontrer que l'ensemble des suites qui n'ont qu'un nombre fini de composantes non nulles n'est pas dense dans l^∞ , et que l'adhérence de cet ensemble est constitué des suites qui tendent vers 0 à l'infini.

Exercice 2.6.6. — Démontrer que l'on a $l^1 \subset l^2 \subset l^\infty$, chaque inclusion étant stricte, et que l'on a $\|u\|_2 \leq \|u\|_1$ pour $u \in l^1$ et $\|u\|_\infty \leq \|u\|_2$ pour $u \in l^2$.

2.6.7. La relation d'équivalence $f = g$ p.p. — En étudiant les espaces $\mathcal{L}^1$, $\mathcal{L}^2$ et $\mathcal{L}^\infty$, nous avons vu que l'on pouvait les munir naturellement de quantités qui sont presque des normes, le seul problème étant que la nullité de la "norme" d'une fonction n'entraîne la nullité de la fonction que presque partout. Pour pouvoir profiter des ressources de l'analyse fonctionnelle, nous allons être amenés à travailler sur des classes d'équivalence de fonctions plutôt que sur les fonctions elles-mêmes.

Sur l'ensemble $\mathcal{F}(\mathbb{R}^n, \mathbb{C})$ des applications de $\mathbb{R}^n$ dans $\mathbb{C}$, on voit facilement que la relation $f \sim g$ définie par $f = g$ p.p. est une relation d'équivalence. L'ensemble quotient $\mathcal{F}(\mathbb{R}^n, \mathbb{C})/(\sim)$ est par définition l'ensemble des classes d'équivalence. Si $f \in \mathcal{F}(\mathbb{R}^n, \mathbb{C})/(\sim)$, au lieu de dire qu'une (vraie) fonction $\tilde{f}$ appartient à f , on préfère dire que $\tilde{f}$ est un représentant de f .

Il est important de faire la liste des propriétés (ou des relations) qui ont un sens pour les éléments $\tilde{f}$ de $\mathcal{F}(\mathbb{R}^n, \mathbb{C})$ et qui restent inchangées lorsqu'on remplace $\tilde{f}$ par un élément équivalent. Elles deviennent alors des propriétés de la classe d'équivalence (on dit qu'elles passent au quotient).

Les opérations algébriques. — Si f et g appartiennent à $\mathcal{F}(\mathbb{R}^n, \mathbb{C})/(\sim)$, et si on en prend des représentants $\tilde{f}$ et $\tilde{g}$, on voit immédiatement que les fonctions $\tilde{f} + \tilde{g}$, $\lambda \tilde{f}$, $\tilde{f}\tilde{g}$, $\Phi \circ \tilde{f}$ ne sont modifiées que sur un ensemble de mesure nulle si on change les représentants de f et g (on a désigné par λ un scalaire et par Φ une application définie sur $\mathbb{C}$). La classe d'équivalence du résultat ne dépend donc que de f et g et sera notée $f + g$, λf , fg ou $\Phi \circ f$.

Il est immédiat de vérifier que $\mathcal{F}(\mathbb{R}^n, \mathbb{C})/(\sim)$ est un espace vectoriel une fois muni des deux premières opérations, et une algèbre si on y adjoint la troisième.

La sommabilité et l'intégrale. — Comme nous l'avons vu, pour f appartenant à $\mathcal{F}(\mathbb{R}^n, \mathbb{C})/(\sim)$, le fait qu'un représentant $\tilde{f}$ soit sommable, et dans ce cas la valeur de l'intégrale, ne dépendent pas du représentant choisi. On dira alors que f est sommable, l'intégrale de f sera aussi appelée l'intégrale de f , et on la notera $\int f(x) dx$.

On voit de même que les expressions f est de carré sommable, f est essentiellement bornée, sont bien définies.

La convergence presque partout. — Soit f_j une suite d'éléments de l'espace quotient $\mathcal{F}(\mathbb{R}^n, \mathbb{C})/(\sim)$ telle que, en choisissant un représentant $\tilde{f}_j$ de chaque f_j , la suite $\tilde{f}_j$ converge presque partout vers un représentant $\tilde{f}$ de f . Alors, pour tout autre choix de représentants $\overline{f}_j$ et $\overline{f}$, la suite $\overline{f}_j$ converge presque partout vers $\overline{f}$.

En effet, posons $A_j = \{x \mid \overline{f}_j(x) \neq \tilde{f}_j(x)\}$, $B = \{x \mid \overline{f}(x) \neq \tilde{f}(x)\}$ et notons C l'ensemble des x où $\tilde{f}_j(x)$ ne converge pas vers $\tilde{f}(x)$. En dehors de l'ensemble $(\bigcup_j A_j) \cup B \cup C$, qui est de mesure nulle, la suite $\overline{f}_j(x)$ converge vers $\overline{f}(x)$.

On dira dans ce cas que la suite f_j converge presque partout vers f .

Dans la pratique, on ne s'astreint pas à distinguer (avec des tildes par exemple) les représentants d'un élément f de $\mathcal{F}(\mathbb{R}^n, \mathbb{C})/(\sim)$ comme nous l'avons fait de manière (peut-être trop) détaillée ci-dessus. On n'hésite pas à désigner par la même lettre l'élément f et l'un quelconque de ses représentants. Cela ne peut avoir aucune conséquence fâcheuse, à condition de ne manipuler que des propriétés et des quantités qui passent au quotient.

Une dernière remarque : une fonction continue sur $\mathbb{R}^n$ définit canoniquement un élément de $\mathcal{F}(\mathbb{R}^n, \mathbb{C})/(\sim)$ (sa classe d'équivalence), et cette application est injective (deux fonctions continues qui sont égales presque partout coïncident). Cela permet d'identifier $C(\mathbb{R}^n)$ à un sous-espace de $\mathcal{F}(\mathbb{R}^n, \mathbb{C})/(\sim)$.

Théorème et Définition 2.6.8 (Espace $L^1(\mathbb{R}^n)$). — *C'est l'espace des classes de fonctions sommables pour la relation d'équivalence $f = g$ p.p.*

(a) **(Fischer-Riesz)** *Muni de la norme*

$$\|f\|_{L^1} = \int_{\mathbb{R}^n} |f(x)| \, dx, \quad (2.12)$$

$L^1(\mathbb{R}^n)$ est un espace de Banach.

(b) *L'espace $C_0(\mathbb{R}^n)$ des fonctions continues à support compact est partout dense dans $L^1(\mathbb{R}^n)$.*

Le membre de droite de (2.12) a bien un sens qui ne dépend que de l'élément $f \in L^1$. Si $\|f\|_{L^1} = 0$, un représentant quelconque de f doit être nul presque partout, ce qui signifie précisément que $f = 0$. Nous avons déjà vu les autres propriétés assurant que $\|\cdot\|_{L^1}$ est une norme. D'autre part, la partie (b) du théorème ne fait que reformuler le théorème 1.5.2. Il nous reste à montrer que L^1 est complet. D'après le théorème 2.4.5, il suffit de montrer que toute série normalement convergente dans L^1 est convergente.

Soit donc (u_n) une série normalement convergente. Nous choisirons pour chaque n un représentant de u_n que nous noterons encore u_n . Posons $h(x) = \sum |u_n(x)|$. D'après le théorème de Beppo-Levi, on a

$$\int h(x) dx = \sum \int |u_n(x)| dx = \sum \|u_n\|_{L^1} < \infty.$$

La fonction positive h est donc sommable. En particulier, l'ensemble A des points où $h(x) = +\infty$ est négligeable. Pour chaque x n'appartenant pas à A , la série numérique $\sum u_n(x)$ est absolument convergente et donc convergente. Posons

$$S_N(x) = \sum_0^N u_n(x) \quad ; \quad S(x) = \sum_0^\infty u_n(x),$$

cette dernière fonction étant définie dans le complémentaire de A . Les fonctions S_N sont majorées en module par la fonction sommable fixe h et la suite S_N converge presque partout vers S . D'après le théorème de Lebesgue, on a donc

$$\int |S(x) - S_N(x)| dx \xrightarrow{N \rightarrow \infty} 0.$$

Cela signifie très précisément que la série de terme général u_n converge vers S dans L^1 et achève donc la démonstration.

Corollaire 2.6.9. — *Si une suite f_j converge vers f dans L^1 , on peut en extraire une suite qui converge presque partout vers f .*

On peut en effet extraire une sous-suite g_j telle que l'on ait $\|g_j - f\|_{L^1} \leq 2^{-j}$. On a alors

$$g_j = g_0 + \sum_{k=0}^{j-1} (g_{k+1} - g_k).$$

La série de terme général $g_{k+1} - g_k$ est normalement convergente, et la démonstration précédente a montré qu'elle converge presque partout.

Remarque 2.6.10. — La démonstration ci-dessus a en fait prouvé une réciproque partielle du théorème de Lebesgue : si une suite f_j converge dans L^1 , on peut en extraire une sous-suite qui converge presque partout et qui est majorée en module par une fonction sommable fixe. Aucune de ces deux propriétés n'est vraie, en général, pour la suite f_j elle-même.

Exercice 2.6.11. — On considère la suite f_j de fonctions sur $[0, 1]$ ainsi définie. L'entier j s'écrit de manière unique sous la forme $j = 2^k + l$ avec $0 \leq l < 2^k$; la fonction f_j est égale à k dans l'intervalle $[l2^{-k}, (l+1)2^{-k}]$ et à 0 ailleurs. Montrer que $\int f_j$ tend vers 0, mais que la suite $f_j(x)$ ne converge en aucun point, et que la seule fonction qui majore toutes les f_j est la constante $+\infty$.

2.6.12. Espace $L^1(A)$. — Si A est un sous-ensemble de mesure > 0 de $\mathbb{R}^n$, on peut de même définir l'espace $L^1(A)$ des classes de fonctions sommables sur A . C'est un espace de Banach pour la norme $\int_A |f(x)| dx$. Il s'identifie au sous-espace fermé de $L^1(\mathbb{R}^n)$ constitué des classes de fonctions nulles (p.p.) dans le complémentaire de A .

Théorème 2.6.13. — *Soit Ω un ouvert de $\mathbb{R}^n$. L'espace des fonctions continues à support compact dans Ω est dense dans $L^1(\Omega)$.*

D'après le théorème 2.3.10, l'ouvert Ω est réunion d'une suite exhaustive de compacts K_j . Soit θ_j une fonction continue, à valeurs dans $[0, 1]$, égale à 1 sur K_j et à support dans K_{j+1} .

Soit f_1 la fonction égale à f dans Ω et à 0 ailleurs. Elle appartient à $L^1(\mathbb{R}^n)$ et il existe donc des φ_j continues et à support compact dans $\mathbb{R}^n$ telles que $\int_{\mathbb{R}^n} |f_1 - \varphi_j| \rightarrow 0$. Les intégrales sur Ω , qui sont inférieures, tendent aussi vers 0 et on a donc $\int_{\Omega} |f - \varphi_j| \rightarrow 0$.

Nous allons montrer que les fonctions $\varphi_j \theta_j$, qui sont à support compact dans Ω , tendent vers f dans $L^1(\Omega)$. On peut écrire

$$\int_{\Omega} |f(x) - \theta_j(x) \varphi_j(x)| dx \leq \int_{\Omega} |f(x) - \theta_j(x) f(x)| dx + \int_{\Omega} \theta_j(x) |f(x) - \varphi_j(x)| dx.$$

La seconde intégrale est majorée par celle de $|f - \varphi_j|$ et tend donc vers 0. Quant à la première, elle tend également vers 0 d'après le théorème de Lebesgue : la fonction à intégrer est majorée par $|f|$ qui est sommable, et elle tend vers 0 presque partout. Cela achève la démonstration.

Exercice 2.6.14. — Démontrer que l'espace $L^1(\Omega)$ est séparable (définition 2.4.3). En utilisant le théorème de Stone-Weierstrass, on montrera que les fonctions $\theta_j P$, où θ_j est comme ci-dessus et où P est un polynôme à coefficients rationnels, sont partout denses.

Théorème et Définition 2.6.15 (Espace $L^2(\mathbb{R}^n)$). — *C'est l'espace des classes de fonctions de carré sommable, pour la relation d'équivalence $f=g$ p.p.*

(a) Muni du produit scalaire

$$(f | g)_{L^2} = \int_{\mathbb{R}^n} f(x) \overline{g(x)} dx, \quad (2.13)$$

l'espace L^2 est un espace de Hilbert séparable.

(b) *L'espace $C_0(\mathbb{R}^n)$ est partout dense dans $L^2(\mathbb{R}^n)$.*

L'expression (2.13) est bien définie et ne dépend que des classes de fonctions f et g . Nous avons vérifié au n°1.5.4 toutes les propriétés d'un produit scalaire à l'exception de la suivante. Si $(f | f) = 0$ chaque représentant de f est nul presque partout, ce qui signifie précisément que $f = 0$.

Montrons que l'espace L^2 est complet, par une démonstration analogue à celle du théorème 2.6.8, en montrant qu'une série (de terme général u_n) normalement convergente dans L^2 est convergente.

En posant $v_n(x) = |u_n(x)|$, on a

$$\left\| \sum_0^N v_n \right\|_{L^2} \leq \sum_0^N \|v_n\|_{L^2} = \sum_0^N \|u_n\|_{L^2} \leq \sum_0^\infty \|u_n\|_{L^2}.$$

En posant $h(x) = (\sum_0^\infty |u_n(x)|)^2$, on a donc $\int h(x) dx \leq (\sum_0^\infty \|u_n\|_{L^2})^2$ d'après le théorème de Beppo-Levi. En particulier, l'ensemble A des points x où $h(x) = +\infty$ est de mesure nulle.

En tout point x du complémentaire de A , la série numérique de terme général $|u_n(x)|$ est absolument convergente et donc convergente. Les fonctions $S_N(x) = \sum_0^N u_n(x)$ convergent donc presque partout vers une fonction S vérifiant $S(x) \leq h(x)^{1/2}$ p.p. Les fonctions $|S(x) - S_N(x)|^2$ sont majorées par la fonction sommable fixe $4h$ et convergent presque partout vers 0. D'après le théorème de Lebesgue, on a donc

$$\int |S(x) - S_N(x)|^2 dx \xrightarrow{N \rightarrow \infty} 0,$$

ce qui exprime le fait que la série $\sum u_n$ est convergente dans L^2 et de somme S . Cela achève la démonstration du fait que L^2 est complet.

Démontrons maintenant la partie (b) et montrons qu'une fonction $f \in L^2$, que l'on peut supposer à valeurs réelles, peut être approchée à ε près par une fonction continue à support compact. Considérons d'abord pour tout j la fonction f_j , égale à 0 si $|x| \geq j$ ou si $|f(x)| \geq j$, et égale à $f(x)$ dans le cas contraire. Les fonctions $|f_j - f|^2$ sont majorées par $|f|^2$ et tendent vers 0 presque partout, et il résulte du théorème de Lebesgue que $\|f_j - f\|_{L^2}$ tend vers 0. On peut donc fixer j tel que $\|f_j - f\|_{L^2} \leq \varepsilon/2$.

La fonction f_j appartient à L^1 . En effet, $\int |f_j(x)| dx = \int |f_j(x)| \mathbf{1}_{B_j} dx$, en notant B_j la boule de rayon j et on a $\int |f_j(x)| dx \leq \mu(B_j)^{1/2} \|f\|_{L^2}$ par l'inégalité de Cauchy-Schwarz. Le nombre $\alpha > 0$ étant à déterminer ultérieurement, on peut trouver une fonction φ continue à support compact telle que $\|f_j - \varphi\|_{L^1} \leq \alpha$.

Posons $\psi(x) = \sup\{-j, \inf\{j, \varphi(x)\}\}$. C'est encore une fonction continue à support compact, et on voit facilement que l'on a $|f_j - \psi| \leq 2j$ et $|f_j - \psi| \leq |f_j - \varphi|$ en tout point. On a donc $\|f_j - \psi\|_{L^1} \leq \alpha$ et

$$\|f_j - \psi\|_{L^2}^2 = \int |f_j(x) - \psi(x)|^2 dx \leq \int 2j |f_j(x) - \psi(x)| dx \leq 2j\alpha.$$

Il suffit de choisir maintenant $\alpha = \varepsilon^2/8j$ et on a trouvé une fonction $\psi \in C_0(\mathbb{R}^n)$ telle que $\|f - \psi\|_{L^2} \leq \varepsilon$.

Il reste à prouver que L^2 est séparable. Nous allons montrer que l'ensemble dénombrable constitué des fonctions de la forme $\mathbf{1}_{B_j} P$, où $j \in \mathbb{N}$ et où P est un polynôme à coefficients rationnels des variables $x_1, \dots, x_n$, est dense dans $L^2(\mathbb{R}^n)$.

Soient donc $f \in L^2$ et $\varepsilon > 0$ donnés. Il est d'abord possible de trouver une fonction φ continue et à support compact telle que $\|f - \varphi\|_{L^2} \leq \varepsilon/2$. Soit j tel que le support de φ soit contenu dans B_j , le théorème de Stone-Weierstrass nous assure que pour tout $\alpha > 0$, il est possible de trouver un polynôme Q tel que $\sup_{x \in B_j} |\varphi(x) - Q(x)| \leq \alpha$. En approchant les

coefficients de Q par des nombres rationnels, il est donc possible de trouver un polynôme P à coefficients rationnels tel que $\sup_{x \in B_j} |\varphi(x) - P(x)| \leq 2\alpha$. On a alors

$$\left\| \varphi - \mathbf{1}_{B_j} P \right\|_{L^2(\mathbb{R}^n)}^2 = \int_{B_j} |\varphi(x) - P(x)|^2 dx \leq 4\alpha^2 \mu(B_j).$$

Il suffit de choisir $\alpha = \varepsilon / (4\sqrt{\mu(B_j)})$ pour terminer la démonstration.

Exercice 2.6.16. — Démontrer que si une suite f_j converge vers f dans L^2 , on peut en extraire une suite qui converge presque partout vers f . On pourra considérer les fonctions $g_j = |f - f_j|^2$.

2.6.17. Espace $L^2(A)$. — Si A est un sous-ensemble de mesure > 0 de $\mathbb{R}^n$, on peut de même définir l'espace $L^2(A)$ des classes de fonctions de carré sommable sur A . C'est un espace de Hilbert pour le produit scalaire $(f|g)_{L^2(A)} = \int_A f(x)g(x) dx$. Il s'identifie au sous-espace fermé de $L^2(\mathbb{R}^n)$ constitué des classes de fonctions nulles (p.p.) dans le complémentaire de A .

L'espace $L^2(A)$, comme d'ailleurs tout sous-espace fermé F d'un espace de Hilbert H séparable, est lui-même séparable. Il suffit de prendre les projections orthogonales sur F des éléments d'un ensemble dénombrable dense dans H pour obtenir un sous-ensemble dénombrable dense dans F . Dans ce cas particulier, la projection de f sur $L^2(A)$ est tout simplement la (classe de) fonction égale à f sur A et à 0 ailleurs, identifiée à $f|_A$.

Théorème 2.6.18. — Soit Ω un ouvert de $\mathbb{R}^n$. L'espace des fonctions continues à support compact dans Ω est dense dans $L^2(\Omega)$.

La démonstration, que nous laissons au lecteur à titre d'exercice, est identique à celle que nous avons donnée dans le cas $\Omega = \mathbb{R}^n$, en remplaçant les boules B_j par une suite exhaustive de compacts de Ω .

Théorème et Définition 2.6.19 (Espace $L^\infty(\mathbb{R}^n)$)

C'est l'espace des classes de fonctions essentiellement bornées pour la relation d'équivalence $f = g$ p.p. Muni de la norme (notée souvent $\|\cdot\|_\infty$)

$$\|f\|_{L^\infty} = \sup_{x \in \mathbb{R}^n} \text{ess } |f(x)|, \quad (2.14)$$

L^∞ est un espace de Banach.

Il est clair, après le passage au quotient, que $\|\cdot\|_{L^\infty}$ est une norme, et il reste à prouver que L^∞ est complet. Soit donc f_j une suite de Cauchy d'éléments de L^∞ , dont nous choisirons pour chaque j un représentant (noté encore f_j). Soit $A_{jk} = \{x \mid |f_j(x) - f_k(x)| > \|f_j - f_k\|_\infty\}$. C'est un ensemble de mesure nulle, et il en est donc de même de $A = \bigcup_{j,k} A_{jk}$. En dehors de A , la suite des f_j est de Cauchy pour la norme uniforme, et converge donc uniformément sur $\mathbb{C}A$ vers une fonction f bornée sur $\mathbb{C}A$. En prenant la classe d'équivalence (notée

encore f) des prolongements de f à $\mathbb{R}^n$, il est clair que l'on a $\|f - f_j\|_\infty \rightarrow 0$ et le résultat.

On peut montrer que l'espace $L^\infty(\mathbb{R}^n)$ n'est pas séparable, et l'exercice suivant montre que les fonctions continues à support compact ne sont pas denses dans L^∞ .

Exercice 2.6.20. — Démontrer que l'adhérence dans L^∞ de l'espace des fonctions continues à support compact est l'espace des (classes des) fonctions continues tendant vers 0 à l'infini.

2.6.21. Espace $L^\infty(A)$. — Si A est un sous-ensemble de mesure > 0 de $\mathbb{R}^n$, on peut de même définir l'espace $L^\infty(A)$ des classes de fonctions essentiellement bornées sur A . C'est un espace de Banach pour la norme $\|f\|_{L^\infty(A)} = \sup_{\text{ess}_A} |f(x)|$. Il s'identifie au sous-espace fermé de $L^\infty(\mathbb{R}^n)$ constitué des classes de fonctions nulles (p.p.) dans le complémentaire de A .

Proposition 2.6.22. — Soit A un sous-ensemble de $\mathbb{R}^n$ vérifiant $\mu(A) < \infty$. On a alors $L^\infty(A) \subset L^2(A) \subset L^1(A)$.

On a en effet

$$\int_A |f(x)|^2 dx \leq \int_A \|f\|_\infty^2 dx = \|f\|_\infty^2 \mu(A),$$

et, en utilisant Cauchy-Schwarz,

$$\int_A |f(x)| dx = \int |f(x)| \mathbf{1}_A(x) dx \leq \|f\|_2 \mu(A)^{1/2}.$$

Comme le montreront les exercices qui vont suivre, ces inclusions sont strictes. Dans $\mathbb{R}^n$ entier (ou dans un ensemble de mesure infinie), on n'a d'inclusion dans aucun sens, et on ne dispose que de la proposition suivante. Toutefois (voir l'exercice 2.6.26), si on se prémunit contre la présence de pics trop pointus, on retrouve la situation que l'on avait pour les espaces de suites ($l^1 \subset l^2 \subset l^\infty$).

Proposition 2.6.23. — Pour tout sous-ensemble (de mesure > 0) de $\mathbb{R}^n$, on a $L^1(A) \cap L^\infty(A) \subset L^2(A)$.

On a en effet

$$\int_A |f(x)|^2 dx \leq \int_A \|f\|_\infty |f(x)| dx \leq \|f\|_\infty \|f\|_1.$$

Exercice 2.6.24. — Montrer que sur $[0, 1]$, en utilisant par exemple des fonction $x \mapsto x^{-\alpha}$, les inclusions de la proposition 2.6.22 sont strictes.

Exercice 2.6.25. — En considérant des suites α_j bien choisies, et en formant une fonction continue f nulle hors des intervalles $[j - \alpha_j, j + \alpha_j]$, linéaire dans chaque intervalle $[j - \alpha_j, j]$ et $[j, j + \alpha_j]$, et vérifiant $f(j) = j$, montrer qu'il existe des fonctions continues appartenant à $L^1(\mathbb{R})$ et pas à $L^2(\mathbb{R})$, ainsi que des fonctions continues appartenant à $L^2(\mathbb{R})$ et pas à $L^\infty(\mathbb{R})$.

Exercice 2.6.26. — Soit $\mathcal{A}$ l'espace des fonctions uniformément continues de $\mathbb{R}^n$ dans $\mathbb{C}$. Démontrer que l'on a $\mathcal{A} \cap L^1 \subset \mathcal{A} \cap L^2 \subset \mathcal{A} \cap L^\infty$.

2.6.27. Espaces locaux. — On dit qu'une fonction f définie dans un ouvert Ω de $\mathbb{R}^n$ est *localement sommable* si pour tout compact K de Ω on a $\int_K |f(x)| dx < \infty$. Il est équivalent de dire (Borel-Lebesgue) que chaque point de Ω possède un voisinage dans lequel f est sommable.

On note $L^1_{\text{loc}}(\Omega)$ le quotient de l'espace des fonctions localement sommables par la relation d'équivalence $f = g$ p.p.

On dit de même qu'une fonction est localement de carré sommable dans Ω , si sa restriction à chaque compact de Ω est de carré sommable. On note $L^2_{\text{loc}}(\Omega)$ le quotient de l'espace des fonctions localement de carré sommable par la même relation d'équivalence.

On définit enfin par le même procédé l'espace $L^\infty_{\text{loc}}(\Omega)$ des classes de fonctions qui sont essentiellement bornées sur chaque compact. D'après la proposition 2.6.22, on a $\mathcal{L}^\infty_{\text{loc}}(\Omega) \subset \mathcal{L}^2_{\text{loc}}(\Omega) \subset \mathcal{L}^1_{\text{loc}}(\Omega)$.

Ces espaces ne peuvent pas être munis de normes raisonnables. Le lecteur pourra voir dans l'appendice C qu'ils peuvent être munis d'une structure d'espace de Fréchet.

2.7. Séries de Fourier

On se donne dans cette section un nombre $T > 0$, et on pose $\omega = 2\pi/T$.

Si f est une fonction définie sur $\mathbb{R}$ et vérifiant $f(t - T) = f(t)$ p.p., il en est de même de toute fonction égale presque partout à f . On dit qu'une classe de fonctions f est T -périodique si l'un quelconque de ses représentants vérifie la propriété ci dessus.

On définit l'espace L^2_T comme étant l'espace des (classes de) fonctions T -périodiques localement de carré sommable. Pour qu'une fonction T -périodique appartienne à L^2_T , il suffit qu'elle soit de carré sommable sur un intervalle période ($[0, T]$ par exemple), tout compact de $\mathbb{R}$ pouvant être recouvert par un nombre fini de translatés $[kT, (k + 1)T]$ de celui-ci. Il en résulte que

l'application de L_T^2 dans $L^2([0, T])$ qui à f fait correspondre sa restriction à $[0, T]$ est un isomorphisme.

L'espace L_T^2 sera muni du produit scalaire

$$(f | g) = \int_0^T f(t) \overline{g(t)} dt / T.$$

A la normalisation de la mesure près, il est isométrique à $L^2([0, T])$ et est donc un espace de Hilbert.

Théorème 2.7.1. — *Les fonctions $e_p(t) = e^{ip\omega t}$, $p \in \mathbb{Z}$, forment une base hilbertienne de L_T^2 .*

Il est très facile de prouver le caractère orthonormal de cet ensemble, et il reste à prouver que le système des e_p est total. Nous en proposons deux démonstrations.

La première s'appuie sur les connaissances antérieures du lecteur sur les séries de Fourier. Nous n'aurons besoin que de l'énoncé très faible suivant : si f est une fonction T -périodique de classe C^∞ , sa série de Fourier converge uniformément vers f (un énoncé plus fort et sans doute connu du lecteur est rappelé dans le théorème 6.3.3).

D'autre part, nous verrons dans le prochain chapitre (théorème 3.5.1) qu'un élément de $L^2([0, T])$ peut être approché en norme L^2 par des fonctions de classe C^∞ à support compact dans $]0, T[$ (nous ne connaissons pour le moment que l'approximation par des fonctions continues à support compact (théorème 2.6.18) ce qui serait insuffisant ici).

Soient maintenant $f \in L_T^2$ et $\varepsilon > 0$. Soit φ une fonction C^∞ à support compact dans $]0, T[$ telle que $\int_0^T |f(t) - \varphi(t)|^2 dt \leq \varepsilon/2$. La fonction $\tilde{\varphi}(t) = \sum_{p \in \mathbb{Z}} \varphi(t - pT)$ est T -périodique de classe C^∞ et coïncide avec φ dans $[0, T]$. On a donc $\|f - \tilde{\varphi}\|_{L_T^2} \leq \varepsilon/2$. D'autre part, il existe une somme partielle S_N de la série de Fourier de $\tilde{\varphi}$ qui vérifie $\sup_t |\tilde{\varphi}(t) - S_N(t)| \leq \varepsilon/2$, ce qui entraîne que $\int_0^T |\tilde{\varphi}(t) - S_N(t)|^2 dt / T \leq \varepsilon^2/4$. On a donc également $\|\tilde{\varphi} - S_N\|_{L_T^2} \leq \varepsilon/2$. Nous avons donc pu approcher, à ε près en norme, la fonction f par la fonction S_N qui est une combinaison linéaire des e_p .

La seconde démonstration n'utilise aucune connaissance préalable sur les séries de Fourier. Étant donnés f et ε comme ci-dessus, on choisit une fonction φ continue à support compact dans $]0, T[$ telle que $\int_0^T |f(t) - \varphi(t)|^2 dt \leq \varepsilon/2$. On construit sa périodisée $\tilde{\varphi}(t) = \sum_{p \in \mathbb{Z}} \varphi(t - pT)$ qui est continue, T -périodique et vérifie $\|f - \tilde{\varphi}\|_{L_T^2} \leq \varepsilon/2$. On s'appuie maintenant sur le théorème de Stone-Weierstrass qui affirme (voir exemple 2.3.22) qu'il existe un polynôme

trigonométrique $P(t) = \sum_{n=-p}^q \lambda_n e_n(t)$ tel que $\sup_t |\tilde{\varphi}(t) - P(t)| \leq \varepsilon/2$. On conclut comme ci-dessus.

En explicitant dans ce cas particulier le théorème 2.5.14 relatif aux bases hilbertiennes, on obtient le résultat suivant.

Théorème 2.7.2. — (a) *Tout élément $f \in L_T^2$ peut se décomposer de façon unique sous la forme*

$$f = \sum_{p \in \mathbb{Z}} c_p(f) e^{ip\omega t},$$

la série convergeant en moyenne quadratique (c'est-à-dire pour la norme de L_T^2). Les composantes $c_p(f)$ sont données par

$$c_p(f) = (f | e_p) = \int_0^T f(t) e^{-ip\omega t} dt / T,$$

et vérifient

$$\int_0^T |f(t)|^2 dt / T = \sum_{p \in \mathbb{Z}} |c_p(f)|^2 \quad (\text{Bessel-Parseval})$$

(b) *Réciproquement, étant donnés des scalaires γ_p vérifiant $\sum_{p \in \mathbb{Z}} |\gamma_p|^2 < \infty$, la série $\sum \gamma_p e^{ip\omega t}$ converge en moyenne quadratique vers une fonction $f \in L_T^2$ telle que l'on ait $c_p(f) = \gamma_p$.*

Ce théorème caractérise complètement les fonctions de carré sommable en termes de leur séries de Fourier. Il ne fournit toutefois que la convergence en moyenne quadratique, et ne dit rien sur la convergence ponctuelle. Un théorème considérablement plus difficile (L. Carleson, 1965) assure que les sommes partielles symétriques de la série de Fourier d'une fonction $f \in L_T^2$ convergent presque partout vers f .

Exercice 2.7.3. — Soit f une fonction T -périodique deux fois continûment dérivable. Démontrer que l'on a $|c_p(f)| \leq C^{\text{te}} |p|^{-2}$ pour $p \neq 0$. En déduire que la série de Fourier de f converge uniformément vers une fonction g continue T -périodique. En utilisant le théorème 2.7.2, démontrer que l'on a $f = g$ et que la série de Fourier converge donc uniformément vers f . (Cela n'apporte rien si on s'est appuyé sur ce résultat pour démontrer que les e_p forment un système total. Par contre, si on a utilisé le théorème de Stone-Weierstrass, on obtient un résultat de convergence uniforme sans aucune analyse du noyau $\sum_{-N}^N e^{ip\omega t}$).

Exercice 2.7.4 (Polynômes et fonctions d'Hermite)

(a) Démontrer que l'on a $(d/dx)^n e^{-x^2} = (-1)^n e^{-x^2} H_n(x)$, où H_n est un polynôme de degré exactement n dont on calculera le terme de plus haut degré.

(b) On pose

$$\Psi_n(x) = c_n e^{-x^2/2} H_n(x),$$

où c_n est une constante qui sera choisie ultérieurement. Montrer que les Ψ_n appartiennent à $L^2(\mathbb{R})$ et calculer les produits scalaires $(\Psi_p | \Psi_q)_{L^2}$. On rappelle que $\int e^{-x^2} dx = \pi^{1/2}$. On pourra intégrer p fois (ou q fois) par parties.

(c) En déduire que, en posant $c_n = \pi^{-1/4} 2^{-n/2} (n!)^{-1/2}$, les fonctions Ψ_n , $n \in \mathbb{N}$, forment un système orthonormal dans L^2 .

(d) On se donne $f \in L^2$ tel que l'on ait $\int e^{-x^2/2} P(x) f(x) dx = 0$ pour tout polynôme P . On pose $g = e^{-x^2/2} f$. Montrer que $g \in L^1(\mathbb{R})$. En anticipant sur la suite du cours (voir le théorème 9.1.5), on admettra que si une fonction sommable g a une transformée de Fourier $\widehat{g}(\xi) = \int e^{-ix\xi} g(x) dx$ identiquement nulle, alors g est nulle presque partout.

(d1) Pour ξ fixé, démontrer qu'il existe une suite de polynômes Q_n telle que l'on ait

$$Q_n(x) \xrightarrow{n \rightarrow \infty} e^{-ix\xi} \quad \text{et} \quad |Q_n(x)| \leq e^{|x\xi|}$$

pour tout x (on développera $e^{-ix\xi}$ en série entière).

(d2) En déduire que $\widehat{g}(\xi) = 0$ et que $f = 0$.

(e) En déduire que les fonctions d'Hermite Ψ_n forment une base hilbertienne de $L^2(\mathbb{R})$.

CHAPITRE 3

FONCTIONS DIFFÉRENTIABLES ET APPROXIMATION

Le point important de ce chapitre est le fait qu'il existe "beaucoup" de fonctions de classe C^∞ à support compact. Il en existe suffisamment pour disposer de partitions de l'unité, pour pouvoir effectuer des troncatures et pour que l'espace $C_0^\infty(\Omega)$ soit dense dans de nombreux espaces fonctionnels.

Le lecteur trouvera dans la section 3.1 un minimum de rappels sur le calcul différentiel dans $\mathbb{R}^n$, assorti de notations très utiles. Un point de vue plus intrinsèque est développé dans l'appendice A.

3.1. Espaces de fonctions différentiables

Rappelons qu'une application f définie dans un ouvert Ω de $\mathbb{R}^n$ et à valeurs réelles ou complexes est dite différentiable au point x s'il existe une forme linéaire L , appelée *différentielle* de f au point x telle que l'on ait

$$f(x+h) = f(x) + L(h) + |h| r(h), \text{ avec } \lim_{h \rightarrow 0} |r(h)| = 0.$$

La forme linéaire L est souvent notée $df(x)$ et $L(h)$ s'écrit alors $df(x) \cdot h$ ou bien $\langle df(x), h \rangle$. C'est un élément de l'espace $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$ (le dual de $\mathbb{R}^n$) ou $\mathcal{L}(\mathbb{R}^n, \mathbb{C})$ selon les cas. Si f est différentiable en x , ses dérivées partielles existent en ce point et on a $df(x) \cdot h = \sum \partial f / \partial x_i(x) h_i$. La réciproque est fausse, mais on a le résultat suivant.

Théorème et Définition 3.1.1. — *Soit f définie dans Ω . Les deux propriétés suivantes sont équivalentes.*

- (a) *Les dérivées partielles $\partial f / \partial x_i(x)$ existent et sont continues dans Ω .*
- (b) *La fonction f est différentiable en tout point, et l'application $x \mapsto df(x)$, de Ω dans $\mathcal{L}(\mathbb{R}^n, \mathbb{C})$ est continue.*

On dit que la fonction f est continûment différentiable — ou de classe C^1 — dans Ω si ces propriétés sont vérifiées, et on note $C^1(\Omega)$ l'espace des fonctions de ce type.

Remarque 3.1.2. — L'expérience et l'histoire montrent que, même en dimension 1, l'espace des fonctions différentiables en tout point doit être considéré comme un "mauvais" espace : son étude est difficile et il possède peu de propriétés intéressantes. Le point de vue qui a prévalu est de s'intéresser principalement ou bien aux fonctions continûment différentiables, ou bien à une notion beaucoup plus faible de dérivée (dérivée au sens des distributions) que nous étudierons plus loin.

Si f est de classe C^1 , ses dérivées partielles qui existent en tout point peuvent être elles-mêmes de classe C^1 , auquel cas on dit que f est de classe C^2 . Par récurrence, on définit l'espace $C^k(\Omega)$ des fonctions f dont les dérivées partielles $\partial f / \partial x_i$ appartiennent à $C^{k-1}(\Omega)$. On dit enfin que f est de classe C^∞ si elle est de classe C^k pour tout k .

Le lecteur connaît bien le lemme de Schwarz qui assure que, pour tout $f \in C^2(\Omega)$, on a

$$\frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right).$$

Il en résulte que dans les dérivées partielles d'ordre k d'une fonction de classe C^k , l'ordre des dérivations n'intervient pas. Les notations suivantes nous seront très utiles.

3.1.3. Multiindices et notations. — On notera $x = (x_1, \dots, x_n)$ le point courant de $\mathbb{R}^n$ et $|x|$ sa norme euclidienne. On appelle *multiindices* les éléments $\alpha = (\alpha_1, \dots, \alpha_n)$ de $\mathbb{N}^n$. Ceux-ci fournissent des notations commodes pour les monômes et les dérivations partielles

$$\begin{aligned} x^\alpha &= x_1^{\alpha_1} \dots x_n^{\alpha_n} \\ \partial_i &= \partial / \partial x_i \\ \partial^\alpha f &= \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n} f = \frac{\partial^{\alpha_1 + \dots + \alpha_n} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}. \end{aligned}$$

La notation ci-dessus doit être strictement réservée aux cas où on sait que l'ordre des dérivations partielles n'intervient pas, c'est à dire soit (lemme de Schwarz) si f est de classe C^k , $k = \alpha_1 + \dots + \alpha_n$, soit, comme nous le verrons plus loin, en prenant les dérivées au sens des distributions.

L'entier $\alpha_1 + \dots + \alpha_n$ est noté $|\alpha|$; on l'appelle *longueur* du multiindice α , c'est le degré du monôme x^α , et l'ordre de la dérivation ∂^α . On notera $\beta \leq \alpha$ la relation d'ordre partiel $\forall i, \beta_i \leq \alpha_i$. Enfin, la notation suivante

pour les factorielles et les coefficients binomiaux généralisés permet d'avoir des expressions concises et très proches des expressions correspondantes en dimension 1.

$$\alpha! = \alpha_1! \dots \alpha_n! \quad \binom{\alpha}{\beta} = \frac{\alpha!}{\beta! (\alpha - \beta)!}.$$

Le lecteur n'éprouvera pas de difficulté à démontrer par récurrence les formules suivantes.

3.1.4. Formule du binôme.

$$(x + y)^\alpha = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} x^{\alpha - \beta} y^\beta.$$

3.1.5. *Formule de Leibniz.* — Pour f et g de classe $C^{|\alpha|}$, on a

$$\partial^\alpha (fg) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \partial^{\alpha - \beta} f \partial^\beta g.$$

3.1.6. *Formule du multinôme.* — Pour $x \in \mathbb{R}^n$ et $k \in \mathbb{N}$, on a

$$(x_1 + x_2 + \dots + x_n)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} x^\alpha.$$

3.1.7. Formule de Taylor. — Nous donnerons ici la forme intégrale du reste, qui a l'avantage d'en donner une expression exacte et qui permet d'en obtenir les majorations les plus précises.

Théorème 3.1.8. — Soit f de classe C^{m+1} dans un ouvert contenant le segment $[a, b]$, on a

$$f(b) = \sum_{|\alpha| \leq m} \frac{(b-a)^\alpha}{\alpha!} \partial^\alpha f(a) + R, \quad (3.1)$$

avec

$$R = (m+1) \sum_{|\alpha|=m+1} \frac{(b-a)^\alpha}{\alpha!} \int_0^1 (1-t)^m \partial^\alpha f(a + t(b-a)) dt.$$

Nous rappelons au lecteur la forme correspondante pour les fonctions d'une variable de classe C^{m+1} dans l'intervalle $[0, 1]$

$$\varphi(1) = \sum_{k=0}^m \frac{1}{k!} \frac{d^k \varphi}{dt^k}(0) + \frac{1}{m!} \int_0^1 (1-t)^m \frac{d^{m+1} \varphi}{dt^{m+1}}(t) dt,$$

qui s'obtient en intégrant m fois par parties l'intégrale du membre de droite. Nous allons appliquer cette formule à la fonction φ définie sur $[0, 1]$ par $\varphi(t) = f(a + th)$ en posant $h = b - a$. On a facilement $\varphi'(t) = \sum_i h_i \partial_i f(a + th)$ et par récurrence

$$\varphi^{(k)}(t) = \sum_{i_1=1}^n \dots \sum_{i_k=1}^n h_{i_1} \dots h_{i_k} \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(a + th).$$

Regroupons tous les termes égaux à $h^\alpha \partial^\alpha f(a + th)$ dont le nombre est égal à $k!/(\alpha_1! \dots \alpha_n!)$ (c'est exactement le même calcul que celui qui établit la formule 3.1.6). On obtient $\varphi^{(k)}(t) = \sum_{|\alpha|=k} (k!/|\alpha|!) h^\alpha \partial^\alpha f(a + th)$ et la formule (3.1) en résulte immédiatement.

Corollaire 3.1.9 (lemme de Hadamard). — Soit f une fonction de classe C^{m+1} définie dans $\mathbb{R}^n$ et s'annulant à l'origine. Il existe des fonctions $g_1, \dots, g_n$ appartenant à C^m telles que l'on ait $f(x) = \sum_1^n x_i g_i(x)$.

La formule de Taylor écrite à l'ordre 0 fournit la décomposition

$$f(x) = f(0) + \sum_i x_i \int_0^1 \partial_i f(tx) dt = 0 + \sum x_i g_i(x).$$

Il reste à montrer que les fonctions g_i sont de classe C^m dans $\mathbb{R}^n$, ce que le lecteur démontrera facilement à titre d'exercice, à l'aide du théorème de Lebesgue et du théorème de dérivation sous le signe somme.

Nous verrons plus loin (exercice 3.2.8) que ce lemme est valable, ainsi que ses extensions ci-dessous, en remplaçant $\mathbb{R}^n$ par un ouvert quelconque contenant l'origine.

Exercice 3.1.10. — Soient $k \leq m$ et $f \in C^{m+1}(\mathbb{R}^n)$ nulle à l'origine ainsi que ses dérivées partielles d'ordre $\leq k$. Démontrer que l'on a la décomposition suivante $f(x) = \sum_{|\alpha|=k+1} x^\alpha g_\alpha(x)$ avec $g_\alpha \in C^{m-k}(\mathbb{R}^n)$.

Exercice 3.1.11. — On se donne maintenant $f(t, x)$ de classe C^{m+1} dans $\mathbb{R}_t^p \times \mathbb{R}_x^n$ vérifiant $f(t, 0) = 0$ pour tout t . Démontrer que l'on a $f(t, x) = \sum_{i=1}^n x_i g_i(t, x)$ avec $g_i \in C^m(\mathbb{R}^{n+p})$. Généraliser au cas où les dérivées partielles jusqu'à l'ordre k s'annulent pour $x = 0$.

3.1.12. Fonctions composées. — Soient Ω un ouvert de $\mathbb{R}^n$ et $x \mapsto \chi(x) = (\chi_1(x), \dots, \chi_p(x))$ une application de Ω dans $\mathbb{R}^p$. On dit que χ est de classe C^k si les applications composantes χ_j sont de classe C^k .

Soit maintenant Ω' un ouvert de $\mathbb{R}^p$ qui contient $\chi(\Omega)$, et soit f une application de classe C^k de Ω' dans $\mathbb{R}$ ou $\mathbb{C}$. L'application $f \circ \chi$ est alors de classe

C^k dans Ω et on a

$$\frac{\partial(f \circ \chi)}{\partial x_i}(x) = \sum_{l=1}^p \frac{\partial f}{\partial y_l}(\chi(x)) \frac{\partial \chi_l}{\partial x_i}(x) \quad i = 1, \dots, n. \quad (3.2)$$

Le membre de droite est une somme de produits de fonctions de classe C^{k-1} , le premier facteur étant lui-même une fonction composée. Pour $k > 1$ il est donc possible de calculer les dérivées d'ordre supérieur de $f \circ \chi$ en utilisant récursivement la formule de Leibniz et la formule (3.2) elle-même. Les formules explicites sont passablement compliquées, même pour $n = p = 1$.

Exercice 3.1.13. — Sous les hypothèses précédentes, pour $m \leq k$ montrer que l'on a

$$\begin{aligned} \frac{\partial^m(f \circ \chi)}{\partial x_{i_1} \dots \partial x_{i_m}}(x) &= \sum_{l_1=1}^p \dots \sum_{l_m=1}^p \frac{\partial^m f}{\partial y_{l_1} \dots \partial y_{l_m}}(\chi(x)) \prod_{r=1}^m \frac{\partial \chi_{l_r}}{\partial x_{i_r}}(x) \\ &\quad + \sum_{\lambda \in \mathbb{N}^p : |\lambda| < m} h_\lambda(x) (\partial^\lambda f) \circ \chi(x). \end{aligned}$$

où les fonctions h_λ , que l'on ne cherchera pas à calculer explicitement, ne dépendent que des χ_l et de leurs dérivées et sont de classe C^{k-m} .

Exercice 3.1.14 (Formule de Faa de Bruno). — Soient f et g deux fonctions définies sur $\mathbb{R}$ de classe C^n . Démontrer que l'on a

$$\frac{d^n}{dx^n} g \circ f(x) = \sum \frac{n!}{m_1! \dots m_q!} g^{(p)}(f(x)) \left(\frac{f'(x)}{1!} \right)^{m_1} \dots \left(\frac{f^{(q)}(x)}{q!} \right)^{m_q},$$

où la somme est étendue à toutes les suites finies d'entiers vérifiant $m_1 + 2m_2 + \dots + qm_q = n$, et où on a posé $p = m_1 + m_2 + \dots + m_q$.

3.2. Partitions de l'unité C^∞

Les espaces suivants seront d'utilisation courante. On désignera par Ω un ouvert de $\mathbb{R}^n$, par K un compact de $\mathbb{R}^n$ et par m , soit un entier ≥ 0 soit le symbole ∞ .

3.2.1. L'espace $C^m(\Omega)$. — C'est l'espace des fonctions m fois continûment différentiables dans Ω . En particulier, l'espace $C^0(\Omega)$ (noté aussi $C(\Omega)$) désigne l'espace des fonctions continues dans Ω et C^∞ désigne l'espace des fonctions indéfiniment dérivables.

3.2.2. L'espace C_K^m . — C'est l'espace des fonctions appartenant à $C^m(\mathbb{R}^n)$ à support dans K (c'est-à-dire nulles hors de K).

3.2.3. L'espace $C_0^m(\Omega)$. — C'est la réunion des espaces C_K^m lorsque K parcourt l'ensemble des compacts contenus dans Ω .

Il n'est pas complètement évident qu'il existe des fonctions de classe C^∞ à support compact. Le lemme suivant nous sera utile.

Lemme 3.2.4. — (a) *La fonction Φ , définie sur $\mathbb{R}$ par $\Phi(t) = e^{-1/t}$ pour $t > 0$ et par 0 sinon, est croissante, de classe C^∞ , nulle pour $t \leq 0$ et strictement positive pour $t > 0$.*

(b) *Il existe une fonction Ψ définie sur $\mathbb{R}$, croissante et de classe C^∞ , telle que l'on ait $\Psi(t) = 0$ pour $t \leq 0$ et $\Psi(t) = 1$ pour $t \geq 1$.*

Il est facile de démontrer, par récurrence sur k , que l'on a, pour $t > 0$, la relation $\Phi^{(k)}(t) = e^{-1/t} P_k(1/t)$ où P_k est un polynôme. Il en résulte que $\lim_{t \rightarrow 0} \Phi^{(k)}(t) = 0$ ce qui assure que $\Phi^{(k-1)}$ est dérivable à droite à l'origine et que cette dérivée est nulle. On obtient donc $\Phi \in C^\infty(\mathbb{R})$ et les autres propriétés de Φ sont évidentes.

Considérons maintenant la fonction g définie par $g(t) = \Phi(t)\Phi(1-t)$. Elle est positive, de classe C^∞ et est nulle en dehors de $[0, 1]$. Posons $I = \int_{\mathbb{R}} g(t) dt$. La fonction

$$\Psi(t) = I^{-1} \int_{-\infty}^t g(s) ds$$

vérifie les conditions de la partie (b).

Corollaire 3.2.5. — *Il existe des fonctions non identiquement nulles appartenant à $C_0^\infty(\mathbb{R}^n)$. Plus précisément, étant donnée une boule ouverte $B(a, r)$, il existe une fonction $\varphi \in C_0^\infty(\mathbb{R}^n)$ strictement positive dans cette boule et à support dans son adhérence.*

Il suffit de poser

$$\varphi(x) = \Phi(r^2 - |x-a|^2).$$

Corollaire 3.2.6. — *Soient Ω un ouvert de $\mathbb{R}^n$ et K un compact contenu dans Ω . Il existe une fonction $f \in C_0^\infty(\Omega)$ à valeurs dans $[0, 1]$ et égale à 1 au voisinage de K .*

Pour chaque $x \in K$, on peut trouver une boule B_x centrée en x dont l'adhérence est contenue dans Ω . On peut lui associer, en multipliant la fonction du corollaire précédent par une constante assez grande, une fonction positive φ_x à support dans $\overline{B_x}$, de classe C^∞ et qui vérifie $\varphi_x(x) > 1$. Les ouverts $V_x = \{y | \varphi_x(y) > 1\}$ constituent un recouvrement ouvert de K , dont on peut extraire un recouvrement fini $V_{x_1}, \dots, V_{x_N}$.

La fonction $h = \sum_1^N \varphi_{x_i}$ est positive, appartient à $C_0^\infty(\Omega)$, et vérifie $h(x) \geq 1$ au voisinage de K . La fonction $f = \Psi \circ h$ vérifie les conditions voulues, Ψ étant la fonction du lemme 3.2.4(b).

Exercice 3.2.7. — Soient $K_1, \dots, K_p$ des compacts disjoints de $\mathbb{R}^n$ et soient $f_1, \dots, f_p$, des éléments de $C^\infty(\mathbb{R}^n)$. Montrer qu'il existe $f \in C_0^\infty(\mathbb{R}^n)$ qui, pour chaque $j = 1, \dots, p$ vérifie $f(x) = f_j(x)$ au voisinage de K_j .

Exercice 3.2.8. — (*Lemme de Hadamard dans un ouvert*) Soient Ω un ouvert de $\mathbb{R}^n$ contenant l'origine et $f \in C^\infty(\Omega)$ s'annulant à l'origine. Montrer qu'il existe des fonctions $g_i \in C^\infty(\Omega)$ telles que l'on ait $f(x) = \sum_1^n x_i g_i(x)$ (on considérera une fonction φ de classe C^∞ égale à 1 au voisinage de 0, à support dans une boule contenue dans Ω et on décomposera séparément la fonction φf , à l'aide du lemme de Hadamard (corollaire 3.1.9), et la fonction $(1 - \varphi)f$, ce qui n'est pas difficile).

Théorème 3.2.9 (Partitions de l'unité). — *Soit K un compact contenu dans la réunion d'un nombre fini d'ouverts $\omega_1, \dots, \omega_N$. Il existe alors des fonctions $\chi_j, j = 1, \dots, N$, appartenant à $C_0^\infty(\omega_j)$, comprises entre 0 et 1, telles que l'on ait*

$$\sum_{j=1}^N \chi_j = 1 \quad \text{au voisinage de } K.$$

Compte tenu du corollaire précédent, la démonstration est essentiellement la même que celle du théorème 2.3.25. D'après la proposition 2.3.24, il existe des compacts $K_j \subset \omega_j$ dont la réunion contient K . D'après le corollaire 3.2.6 il existe des éléments ψ_j de $C_0^\infty(\omega_j)$ compris entre 0 et 1, et égaux à 1 au voisinage de K_j . On a donc $\sum_{j=1}^N \psi_j(x) > 0$ dans un voisinage V de K . Soit encore $\theta \in C_0^\infty(V)$ comprise entre 0 et 1, égale à 1 au voisinage de K et posons $\psi_0 = 1 - \theta$. On voit facilement que les fonctions

$$\chi_j = \frac{\psi_j}{\sum_{i=0}^N \psi_i}, \quad j = 1, \dots, N$$

vérifient les propriétés voulues, le dénominateur étant une fonction de classe C^∞ strictement positive dans tout $\mathbb{R}^n$.

On peut en fait démontrer le résultat suivant, où Ω est un ouvert de $\mathbb{R}^n$, une hypersurface ou une sous-variété de $\mathbb{R}^n$ (voir appendice A), ou plus généralement une variété différentielle qui est réunion dénombrable de compacts. Étant donnée une famille non nécessairement finie d'ouverts $(\omega_i)_{i \in I}$, dont la réunion est Ω entier, il existe une famille $(\varphi_\alpha)_{\alpha \in A}$ d'éléments de $C_0^\infty(\Omega)$ telles que :

— Pour chaque α , le support de φ_α est contenu dans l'un des ω_i .

— La famille des supports des φ_α est localement finie : chaque point possède un voisinage ne rencontrant qu'un nombre fini de ces supports.

— la somme des fonctions φ_α , qui est alors bien définie, est égale à 1 dans tout Ω .

Ces partitions de l'unité permettent de définir ou de construire des "objets" globaux à partir de données locales. Par exemple, la construction de l'intégrale de surface que nous donnons dans la section A.3 dans le cas compact s'étendrait au cas d'une hypersurface quelconque à l'aide du résultat ci-dessus.

Exercice 3.2.10. — Soit (a_j) une suite numérique quelconque. Démontrer qu'il existe une fonction $f \in C_0^\infty(\mathbb{R})$ vérifiant $(d/dt)^j f(0) = a_j$ pour tout entier j (théorème de Borel). On considérera une fonction $\Phi \in C_0^\infty(\mathbb{R})$ égale à 1 au voisinage de l'origine, et on posera

$$f = \sum_j g_j, \quad g_j(t) = \Phi(t/\varepsilon_j) a_j t^j / j!.$$

On montrera qu'il est possible de choisir les ε_j strictements positifs tels que l'on ait $|(d/dt)^k g_j(t)| \leq 2^{-j}$ pour tout t et pour $k \leq j-1$.

Plus généralement, étant données des fonctions $a_j(x)$ de classe C^∞ dans $\mathbb{R}^n$, montrer qu'il existe une fonction $f(t, x)$ de classe C^∞ dans $\mathbb{R}^{n+1}$ vérifiant $\partial_t^j f(0, x) = a_j(x)$ quels que soient j et x .

Exercice 3.2.11. — Soit F un fermé quelconque de $\mathbb{R}^n$. Démontrer qu'il existe une fonction $f \in C^\infty(\mathbb{R}^n)$ telle que $f^{-1}(0) = F$. On commencera par montrer qu'il existe une suite B_j de boules ouvertes de $\mathbb{R}^n$ dont la réunion est $\mathbb{C}F$ (prendre toutes les boules de rayon rationnel et dont le centre a des coordonnées rationnelles qui sont contenues dans $\mathbb{C}F$). On prendra ensuite pour chaque j la fonction φ_j fournie par le corollaire 3.2.5 et on montrera que l'on peut choisir des $\varepsilon_j > 0$ tels que la fonction $f = \sum \varepsilon_j \varphi_j$ soit de classe C^∞ . Comme dans l'exercice précédent, on montrera qu'il suffit d'avoir $\varepsilon_j \|\partial^\alpha \varphi_j\|_{L^\infty} \leq 2^{-j}$ pour $j \geq |\alpha|$, et que cela n'impose qu'un nombre fini de conditions sur chaque ε_j .

3.3. Convolution

Théorème et Définition 3.3.1. — Si f et g appartiennent à $L^1(\mathbb{R}^n)$, le produit de convolution $f \star g$ est défini par

$$(f \star g)(x) = \int f(x-y)g(y) dy = \int g(x-y)f(y) dy, \quad (3.3)$$

l'intégrale étant bien définie pour presque tout x . En outre, la fonction $f \star g$ appartient à L^1 et on a

$$\|f \star g\|_{L^1} \leq \|f\|_{L^1} \|g\|_{L^1}. \quad (3.4)$$

Pour x fixé, le membre de droite de (3.3) est l'intégrale du produit de deux fonctions sommables (de la variable y), produit qui n'a en général aucune raison d'être sommable. C'est le théorème de Fubini qui nous assurera de la sommabilité pour presque tout x .

Montrons d'abord que la fonction $f(x-y)g(y)$ est sommable dans $\mathbb{R}^{2n}$. D'après le théorème de Fubini pour les fonctions positives, on a

$$\iint |f(x-y)g(y)| \, dx \, dy = \int |g(y)| \left\{ \int |f(x-y)| \, dx \right\} dy. \quad (3.5)$$

Le changement de variable $z = x - y$, à y fixé, a un jacobien égal à 1, et on a donc

$$\int |f(x-y)| \, dx = \int |f(z)| \, dz = \|f\|_{L^1}.$$

La constante précédente sort de l'intégrale (3.5), et on obtient

$$\iint |f(x-y)g(y)| \, dx \, dy = \|f\|_{L^1} \|g\|_{L^1} < +\infty.$$

La fonction $f(x-y)g(y)$ est donc sommable dans $\mathbb{R}^{2n}$, et le théorème de Fubini nous assure que la fonction $y \mapsto f(x-y)g(y)$ est sommable pour presque tout x et que la fonction $(f \star g)(x)$ définie par (3.3) est elle-même sommable. On a enfin

$$\|f \star g\|_{L^1} = \int |(f \star g)(x)| \, dx \leq \iint |f(x-y)g(y)| \, dx \, dy = \|f\|_{L^1} \|g\|_{L^1}.$$

Il ne reste à vérifier que l'égalité de droite dans (3.3) (commutativité du produit de convolution). Or, pour presque chaque x , la fonction $f(x-y)g(y)$ est sommable en y , et on peut effectuer le changement de variable $z = x - y$, dont le jacobien vaut $(-1)^n$. On obtient donc

$$\int f(x-y)g(y) \, dy = \int f(z)g(x-z) \, dz,$$

ce qui achève la démonstration.

Remarque 3.3.2. — Pour tout fermé F , on posera $F_\rho = \{x \mid d(x, F) \leq \rho\}$. Les propriétés suivantes résultent immédiatement de la définition (3.3).

— Si f est nulle (p.p.) hors d'un fermé F et si g est nulle hors d'une boule $B(0, \rho)$, la fonction $f \star g$ est nulle hors de F_ρ .

— Si g est nulle hors de $B(0, \rho)$, la valeur de $f \star g$ en un point x ne dépend que des valeurs de f dans $B(x, \rho)$.

Corollaire 3.3.3. — Soit $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ et soit $g \in L^1(\mathbb{R}^n)$ nulle en dehors d'un compact. La fonction $y \mapsto f(x-y)g(y)$ est alors sommable pour presque tout x et la formule (3.3) définit un élément $f \star g$ de $L^1_{\text{loc}}(\mathbb{R}^n)$.

Soit $\rho > 0$ tel que g soit nulle hors de $B(0, \rho)$. Pour $|x| \leq R$, on a

$$\int_{\mathbb{R}^n} g(x-y)f(y) dy = \int_{B(0, R+\rho)} g(x-y)f(y) dy = ((1_{B(0, R+\rho)}f) \star g)(x),$$

la fonction à intégrer étant nulle hors de $B(0, R+\rho)$. Le membre de droite, qui est le produit de convolution de deux fonctions de $L^1(\mathbb{R}^n)$ est défini pour presque tout x et est sommable. Il en résulte que $f \star g$ est sommable sur toute boule $B(0, R)$ et appartient donc à $L^1_{\text{loc}}(\mathbb{R}^n)$.

Théorème 3.3.4. — Soient $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ et $\varphi \in C_0^m(\mathbb{R}^n)$, $m = 0, 1, \dots, \infty$. La fonction $f \star \varphi$ appartient alors à $C^m(\mathbb{R}^n)$. On a en outre, pour $|\alpha| \leq m$

$$\partial^\alpha (f \star \varphi) = f \star \partial^\alpha \varphi. \quad (3.6)$$

On peut se borner à démontrer le théorème lorsque $f \in L^1(\mathbb{R}^n)$, ce que nous supposerons par la suite. En effet, il suffit de prouver que $f \star \varphi$ est de classe C^m dans chaque boule fixée $B(0, R)$ de $\mathbb{R}^n$, et la démonstration précédente, dont nous gardons les notations, montre que dans cette boule, la fonction $f \star \varphi$ coïncide avec le produit de convolution de φ et de la fonction sommable $1_{B(0, R+\rho)}f$.

Considérons d'abord le cas où $m = 0$. Pour prouver la continuité de $f \star \varphi$, appliquons le théorème de la convergence dominée (théorème 1.2.4) à

$$(f \star \varphi)(x_j) = \int \varphi(x_j - y)f(y) dy, \quad (3.7)$$

où x_j est une suite tendant vers un point x_0 . La fonction φ est continue à support compact et donc bornée. Les fonctions à intégrer dans (3.7) sont donc majorées par la fonction sommable fixe $(\max |\varphi|)|f(y)|$. D'autre part, elles convergent pour chaque y vers $\varphi(x_0 - y)f(y)$ par continuité de φ . On a donc

$$\lim_{j \rightarrow \infty} (f \star \varphi)(x_j) = (f \star \varphi)(x_0),$$

ce qui démontre le théorème pour $m = 0$.

Montrons maintenant (3.6) pour $|\alpha| = 1$. La continuité de $\partial_i(f \star \varphi) = f \star \partial_i \varphi$ résultera alors du cas $m = 0$ déjà étudié, et le cas $|\alpha| \leq m$ quelconque s'obtient par une récurrence évidente. Appliquons le théorème de dérivation sous le signe somme à

$$\frac{\partial}{\partial x_i} \int \varphi(x-y)f(y) dy.$$

La fonction à intégrer est dérivable partout par rapport au “paramètre” x_i et la fonction

$$\partial/\partial x_i \{ \varphi(x-y)f(y) \} = \partial_i \varphi(x-y) f(y) \quad (3.8)$$

est majorée en module par la fonction $(\max |\partial_i \varphi|) |f(y)|$ qui est sommable en y et indépendante de x . On peut donc appliquer le théorème 1.2.6 à (3.8) et on obtient (3.6) pour $|\alpha| = 1$. Cela achève la démonstration du théorème 3.3.4.

Exercice 3.3.5. — Montrer que, pour $f \in C^p(\mathbb{R}^n)$ et $g \in C_0^q(\mathbb{R}^n)$, on a $f \star g \in C^{p+q}(\mathbb{R}^n)$ et que, pour $|\alpha| \leq p$ et $|\beta| \leq q$, on a $\partial^{\alpha+\beta}(f \star g) = \partial^\alpha f \star \partial^\beta g$.

Exercice 3.3.6. — (a) Montrer que, pour deux fonctions f et g appartenant à $L^2(\mathbb{R}^n)$, la formule (3.3) fournit une fonction bornée définie en chaque point, et que l'on a $\|f \star g\|_{L^\infty} \leq \|f\|_{L^2} \|g\|_{L^2}$.

(b) Démontrer, toujours pour f et g dans L^2 , que la fonction $f \star g$ est continue et tend vers 0 à l'infini (en utilisant le théorème 2.6.15(b), on montrera que $f \star g$ est limite uniforme de fonctions continues à support compact).

Exercice 3.3.7. — (a) Montrer que, si $f \in L^2(\mathbb{R}^n)$ et $g \in L^1(\mathbb{R}^n)$, on a $f \star g \in L^2(\mathbb{R}^n)$ et $\|f \star g\|_{L^2} \leq \|f\|_{L^2} \|g\|_{L^1}$. On montrera d'abord que

$$\iiint |f(x-y)| |f(x-z)| |g(y)| |g(z)| \, dx \, dy \, dz \leq \|f\|_{L^2}^2 \|g\|_{L^1}^2$$

en utilisant Cauchy-Schwarz.

(b) Montrer que, si $f \in L^\infty(\mathbb{R}^n)$ et $g \in L^1(\mathbb{R}^n)$, on a $f \star g \in L^\infty(\mathbb{R}^n)$ et $\|f \star g\|_{L^\infty} \leq \|f\|_{L^\infty} \|g\|_{L^1}$.

3.4. Régularisation

Comme nous l'avons vu, si une fonction g est nulle hors de $B(0, \varepsilon)$, la valeur de $f \star g$ au point x ne dépend que des valeurs de f dans $B(x, \varepsilon)$. Si de plus la fonction g est d'intégrale 1 et positive, $f \star g(x)$ devient une moyenne pondérée des valeurs de f dans $B(x, \varepsilon)$. Pour ε petit, on peut donc espérer que $f \star g$ soit proche de f .

3.4.1. Approximations de l'identité. — Soit χ une fonction de classe C^∞ , positive, à support dans la boule unité et d'intégrale égale à 1 (il suffit de multiplier la fonction fournie par le corollaire 3.2.5 par une constante convenable pour ajuster son intégrale). Nous dirons que la famille de fonctions suivante

$$\chi_\varepsilon(x) = \varepsilon^{-n} \chi(x/\varepsilon).$$

est une approximation de l'identité. Les fonctions χ_ε appartiennent à C_0^∞ , sont positives, d'intégrale égale à 1 et ont leur support dans la boule de rayon ε centrée à l'origine.

La terminologie est classique, mais “approximation régulière de δ ” serait une dénomination plus appropriée. Comme nous allons le voir, ce sont les opérateurs de convolution par χ_ε qui convergent vers l’opérateur identité. Nous montrerons plus loin que les χ_ε eux-même convergent au sens des distributions vers la masse de Dirac δ .

Exercice 3.4.2. — Soit K un compact contenu dans un ouvert Ω . Pour ε assez petit, on considère la fonction g égale à 1 si $d(x, K) \leq 2\varepsilon$ et à 0 sinon, et on pose $f = g \star \chi_\varepsilon$. Montrer que l’on obtient ainsi une autre démonstration du corollaire 3.2.6.

Théorème 3.4.3. — Soit χ_ε une approximation de l’identité.

(a) Si f est une fonction uniformément continue dans $\mathbb{R}^n$ (et notamment si f est continue à support compact), les fonctions $f \star \chi_\varepsilon$ convergent uniformément vers f dans $\mathbb{R}^n$.

(b) Pour toute $f \in L^1(\mathbb{R}^n)$, les fonctions $f \star \chi_\varepsilon$ convergent vers f en norme dans L^1 .

(c) Pour toute $f \in L^2(\mathbb{R}^n)$, les fonctions $f \star \chi_\varepsilon$ convergent vers f en norme dans L^2 .

Remarque 3.4.4. — Les fonctions $f \star \chi_\varepsilon$ appartiennent à C^∞ d’après le théorème 3.3.4. Les énoncés ci-dessus, ainsi que ceux du théorème suivant assurent non seulement que les fonctions de classe C^∞ sont denses dans les espaces considérés, mais ils fournissent un procédé systématique d’approximation connu sous le nom de *régularisation*.

On ne cherchera pas à approcher en norme un élément quelconque de L^∞ par ses régularisées, une limite uniforme de fonctions C^∞ étant nécessairement continue.

3.4.5. Démonstration du théorème 3.4.3. — L’intégrale de χ_ε étant égale à 1, on peut écrire

$$(f \star \chi_\varepsilon)(x) - f(x) = \int (f(x-y) - f(x)) \chi_\varepsilon(y) dy.$$

La fonction à intégrer étant nulle hors de $B(0, \varepsilon)$, on a

$$|(f \star \chi_\varepsilon)(x) - f(x)| \leq \sup_{x, z \in \mathbb{R}^n; |x-z| \leq \varepsilon} |f(z) - f(x)|.$$

La fonction f étant uniformément continue, le membre de droite tend vers 0 avec ε , ce qui exprime que $f \star \chi_\varepsilon$ converge vers f uniformément et achève la démonstration de la partie (a).

Nous utiliserons le fait (voir théorème 2.6.8) que pour $f \in L^1$ il existe une suite f_j de fonctions continues à support compact qui converge vers f en norme

dans L^1 . On a

$$\|f \star \chi_\varepsilon - f\|_{L^1} \leq \|(f - f_j) \star \chi_\varepsilon\|_{L^1} + \|f_j \star \chi_\varepsilon - f_j\|_{L^1} + \|f_j - f\|_{L^1}. \quad (3.9)$$

Un nombre $\beta > 0$ étant donné, on peut choisir j assez grand pour que $\|f_j - f\|_{L^1}$ soit inférieur à $\beta/3$. On a alors, d'après (3.4)

$$\|(f - f_j) \star \chi_\varepsilon\|_{L^1} \leq \|f - f_j\|_{L^1} \|\chi_\varepsilon\|_{L^1} = \|f - f_j\|_{L^1} \leq \beta/3.$$

L'indice j étant maintenant fixé, soit K le support de f_j et soit K' l'ensemble des points dont la distance à K est inférieure ou égale à 1. Pour $\varepsilon < 1$ les fonctions $f_j \star \chi_\varepsilon - f_j$ sont à support dans le compact K' et convergent vers 0 uniformément d'après la partie (a) du théorème. Elles convergent donc vers 0 en norme L^1 et on a

$$\|f_j \star \chi_\varepsilon - f_j\|_{L^1} \leq \beta/3$$

pour ε assez petit. Le membre de droite de (3.9) peut être rendu inférieur à tout nombre $\beta > 0$ pour ε assez petit, ce qui achève la démonstration de la partie (b) du théorème.

La démonstration de la partie (c) est identique et est laissée au lecteur à titre d'exercice. On utilisera le fait que l'espace C_0 est dense dans L^2 (théorème 2.6.15(b)) et l'inégalité $\|f \star g\|_{L^2} \leq \|f\|_{L^2} \|g\|_{L^1}$ (exercice 3.3.7).

Théorème 3.4.6. — Soit χ_ε une approximation de l'identité.

(a) Soit f une fonction appartenant à $C^m(\mathbb{R}^n)$, $m = 0, 1, \dots, +\infty$. Pour tout compact K de $\mathbb{R}^n$, et tout multiindice α de longueur $\leq m$, les fonctions $\partial^\alpha(f \star \chi_\varepsilon)$ convergent vers $\partial^\alpha f$ uniformément sur K .

(b) Quels que soient la fonction $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ et le compact $K \subset \mathbb{R}^n$, on a $\|f \star \chi_\varepsilon - f\|_{L^1(K)} \rightarrow 0$ pour $\varepsilon \rightarrow 0$.

(c) Quels que soient la fonction $f \in L^2_{\text{loc}}(\mathbb{R}^n)$ et le compact $K \subset \mathbb{R}^n$, on a $\|f \star \chi_\varepsilon - f\|_{L^2(K)} \rightarrow 0$ pour $\varepsilon \rightarrow 0$.

Considérons d'abord le cas $m = 0$. Soit donc K un compact de $\mathbb{R}^n$ et soit K' l'ensemble des points dont la distance à K est inférieure ou égale à 1. Soit $\psi \in C_0^\infty(\mathbb{R}^n)$ une fonction égale à 1 sur K' , et posons $\tilde{f} = \psi f$.

Pour $x \in K$, on a

$$(f \star \chi_\varepsilon)(x) - f(x) = \int_{|y| \leq 1} (f(x-y) - f(x)) \chi_\varepsilon(y) dy = (\tilde{f} \star \chi_\varepsilon)(x) - \tilde{f}(x) \quad (3.10)$$

en supposant $\varepsilon \leq 1$, les points x et $x - y$ appartenant alors à K' , où les fonctions f et $\tilde{f}$ coïncident. D'après le théorème 3.4.3(a), le membre de droite tend uniformément vers 0 avec ε , ce qui montre que $f \star \chi_\varepsilon$ converge vers f uniformément sur K .

Pour m quelconque, et $|\alpha| \leq m$, on a $\partial^\alpha(f \star \chi_\varepsilon) = (\partial^\alpha f) \star \chi_\varepsilon$ d'après l'exercice 3.3.5. Le résultat démontré pour $m = 0$ appliqué à $\partial^\alpha f$ fournit la convergence uniforme sur K des dérivées de $f \star \chi_\varepsilon$ vers les dérivées correspondantes de f . Cela achève la démonstration de la partie (a).

Soient maintenant $f \in L^1_{\text{loc}}$ et K un compact de $\mathbb{R}^n$, et définissons comme ci-dessus le compact K' et la fonction $\tilde{f}$ qui appartient à $L^1(\mathbb{R}^n)$. Pour $x \in K$, on a $f(x) = \tilde{f}(x)$ et, pourvu que $\varepsilon < 1$, on a $(f \star \chi_\varepsilon)(x) = (\tilde{f} \star \chi_\varepsilon)(x)$. On a donc

$$\|f \star \chi_\varepsilon - f\|_{L^1(K)} = \|\tilde{f} \star \chi_\varepsilon - \tilde{f}\|_{L^1(K)} \leq \|\tilde{f} \star \chi_\varepsilon - \tilde{f}\|_{L^1(\mathbb{R}^n)} \xrightarrow{\varepsilon \rightarrow 0} 0,$$

en utilisant le théorème 3.4.3 pour $\tilde{f}$. Cela achève la démonstration du théorème, le cas de L^2 se traitant de manière absolument identique.

Exercice 3.4.7. — Soit χ une fonction sommable, avec $\int \chi(x) dx = 1$, qu'on ne suppose ni positive, ni à support compact. On pose $\chi_\varepsilon(x) = \varepsilon^{-n} \chi(x/\varepsilon)$, et $f_\varepsilon = f \star \chi_\varepsilon$. Montrer que l'on a $f_\varepsilon \rightarrow f$ dans L^1 . On remarquera que $f_\varepsilon(x) = \int f(x - \varepsilon s) \chi(s) ds$, et on traitera d'abord le cas où $f \in C^0_0$.

Exercice 3.4.8 (distance régularisée de Whitney). — Soit F un fermé de $\mathbb{R}^n$. On considère une fonction χ positive, C^∞ , d'intégrale 1, à support dans la boule $B(0, 1/2)$ et vérifiant $\chi(0) \neq 0$. On pourra introduire $a > 0$ tel que l'on ait $\chi(x) \geq a$ pour $|x| \leq a$. Pour $x \in \mathbb{C}F$, on pose

$$\delta(x) = \int \chi\left(\frac{x-y}{d(y, F)}\right) d(y, F)^{-n+1} dy.$$

(a) Démontrer que δ est de classe C^∞ dans $\mathbb{C}F$ et qu'il existe des constantes $C_k > 0$ telles que l'on ait $C_0^{-1} d(x, F) \leq \delta(x) \leq C_0 d(x, F)$, et $|\partial^\alpha \delta(x)| \leq C_{|\alpha|} d(x, F)^{1-|\alpha|}$.

(b) La fonction Φ étant celle du lemme 3.2.4, montrer que la fonction égale à 0 sur F et à $\Phi \circ \delta$ sur $\mathbb{C}F$ est de classe C^∞ dans $\mathbb{R}^n$ et qu'elle fournit une autre solution à l'exercice 3.2.11.

3.5. Approximation dans un ouvert

Soit Ω un ouvert de $\mathbb{R}^n$ (qui pourra être $\mathbb{R}^n$ lui-même). Nous allons montrer que l'espace $C^\infty_0(\Omega)$ est dense dans de nombreux espaces fonctionnels sur Ω . Il n'est pas possible d'utiliser directement la régularisation : si f est définie dans Ω , la formule $\int_{|y| < \varepsilon} f(x-y) \chi_\varepsilon(y) dy$ n'a de sens que pour les x vérifiant $d(x, \mathbb{C}\Omega) \geq \varepsilon$. Nous devons combiner ce procédé avec des "troncatures" permettant de ne pas déborder de Ω .

Soit K_j une suite exhaustive de compacts de Ω , le corollaire 3.2.6 montre l'existence pour chaque j d'une fonction θ_j de classe C^∞ , vérifiant $0 \leq \theta_j(x) \leq 1$, égale à 1 au voisinage de K_j , et à support dans $\overset{\circ}{K}_{j+1}$. Soit maintenant (ε_j) une suite tendant vers 0 avec $\varepsilon_j > 0$ inférieur à la distance de K_{j+1} à $\mathbb{C}\overset{\circ}{K}_{j+2}$. Posons pour $f \in L^1_{\text{loc}}(\Omega)$

$$R_j f = (f\theta_j) \star \chi_{\varepsilon_j},$$

où χ_ε est une approximation de l'identité. La fonction $R_j f$ est bien définie, de classe C^∞ et à support dans K_{j+2} .

Théorème 3.5.1. — *L'espace $C_0^\infty(\Omega)$ est dense dans $L^1(\Omega)$ et dans $L^2(\Omega)$. Plus précisément, pour $f \in L^1(\Omega)$ [resp. $f \in L^2(\Omega)$], les fonctions $R_j f$ ci-dessus convergent vers f en norme dans L^1 [resp. L^2].*

La fonction $\tilde{f}$ égale à f dans Ω et à 0 dans le complémentaire appartient à $L^1(\mathbb{R}^n)$, et le théorème 3.4.3 assure que $\|\tilde{f} - \tilde{f} \star \chi_\varepsilon\|_{L^1}$ tend vers 0 avec ε . On a

$$\|f - R_j f\|_{L^1(\Omega)} \leq \int_{\Omega} |f - \tilde{f} \star \chi_{\varepsilon_j}| + \int_{\Omega} |(\tilde{f} - f\theta_j) \star \chi_{\varepsilon_j}|.$$

Nous avons vu que la première intégrale, où on peut remplacer f par $\tilde{f}$, tend vers 0. Quant à la seconde, qui est inférieure à l'intégrale sur $\mathbb{R}^n$ entier, elle est majorée d'après (3.4) par $\|\tilde{f} - f\theta_j\|_{L^1} \|\chi_{\varepsilon_j}\|_{L^1}$. Le second facteur est égal à 1, et le premier tend vers 0 d'après le théorème de Lebesgue, ce qui achève la démonstration.

Le cas où $f \in L^2(\Omega)$, qui se traite de manière identique, est laissé au lecteur à titre d'exercice.

Théorème 3.5.2. — (a) *Soit $f \in C^m(\Omega)$, $m = 0, 1, \dots, \infty$. Pour tout compact $K \subset \Omega$ et tout multiindice α de longueur $\leq m$, les $\partial^\alpha(R_j f)$ convergent vers $\partial^\alpha f$ uniformément sur K .*

(b) *Soit $f \in L^1_{\text{loc}}(\Omega)$ [resp. $L^2_{\text{loc}}(\Omega)$]. Alors pour tout compact $K \subset \Omega$, on a $\|f - R_j f\|_{L^1(K)} \rightarrow 0$ [resp. $\|f - R_j f\|_{L^2(K)} \rightarrow 0$].*

Le compact K étant inclus dans un des compacts K_{j_0} de la suite exhaustive, posons $\tilde{f} = \theta_{j_0} f$. Il est clair que l'on a $f = \tilde{f}$ au voisinage de K et il n'est pas difficile de voir que l'on a, pour j assez grand, $R_j f = \tilde{f} \star \chi_{\varepsilon_j}$ sur K . On est alors ramené aux énoncés du théorème 3.4.6.

CHAPITRE 4

LES DISTRIBUTIONS

4.1. Introduction

La théorie des distributions a été introduite par L. Schwartz en 1945. Il s'agit de "fonctions généralisées" auxquelles on peut étendre nombre de concepts de l'analyse, notamment le calcul différentiel. Une des manières de se convaincre de l'utilité et même de la nécessité d'une telle extension est d'avoir une réflexion critique sur l'utilisation des fonctions en physique.

La représentation systématique, introduite au XVIII^e siècle, des phénomènes physiques étendus dans l'espace ou l'espace-temps par des fonctions de plusieurs variables, et l'expression des lois physiques en termes d'équations aux dérivées partielles, ont été un progrès considérable auquel il n'est bien sûr pas question de renoncer. Cependant, cette représentation par une fonction au sens mathématique du terme, assignant une valeur en chaque point, pose problème.

On sait bien, par exemple, que la température en un point, ou même la température moyenne dans un volume de petites dimensions devant le libre parcours moyen des molécules, est dépourvue de sens. Néanmoins, la propagation de la chaleur est régie par une équation aux dérivées partielles que nous étudierons et qui fournit, à l'échelle macroscopique, des résultats conformes à l'expérience.

Une autre critique, qui nous fera progresser, est la suivante. Sans remettre en cause la représentation des phénomènes physiques par des fonctions, il est clair que les valeurs ponctuelles de celles-ci sont inaccessibles à l'expérience. Un appareil de mesure, qui a nécessairement une certaine étendue spatiale, ne pourra jamais fournir la valeur $f(x_0)$ d'une fonction f en un point. Le mieux que l'on puisse en espérer est de nous fournir une moyenne pondérée $\int f(x)\varphi(x)dx$ au lieu de $f(x_0)$, la fonction φ qui caractérise l'appareil ayant

un support “proche” de x_0 et une intégrale proche de 1 pour un appareil précis et bien réglé.

Il apparaît donc raisonnable de considérer comme importantes les quantités suivantes, où φ parcourt l'espace C_0^∞ des fonctions indéfiniment dérivables à support compact (espace des fonctions d'essai) :

$$\langle f, \varphi \rangle = \int f(x) \varphi(x) dx, \quad (4.1)$$

et de faire passer au second plan les valeurs ponctuelles de f . Ce changement de point de vue, si on veut rester cohérent, en induit d'autres :

— Il faut se limiter à des fonctions f localement sommables, pour que les expressions (4.1) aient un sens.

— Il est très raisonnable d'identifier, comme nous l'avons fait dans la section 2.6, deux fonctions f et g égales presque partout. En effet, les expressions (4.1) relatives à f et g sont toujours égales.

— Enfin la vieille notion de convergence ponctuelle $\forall x, f_j(x) \rightarrow f(x)$ doit céder la place à la suivante

$$\forall \varphi \in C_0^\infty, \langle f_j, \varphi \rangle \rightarrow \langle f, \varphi \rangle \quad (4.2)$$

qui représente la convergence des quantités accessibles à l'expérience.

Nous allons maintenant procéder à deux expériences mathématiques qui vont nous amener à élargir le cadre des fonctions.

Exemple 4.1.1. — Considérons une suite de fonctions f_j , positives et d'intégrale égale à 1, à support dans des boules centrées à l'origine de rayon α_j , avec $\lim_{j \rightarrow \infty} \alpha_j = 0$, et examinons leur comportement du point de vue de cette nouvelle notion de convergence. On a :

$$\langle f_j, \varphi \rangle = \int f_j(x) \varphi(0) dx + \int f_j(x) (\varphi(x) - \varphi(0)) dx.$$

La seconde intégrale est majorée en module par $\alpha_j \max |\nabla \varphi|$ et tend donc vers 0, tandis que la première est constante et égale à $\varphi(0)$. Il en résulte donc que, pour toute fonction d'essai φ , on a

$$\lim_{j \rightarrow \infty} \langle f_j, \varphi \rangle = \varphi(0).$$

Du point de vue de cette nouvelle notion de convergence, la suite f_j semble tendre vers “quelque chose”. Toutefois, il n'existe aucune fonction f telle que l'on ait $\langle f, \varphi \rangle = \varphi(0)$ pour tout φ (voir l'exercice 4.4.2). Les notions usuelles de convergence pour les fonctions ne nous apprennent rien : la suite des valeurs $f_j(x)$ tend vers 0 pour tout x différent de 0, et peut avoir n'importe quel comportement pour $x = 0$. Par contre, un physicien, interprétant les f_j comme des densités de masse, aura une idée très claire de la limite : une masse

totale égale à un, concentrée dans des boules de plus en plus petites, converge vers la masse unité située à l'origine.

Nous allons définir l'être mathématique δ (la distribution de Dirac) comme la forme linéaire sur l'espace des fonctions d'essai :

$$\forall \varphi \in C_0^\infty, \langle \delta, \varphi \rangle = \varphi(0).$$

C'est ce concept mathématique, qui n'est pas une fonction mais qui peut être approché par des fonctions, qui correspondra au concept physique de masse ponctuelle.

Exemple 4.1.2. — Nous allons maintenant faire la même expérience, en dimension 1, avec la suite de fonctions g_j définie comme suit : $g_j(x)$ est égal à j^2 dans l'intervalle $]0, 1/j[$, à $-j^2$ dans $] -1/j, 0[$ et à 0 sinon. On a

$$\begin{aligned} \langle g_j, \varphi \rangle &= \int \varphi(0) g_j(x) dx + \int x \varphi'(0) g_j(x) dx \\ &\quad + \int (\varphi(x) - \varphi(0) - x \varphi'(0)) g_j(x) dx. \end{aligned}$$

La première intégrale est nulle, et la seconde est égale à $\varphi'(0)$. Quant à la troisième, elle est majorée en module par $\max |\varphi''|/3j$ et tend donc vers 0. On a

$$\lim_{j \rightarrow \infty} \langle g_j, \varphi \rangle = \varphi'(0).$$

Là encore, on a convergence vers un nouvel être mathématique : la forme linéaire qui à toute fonction d'essai φ associe $\varphi'(0)$. Là encore, la convergence usuelle des fonctions ne donne rien d'intéressant : pour tout x , la suite $g_j(x)$ tend vers 0. Enfin un physicien, placé devant une masse positive égale à j dont le barycentre est le point $1/2j$, et une masse négative $-j$ dont le barycentre est le point $-1/2j$, dira immédiatement que l'objet limite est un dipôle de moment 1 à l'origine.

En résumé, il apparaît que des concepts physiques qui ne se laissent pas représenter par des fonctions, mais qui sont "limites" de tels concepts, se laissent très raisonnablement représenter par des formes linéaires sur l'espace des fonctions d'essai. Une telle forme linéaire u , qui moyennant une propriété de continuité s'appellera une distribution, ne possèdera pas de valeurs ponctuelles, mais aura des "valeurs moyennes pondérées" : les quantités $\langle u, \varphi \rangle$.

4.2. Définition et convergence

Soit Ω un ouvert de $\mathbf{R}^n$. On appelle fonctions d'essai⁽¹⁾ dans Ω les éléments de l'espace $C_0^\infty(\Omega)$ des fonctions indéfiniment dérivables et à support compact dans Ω . Rappelons (n°3.2.2) que, pour K compact dans Ω , on note C_K^∞ l'espace des fonctions d'essai à support dans K (c'est-à-dire nulles hors de K).

Définition 4.2.1. — On dit que u est une distribution dans l'ouvert Ω si u est une forme linéaire sur $C_0^\infty(\Omega)$ qui vérifie la propriété de continuité suivante : pour tout compact K de Ω il existe un entier p et une constante C tels que

$$\forall \varphi \in C_K^\infty, \quad |\langle u, \varphi \rangle| \leq C \sup_{x \in K, |\alpha| \leq p} |\partial^\alpha \varphi(x)|. \quad (4.3)$$

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

Lorsque l'entier p peut être choisi indépendamment de K on dit que la distribution u est d'ordre fini, et la plus petite valeur de p possible est appelée l'ordre de u .

La majoration (4.3) exprime bien une continuité de u : pour une suite (φ_j) de fonctions d'essai ayant leur support dans un même compact K , si φ_j converge vers φ uniformément ainsi que toutes ses dérivées, il résulte facilement de la majoration (4.3) que $\langle u, \varphi_j \rangle$ converge vers $\langle u, \varphi \rangle$. La réciproque est vraie, on pourra le démontrer à titre d'exercice, ou se reporter à l'appendice C (théorème C.1.7 et n°C.2.2).

Nous utiliserons systématiquement le crochet de dualité pour marquer l'action de la distribution u sur la fonction d'essai φ . Nous écrirons parfois $\langle u(x), \varphi(x) \rangle$, où x est une lettre muette, au lieu de $\langle u, \varphi \rangle$, bien que les valeurs ponctuelles $u(x)$ n'aient en soi aucun sens, par exemple lorsque φ dépend aussi d'autres variables jouant le rôle de paramètres. Nous n'utiliserons pas la notation $\int u(x)\varphi(x) dx$, pourtant courante en physique et aussi en mathématiques, mais qui nécessite une bonne expérience pour être utilisée sans erreur.

Définition 4.2.2 (Convergence des distributions)

On dit qu'une suite u_j de distributions dans Ω converge (dans $\mathcal{D}'(\Omega)$) vers u si, pour toute fonction φ dans $C_0^\infty(\Omega)$, on a

$$\lim_{j \rightarrow \infty} \langle u_j, \varphi \rangle = \langle u, \varphi \rangle$$

Bien entendu, on peut définir la limite u , lorsque λ tendant vers λ_0 , d'une famille u_λ de distributions, où λ parcourt un espace métrique. Cela signifie

⁽¹⁾ On dit parfois fonction-test (test-function en anglais). L'espace $C_0^\infty(\Omega)$ est souvent noté $\mathcal{D}(\Omega)$.

que pour toute suite λ_j tendant vers λ_0 la suite u_{λ_j} tend vers u dans $\mathcal{D}'$, c'est-à-dire que $\langle u_{\lambda_j}, \varphi \rangle \rightarrow \langle u, \varphi \rangle$ pour tout φ . Nous ne chercherons pas ici à définir de limite pour des espaces topologiques plus généraux.

La vérification de la condition (4.3) peut parfois paraître pénible. La propriété suivante, que nous admettrons, permet souvent d'en faire l'économie.

Théorème 4.2.3. — Soit u_j une suite de distributions dans Ω . On suppose que pour tout $\varphi \in C_0^\infty(\Omega)$ la suite numérique $\langle u_j, \varphi \rangle$ converge vers une limite. Alors, si on note $\langle u, \varphi \rangle$ cette limite, u appartient à $\mathcal{D}'(\Omega)$.

Il est clair que la linéarité de u ne présente pas de difficulté. Il n'en est pas de même de la continuité (aucune uniformité n'est requise a priori sur les constantes C et p de (4.3) relatives aux u_j), pour laquelle nous renvoyons au n°C.3.4.

4.2.4. Premiers exemples

Fonctions localement sommables. — Conformément à ce que nous avons vu dans l'introduction, nous associerons à une telle fonction f , définie dans Ω , une distribution que nous noterons encore f par

$$\langle f, \varphi \rangle = \int f(x) \varphi(x) dx.$$

On voit facilement que, pour K compact dans Ω , et pour $\varphi \in C_K^\infty$, on a

$$|\langle f, \varphi \rangle| \leq \left(\int_K |f(x)| dx \right) \left(\sup_{x \in K} |\varphi(x)| \right).$$

Cela prouve que f définit une distribution, et que celle-ci est d'ordre 0. Il faut bien voir que l'utilisation de la même notation pour la fonction et la distribution qu'elle définit serait très dangereuse s'il advenait que deux fonctions non égales (presque partout) définissent la même distribution. Nous verrons (théorème 4.4.1) que ce n'est heureusement pas le cas.

Masses ponctuelles. — La distribution de Dirac au point a de $\mathbf{R}^n$ est définie par

$$\langle \delta_a, \varphi \rangle = \varphi(a).$$

Il est clair que c'est une distribution d'ordre 0. On note plus simplement δ cette distribution lorsque a est l'origine.

Dipôles et multipôles. — Il s'agit des distributions du type suivant, ou de combinaisons linéaires de celles-ci :

$$\langle u, \varphi \rangle = \partial^\alpha \varphi(a)$$

où a appartient à $\mathbb{R}^n$ et α à $\mathbb{N}^n$. On voit facilement que u est une distribution d'ordre inférieur ou égal à $|\alpha|$.

Exercice 4.2.5. — Démontrer que l'ordre de la distribution u ci-dessus est exactement $|\alpha|$. On pourra considérer les fonctions $\varphi_\varepsilon(x) = x^\alpha \psi((x - a)/\varepsilon)$, où $\psi \in C_0^\infty$ et est égale à 1 au voisinage de l'origine.

Nous avons vu dans l'exemple 4.1.2 une suite de fonctions convergeant vers un dipôle. L'exemple 4.1.1 démontrait en fait un résultat utile de convergence vers δ que nous énonçons ci-dessous.

Théorème 4.2.6. — *Soit f_j une suite de fonctions sommables positives, d'intégrale 1, dont les supports sont contenus dans des boules centrées à l'origine et de rayon tendant vers 0. On a alors $\lim_{j \rightarrow \infty} f_j = \delta$ au sens des distributions.*

Nous allons dès maintenant mettre en évidence la raison essentielle de l'introduction des distributions. On peut non seulement étendre le calcul différentiel à ces nouveaux objets, mais on obtient un calcul beaucoup plus facile à manier qu'il ne l'était sur les fonctions.

4.3. Dérivées

Considérons une fonction f de classe C^1 dans un ouvert Ω de $\mathbb{R}^n$, et regardons l'action de ses dérivées partielles sur les fonctions d'essai. On a

$$\langle \partial f / \partial x_i, \varphi \rangle = \int \frac{\partial f}{\partial x_i}(x) \varphi(x) dx = - \int f(x) \frac{\partial \varphi}{\partial x_i}(x) dx = - \langle f, \partial \varphi / \partial x_i \rangle.$$

Cette dernière formule conserve un sens si on remplace f par une distribution. Nous allons donc pouvoir définir les dérivées d'une distribution quelconque u , ces dérivées gardant leur sens habituel dans le cas où u est en fait une fonction de classe C^1 .

Définition 4.3.1. — *Soit u appartenant à $\mathcal{D}'(\Omega)$. On note $\partial u / \partial x_i$ la distribution définie par*

$$\langle \partial u / \partial x_i, \varphi \rangle = - \langle u, \partial \varphi / \partial x_i \rangle. \quad (4.4)$$

Il faut vérifier que la forme linéaire $\partial u / \partial x_i$ est bien une distribution, c'est à dire que l'on a l'estimation (4.3). En fait, pour $\varphi \in C_R^\infty(\Omega)$, on a

$$|\langle \partial u / \partial x_i, \varphi \rangle| \leq C \sup_{x \in K, |\alpha| \leq p} \left| \partial^\alpha \frac{\partial \varphi}{\partial x_i}(x) \right| \leq C \sup_{x \in K, |\beta| \leq p+1} \left| \partial^\beta \varphi(x) \right| \quad (4.5)$$

où C et p sont les constantes de (4.3) relatives à u et K . La relation (4.5) exprime que $\partial u / \partial x_i$ est une distribution, et en outre qu'elle est d'ordre au plus $p + 1$ lorsque u est d'ordre fini p .

Remarque 4.3.2. — Il est naturel de s'étonner du fait que la dérivation soit toujours possible, et cet étonnement devrait grandir au vu du théorème 4.3.4 ci-dessous. On sait bien qu'il existe des fonctions localement sommables et même continues qui ne sont pas dérivables. Pour le moment la réponse est facile : dans un tel cas, les dérivées sont des distributions qui ne sont pas des fonctions. Cela dit, il faudra apprendre à calculer ces dérivées, au moins dans des cas simples. Ce sera l'objet de l'exemple 4.3.5 ci-dessous, et des sections 5.3 et 5.4.

4.3.3. Dérivées successives. — Si on prend la dérivée par rapport à x_j de $\partial u / \partial x_i$ on obtient le même résultat qu'en permutant i et j . En effet

$$\langle \partial^2 u / \partial x_i \partial x_j, \varphi \rangle = \langle u, \partial^2 \varphi / \partial x_j \partial x_i \rangle,$$

et, d'après le lemme de Schwarz, le membre de droite est invariant par permutation de i et j . Plus généralement, on a

$$\langle \partial^\alpha u, \varphi \rangle = (-1)^{|\alpha|} \langle u, \partial^\alpha \varphi \rangle.$$

Par exemple, le "multipôle" u considéré précédemment, défini par $\langle u, \varphi \rangle = \partial^\alpha \varphi(a)$ est égal à $(-1)^{|\alpha|} \partial^\alpha \delta_a$.

Théorème 4.3.4. — *Si une suite u_j d'éléments de $\mathcal{D}'(\Omega)$ converge vers u (au sens des distributions), alors la suite $\partial^\alpha u_j$ converge vers $\partial^\alpha u$. On peut en particulier dériver terme à terme toute série convergente (c'est-à-dire dont les sommes partielles convergent) au sens des distributions.*

On a en effet, pour toute fonction $\varphi \in C_0^\infty(\Omega)$,

$$\langle \partial^\alpha u_j, \varphi \rangle = (-1)^{|\alpha|} \langle u_j, \partial^\alpha \varphi \rangle.$$

Le membre de droite, pour $j \rightarrow \infty$, tend vers $(-1)^{|\alpha|} \langle u, \partial^\alpha \varphi \rangle$ qui est égal à $\langle \partial^\alpha u, \varphi \rangle$. La convergence du membre de gauche vers cette quantité exprime exactement la convergence des $\partial^\alpha u_j$ vers $\partial^\alpha u$ au sens des distributions.

Cette propriété est d'une importance capitale. La dérivation devient une opération toujours définie et continue, alors qu'il est bien connu qu'une telle propriété est complètement fautive pour la dérivation usuelle des fonctions. L'exemple très simple qui va suivre explique bien pourquoi : les dérivées usuelles peuvent laisser échapper l'essentiel (une masse de Dirac par exemple) de la "vraie dérivée".

On a représenté ci-contre une fonction régulière voisine de la fonction de Heaviside et ses deux premières dérivées.

La première présente un pic étroit, de masse totale égale à 1, et constitue une bonne approximation de la masse de Dirac. La seconde montre un pic positif situé à proximité et à gauche d'un pic négatif. Un œil exercé n'a aucun mal à y voir un dipôle de moment négatif.

Seules, la convergence et la dérivation au sens des distributions peuvent rendre compte de la situation limite.

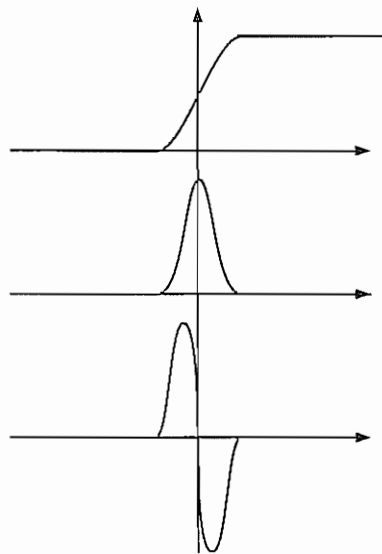


FIGURE 1. approximations de H et de ses dérivées

Exemple 4.3.5. — Calculons les dérivées d'ordre 1 et 2 de la *fonction de Heaviside* H définie comme suit : $H(x) = 1$ pour $x > 0$, et $H(x) = 0$ sinon. On a

$$\langle H', \varphi \rangle = - \int H(x) \varphi'(x) dx = - \int_0^{\infty} \varphi'(x) dx = \varphi(0), \quad (4.6)$$

$$\langle H'', \varphi \rangle = + \int H(x) \varphi''(x) dx = -\varphi'(0). \quad (4.7)$$

On a donc $\partial H / \partial x = \delta$ d'après (4.6). Quant à l'équation (4.7), elle exprime précisément que la distribution $H'' = \delta'$ est un dipôle dont le moment est égal à -1 . La figure 1 montre en quoi ces résultats sont naturels. La dérivée de H au sens des fonctions (0 hors de l'origine et, si on veut, $+\infty$ à l'origine) ne rend pas compte de la variation de H . Par contre, affirmer que la dérivée de H est δ et que celle de $2H$ est 2δ donne tous les renseignements sur leurs variations.

4.4. Exemples de distributions

A. Fonctions localement sommables

Le théorème suivant va nous permettre d'identifier complètement l'espace L^1_{loc} à un sous espace de $\mathcal{D}'$.

Théorème 4.4.1. — Soient f et g deux fonctions localement sommables dans un ouvert Ω de $\mathbb{R}^n$. Les deux propriétés suivantes sont équivalentes.

- (a) $f(x) = g(x)$ presque partout dans Ω .
- (b) $\int f(x)\varphi(x) dx = \int g(x)\varphi(x) dx$ pour tout $\varphi \in C_0^\infty(\Omega)$.

En introduisant $h = f - g$, il faut donc montrer que h est nulle presque partout, sachant que le produit de h par un élément de C_0^∞ est toujours d'intégrale nulle.

Soit K un compact de Ω . D'après le théorème 3.5.2, on peut approcher h en norme L^1 sur K par les fonctions $R_j h$, et celles-ci vérifient pour $x \in K$ et pourvu que j soit assez grand

$$R_j h(x) = \int h(x-y)\chi_{\varepsilon(j)}(y) dy = \int h(z)\chi_{\varepsilon(j)}(x-z) dz.$$

Cette dernière intégrale du produit de $h(z)$ par un élément de C_0^∞ est nulle par hypothèse. La fonction h qui est limite en norme dans $L^1(K)$ de fonctions nulles sur K est donc nulle presque partout sur K et, Ω étant réunion dénombrable de tels compacts, nulle presque partout sur Ω .

Exercice 4.4.2. — Montrer qu'il n'existe pas de fonction localement sommable f telle que l'on ait $\langle f, \varphi \rangle = \varphi(0)$ pour tout $\varphi \in C_0^\infty$.

Remarque 4.4.3. — La relation d'équivalence " $f = g$ p.p." est exactement la même que la relation d'équivalence " f et g définissent la même distribution". Nous identifierons toujours par la suite l'espace quotient L_{loc}^1 avec l'espace des distributions qu'il définit, de même que nous identifierons l'espace C^0 avec son image dans L_{loc}^1 (deux fonctions continues égales presque partout sont égales). On a donc, par exemple :

$$C_0^\infty(\Omega) \subset C^0(\Omega) \subset L_{loc}^1(\Omega) \subset \mathcal{D}'(\Omega),$$

ce qui donne un sens précis à des expressions du genre : telle distribution u est de classe $C^2 \dots$. Une situation de même nature est bien connue : on écrit habituellement $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$, plutôt que les homomorphismes injectifs canoniques résultant des diverses constructions.

Théorème 4.4.4. — Soit f_j une suite d'éléments de $L_{loc}^1(\Omega)$ convergeant vers f en norme L^1 sur chaque compact. La suite f_j converge alors vers f au sens des distributions.

En particulier, si $f_j(x) \rightarrow f(x)$ p.p. et s'il existe une fonction $g \in L_{loc}^1(\Omega)$ positive telle que l'on ait $|f_j(x)| \leq g(x)$ p.p. pour tout j , la suite f_j tend vers f au sens des distributions.

On a en effet, pour $\varphi \in C_K^\infty$,

$$|\langle f_j, \varphi \rangle - \langle f, \varphi \rangle| \leq \max |\varphi(x)| \int_K |f_j(x) - f(x)| dx \longrightarrow 0.$$

La seconde partie résulte du théorème de Lebesgue : sur chaque compact K , les f_j convergent presque partout et sont majorées par la fonction sommable $1_K g$, et convergent donc dans $L^1(K)$.

Rappelons enfin, pour des suites de fonctions appartenant à L_{loc}^2 (ou L_{loc}^∞), que la convergence dans L^2 (ou L^∞) sur chaque compact entraîne la convergence dans L^1 sur chaque compact, et donc la convergence au sens des distributions.

B. Mesures de Radon

Théorème et Définition 4.4.5. — Soit $\varphi \mapsto \langle u, \varphi \rangle$ une forme linéaire positive sur $C_0^\infty(\Omega)$ (c'est-à-dire telle que $\varphi \geq 0$ entraîne $\langle u, \varphi \rangle \geq 0$). Alors u est une distribution d'ordre 0 sur Ω .

On appelle mesure de Radon positive une telle distribution, et on appelle mesure de Radon une combinaison linéaire $u_1 - u_2 + iu_3 - iu_4$ de telles mesures positives.

Remarquons d'abord qu'une forme linéaire positive est croissante : si $\varphi \leq \psi$, on a $\langle u, \varphi \rangle \leq \langle u, \psi \rangle$. Soit maintenant K un compact quelconque de Ω et (théorème 3.2.6) soit $\psi \in C_0^\infty(\Omega)$ positive et égale à 1 sur K . Pour $\varphi \in C_K^\infty$, à valeurs réelles, on a

$$-(\max |\varphi|)\psi \leq \varphi \leq (\max |\varphi|)\psi$$

et donc

$$-(\max |\varphi|) \langle u, \psi \rangle \leq \langle u, \varphi \rangle \leq (\max |\varphi|) \langle u, \psi \rangle.$$

Pour chaque compact K nous avons donc trouvé une constante $C = \langle u, \psi \rangle$ telle que

$$\forall \varphi \in C_K^\infty, |\langle u, \varphi \rangle| \leq C \max |\varphi|. \quad (4.8)$$

Cette inégalité, qui s'étend immédiatement aux fonctions à valeurs complexes, exprime précisément (voir définition 4.2.1) que u est une distribution d'ordre 0.

4.4.6. Mesures boréliennes et mesures de Radon. — La définition donnée ci-dessus n'est pas la définition classique (les travaux de F. Riesz et de Radon datent de 1909–1913), mais elle lui est équivalente. Remarquons d'abord qu'il est équivalent de se donner une distribution d'ordre 0, ou une forme linéaire sur l'espace $C_0^0(\Omega)$ vérifiant la propriété de continuité suivante : pour tout compact K il existe C tel que

$$\forall \varphi \in C_K^0, |\langle u, \varphi \rangle| \leq C \max |\varphi|. \quad (4.9)$$

En effet, une telle forme linéaire définit évidemment une distribution d'ordre 0 par restriction à C_0^∞ . Il suffit donc de montrer qu'une distribution u d'ordre 0 se prolonge de manière unique en une forme linéaire sur C_0^0 possédant la propriété ci-dessus. Pour K compact, et pour K' voisinage compact de K , toute fonction $f \in C_K^0$ peut être approchée uniformément (régularisation) par une suite de fonctions $\varphi_j \in C_{K'}^\infty$. La suite numérique $\langle u, \varphi_j \rangle$ est alors une suite de Cauchy (utiliser (4.8) dans K') et est donc convergente. Il reste à vérifier que la limite obtenue ne dépend pas de la suite φ_j et que l'on obtient ainsi une forme linéaire continue au sens de (4.9). Cela est facile, toute la démonstration n'étant qu'une petite variation (passage de K à K') sur le thème du prolongement des applications uniformément continues définies sur un sous-ensemble dense.

Soit maintenant μ une mesure borélienne (fonction d'ensemble dénombrablement additive définie sur la plus petite tribu contenant les ouverts) positive sur Ω telle que l'on ait $\mu(K) < \infty$ pour tout compact K . Les fonctions continues à support compact sont alors μ -sommables. La mesure μ définit ainsi la forme linéaire positive suivante sur l'espace C_0^0

$$\langle \mu, f \rangle = \int f(x) d\mu(x).$$

Un résultat célèbre (le "théorème de représentation" de F. Riesz) assure que la réciproque est vraie : toute forme linéaire positive sur $C_0^0(\Omega)$ provient d'une mesure positive finie sur les compacts, et toute forme linéaire sur cet espace possédant la propriété (4.9) provient de $\mu_1 - \mu_2 + i\mu_3 - i\mu_4$ avec les μ_j comme ci-dessus.

Il est donc équivalent de parler de distributions d'ordre 0, de formes linéaires continues sur C_0^0 au sens de (4.9) (ce que l'on appelle habituellement mesure de Radon) et de mesures ensemblistes finies sur les compacts. Nous n'aurons à utiliser que le résultat (facile) du théorème 4.4.5.

Comme exemple de mesure de Radon nous avons déjà rencontré les masses ponctuelles (mesures de Dirac) et leurs combinaisons linéaires qui sont les mesures les plus "concentrées". À l'inverse, les mesures les plus "étalées" sont celles qui ont une densité f (localement sommable) par rapport à la mesure de Lebesgue : elles sont définies par $\langle \mu, \varphi \rangle = \int f(x)\varphi(x) dx$ et ne se distinguent donc pas de la fonction f (identifiée à une distribution) elle-même. Les cas intermédiaires de mesures de Radon concentrées sur des sous-variétés (courbes, surfaces, ...) sont intéressants.

4.4.7. Distributions de simple couche. — Considérons par exemple le plan d'équation $z = 0$ dans $\mathbb{R}^3$ et soit $f(x, y)$ une fonction localement sommable dans $\mathbb{R}^2$. On appelle distribution de simple couche de densité superficielle f la distribution $\mu \in \mathcal{D}'(\mathbb{R}^3)$ définie par

$$\forall \varphi \in C_0^\infty(\mathbb{R}^3), \quad \langle \mu, \varphi \rangle = \iint_{\mathbb{R}^2} f(x, y)\varphi(x, y, 0) dx dy. \quad (4.10)$$

Il est clair que μ est positive si f est positive, sinon il suffit de décomposer f en $f_1 - f_2 + if_3 - if_4$ pour voir que μ est une mesure de Radon.

Plus généralement, si Σ est une hypersurface de $\mathbb{R}^n$ dont on note $d\sigma$ la mesure de surface et si g est une fonction définie sur Σ telle que l'on ait

$\int_{\Sigma \cap K} |g(m)| d\sigma_m < \infty$ pour tout compact K de $\mathbb{R}^n$, on définit la distribution de simple couche notée $g d\sigma$ par

$$\langle g d\sigma, \varphi \rangle = \int_{\Sigma \cap \text{Supp}(\varphi)} g(m) \varphi(m) d\sigma_m.$$

On peut également définir, par exemple, la mesure ν ayant la densité $g(z)$ (localement sommable sur $\mathbb{R}$) par rapport à la mesure de longueur, portée par l'axe des z :

$$\forall \varphi \in C_0^\infty(\mathbb{R}^3), \quad \langle \nu, \varphi \rangle = \int g(z) \varphi(0, 0, z) dz,$$

et on peut remplacer l'axe des z par une courbe quelconque.

Il existe, même en dimension 1, des mesures de Radon nettement plus compliquées, mais nous n'aurons pas à les utiliser dans la suite de ce cours.

C. Multipôles, couches multiples

D'après le théorème 4.4.5, des “masses” positives ne peuvent constituer que des distributions d'ordre 0. Les distributions plus singulières feront apparaître des compensations entre masses infinies de signe contraire. Nous avons déjà rencontré au n°4.2.4 les multipôles définis comme les combinaisons de distributions u du type $\langle u, \varphi \rangle = \partial^\alpha \varphi(a)$, et nous savons maintenant que l'on a $u = (-1)^{|\alpha|} \partial^\alpha \delta_a$. Il est possible de définir des “répartitions de multipôles” sur des sous-variétés, nous ne décrivons qu'un cas simple à titre d'exemple.

4.4.8. Distribution de double couche. — Considérons le plan P d'équation $z = 0$ dans $\mathbb{R}^3$, et soit $f(x, y)$ localement sommable dans $\mathbb{R}^2$. La distribution u suivante représente une répartition de dipôles orientés selon l'axe des z , portés par P avec la densité f

$$\forall \varphi \in C_0^\infty(\mathbb{R}^3), \quad \langle u, \varphi \rangle = \iint f(x, y) \frac{\partial \varphi}{\partial z}(x, y, 0) dx dy.$$

Exercice 4.4.9. — Vérifier que u est une distribution d'ordre 1, et que u est limite (ce qui justifie la terminologie “double couche”), pour $\varepsilon \rightarrow 0$, des distributions constituées d'une simple couche de densité $\varepsilon^{-1}f$ portée par le plan $z = \varepsilon$, et d'une simple couche de densité $-\varepsilon^{-1}f$ portée par le plan $z = 0$.

Exercice 4.4.10. — Montrer que l'on a $u = -\partial \mu / \partial z$, où μ est la simple couche définie par (4.10). En supposant la densité f de classe C^1 montrer que, au contraire, les dérivées tangentielles $\partial \mu / \partial x$ et $\partial \mu / \partial y$ sont des distributions de simple couche.

D. Valeurs principales et parties finies

Une fonction non localement sommable n'est pas une distribution. Toutefois, certaines distributions dont nous verrons l'intérêt par la suite sont associées, par des procédures qu'il faut expliciter, à des fonctions non localement sommables. Pour simplifier, nous ne considérerons que le cas de la dimension 1.

4.4.11. La distribution $\text{vp}(1/x)$. — La fonction f_ε , égale à $1/x$ pour $|x| \geq \varepsilon$ et à 0 sinon, est localement sommable et définit donc une distribution sur $\mathbb{R}$. Montrons que les f_ε convergent au sens des distributions pour $\varepsilon \rightarrow 0$. Pour $R > 0$ et pour $\varphi \in C_0^\infty$ à support dans $[-R, +R]$, on a

$$\langle f_\varepsilon, \varphi \rangle = \int_{\varepsilon \leq |x| \leq R} \frac{\varphi(x)}{x} dx = \int_{\varepsilon \leq |x| \leq R} \frac{\varphi(x) - \varphi(0)}{x} dx.$$

Dans le membre de droite, la fonction à intégrer est majorée en module par $\max |\varphi'(x)|$, qui est sommable sur $[-R, +R]$. L'intégrale a donc une limite (à savoir l'intégrale de la même expression sur $[-R, +R]$) lorsque ε tend vers 0. On définit ainsi une forme linéaire

$$\left\langle \text{vp} \left(\frac{1}{x} \right), \varphi \right\rangle = \lim_{\varepsilon \rightarrow 0} \int_{|x| \geq \varepsilon} \frac{\varphi(x)}{x} dx$$

qui, d'après le théorème 4.2.3, est automatiquement une distribution, et que l'on appelle *valeur principale* de $1/x$. L'exercice suivant montre qu'il est indispensable, dans la définition, de préciser que l'on intègre dans le complémentaire de voisinages *symétriques* de l'origine. La considération d'intégrales en valeurs principales remonte à Cauchy.

Exercice 4.4.12. — Soit g_ε [resp. h_ε] la fonction égale à 0 dans l'intervalle $[-\varepsilon, 2\varepsilon]$ [resp. $[-\varepsilon, \varepsilon^2]$] et à $1/x$ en dehors. Démontrer que, lorsque ε tend vers 0, les g_ε tendent vers une limite que l'on déterminera alors que les h_ε n'ont pas de limite dans $\mathcal{D}'(\mathbb{R})$.

Exercice 4.4.13. — Démontrer que la distribution $\text{vp}(1/x)$ est d'ordre ≤ 1 , puis qu'elle est d'ordre 1 exactement (on pourra calculer son action sur la suite $\varphi_j(x) = \varphi_0(x) \arctg(jx)$, où φ_0 appartient à C_0^∞ , est paire, et vaut 1 près de 0).

4.4.14. Exemples de parties finies. — L'utilisation de parties finies d'intégrales divergentes remonte à Liouville, et Hadamard a montré leur importance en théorie des équations aux dérivées partielles. Nous allons voir comment associer des distributions aux fonctions x_+^α (en notant $x_+ = \max(x, 0)$) qui ne sont pas localement sommables pour $\alpha \leq -1$.

Exercice 4.4.15. — Pour $-2 < \alpha < -1$ montrer que, pour $\varphi \in C_0^\infty(\mathbb{R})$, on a

$$\int_{\varepsilon}^{\infty} x^{\alpha} \varphi(x) dx = A\varepsilon^{\alpha+1} + R_{\varepsilon}$$

où A dépend de φ mais pas de ε , et où R_{ε} tend vers une limite pour $\varepsilon \rightarrow 0$. On pose

$$\langle \text{pf}(x_+^{\alpha}), \varphi \rangle = \lim_{\varepsilon \rightarrow 0} R_{\varepsilon}. \quad (4.11)$$

Montrer que $\text{pf}(x_+^{\alpha})$ est une distribution d'ordre 1.

Exercice 4.4.16. — Montrer que, plus généralement, pour tout $\alpha \leq -1$, on peut écrire

$$\int_{\varepsilon}^{\infty} x^{\alpha} \varphi(x) dx = P(\varepsilon) + R_{\varepsilon}$$

où $P(\varepsilon)$ est une combinaison linéaire de puissances strictement négatives de ε (et, dans le cas particulier où α est un entier négatif, de $\log \varepsilon$), et où R_{ε} tend vers une limite pour $\varepsilon \rightarrow 0$. Montrer que $\text{pf}(x_+^{\alpha})$ définie encore par (4.11) est une distribution. (On pourra montrer que son ordre est égal à la partie entière de $-\alpha$).

CHAPITRE 5

OPÉRATIONS SUR LES DISTRIBUTIONS

Il est possible d'étendre aux distributions, de manière raisonnable, beaucoup d'opérations définies sur les fonctions. La première exigence pour une telle extension est bien sûr que la nouvelle définition coïncide avec l'ancienne pour les fonctions. Quant au caractère "raisonnable" de l'extension, il sera garanti par la continuité au sens des distributions des applications ainsi définies.

5.1. Opérations élémentaires

5.1.1. Restriction. — S'il est possible de définir la restriction d'une fonction à un sous-ensemble A quelconque, la restriction d'un élément de L^1_{loc} à A n'est définie que modulo les fonctions nulles presque partout sur A , concept dépourvu d'intérêt lorsque A est de mesure nulle. La restriction d'une distribution ne pourra en général être définie que sur un ensemble ouvert.

Soient donc $\omega \subset \Omega$ deux ouverts de $\mathbb{R}^n$, et $u \in \mathcal{D}'(\Omega)$. Les éléments de $C_0^\infty(\omega)$ appartenant à $C_0^\infty(\Omega)$, on définit la restriction $u|_\omega \in \mathcal{D}'(\omega)$ par

$$\forall \varphi \in C_0^\infty(\omega), \quad \langle u|_\omega, \varphi \rangle = \langle u, \varphi \rangle.$$

5.1.2. Conjugaison complexe. — Si $u \in \mathcal{D}'(\Omega)$, on définit la distribution complexe conjuguée $\bar{u}$ par

$$\forall \varphi \in C_0^\infty(\Omega), \quad \langle \bar{u}, \varphi \rangle = \overline{\langle u, \bar{\varphi} \rangle}.$$

Il est clair en effet que cette définition est la bonne lorsque u est en fait une fonction localement sommable. On dit que u est réelle si $u = \bar{u}$ et on définit les parties réelle et imaginaire d'une distribution par $\text{Re } u = (u + \bar{u})/2$ et $\text{Im } u = (u - \bar{u})/2i$.

Le lecteur pourra vérifier que les distributions réelles ainsi définies s'identifient aux formes $\mathbb{R}$ -linéaires sur l'espace des fonctions à valeurs réelles de C_0^∞ qui satisfont à la condition de continuité (4.3).

5.1.3. Translation. — Pour une fonction f , sa translatée de vecteur $a \in \mathbb{R}^n$ est la fonction $\tau_a f(x) = f(x - a)$. Compte tenu de la relation évidente $\int f(x - a)\varphi(x) dx = \int f(x)\varphi(x + a) dx$, il devient naturel de définir la translatée d'une distribution $u \in \mathcal{D}'(\mathbb{R}^n)$ comme suit :

$$\forall \varphi \in C_0^\infty(\mathbb{R}^n), \langle \tau_a u, \varphi \rangle = \langle u, \tau_{-a} \varphi \rangle.$$

5.1.4. Dilatation. — La fonction dilatée de f dans le rapport $\lambda \neq 0$ est la fonction f_λ définie par $f_\lambda(x) = f(x/\lambda)$. L'intégrale $\int f(x/\lambda)\varphi(x) dx$ est égale à $\int f(y)\varphi(\lambda y) |\lambda|^n dy$, expression généralisable aux distributions quelconques. Pour $u \in \mathcal{D}'(\mathbb{R}^n)$ on notera u_λ (ou symboliquement $u(x/\lambda)$), la distribution définie par

$$\langle u_\lambda, \varphi \rangle = |\lambda|^n \langle u, \varphi_{1/\lambda} \rangle.$$

Un cas particulier important est celui de la symétrie par rapport à l'origine ($\lambda = -1$). On pose alors $\tilde{f}(x) = f(-x)$ et on définit la *symétrique* $\tilde{u}$ d'une distribution u par $\langle \tilde{u}, \varphi \rangle = \langle u, \tilde{\varphi} \rangle$. On dit que u est *paire* [resp. *impaire*] si on a $u = \tilde{u}$ [resp. $u = -\tilde{u}$].

Nous laissons au lecteur le soin de vérifier que les opérations précédentes définissent bien des distributions, et que si $u_j \rightarrow u$ au sens des distributions, le résultat des opérations ci-dessus relatif à u_j converge de même vers le résultat correspondant pour u .

Exercice 5.1.5. — Montrer que, pour $u \in \mathcal{D}'(\mathbb{R})$, sa dérivée u' est la limite, pour $h \rightarrow 0$, des distributions $h^{-1}(\tau_{-h}u - u)$. Énoncer et démontrer l'analogue pour les dérivées partielles dans $\mathbb{R}^n$.

Exercice 5.1.6. — Soit F une fonction croissante définie sur $\mathbb{R}$. Démontrer que sa dérivée au sens des distributions F' est une mesure de Radon positive.

Exercice 5.1.7. — On dit qu'une distribution dans $\mathbb{R}^n$ est (*positivement*) *homogène de degré k* si on a $u_\lambda = \lambda^{-k}u$ pour tout $\lambda > 0$ (ou, avec l'autre notation des distributions dilatées, $u(\mu x) = \mu^k u(x)$). Démontrer que δ est homogène de degré $-n$ dans $\mathbb{R}^n$ et que, sur $\mathbb{R}$, la distribution $\text{vp}(1/x)$ est homogène de degré -1 .

Montrer que, sur $\mathbb{R}$, la distribution $\text{pf}(x_+^\alpha)$ est homogène de degré α à condition que α ne soit pas un entier négatif.

5.2. Multiplication par les fonctions C^∞

Il est impossible de définir en général un produit dans $\mathcal{D}' \times \mathcal{D}'$ qui prolonge continûment le produit des fonctions. Par exemple, les fonctions f_j égales à j dans $[0, 1/j]$ et à 0 ailleurs convergent vers δ dans $\mathcal{D}'(\mathbb{R})$ tandis que leurs carrés n'ont pas de limite. On peut par contre définir naturellement le produit d'une distribution par une fonction de classe C^∞ .

Théorème et Définition 5.2.1. — Soient $u \in \mathcal{D}'(\Omega)$ et $f \in C^\infty(\Omega)$. On définit la distribution produit fu par

$$\langle fu, \varphi \rangle = \langle u, f\varphi \rangle.$$

Il faut vérifier que la forme linéaire ainsi définie sur $C_0^\infty(\Omega)$ possède la propriété de continuité (4.3). On sait que, pour chaque compact K , il existe C et p tels que l'on ait

$$|\langle u, f\varphi \rangle| \leq C \sup_{|\alpha| \leq p, x \in K} |\partial^\alpha (f(x)\varphi(x))|$$

dès que φ (et donc $f\varphi$) appartiennent à C_K^∞ . En développant $\partial^\alpha (f\varphi)$ par la formule de Leibniz, on obtient une majoration

$$|\langle u, f\varphi \rangle| \leq C' \sup_{|\beta| \leq p, x \in K} |\partial^\beta (\varphi(x))|,$$

en prenant comme constante C' le produit de C , du maximum sur K des dérivées d'ordre au plus p de f , et de la somme des coefficients binomiaux concernés. L'existence de p et de C' pour chaque K exprime précisément que fu est une distribution. On voit en outre que, si u est d'ordre fini p , alors l'ordre de fu est inférieur ou égal à p .

Théorème 5.2.2. — Soient $u_j \in \mathcal{D}'(\Omega)$ et $f_j \in C^\infty(\Omega)$ deux suites telles que $u_j \rightarrow u$ au sens des distributions dans Ω et que $f_j \rightarrow f$ uniformément ainsi que chacune de ses dérivées sur tout compact de Ω . Alors $f_j u_j \rightarrow fu$ dans $\mathcal{D}'(\Omega)$

Il s'agit d'une situation qui se reproduira plusieurs fois dans ce cours. Nous allons démontrer ci-dessous, ce qui est facile, la *continuité séparée* de l'application bilinéaire, c'est-à-dire que $f_j u \rightarrow fu$ et que $f u_j \rightarrow fu$ sous les hypothèses ci-dessus. Par contre, pour la continuité énoncée dans le théorème, qui est conséquence d'un résultat difficile sur la continuité des applications bilinéaires, nous renvoyons au n°C.4.2.

On a $\langle f u_j, \varphi \rangle = \langle u_j, f\varphi \rangle \rightarrow \langle u, f\varphi \rangle = \langle fu, \varphi \rangle$. Par définition de la convergence au sens des distributions, cela prouve que $f u_j \rightarrow fu$. D'autre part, on a

$$\langle f_j u - fu, \varphi \rangle = \langle u, (f_j - f)\varphi \rangle.$$

Par définition, il existe C et p ne dépendant que du support K de φ tels que

$$|\langle f_j u - f u, \varphi \rangle| \leq C \sup_{|\alpha| \leq p, x \in K} |\partial^\alpha ((f_j(x) - f(x))\varphi(x))|.$$

En utilisant la formule de Leibniz, on voit facilement que la convergence uniforme sur K des dérivées d'ordre $\leq p$ de f_j vers les dérivées correspondantes de f entraîne la convergence vers 0 du membre de droite. On a donc $f_j u \rightarrow f u$ au sens des distributions.

Exercice 5.2.3. — Montrer que, pour $f \in C^\infty(\Omega)$ et pour $u \in \mathcal{D}'(\Omega)$, on a la formule de Leibniz :

$$\partial^\alpha (f u) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \partial^\beta f \partial^{\alpha-\beta} u.$$

Exercice 5.2.4. — Montrer que, pour f de classe C^∞ dans un ouvert de $\mathbb{R}^n$ contenant a , on a $f \delta_a = f(a) \delta_a$. Montrer que, sur $\mathbb{R}$, on a $x \delta' = -\delta$ et $x \operatorname{vp}(1/x) = 1$.

Exercice 5.2.5. — Soit $f \in C^\infty(\mathbb{R})$. Montrer qu'il existe $\alpha \in \mathbb{C}$ et $g \in C^\infty(\mathbb{R})$ tels que le produit $f \operatorname{vp}(1/x)$ soit égal à $\alpha \operatorname{vp}(1/x) + g$.

5.2.6. Équations $xu = 0$ et $xu = 1$. — Soit $u \in \mathcal{D}'(\mathbb{R})$ vérifiant $xu = 0$. Choisissons une fonction $\chi \in C_0^\infty(\mathbb{R})$ vérifiant $\chi(0) = 1$. Pour toute $\varphi \in C_0^\infty(\mathbb{R})$ la fonction $\varphi - \varphi(0)\chi$ s'annule à l'origine, et d'après le lemme de Hadamard (exercice 3.1.9), il existe $\psi \in C_0^\infty(\mathbb{R})$ telle que

$$\varphi = \varphi(0)\chi + x\psi.$$

On a donc

$$\langle u, \varphi \rangle = \varphi(0) \langle u, \chi \rangle + \langle u, x\psi \rangle.$$

Par hypothèse, on a $\langle u, x\psi \rangle = \langle xu, \psi \rangle = 0$. En appelant C la constante $\langle u, \chi \rangle$, on a donc $\langle u, \varphi \rangle = C\varphi(0)$ pour toute φ et donc $u = C\delta$. Réciproquement (exercice précédent) on a $x\delta = 0$ et les distributions solution de $xu = 0$ sont exactement les distributions de la forme $C\delta$.

D'après l'exercice ci-dessus, pour que u soit solution de $xu = 1$, il faut et il suffit que $x(u - \operatorname{vp}(1/x)) = 0$. Les solutions de cette équation sont donc les distributions de la forme $u = \operatorname{vp}(1/x) + C\delta$.

5.3. Dérivation (dimension 1)

Nous avons défini la dérivation des distributions dans la section 4.3. Nous allons maintenant essayer de calculer en pratique les dérivées des distributions qui ne sont pas des fonctions de classe C^1 .

Définition 5.3.1. — On dit qu'une fonction f définie dans un intervalle $]a, b[$ est de classe C^m par morceaux s'il existe un nombre fini de points $a = a_0 < a_1 < \dots < a_N = b$ tels que, dans chacun des intervalles $]a_i, a_{i+1}[$, les dérivées de f jusqu'à l'ordre m existent, soient continues, et se prolongent continûment dans les intervalles $]a_0, a_1]; [a_i, a_{i+1}]; [a_{N-1}, a_N[$.

Pour $0 \leq k \leq m$, on note $f^{(k)}(a_i \pm 0)$ les limites à droite et à gauche de $f^{(k)}(x)$ au point a_i .

Théorème 5.3.2 (Formule des sauts). — Soit f de classe C^1 par morceaux dans $]a, b[$. On a alors, avec les notations ci-dessus

$$f' = \{f'\} + \sum_{i=1}^{N-1} [f(a_i + 0) - f(a_i - 0)] \delta_{a_i}$$

où on a noté f' la dérivée au sens des distributions, et $\{f'\}$ la fonction continue par morceaux dérivée usuelle en dehors des points a_i .

Par définition, pour $\varphi \in C_0^\infty(]a, b[)$, on a

$$\langle f', \varphi \rangle = - \int f(x) \varphi'(x) dx.$$

En intégrant par parties dans chacun des intervalles $[a_i, a_{i+1}]$, on obtient d'une part des termes dont la somme vaudra $\int \{f'\}(x) \varphi(x) dx$ et les termes $f(a_{i+1} - 0) \varphi(a_{i+1}) - f(a_i + 0) \varphi(a_i)$ qui regroupés donneront les $\langle \delta_{a_i}, \varphi \rangle$ avec un coefficient égal au saut de f en a_i .

Dérivées successives. — Lorsque la fonction f est de classe C^2 par morceaux, on peut calculer la dérivée seconde de f au sens des distributions en appliquant le théorème précédent à $\{f'\}$. On obtient

$$f'' = \{f''\} + \sum (f'(a_i + 0) - f'(a_i - 0)) \delta_{a_i} + \sum (f(a_i + 0) - f(a_i - 0)) \delta'_{a_i}.$$

5.3.3. Dérivée de $\log |x|$. — La fonction $\log |x|$ est localement sommable sur $\mathbb{R}$ et définit donc une distribution. Pour calculer la dérivée de celle-ci, nous allons l'approcher par des fonctions dont nous savons calculer la dérivée, et utiliser la continuité de la dérivation (théorème 4.3.4). Soit f_ε la fonction égale à $\log |x|$ pour $|x| \geq \varepsilon$ et à $\log \varepsilon$ sinon. C'est une fonction de classe C^1 par morceaux dont les sauts sont nuls, et sa dérivée est donc la fonction égale à $1/x$ pour $|x| \geq \varepsilon$ et à 0 sinon.

On a $f_\varepsilon(x) \rightarrow \log|x|$ p.p. et donc (théorème 4.4.4) convergence au sens des distributions, ces fonctions étant majorées en module par la fonction localement sommable fixe $\log|x|$. On a

$$\langle (\log|x|)', \varphi \rangle = \lim \int_{|x| \geq \varepsilon} \frac{\varphi(x)}{x} dx$$

et donc

$$\frac{d}{dx} \log|x| = \text{vp} \left(\frac{1}{x} \right).$$

Exercice 5.3.4. — Démontrer que l'on a

$$\lim_{\varepsilon > 0, \varepsilon \rightarrow 0} \frac{1}{x \pm i\varepsilon} = \text{vp} \left(\frac{1}{x} \right) \mp i\pi\delta.$$

(a) Par une démonstration directe, en posant $\varphi(x) = \varphi(0) + (\varphi(x) - \varphi(0))$ pour $\varphi \in C_0^\infty$.

(b) En déterminant les primitives des fonctions $1/(x \pm i\varepsilon)$ s'annulant à l'origine et leur limite pour $\varepsilon \rightarrow 0$, et en utilisant la continuité de la dérivation (théorème 4.3.4).

Exercice 5.3.5. — Montrer que la dérivée de x_+^α est, pour $\alpha > 0$, la fonction $\alpha x_+^{\alpha-1}$. Pour $-1 < \alpha < 0$ montrer que sa dérivée est $\alpha \text{pf}(x_+^{\alpha-1})$. Plus généralement, montrer que, pour $\alpha < -1$ non entier, la dérivée de $\text{pf}(x_+^\alpha)$ est $\alpha \text{pf}(x_+^{\alpha-1})$.

Le lecteur désireux de bien comprendre ce qu'est une partie finie aura tout intérêt à faire pour la fonction $x_+^{1/2}$ ce que nous avons fait dans la figure 1 pour la fonction de Heaviside : dessiner une fonction régulière "voisine" et ses deux premières dérivées, évaluer les intégrales des pics qui apparaissent, comprendre pourquoi la première dérivée converge vers une fonction, et analyser le mécanisme de compensation qui fait que la dérivée seconde converge dans $\mathcal{D}'$.

Remarque. — Pour une fonction (localement sommable) f , les rapports entre dérivée au sens des distributions et, lorsqu'elle existe, dérivée traditionnelle peuvent être complexes. Ce n'est que pour les fonctions de classe C^1 qu'il y a identité évidente des deux concepts. Il peut y avoir coïncidence dans d'autres cas (voir exercice ci-dessus), mais cela nécessite une démonstration.

Il faut s'habituer à l'idée que, en cas de conflit, la "véritable" dérivée, celle qui rend compte complètement de la variation de f est celle au sens des distributions. Pour les fonctions de classe C^1 par morceaux, la dérivée usuelle qui existe presque partout laisse échapper l'essentiel de la variation : les sauts.

Le cas où f est dérivable en tout point est plus subtil. Dans ce cas la fonction dérivée contient aussi toute l'information sur la variation de f , mais on ne dispose guère que du théorème des accroissements finis, qui fait intervenir un point inconnu, pour l'utiliser. Lorsque la fonction dérivée est localement sommable, on peut montrer qu'elle coïncide avec

la dérivée au sens des distributions. Sinon (penser à $x^2 \sin(1/x^{1000})$) cette dernière en est une "sorte de partie finie" qui peut être fort complexe.

Exercice 5.3.6. — Soit f une fonction localement sommable sur $\mathbb{R}$ et posons $F(x) = \int_0^x f(t) dt$. Démontrer que f est la dérivée de F au sens des distributions.

Théorème 5.3.7. — Soit I un intervalle ouvert.

(a) Les distributions u sur I vérifiant $u' = 0$ sont les fonctions constantes.

(b) Pour toute $v \in \mathcal{D}'(I)$, il existe $u \in \mathcal{D}'(I)$ telle que $u' = v$.

Remarquons d'abord que $\varphi \in C_0^\infty(I)$ possède une primitive dans $C_0^\infty(I)$ si et seulement si $\int \varphi(x) dx = 0$. En effet, la seule primitive s'annulant à gauche du support de φ est $\int_{-\infty}^x \varphi(t) dt$ et celle-ci ne s'annule à droite du support de φ que si l'intégrale est nulle.

Soit $\theta \in C_0^\infty(I)$ telle que $\int \theta(x) dx = 1$. En retranchant d'une fonction $\varphi \in C_0^\infty(I)$ le produit de θ par l'intégrale de φ on obtient une fonction d'intégrale nulle, et il existe donc une unique $\psi \in C_0^\infty(I)$ telle que

$$\varphi = \left(\int \varphi(x) dx \right) \theta + \frac{d\psi}{dx}. \quad (5.1)$$

Soit maintenant u vérifiant $u' = 0$. On a

$$\langle u, \varphi \rangle = \left(\int \varphi(x) dx \right) \langle u, \theta \rangle + \langle u, \psi' \rangle.$$

Le dernier terme, qui vaut $-\langle u', \psi \rangle$ est nul, et on obtient, en appelant C la constante $\langle u, \theta \rangle$,

$$\langle u, \varphi \rangle = C \int \varphi(x) dx,$$

qui exprime que u est égale à la constante C .

Pour trouver une primitive u de v , on pose

$$\langle u, \varphi \rangle = -\langle v, \psi \rangle$$

où ψ est l'unique fonction de $C_0^\infty(I)$ associée à φ par la relation (5.1). Nous laissons au lecteur, à titre d'exercice, le soin de prouver la linéarité et la continuité de u et le fait qu'on a bien $u' = v$.

Exercice 5.3.8. — Résoudre, dans $\mathcal{D}'(\mathbb{R})$ l'équation différentielle $xu' + u = 0$.

5.4. Dérivation (dimension quelconque)

La démonstration de la formule des sauts en dimension 1 reposait sur l'intégration par parties. En dimension supérieure, nous aurons besoin de la formule de Stokes. Cette dernière a parfois mauvaise réputation : les démonstrations en seraient d'une rigueur douteuse ou très difficiles. Ce n'est pas exact, mais il faut être bien conscient du point suivant. Si on prétend prouver une formule faisant intervenir la normale extérieure à un ouvert et la mesure de surface, il est nécessaire d'avoir défini mathématiquement ces concepts. Sinon, la "démonstration" devra faire appel, en un point plus ou moins bien caché, à l'intuition physique.

Si, par contre, on a défini rigoureusement ces concepts et prouvé leur caractère invariant, c'est-à-dire leur indépendance par rapport aux coordonnées choisies, les partitions de l'unité assorties d'un choix judicieux des coordonnées permettent de se ramener au cas où l'ouvert est un surgraphe, et la démonstration devient alors très facile.

Dans le cas d'un surgraphe, puis dans un cas raisonnablement général, nous allons donner des définitions et des énoncés corrects, en indiquant brièvement le principe des démonstrations.

Le lecteur trouvera dans la section A.3 la construction de l'intégrale de surface. Cette lecture, quoique conseillée, n'est pas indispensable, à condition d'admettre les propriétés énoncées ci-dessous.

A. Formule de Stokes (cas d'un surgraphe)

Nous nous placerons dans $\mathbb{R}^n$, en notant (x', x_n) le point courant, avec $x' = (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$. Soit ω un ouvert de $\mathbb{R}^{n-1}$, soient $]a, b[$ un intervalle et Φ une fonction de classe C^∞ dans ω , à valeurs dans $]a, b[$. On pose

$$\begin{aligned}\Omega &= \{x \in \omega \times]a, b[\mid x_n > \Phi(x')\} \\ \tilde{\Omega} &= \{x \in \omega \times]a, b[\mid x_n \geq \Phi(x')\} \\ \partial\Omega &= \{x \in \omega \times]a, b[\mid x_n = \Phi(x')\}\end{aligned}$$

en notant $\tilde{\Omega}$ l'adhérence de Ω dans le "cylindre" $\omega \times]a, b[$ pour éviter toute confusion avec l'adhérence dans $\mathbb{R}^n$.

5.4.1. Mesure de surface. — Soit f une fonction définie sur $\partial\Omega$, continue et à support compact. On pose

$$\int_{\partial\Omega} f(x) d\sigma(x) = \int_{\omega} f(x', \Phi(x')) \sqrt{1 + |\nabla' \Phi(x')|^2} dx' \quad (5.2)$$

où on a noté $\nabla'\Phi$ le gradient $(n-1)$ -dimensionnel $(\partial_1\Phi, \dots, \partial_{n-1}\Phi) \in \mathbb{R}^{n-1}$. Il s'agit bien d'une forme linéaire positive sur l'espace des fonctions continues à support compact sur $\partial\Omega$.

On peut bien sûr définir directement l'intégrale de fonctions plus générales. Le membre de droite dans (5.2) a un sens, fini ou infini, si f est une fonction positive sur $\partial\Omega$, et définit un nombre complexe si f est majorée en module par une fonction positive d'intégrale finie.

5.4.2. Normale extérieure. — La normale extérieure unitaire en un point $(x', \Phi(x'))$ est le vecteur ν défini par

$$\nu = \frac{1}{\sqrt{1 + |\nabla'\Phi(x')|^2}} \begin{pmatrix} \partial_1\Phi \\ \vdots \\ \partial_{n-1}\Phi \\ -1 \end{pmatrix}. \quad (5.3)$$

Théorème et Définition 5.4.3. — Soit f une fonction définie et continue sur $\tilde{\Omega}$. Les deux propriétés suivantes sont équivalentes.

(a) La fonction f est de classe C^m dans Ω et ses dérivées jusqu'à l'ordre m se prolongent continûment à $\tilde{\Omega}$.

(b) Il existe une fonction appartenant à $C^m(\omega \times]a, b[)$ qui coïncide avec f dans $\tilde{\Omega}$.

On dit que f est de classe C^m jusqu'au bord, ce qu'on note $f \in C^m(\tilde{\Omega})$ si ces conditions sont vérifiées.

En posant $g(x', t) = f(x', \Phi(x') + t)$, on se ramène au problème de trouver un prolongement $\tilde{g}$, pour t quelconque, d'une fonction g définie pour $t \geq 0$ et de classe C^m jusqu'au bord au sens de (a). Pour $m = 1$, le lecteur vérifiera que $\tilde{g}(x', -t) = g(x', 0) - t\partial_t g(x', 0)$ ne convient pas, mais que $\tilde{g}(x', -t) = 2g(x', 0) - g(x', t)$ convient. Il constatera également que $\tilde{g}(x', -t) = 3g(x', 0) - 3g(x', t) + g(x', 2t)$ résout le problème pour $m = 2$, et pourra généraliser à m quelconque. Les prolongements obtenus ne sont définis qu'au voisinage de $\tilde{\Omega}$, mais en les multipliant par des fonctions C^∞ à support dans ce voisinage et égales à 1 dans un voisinage plus petit de $\tilde{\Omega}$, on obtient un prolongement dans $\omega \times]a, b[$.

Pour $m = +\infty$, les $\partial_t^k g(x', 0)$ sont de classe C^∞ , et on prolonge pour $t < 0$ par le procédé de Borel (exercice 3.2.10).

Théorème 5.4.4 (Formule de Stokes). — Soit X un champ de vecteurs défini sur $\tilde{\Omega}$, dont le support est un compact de $\tilde{\Omega}$, et dont les composantes appartiennent à $C^1(\tilde{\Omega})$. On a⁽¹⁾

$$\int_{\partial\Omega} X \cdot \nu \, d\sigma = \int_{\Omega} \operatorname{div} X \, dx.$$

⁽¹⁾On a pris l'habitude de désigner sous le nom générique de "formule de Stokes" toutes les formules d'intégration par parties en plusieurs dimensions, ce qui est quelque peu injuste pour Gauss, Green, Riemann, Ostrogradski, ...

en notant $\operatorname{div} X = \partial_1 X_1 + \dots + \partial_n X_n$ et $X \cdot \nu$ le produit scalaire. en un point du bord, de la valeur de X et de la normale extérieure unitaire en ce point.

Rappelons brièvement le principe de la démonstration que le lecteur connaît sans doute déjà. En intégrant en x_n , on obtient

$$\int_{\Omega} \frac{\partial X_n}{\partial x_n}(x_1, \dots, x_n) dx = - \int_{\omega} X_n(x', \Phi(x')) dx'.$$

Pour $i = 1, \dots, n-1$, on pose $X_i(x', x_n) = h(x', x_n - \Phi(x'))$, la fonction $h(x, t)$ étant définie pour $t \geq 0$. On obtient

$$\partial_i X_i(x', x_n) = \partial_i h(x', x_n - \Phi(x')) - \partial_n h(x', x_n - \Phi(x')) \partial_i \Phi(x').$$

En faisant le changement de variables $x' = x'$, $t = x_n - \Phi(x')$ dont le jacobien est égal à 1, on a

$$\int_{\Omega} \partial_i X_i(x', x_n) dx = \int_{\omega \times [0, \infty[} \partial_i h(x', t) dx' dt - \int_{\omega \times [0, \infty[} \partial_n h(x', t) \partial_i \Phi(x') dx' dt.$$

La première intégrale du membre de droite est nulle, comme on le voit en intégrant d'abord en x_i . En intégrant la seconde en t , on a

$$\int_{\Omega} \partial_i X_i(x', x_n) dx = \int_{\omega} h(x', 0) \partial_i \Phi(x') dx' = \int_{\omega} X_i(x', \Phi(x')) \partial_i \Phi(x') dx'.$$

On obtient donc, pour $\int_{\Omega} \operatorname{div} X dx$, l'intégrale sur $\partial\Omega$ (pour la mesure dx') du produit scalaire de X et du vecteur de composantes $\partial_1 \Phi(x'), \dots, \partial_{n-1} \Phi(x'), -1$, ce qui est précisément le résultat voulu.

B. Formule de Stokes (cas d'un ouvert régulier)

Définition 5.4.5. — On appellera ouvert régulier de $\mathbb{R}^n$ un ouvert Ω tel que, pour tout point $m \in \partial\Omega$, on peut trouver

- (a) un système de coordonnées orthonormales $(y_1, \dots, y_n)$ (on désignera par m' [resp. m_n] les $n-1$ premières [resp. la dernière] coordonnées de m),
- (b) un ouvert $\omega \in \mathbb{R}^{n-1}$ contenant m' et un intervalle $]a, b[$ contenant m_n (on désignera par B le "cylindre" de $\mathbb{R}^n$ s'écrivant $\omega \times]a, b[$ dans les coordonnées (y)),
- (c) une application Φ de classe C^∞ de ω dans $]a, b[$, tels que, dans le système de coordonnées (y) , on ait

$$\Omega \cap B = \{y \in \omega \times]a, b[\mid y_n > \Phi(y')\}.$$

Cette définition exprime que Ω se comporte comme le surgraphe d'une fonction régulière au voisinage de chaque point de sa frontière. Elle affirme notamment que $\partial\Omega$ est une hypersurface de $\mathbb{R}^n$ (voir la section A.2), mais elle assure en outre que Ω est situé localement du même côté de sa frontière, propriété bien utile pour parler de normale extérieure.

Théorème et Définition 5.4.6 (Normale extérieure)

Soient Ω un ouvert régulier et m un point de $\partial\Omega$. Avec les notations de la définition 5.4.5. soit ν le vecteur dont les composantes dans le système de

coordonnées (y) sont données par (5.3). Ce vecteur est indépendant du choix de (y) et est appelé normale extérieure unitaire au point m .

Soit maintenant Ω un ouvert régulier borné. Sa frontière $\partial\Omega$ étant compacte, on peut la recouvrir par un nombre fini de "cylindres" $B^\varkappa$, $\varkappa = 1, \dots, N$, vérifiant les conditions de la définition 5.4.5 pour des données $y^\varkappa$, $\Phi^\varkappa$, ... D'après le théorème 3.2.9, on peut trouver une partition de l'unité ($\sum \chi^\varkappa = 1$ au voisinage de $\partial\Omega$), par des fonctions $\chi^\varkappa$ à support dans les $B^\varkappa$.

Soit f une fonction continue sur $\partial\Omega$. Nous pouvons alors calculer l'intégrale $\int f(x) \chi^\varkappa(x) d\sigma(x)$ par la formule (5.2), où x et Φ sont remplacés par $y^\varkappa$ et $\Phi^\varkappa$.

Théorème et Définition 5.4.7 (Mesure de surface)

La somme des N intégrales $\int f(x) \chi^\varkappa(x) d\sigma(x)$, calculées comme précédemment, ne dépend que de f et non du choix des $B^\varkappa$, $y^\varkappa$, $\Phi^\varkappa$ et $\chi^\varkappa$. On note cette quantité

$$\langle d\sigma, f \rangle = \int_{\partial\Omega} f(x) d\sigma(x).$$

On définit ainsi la mesure de Radon positive $d\sigma$ appelée mesure de surface sur $\partial\Omega$.

Rappelons (n°4.4.7) que, plus généralement, on note $g d\sigma$ la distribution de simple couche définie par

$$\langle g d\sigma, \varphi \rangle = \int_{\partial\Omega} g(x) \varphi(x) d\sigma(x)$$

pour g sommable par rapport à $d\sigma$ sur $\partial\Omega$.

La démonstration des propriétés ci-dessus est un peu longue, mais elle se réduit à une succession de vérifications reposant sur le théorème d'inversion locale et la formule de changement de variables dans les intégrales multiples. Le lecteur trouvera dans la section A.3 la démonstration du théorème précédent et, en contrôlant le signe des déterminants qui y interviennent, pourra également prouver l'invariance de la définition de la normale extérieure.

Par contre, une fois établis ces théorèmes d'invariance, la démonstration des résultats suivants est facile. Pour prouver la formule de Stokes, on décompose, grâce à une partition de l'unité, le champ de vecteurs X en la somme de champs $X^\varkappa$ ayant leurs supports dans les $B^\varkappa$ et d'un champ X_0 à support compact dans Ω . Il suffit de vérifier l'égalité des deux membres pour chacun de ces champs, ce qui est immédiat pour X_0 et déjà démontré (théorème 5.4.4) pour les $X_\varkappa$ puisque l'on peut utiliser les coordonnées $y^\varkappa$.

D'une manière générale, pour tous les énoncés du type "formule de Stokes" sur une sous-variété de $\mathbb{R}^n$ ou sur une variété différentielle abstraite, le point important est la définition des concepts mis en œuvre et leur caractère invariant par changement de coordonnées. La démonstration elle-même se ramène toujours, par partitions de l'unité et choix de coordonnées locales, à une simple intégration par parties dans un demi-espace de $\mathbb{R}^n$.

Théorème 5.4.8. — Soient Ω un ouvert borné régulier, $m = 0, 1, \dots, \infty$ et f continue sur $\overline{\Omega}$. Les deux propriétés suivantes sont équivalentes.

(a) f est de classe C^m dans Ω et les dérivées de f jusqu'à l'ordre m se prolongent continûment à $\overline{\Omega}$.

(b) Il existe une fonction appartenant à $C^m(\mathbb{R}^n)$ qui coïncide avec f dans $\overline{\Omega}$.

On dit que f est de classe C^m jusqu'au bord, ce qu'on note $f \in C^m(\overline{\Omega})$ si ces conditions sont vérifiées.

Théorème 5.4.9 (Formule de Stokes). — Soit Ω un ouvert borné régulier et X un champ de vecteurs défini sur $\overline{\Omega}$ dont les composantes appartiennent à $C^1(\overline{\Omega})$. On a

$$\int_{\partial\Omega} X \cdot \nu \, d\sigma = \int_{\Omega} \operatorname{div} X \, dx.$$

Remarque 5.4.10. — Si on appelle Ω' le complémentaire de $\overline{\Omega}$, c'est encore un ouvert régulier, non borné mais de frontière $\partial\Omega$ compacte. Le théorème 5.4.8 est valable mot pour mot et permet de définir $C^m(\overline{\Omega}')$. La formule de Stokes est encore valable pour un champ de vecteurs dont les coefficients appartiennent à $C^1(\overline{\Omega}')$ et sont à support compact. Il ne faut bien sûr pas oublier que les normales extérieures à Ω et Ω' sont opposées en un même point de $\partial\Omega$.

Corollaire 5.4.11 (Intégration par parties). — Soient Ω un ouvert borné régulier, et f et g appartenant à $C^1(\overline{\Omega})$. On a

$$\int_{\Omega} g(x) \partial_i f(x) \, dx = \int_{\partial\Omega} f(x) g(x) \cos(\nu, e_i) \, d\sigma(x) - \int_{\Omega} f(x) \partial_i g(x) \, dx$$

où les $\nu_i = \cos(\nu, e_i)$ sont les composantes du vecteur normal extérieur unitaire au point x .

Il suffit en effet d'appliquer la formule de Stokes au champ de vecteurs $x \mapsto f(x)g(x)e_i$.

Corollaire 5.4.12 (Formule de Green). — Soient Ω un ouvert borné régulier, u et v appartenant à $C^2(\overline{\Omega})$. On a alors

$$\int_{\Omega} (u\Delta v - v\Delta u) \, dx = \int_{\partial\Omega} \left(u \frac{\partial v}{\partial \nu} - v \frac{\partial u}{\partial \nu} \right) d\sigma$$

en notant $\partial u / \partial \nu = \nabla u \cdot \nu$.

Appliquons la formule de Stokes au champ de vecteurs $u\nabla v$. On obtient

$$\int_{\Omega} (\nabla u \cdot \nabla v + u\Delta v) \, dx = \int_{\partial\Omega} u (\nabla v \cdot \nu) \, d\sigma.$$

Il suffit de retrancher la formule symétrique en u et v pour avoir le résultat.

C. Formule des sauts dans l'espace

Soit Ω un ouvert borné régulier et Ω' le complémentaire de $\overline{\Omega}$. Soit f une fonction définie dans $\mathbb{R}^n$ telle que ses restrictions à Ω et Ω' se prolongent par continuité en des éléments de $C^1(\overline{\Omega})$ et $C^1(\overline{\Omega}')$. Pour $x \in \partial\Omega$, on notera $f_{\text{int}}(x)$ et $f_{\text{ext}}(x)$ les valeurs respectives de ces prolongements.

Théorème 5.4.13. — *Avec les notations ci-dessus, on a*

$$\partial f / \partial x_i = \{ \partial f / \partial x_i \} + [f_{\text{ext}} - f_{\text{int}}] \cos(\nu, e_i) d\sigma, \quad (5.4)$$

où $\{ \partial f / \partial x_i \}$ est la fonction définie hors de $\partial\Omega$ dérivée usuelle de f et où le second terme est la distribution de simple couche ayant pour densité par rapport à la mesure de surface le produit du saut de f par le cosinus de l'angle (ν, e_i) .

Soit en effet $\varphi \in C_0^\infty(\mathbb{R}^n)$. Nous pouvons alors appliquer le corollaire de la formule de Stokes dans $\overline{\Omega}$ au produit par φ du prolongement par continuité de $f|_\Omega$, et dans $\overline{\Omega}'$ (voir remarque 5.4.10), au produit par φ du prolongement de $f|_{\Omega'}$. On obtient

$$\begin{aligned} - \int_{\Omega} f \partial_i \varphi dx &= \int_{\Omega} \varphi \partial_i f dx - \int_{\partial\Omega} \varphi f_{\text{int}} \cos(\nu, e_i) d\sigma \\ - \int_{\Omega'} f \partial_i \varphi dx &= \int_{\Omega'} \varphi \partial_i f dx - \int_{\partial\Omega} \varphi f_{\text{ext}} \cos(-\nu, e_i) d\sigma. \end{aligned}$$

La somme des membres de gauche vaut par définition $\langle \partial_i f, \varphi \rangle$. Les premiers termes des membres de droite donnent $\langle \{ \partial_i f \}, \varphi \rangle$, tandis que les termes restants donnent

$$\int_{\partial\Omega} [f_{\text{ext}} - f_{\text{int}}] \cos(\nu, e_i) \varphi d\sigma,$$

qui est précisément l'action sur φ de la mesure de Radon dont la densité par rapport à $d\sigma$ est $[f_{\text{ext}} - f_{\text{int}}] \cos(\nu, e_i)$.

5.4.14. Forme locale de la formule des sauts. — Soit $B = \omega \times]a, b[$ un "cylindre" dans $\mathbb{R}^n$, soit Φ une application de classe C^∞ de ω dans $]a, b[$, et soient B_1 et B_2 les sous-ouverts de B définis par $x_n < \Phi(x')$ et $x_n > \Phi(x')$. On note ν le vecteur normal unitaire (extérieur à B_1 ou B_2 , peu importe à condition de compter les sauts dans le bon sens). La formule suivante se démontre de même à partir du théorème 5.4.4

$$\partial f / \partial x_i = \{ \partial f / \partial x_i \} + [f(x+0\nu) - f(x-0\nu)] \cos(\nu, e_i) d\sigma, \quad (5.5)$$

où f est définie dans B et est telle que ses restrictions à B_j , $j = 1, 2$ se prolongent en des éléments de $C^1(\widetilde{B_j})$.

D. Applications

5.4.15. *Condition de Rankine-Hugoniot.* — Les lois de conservation en mécanique et en physique se traduisent fréquemment par des (systèmes d') équations aux dérivées partielles du type suivant

$$\partial u / \partial t + \operatorname{div} F(u) = 0$$

où $u(t, x)$ est une application inconnue de $\mathbb{R}^{n+1}$ dans $\mathbb{R}$, et où F est de classe C^∞ à valeurs dans $\mathbb{R}^n$. Cette équation s'écrit donc

$$\partial u / \partial t + \sum_1^n \partial / \partial x_i \{ f_i(u(t, x)) \} = 0, \quad (5.6)$$

en notant f_i les composantes de F .

Lorsque la fonction u est assez régulière, il est équivalent de demander que u vérifie

$$\partial u / \partial t + \sum_1^n f'_i(u(t, x)) \partial u / \partial x_i = 0. \quad (5.7)$$

Par contre, lorsque la fonction u n'est pas régulière, il y a une grande différence entre la forme (5.6), dite "équation sous forme conservative", et la forme (5.7). Ainsi, si u est une fonction bornée, les fonctions $f_i \circ u$ sont bornées et possèdent des dérivées au sens des distributions. Demander que la distribution figurant au membre de gauche de (5.6) soit nulle a un sens parfaitement défini. Au contraire, le membre de gauche de (5.7) contiendrait le produit d'une fonction bornée par la dérivée d'une fonction bornée, qui n'a aucune raison d'être défini.

Un cas particulier important est celui où on cherche des solutions u du type décrit dans le numéro 5.4.14, ayant une discontinuité le long d'une hypersurface Σ (ondes de choc). Dire que u est une solution de (5.6) s'exprime alors par les deux conditions suivantes :

- les termes correspondant aux dérivées au sens des fonctions dans la formule (5.5) doivent s'annuler, ce qui veut dire que, en dehors de Σ , la fonction u vérifie (5.7),
- les distributions de simple couche apparaissant dans les dérivées doivent également s'annuler. La formule (5.5) permet d'en calculer la densité par rapport à $d\sigma$. On doit donc avoir

$$\nu_0 [u(x+0\nu) - u(x-0\nu)] + \sum \nu_i [f_i(u(x+0\nu)) - f_i(u(x-0\nu))] = 0$$

en notant ν_0 et ν_i les composantes temporelle et spatiales du vecteur normal à Σ .

C'est la célèbre condition de Rankine-Hugoniot que, dans la formulation classique, il fallait adjoindre aux équations (écrites seulement en dehors de Σ) pour décrire la conservation du flux à l'endroit même du choc. Par contre, en écrivant les équations sous forme conservative et en considérant les dérivées au sens des distributions, il n'y a aucune condition supplémentaire à introduire.

Exercice 5.4.16. — Soit u une fonction du type ci-dessus ayant une discontinuité le long de Σ . On suppose qu'il existe une suite u_j de fonctions de classe C^1 qui sont des solutions (usuelles) de l'équation (5.7). On suppose de plus que les u_j sont uniformément bornées et que $u_j(x) \rightarrow u(x)$ p.p. Montrer que u est solution de (5.6) au sens des distributions et qu'elle vérifie donc en particulier la condition de Rankine-Hugoniot.

5.4.17. Calcul du Laplacien de $1/r$. — On se place dans $\mathbb{R}^3$, en posant $r = \sqrt{x^2 + y^2 + z^2}$. La fonction $1/r$ est localement sommable et définit donc une distribution dont on peut calculer le Laplacien. Soient f_ε les fonctions égales à $1/r$ pour $r \geq \varepsilon$ et à $1/\varepsilon$ sinon. Ces fonctions sont majorées par la fonction localement sommable fixe $1/r$, elles convergent vers $1/r$ presque partout et donc au sens des distributions. Par continuité de la dérivée, on a donc $\Delta(1/r) = \lim(\Delta f_\varepsilon)$.

On peut appliquer la formule des sauts à la fonction f_ε , l'ouvert régulier étant la boule de rayon ε . Comme le saut est nul, $\partial f_\varepsilon / \partial x$ est la fonction valant $-x/r^3$ pour $r > \varepsilon$ et 0 sinon. Par contre, dans le calcul de $\partial^2 f_\varepsilon / \partial x^2$, il apparaît d'une part la fonction valant $-1/r^3 + 3x^2/r^5$ pour $r > \varepsilon$ et 0 sinon, d'autre part le terme de saut. Celui-ci est une mesure ayant pour densité par rapport à la mesure de surface le produit de la valeur du saut $-x/\varepsilon^3$ par le cosinus x/ε de l'angle de la normale avec l'axe des x . On a donc, la somme des fonctions étant nulle,

$$\Delta f_\varepsilon = -(x^2 + y^2 + z^2)/\varepsilon^4 d\sigma_\varepsilon = -\varepsilon^{-2} d\sigma_\varepsilon,$$

en notant $d\sigma_\varepsilon$ la mesure de surface sur la sphère S_ε de rayon ε . Il suffit maintenant de reprendre l'argument de l'exemple 4.1.1. On a

$$\langle \Delta f_\varepsilon, \varphi \rangle = -\varepsilon^{-2} \int_{S_\varepsilon} \varphi(0) d\sigma_\varepsilon - \varepsilon^{-2} \int_{S_\varepsilon} (\varphi(x) - \varphi(0)) d\sigma_\varepsilon.$$

Le premier terme est constant et égal à $-4\pi\varphi(0)$, et on montre facilement que le second tend vers 0 avec ε . On a donc

$$\langle \Delta(1/r), \varphi \rangle = -4\pi\varphi(0),$$

ce qui prouve que l'on a

$$\Delta(1/r) = -4\pi\delta.$$

5.4.18. *Autre méthode, exercice.* — Approcher $1/r$ par des fonctions g_ε , égales à $1/r$ pour $r > \varepsilon$, et à $a_\varepsilon r^2 + b_\varepsilon$ sinon. Choisir a_ε et b_ε pour que g_ε et ses dérivées premières soient continues sur S_ε . Conclure par un calcul très simple.

Calculer de même $\Delta(1/r^{n-2})$ dans $\mathbb{R}^n$ pour $n \geq 3$, et $\Delta(\log r)$ dans $\mathbb{R}^2$.

Exercice 5.4.19. — Soit, dans $\mathbb{R}^3$, la fonction g_R égale à $1/r - 1/R$ pour $r \leq R$ et à 0 sinon. Montrer que l'on a $\Delta g_R = R^{-2} d\sigma_R - 4\pi\delta$.

Exercice 5.4.20. — Calculer $\bar{\partial}(1/z)$ dans le plan complexe, où on a posé $\bar{\partial} = \partial/\partial x + i\partial/\partial y$. On pourra approcher $1/z$ par des fonctions égales à $1/z$ pour $|z| > \varepsilon$ et à 0 (première méthode) ou à $\bar{z}/\varepsilon^2$ (seconde méthode) pour $|z| \leq \varepsilon$.

CHAPITRE 6

ESPACES PARTICULIERS DE DISTRIBUTIONS

6.1. Distributions à support compact

Théorème et Définition 6.1.1 (Support d'une distribution)

Soit u appartenant à $\mathcal{D}'(\Omega)$. Il existe un plus grand sous-ouvert ω de Ω tel que la restriction $u|_{\omega}$ soit nulle. Le complémentaire de cet ouvert est appelé le support de u et est noté $\text{Supp}(u)$.

Le support de u est donc un sous-ensemble fermé F de Ω tel que, pour $\varphi \in C_0^\infty(\Omega)$ nulle au voisinage de F on ait $\langle u, \varphi \rangle = 0$, et c'est le plus petit ensemble fermé jouissant de cette propriété.

Considérons l'ensemble des sous-ouverts ω de Ω tels que $u|_{\omega} = 0$ et notons ω_0 leur réunion. Il suffit de démontrer que $u|_{\omega_0} = 0$ et donc de montrer que, pour chaque $\varphi \in C_0^\infty(\omega_0)$, on a $\langle u, \varphi \rangle = 0$.

Le compact $\text{Supp}(\varphi)$ est alors recouvert par des ouverts ω auxquels la restriction de u est nulle et, d'après le théorème de Borel-Lebesgue, on peut en trouver un nombre fini ω_i , $i = 1, \dots, k$ tels que $\text{Supp}(\varphi) \subset \omega_1 \cup \dots \cup \omega_k$.

Considérons une partition de l'unité au voisinage de $\text{Supp}(\varphi)$ par des fonctions ψ_i à support dans ω_i . On a

$$\langle u, \varphi \rangle = \sum_{i=1}^k \langle u, \psi_i \varphi \rangle = 0,$$

ce qui achève la démonstration.

Remarque 6.1.2. — Pour une fonction continue, la définition coïncide avec la définition usuelle : adhérence de l'ensemble des points où la fonction est non nulle. Pour une fonction localement sommable, c'est le complémentaire du plus grand ouvert dans lequel la fonction est nulle (presque partout).

Il ne suffit pas que φ soit nulle sur $\text{Supp}(u)$ pour avoir $\langle u, \varphi \rangle = 0$. Par exemple, on voit facilement que $\text{Supp}(\delta') = \{0\}$ et que φ peut être nulle à l'origine sans que sa dérivée le soit.

Il résulte immédiatement de la définition que, pour $f \in C^\infty(\Omega)$, on a

$$\text{Supp}(fu) \subset \text{Supp}(f) \cap \text{Supp}(u). \quad (6.1)$$

Théorème 6.1.3. — On note $\mathcal{E}'(\Omega)$ l'espace des distributions dans Ω à support compact⁽¹⁾.

(a) Toute distribution $u \in \mathcal{E}'(\Omega)$ est d'ordre fini.

(b) Plus précisément, en notant p l'ordre de u , pour tout voisinage compact K de $\text{Supp}(u)$, il existe une constante C telle que

$$\forall \varphi \in C_0^\infty(\Omega), |\langle u, \varphi \rangle| \leq C \sup_{|\alpha| \leq p, x \in K} |\partial^\alpha \varphi(x)|. \quad (6.2)$$

Soit donc K un voisinage compact de $\text{Supp}(u)$, et soit χ une fonction de classe C^∞ à support dans K et égale à 1 au voisinage de $\text{Supp}(u)$. Pour $\varphi \in C_0^\infty(\Omega)$, la fonction $\varphi - \chi\varphi$ est nulle au voisinage de $\text{Supp}(u)$, et on a donc $\langle u, \varphi \rangle = \langle u, \chi\varphi \rangle$. Appelons p le plus petit entier tel qu'il existe C_1 avec (voir définition 4.2.1)

$$\forall \psi \in C_K^\infty, |\langle u, \psi \rangle| \leq C_1 \sup_{|\alpha| \leq p} |\partial^\alpha \psi|. \quad (6.3)$$

On a donc

$$\forall \varphi \in C_0^\infty(\Omega), |\langle u, \varphi \rangle| \leq C_1 \sup_{|\alpha| \leq p} |\partial^\alpha (\chi\varphi)|.$$

En développant $\partial^\alpha (\chi\varphi)$ par la formule de Leibniz, on obtient l'estimation (6.2) avec une constante C faisant intervenir C_1 , les bornes supérieures des dérivées de χ d'ordre $\leq p$, et des coefficients binomiaux.

Cela prouve que u est d'ordre fini inférieur ou égal à p . Mais d'autre part, le fait que p soit le plus petit entier tel que l'estimation (6.3) ait lieu entraîne que l'ordre est $\geq p$, ce qui achève la démonstration du théorème.

Corollaire 6.1.4. — Soit $u \in \mathcal{E}'(\Omega)$ et soit p son ordre. Pour toute fonction $\varphi \in C_0^\infty(\Omega)$, nulle sur $\text{Supp}(u)$ ainsi que toutes ses dérivées d'ordre $\leq p$, on a $\langle u, \varphi \rangle = 0$.

On notera K le support de u et K_ρ l'ensemble des points dont la distance à K est $\leq \rho$. Soit g_ε la fonction égale à 1 sur $K_{2\varepsilon}$ et à 0 ailleurs et posons $\psi_\varepsilon = g_\varepsilon \star \chi_\varepsilon$, où les χ_ε forment une approximation de l'identité. La fonction

⁽¹⁾ Les notations originales de L. Schwartz, encore souvent utilisées, sont $\mathcal{D}$ et $\mathcal{E}$ pour les espaces que nous notons C_0^∞ et C^∞ , d'où les notations (universellement utilisées) de leurs duaux.

ψ_ε est de classe C^∞ , comprise entre 0 et 1, à support dans $K_{3\varepsilon}$, égale à 1 sur K_ε et elle vérifie

$$\partial^\alpha \psi_\varepsilon(x) = \int g_\varepsilon(x-y) \partial^\alpha \chi_\varepsilon(y) dy. \quad (6.4)$$

On a $\partial^\alpha \chi_\varepsilon(x) = \varepsilon^{-n-|\alpha|} (\partial^\alpha \chi)(x/\varepsilon)$, et donc $\|\partial^\alpha \chi_\varepsilon\|_{L^1} = C_\alpha \varepsilon^{-|\alpha|}$, en notant C_α la norme de $\partial^\alpha \chi$ dans L^1 . Il résulte donc de (6.4) que

$$\sup |\partial^\alpha \psi_\varepsilon(x)| \leq C_\alpha \varepsilon^{-|\alpha|}. \quad (6.5)$$

Soit maintenant $\varphi \in C_0^\infty(\Omega)$ nulle sur K ainsi que ses dérivées d'ordre $\leq p$. Il est facile de montrer que l'on a, pour $|\beta| \leq p$, et avec une constante C faisant intervenir le maximum des dérivées d'ordre au plus $p+1$ de φ ,

$$\sup_{x \in K_{3\varepsilon}} |\partial^\beta \varphi(x)| \leq C \varepsilon^{p+1-|\beta|} \quad (6.6)$$

(on écrira le développement de Taylor à l'ordre $p-|\beta|$ de la fonction $\partial^\beta \varphi$ en un point $x_0 \in K$ tel que $|x-x_0| \leq 3\varepsilon$).

Nous avons alors pour tout ε

$$\langle u, \varphi \rangle = \langle u, \varphi - \varphi \psi_\varepsilon \rangle + \langle u, \varphi \psi_\varepsilon \rangle.$$

Le premier terme est nul, la fonction étant nulle dans le voisinage K_ε du support de u . Le second terme est majoré en module par une constante fois la somme des bornes supérieures des $\partial^\alpha (\varphi \psi_\varepsilon)$ pour $|\alpha| \leq p$. Ces fonctions étant nulles hors de $K_{3\varepsilon}$, la formule de Leibniz, (6.5) et (6.6) donnent

$$|\partial^\alpha (\varphi \psi_\varepsilon)| \leq \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} C \varepsilon^{p+1-|\beta|} C_{\alpha-\beta} \varepsilon^{-|\alpha|+|\beta|}.$$

On a donc $|\langle u, \varphi \rangle| \leq C' \varepsilon$ pour tout ε , ce qui achève la démonstration du corollaire.

6.1.5. Distributions à support réduit à un point. — Soit $u \in \mathcal{D}'(\mathbb{R}^n)$ dont le support est l'origine (par exemple), et soit p son ordre. Soit $\psi \in C_0^\infty(\mathbb{R}^n)$ égale à 1 au voisinage de l'origine. On peut alors écrire, pour toute fonction $\varphi \in C_0^\infty(\mathbb{R}^n)$

$$\varphi(x) = \sum_{|\alpha| \leq p} \frac{1}{\alpha!} \partial^\alpha \varphi(0) x^\alpha \psi(x) + r(x),$$

où la différence $r(x)$ est nulle à l'origine ainsi que ses dérivées d'ordre $\leq p$. On a $\langle u, r \rangle = 0$ d'après le corollaire précédent, et donc, en notant b_α les constantes $\langle u, \psi(x) x^\alpha \rangle / \alpha!$,

$$\langle u, \varphi \rangle = \sum_{|\alpha| \leq p} b_\alpha \partial^\alpha \varphi(0),$$

ou encore, en posant $c_\alpha = (-1)^{|\alpha|} b_\alpha$,

$$u = \sum_{|\alpha| \leq p} c_\alpha \partial^\alpha \delta. \quad (6.7)$$

En d'autres termes, les distributions dont le support est réduit à un point sont exactement les combinaisons linéaires de dérivées de la masse de Dirac en ce point.

La décomposition (6.7) est unique, les distributions $\partial^\alpha \delta$ étant linéairement indépendantes sur $\mathbb{C}$. Ce dernier point résulte du fait que $\langle \partial^\alpha \delta, x^\beta \psi \rangle = (-1)^{|\alpha|} \alpha! \delta_\beta^\alpha$ (le dernier delta étant celui de Kronecker, égal à 1 si les multiindices α et β sont égaux, et à 0 sinon).

Définition 6.1.6 (Extension de la dualité). — Soient $u \in \mathcal{E}'(\Omega)$ et $\varphi \in C^\infty(\Omega)$. On pose

$$\langle u, \varphi \rangle_{\mathcal{E}', C^\infty} = \langle u, \theta \varphi \rangle,$$

où θ appartient à $C_0^\infty(\Omega)$ et est égale à 1 au voisinage du support de u , le résultat étant indépendant de la fonction θ choisie.

Il est clair que le changement de fonction θ ne modifie $\theta \varphi$ que par une fonction nulle au voisinage du support de u et ne change donc pas le résultat. Pour la même raison, cette définition coïncide avec l'ancienne lorsque φ est à support compact. On omettra parfois, lorsqu'il n'y a pas d'ambiguïté, l'indice du crochet de dualité.

Remarque 6.1.7. — L'estimation (6.2) est valable, avec les mêmes constantes C , pour $\varphi \in C^\infty(\Omega)$. Il suffit en effet de choisir la fonction θ ci-dessus égale à 1 au voisinage de K , les dérivées de φ et de $\theta \varphi$ coïncidant alors sur K .

Cette extension identifie $\mathcal{E}'(\Omega)$ à l'espace des formes linéaires L sur $C^\infty(\Omega)$ vérifiant la propriété de continuité suivante : il existe un compact $K \subset \Omega$, une constante $C > 0$ et un entier p tels que l'on ait

$$|L(\varphi)| \leq C \sup_{|\alpha| \leq p, x \in K} |\partial^\alpha \varphi(x)| \quad (6.8)$$

pour toute fonction $\varphi \in C^\infty(\Omega)$. Remarquons d'abord que, pour $u \in \mathcal{E}'(\Omega)$, la forme linéaire $\varphi \mapsto \langle u, \varphi \rangle_{\mathcal{E}', C^\infty}$ vérifie la propriété ci-dessus. C'est précisément ce qu'affirme l'extension de l'estimation (6.2) en prenant pour K n'importe quel voisinage compact de $\text{Supp}(u)$.

Réciproquement, si L est une telle forme linéaire, en posant $\langle u, \varphi \rangle = L(\varphi)$ pour $\varphi \in C_0^\infty(\Omega)$, il résulte facilement de (6.8) que u est une distribution. En outre, si $\text{Supp}(\varphi) \cap K = \emptyset$, l'estimation donne $\langle u, \varphi \rangle = 0$ et u est donc à support dans K . Enfin, la même estimation (6.8) donne $L(\varphi - \theta \varphi) = 0$ pour $\varphi \in C^\infty$ et θ égale à 1 au voisinage de K , ce qui assure que L est bien l'extension de l'application $\varphi \mapsto \langle u, \varphi \rangle$ définie ci-dessus.

6.1.8. *Autre extension de la dualité.* — Plus généralement, soient F et G deux fermés de Ω dont l'intersection est compacte. Pour $u \in \mathcal{D}'(\Omega)$ et $\varphi \in C^\infty(\Omega)$ vérifiant $\text{Supp}(u) \subset F$, $\text{Supp}(\varphi) \subset G$, on peut poser

$$\langle u, \varphi \rangle = \langle u, \theta \varphi \rangle,$$

où $\theta \in C_0^\infty(\Omega)$ est égale à 1 au voisinage de $F \cap G$, le résultat ne dépendant pas de la fonction θ choisie. Cette définition coïncide avec les précédentes lorsque u ou φ est à support compact.

6.2. Espaces de Sobolev d'ordre entier

A. Notions de régularité

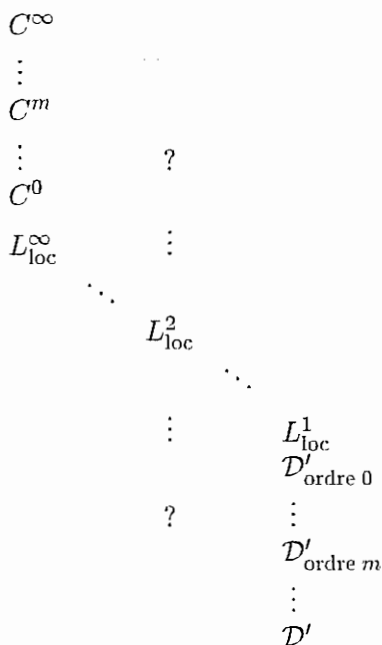


FIGURE 1. Une échelle bancale

La notion intuitive de régularité d'une fonction ou d'une distribution peut se traduire par un grand nombre de concepts mathématiques. On peut dire qu'une fonction très dérivable est plus régulière qu'une fonction peu dérivable, qui est elle-même plus régulière qu'une distribution d'ordre élevé. Mais on peut dire aussi qu'une fonction localement bornée est plus régulière qu'une fonction localement sommable, cette dernière pouvant avoir des "pics" en $1/|x|^{n-\varepsilon}$, tandis que les fonctions localement de carré sommable occupent une position intermédiaire.

La figure 1 illustre ce qui précède, les espaces que nous avons utilisés jusqu'ici forment une "échelle" où la régularité décroît de haut en bas, mais selon deux types de critères assez différents : perte de dérivabilité ou perte du caractère localement borné. Elle montre également le caractère central de l'espace L^2 . Le but de cette section est d'introduire une échelle bien rectiligne d'espaces, qui prendront la place des points d'interrogation dans la figure 1, construite à partir de l'espace L^2 et de la dérivation au sens des distributions.

L'appartenance à un espace fonctionnel traduit parfois une propriété purement locale ($C^\infty, \mathcal{D}', L^1_{\text{loc}} \dots$), que l'on peut interpréter comme une question de régularité, et, dans d'autres cas ($C^\infty_0, L^1, L^2, L^\infty \dots$), ajoute à cette condition de régularité une restriction sur le comportement à l'infini. C'est ce dernier cas qui se produira pour les espaces de Sobolev que nous allons introduire.

B. Définition et propriétés

Définition 6.2.1 (Espaces de Sobolev). — Soit m un entier positif ou nul. On dit que $u \in H^m(\mathbb{R}^n)$ si $u \in L^2(\mathbb{R}^n)$ et si les dérivées de u , au sens des distributions, jusqu'à l'ordre m appartiennent également à $L^2(\mathbb{R}^n)$.

Si l'on utilise souvent les espaces de Sobolev, en apparence moins naturels, au lieu des espaces C^m , ce n'est pas sans raison. Le théorème suivant fait apparaître une supériorité importante des espaces H^m , dont la norme est en outre étroitement reliée au concept d'énergie. Enfin, en théorie des équations aux dérivées partielles, nombre de propriétés que l'on souhaiterait voir vérifiées sont fausses dans les espaces C^m (de justesse, mais fausses quand même) alors que leurs homologues dans les espaces de Sobolev sont vraies.

Théorème 6.2.2. — Les espaces $H^m(\mathbb{R}^n)$ sont hilbertisables : munis du produit scalaire

$$(u | v)_m = \sum_{|\alpha| \leq m} \int \partial^\alpha u(x) \overline{\partial^\alpha v(x)} dx \quad (6.9)$$

ou de tout autre produit scalaire donnant une norme équivalente à la norme

$$\|u\|_m = \sqrt{\sum_{|\alpha| \leq m} \|\partial^\alpha u\|_{L^2}^2},$$

ce sont des espaces de Hilbert.

On voit facilement que l'expression (6.9) est un produit scalaire. Il suffit de vérifier que l'espace est complet pour la norme associée, et il le sera évidemment pour toute norme équivalente. Soit donc u_j une suite de Cauchy pour cette

norme. On a

$$\lim_{j,k \rightarrow \infty} \sum_{|\alpha| \leq m} \|\partial^\alpha u_j - \partial^\alpha u_k\|_{L^2}^2 = 0.$$

ce qui exprime que, pour chaque α de longueur inférieure ou égale à m , la suite $\partial^\alpha u_j$ est une suite de Cauchy dans L^2 . Ce dernier espace étant complet, il existe donc des v_α tels que $\partial^\alpha u_j \rightarrow v_\alpha$ dans L^2 . La convergence dans L^2 impliquant la convergence au sens des distributions, on a également $u_j \rightarrow v_0$ et par continuité, $\partial^\alpha u_j \rightarrow \partial^\alpha v_0$ au sens des distributions. On a donc $\partial^\alpha v_0 = v_\alpha \in L^2$, ce qui prouve que $v_0 \in H^m$.

Il reste à prouver que $u_j \rightarrow v_0$ pour la norme de H^m . On a

$$\|v_0 - u_j\|_m^2 = \sum \|v_\alpha - \partial^\alpha u_j\|_{L^2}^2$$

et les normes figurant dans le membre de droite tendant vers 0 par définition même des v_α , cela achève la démonstration.

Théorème 6.2.3. — *L'espace $C_0^\infty(\mathbb{R}^n)$ est dense dans H^m .*

Nous verrons au chapitre 10 une démonstration de ce résultat très important dans un cadre plus général. Nous proposons ci-dessous, à titre d'exercice, des indications sur une démonstration directe par troncature et régularisation.

Nous admettrons d'abord que, pour $\chi \in C_0^\infty$, la relation $\partial^\alpha (u * \chi) = (\partial^\alpha u) * \chi$, valable (théorème 3.3.4) pour $u \in C^m$ est encore valable pour $u \in H^m$. Nous verrons que ce résultat est toujours vrai (pour $u \in \mathcal{D}'$ quelconque) dans le chapitre 7.

Soit χ_ε une approximation de l'identité. D'après le théorème 3.4.3, on a alors $(\partial^\alpha u) * \chi_\varepsilon \rightarrow \partial^\alpha u$ dans L^2 pour $|\alpha| \leq m$, et donc

$$\|u - u * \chi_\varepsilon\|_{H^m} \rightarrow 0.$$

Un nombre $\beta > 0$ étant fixé, on peut choisir ε tel que, en posant $v = u * \chi_\varepsilon$, on ait $\|u - v\|_m \leq \beta/2$. La fonction v appartient à la fois à H^m et à C^∞ .

Soit maintenant $\psi \in C_0^\infty$, égale à 1 sur la boule de rayon 1, et posons $v_R(x) = v(x)\psi(x/R)$. On a

$$\partial^\alpha (v_R - v)(x) = \partial^\alpha v(x) \left(\psi\left(\frac{x}{R}\right) - 1 \right) + \sum_{\gamma \leq \alpha; |\gamma| > 0} \binom{\alpha}{\gamma} \frac{1}{R^{|\gamma|}} (\partial^{\alpha-\gamma} v(x)) (\partial^\gamma \psi(\frac{x}{R})).$$

Le lecteur montrera que la norme L^2 du premier terme tend vers 0 lorsque R tend vers ∞ (utiliser le théorème de Lebesgue), et que les autres termes contiennent une puissance négative de R en facteur d'une fonction de L^2 dont la norme est bornée. En choisissant R assez grand pour que $\|v - v_R\|_m \leq \beta/2$, on obtient $\|u - v_R\|_m \leq \beta$. La fonction v_R appartenant à C_0^∞ , cela achève la démonstration.

Définition 6.2.4. — *Soit $m \in \mathbb{N}$. On note $H^{-m}(\mathbb{R}^n)$ l'espace des distributions u telles qu'il existe une constante C avec*

$$\forall \varphi \in C_0^\infty(\mathbb{R}^n), |\langle u, \varphi \rangle| \leq C \|\varphi\|_m. \quad (6.10)$$

Cette définition, où on peut remplacer la norme $\|\cdot\|_m$ par n'importe quelle norme équivalente, exprime que la forme linéaire $\varphi \mapsto \langle u, \varphi \rangle$ est continue lorsqu'on munit C_0^∞ de la norme de l'espace H^m . Pour $m = 0$, cette définition, comme la définition 6.2.1, redonne l'espace L^2 . C'est si on veut un cas particulier très simple du théorème 6.2.7 ci-dessous.

Théorème 6.2.5 (Extension de la dualité). — (a) Pour tout $u \in H^{-m}$, l'application $\varphi \mapsto \langle u, \varphi \rangle$ se prolonge en une forme linéaire continue sur H^m . On notera

$$\langle u, v \rangle_{H^{-m}, H^m}, \quad u \in H^{-m}, \quad v \in H^m,$$

ce prolongement (en omettant parfois l'indice du crochet de dualité).

(b) L'extension précédente identifie H^{-m} au dual de H^m : pour toute forme linéaire L continue sur H^m , il existe un et un seul $u \in H^{-m}$ tel que

$$\forall v \in H^m, \quad L(v) = \langle u, v \rangle_{H^{-m}, H^m}.$$

L'application linéaire $\varphi \mapsto \langle u, \varphi \rangle$ est définie sur le sous-espace dense C_0^∞ de H^m , et l'hypothèse assure qu'elle est continue, et donc uniformément continue. Elle admet donc un prolongement unique en une forme linéaire continue sur H^m .

Soit maintenant L une forme linéaire continue sur H^m . En particulier, pour K compact et $\varphi \in C_K^\infty$, on a

$$|L(\varphi)| \leq C \|\varphi\|_m \leq C' \sup_{x \in K, |\alpha| \leq m} |\partial^\alpha \varphi(x)|,$$

l'égalité de droite résultant de

$$\|\partial^\alpha \varphi\|_{L^2} \leq \sqrt{\mu(K)} \|\partial^\alpha \varphi\|_{L^\infty}$$

en notant $\mu(K)$ la mesure de Lebesgue de K . On peut en effet prendre comme constante C' le produit de C , de $\sqrt{\mu(K)}$ et du nombre de multiindices concernés. Cela prouve que la restriction de L à C_0^∞ est une distribution d'ordre $\leq m$ que l'on notera u . On a $|\langle u, \varphi \rangle| = |L(\varphi)| \leq C \|\varphi\|_m$ pour $\varphi \in C_0^\infty$, ce qui montre que $u \in H^{-m}$.

Les deux formes linéaires $v \mapsto L(v)$ et $v \mapsto \langle u, v \rangle_{H^{-m}, H^m}$ sont définies et continues sur H^m , et elles coïncident sur l'ensemble dense C_0^∞ . Elles coïncident donc partout, ce qui achève la démonstration du théorème.

Remarque 6.2.6 (Abondance de biens ne nuit pas). — Nous disposons de plusieurs espaces auxquels s'identifie le dual de l'espace hilbertisable H^m . Le théorème précédent l'identifie à l'espace de distributions H^{-m} . Cette identification sera dite *canonique* : elle résulte du prolongement par continuité de la dualité usuelle entre distributions et fonctions, et est indépendante du produit scalaire que l'on peut choisir sur H^m .

Pour chaque choix d'un produit scalaire (noté $(\cdot | \cdot)_*$) sur H^m , on dispose en outre de l'identification de l'espace avec son dual qui est valable dans tout espace de Hilbert (théorème 2.5.10) : pour toute forme linéaire continue L sur H^m , il existe un unique $u \in H^m$ tel que l'on ait l'égalité $L(v) = (v | u)_*$ pour tout $v \in H^m$.

Nous verrons ci-dessous que, loin d'être une gêne, l'existence de plusieurs espaces s'identifiant au dual de H^m nous permettra d'obtenir des théorèmes importants d'existence de solutions.

Le résultat suivant fournit une caractérisation très simple des distributions appartenant à H^{-m} .

Théorème 6.2.7. — *Pour qu'une distribution u appartienne à H^{-m} , il faut et il suffit qu'elle soit somme de dérivées d'ordre $\leq m$ d'éléments de L^2 .*

Il est facile de voir que, si $v \in L^2$ et si $|\alpha| \leq m$, on a $\partial^\alpha v \in H^{-m}$. On a en effet, pour $\varphi \in C_0^\infty$,

$$|\langle \partial^\alpha v, \varphi \rangle| = |\langle v, \partial^\alpha \varphi \rangle| \leq \|v\|_{L^2} \|\partial^\alpha \varphi\|_{L^2}.$$

Le membre de droite est majoré par $C \|\varphi\|_m$, où C est la norme de v dans L^2 , ce qui assure précisément que $\partial^\alpha v \in H^{-m}$. La réciproque est un peu plus délicate.

Soit $A = \{\alpha \in \mathbb{N}^n \mid |\alpha| \leq m\}$, et considérons l'ensemble $(L^2)^A$ des familles $U = (u_\alpha)_{\alpha \in A}$ d'éléments de L^2 . On voit facilement que c'est un espace de Hilbert si on le munit du produit scalaire

$$(U | V) = \sum_{\alpha \in A} \int u_\alpha(x) \overline{v_\alpha(x)} dx.$$

Considérons d'autre part l'application D de H^m dans $(L^2)^A$ qui à u associe $Du = (\partial^\alpha u)_{\alpha \in A}$. Vu notre choix des produits scalaires, c'est une isométrie de H^m dans $(L^2)^A$. Si on note F l'image de D , l'application D est bijective de H^m sur F , et on notera $D^{-1} : F \rightarrow H^m$ l'isométrie inverse. Enfin, F étant isométrique à H^m et donc complet, c'est un sous-espace fermé de $(L^2)^A$. On notera P le projecteur orthogonal sur ce sous-espace.

Soit maintenant u un élément quelconque de H^{-m} . Considérons la forme linéaire suivante sur $(L^2)^A$:

$$L(W) = \langle u, D^{-1}PW \rangle_{H^{-m}, H^m}. \quad (6.11)$$

Elle est continue, comme composée de trois applications continues : $(L^2)^A \rightarrow F \rightarrow H^m \rightarrow \mathbb{C}$. Il existe donc (théorème 2.5.10) un élément V de l'espace de Hilbert $(L^2)^A$ tel que l'on ait

$$\forall W \in (L^2)^A, \quad L(W) = (W | V)_{(L^2)^A}. \quad (6.12)$$

Explicitons maintenant l'égalité des membres de droite dans (6.11) et (6.12) lorsque W est de la forme $(\partial^\alpha \varphi)_{\alpha \in A}$ avec $\varphi \in C_0^\infty$. On a alors $D^{-1}PW = \varphi$, et donc

$$\langle u, \varphi \rangle = \sum_{\alpha \in A} \int \partial^\alpha \varphi(x) \overline{v_\alpha(x)} dx,$$

où les composantes v_α appartiennent à L^2 . On a donc

$$\langle u, \varphi \rangle = \sum_{\alpha \in A} \langle \overline{v_\alpha}, \partial^\alpha \varphi \rangle = \langle \sum (-1)^{|\alpha|} \partial^\alpha \overline{v_\alpha}, \varphi \rangle.$$

Cela étant valable pour tout $\varphi \in C_0^\infty$, on a

$$u = \sum_{\alpha \in A} (-1)^\alpha \partial^\alpha \overline{v_\alpha},$$

ce qui achève la démonstration du théorème.

Le point important de cette démonstration a été, outre bien sûr le théorème de Riesz sur l'identification d'un espace de Hilbert avec son dual, l'argument déjà vu dans l'exercice 2.5.9 : prolongement à tout l'espace $(L^2)^A$ de la forme linéaire continue $W \mapsto \langle u, D^{-1}W \rangle$ qui n'est définie que sur l'image de D . On trouvera une autre démonstration de ce théorème dans l'exercice 6.2.12.

C. Applications

Considérons dans $\mathbb{R}^n$ l'équation aux dérivées partielles suivante

$$\Delta u - \lambda u = f$$

où Δ est le Laplacien et λ une constante > 0 , ou plus généralement l'équation suivante

$$\sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left\{ a_{ij}(x) \frac{\partial u}{\partial x_j}(x) \right\} - \lambda u(x) = f(x). \quad (6.13)$$

On suppose que les fonctions a_{ij} sont bornées dans $\mathbb{R}^n$ et que la matrice (a_{ij}) est symétrique réelle et uniformément définie positive, c'est-à-dire qu'il existe $c > 0$ tel que

$$\forall \xi \in \mathbb{C}^n, \forall x \in \mathbb{R}^n, \sum_{i,j=1}^n a_{ij}(x) \xi_i \overline{\xi_j} \geq c |\xi|^2. \quad (6.14)$$

Si on avait pris $\lambda = 0$, la première équation (qui correspond au cas où la matrice (a_{ij}) est la matrice unité) serait l'équation de Laplace-Poisson, qui intervient dans une grande quantité de phénomènes physiques dans un espace isotrope et homogène. On doit introduire une matrice définie positive générale dans un milieu non isotrope, et la matrice doit varier avec x dans un milieu non homogène. Enfin, il n'est pas sans intérêt d'autoriser le cas de coefficients discontinus (juxtaposition de deux milieux différents).

Nous verrons au chapitre 10 que les arguments qui vont suivre s'étendront au cas $\lambda = 0$ pour résoudre des équations du même type dans un ouvert borné de $\mathbb{R}^n$. Le cas $\lambda > 0$ correspond physiquement à un terme d'amortissement. Par exemple, la température [*resp.* le potentiel] au point x , dans une plaque conductrice de la chaleur [*resp.* de l'électricité] imparfaitement isolée du milieu

extérieur est, en régime permanent, solution de l'équation (6.13), le second membre f représentant les sources calorifiques [*resp.* électriques].

Nous désignerons par A l'application de $H^1(\mathbb{R}^n)$ dans $H^{-1}(\mathbb{R}^n)$ définie par

$$Au = \sum_{i,j=1}^n \partial_i (a_{ij} \partial_j u). \quad (6.15)$$

En effet on a alors $\partial_j u \in L^2$, son produit par la fonction bornée a_{ij} est encore dans L^2 , et la dérivée de ce produit appartient à H^{-1} .

Dans le cas où les coefficients a_{ij} sont réguliers, on peut bien sûr développer la dérivée du produit pour obtenir $Au = \sum a_{ij} \partial_i \partial_j u + \sum (\partial a_{ij} / \partial x_i) \partial_j u$. Par contre, lorsque les coefficients sont seulement bornés, seule la première forme a un sens. L'équation (6.13) est dite "sous forme variationnelle", nous verrons plus loin pourquoi. Pour des raisons tout à fait analogues à celles du n°5.4.15 ce n'est qu'en écrivant ces équations sous forme variationnelle, et en interprétant les dérivées au sens des distributions, que l'on a toute l'information venant des principes physiques (ici, la minimisation de l'énergie) qui ont servi à les établir.

Notre objectif est la démonstration du théorème suivant, qui contient comme cas particulier le cas où le second membre f appartient à L^2 .

Théorème 6.2.8. — *Pour $\lambda > 0$ et pour tout $f \in H^{-1}(\mathbb{R}^n)$, il existe un et un seul $u \in H^1(\mathbb{R}^n)$ tel que*

$$(A - \lambda)u = f$$

Nous aurons besoin du lemme suivant.

Lemme 6.2.9. — *L'expression*

$$(u | v)_* = \sum_{i,j=1}^n \int a_{ij} \partial_i u \overline{\partial_j v} dx + \int \lambda u \overline{v} dx$$

est un produit scalaire sur H^1 qui définit une norme équivalente à la norme usuelle.

Il est clair que l'expression est bien définie pour u et v dans H^1 , et que c'est un produit scalaire. D'autre part, les coefficients étant bornés, il existe une constante M telle que l'on ait $\sum a_{ij}(x) \xi_i \overline{\xi_j} \leq M |\xi|^2$. En appliquant cette dernière relation et (6.14) à $\xi = \nabla u(x)$, on obtient

$$c |\nabla u(x)|^2 \leq \sum a_{ij}(x) \partial_i u(x) \overline{\partial_j u(x)} \leq M |\nabla u(x)|^2.$$

En intégrant par rapport à x , et en ajoutant $\lambda \|u\|_{L^2}^2$, on obtient

$$\min(c, \lambda) \|u\|_1^2 \leq \|u\|_*^2 \leq \max(M, \lambda) \|u\|_1^2$$

ce qui prouve bien l'équivalence des normes.

Démonstration du théorème 6.2.8. — Donnons-nous $f \in H^{-1}$, et considérons la forme linéaire continue $v \mapsto \langle f, v \rangle_{H^{-1}, H^1}$. En vertu du théorème de Riesz (théorème 2.5.10) il existe donc $w \in H^1$ tel que

$$\forall v \in H^1, \langle f, v \rangle_{H^{-1}, H^1} = (v | w)_*.$$

En particulier, pour $\varphi \in C_0^\infty$, on a

$$\langle f, \varphi \rangle = \int \sum a_{ij} \frac{\partial \varphi}{\partial x_i} \frac{\partial \bar{w}}{\partial x_j} dx + \lambda \int \varphi \bar{w} dx,$$

et donc

$$\sum \langle a_{ij} \partial_j \bar{w}, \partial_i \varphi \rangle + \lambda \langle \bar{w}, \varphi \rangle = \left\langle - \sum \partial_i (a_{ij} \partial_j \bar{w}) + \lambda \bar{w}, \varphi \right\rangle = \langle f, \varphi \rangle.$$

Cela signifie que, en posant $u = -\bar{w}$, on a $(A - \lambda)u = f$ au sens des distributions.

L'unicité de u se démontre facilement. En effet, la différence u_0 de deux solutions vérifie $(A - \lambda)u_0 = 0$, et le calcul ci-dessus montre que l'on a alors, pour $\varphi \in C_0^\infty$,

$$(\varphi | \bar{u}_0)_* = 0.$$

L'élément $\bar{u}_0$ de l'espace de Hilbert H^1 est orthogonal au sous-espace dense C_0^∞ et est donc nul.

Remarque 6.2.10. — La démonstration précédente utilise un argument devenu très classique, dont des variantes sont connues sous le nom de *théorème de Lax-Milgram*, mais le lecteur non encore blasé peut trouver qu'elle tient du miracle. Elle est en fait étroitement reliée à des concepts du calcul des variations.

Considérons la fonctionnelle suivante, qui représente physiquement l'énergie du champ de potentiel u , en présence du terme d'amortissement représenté par λ et des sources représentées par f (en se limitant à des fonctions ou distributions réelles pour simplifier)

$$J(u) = \frac{1}{2} \int \sum_{ij} a_{ij} \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j} dx + \frac{1}{2} \lambda \int u^2 dx - \langle f, u \rangle.$$

Si u est un élément de H^1 tel que $J(u)$ soit égal au minimum de J sur H^1 , on a pour tout $\varphi \in C_0^\infty$

$$\begin{aligned} 0 \leq J(u + t\varphi) - J(u) &= t \left\{ \int \left(\sum a_{ij} \frac{\partial u}{\partial x_i} \frac{\partial \varphi}{\partial x_j} + \lambda u \varphi \right) dx - \langle f, \varphi \rangle \right\} \\ &\quad + \frac{t^2}{2} \left\{ \int \left(\sum a_{ij} \frac{\partial \varphi}{\partial x_i} \frac{\partial \varphi}{\partial x_j} + \lambda \varphi^2 \right) dx \right\}. \end{aligned}$$

Le minimum devant être atteint pour $t = 0$, le coefficient de t doit être nul, et ceci pour chaque $\varphi \in C_0^\infty$, ce qui signifie précisément que, au sens des distributions, on a

$$\sum \frac{\partial}{\partial x_i} \left\{ a_{ij} \frac{\partial u}{\partial x_j} \right\} - \lambda u = f.$$

Ce calcul est un cas typique de mise en évidence de ce que l'on appelle les équations d'Euler-Lagrange d'un problème de calcul des variations. Notre équation apparaît naturellement sous forme variationnelle (ce qui explique la terminologie que nous avons utilisée), les dérivées devant être entendues au sens des distributions.

Si on peut prouver que J atteint effectivement son minimum en un point de H^1 , cela démontrera l'existence d'une solution. Or, en introduisant l'élément u_0 de H^1 tel que $\langle f, v \rangle = (u_0 | v)_*$, pour tout $v \in H^1$, on a

$$J(u) = \frac{1}{2} (u | u)_* - \langle f, u \rangle = \frac{1}{2} (u | u)_* - (u_0 | u)_*,$$

et donc

$$J(u) = \frac{1}{2} \{ (u - u_0 | u - u_0)_* - (u_0 | u_0)_* \}.$$

Il est alors clair que le minimum est atteint pour $u = u_0$.

C'est le théorème de Riesz assurant qu'une forme linéaire continue sur un espace de Hilbert est donnée par le produit scalaire avec un élément de l'espace qui joue un rôle clef. Mais, comme nous l'avons vu, c'est un corollaire simple du théorème d'existence de la projection sur un convexe (un hyperplan en l'occurrence) fermé. Et ce théorème est l'un des précieux résultats dont on dispose pour prouver que des fonctionnelles (ici la distance à un convexe fermé) atteignent leur minimum.

Les fonctionnelles ne sont pas toujours quadratiques, et les équations d'Euler-Lagrange sont souvent non-linéaires. Il importe dans chaque cas de bien préciser l'espace fonctionnel dans lequel on travaille, de bien préciser en quel sens un extremum est solution "faible" des équations d'Euler-Lagrange, et de disposer de théorèmes (liés en général à la compacité dans l'espace fonctionnel utilisé) assurant l'existence d'extremums (ou de points critiques).

Exercice 6.2.11. — (a) En reprenant la démonstration du théorème 6.2.8 démontrer qu'il existe C tel que l'unique solution de

$$(\Delta - I)u = \partial f / \partial x_i \quad ; \quad f \in L^2$$

vérifie $\|u\|_{H^1} \leq C \|f\|_{L^2}$.

(b) En déduire que, pour $u \in H^2$, on a $\|u\|_{H^2} \leq C^{\text{te}} (\|\Delta u\|_{L^2} + \|u\|_{L^2})$. (On posera $f = (\Delta - I)u$ et on écrira l'équation satisfaite par $\partial u / \partial x_i$).

(c) En déduire que, pour $\lambda > 0$, l'expression

$$\sum_i \sum_j \int \frac{\partial^2 u}{\partial x_i^2} \frac{\partial^2 \bar{v}}{\partial x_j^2} dx + \lambda \int u \bar{v} dx$$

est un produit scalaire sur H^2 et que la norme correspondante est équivalente à la norme usuelle de H^2 .

(c) Démontrer qu'il existe, pour $\lambda > 0$ et pour tout $f \in H^{-2}$, une et une seule solution $u \in H^2$ de l'équation

$$(\Delta^2 + \lambda)u = f.$$

Exercice 6.2.12. — On considère l'opérateur $A = \sum_{|\alpha| \leq m} (-1)^{|\alpha|} \partial^{2\alpha}$, où m est un entier positif. Montrer que l'on a, pour $u \in H^m$ et $\varphi \in C_0^\infty$, $\langle Au, \bar{\varphi} \rangle = (u | \varphi)_m$, le produit scalaire étant défini par (6.9). Démontrer que A est un isomorphisme de H^m sur H^{-m} et en déduire une nouvelle démonstration du théorème 6.2.7.

6.3. Distributions périodiques

Dans ce qui suit, nous nous donnerons un nombre $T > 0$, et nous poserons $\omega = 2\pi/T$. On dit qu'une distribution u est *T-périodique* si sa translatée $\tau_T u$ est égale à u .

Théorème 6.3.1. — Soit $(\gamma_p)_{p \in \mathbb{Z}}$ une suite à croissance lente, c'est-à-dire vérifiant une estimation du type

$$\forall p \in \mathbb{Z}, |\gamma_p| \leq C(1 + |p|)^N.$$

La série

$$\sum_{p=-\infty}^{+\infty} \gamma_p e^{ip\omega t} \quad (6.16)$$

converge alors au sens des distributions et définit une distribution *T-périodique*.

Considérons en effet la série $\sum_{p \neq 0} (\gamma_p / (ip\omega)^{N+2}) e^{ip\omega t}$. Elle est uniformément convergente, la norme uniforme du terme général est $O(p^{-2})$. La série converge donc au sens des distributions et, d'après le théorème 4.3.4, elle peut être dérivée terme à terme. Il suffit de répéter $N + 2$ fois l'opération pour obtenir le résultat. D'autre part, chaque somme partielle S vérifiant $\tau_T S = S$, il en est de même de la somme de la série, les translations étant continues dans $\mathcal{D}'$.

Remarque 6.3.2. — Le terme général de la série tend bien sûr vers 0 dans $\mathcal{D}'$ et on peut s'étonner de voir la suite $p^{1000} e^{ip\omega t}$ tendre vers 0 en un sens raisonnable. Cela reflète le fait qu'un appareil physique (dont le comportement reste linéaire) est insensible aux très hautes fréquences, même de grande amplitude.

L'objectif de la série d'exercices qui vont suivre est de démontrer que, réciproquement, toute distribution périodique est somme d'une unique série de Fourier, dont les coefficients sont à croissance lente. Nous rappelons le théorème classique suivant.

Théorème 6.3.3. — Soit f une fonction *T-périodique* de classe C^1 par morceaux. Si on pose $c_p(f) = \int_a^{a+T} f(t) e^{-ip\omega t} dt / T$, les sommes partielles

symétriques $\sum_{p=-k}^k c_p(f) e^{ip\omega t}$ de la série de Fourier convergent uniformément vers f dans tout intervalle fermé formé de points où f est continue, et convergent vers $1/2(f(t-0) + f(t+0))$ en tout point de discontinuité.

Exercice 6.3.4. — Dédurre de ce qui précède que, si f est T -périodique et de classe C^∞ , la suite des $c_p(f)$ est à décroissance rapide, c'est-à-dire que l'on a $|c_p(f)| \leq C_N(1 + |p|)^{-N}$ pour tout N et que la différence

$$f - \sum_{-k}^{k'} c_p(f) e^{ip\omega t}$$

tend vers 0 uniformément ainsi que chacune de ses dérivées lorsque k et k' tendent vers $+\infty$.

Exercice 6.3.5. — Montrer qu'il existe $\chi \in C_0^\infty(\mathbb{R})$ telle que l'on ait $\sum_{k \in \mathbb{Z}} \tau_{kT} \chi = 1$. On pourra prendre d'abord $\varphi \in C_0^\infty$ positive partout et strictement positive sur $[0, T]$, et poser $\chi = \varphi / (\sum \tau_{kT} \varphi)$.

Exercice 6.3.6. — Soit $\alpha \in C_0^\infty(\mathbb{R})$ et posons $\tilde{\alpha} = \sum \tau_{kT} \alpha$. Montrer que $c_p(\tilde{\alpha}) = \int_{\mathbb{R}} \alpha(t) e^{-ip\omega t} dt / T$.

Exercice 6.3.7. — Montrer que pour (γ_p) à croissance lente et α comme ci-dessus, on a

$$\langle \sum \gamma_p e^{ip\omega t}, \alpha \rangle = \sum \gamma_p c_{-p}(\tilde{\alpha}),$$

et en déduire que l'application qui à une suite à croissance lente γ_p associe la distribution $\sum \gamma_p e^{ip\omega t}$ est injective (considérer $\alpha = \chi e^{ip_0\omega t}$).

Exercice 6.3.8. — Soit U une distribution T -périodique. Montrer qu'il existe $u \in \mathcal{E}'(\mathbb{R})$ telle que $U = \sum \tau_{kT} u$. En posant

$$\beta_p = \frac{1}{T} \langle u, e^{-ip\omega t} \rangle_{\mathcal{E}', C^\infty},$$

montrer que la suite β_p est à croissance lente et que l'on a

$$\langle U, \alpha \rangle = \langle u, \tilde{\alpha} \rangle = \sum \beta_p c_{-p}(\tilde{\alpha}).$$

En déduire que $U = \sum \beta_p e^{ip\omega t}$.

Cela prouve donc le résultat suivant.

Théorème 6.3.9. — Toute distribution T -périodique U peut s'écrire de manière unique sous forme d'une série de Fourier convergente dans $\mathcal{D}'$

$$U = \sum_{p \in \mathbb{Z}} c_p(U) e^{ip\omega t},$$

où la suite $(c_p(U))$ est à croissance lente. Les coefficients de Fourier peuvent se calculer par $c_p(U) = T^{-1} \langle u, e^{-ip\omega t} \rangle$, en choisissant une distribution u à support compact telle que $U = \sum \tau_{kT} u$.

On a ainsi obtenu, comme dans le cas des fonctions de carré sommable, une caractérisation complète des distributions périodiques en termes de leur série de Fourier. Dans la plupart des espaces fonctionnels, et notamment pour les fonctions périodiques sommables ou continues, la situation est très loin d'être aussi simple, et de nombreux problèmes sont encore ouverts.

Exercice 6.3.10. — Démontrer que l'on a l'égalité suivante, les séries convergeant au sens des distributions,

$$\sum_{k=-\infty}^{+\infty} \delta_{kT} = \frac{1}{T} \sum_{p=-\infty}^{+\infty} e^{ip\omega t}.$$

En déduire que, pour $\varphi \in C_0^\infty(\mathbb{R})$ (cette hypothèse peut être notablement affaiblie), on a

$$\sum_{k=-\infty}^{+\infty} \varphi(2k\pi) = \frac{1}{2\pi} \sum_{p=-\infty}^{+\infty} \widehat{\varphi}(p),$$

où on a noté $\widehat{\varphi}(p) = \int e^{-ip t} \varphi(t) dt$ la transformée de Fourier de φ . Cette relation est connue sous le nom de *formule sommatoire de Poisson*.

CHAPITRE 7

CONVOLUTION

7.1. Préliminaires

Les deux théorèmes qui vont suivre se réduisent respectivement au théorème de dérivation sous le signe somme et au théorème de Fubini lorsque la distribution u est en fait une fonction localement sommable.

Théorème 7.1.1 (Dérivation sous le crochet). — Soit $u \in \mathcal{E}'(\mathbb{R}^p)$. et soit $\varphi \in C^\infty(\mathbb{R}^{p+q})$. Alors la fonction $y \mapsto \langle u(x), \varphi(x, y) \rangle$ est de classe C^∞ dans $\mathbb{R}^q$, et on a

$$\partial_y^\alpha \langle u(x), \varphi(x, y) \rangle = \langle u(x), \partial_y^\alpha \varphi(x, y) \rangle. \quad (7.1)$$

Posons $f(y) = \langle u(x), \varphi(x, y) \rangle$ et montrons d'abord la continuité de f . Si K est un voisinage compact du support de u , et si y_j est une suite tendant vers y_0 , on a (voir remarque 6.1.7)

$$|f(y_j) - f(y_0)| \leq C \sup_{x \in K, |\beta| \leq p} \left| \partial_x^\beta (\varphi(x, y_j) - \varphi(x, y_0)) \right| \quad (7.2)$$

où p est l'ordre de u . Les dérivées de φ étant uniformément continues sur le produit de K par un voisinage compact de y_0 , le membre de droite tend vers 0 pour $j \rightarrow \infty$, ce qui prouve la continuité de f .

Montrons maintenant l'existence des dérivées partielles $\partial f / \partial y_i$ en un point y fixé. En appelant e_i le i^e vecteur de base, on a

$$\frac{1}{h} (f(y + h e_i) - f(y)) - \langle u(x), \frac{\partial \varphi}{\partial y_i}(x, y) \rangle = \langle u(x), \psi_h(x, y) \rangle, \quad (7.3)$$

en posant

$$\psi_h(x, y) = \frac{1}{h} (\varphi(x, y + h e_i) - \varphi(x, y)) - \frac{\partial \varphi}{\partial y_i}(x, y).$$

Il est facile de montrer que ψ_h ainsi que ses dérivées en x d'ordre $\leq p$ convergent uniformément vers 0 pour $x \in K$. L'estimation utilisée ci-dessus dans (7.2) montre alors que le membre de gauche de (7.3) tend vers 0 avec h . Cela prouve l'existence des dérivées partielles premières, avec

$$\partial f / \partial y_i(y) = \langle u(x), \partial \varphi / \partial y_i(x, y) \rangle.$$

En remplaçant φ par $\partial \varphi / \partial y_i$, l'argument qui nous a permis de prouver la continuité de f montre la continuité de $\partial f / \partial y_i$. Nous avons ainsi démontré que f est de classe C^1 et que la formule (7.1) est valable pour les dérivées d'ordre 1. On conclut par récurrence.

L'énoncé du théorème qui va suivre nous suffira, mais les hypothèses pourraient en être notablement affaiblies : la dérivabilité de φ par rapport à y n'a rien à voir dans l'affaire, et les pavés n'ont pas de vertus particulières pour l'intégration.

Théorème 7.1.2 (Intégration sous le crochet). — Soient $u \in \mathcal{E}'(\mathbb{R}^p)$ et $\varphi \in C^\infty(\mathbb{R}^{p+q})$. Soit Q un pavé (produit d'intervalles) compact de $\mathbb{R}^q$. On a alors

$$\langle u(x), \int_Q \varphi(x, y) dy \rangle = \int_Q \langle u(x), \varphi(x, y) \rangle dy.$$

Il suffit de prouver le théorème lorsque $q = 1$ et $Q = [a, b]$, ce que nous supposons désormais, le cas général s'y ramenant en intégrant successivement par rapport à $y_1, \dots, y_q$. Posons

$$F(y) = \langle u(x), \int_a^y \varphi(x, t) dt \rangle.$$

En appliquant le théorème précédent, on obtient

$$F'(y) = \langle u(x), \varphi(x, y) \rangle,$$

et donc

$$F(b) - 0 = \int_a^b \langle u(x), \varphi(x, y) \rangle dy,$$

ce qui établit le résultat.

Pour l'étude du produit de convolution, les concepts suivants nous permettront d'étendre systématiquement à des situations beaucoup plus générales des résultats obtenus sous des conditions de supports compacts.

Définition 7.1.3 (Ensembles convolutifs). — On dit que deux ensembles fermés F et G de $\mathbb{R}^n$ sont convolutifs (on dit aussi que le couple (F, G) est convolutif) si, pour tout $R > 0$, il existe $\rho(R)$ tel que

$$(x \in F, y \in G, |x + y| \leq R) \Rightarrow (|x| \leq \rho(R), |y| \leq \rho(R)).$$

On dit de même qu'une famille finie $(F_i)_{i \in I}$ d'ensembles fermés est convolutive si pour tout sous-ensemble J de I et pour tout R , il existe $\rho(R)$ tel que

$$(x_j \in F_j, \left| \sum_{j \in J} x_j \right| \leq R) \Rightarrow (\forall j \in J, |x_j| \leq \rho(R)).$$

Le lecteur vérifiera facilement les propriétés suivantes

1. Si les $F_1, \dots, F_p$ sont fermés et si tous, sauf peut-être un, sont compacts, ils sont convolutifs.
2. En adjoignant q ensembles compacts à un p -uple convolutif, on obtient un $(p+q)$ -uple convolutif.
3. Sur $\mathbb{R}$, les intervalles $[a_1, +\infty[, \dots, [a_p, +\infty[$ forment une famille convolutive.
4. Dans l'espace-temps $\mathbb{R} \times \mathbb{R}^3$, dont on notera $(t, \mathbf{r})$ le point courant, avec $r = |\mathbf{r}|$, le "cône d'avenir" $G = \{(t, \mathbf{r}) | t \geq r\}$ et $\mathbb{R}_+^4 = \{(t, \mathbf{r}) | t \geq 0\}$ sont convolutifs.
5. Si F et G contiennent respectivement deux demi-droites parallèles et de direction opposées, ils ne sont pas convolutifs.

Proposition 7.1.4. — Si F et G sont convolutifs, l'ensemble somme

$$F + G = \{x \mid \exists y \in F, \exists z \in G, x = y + z\}$$

est un sous-ensemble fermé de $\mathbb{R}^n$.

Il suffit de montrer que, pour toute suite x_j d'éléments de $F + G$ qui converge vers un point x_0 , on a $x_0 \in F + G$. Il existe alors $y_j \in F$ et $z_j \in G$ tels que $y_j + z_j = x_j$. La suite convergente x_j est majorée en norme par un certain nombre R , et l'hypothèse assure alors que les y_j et z_j sont majorés en norme par $\rho(R)$.

En extrayant une première, puis une seconde sous-suite (nous omettrons de changer les noms des indices) on peut supposer que $y_j \rightarrow y_0$ et $z_j \rightarrow z_0$. On a $y_0 \in F$ et $z_0 \in G$, ces ensembles étant fermés. On a enfin par continuité $x_0 = y_0 + z_0$, ce qui achève la démonstration.

7.1.5. Utilisation des ensembles convolutifs. — Nous aurons à répéter plusieurs fois, dans des contextes différents, l'argument qui va suivre. C'est pourquoi nous en donnons une présentation abstraite.

Soient $\mathcal{A}$ et $\mathcal{B}$ deux sous-espaces vectoriels de $\mathcal{D}'(\mathbb{R}^n)$ qui sont stables par multiplication par C_0^∞ (c'est-à-dire que $u \in \mathcal{A}$ et $\varphi \in C_0^\infty$ entraîne $\varphi u \in \mathcal{A}$). Dans les applications, ces espaces seront L_{loc}^1 ou C^∞ ou $\mathcal{D}'$. Supposons de plus que l'on ait une application bilinéaire (notons-la $(u, v) \mapsto u * v$) de $(\mathcal{A} \cap \mathcal{E}') \times (\mathcal{B} \cap \mathcal{E}')$ dans $\mathcal{E}'$ vérifiant

$$\text{Supp}(u * v) \subset \text{Supp}(u) + \text{Supp}(v). \quad (7.4)$$

On a alors

1. L'application $(u, v) \mapsto u * v$ s'étend naturellement aux couples $(u, v) \in \mathcal{A} \times \mathcal{B}$ dont les supports sont convolutifs.
2. Pour de tels couples, on a encore la relation (7.4).

L'extension de l'opération $*$ est définie de la façon suivante. On prend une suite de fonctions $\theta_j \in C_0^\infty$, telles que l'on ait $\theta_j = 1$ pour $|x| \leq j$. Pour chaque $\varphi \in C_0^\infty$, on pose

$$\langle u * v, \varphi \rangle = \lim_{j \rightarrow \infty} \langle (\theta_j u) * (\theta_j v), \varphi \rangle, \quad (7.5)$$

la suite figurant au membre de droite étant convergente au sens le plus fort qui soit : elle est constante à partir d'un certain rang.

Démontrons ce dernier point. Soit R tel que le support de φ soit contenu dans la boule de rayon R . Nous allons montrer que dès que l'on a $j \geq \rho(R)$, où ρ est la fonction de la définition 7.1.3, le membre de droite de (7.5) reste constant.

Soient en effet j et k supérieurs à $\rho(R)$. On a

$$\begin{aligned} \langle (\theta_j u) * (\theta_j v) - (\theta_k u) * (\theta_k v), \varphi \rangle &= \langle (\theta_j u - \theta_k u) * (\theta_j v), \varphi \rangle \\ &\quad + \langle (\theta_k u) * (\theta_j v - \theta_k v), \varphi \rangle. \end{aligned}$$

Le premier terme (le raisonnement est identique pour le second) ne pourrait être non nul que si l'intersection de $\text{Supp}((\theta_j - \theta_k)u) + \text{Supp}(\theta_j v)$ et du support de φ était non vide. Il existerait alors un point $x = y + z$ avec $x \in \text{Supp}(\varphi)$, $y \in \text{Supp}((\theta_j - \theta_k)u)$ et $z \in \text{Supp}(\theta_j v)$. On aurait alors $|y + z| < R$, $y \in \text{Supp}(u)$, $|y| > \rho(R)$, $z \in \text{Supp}(v)$ et c'est précisément ce qu'interdit l'hypothèse de convolutivité.

On voit facilement (directement ou par le théorème 4.2.3) que la forme linéaire $u * v$ définie par (7.5) est une distribution. Il reste à vérifier la propriété de support, et donc à montrer que, si $\text{Supp}(\varphi)$ ne rencontre pas $\text{Supp}(u) + \text{Supp}(v)$, on a $\langle u * v, \varphi \rangle = 0$. Or, sous cette hypothèse, $\text{Supp}(\varphi)$ ne rencontre pas a fortiori $\text{Supp}(\theta_j u) + \text{Supp}(\theta_j v)$, on a donc $\langle (\theta_j u) * (\theta_j v), \varphi \rangle = 0$ et le résultat.

Remarque 7.1.6. — La démonstration précédente montre bien sûr que la définition de $u * v$ ne dépend pas de la suite θ_j choisie. En fait, la restriction de $u * v$ à la boule de rayon R est égale à celle de $(\theta u) * (\theta v)$ dès que la fonction $\theta \in C_0^\infty$ est égale à 1 sur la boule de rayon $\rho(R)$.

Exercice 7.1.7. — Soient f et g localement sommables dans $\mathbb{R}^n$ et dont les supports sont convolutifs. Montrer que le produit de convolution $f * g$ peut

être défini par la méthode ci-dessus, et qu'il peut également être défini par

$$(f \star g)(x) = \int f(x-y)g(y) dy,$$

l'intégrale étant convergente pour presque tout x .

7.2. Convolution d'une distribution et d'une fonction C^∞

Lorsque u est une fonction sommable, on peut écrire, pour $\varphi \in C_0^\infty$,

$$(u \star \varphi)(x) = \int \varphi(x-y)u(y) dy = \langle u(y), \varphi(x-y) \rangle.$$

Cette dernière expression conserve un sens lorsque u est une distribution et va nous servir à définir le produit de convolution dans ce cas. La nouvelle définition, appliquée à une fonction sommable, coïncidera donc avec l'ancienne.

Théorème et Définition 7.2.1. — Soient $u \in \mathcal{E}'(\mathbb{R}^n)$ et $\varphi \in C_0^\infty(\mathbb{R}^n)$. Leur produit de convolution est défini en chaque point x par

$$(u \star \varphi)(x) = \langle u(y), \varphi(x-y) \rangle.$$

La fonction $u \star \varphi$ appartient à C_0^∞ , et on a

$$\partial^\alpha (u \star \varphi) = u \star (\partial^\alpha \varphi) = (\partial^\alpha u) \star \varphi, \quad (7.6)$$

$$\text{Supp}(u \star \varphi) \subset \text{Supp}(u) + \text{Supp}(\varphi).$$

Le théorème de dérivation sous le crochet 7.1.1 montre immédiatement que $u \star \varphi$ est de classe C^∞ et que

$$\partial^\alpha (u \star \varphi)(x) = \langle u(y), \partial_x^\alpha \varphi(x-y) \rangle,$$

ce qui prouve la première égalité. D'autre part, en remarquant que $\partial_x^\alpha \varphi(x-y) = (-1)^{|\alpha|} \partial_y^\alpha \varphi(x-y)$, le second membre de (7.6) est égal, par définition de la dérivation des distributions, à $\langle \partial^\alpha u(y), \varphi(x-y) \rangle$, ce qui prouve la seconde égalité.

Enfin, le support de $u \star \varphi$ est compact. Plus précisément, il est contenu dans $\text{Supp}(u) + \text{Supp}(\varphi)$. En effet, si x n'appartient pas à cet ensemble, les supports de u et de $y \mapsto \varphi(x-y)$ sont disjoints.

Théorème 7.2.2. — Soient $u \in \mathcal{E}'(\mathbb{R}^n)$, φ et $\psi \in C_0^\infty(\mathbb{R}^n)$. On a alors

$$(u \star \varphi) \star \psi = u \star (\varphi \star \psi)$$

et

$$\langle u \star \varphi, \psi \rangle = \langle u, \tilde{\varphi} \star \psi \rangle \quad (7.7)$$

Le théorème d'intégration sous le crochet donne immédiatement

$$\left\langle u(y), \int_Q \varphi(x-y-z)\psi(z) dz \right\rangle = \int_Q \langle u(y), \varphi(x-y-z) \rangle \psi(z) dz,$$

où Q est un pavé compact, que l'on choisira contenant le support de ψ . Les intégrales sur Q sont alors égales aux intégrales sur $\mathbb{R}^n$. Le membre de gauche est égal à $(u \star (\varphi \star \psi))(x)$, tandis que le membre de droite vaut $((u \star \varphi) \star \psi)(x)$, ce qui démontre la première partie du théorème.

On a de même

$$\left\langle u(x), \int \varphi(y-x)\psi(y) dy \right\rangle = \int \langle u(x), \varphi(y-x) \rangle \psi(y) dy. \quad (7.8)$$

En notant que $\tilde{\varphi}(x-y) = \varphi(y-x)$, on voit que le membre de gauche de (7.8) est égal au membre de droite de (7.7) et vice versa, d'où le résultat.

Corollaire 7.2.3. — Soient $u \in \mathcal{D}'(\mathbb{R}^n)$ et $\varphi \in C^\infty(\mathbb{R}^n)$ tels que leurs supports soient convolutifs. Le procédé du n° 7.1.5 permet de définir $u \star \varphi$. C'est une fonction appartenant à $C^\infty(\mathbb{R}^n)$, à support dans $\text{Supp}(u) + \text{Supp}(\varphi)$.

Si on se donne en outre $\psi \in C^\infty(\mathbb{R}^n)$, et si les supports de u, φ, ψ forment un triplet convolutif, on a

$$(u \star \varphi) \star \psi = u \star (\varphi \star \psi).$$

Les conditions d'application du n° 7.1.5 sont réalisées, avec $\mathcal{A} = \mathcal{D}'$ et $\mathcal{B} = C^\infty$, et $w = u \star \varphi$ est donc une distribution bien définie, dont le support est contenu dans la somme de ceux de u et φ . On sait en outre que la restriction de w à une boule quelconque $B(R)$ est égale à la restriction de $\theta u \star \theta \varphi$, pour une fonction $\theta \in C_0^\infty$ convenable. La restriction de w à chaque boule étant de classe C^∞ , on a $w \in C^\infty(\mathbb{R}^n)$.

Lorsque les supports de u, φ, ψ forment un triplet convolutif (on notera $\rho(R)$ la fonction associée), nous laissons au lecteur le soin de montrer que, si u_1, φ_1, ψ_1 sont les produits respectifs de u, φ, ψ par une fonction $\theta \in C_0^\infty$ égale à 1 au voisinage de $\overline{B(\rho(R))}$, on a

$$\begin{aligned} \forall x \in B(R), \quad (u \star \varphi) \star \psi(x) &= (u_1 \star \varphi_1) \star \psi_1(x) \\ u \star (\varphi \star \psi)(x) &= u_1 \star (\varphi_1 \star \psi_1)(x). \end{aligned}$$

Les membres de gauche coïncident dans toute boule et donc partout.

Remarque 7.2.4. — Les cas les plus fréquents d'utilisation de la condition précédente sont d'une part le cas $u \in \mathcal{E}', \varphi \in C^\infty$, et d'autre part le cas $u \in \mathcal{D}', \varphi \in C_0^\infty$. Dans les deux cas (vérification facile), on a encore

$$(u \star \varphi)(x) = \langle u(y), \varphi(x-y) \rangle,$$

le membre de droite étant bien défini.

On pourra, à titre d'exercice, remarquer que, lorsque $\text{Supp}(u)$ et $\text{Supp}(\varphi)$ sont convolutifs, l'expression $\langle u(y), \varphi(x - y) \rangle$ est bien définie grâce à l'extension de la dualité du n°6.1.8, et qu'elle est égale à $u \star \varphi(x)$ telle que nous l'avons définie.

Théorème 7.2.5. — Soient Ω un ouvert de $\mathbb{R}^n$ et $u \in \mathcal{D}'(\Omega)$. Il existe alors une suite de fonctions $f_j \in C_0^\infty(\Omega)$ qui converge vers u au sens des distributions.

Nous allons maintenant pouvoir procéder par troncature et régularisation comme nous l'avons fait pour les fonctions dans le théorème 3.5.2. Soient donc ψ_j une suite d'éléments de $C_0^\infty(\Omega)$, égaux à 1 sur une suite exhaustive de compacts, et $\varepsilon(j)$ une suite décroissante, tendant vers 0, telle que $\varepsilon(j)$ soit inférieur à la moitié de la distance de $\text{Supp}(\psi_j)$ au complémentaire de Ω . Soit enfin χ_ε une approximation de l'identité, et posons

$$f_j = (\psi_j u) \star \chi_{\varepsilon(j)}.$$

Pour chaque fonction d'essai φ dans Ω , on a

$$\langle f_j, \varphi \rangle = \langle \psi_j u, \tilde{\chi}_{\varepsilon(j)} \star \varphi \rangle, \quad (7.9)$$

d'après (7.7). La suite $\tilde{\chi}_\varepsilon$ étant elle aussi une approximation de l'identité, la suite $\tilde{\chi}_{\varepsilon(j)} \star \varphi$ garde son support dans un compact fixe K de Ω pour j assez grand, et converge vers φ uniformément ainsi que chacune de ses dérivées (voir théorème 3.4.3).

On a donc $\langle u, \tilde{\chi}_{\varepsilon(j)} \star \varphi \rangle \rightarrow \langle u, \varphi \rangle$. Mais d'autre part, les fonctions ψ_j étant égales à 1 au voisinage de K pour j assez grand, on a également

$$\langle \psi_j u, \tilde{\chi}_{\varepsilon(j)} \star \varphi \rangle = \langle u, \tilde{\chi}_{\varepsilon(j)} \star \varphi \rangle \rightarrow \langle u, \varphi \rangle. \quad (7.10)$$

Il résulte de (7.9) et de (7.10) que $\langle f_j, \varphi \rangle \rightarrow \langle u, \varphi \rangle$, et donc le résultat.

Remarque 7.2.6. — Ce théorème important permettra de démontrer des propriétés sur les distributions en passant à la limite à partir de propriétés connues pour les fonctions. Il signifie également que nous n'avons pas introduit un espace "trop gros" de fonctions généralisées. Comme nous l'avons suggéré dans la section 4.1, on peut aussi définir $\mathcal{D}'$ comme le plus petit espace contenant les fonctions et leurs limites (au sens des distributions).

En reprenant (sans les troncatures) la démonstration ci-dessus, on montre facilement le résultat suivant.

Corollaire 7.2.7. — Si u appartient à $\mathcal{D}'(\mathbb{R}^n)$, ses régularisées $u \star \chi_\varepsilon$ sont des fonctions de classe C^∞ qui convergent vers u au sens des distributions.

Exercice 7.2.8. — Il s'agit de démontrer le résultat suivant : si toutes les dérivées $\partial_i u$ (au sens des distributions) d'une distribution $u \in \mathcal{D}'(\mathbb{R}^n)$ sont des fonctions continues, alors u est en fait une fonction de classe C^1 .

(a) Démontrer, en utilisant le théorème 5.3.7, que la propriété est vraie en dimension 1, et se convaincre du fait que le résultat n'est pas évident en dimension supérieure.

(b) Soient χ_ε une approximation de l'identité, et $f_\varepsilon = u \star \chi_\varepsilon$. On écrira

$$f_\varepsilon(x) = f_\varepsilon(0) + \sum_1^n \int_0^1 x_i \partial_i f_\varepsilon(tx) dt = f_\varepsilon(0) + g_\varepsilon(x).$$

Démontrer successivement que $\partial_i f_\varepsilon$ converge uniformément sur tout compact vers $\partial_i u$ et que g_ε converge uniformément sur tout compact vers une fonction continue g .

(c) Soit φ une fonction de C_0^∞ d'intégrale non nulle. Montrer que $\langle f_\varepsilon, \varphi \rangle \rightarrow \langle u, \varphi \rangle$ et en déduire que $f_\varepsilon(0)$ tend vers une limite. Montrer que f_ε converge uniformément sur tout compact vers une fonction continue et conclure.

7.3. Convolution et translations

A. Propriété caractéristique de la convolution

Il résulte immédiatement de la définition que le produit de convolution commute avec les translations. On a en effet, pour $u \in \mathcal{E}'(\mathbb{R}^n)$ et pour $\varphi \in C_0^\infty(\mathbb{R}^n)$

$$\langle u(y), \varphi(x - (a + y)) \rangle = \langle \tau_a u(y), \varphi(x - y) \rangle,$$

et le premier membre peut aussi bien s'interpréter comme la valeur en $x - a$ de $u \star \varphi$ que comme la valeur en x de $u \star (\tau_a \varphi)$. On a donc

$$\tau_a(u \star \varphi) = (\tau_a u) \star \varphi = u \star (\tau_a \varphi). \quad (7.11)$$

Une remarque qui nous sera utile est la suivante. On a $(u \star \varphi)(0) = \langle u(y), \varphi(-y) \rangle$, et donc, en remplaçant φ par $\tilde{\varphi}$

$$\langle u, \varphi \rangle = (u \star \tilde{\varphi})(0). \quad (7.12)$$

Les opérateurs de convolution peuvent être caractérisés par le fait qu'ils commutent avec les translations, si on adjoint la propriété de continuité suivante.

Théorème 7.3.1. — Soit $u \in \mathcal{E}'(\mathbb{R}^n)$ et soit U l'opérateur appliquant C_0^∞ dans lui-même défini par $U\varphi = u \star \varphi$. On a

$\forall K$ compact, $\forall p \in \mathbb{N}$, $\exists L$ compact, $\exists q \in \mathbb{N}$, $\exists C$,

$$\forall \varphi \in C_0^\infty(\mathbb{R}^n), \quad \sup_{x \in K, |\alpha| \leq p} |\partial^\alpha (U\varphi)(x)| \leq C \sup_{x \in L, |\beta| \leq q} |\partial^\beta \varphi(x)|. \quad (7.13)$$

Cette expression riche en quantificateurs signifie simplement que l'application U est continue (voir le théorème C.1.7) lorsqu'on munit le sous-espace C_0^∞ de $C^\infty(\mathbb{R}^n)$ de la topologie définie au n° C.2.6.

On a en effet, d'après (7.6)

$$\partial^\alpha (u \star \varphi)(x) = \langle u(y), (\partial^\alpha \varphi)(x - y) \rangle.$$

Si r est l'ordre de u , et si H est un voisinage compact du support de u , on peut d'après le théorème 6.1.3 estimer le membre de droite par une constante fois la borne supérieure sur H des dérivées d'ordre $\leq r$ de $y \mapsto \partial^\alpha \varphi(x - y)$. On a donc, $x - y$ parcourant le compact $K - H$ (différence algébrique) lorsque x parcourt K ,

$$\sup_{x \in K} |\partial^\alpha (u \star \varphi)(x)| \leq C \sup_{z \in K - H, |\beta| \leq |\alpha| + r} \left| \partial^\beta \varphi(z) \right|.$$

La propriété (7.13) en résulte immédiatement, avec $L = K - H$ et $q = p + r$.

Théorème 7.3.2 (Propriété caractéristique). — *Soit U une application linéaire de $C_0^\infty(\mathbb{R}^n)$ dans lui-même, qui commute avec les translations (c'est-à-dire telle que $U\tau_a\varphi = \tau_a U\varphi$) et qui possède la propriété de continuité (7.13). Il existe alors une unique distribution u à support compact telle que*

$$\forall \varphi \in C_0^\infty(\mathbb{R}^n), \quad U\varphi = u \star \varphi.$$

Supposons d'abord le problème résolu, on doit alors avoir

$$U\varphi(0) = (u \star \varphi)(0) = \langle u, \check{\varphi} \rangle = \langle \tilde{u}, \varphi \rangle.$$

Il en résulte que la distribution $\tilde{u}$ (et donc u elle-même) est, si elle existe, uniquement déterminée par U .

Considérons maintenant la forme linéaire v sur $C_0^\infty(\mathbb{R}^n)$ définie par $\langle v, \varphi \rangle = U\varphi(0)$. En utilisant la propriété (7.13) dans le cas $p = 0$, $K = \{0\}$, on obtient l'existence de q, L, C tels que

$$\forall \varphi \in C_0^\infty, \quad |\langle v, \varphi \rangle| = |U\varphi(0)| \leq C \sup_{|\alpha| \leq q, x \in L} |\partial^\alpha \varphi(x)|.$$

Cela exprime que v est une distribution, et qu'elle est à support dans L , le membre de droite étant nul si le support de φ est disjoint de L .

Pour $a \in \mathbb{R}^n$, la valeur de $U\varphi$ au point a est égale à la valeur de $\tau_{-a}U\varphi = U\tau_a\varphi$ à l'origine. On a donc

$$U\varphi(a) = \langle v, \tau_{-a}\varphi \rangle = \langle \tilde{v}, \tau_a\check{\varphi} \rangle = \langle \tilde{v}(y), \varphi(a - y) \rangle.$$

On a donc $U\varphi = \tilde{v} \star \varphi$, ce qui achève la démonstration.

Exercice 7.3.3. — Démontrer de même qu'une application linéaire U de $C_0^\infty(\mathbb{R}^n)$ dans $C^\infty(\mathbb{R}^n)$ qui commute avec les translations est nécessairement de la forme $U\varphi = u \star \varphi$ avec $u \in \mathcal{D}'(\mathbb{R}^n)$ dès qu'elle possède la propriété de continuité suivante :

$$\forall K, \exists p, \exists C, \forall \varphi \in C_K^\infty, |U\varphi(0)| \leq C \sup_{x \in K, |\alpha| \leq p} |\partial^\alpha \varphi(x)|.$$

B. Interprétation physique

Considérons un système physique, que nous nous représenterons comme une “boîte noire” (on peut penser si on veut à un amplificateur), qui lorsqu'on l'excite avec un signal $s(t)$ produit une réponse $r(t)$. On fait fréquemment en physique les hypothèses suivantes.

- Le principe de superposition : si r_1 et r_2 sont les réponses aux signaux s_1 et s_2 , alors la réponse au signal $\lambda_1 s_1 + \lambda_2 s_2$ est $\lambda_1 r_1 + \lambda_2 r_2$.
- L'homogénéité dans le temps : la réponse au signal s décalé de T secondes est la réponse r décalée de T secondes.
- Une certaine stabilité : des signaux très voisins ne produisent pas des réponses très différentes.

Si on considère l'application (de $C_0^\infty(\mathbb{R})$ dans $C^\infty(\mathbb{R})$) qui à s fait correspondre r , la première propriété assure qu'elle est linéaire, et la seconde qu'elle commute avec les translations. Si on interprète la troisième comme la condition de continuité de l'exercice 7.3.3 (c'est une condition extrêmement faible : on accepte de perdre 1000 dérivées entre le contrôle que l'on a sur s et celui que l'on obtient sur r), il en résulte que la réponse est donnée à partir du signal par

$$r = k \star s$$

où k est une distribution sur $\mathbb{R}$. Ce résultat très général explique l'intervention de la convolution dans tant de domaines de la physique. Il n'est bien sûr correct que dans les conditions de validité des hypothèses ci-dessus : un système qui sature, ou dont les composants évoluent, ne rentre pas dans ce cadre.

En anticipant un peu sur la suite du cours, où nous verrons que δ est l'élément neutre de la convolution, on obtient que k est la réponse à l'excitation δ (on appelle souvent k la réponse impulsive), et que la réponse à la fonction de Heaviside $H(t)$ (facile à réaliser en fermant un interrupteur) est une primitive de k .

Si on fait l'hypothèse de causalité (si $s(t)$ est nul jusqu'à l'instant T , il en est de même de r), il en résulte que k est à support dans $[0, +\infty[$, et on obtient la forme familière $r(t) = \int_{-\infty}^t k(t - \tau)s(\tau) d\tau$ lorsque k est une fonction.

En admettant que les résultats précédents, que nous avons implicitement étendus à des signaux $s \in \mathcal{E}'$, se laissent aussi étendre à des signaux de type sinusoïdaux (ils ne sont pas à support compact, mais si k est une fonction assez petite à l'infini, le résultat est encore valable par passage à la limite), on obtient, pour $s(t) = e^{i\omega t}$ la réponse

$$r(t) = \int e^{i\omega(t-\tau)} k(\tau) d\tau = \widehat{k}(\omega) e^{i\omega t}$$

où on a noté $\widehat{k}$ la transformée de Fourier de k

$$\widehat{k}(\omega) = \int e^{-i\omega\tau} k(\tau) d\tau.$$

Il en résulte qu'un signal purement sinusoïdal de pulsation ω produit une réponse de même type, amplifiée dans le rapport $|\widehat{k}(\omega)|$ (le "gain" à la fréquence $\omega/2\pi$) et déphasée de $\arg(\widehat{k}(\omega))$. Nous verrons par la suite que la transformation de Fourier et la convolution sont très étroitement reliées.

Il ne faut pas confondre la linéarité de la courbe de réponse, tant vantée par les constructeurs de matériel haute fidélité, avec la linéarité de l'application $s \mapsto r$. La première signifie, si l'on en croit lesdits constructeurs, que $\widehat{k}$ (en fait son module, et dans un intervalle de fréquences) est presque une constante, ce qui équivaut à dire comme nous le verrons que k est proche d'un multiple de δ . Par contre, la linéarité de $s \mapsto r$, qui signifie en gros l'absence de saturation, est parfaitement compatible avec le fait que k ait une amplitude appréciable sur un grand intervalle de temps, ou que $\widehat{k}(\omega)$ varie beaucoup avec ω , toutes choses désastreuses pour l'audition musicale.

Des arguments du même type sont également valables en toute dimension. Ainsi par exemple, dès que l'on affirme en électrostatique que toute répartition de charges μ produit un potentiel V , que le principe de superposition est valable, et que si on translate les charges, le potentiel subit la même translation, l'argument ci-dessus entraîne que l'on a nécessairement $V = k \star \mu$, où k est le potentiel créé par la masse de Dirac. Nous avons bien sûr supposé une certaine continuité de l'application $\mu \mapsto V$, des hypothèses de ce type étant (presque) toujours faites et (presque) toujours implicites en physique.

Remarque 7.3.4. — On peut s'intéresser aux opérateurs linéaires qui ne vérifient pas la propriété de commutation avec les translations. Par exemple, si une fonction $k(x, y)$ est localement sommable dans $\mathbb{R}^p \times \mathbb{R}^q$, on peut lui associer l'opérateur intégral $\varphi \mapsto \int k(x, y) \varphi(y) dy$ qui applique $C_0^0(\mathbb{R}^q)$ dans $L_{loc}^1(\mathbb{R}^p)$. Ce dernier n'est un opérateur de convolution que si $p = q$ et si k ne dépend que de $x - y$.

On peut généraliser ce qui précède comme suit. Si $K \in \mathcal{D}'(\mathbb{R}^{p+q})$, on peut lui associer l'opérateur linéaire $\varphi \mapsto K\varphi$ de $C_0^\infty(\mathbb{R}^q)$ dans $\mathcal{D}'(\mathbb{R}^p)$ défini par

$$\langle K\varphi(x), \psi(x) \rangle = \langle K(x, y), \psi(x)\varphi(y) \rangle. \quad (7.14)$$

On vérifie facilement que $\mathcal{K}$ est linéaire, et que si une suite φ_j converge vers 0 ainsi que toutes ses dérivées, les supports des φ_j restant dans un compact fixe, alors la suite $\mathcal{K}\varphi_j$ tend vers 0 au sens des distributions.

Un résultat célèbre et difficile : le “théorème des noyaux” de L. Schwartz, assure que la réciproque est vraie. Toute application linéaire $\mathcal{K}$ vérifiant la propriété de continuité ci-dessus est de la forme (7.14), où la distribution K est déterminée de manière unique par $\mathcal{K}$.

7.4. Convolution des distributions

Nous allons définir ci-dessous le produit de convolution de deux distributions u et v à support compact. Il existe déjà deux cas où ce produit est défini : pour u et v dans L^1 d'une part, et pour $u \in \mathcal{E}'$, $v \in C_0^\infty$. Dans les deux cas, on sait que, pour $\varphi \in C_0^\infty$, on a

$$(u \star v) \star \varphi = u \star (v \star \varphi).$$

Dans le premier, c'est une conséquence simple du théorème de Fubini, dans le second, c'est le contenu du théorème 7.2.2. Nous allons prendre en fait le membre de droite comme définition du produit de convolution de deux distributions, ce qui nous garantit qu'elle prolonge les définitions antérieures.

Théorème et Définition 7.4.1 (Supports compacts)

Soient u et v appartenant à $\mathcal{E}'(\mathbb{R}^n)$. Il existe alors un et un seul $w \in \mathcal{E}'(\mathbb{R}^n)$ tel que

$$\forall \varphi \in C_0^\infty, \quad u \star (v \star \varphi) = w \star \varphi.$$

Cette distribution w est notée $u \star v$ et on a

$$\text{Supp}(u \star v) \subset \text{Supp}(u) + \text{Supp}(v).$$

Soient U et V les opérateurs de convolution par u et v respectivement, et posons $W = U \circ V$. L'opérateur W est linéaire de C_0^∞ dans lui-même et commute avec les translations, comme composé d'opérateurs possédant ces propriétés. En outre, il possède également la propriété de continuité (7.13). En effet, K et p étant donnés, on peut trouver L_1 et q_1 tels qu'on ait une majoration du type (7.13) pour U , puis, L_1 et q_1 étant déterminés, on peut trouver L et q tels qu'on ait une majoration du type (7.13) pour V . On obtient alors

$$\sup_{x \in K, |\alpha| \leq p} |\partial^\alpha U(V\varphi)| \leq C_1 \sup_{x \in L_1, |\alpha| \leq q_1} |\partial^\alpha V\varphi| \leq CC_1 \sup_{x \in L, |\alpha| \leq q} |\partial^\alpha \varphi|.$$

D'après le théorème 7.3.2, il existe donc une unique distribution w telle que $W\varphi = w \star \varphi$, ce qui permet de définir $u \star v$.

Il reste à démontrer la propriété sur les supports. Il faut donc prouver que l'on a $\langle (u \star v), \varphi \rangle = 0$ pour $\varphi \in C_0^\infty$ telle que

$$\text{Supp}(\varphi) \cap (\text{Supp}(u) + \text{Supp}(v)) = \emptyset. \quad (7.15)$$

On a

$$\langle u \star v, \varphi \rangle = ((u \star v) \star \tilde{\varphi})(0) = (u \star (v \star \tilde{\varphi}))(0)$$

par définition de $u \star v$. En utilisant deux fois le théorème 7.2.1, on obtient que le support de $u \star (v \star \tilde{\varphi})$ est contenu dans $\text{Supp}(u) + \text{Supp}(v) + \text{Supp}(\tilde{\varphi})$. Si 0 appartenait à cet ensemble, on aurait $x + y - z = 0$, avec x, y, z dans les supports respectifs de u, v, φ , ce qui contredirait (7.15). La fonction $u \star (v \star \tilde{\varphi})$ est donc nulle au voisinage de 0, ce qui achève la démonstration du théorème.

Remarque 7.4.2. — La définition précédente, et celle de l'extension qui va suivre, ne fournissent pas un procédé très effectif de calcul de $u \star v$. Il est souvent préférable d'utiliser la formule (7.17) ci-dessous, ou bien de passer à la limite (Théorème 7.4.9) à partir de produits de convolution facilement calculables (cas de deux éléments de L^1 , ou d'une distribution et d'une fonction C^∞).

Corollaire 7.4.3 (Supports convolutifs). — Soient u et v deux distributions sur $\mathbb{R}^n$ tels que $\text{Supp}(u)$ et $\text{Supp}(v)$ soient convolutifs. Le procédé du n° 7.1.5 permet de définir le produit de convolution $u \star v \in \mathcal{D}'(\mathbb{R}^n)$ et on a $\text{Supp}(u \star v) \subset \text{Supp}(u) + \text{Supp}(v)$.

On est en effet dans les conditions d'applications du n° 7.1.5, où les espaces $\mathcal{A}$ et $\mathcal{B}$ sont tous deux égaux à $\mathcal{D}'$.

Théorème 7.4.4 (Associativité sous condition). — Soient u, v, w des distributions dans $\mathbb{R}^n$ telles que $\text{Supp}(u), \text{Supp}(v)$ et $\text{Supp}(w)$ forment un triplet convolutif. On a alors

$$(u \star v) \star w = u \star (v \star w). \quad (7.16)$$

Dans le cas où les trois distributions sont à support compact, il suffit d'écrire la définition. En notant U, V, W les opérateurs de convolution par u, v, w respectivement, le membre de gauche [*resp.* de droite] de (7.16) est l'unique distribution qui, convoluée avec φ , donne $(U \circ V) \circ W\varphi$ [*resp.* $U \circ (V \circ W)\varphi$].

Lorsque les supports forment un triplet convolutif, on vérifie facilement que la restriction des deux membres de (7.16) à $B(R)$ ne change pas si on remplace u, v, w par leurs produits par une fonction θ à support compact et égale à 1 au voisinage de $B(\rho(R))$ (pour la fonction ρ du triplet). On est alors ramené au cas précédent.

7.4.5. Remarque importante. — Il est possible que les deux membres de (7.16) soient parfaitement définis, et qu'ils soient cependant différents. Il peut arriver en effet que les supports de u et v , de v et w , de $u \star v$ et w , de u et $v \star w$ forment quatre couples convolutifs, sans que le triplet $(\text{Supp}(u), \text{Supp}(v), \text{Supp}(w))$ le soit. Nous verrons ci-dessous un exemple de ce phénomène.

Il ne s'agit pas du tout de ces contre-exemples que l'on ne rencontre que dans les listes d'exercices. Dans la résolution des équations de convolution, l'usage injustifié de l'associativité conduit très facilement à des résultats faux.

Théorème 7.4.6 (Commutativité). — Soient u et v appartenant à $\mathcal{D}'(\mathbb{R}^n)$ dont les supports sont convolutifs. On a alors $u \star v = v \star u$.

Il suffit de démontrer le théorème lorsque u et v sont à support compact, le cas général s'y ramenant comme précédemment en remplaçant u et v par θu et θv .

D'après la définition 7.4.1, il suffit de prouver que $u \star (v \star \varphi) = v \star (u \star \varphi)$, et donc de montrer que

$$\langle u \star v \star \varphi, \psi \rangle = \langle v \star u \star \varphi, \psi \rangle,$$

quels que soient φ, ψ dans C_0^∞ . Le premier membre valant $(u \star v \star \varphi \star \tilde{\psi})(0)$, il suffit donc de démontrer que l'on a

$$u \star v \star \varphi \star \tilde{\psi} = v \star u \star \varphi \star \tilde{\psi}.$$

Nous allons utiliser l'associativité (tout est à support compact), et le fait que la commutativité est connue pour le produit de convolution de deux fonctions. On a

$$\begin{aligned} u \star (v \star \varphi) \star \tilde{\psi} &= u \star \tilde{\psi} \star (v \star \varphi) = (u \star \tilde{\psi}) \star (v \star \varphi) = (v \star \varphi) \star (u \star \tilde{\psi}) = \\ &= v \star \varphi \star (u \star \tilde{\psi}) = v \star (u \star \tilde{\psi}) \star \varphi = v \star u \star \tilde{\psi} \star \varphi = v \star u \star \varphi \star \tilde{\psi}, \end{aligned}$$

ce qui achève la démonstration.

Théorème 7.4.7. — (a) Soient u et v deux distributions dont l'une est à support compact. On a alors

$$\forall \varphi \in C_0^\infty, \quad \langle u \star v, \varphi \rangle = \langle u, \tilde{v} \star \varphi \rangle \quad (7.17)$$

(b) Pour toute distribution u , on a

$$\begin{aligned} \delta \star u &= u \\ \delta_a \star u &= \tau_a u \\ (\partial^\alpha \delta) \star u &= \partial^\alpha u. \end{aligned}$$

Pour dériver (ou translater) un produit de convolution, il suffit de dériver (ou translater) l'un des facteurs.

En utilisant (7.12) et l'associativité (deux des trois facteurs sont à support compact), on obtient

$$\langle u \star v, \varphi \rangle = ((u \star v) \star \check{\varphi})(0) = (u \star (v \star \check{\varphi}))(0).$$

Il suffit de remarquer que $(v \star \check{\varphi})^\sim = \check{v} \star \varphi$ et d'utiliser encore (7.12) pour obtenir la partie (a) du théorème.

Pour $\varphi \in C_0^\infty$, la fonction $\delta_a \star \varphi$ est égale à $\tau_a \varphi$. On a en effet $(\delta_a \star \varphi)(x) = \langle \delta_a(y), \varphi(x-y) \rangle = \varphi(x-a)$. On a donc d'après (7.17)

$$\langle u \star \delta_a, \varphi \rangle = \langle u, \delta_{-a} \star \varphi \rangle = \langle u, \tau_{-a} \varphi \rangle = \langle \tau_a u, \varphi \rangle.$$

Cela montre que $\delta_a \star u = \tau_a u$, et en particulier, pour $a = 0$, que δ est l'élément neutre de la convolution.

On a de même $(\partial^\alpha \delta) \star \varphi = \partial^\alpha \varphi$. En effet, on a

$$((\partial^\alpha \delta) \star \varphi)(x) = \langle \partial^\alpha \delta(y), \varphi(x-y) \rangle = (-1)^{|\alpha|} \langle \delta(y), \partial_y^\alpha \varphi(x-y) \rangle = \partial^\alpha \varphi(x).$$

On déduit de (7.17), la distribution $\partial^\alpha \delta$ étant paire ou impaire selon la parité de $|\alpha|$

$$\langle u \star (\partial^\alpha \delta), \varphi \rangle = (-1)^{|\alpha|} \langle u, (\partial^\alpha \delta) \star \varphi \rangle = (-1)^{|\alpha|} \langle u, \partial^\alpha \varphi \rangle = \langle \partial^\alpha u, \varphi \rangle,$$

ce qui démontre que $u \star (\partial^\alpha \delta)$ est égal à $\partial^\alpha u$.

Si les supports de u et v sont convolutifs, on obtient un triplet convolutif en leur adjoignant un compact. On peut donc écrire, en utilisant associativité et commutativité :

$$\delta_a \star (u \star v) = (\delta_a \star u) \star v = u \star (\delta_a \star v) \quad (7.18)$$

d'où

$$\tau_a(u \star v) = (\tau_a u) \star v = u \star (\tau_a v). \quad (7.19)$$

De même, en groupant différemment les facteurs de $(\partial^\alpha \delta) \star u \star v$, on obtient

$$\partial^\alpha (u \star v) = (\partial^\alpha u) \star v = u \star (\partial^\alpha v) = (\partial^\beta u) \star (\partial^\gamma v),$$

pour $\alpha = \beta + \gamma$.

7.4.8. Continuation de la remarque 7.4.5. — On a sur $\mathbb{R}$

$$(1 \star \delta') \star H = 0 \star H = 0,$$

$$1 \star (\delta' \star H) = 1 \star \delta = 1.$$

Chaque produit de convolution écrit a un sens (l'un des facteurs étant à support compact). On a utilisé le fait que, pour toute distribution u , on a $\delta' \star u = u'$.

Théorème 7.4.9 (Continuité). — Soient F et G deux fermés convolutifs. Soient u_j et v_j deux suites de distributions, respectivement à supports dans F et G , et convergentes au sens des distributions. On a alors

$$\lim_{j \rightarrow \infty} (u_j \star v_j) = \left(\lim_{j \rightarrow \infty} u_j \right) \star \left(\lim_{j \rightarrow \infty} v_j \right)$$

au sens des distributions.

Nous nous bornerons à démontrer la propriété plus faible de continuité séparée, c'est-à-dire que, pour v à support dans G et u_j comme ci-dessus convergeant vers u_0 , la suite $u_j \star v$ converge vers $u_0 \star v$.

Pour étudier la convergence de $\langle u_j \star v, \varphi \rangle$, il suffit de connaître les restrictions des distributions en question à une boule contenant le support de φ , et on se ramène, grâce à l'hypothèse de convolutivité, au cas où les u_j et v sont à support compact. On a alors d'après (7.17),

$$\langle u_j \star v, \varphi \rangle = \langle u_j, \tilde{v} \star \varphi \rangle.$$

La fonction $\tilde{v} \star \varphi$ appartenant à C_0^∞ , le dernier membre converge vers $\langle u_0, \tilde{v} \star \varphi \rangle$ qui est égal, par le même calcul, à $\langle u_0 \star v, \varphi \rangle$.

Cela termine la preuve de la continuité séparée, qui est très souvent suffisante. On trouvera une démonstration de la continuité énoncée dans le théorème au n°C.4.3.

7.5. Mode d'emploi

Ce résumé ne dispense pas de lire les sections précédentes, il en présente les principaux résultats dans un ordre différent.

L'ordre logique de l'exposition a en effet conduit à introduire successivement des définitions de plus en plus générales du produit de convolution. Au niveau de l'utilisation, c'est bien entendu le concept le plus général qui doit être mis au premier plan, les autres n'apparaissant plus que comme des cas particuliers. D'autre part, il faut bien distinguer, dans ce chapitre comme dans les autres, les définitions théoriques et les propriétés — que nous privilégions ci-dessous — directement utilisables pour les démonstrations et les calculs.

A. Conditions de Définition

Pour tout couple de distributions $u, v \in \mathcal{D}'(\mathbb{R}^n)$ dont les supports sont convolutifs [cas particulier important : l'un des deux supports est compact], le produit de convolution $u \star v$ est défini et appartient à $\mathcal{D}'(\mathbb{R}^n)$.

Il n'est pas utile ici de réécrire la définition. Par contre, les cinq propriétés suivantes (auxquelles il faut adjoindre bien sûr la bilinéarité) sont d'une importance capitale.

B. Propriétés fondamentales

7.5.1. Support. — $\text{Supp}(u \star v) \subset \text{Supp}(u) + \text{Supp}(v)$.

7.5.2. Commutativité. — $u \star v = v \star u$.

7.5.3. Associativité sous condition. — Si les trois fermés $\text{Supp}(u)$, $\text{Supp}(v)$ et $\text{Supp}(w)$ forment un triplet convolutif [*cas particulier important : deux des trois supports sont compacts*], on a :

$$(u \star v) \star w = u \star (v \star w).$$

7.5.4. Dérivation et translation. — $\delta_a \star u = \tau_a u$, $(\partial^\alpha \delta) \star u = \partial^\alpha u$.

7.5.5. Continuité. — Soient F et G deux fermés convolutifs. Soient u_j et v_j deux suites de distributions, respectivement à supports dans F et G , et convergentes au sens des distributions. On a alors

$$\lim_{j \rightarrow \infty} (u_j \star v_j) = (\lim_{j \rightarrow \infty} u_j) \star (\lim_{j \rightarrow \infty} v_j)$$

au sens des distributions.

C. Modes de calcul

7.5.6. Deux fonctions localement sommables. — Si u et v appartiennent à $L^1_{\text{loc}}(\mathbb{R}^n)$, l'un des deux supports étant compact⁽¹⁾, on a $u \star v \in L^1_{\text{loc}}(\mathbb{R}^n)$ avec

$$(u \star v)(x) = \int u(x-y)v(y) dy.$$

7.5.7. Une distribution et une fonction C^∞ . — Si $u \in \mathcal{D}'(\mathbb{R}^n)$ et $v \in C^\infty(\mathbb{R}^n)$, l'un des deux supports étant compact⁽²⁾, on a $u \star v \in C^\infty(\mathbb{R}^n)$ avec

$$(u \star v)(x) = \langle u(y), v(x-y) \rangle.$$

Le crochet ci-dessus désigne, selon les cas, la dualité usuelle entre $\mathcal{D}'$ et C^∞_0 ou bien la dualité entre $\mathcal{E}'$ et C^∞ .

⁽¹⁾Le résultat est encore valable en supposant seulement les supports convolutifs, voir l'exercice 7.1.7.

⁽²⁾Le résultat est encore valable en supposant seulement les supports convolutifs, voir la seconde partie de la remarque 7.2.4.

7.5.8. Utilisation des propriétés fondamentales et des cas précédents. — Un cas typique est l'utilisation de la continuité : pour calculer $u \star v$, on peut par exemple approcher u par des fonctions $u_j \in C^\infty$ (c'est toujours possible) qui soient suffisamment explicites (cela arrive), on est alors ramené au calcul de $u_j \star v$ (c'est le cas précédent) et, en supposant les conditions de support satisfaites, à celui de la limite des $u_j \star v$.

De même, en dimension 1 pour fixer les idées, si u est la dérivée de U , il peut être plus simple de calculer $U \star v$ et d'utiliser le fait que $u \star v = (U \star v)'$. Si au contraire c'est u' dont l'expression est particulièrement simple, on pourra calculer d'abord $u' \star v$ et $u \star v$ sera déterminé à une constante additive près par le fait que c'en est une primitive. Le fait que $\text{Supp}(u \star v) \subset \text{Supp}(u) + \text{Supp}(v)$ permettra souvent de déterminer la constante.

7.5.9. En désespoir de cause. — Si l'un des deux supports est compact⁽³⁾, on a

$$\forall \varphi \in C_0^\infty, \quad \langle u \star v, \varphi \rangle = \langle u, \tilde{v} \star \varphi \rangle.$$

Ce mode d'emploi est loin de résumer tout le chapitre. Les théorèmes de dérivation et d'intégration sous le crochet ont leur intérêt propre. L'approximation des distributions par régularisation est très importante. Enfin, nous avons souligné l'intérêt conceptuel, sur les plans mathématique et extra-mathématique, que présente la propriété caractéristique de la convolution.

⁽³⁾ A titre d'exercice, on pourra montrer que la formule est valable dans le cas de supports convolutifs, le crochet du membre de droite étant bien défini grâce à l'extension de la dualité du n° 6.1.8.

CHAPITRE 8

QUELQUES ÉQUATIONS DE LA PHYSIQUE MATHÉMATIQUE

8.1. Généralités sur les équations de convolution

Une équation de convolution est une équation de la forme

$$A \star u = f, \tag{8.1}$$

où A et f sont des distributions données, et où u est une distribution inconnue. Un grand nombre d'équations rentrent dans ce cadre.

— Les équations aux dérivées partielles linéaires à coefficients constants. En effet, l'équation

$$\sum_{|\alpha| \leq m} a_\alpha \partial^\alpha u = f$$

peut être mise sous la forme (8.1) avec $A = \sum a_\alpha \partial^\alpha \delta$.

— Des équations aux différences finies. Par exemple, en dimension 1, l'équation $u(x+h) - u(x) = f(x)$ se met sous la forme (8.1) avec $A = (\delta_{-h} - \delta)$.

— Des équations intégrales du type $\int k(x-y)u(y) dy = f(x)$, mais aussi des équations intégrales en valeurs principales ou en parties finies.

— Des combinaisons linéaires des cas précédents : équations aux dérivées partielles avec retard, équations intégro-différentielles, ...

Le concept suivant joue un rôle fondamental dans la résolution de ces équations.

Définition 8.1.1. — Soit $A \in \mathcal{E}'(\mathbb{R}^n)$. On dit qu'une distribution $E \in \mathcal{D}'(\mathbb{R}^n)$ est une solution élémentaire⁽¹⁾ de A (on dit parfois solution élémentaire de l'équation (8.1)) si on a $A \star E = \delta$.

Il n'existe pas toujours de solution élémentaire (démontrer à titre d'exercice qu'un élément $A \in C_0^\infty$ n'en possède jamais). S'il existe une solution élémentaire E , on obtient toutes les autres en ajoutant à E une solution quelconque v de l'équation homogène $A \star v = 0$.

Un théorème célèbre de Malgrange et Ehrenpreis (1955) assure que toute équation aux dérivées partielles à coefficients constants (non tous nuls) possède une solution élémentaire (et en fait une infinité). On peut prouver qu'il n'en existe jamais qui soient à support compact dès que le degré de l'équation est au moins égal à 1.

Une solution élémentaire résout l'équation (8.1) lorsque le second membre est l'élément neutre δ . Elle jouit d'une partie des propriétés que posséderait l'inverse de A dans une algèbre associative, mais d'une partie seulement, l'associativité n'étant valable que sous certaines conditions. Par exemple, si E_1 et E_2 sont deux solutions élémentaires de A , on a $(E_1 \star A) \star E_2 = E_2$ et $E_1 \star (A \star E_2) = E_1$, et il ne faut surtout pas vouloir en déduire l'égalité de E_1 et E_2 .

Théorème 8.1.2. — Soit A une distribution à support compact possédant une solution élémentaire E .

- (a) Pour tout $f \in \mathcal{E}'(\mathbb{R}^n)$, il existe au moins une solution de (8.1) appartenant à $\mathcal{D}'(\mathbb{R}^n)$ et $u = E \star f$ en est une.
- (b) Pour tout $f \in \mathcal{E}'(\mathbb{R}^n)$, il existe au plus une solution de (8.1) appartenant à $\mathcal{E}'(\mathbb{R}^n)$ et, s'il en existe une, c'est $E \star f$.

En effet, les supports de A et f étant compacts, on peut écrire

$$A \star (E \star f) = (A \star E) \star f = \delta \star f = f,$$

ce qui démontre le premier point. Si on suppose que u est une solution à support compact, le triplet $(\text{Supp}(u), \text{Supp}(A), \text{Supp}(E))$ est convolutif, et on a

$$u = (E \star A) \star u = E \star (A \star u) = E \star f,$$

ce qui achève la démonstration.

⁽¹⁾On dit aussi *solution fondamentale*. Les physiciens disent le plus souvent *fonction de Green* alors que les mathématiciens n'utilisent en général cette expression que pour l'équation de Laplace.

8.2. Équations de Laplace et de Poisson

Nous noterons $\mathbf{r} = (x, y, z)$ le point courant de $\mathbb{R}^3$, et nous poserons $r = |\mathbf{r}|$. Nous nous intéressons à l'équation de Poisson

$$\Delta u = f, \quad (8.2)$$

où Δ est le Laplacien et où f est une distribution donnée à support compact dans $\mathbb{R}^3$. Lorsque le second membre f est nul, on dit que u est une *fonction harmonique* (nous allons voir que c'est effectivement toujours une fonction).

Pour cette équation, comme pour les autres équations aux dérivées partielles, on ne cherche pas en général à obtenir toutes les solutions — il y en a beaucoup trop — mais on cherche à obtenir une solution (si possible unique) satisfaisant à des conditions additionnelles (conditions aux limites, ou à l'infini, ou à l'instant initial pour les problèmes d'évolution, ou bien invariance par les groupes de symétrie de l'équation, ...). Nous demanderons ici, guidés par l'interprétation physique (électrostatique par exemple) que u tende vers 0 à l'infini.

Nous connaissons déjà (voir n°5.4.17) une solution élémentaire du Laplacien : en posant

$$E = -\frac{1}{4\pi r},$$

on a $\Delta E = \delta$. Il existe bien d'autres solutions élémentaires (on peut ajouter à E n'importe quelle fonction harmonique), mais nous allons voir que c'est $E \star f$ qui possède les meilleures propriétés à l'infini.

Théorème 8.2.1. — *Soit $f \in \mathcal{E}'(\mathbb{R}^3)$. Alors la distribution $u = E \star f$ est une solution de (8.2). En dehors du support de f , c'est une fonction de classe C^∞ qui tend vers 0 à l'infini.*

Le fait que u soit une solution est un cas particulier du théorème 8.1.2. Soit θ_ε une fonction de classe C^∞ , égale à 1 hors de la boule B_ε de rayon ε centrée à l'origine, et égale à 0 au voisinage de l'origine. La fonction $E_\varepsilon = \theta_\varepsilon E$ est alors de classe C^∞ dans $\mathbb{R}^3$ et on a $\text{Supp}(E - E_\varepsilon) \subset B_\varepsilon$.

On a

$$u = E_\varepsilon \star f + (E - E_\varepsilon) \star f.$$

Le premier terme est une fonction C^∞ et le second est une distribution dont le support est contenu dans $\text{Supp}(f) + B_\varepsilon$. Si $\mathbf{r}$ est un point du complémentaire du support de f , en choisissant ε égal à la moitié de la distance de $\mathbf{r}$ à $\text{Supp}(f)$, on en déduit que u est de classe C^∞ au voisinage de $\mathbf{r}$.

Soit m l'ordre de f , choisissons K voisinage compact de $\text{Supp}(f)$ et $\varepsilon > 0$. On a $(E_\varepsilon \star f)(\mathbf{r}) = \langle f(\mathbf{r}') \cdot E_\varepsilon(\mathbf{r} - \mathbf{r}') \rangle$ et on peut majorer $(E_\varepsilon \star f)(\mathbf{r})$, d'après le théorème 6.1.3 et la remarque 6.1.7, par la borne supérieure sur K des dérivées d'ordre $\leq m$ de la fonction $\mathbf{r}' \mapsto E(\mathbf{r} - \mathbf{r}')$. On a donc

$$|(E_\varepsilon \star f)(\mathbf{r})| \leq C \sup_{\mathbf{r}' \in K, |\alpha| \leq m} |\partial_{\mathbf{r}'}^\alpha E_\varepsilon(\mathbf{r} - \mathbf{r}')|, \quad (8.3)$$

et donc, dès que la distance $d(\mathbf{r}, K)$ est supérieure à ε ,

$$|u(\mathbf{r})| \leq C \sup_{\mathbf{r}' \in K, |\alpha| \leq m} |\partial_{\mathbf{r}'}^\alpha E(\mathbf{r} - \mathbf{r}')|,$$

C'est un exercice facile de montrer que la fonction $\mathbf{r} \mapsto (1/r)$ et toutes ses dérivées sont majorées en $O(1/r)$ pour $\mathbf{r} \rightarrow \infty$. Il en résulte que le membre de droite de (8.3) est majoré par $C'/d(\mathbf{r}, K)$, et tend donc vers 0 lorsque $\mathbf{r}$ tend vers l'infini.

Remarque 8.2.2. — Pour démontrer que $E \star f$ est de classe C^∞ en dehors du support de f , nous n'avons pas utilisé la forme particulière de E , mais seulement le fait que E est de classe C^∞ dans le complémentaire de l'origine.

Le corollaire suivant assure que les seules distributions harmoniques (c'est-à-dire solutions de $\Delta u = 0$) sont les fonctions harmoniques usuelles.

Corollaire 8.2.3. — Soit Ω un ouvert de $\mathbb{R}^3$, et u une distribution dans Ω vérifiant $\Delta u = 0$. Alors u est une fonction de classe C^∞ .

Il suffit de prouver que u est de classe C^∞ au voisinage de chaque point $\mathbf{r}_0 \in \Omega$; et donc de montrer que, pour toute fonction $\varphi \in C_0^\infty(\Omega)$, égale à 1 au voisinage de $\mathbf{r}_0$, on a $\varphi u \in C^\infty$ au voisinage de $\mathbf{r}_0$.

On a $\varphi u = E \star (\Delta \delta) \star (\varphi u) = E \star (\Delta(\varphi u))$, deux des trois supports étant compacts. On a d'autre part $\Delta(\varphi u) = u \Delta \varphi + 2 \nabla u \cdot \nabla \varphi$ et, les dérivées de φ étant nulles près de $\mathbf{r}_0$, le point $\mathbf{r}_0$ n'appartient pas au support de $\Delta(\varphi u)$. Le théorème précédent assure alors que φu est de classe C^∞ près de $\mathbf{r}_0$, ce qui achève la démonstration.

Exercice 8.2.4. — Démontrer que, plus généralement, si P est un opérateur différentiel à coefficients constants dans $\mathbb{R}^n$ qui possède une solution élémentaire de classe C^∞ en dehors de l'origine, on a la propriété suivante : toute distribution u dans un ouvert $\Omega \subset \mathbb{R}^n$ qui vérifie $Pu = 0$ (ou même $Pu \in C^\infty(\Omega)$) est de classe C^∞ .

Nous savons déjà que le Laplacien en toute dimension, et l'opérateur $\partial/\partial x + i\partial/\partial y$ dans le plan, possèdent cette propriété (voir exercice 5.4.20). Nous verrons qu'il en est de même pour l'équation de la chaleur.

Théorème 8.2.5 (Propriété de moyenne). — Soit Ω un ouvert de $\mathbb{R}^3$ contenant la boule fermée de centre $\mathbf{r}_0$ et de rayon R , et soit u une fonction harmonique dans Ω . On a alors

$$u(\mathbf{r}_0) = \frac{1}{4\pi R^2} \int_{S_R} u \, d\sigma_R,$$

en notant S_R la sphère de centre $\mathbf{r}_0$ et de rayon R , et $d\sigma_R$ la mesure de surface.

On peut supposer que le point $\mathbf{r}_0$ est l'origine. Considérons alors la fonction g_R , égale à $1/r - 1/R$ pour $r \leq R$, et à 0 sinon. Le lecteur sait certainement déjà (c'est le contenu de l'exercice 5.4.19) que l'on a

$$\Delta g_R = R^{-2} d\sigma_R - 4\pi \delta.$$

Soit $\tilde{u}$ une fonction appartenant à $C_0^\infty(\mathbb{R}^n)$ coïncidant avec u au voisinage de la boule de rayon R . On a alors

$$\langle \Delta g_R, \tilde{u} \rangle = \langle g_R, \Delta \tilde{u} \rangle.$$

La fonction $\Delta \tilde{u}$ étant nulle au voisinage du support de g_R , le membre de droite est nul, et il suffit d'expliciter le membre de gauche pour avoir le résultat.

Corollaire 8.2.6. — Si u est une fonction harmonique dans $\mathbb{R}^3$ qui tend vers 0 à l'infini, on a $u = 0$.

Il suffit en effet, pour chaque $\mathbf{r}_0 \in \mathbb{R}^3$, d'écrire que $u(\mathbf{r}_0)$ est la moyenne de u sur la sphère de centre $\mathbf{r}_0$ et de rayon R , et de faire tendre R vers l'infini.

Exercice 8.2.7 (Forme forte du principe du maximum)

Démontrer qu'une fonction harmonique dans une boule qui atteint son maximum au centre est constante. En déduire qu'une fonction harmonique dans un ouvert connexe qui atteint son maximum en un point de l'ouvert est nécessairement constante.

Nous pouvons maintenant énoncer le résultat essentiel de cette section. Le lecteur pourra montrer que le même énoncé (en modifiant la forme explicite de la solution élémentaire) est valable en toute dimension $n \geq 3$.

Théorème 8.2.8. — Pour toute distribution f à support compact dans $\mathbb{R}^3$, il existe une infinité de solutions $u \in \mathcal{D}'(\mathbb{R}^3)$ de l'équation $\Delta u = f$. Toutes ces solutions sont de classe C^∞ en dehors du support de f . Parmi celles-ci, il en existe une et une seule qui tende vers 0 à l'infini, et elle est donnée par

$$u = \left(-\frac{1}{4\pi r} \right) \star f.$$

Nous savons déjà que $E \star f$ est une solution C^∞ en dehors de $\text{Supp}(f)$ et tendant vers 0 à l'infini. Dire que u est une autre solution de $\Delta u = f$ équivaut à dire que $u - E \star f$ est une distribution harmonique et donc, d'après le corollaire 8.2.3, une fonction harmonique C^∞ . Si elle tend vers 0 à l'infini, elle est identiquement nulle d'après le corollaire précédent, et cela achève la démonstration.

8.3. Équation des ondes

On notera $(t, \mathbf{r}) = (t, x, y, z)$ les coordonnées d'un point dans l'espace-temps, et $\square$ le d'Alembertien

$$\square = \partial_t^2 - \Delta.$$

On notera Γ la surface du (demi)-cône d'onde d'avenir définie par $t = r$. Bien que Γ ne soit pas une surface régulière à l'origine, on peut facilement définir sa mesure de surface par

$$\int_{\Gamma} f \, d\sigma = \sqrt{2} \int_{\mathbb{R}^3} f(r, \mathbf{r}) \, d\mathbf{r},$$

pour $f(t, \mathbf{r})$ définie sur Γ , continue et à support compact.

On notera enfin $\rho = \sqrt{t^2 + r^2}$. Nous admettrons provisoirement (voir le n°9.6.5) le théorème suivant. La transformation de Fourier nous donnera ultérieurement des méthodes générales pour chercher des solutions élémentaires.

Théorème 8.3.1. — *La distribution de simple couche $d\sigma/(4\pi\rho)$ est une solution élémentaire de l'équation des ondes :*

$$\square \left(\frac{d\sigma}{\rho} \right) = 4\pi\delta. \quad (8.4)$$

Il s'agit d'une mesure de Radon positive bien définie : en remarquant que l'on a $\rho = r\sqrt{2}$ sur Γ , on a

$$\left\langle \frac{d\sigma}{\rho}, \varphi \right\rangle = \int_{\mathbb{R}^3} \frac{\varphi(r, \mathbf{r})}{r} \, d\mathbf{r}. \quad (8.5)$$

La fonction $1/r$ étant localement sommable dans $\mathbb{R}^3$, l'intégrale figurant au membre de droite est finie, et la forme linéaire ainsi définie est positive.

Si on veut seulement vérifier que la formule (8.4) est valable, il s'agit de prouver que, pour tout $\varphi \in C_0^\infty(\mathbb{R}^4)$, on a $\int_{\Gamma} \square \varphi \, d\sigma / \rho = 4\pi\varphi(0)$, et le lecteur soigneux pourra s'y essayer.

On aura intérêt à utiliser les coordonnées polaires $\mathbf{r} = r\boldsymbol{\theta}$, $\boldsymbol{\theta} \in \mathbb{S}^2$ de $\mathbb{R}^3$, et à écrire le Laplacien en coordonnées polaires : $\partial_r^2 + (2/r)\partial_r + r^{-2}\Delta_{\boldsymbol{\theta}}$. On note $d\boldsymbol{\theta}$ la mesure de surface et $\Delta_{\boldsymbol{\theta}}$ le Laplacien de la sphère $\mathbb{S}^2$, la seule propriété utile ici étant la conséquence suivante de la formule de Green sur la sphère : $\int_{\mathbb{S}^2} \Delta_{\boldsymbol{\theta}} \psi \, d\boldsymbol{\theta} = 0$ pour ψ de classe C^2 définie sur la sphère.

On pourra ensuite remplacer les coordonnées (t, r) par $(t + r, t - r)$ au voisinage de Γ , et faire les intégrations par parties qui s'imposent pour $(t + r) \in [\varepsilon, \infty]$.

Nous dirons qu'une distribution u dans $\mathbb{R}^4$ est *nulle dans le passé*, s'il existe $T_0 \in \mathbb{R}$ tel que le support de u soit contenu dans $[T_0, \infty[\times \mathbb{R}^3$.

Théorème 8.3.2. — *Soit $f \in \mathcal{D}'(\mathbb{R}^4)$ nulle dans le passé. Il existe alors une et une seule solution $u \in \mathcal{D}'(\mathbb{R}^4)$ de $\square u = f$ qui soit nulle dans le passé, et on a*

$$u = \left(\frac{d\sigma}{4\pi\rho} \right) \star f.$$

Le support de u est contenu dans l'ensemble des $(t, \mathbf{r})$ tels qu'il existe $(t_0, \mathbf{r}_0) \in \text{Supp}(f)$ avec $t \geq t_0$ et $t - t_0 = |\mathbf{r} - \mathbf{r}_0|$.

L'observation importante est ici que Γ et $[T, \infty[\times \mathbb{R}^3$ sont des ensembles convolutifs. Si en effet $(t_1, \mathbf{r}_1)$ et $(t_2, \mathbf{r}_2)$ appartiennent respectivement à ces deux ensembles, avec $|t_1 + t_2| \leq R$ et $|\mathbf{r}_1 + \mathbf{r}_2| \leq R$, on en déduit d'abord que $|t_1|$ et $|t_2|$ sont majorés par $|T| + R$, puis qu'il en est de même de $|\mathbf{r}_1| = |t_1|$, et qu'on a donc $|\mathbf{r}_2| \leq |T| + 2R$.

Il en résulte d'abord que, en posant $E = d\sigma/(4\pi\rho)$, la distribution $u = E \star f$ est bien définie, et que

$$\square u = (\square \delta) \star E \star f = f.$$

Si maintenant v est une solution de la même équation nulle dans le passé, les supports de E et v sont convolutifs, et on a

$$v = E \star (\square \delta) \star v = E \star f.$$

Enfin, la description du support ne fait que paraphraser la relation $\text{Supp}(E \star f) \subset \text{Supp}(f) + \Gamma$.

Physiquement, f représente les causes (sources sonores, lumineuses, ...) des ondes décrites par u . Il est assez remarquable que, à partir d'une condition de causalité très faible (des causes nulles dans le passé produisent des effets nuls dans le passé), on en déduise la relation de causalité beaucoup plus forte fournie par la majoration du support de u : les causes ne produisent d'effets que par propagation à la vitesse 1. Ces effets sont postérieurs aux causes, non seulement dans le repère initial, mais aussi dans tout repère qui s'en déduit par une transformation de Lorentz respectant le sens du temps.

En particulier, $d\sigma/4\pi\rho$ est la seule solution élémentaire qui soit nulle dans le passé.

8.3.3. Problème de Cauchy. — Pour les équations comportant une variable de temps, on se pose souvent la question de déterminer l'évolution (libre ou avec second membre) connaissant la situation à l'instant initial. Ce problème prend ici la forme suivante, où h_0 et h_1 sont des fonctions données dans $\mathbb{R}^3$, et où u est inconnue.

$$\begin{aligned}\square u(t, \mathbf{r}) &= 0 && \text{pour } (t, \mathbf{r}) \in [0, \infty[\times \mathbb{R}^3 \\ u(0, \mathbf{r}) &= h_0(\mathbf{r}) && \text{pour } \mathbf{r} \in \mathbb{R}^3 \\ \partial_t u(0, \mathbf{r}) &= h_1(\mathbf{r}) && \text{pour } \mathbf{r} \in \mathbb{R}^3.\end{aligned}$$

L'équation des ondes étant invariante par renversement du temps, il se trouve que l'on peut en fait résoudre le problème dans $\mathbb{R}^4$ entier.

Nous supposons h_0 et h_1 de classe C^∞ . Dans un premier temps, en admettant l'existence d'une solution de classe C^2 , nous allons montrer qu'elle est nécessairement donnée par une formule bien précise. Puis nous montrerons que cette formule donne effectivement une solution de classe C^∞ du problème de Cauchy. Le lecteur pourra montrer que les conditions de régularité imposées aux h_i peuvent être affaiblies. Nous verrons plus loin que c'est dans le cadre des espaces de Sobolev que l'on peut obtenir les résultats les plus précis sur ce point.

8.3.4. Conditions nécessaires. — Supposons donc qu'il existe une solution u de classe C^2 , et considérons la fonction $H(t)u(t, \mathbf{r})$. Il s'agit d'une fonction de classe C^2 par morceaux, dont nous pouvons calculer les dérivées d'ordre 1 et 2 par la formule des sauts dans l'espace. Pour les dérivées spatiales, les termes de saut sont nuls, l'angle de la normale à l'hyperplan $t = 0$ avec les directions de dérivation étant égal à $\pi/2$. On a donc par exemple $\partial_x^2(H(t)u) = H(t)\partial_x^2 u$.

Par contre, on a $\partial_t(H(t)u) = H(t)\partial_t u + \mu_0$, en notant μ_0 la distribution de simple couche portée par l'hyperplan $t = 0$ de densité $h_0(\mathbf{r})$. De même, on obtient

$$\partial_t^2(H(t)u) = H(t)\partial_t^2 u + \mu_1 + \partial_t \mu_0,$$

en notant μ_1 la simple couche de densité h_1 .

En sommant ces dérivées, et en utilisant le fait que $\square u = 0$, on obtient

$$\square(H(t)u) = \mu_1 + \partial_t \mu_0. \quad (8.6)$$

La fonction $H(t)u$ est nulle dans le passé, et il en est de même de la distribution figurant au second membre de (8.6), dont le support est contenu dans l'hyperplan $t = 0$. Le théorème 8.3.2 s'applique alors, et on obtient

$$H(t)u = E \star (\mu_1 + \partial_t \mu_0) = E \star \mu_1 + \partial_t(E \star \mu_0), \quad (8.7)$$

en notant toujours $E = d\sigma/(4\pi\rho)$. Cela prouve déjà que la solution u , si elle existe, est nécessairement donnée, pour $t > 0$ (la même preuve s'appliquerait à $t < 0$), par la formule ci-dessus.

8.3.5. Explicitation de $(d\sigma/(4\pi\rho)) \star \mu$. — On notera μ une distribution de simple couche de densité $h(\mathbf{r}) \in C^\infty(\mathbb{R}^3)$ portée par $t = 0$. Nous ne disposons pas de formule explicite pour calculer le produit de convolution de deux simples couches, dont l'une de surcroît est portée par une hypersurface non régulière. La méthode consiste à approcher l'une d'entre elles par une fonction C^∞ , à calculer explicitement le produit de convolution, et à utiliser la continuité de celui-ci.

Soit $\chi_\varepsilon(t)$ une approximation de l'identité sur $\mathbb{R}$, et considérons les fonctions $\mu_\varepsilon(t, \mathbf{r}) = \chi_\varepsilon(t)h(\mathbf{r})$. Elles sont de classe C^∞ , et c'est un exercice facile de montrer que $\mu_\varepsilon \rightarrow \mu$ au sens des distributions. D'autre part, pour $\varepsilon \leq 1$, elles ont toutes leur support dans $[-1, \infty[\times\mathbb{R}^3$, qui forme un couple convolutif avec Γ . On peut donc appliquer le théorème 7.4.9, et on obtient

$$E \star \mu = \lim_{\varepsilon \rightarrow 0} E \star \mu_\varepsilon.$$

Les fonctions $E \star \mu_\varepsilon$ sont données, pour chaque point $(t, \mathbf{r})$ fixé, par

$$(E \star \mu_\varepsilon)(t, \mathbf{r}) = \langle E(t', \mathbf{r}'), \mu_\varepsilon(t-t', \mathbf{r}-\mathbf{r}') \rangle,$$

formule que l'on peut expliciter en utilisant (8.5). On obtient

$$(E \star \mu_\varepsilon)(t, \mathbf{r}) = \int_{\mathbb{R}^3} \frac{\chi_\varepsilon(t-r')h(\mathbf{r}-\mathbf{r}')}{4\pi r'} d\mathbf{r}'.$$

Il sera commode d'utiliser les coordonnées polaires $\mathbf{r}' = r'\boldsymbol{\theta}$, $\boldsymbol{\theta} \in \mathbb{S}^2$, en notant $\mathbb{S}^2$ la sphère unité de $\mathbb{R}^3$, et dS la mesure de surface de celle-ci. On a

$$(E \star \mu_\varepsilon)(t, \mathbf{r}) = \iint_{[0, \infty] \times \mathbb{S}^2} \chi_\varepsilon(t-r')h(\mathbf{r}-r'\boldsymbol{\theta})r' dr' dS/4\pi.$$

L'intégration en $\boldsymbol{\theta}$ amène à considérer la quantité suivante, définie dans $\mathbb{R} \times \mathbb{R}^3$

$$M(a, \mathbf{r}) = \int_{\mathbb{S}^2} h(\mathbf{r}-a\boldsymbol{\theta}) dS/4\pi,$$

égale à la moyenne de h sur la sphère de centre $\mathbf{r}$ et de rayon $|a|$. On obtient

$$(E \star \mu_\varepsilon)(t, \mathbf{r}) = \int_{[0, \infty[} \chi_\varepsilon(t-r')r' M(r', \mathbf{r}) dr'.$$

Nous avons pu appliquer le théorème de Fubini, la fonction à intégrer étant bornée, et le domaine d'intégration compact grâce au terme $\chi_\varepsilon(t-r')$.

La suite de fonctions $r' \mapsto \chi_\varepsilon(t-r')$ converge vers δ_t , et la fonction M est de classe C^∞ (voir ci-dessous), on a donc convergence vers 0 pour $t < 0$ et vers

la valeur en $r' = t$ pour $t > 0$. La convergence presque partout de fonctions uniformément bornées entraînant la convergence au sens des distributions, on obtient

$$(E \star \mu)(t, \mathbf{r}) = tH(t)M(t, \mathbf{r}) = tH(t) \int_{\mathbb{S}^2} h(\mathbf{r} - t\boldsymbol{\theta}) d\boldsymbol{\theta}/4\pi.$$

8.3.6. Vérification. — En reportant le calcul ci-dessus dans l'expression (8.7), il reste à prouver que nous avons effectivement obtenu une fonction C^∞ vérifiant les conditions initiales voulues.

Dans l'expression intégrale définissant $M(a, \mathbf{r})$, la fonction à intégrer $h(\mathbf{r} - a\boldsymbol{\theta})$ admet des dérivées de tous les ordres par rapport au "paramètre" $(a, \mathbf{r})$, dérivées qui restent uniformément bornées lorsque $(a, \mathbf{r})$ reste dans un compact. D'après le théorème de dérivation sous le signe somme, la fonction $M(a, \mathbf{r})$ est donc de classe C^∞ dans $\mathbb{R}^4$.

Par définition même de M , on a $M(0, \mathbf{r}) = h(\mathbf{r})$. La fonction $tM(t, \mathbf{r})$ est donc une fonction C^∞ , nulle sur l'hyperplan $t = 0$, et dont la dérivée par rapport à t y vaut $h(\mathbf{r})$. La dérivée seconde y est également nulle, la fonction étant impaire par rapport à t .

Si on revient au problème initial, et si on pose

$$u(t, \mathbf{r}) = t \int_{\mathbb{S}^2} h_1(\mathbf{r} - t\boldsymbol{\theta}) d\boldsymbol{\theta}/4\pi + \partial_t \left\{ t \int_{\mathbb{S}^2} h_0(\mathbf{r} - t\boldsymbol{\theta}) d\boldsymbol{\theta}/4\pi \right\},$$

on obtient donc une fonction de classe C^∞ , égale à h_0 pour $t = 0$, et telle que $\partial_t u$ est égal à h_1 sur ce même hyperplan.

Il est inutile de refaire un calcul pour montrer que l'on a $\square u = 0$. La partie "*Explicitation...*" ci-dessus a prouvé que $E \star (\mu_1 + \partial_t \mu_0)$ est égal au produit de $H(t)$ par la fonction u ci-dessus. Le support de $\square(H(t)u)$ est contenu dans $t = 0$, et on a donc $\square u = \square(H(t)u) = 0$ pour $t > 0$. Par continuité de $\square u$, c'est encore valable pour $t = 0$.

Remarque 8.3.7. — La résolution du problème de Cauchy non homogène ($\square u = f$ avec les mêmes conditions pour $t = 0$) se ramène en principe aux deux problèmes précédents. On résout d'abord $\square u_0 = f$ avec u_0 nulle dans le passé. Dans les bons cas, on peut définir k_0 et k_1 comme les traces sur l'hyperplan $t = 0$ de u_0 et de $\partial_t u_0$. Si on appelle v la solution du problème de Cauchy homogène, associé aux données $h_0 - k_0$ et $h_1 - k_1$, on obtient la solution en posant $u = u_0 + v$.

Exercice 8.3.8 (Équation des ondes en dimension 2 d'espace)

On considère dans $\mathbb{R} \times \mathbb{R}^2$, dont on notera (t, x, y) le point courant, le cône plein G défini par $t \geq 0$, $x^2 + y^2 \leq t^2$. Soit E la fonction définie par

$$E(t, x, y) = \begin{cases} \frac{1}{2\pi} \frac{1}{\sqrt{t^2 - x^2 - y^2}} & \text{pour } (t, x, y) \in G \\ 0 & \text{pour } (t, x, y) \notin G \end{cases}$$

(a) Démontrer que la fonction E est localement sommable. On admettra que l'on a

$$\left(\frac{\partial}{\partial t^2} - \frac{\partial}{\partial x^2} - \frac{\partial}{\partial y^2} \right) E = \delta.$$

(b) Énoncer et démontrer l'analogie du théorème 8.3.2 dans ce cas. On montrera que, comme en dimension 3, les causes ne produisent pas d'effet se propageant à une vitesse > 1 mais que, à la différence de la dimension 3, elles en produisent se propageant à des vitesses < 1 .

(c) Donner les formules permettant de résoudre le problème de Cauchy.

8.4. Équations différentielles et intégrales

Nous allons considérer l'espace

$$\mathcal{D}'_+ = \{u \in \mathcal{D}'(\mathbb{R}) \mid \text{Supp}(u) \subset [0, \infty[\}.$$

Il est clair que p exemplaires de $[0, \infty[$ constituent un p -uplet convolutif. Le produit de convolution est une loi interne dans $\mathcal{D}'_+$, associative et commutative. Pour $u \in \mathcal{D}'_+$, il existe au plus un $v \in \mathcal{D}'_+$ tel que $u \star v = \delta$. S'il existe, on notera $v = u^{\star-1}$. Plus généralement, on notera $u^{\star n}$, pour $n \in \mathbb{Z}$, le produit de convolution de $|n|$ exemplaires de u ou v selon le signe de n , en désignant comme il se doit par $u^{\star 0}$ l'élément neutre δ .

Exemple 8.4.1. — On montrera à titre d'exercice que, pour $\lambda \in \mathbb{C}$, la distribution $\delta' - \lambda\delta$ est inversible, et que l'on a

$$\begin{aligned} (\delta' - \lambda\delta)^{\star-1} &= H(t)e^{\lambda t}, \\ (\delta' - \lambda\delta)^{\star-n} &= H(t) \frac{t^{n-1}e^{\lambda t}}{(n-1)!}. \end{aligned}$$

Il en résulte qu'une équation différentielle linéaire à coefficients constants (non tous nuls) possède toujours une solution élémentaire appartenant à $\mathcal{D}'_+$. Il suffit de factoriser $A = \sum_{k=0}^n a_k \delta^k$ sous la forme

$$A = a_n (\delta' - \lambda_1 \delta)^{\star \alpha_1} \star \cdots \star (\delta' - \lambda_l \delta)^{\star \alpha_l},$$

où les λ_j sont les racines du polynôme $\sum a_k t^k$ et les α_j leurs multiplicités. La solution élémentaire est

$$A^{*-1} = a_n^{-1} \prod (\delta' - \lambda_j \delta)^{*- \alpha_j}.$$

Remarque 8.4.2. — Bien entendu, un élément à support compact et inversible de $\mathcal{D}'_+$ peut posséder d'autres solutions élémentaires dans $\mathcal{D}'(\mathbb{R})$, mais il n'en existe qu'une seule à support dans $[0, \infty[$.

Il est beaucoup plus agréable de travailler comme ci-dessus dans une algèbre (associative) de convolution mais, nous l'avons vu, ce n'est pas toujours possible. En dimension quelconque, $\mathcal{E}'(\mathbb{R}^n)$ en est une, mais son intérêt est limité car très peu d'éléments y sont inversibles. Dans l'espace temps $\mathbb{R}^{n+1}$, l'algèbre de convolution formée des distributions à support dans $\{(t, \mathbf{r}) \mid t \geq r\}$ est intéressante, et présente beaucoup d'analogies avec $\mathcal{D}'_+$.

8.4.3. Équations intégrales de seconde espèce. — Le problème est souvent formulé de la manière suivante : trouver $u(t)$ sur $[0, \infty[$ vérifiant

$$u(t) + \int_0^t k(t-s)u(s) ds = f(t). \quad (8.8)$$

Nous supposons k et f localement sommables (y compris au voisinage de l'origine), et reformulerons le problème comme suit : trouver $u \in \mathcal{D}'_+$ vérifiant

$$(\delta + Hk) \star u = g, \quad (8.9)$$

pour $g \in \mathcal{D}'_+$ (par exemple, $g = Hf$). La résolution repose sur le résultat suivant.

Exercice 8.4.4. — Démontrer que la série $\sum_{n=1}^{\infty} (-1)^n (Hk)^{\star n}$ converge en norme L^1 sur chaque compact. On pourra démontrer d'abord le résultat lorsque Hk est à support dans $[\varepsilon, \infty[$, puis écrire Hk dans le cas général comme somme d'une fonction du type précédent, et d'une fonction dont la norme dans L^1 est strictement inférieure à 1.

Si nous notons $H(t)l(t)$ la fonction localement sommable somme de la série précédente, en appliquant le théorème 7.4.9 sur la continuité du produit de convolution (avec $F = G = [0, \infty[$), on obtient

$$(\delta + Hk) \star (\delta + Hl) = \delta,$$

ce qui fournit l'inverse de convolution de $\delta + Hk$. Le problème (8.9) est donc résolu par $u = (\delta + Hl) \star g$. En particulier, le problème initial (8.8) se résout par

$$u(t) = f(t) + \int_0^t l(t-s)f(s) ds, \quad t \geq 0.$$

8.4.5. *Équations intégrales de première espèce.* — Il s'agit d'équations que l'on écrit souvent

$$\int_0^t k(t-s)u(s) ds = f(t), \quad t \geq 0 \quad (8.10)$$

et qu'il est nécessaire cette fois-ci (ce n'était qu'utile dans le cas précédent) de reformuler en termes de distributions : pour $g \in \mathcal{D}'_+$, trouver u dans le même espace vérifiant $(Hk) \star u = g$.

Nous supposons k de classe C^1 , avec $k(0) \neq 0$. L'équation précédente est équivalente à l'équation obtenue en dérivant les deux membres (on passe de l'une à l'autre en convoluant par δ' ou par son inverse H)

$$(k(0)\delta + Hk') \star u = g'$$

où on a utilisé la formule des sauts. En posant $a = k(0)$, il s'agit de déterminer l'inverse de $\delta + a^{-1}Hk'$. La résolution des équations de seconde espèce nous a appris que cet inverse existe et est de la forme $\delta + Hl$. On obtient donc la solution de $(Hk) \star u = g$ par $u = a^{-1}(g' + Hl \star g')$.

Revenons maintenant au problème (8.10), qui correspond au cas $g = Hf$, en supposant la fonction f de classe C^1 . On obtient donc

$$\begin{aligned} u &= a^{-1} ((\delta + Hl) \star (f(0)\delta + Hf')) \\ &= a^{-1} \left(f(0)\delta + H(t) \left\{ l(t) + f'(t) + \int_0^t l(t-s)f'(s) ds \right\} \right). \end{aligned}$$

Si $f(0) = 0$, le problème formulé par (8.10) possède une solution donnée par la fonction entre accolades ci-dessus multipliée par a^{-1} . Dans le cas contraire, il n'y a pas de fonction solution (ce qui était prévisible dès le début en regardant les limites pour $t \rightarrow 0$), mais il y a une distribution solution qui est somme d'une fonction et d'une masse ponctuelle à l'origine.

Exercice 8.4.6. — Résoudre la même équation en supposant k de classe C^2 , avec $k(0) = 0$ et $k'(0) \neq 0$.

8.4.7. *Calcul symbolique de Heaviside.* — Les distributions dont le support est réduit à l'origine sont de la forme $\sum_{k=0}^n a_k \delta^k = \sum_{k=0}^n a_k (\delta')^{\star k}$. Elles forment un anneau $\mathcal{P}$ pour la convolution, isomorphe à l'anneau des polynômes. Comme nous l'avons vu plus haut, tout élément non nul de $\mathcal{P}$ admet un inverse dans $\mathcal{D}'_+$, et le corps des fractions de $\mathcal{P}$ s'identifie donc à un sous-corps $\mathcal{K}$ de $\mathcal{D}'_+$, isomorphe au corps des fractions rationnelles (pour la multiplication usuelle) en une variable.

Il est traditionnel de noter p la variable en question. L'isomorphisme précédent est facile à décrire. À la distribution δ correspond l'élément neutre 1, à δ' correspond p , à $\delta^k = (\delta')^{\star k}$ correspond p^k , et à la fonction

$H(t)e^{\lambda t}t^{k-1}/(k-1)! = (\delta' - \lambda\delta)^{\star-k}$ correspond $(p - \lambda)^{-k}$. La décomposition des fractions rationnelles en éléments simples nous assure que $\mathcal{K}$ est constitué des combinaisons $\mathbb{C}$ -linéaires des distributions précédentes.

Si maintenant on veut résoudre dans $\mathcal{D}'_+$ une équation de convolution

$$A \star U = B$$

où A et B appartiennent à $\mathcal{K}$, on écrit les fractions rationnelles $\mathcal{A}(p)$ et $\mathcal{B}(p)$ correspondantes. La solution doit correspondre à $\mathcal{U}(p) = \mathcal{B}(p)/\mathcal{A}(p)$. On décompose cette fraction rationnelle en éléments simples, et les correspondances ci-dessus donnent u sans aucun calcul supplémentaire.

Cette méthode fournit notamment les solutions élémentaires à support dans $[0, \infty[$ de toutes les équations différentielles à coefficients constants, mais aussi d'équations intégro-différentielles.

Exercice 8.4.8. — Résoudre $u'(t) + \int_0^t \cos(t-s)u(s)ds = f(t)$ pour $t \geq 0$, ou plus précisément, pour $f \in L^1_{\text{loc}}(\mathbb{R})$, résoudre dans $\mathcal{D}'_+$

$$(\delta' + H(t) \cos(t)) \star u = Hf.$$

Nous verrons plus loin que la transformation de Laplace nous donnera une autre interprétation de l'isomorphisme ci-dessus, et nous permettra de l'étendre à un ensemble de distributions beaucoup plus gros que $\mathcal{K}$.

Ce calcul symbolique a été introduit, avec comme on l'imagine des “justifications” très différentes, par l'ingénieur anglais Heaviside en 1893.

CHAPITRE 9

TRANSFORMATION DE FOURIER

La transformation de Fourier nous permettra, grâce au théorème d'inversion, d'écrire une fonction sommable et plus généralement une distribution tempérée comme une superposition de fonctions exponentielles complexes. C'est un outil très puissant pour étudier les équations de convolution, la résolution de celles-ci se ramenant à des problèmes de division.

Comme on l'imagine, il est important de pouvoir utiliser cette transformation dans des conditions les plus générales possibles. Nous la définirons d'abord dans le cadre, trop restreint, des fonctions sommables. Pour l'étendre, nous aurons ensuite recours à une stratégie dont l'efficacité n'est plus à démontrer : reculer pour mieux sauter. Nous introduirons un espace $\mathcal{S}$ beaucoup plus petit que L^1 — mais dont le dual $\mathcal{S}'$ est beaucoup plus gros — qui jouit de bonnes propriétés d'invariance pour la transformation de Fourier. Par dualité, nous pourrons enfin étendre cette transformation à $\mathcal{S}'$.

Il n'est malheureusement pas possible de définir la transformée de Fourier d'une distribution quelconque u . Cela dit, la limitation $u \in \mathcal{S}'$ n'est pas très restrictive : en un sens qu'il faudra préciser, les éléments de $\mathcal{S}'$ sont des distributions quelconques à distance finie, mais dont la croissance doit être au plus polynomiale à l'infini.

9.1. Transformation de Fourier des fonctions sommables

Théorème et Définition 9.1.1. — Soit $f \in L^1(\mathbb{R}^n)$. On appelle transformée de Fourier de f la fonction, notée $\widehat{f}$ ou $\mathcal{F}(f)$, définie pour $\xi \in \mathbb{R}^n$ par

$$\widehat{f}(\xi) = \int e^{-i \cdot x \cdot \xi} f(x) dx, \quad (9.1)$$

en notant $x \cdot \xi$ le produit scalaire de $\mathbb{R}^n$. La fonction $\widehat{f}$ est continue, tend vers 0 à l'infini, et vérifie

$$\|\widehat{f}\|_{L^\infty} \leq \|f\|_{L^1}. \quad (9.2)$$

Il est clair que la valeur de $\widehat{f}$ en chaque point est majorée en module par $\int |f(x)| dx$, d'où l'inégalité (9.2). D'autre part, si une suite ξ_j converge vers ξ_0 , les fonctions $e^{-ix \cdot \xi_j} f(x)$ convergent en chaque point vers $e^{-ix \cdot \xi_0} f(x)$, et sont majorées en module par la fonction sommable fixe $|f(x)|$. La continuité de $\widehat{f}$ résulte du théorème de Lebesgue.

Nous verrons ci-dessous (théorème 9.2.4) que, si $\varphi \in C_0^\infty(\mathbb{R}^n)$, la fonction $\widehat{\varphi}$ tend vers 0 à l'infini. Pour chaque $f \in L^1$, on peut trouver une suite φ_j d'éléments de C_0^∞ telle que $\|f - \varphi_j\|_{L^1}$ tende vers 0 (voir théorème 3.5.1). Il résulte de (9.2) que $\|\widehat{f} - \widehat{\varphi_j}\|_{L^\infty}$ tend vers 0, et la fonction $\widehat{f}$, qui est limite uniforme de fonctions tendant vers 0 à l'infini, possède la même propriété.

9.1.2. Transformées des fonctions gaussiennes. — Nous présentons ci-dessous, sous forme d'une suite d'exercices élémentaires, la démonstration du résultat classique

$$\mathcal{F}\left(e^{-a|x|^2}\right) = \left(\frac{\pi}{a}\right)^{n/2} e^{-|\xi|^2/4a}, \quad (9.3)$$

où a est une constante > 0 .

Exercice 9.1.3. — En dimension 1, posons $g(\xi) = \int e^{-ix\xi} e^{-x^2} dx$. Démontrer que g est de classe C^1 , et déduire de l'expression de g' que g est solution de l'équation différentielle $2g'(\xi) + \xi g(\xi) = 0$.

Démontrer que l'on a $g(\xi) = \sqrt{\pi} e^{-\xi^2/4}$ (on utilisera $\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}$).

Exercice 9.1.4. — Montrer que, si $f \in L^1(\mathbb{R})$, la transformée de Fourier de la fonction $f(x/\lambda)$ est la fonction $|\lambda| \widehat{f}(\lambda\xi)$, et en déduire (9.3) pour $n = 1$.

Soient f_j , $j = 1, \dots, n$, des éléments de $L^1(\mathbb{R})$. Démontrer que la transformée de Fourier dans $\mathbb{R}^n$ de la fonction $F(x) = f_1(x_1) \dots f_n(x_n)$ est la fonction $\widehat{F}(\xi) = \widehat{f_1}(\xi_1) \dots \widehat{f_n}(\xi_n)$. En déduire (9.3) dans le cas général.

Théorème 9.1.5 (Inversion de Fourier). — Soit $f \in L^1(\mathbb{R}^n)$ telle que $\widehat{f} \in L^1(\mathbb{R}^n)$. On a alors

$$f = (2\pi)^{-n} \overline{\mathcal{F}}(\widehat{f}), \quad (9.4)$$

où on a noté $\overline{\mathcal{F}}$ l'analogue de la transformation de Fourier obtenue en remplaçant i par $-i$ dans (9.1).

La valeur au point x de la fonction figurant au membre de droite peut s'écrire $(2\pi)^{-n} \int (\int e^{i(x-y)\cdot\xi} f(y) dy) d\xi$, mais il est impossible d'appliquer le théorème de Fubini, la fonction à intégrer n'étant pas sommable dans $\mathbb{F}^{2n}$. Nous allons la multiplier par la fonction $e^{-\varepsilon^2|\xi|^2/4}$ qui tend vers 1 pour $\varepsilon \rightarrow 0$, et évaluer de deux manières différentes l'intégrale.

Posons donc

$$I_\varepsilon(x) = (2\pi)^{-n} \iint_{\mathbb{F}^{2n}} e^{i(x-y)\cdot\xi} e^{-\varepsilon^2|\xi|^2/4} f(y) dy d\xi. \quad (9.5)$$

Cette fois-ci, le théorème de Fubini est applicable, et on obtient en intégrant d'abord par rapport à y

$$I_\varepsilon(x) = (2\pi)^{-n} \int e^{ix\cdot\xi} e^{-\varepsilon^2|\xi|^2/4} \widehat{f}(\xi) d\xi.$$

Lorsque $\varepsilon \rightarrow 0$, la fonction à intégrer reste majorée en module par la fonction sommable fixe $|\widehat{f}(\xi)|$. Le théorème de Lebesgue entraîne que l'on a, pour chaque x

$$\lim_{\varepsilon \rightarrow 0} I_\varepsilon(x) = (2\pi)^{-n} \int e^{ix\cdot\xi} \widehat{f}(\xi) d\xi.$$

Si nous calculons maintenant (9.5) en intégrant d'abord par rapport à ξ , on obtient

$$I_\varepsilon(x) = \int G_\varepsilon(x-y) f(y) dy,$$

c'est-à-dire $I_\varepsilon = G_\varepsilon \star f$, où on a posé

$$G_\varepsilon(z) = (2\pi)^{-n} \int e^{iz\cdot\xi} e^{-\varepsilon^2|\xi|^2/4} d\xi.$$

En utilisant (9.3) avec $a = \varepsilon^2/4$, on obtient

$$G_\varepsilon(z) = \varepsilon^{-n} G_1(z/\varepsilon)$$

avec

$$G_1(z) = \pi^{-n/2} e^{-|z|^2}.$$

En notant que la fonction G_1 est d'intégrale 1, on voit que la famille des G_ε a toutes les propriétés d'une approximation de l'identité, à l'exception du fait que les gaussiennes ne sont pas à support compact. Cela n'empêche pas les $G_\varepsilon \star f$ de converger en norme L^1 vers f pour toute fonction sommable f (voir exercice (3.4.7)).

En résumé, les fonctions I_ε convergent vers $(2\pi)^{-n} \overline{\mathcal{F}}(\widehat{f})$ en chaque point x , et donc au sens des distributions puisqu'elles sont majorées par la constante $(2\pi)^{-n} \|\widehat{f}\|_{L^1}$. D'autre part, les I_ε convergent vers f en norme dans L^1 et

donc au sens des distributions. Les fonctions f et $(2\pi)^{-n}\widehat{\mathcal{F}}(\widehat{f})$ sont égales en tant que distributions, et donc en tant qu'éléments de L^1 .

Remarque 9.1.6. — A posteriori, on voit que les hypothèses du théorème ne peuvent être remplies que si f est (égale presque partout à) une fonction continue tendant vers 0 à l'infini, $\widehat{\mathcal{F}}$ transformant les fonctions sommables en des fonctions de ce type. Cela écarte des fonctions d'un usage tout à fait courant en mathématiques comme en physique.

La nécessité de sortir du cadre des fonctions sommables est en fait apparue très tôt. Les grandes étapes en sont la définition de la transformée de Fourier des fonctions de carré sommable par Plancherel, et l'extension par L. Schwartz aux distributions tempérées.

Exercice 9.1.7. — Déterminer, sur $\mathbb{R}$, la transformée de Fourier de la fonction égale à 1 sur $[-a, a]$ et à 0 ailleurs.

Déterminer, pour $\lambda > 0$, les transformées des fonctions $H(x)e^{-\lambda x}$ et $e^{-\lambda|x|}$. En déduire la transformée de Fourier de la fonction $1/(x^2 + \lambda^2)$.

Remarque 9.1.8 (Questions d'invariance). — La structure euclidienne de $\mathbb{R}^n$ n'est nullement nécessaire pour définir la transformation de Fourier. Sur un espace vectoriel réel E de dimension n , la transformée de Fourier d'une fonction sommable f est définie de manière naturelle comme une fonction sur le dual E^* par

$$\forall \xi \in E^*, \quad \widehat{f}(\xi) = \int e^{-i\langle x, \xi \rangle} f(x) dx,$$

le seul choix arbitraire est celui de l'élément de volume dx qui n'est défini par la structure vectorielle de E qu'à un facteur près.

Plus généralement, sur un groupe abélien localement compact G , on introduit le *groupe dual* $\widehat{G}$ constitué des *caractères* de G , c'est-à-dire des homomorphismes continus de G dans le groupe multiplicatif des nombres complexes de module 1. Il existe d'autre part une mesure dg sur G invariante par translation, unique à un facteur près. La transformée de Fourier d'une fonction sommable sur G est la fonction définie sur $\widehat{G}$ par

$$\widehat{f}(\gamma) = \int_G \gamma(g) f(g) dg.$$

Le groupe abélien $\widehat{G}$ est naturellement muni d'une topologie d'espace localement compact, son groupe dual est canoniquement isomorphe à G , et la formule d'inversion de Fourier est valable dans ce cas.

Le groupe dual $\widehat{E}$ d'un espace vectoriel E s'identifie à son espace vectoriel dual E^* par l'application qui à $\xi \in E^*$ fait correspondre le caractère $x \mapsto e^{i\langle x, \xi \rangle}$; on retrouve la définition ci-dessus. Le groupe dual de $\mathbb{R}/(2\pi\mathbb{Z})$ s'identifie à $\mathbb{Z}$ par l'application qui à l'entier p fait correspondre le caractère $x \mapsto e^{ipx}$; on retrouve la théorie des séries de Fourier.

Remarque 9.1.9 (Questions de normalisation). — Il existe des variantes dans la définition de la transformation de Fourier dans $\mathbb{R}^n$. La plus courante est

la suivante, qui revient à choisir la fréquence au lieu de la pulsation comme variable,

$$\mathcal{F}_1 f(\xi) = \int e^{-2i\pi x \cdot \xi} f(x) dx.$$

L'avantage est que l'on a $\mathcal{F}_1^{-1} = \overline{\mathcal{F}_1}$, et que $\mathcal{F}_1$ est une isométrie de L^2 . L'inconvénient est que les facteurs 2π réapparaissent lorsqu'on écrit la transformée de Fourier d'une dérivée.

On utilise aussi la transformation $f \mapsto (2\pi)^{-n/2} \mathcal{F}(f)$, qui est une isométrie de L^2 , mais pour qui la convolution n'est transformée en multiplication qu'à un facteur près.

Il est important, en lisant un ouvrage ou une table de transformées de Fourier de s'assurer de la normalisation utilisée. Il est bien sûr facile de passer de l'une à l'autre des conventions. Le lecteur, rompu aux problèmes de changement d'unités en Physique, trouvera là un nouveau champ d'application à son activité favorite.

9.2. L'espace $\mathcal{S}$ de Schwartz

Définition 9.2.1. — On dit que φ appartient à $\mathcal{S}(\mathbb{R}^n)$ si $\varphi \in C^\infty$ et si φ et toutes ses dérivées sont "à décroissance rapide", c'est-à-dire que leur produit par un polynôme quelconque est une fonction bornée.

Il est équivalent de dire que les quantités suivantes

$$\mathcal{N}_p(\varphi) = \sum_{|\alpha| \leq p, |\beta| \leq p} \left\| x^\alpha \partial^\beta \varphi(x) \right\|_{L^\infty}$$

sont finies pour tout p .

Remarque 9.2.2. — L'espace $\mathcal{S}$ contient C_0^∞ . En fait, alors qu'il n'est pas totalement évident que C_0^∞ contienne des fonctions non nulles, il est beaucoup plus facile d'exhiber des éléments de $\mathcal{S}$. Par exemple, le produit d'une gaussienne par un polynôme appartient à $\mathcal{S}$.

Proposition 9.2.3. — L'espace $\mathcal{S}$ est stable par dérivation et par multiplication par les polynômes. Les éléments de $\mathcal{S}$ sont des fonctions sommables tendant vers 0 à l'infini, et il existe des constantes C_p telles que

$$\forall \varphi \in \mathcal{S}, \quad \sum_{|\alpha| \leq p, |\beta| \leq p} \left\| x^\alpha \partial^\beta \varphi(x) \right\|_{L^1} \leq C_p \mathcal{N}_{p+n+1}(\varphi). \quad (9.6)$$

Il s'agit de conséquences immédiates de la définition. La dernière majoration provient du fait que, pour $|\alpha| \leq p$, $|\beta| \leq p$, on a

$$\left\| (1 + \sum |x_i|^{n+1}) x^\alpha \partial^\beta \varphi(x) \right\|_{L^\infty} \leq \mathcal{N}_{p+n+1}(\varphi).$$

On obtient donc

$$\left\| x^\alpha \partial^\beta \varphi(x) \right\|_{L^1} \leq \mathcal{N}_{p+n+1}(\varphi) \int \frac{1}{1 + \sum_i |x_i|^{n+1}} dx,$$

et, l'intégrale étant finie, l'estimation (9.6).

Théorème 9.2.4. — (a) *La transformation de Fourier applique l'espace $\mathcal{S}$ dans lui-même, et il existe des constantes C_p telles que*

$$\mathcal{N}_p(\widehat{\varphi}) \leq C_p \mathcal{N}_{p+n+1}(\varphi). \quad (9.7)$$

(b) *La transformation de Fourier $\mathcal{F}$ est un isomorphisme de $\mathcal{S}$ sur lui-même, d'inverse $\mathcal{F}^{-1} = (2\pi)^{-n} \overline{\mathcal{F}}$.*

Il suffit de démontrer la partie (a). En effet, le théorème d'inversion, applicable lorsqu'on sait que φ et $\widehat{\varphi}$ sont dans $\mathcal{S}$ et donc sommables, assure que $(2\pi)^{-n} \overline{\mathcal{F}} \mathcal{F}$ et $(2\pi)^{-n} \mathcal{F} \overline{\mathcal{F}}$ coïncident avec l'identité sur $\mathcal{S}$. La partie (a) résultera des deux lemmes suivants.

Lemme 9.2.5. — *Soit $\varphi \in \mathcal{S}$. Alors $\widehat{\varphi}$ est de classe C^1 , et on a*

$$\partial_j(\mathcal{F}\varphi) = \mathcal{F}\{(-ix_j)\varphi(x)\}.$$

Il suffit d'appliquer le théorème de dérivation sous le signe somme à l'intégrale $\int e^{-ix \cdot \xi} \varphi(x) dx$. La dérivée par rapport à ξ_j donne $e^{-ix \cdot \xi} (-ix_j) \varphi(x)$, fonction majorée en module par la fonction fixe $|x_j \varphi(x)|$. La fonction $x_j \varphi$ appartient à $\mathcal{S}$ et donc à L^1 , d'où le résultat.

Lemme 9.2.6. — *Si $\varphi \in \mathcal{S}$, on a*

$$\mathcal{F}(\partial_j \varphi) = i \xi_j \widehat{\varphi}(\xi). \quad (9.8)$$

Pour simplifier les notations, nous supposons $j = n$, et nous poserons $x' = (x_1, \dots, x_{n-1})$. Pour chaque x' , on a

$$\int_{\mathbb{R}} e^{-ix_n \xi_n} \partial_n \varphi(x', x_n) dx_n = \int_{\mathbb{R}} (i \xi_n) e^{-ix_n \xi_n} \varphi(x', x_n) dx_n$$

par intégration par parties, la fonction tendant vers 0 pour $x_n \rightarrow \pm\infty$. Il suffit maintenant de multiplier chacun des membres par $e^{-ix' \cdot \xi'}$ et d'intégrer par rapport à x' pour obtenir (9.8). Nous avons pu appliquer le théorème de Fubini : les fonctions φ et $\partial_n \varphi$ appartiennent à $\mathcal{S}$ et sont donc sommables dans $\mathbb{R}^n$.

9.2.7. *Fin de la démonstration du théorème 9.2.4.* — En appliquant $|\beta|$ fois de suite le lemme 9.2.5, on obtient $\partial_\xi^\beta \widehat{\varphi} = (-i)^{|\beta|} \mathcal{F}(x^\beta \varphi)$. En appliquant ensuite $|\alpha|$ fois le lemme 9.2.6, on a

$$\left| \xi^\alpha \partial_\xi^\beta \widehat{\varphi}(\xi) \right| = \left| \mathcal{F} \left\{ \partial_x^\alpha (x^\beta \varphi(x)) \right\} \right|.$$

A l'aide de l'estimation (9.2), on en déduit

$$\mathcal{N}_p(\widehat{\varphi}) \leq \sum_{|\alpha| \leq p, |\beta| \leq p} \left\| \partial_x^\alpha (x^\beta \varphi) \right\|_{L^1} \leq C'_p \sum_{|\alpha'| \leq p, |\beta'| \leq p} \left\| x^{\alpha'} \partial_x^{\beta'} \varphi \right\|_{L^1}$$

en utilisant la formule de Leibniz. Pour conclure, il ne reste plus qu'à majorer le membre de droite par $C''_p \mathcal{N}_{p+n+1}(\varphi)$ à l'aide de la proposition 9.2.3.

Remarque 9.2.8. — On voit apparaître ici pour la première fois un phénomène très important : la transformation de Fourier échange régularité et décroissance à l'infini. C'est parce que les produits de φ par les monômes de degré $\leq p$ sont sommables que $\widehat{\varphi}$ est p fois dérivable, et c'est parce que les dérivées de φ jusqu'à l'ordre p sont sommables que $\widehat{\varphi}(\xi)$ est un $O(|\xi|^{-p})$ à l'infini.

Le lecteur pourra vérifier que les $\mathcal{N}_p$ sont des normes sur $\mathcal{S}$ et, s'il a lu l'appendice C, verra que la majoration (9.7) exprime que l'opérateur $\mathcal{F}$ est continu de $\mathcal{S}$ dans lui-même, lorsque ce dernier espace est muni de la structure d'espace de Fréchet associée aux $\mathcal{N}_p$. De même, le théorème suivant signifie que C_0^∞ est dense dans $\mathcal{S}$.

Théorème 9.2.9. — *Soit $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Il existe alors une suite φ_j d'éléments de $C_0^\infty(\mathbb{R}^n)$ telle que l'on ait pour tout p*

$$\lim_{j \rightarrow \infty} \mathcal{N}_p(\varphi - \varphi_j) = 0.$$

Soit $\Psi \in C_0^\infty$ une fonction égale à 1 dans la boule de centre 0 et de rayon 1, et posons $\varphi_j(x) = \varphi(x)\Psi(x/j)$. Ces fonctions sont à support compact, et coïncident avec φ sur la boule de rayon j . D'après la formule de Leibniz, on a

$$\partial^\beta (\varphi - \varphi_j)(x) = \partial^\beta \varphi(x) (1 - \Psi(x/j)) - \sum_{\substack{|\gamma| \geq 1 \\ \gamma \leq \beta}} \binom{\beta}{\gamma} \frac{1}{j^{|\gamma|}} \partial^{\beta-\gamma} \varphi(x) \partial^\gamma \Psi(x/j).$$

En multipliant par x^α la relation précédente, et en prenant les bornes supérieures, on en déduit

$$\left\| x^\alpha \partial^\beta (\varphi - \varphi_j) \right\|_{L^\infty} \leq \max_{|x| \geq j} \left| x^\alpha \partial^\beta \varphi(x) \right| + \frac{C}{j} \sum_{\gamma \leq \beta} \left\| x^\alpha \partial^{\beta-\gamma} \varphi \right\|_{L^\infty}.$$

Le premier terme tend vers 0 pour $j \rightarrow \infty$, la fonction $x^\alpha \partial^\beta \varphi$ tendant vers 0 à l'infini, et le second terme contient $1/j$ en facteur d'une quantité finie. L'expression $\mathcal{N}_p(\varphi - \varphi_j)$ est ainsi une somme finie de quantités tendant vers 0, d'où le résultat.

9.3. L'espace $\mathcal{S}'$ des distributions tempérées

Définition 9.3.1. — Soit $u \in \mathcal{D}'(\mathbb{R}^n)$. On dit que u est une distribution tempérée, ce qu'on note $u \in \mathcal{S}'(\mathbb{R}^n)$, s'il existe $p \in \mathbb{N}$ et $C \geq 0$ tels que

$$\forall \varphi \in C_0^\infty(\mathbb{R}^n), \quad |\langle u, \varphi \rangle| \leq C \mathcal{N}_p(\varphi). \quad (9.9)$$

Théorème 9.3.2 (Extension de la dualité). — Soit $u \in \mathcal{S}'$. L'application $\varphi \mapsto \langle u, \varphi \rangle$, définie sur C_0^∞ , se prolonge de manière unique en une forme linéaire sur $\mathcal{S}$ (que l'on notera encore $\varphi \mapsto \langle u, \varphi \rangle$) qui vérifie

$$\forall \varphi \in \mathcal{S}(\mathbb{R}^n), \quad |\langle u, \varphi \rangle| \leq C \mathcal{N}_p(\varphi). \quad (9.10)$$

Cette extension de la dualité identifie $\mathcal{S}'$ à l'espace des formes linéaires sur $\mathcal{S}$ qui vérifient une estimation du type précédent.

Cette situation (existence et unicité du prolongement continu d'une forme linéaire continue définie sur un sous-espace dense) est bien connue du lecteur, au moins dans le cas des espaces normés. On utilise ici le théorème 9.2.9 : pour tout $\varphi \in \mathcal{S}$, on peut trouver une suite φ_j dans C_0^∞ telle que $\mathcal{N}_p(\varphi - \varphi_j) \rightarrow 0$ pour tout p . Si $u \in \mathcal{S}'$, il résulte de la majoration (9.9) que la suite numérique $\langle u, \varphi_j \rangle$ est de Cauchy, et a donc une limite. La même estimation assure que cette limite, que l'on notera $\langle u, \varphi \rangle$ ne dépend pas de la suite φ_j choisie, et que l'on a l'estimation (9.10). Enfin, si un autre prolongement u_1 vérifiait (9.10), on aurait $\langle u_1, \varphi - \varphi_j \rangle \rightarrow 0$, et donc $u_1 = u$.

Si on se donne maintenant une forme linéaire sur $\mathcal{S}$ vérifiant une majoration du type (9.10), elle définit évidemment par restriction à C_0^∞ une distribution tempérée u , et l'unicité précédemment démontrée prouve qu'elle coïncide avec l'extension de u .

9.3.3. Exemples importants

— Toute fonction f localement sommable et majorée par un polynôme est une distribution tempérée. Si en effet on a $f(x) \leq C(1 + |x|^N)$, on voit facilement que $\int f(x)\varphi(x) dx$ est majoré en module par $C' \mathcal{N}_{N+n+1}(\varphi)$.

— On montrera de même, à titre d'exercice, que toute fonction sommable, ou de carré sommable, appartient à $\mathcal{S}'$.

— Toute distribution u à support compact est tempérée. On sait en effet, d'après le théorème 6.1.3, que, si K est un voisinage de $\text{Supp}(u)$, il existe C et p tels que l'on ait

$$|\langle u, \varphi \rangle| \leq C \sup_{x \in K, |\alpha| \leq p} |\partial^\alpha \varphi(x)|,$$

et on majore facilement le membre de droite par $C' \mathcal{N}_p(\varphi)$.

— Une méthode courante pour prouver qu'une distribution est tempérée consiste à l'écrire comme une somme de distributions vérifiant l'un des critères précédents.

Exercice 9.3.4. — Démontrer que $u = \text{vp}(1/x)$ appartient à $\mathcal{S}'(\mathbb{R})$. On écrira $u = \varphi u + (1 - \varphi)u$, où φ appartient à C_0^∞ et vaut 1 près de l'origine.

Exercice 9.3.5. — Démontrer que, pour tout $\varphi \in C_0^\infty$ et pour chaque entier p , il existe une constante C telle que $\mathcal{N}_p(\tau_a \varphi) \leq C(1 + |a|)^p$. En déduire que la fonction e^x n'appartient pas à $\mathcal{S}'(\mathbb{R})$.

Exercice 9.3.6. — Démontrer que la distribution $u = \sum_{k=-\infty}^{+\infty} a_k \delta_k$ appartient à $\mathcal{S}'(\mathbb{R})$ si et seulement si la suite a_k est à croissance lente (c'est-à-dire s'il existe C et N avec $|a_k| \leq C(1 + |k|)^N$).

Exercice 9.3.7. — Soient m un entier positif et u un élément de l'espace de Sobolev H^{-m} . Démontrer que $u \in \mathcal{S}'$.

Définition 9.3.8 (Convergence dans $\mathcal{S}'$). — On dit que la suite u_j d'éléments de $\mathcal{S}'(\mathbb{R}^n)$ converge vers u dans $\mathcal{S}'(\mathbb{R}^n)$, si on a

$$\forall \varphi \in \mathcal{S}(\mathbb{R}^n), \quad \lim_{j \rightarrow \infty} \langle u_j, \varphi \rangle = \langle u, \varphi \rangle.$$

Remarque 9.3.9. — Il est clair que, si $u_j \rightarrow u$ dans $\mathcal{S}'$, alors $u_j \rightarrow u$ dans $\mathcal{D}'$, les éléments de C_0^∞ appartenant à $\mathcal{S}$.

D'autre part, si les u_j ont leur support dans un même compact K , la convergence dans $\mathcal{D}'$ est équivalente à la convergence dans $\mathcal{S}'$. Il suffit de choisir $\theta \in C_0^\infty$ égale à 1 au voisinage de K , et on a, pour $\varphi \in \mathcal{S}$,

$$\langle u_j, \varphi \rangle = \langle u_j, \theta \varphi \rangle \rightarrow \langle u, \theta \varphi \rangle = \langle u, \varphi \rangle$$

dès que l'on a convergence dans $\mathcal{D}'$.

La différence entre les deux notions de convergence n'apparaît donc qu'à l'infini, et est bien illustrée par l'exercice suivant.

Exercice 9.3.10. — Soit a_k une suite numérique. Montrer que l'on a toujours $a_k \delta_k \rightarrow 0$ dans $\mathcal{D}'(\mathbb{R})$, mais que cette même suite tend vers 0 dans $\mathcal{S}'(\mathbb{R})$ si et seulement si la suite a_k est à croissance lente.

9.3.11. Exemples importants. — Soient $f_j, j \in \mathbb{N}$ et f des fonctions localement sommables. Le lecteur démontrera à titre d'exercice que sous l'une des hypothèses suivantes, on a $f_j \rightarrow f$ dans $\mathcal{S}'$.

— Les f_j sont sommables et $\|f - f_j\|_{L^1} \rightarrow 0$

..... Les f_j sont de carré sommables et $\|f - f_j\|_{L^2} \rightarrow 0$

..... Les f_j sont bornées et $\|f - f_j\|_{L^\infty} \rightarrow 0$

— $f_j \rightarrow f$ presque partout, et les $|f_j(x)|$ sont majorées par un polynôme fixe.

Théorème 9.3.12. — Si $u \in S'(\mathbb{R}^n)$, alors toutes ses dérivées appartiennent à $S'(\mathbb{R}^n)$. De plus, si $u_j \rightarrow u$ dans S' , on a $\partial^\alpha u_j \rightarrow \partial^\alpha u$ dans S' .

Pour $\varphi \in C_0^\infty$, on a en effet

$$|\langle \partial_i u, \varphi \rangle| = |\langle u, \partial_i \varphi \rangle| \leq C N_p(\partial_i \varphi) \leq C N_{p+1}(\varphi),$$

ce qui prouve que $\partial_i u$ appartient à S' s'il en est de même de u .

Il résulte immédiatement de l'unicité dans le théorème 9.3.2 que la relation $\langle \partial_i u, \varphi \rangle = -\langle u, \partial_i \varphi \rangle$ est encore valable pour $\varphi \in S$. Si $u_j \rightarrow u$ dans S' , on a alors pour $\varphi \in S$

$$\langle \partial_i u_j, \varphi \rangle = -\langle u_j, \partial_i \varphi \rangle \rightarrow -\langle u, \partial_i \varphi \rangle = \langle \partial_i u, \varphi \rangle,$$

ce qui démontre le résultat.

Exercice 9.3.13. — Démontrer que la fonction $e^x (\cos(e^x) + i \sin(e^x))$ appartient à $S'(\mathbb{R})$ (on en cherchera une primitive). Le lecteur comparera avec le résultat de l'exercice 9.3.5 et y verra une nouvelle illustration du fait qu'une fonction qui oscille beaucoup est "petite" au sens des distributions (voir aussi la remarque 6.3.2).

Définition 9.3.14. — On dit qu'une fonction f est à croissance lente ainsi que toute ses dérivées, ce qu'on note $f \in \mathcal{O}_M(\mathbb{R}^n)$, si f est de classe C^∞ , et si pour tout $\beta \in \mathbb{N}^n$, il existe C_β et m_β tels que

$$|\partial^\beta f(x)| \leq C_\beta (1 + |x|)^{m_\beta}. \quad (9.11)$$

$$\begin{array}{ccccccc} C_0^\infty & \longrightarrow & S & \longrightarrow & \mathcal{O}_M & \longrightarrow & C^\infty \\ \downarrow & & & & \downarrow & & \downarrow \\ \mathcal{E}' & \longrightarrow & & & S' & \longrightarrow & \mathcal{D}' \end{array}$$

FIGURE 1. Les flèches représentent des inclusions

Théorème 9.3.15. — Soit $f \in \mathcal{O}_M$.

(a) Pour tout $\varphi \in S$, on a $f\varphi \in S$.

(b) Pour tout $u \in S'$, on a $fu \in S'$. Si $u_j \rightarrow u$ dans S' , alors $fu_j \rightarrow fu$ dans S' .

Il faut prouver que les quantités $N_p(f\varphi)$ sont finies pour $\varphi \in S$. En utilisant la formule de Leibniz, on a

$$\sum_{|\alpha| \leq p, |\beta| \leq p} |x^\alpha \partial^\beta (f\varphi)(x)| \leq C_p \sum_{|\alpha| \leq p, |\beta| \leq p, |\gamma| \leq p} |x^\alpha \partial^\beta f(x) \partial^\gamma \varphi(x)|,$$

où C_p s'exprime à partir d'un nombre fini de coefficients binomiaux. Si on désigne par M_p le plus grand des m_β intervenant dans (9.11) pour $|\beta| \leq p$, on obtient avec une autre constante C'_p indépendante de φ

$$\mathcal{N}_p(f\varphi) \leq C'_p \mathcal{N}_{p+M_p}(\varphi).$$

On en déduit la partie (a) du théorème.

Si maintenant $u \in \mathcal{S}'$, il existe C et p tels que l'on ait, pour $\varphi \in \mathcal{C}_0^\infty$,

$$|\langle fu, \varphi \rangle| = |\langle u, f\varphi \rangle| \leq C \mathcal{N}_p(f\varphi) \leq CC'_p \mathcal{N}_{p+M_p}(\varphi).$$

Cela prouve que la distribution fu est tempérée. Enfin, si $u_j \rightarrow u$ dans $\mathcal{S}'$, on a $\langle fu_j, \varphi \rangle = \langle u_j, f\varphi \rangle \rightarrow \langle u, f\varphi \rangle = \langle fu, \varphi \rangle$ lorsque φ et donc $f\varphi$ appartiennent à $\mathcal{S}$, ce qui achève la démonstration.

9.4. Transformation de Fourier des distributions tempérées

A. Résultats généraux

Nous avons vu qu'une fonction sommable est une distribution tempérée. Si $f \in L^1$ et si $\varphi \in \mathcal{S}$, on a

$$\langle f, \widehat{\varphi} \rangle = \int f(x) \widehat{\varphi}(x) dx = \int f(x) \left\{ \int e^{-ix \cdot y} \varphi(y) dy \right\} dx.$$

La fonction $|f(x)\varphi(y)|$ étant sommable dans $\mathbb{R}^{2n}$, on peut appliquer le théorème de Fubini, et on obtient

$$\langle f, \widehat{\varphi} \rangle = \int \widehat{f}(y) \varphi(y) dy = \langle \widehat{f}, \varphi \rangle. \quad (9.12)$$

Nous allons utiliser le membre de gauche de la relation ci-dessus, qui a un sens même pour $f \in \mathcal{S}'$, comme définition du membre de droite. Cela nous garantit que la nouvelle définition coïncidera avec l'ancienne pour les fonctions sommables.

Théorème et Définition 9.4.1. — Soit $u \in \mathcal{S}'(\mathbb{R}^n)$. La transformée de Fourier de u est la distribution tempérée notée $\widehat{u}$ ou $\mathcal{F}u$ définie par

$$\forall \varphi \in \mathcal{S}(\mathbb{R}^n), \quad \langle \widehat{u}, \varphi \rangle = \langle u, \widehat{\varphi} \rangle.$$

Il faut montrer que la forme linéaire ainsi définie sur $\mathcal{S}$ vérifie bien une majoration du type (9.10). On a

$$|\langle u, \widehat{\varphi} \rangle| \leq C \mathcal{N}_p(\widehat{\varphi})$$

où C et p ne dépendent que de u . En utilisant le théorème 9.2.4, on obtient

$$|\langle u, \widehat{\varphi} \rangle| \leq C' \mathcal{N}_{p+n+1}(\varphi),$$

ce qui assure que $\widehat{u} \in S'$.

Théorème 9.4.2. — *La transformation de Fourier est un isomorphisme de $S'(\mathbb{R}^n)$ sur lui-même, d'inverse $\mathcal{F}^{-1} = (2\pi)^{-n} \overline{\mathcal{F}}$.*

On définit bien entendu $\overline{\mathcal{F}}u$ par la relation $\langle \overline{\mathcal{F}}u, \varphi \rangle = \langle u, \overline{\mathcal{F}}\varphi \rangle$ (on voit facilement qu'on peut aussi le définir par $\overline{\mathcal{F}}u = \overline{\mathcal{F}\overline{u}}$).

On a en effet par définition $\langle \mathcal{F}\overline{\mathcal{F}}u, \varphi \rangle = \langle \overline{\mathcal{F}}u, \mathcal{F}\varphi \rangle = \langle u, \overline{\mathcal{F}}\mathcal{F}\varphi \rangle = (2\pi)^n \langle u, \varphi \rangle$. On en déduit donc que $(2\pi)^{-n} \mathcal{F}\overline{\mathcal{F}}$, (et pour la même raison $(2\pi)^{-n} \overline{\mathcal{F}}\mathcal{F}$), coïncident avec l'identité.

Théorème 9.4.3 (Continuité). — *Si $u_j \rightarrow u$ dans S' , alors $\widehat{u}_j \rightarrow \widehat{u}$ dans S' .*

La démonstration est immédiate. Pour chaque $\varphi \in \mathcal{S}$, on a

$$\langle \widehat{u}_j, \varphi \rangle = \langle u_j, \widehat{\varphi} \rangle \rightarrow \langle u, \widehat{\varphi} \rangle = \langle \widehat{u}, \varphi \rangle.$$

Remarque 9.4.4. — La définition de la transformée de Fourier d'une distribution u ne constitue pas un moyen de calcul effectif. On dispose de formules explicites lorsque u est une fonction sommable et, comme nous le verrons ci-dessous, lorsque u est une distribution à support compact. A partir de résultats déjà connus, on peut en déduire d'autres par inversion de Fourier, par passage à la limite, et (voir plus loin) par dérivation, convolution, ...

9.4.5. Exemple : transformée de Fourier de 1. — Nous avons déjà calculé (voir n°9.1.2) les transformées de Fourier des gaussiennes, on a

$$\mathcal{F}\left(e^{-\varepsilon^2|x|^2/4}\right) = \left(\frac{4\pi}{\varepsilon^2}\right)^{n/2} e^{-|\xi|^2/\varepsilon^2}.$$

Le membre de droite peut se mettre sous la forme $(2\pi)^n \varepsilon^{-n} G(\xi/\varepsilon)$, en posant $G(z) = \pi^{-n/2} e^{-|z|^2}$. Cette dernière fonction étant d'intégrale 1, on sait que $\varepsilon^{-n} G(\xi/\varepsilon)$ converge vers δ dans $\mathcal{D}'$.

Il est facile de voir (remarque 9.3.11(d)) que les fonctions $e^{-\varepsilon^2|x|^2/4}$ convergent vers 1 dans S' pour $\varepsilon \rightarrow 0$. Il en résulte que leurs transformées de Fourier convergent vers $\mathcal{F}1$ dans S' et donc dans $\mathcal{D}'$. Or elles convergent aussi vers $(2\pi)^n \delta$. On a donc

$$\mathcal{F}1 = (2\pi)^n \delta.$$

Nous aurons l'occasion de voir des applications plus convaincantes de cette méthode. Nous verrons en effet que la formule $\mathcal{F}\delta = 1$ est une conséquence immédiate du théorème 9.4.7, et la formule d'inversion de Fourier permet d'en déduire immédiatement $\mathcal{F}1$.

9.4.6. Propriétés élémentaires. — Nous laissons au lecteur la démonstration des propriétés suivantes. Elles se démontrent d'abord pour les fonctions sommables, et en particulier pour les éléments de $\mathcal{S}$, par un simple changement de variable dans l'intégrale définissant la transformation de Fourier. Elles s'étendent ensuite aux distributions tempérées en utilisant directement la définition 9.4.1.

Conjugaison complexe et parité. — Si u possède la symétrie hermitienne (c'est-à-dire $\tilde{u} = \overline{u}$), alors $\hat{u}$ est réelle et vice versa. Si u est paire [*resp.* impaire], alors $\hat{u}$ est paire [*resp.* impaire], et on a plus généralement la formule suivante

$$\hat{\tilde{u}} = \tilde{\hat{u}},$$

qui pourrait figurer dans un chapitre célèbre d'Hippocrate. Si u est réelle et paire, il en est de même de $\hat{u}$.

Translation. — On a

$$\mathcal{F}(\tau_a u) = e^{-ia \cdot \xi} \mathcal{F}u.$$

Par inversion de Fourier, on en déduit

$$\mathcal{F}\{e^{ia \cdot x} u\} = \tau_a \hat{u}.$$

Dilatation. — En notant symboliquement $u(x/\lambda)$ la transformée de u dans la dilatation de rapport λ (voir n°5.1.4), on a

$$\mathcal{F}\{u(x/\lambda)\} = |\lambda|^n \hat{u}(\lambda \xi).$$

Il est important de remarquer qu'à la dilatation de rapport λ sur u correspond une dilatation de rapport inverse sur $\hat{u}$. Pour λ petit, $u(x/\lambda)$ est très concentrée, alors que sa transformée de Fourier est très étalée. Il s'agit de l'un des (nombreux) théorèmes correspondant au principe d'incertitude.

B. Transformation de Fourier dans $\mathcal{E}'$

Lorsqu'une distribution u est à support compact, on peut définir $\langle u, \varphi \rangle$ pour toute fonction $\varphi \in C^\infty$, et en particulier pour la fonction $x \mapsto e^{-ix \cdot \xi}$. Cela fournit un mode de calcul explicite des transformées de Fourier.

Théorème 9.4.7. — Si $u \in \mathcal{E}'(\mathbb{R}^n)$, sa transformée de Fourier appartient à $\mathcal{O}_M$ (fonctions C^∞ à croissance lente ainsi que toutes leurs dérivées), et on a

$$\hat{u}(\xi) = \left\langle u(x), e^{-ix \cdot \xi} \right\rangle. \quad (9.13)$$

Désignons par $v(\xi)$ la fonction figurant au membre de droite de (9.13). D'après le théorème de dérivation sous le crochet, c'est une fonction de classe C^∞ , et on a

$$\partial^\alpha v(\xi) = \left\langle u(x), (-ix)^\alpha e^{-ix \cdot \xi} \right\rangle.$$

Si K est un voisinage compact du support de u et si p est l'ordre de u , on a d'après (6.2)

$$|\partial^\alpha v(\xi)| \leq C \sup_{x \in K, |\beta| \leq p} \left| \partial_x^\beta (x^\alpha e^{-ix \cdot \xi}) \right|.$$

On obtient donc $|\partial^\alpha v(\xi)| \leq C_\alpha (1 + |\xi|)^p$, où C_α ne dépend que du maximum de $|x|$ sur K , ce qui prouve que $v \in \mathcal{O}_M$.

Il reste à prouver que $v = \widehat{u}$. Pour $\varphi \in C_0^\infty$, on a d'après le théorème d'intégration sous le crochet, en intégrant sur un pavé contenant le support de φ ,

$$\langle v, \varphi \rangle = \int \left\langle u(x), e^{-ix \cdot \xi} \varphi(\xi) \right\rangle d\xi = \left\langle u(x), \int e^{-ix \cdot \xi} \varphi(\xi) d\xi \right\rangle,$$

et donc

$$\langle v, \varphi \rangle = \langle u, \widehat{\varphi} \rangle = \langle \widehat{u}, \varphi \rangle,$$

ce qui achève la démonstration.

9.4.8. Applications. — Un calcul immédiat donne les transformées de Fourier de δ , de ses dérivées et de ses translatées.

$$\begin{aligned} \mathcal{F}(\delta) &= 1 \\ \mathcal{F}(\partial^\alpha \delta) &= i^{|\alpha|} \xi^\alpha \\ \mathcal{F}(\delta_a) &= e^{-ia \cdot \xi}, \end{aligned}$$

d'où l'on déduit par inversion de Fourier

$$\begin{aligned} \mathcal{F}(1) &= (2\pi)^n \delta \\ \mathcal{F}(e^{ia \cdot x}) &= (2\pi)^n \delta_a \\ \mathcal{F}(x^\alpha) &= (2\pi)^n i^{|\alpha|} \partial^\alpha \delta. \end{aligned}$$

9.4.9. Transformée de Fourier de la mesure de surface sur la sphère

Notons $d\sigma_R$ la mesure de surface de la sphère de rayon R dans $\mathbb{R}^3$. On a

$$\widehat{d\sigma_R}(\xi) = \int_{|x|=R} e^{-ix \cdot \xi} d\sigma_R.$$

En introduisant l'angle φ des vecteurs x et ξ , la fonction à intégrer ne dépend que de φ , et on est ramené à l'intégrale simple

$$\widehat{d\sigma_R}(\xi) = \int_0^\pi e^{iR|\xi|\cos\varphi} 2\pi R \sin\varphi \, d\varphi.$$

En prenant $\cos\varphi$ comme nouvelle variable, on obtient $4\pi R(\sin R|\xi|)/|\xi|$, ou encore, pour la masse unité uniformément répartie sur la sphère,

$$\mathcal{F}\left(\frac{d\sigma_R}{4\pi R^2}\right) = \frac{\sin R|\xi|}{R|\xi|}.$$

C. Transformation de Fourier dans L^2

Les fonctions de carré sommable ne sont pas en général sommables, et on ne peut pas définir leur transformée de Fourier par la formule intégrale (9.1). Par contre ce sont des distributions tempérées, et on peut donc définir leurs transformées de Fourier qui appartiennent a priori à $\mathcal{S}'$. Elles sont en fait de carré sommable et, comme pour les séries de Fourier, on obtient une isométrie d'espaces de Hilbert.

Théorème 9.4.10. — *L'application $u \mapsto (2\pi)^{-n/2}\mathcal{F}u$ est une isométrie bijective de $L^2(\mathbb{R}^n)$ sur lui-même, d'inverse $(2\pi)^{-n/2}\overline{\mathcal{F}}$.*

Montrons d'abord que $(2\pi)^{-n/2}\mathcal{F}$ est une isométrie de $\mathcal{S}$ sur lui-même lorsque celui-ci est muni de la norme de L^2 . Soient f et g appartenant à $\mathcal{S}$, et posons $h = \widehat{\overline{g}}$. On a $\int f(x)\widehat{h}(x) \, dx = \int \widehat{f}(x)h(x) \, dx$ d'après (9.12). En remarquant que

$$\mathcal{F}h = \mathcal{F}\widehat{\overline{g}} = \overline{\widehat{\mathcal{F}g}} = (2\pi)^n \overline{g},$$

on obtient immédiatement

$$(2\pi)^n \int f(x)\overline{g(x)} \, dx = \int \widehat{f}(x)\overline{\widehat{g(x)}} \, dx,$$

ce qui montre que $(2\pi)^{-n/2}\mathcal{F}$ conserve le produit scalaire, et donc la norme L^2 pour les éléments de $\mathcal{S}$.

En utilisant le fait (théorème 3.5.1) que $\mathcal{S}$, qui contient C_0^∞ , est dense dans L^2 , le théorème va en résulter aisément. Soit $u \in L^2$, et soit f_j une suite d'éléments de $\mathcal{S}$ telle que $\|u - f_j\|_{L^2}$ tende vers 0. La suite f_j est donc de Cauchy et, par isométrie, $\|f_j - f_k\|_{L^2}$ tend vers 0 lorsque j et k tendent vers l'infini. La suite de Cauchy $\widehat{f_j}$ converge donc vers un élément $g \in L^2$. La convergence dans L^2 impliquant la convergence dans $\mathcal{S}'$, on obtient d'une part que $\widehat{f_j} \rightarrow g$ dans $\mathcal{S}'$, d'autre part que $f_j \rightarrow u$ et donc (continuité de $\mathcal{F}$) que $\widehat{f_j} \rightarrow \widehat{u}$ dans $\mathcal{S}'$.

On a donc $\widehat{u} = g \in L^2$ pour un élément quelconque u de L^2 . En outre, à partir de $\|(2\pi)^{-n/2}\widehat{f}_j\|_{L^2} = \|f_j\|_{L^2}$, et par continuité de la norme, on a

$$\|(2\pi)^{-n/2}\widehat{u}\|_{L^2} = \|u\|_{L^2}.$$

Ce qui précède s'applique également à $(2\pi)^{-n/2}\overline{\mathcal{F}}$. Nous avons donc deux isométries de L^2 dans lui-même dont les composées à droite et à gauche sont égales à l'identité. Cela termine la démonstration du théorème.

Remarque 9.4.11. — Soit $u \in L^2(\mathbb{R}^n)$, et notons u_R la fonction (sommable) égale à $u(x)$ pour $|x| \leq R$ et à 0 sinon. On a $u_R \rightarrow u$ dans L^2 lorsque $R \rightarrow +\infty$, et $\widehat{u}$ est donc la limite en norme dans L^2 des fonctions

$$\widehat{u_R}(\xi) = \int_{|x| \leq R} e^{-ix \cdot \xi} u(x) dx.$$

Cette propriété est en fait la définition originale de la transformation de Fourier-Plancherel.

Le corollaire suivant donne un mode de calcul pratique en termes d'intégrales semi-convergentes.

Corollaire 9.4.12. — Soit $u \in L^2(\mathbb{R}^n)$ et supposons que, pour presque tout ξ , on ait

$$\int_{|x| \leq R} e^{-ix \cdot \xi} u(x) dx \xrightarrow{R \rightarrow \infty} g(\xi).$$

Alors, on a $\widehat{u}(\xi) = g(\xi)$ p.p.

Si R_ν tend vers l'infini avec ν , on sait que $\widehat{u_{R_\nu}}$ converge vers $\widehat{u}$ dans L^2 . Quitte à extraire une sous-suite (cf. exercice 2.6.16), on peut supposer que l'on a de plus $\widehat{u_{R_\nu}}(\xi) \rightarrow \widehat{u}(\xi)$ p.p. Cette même sous-suite converge aussi vers g presque partout, d'où le résultat.

9.5. Les propriétés fondamentales

A. L'échange de la convolution et de la multiplication

Le théorème suivant, dont il existe de nombreuses variantes, assure que la transformée de Fourier d'un produit de convolution est égal au produit ordinaire des transformées de Fourier. Comme nous le verrons, ce résultat permet de ramener la résolution des équations de convolution (et donc des équations aux dérivées partielles à coefficients constants) à des problèmes de division.

Théorème 9.5.1. — *La relation*

$$\mathcal{F}(u \star v) = (\mathcal{F}u)(\mathcal{F}v) \quad (9.14)$$

est valable dans les deux cas suivants.

(a) *Les fonctions u et v sont sommables.*

(b) *On a $u \in \mathcal{E}'$ et $v \in \mathcal{S}'$. On a alors $u \star v \in \mathcal{S}'$ et $\mathcal{F}u \in \mathcal{O}_M$ ce qui assure que les deux membres de (9.14) sont définis.*

Dans le premier cas, on a

$$\widehat{u \star v}(\xi) = \int e^{-ix \cdot \xi} \left\{ \int u(x-y)v(y) dy \right\} dx.$$

On peut appliquer le théorème de Fubini, la fonction à intégrer étant sommable dans $\mathbb{R}^{2n}$ (son module vaut $|u(x-y)||v(y)|$). En faisant le changement de variable $(x, y) \mapsto (x-y, y)$, on obtient

$$\begin{aligned} \widehat{u \star v}(\xi) &= \iint e^{-i(y+z) \cdot \xi} u(z)v(y) dy dz \\ &= \left\{ \int e^{-iz \cdot \xi} u(z) dz \right\} \left\{ \int e^{-iy \cdot \xi} v(y) dy \right\}. \end{aligned}$$

Cela achève la démonstration de la partie (a).

Considérons maintenant le cas où les deux distributions u et v sont à support compact. Les transformées de Fourier de u , v , $u \star v$ sont alors des fonctions, et on a, en notant φ_ξ la fonction $x \mapsto e^{-ix \cdot \xi}$

$$\widehat{u \star v}(\xi) = \langle u \star v, \varphi_\xi \rangle = \langle u, \tilde{v} \star \varphi_\xi \rangle. \quad (9.15)$$

D'autre part, on a

$$(\tilde{v} \star \varphi_\xi)(x) = \langle \tilde{v}(y), \varphi_\xi(x-y) \rangle = \langle v(y), \varphi_\xi(x+y) \rangle,$$

et donc

$$(\tilde{v} \star \varphi_\xi)(x) = \langle v(y), e^{-i(x+y) \cdot \xi} \rangle = \widehat{v}(\xi) e^{-ix \cdot \xi}.$$

En reportant dans le membre de droite de (9.15), on obtient

$$\widehat{u \star v}(\xi) = \langle u(x), \widehat{v}(\xi) e^{-ix \cdot \xi} \rangle = \widehat{v}(\xi) \widehat{u}(\xi).$$

Pour terminer la démonstration de la partie (b) du théorème, nous aurons besoin du lemme suivant.

Lemme 9.5.2. — *Pour tout $u \in \mathcal{S}'$, il existe une suite u_j d'éléments de $\mathcal{E}'$ vérifiant $u_j \rightarrow u$ dans $\mathcal{S}'$.*

Nous avons montré (théorème 9.2.9) que, si ψ est une fonction d'essai égale à 1 au voisinage de l'origine, on a

$$\forall p, \quad \lim_{j \rightarrow \infty} \mathcal{N}_p(\varphi(x) - \psi(x/j)\varphi(x)) = 0$$

pour toute fonction $\varphi \in \mathcal{S}$. Si maintenant u appartient à $\mathcal{S}'$, et si on pose $u_j = \psi(x/j)u$, on a pour $\varphi \in \mathcal{S}$

$$|\langle u - u_j, \varphi \rangle| = |\langle u, (\varphi(x) - \psi(x/j)\varphi(x)) \rangle| \leq C \mathcal{N}_p(\varphi(x) - \psi(x/j)\varphi(x)),$$

où C et p ne dépendent que de u . On a donc $\langle u - u_j, \varphi \rangle \rightarrow 0$ pour tout $\varphi \in \mathcal{S}$ ce qui achève la démonstration du lemme.

9.5.3. Fin de la démonstration du théorème 9.5.1(b). — Soient donc $u \in \mathcal{E}'$, $v \in \mathcal{S}'$ et une suite v_j d'éléments de $\mathcal{E}'$ convergeant vers v dans $\mathcal{S}'$. Nous avons vu que $\mathcal{F}(u \star v_j) = \widehat{u} \widehat{v}_j$. Lorsque j tend vers l'infini, on a $\widehat{v}_j \rightarrow \widehat{v}$ dans $\mathcal{S}'$, et (théorème 9.4.7) la fonction $\widehat{u}$ appartenant à $\mathcal{O}_M$, on a $\widehat{u} \widehat{v}_j \rightarrow \widehat{u} \widehat{v}$ dans $\mathcal{S}'$ d'après le théorème 9.3.15.

Nous avons ainsi montré que $\mathcal{F}(u \star v_j) \rightarrow \widehat{u} \widehat{v}$ dans $\mathcal{S}'$, et donc, par continuité de $\mathcal{F}^{-1}$ que $u \star v_j$ converge dans $\mathcal{S}'$ vers une distribution tempérée w vérifiant $\widehat{w} = \widehat{u} \widehat{v}$. La convergence dans $\mathcal{S}'$ entraînant la convergence dans $\mathcal{D}'$, il en résulte que $u \star v_j \rightarrow w$ dans $\mathcal{D}'$. Mais d'autre part, d'après le théorème 7.4.9, on a $u \star v_j \rightarrow u \star v$ dans $\mathcal{D}'$. On a donc $u \star v = w \in \mathcal{S}'$ et $\mathcal{F}(u \star v) = \widehat{u} \widehat{v}$, ce qui achève la démonstration.

Exercice 9.5.4. — Démontrer directement que, pour $u \in \mathcal{E}'$ et $\varphi \in \mathcal{S}$, on a $u \star \varphi \in \mathcal{S}$, et que pour chaque p , il existe q et C tels que l'on ait $\mathcal{N}_p(u \star \varphi) \leq C \mathcal{N}_q(\varphi)$. Retrouver ainsi le fait que, pour $v \in \mathcal{S}'$, on a $u \star v \in \mathcal{S}'$.

Exercice 9.5.5. — Démontrer que, pour φ et ψ appartenant à $\mathcal{S}$, on a $\varphi \star \psi \in \mathcal{S}$. Montrer que, pour $u \in \mathcal{S}'$ et $\varphi \in \mathcal{S}$, la fonction $x \mapsto \langle u(y), \varphi(x - y) \rangle$, que l'on notera encore $u \star \varphi$ appartient à $\mathcal{S}'$. Montrer que l'on a encore $\mathcal{F}(u \star \varphi) = \widehat{\varphi} \widehat{u}$ sous ces hypothèses.

Remarque 9.5.6. — Plus généralement, L. Schwartz a introduit un sous-espace $\mathcal{O}'_C$ de $\mathcal{S}'$, constitué des distributions qui sont "à décroissance rapide" à l'infini. La définition précise est que $u \in \mathcal{O}'_C$ si toutes ses régularisées $u \star \varphi$ avec $\varphi \in C_0^\infty$ appartiennent à $\mathcal{S}$. Il est clair que cet espace contient $\mathcal{E}'$ et $\mathcal{S}$.

L'espace $\mathcal{O}'_C$ opère sur $\mathcal{S}'$ par convolution (de même que $\mathcal{O}_M$ opère par multiplication, cela justifie les indices) c'est-à-dire que l'on peut définir de manière raisonnable le produit de convolution d'un élément de $\mathcal{O}'_C$ et d'un élément de $\mathcal{S}'$. La transformation de Fourier est une bijection de $\mathcal{O}_M$ sur $\mathcal{O}'_C$ et on a, pour $u \in \mathcal{O}'_C$ et $v \in \mathcal{S}'$, l'identité $\mathcal{F}(u \star v) = \widehat{u} \widehat{v}$.

Au vu des exercices et de la remarque qui précèdent, le lecteur souhaiterait sans doute disposer d'un critère général permettant de savoir dans quels cas le produit de convolution est défini. La remarque 9.5.14 à la fin de cette

section lui montrera qu'un tel espoir est vain. Dans la pratique, chaque fois que l'on peut donner un sens raisonnable à l'un des membres de (9.14), on peut donner un sens à l'autre membre et l'égalité a lieu, mais cela nécessite une démonstration dans chaque cas.

Les dérivations et les translations sont des cas particuliers de convolution, et on obtient les résultats suivants qui généralisent les formules obtenues pour les fonctions dans la section 9.2.

Corollaire 9.5.7. — (a) Pour $u \in \mathcal{S}'$, on a

$$\mathcal{F}(\partial^\alpha u) = i^{|\alpha|} \xi^\alpha \mathcal{F}u,$$

$$\mathcal{F}(\tau_a u) = e^{-ia \cdot \xi} \mathcal{F}u,$$

$$\mathcal{F}(x^\alpha u) = i^{|\alpha|} \partial^\alpha \mathcal{F}u.$$

(b) Si $u \in \mathcal{S}'$ et si $\varphi \in \mathcal{O}_M$ est telle que sa transformée de Fourier est à support compact, on a

$$\mathcal{F}(\varphi u) = (2\pi)^{-n} \widehat{\varphi} \star \widehat{u}.$$

Les deux premières relations sont des conséquences simples de (9.14), compte tenu des expressions des transformées de Fourier de $\partial^\alpha \delta$ et δ_a obtenues au n°9.4.8. Quant à la troisième relation et à la partie (b), elles résultent de la formule d'inversion de Fourier. La restriction sur le support de $\widehat{\varphi}$ pourrait être levée en utilisant l'espace $\mathcal{O}'_C$.

Exercice 9.5.8. — Déterminer la transformée de Fourier de $\text{vp}(1/x)$. On pourra remarquer que $\text{vp}(1/x)$ est l'unique distribution u qui soit impaire et qui vérifie $xu = 1$.

B. Équations de convolution

Soient $A \in \mathcal{E}'$ et $f \in \mathcal{S}'$. La transformation de Fourier donne une méthode simple pour déterminer les solutions *tempérées* de l'équation

$$A \star u = f. \quad (9.16)$$

Ce sont les $u \in \mathcal{S}'$ telles que l'on ait $\widehat{A}(\xi)\widehat{u}(\xi) = \widehat{f}(\xi)$. Cette dernière équation est immédiate à résoudre lorsque la fonction $\widehat{A}$ ne s'annule pas (il suffit de multiplier $\widehat{f}$ par la fonction $1/\widehat{A}(\xi)$ qui est de classe C^∞). Dans le cas contraire, se pose le problème de la division d'une distribution par une fonction C^∞ , qui amène souvent à définir des parties finies plus ou moins complexes.

Il faut remarquer que cette méthode ne dit rien sur les éventuelles solutions de (9.16) dans $\mathcal{D}'$ qui ne sont pas tempérées.

9.5.9. Équation de Laplace. — Cherchons les distributions tempérées u solutions de $\Delta u = 0$ dans $\mathbb{R}^n$. On a $\mathcal{F}(\Delta \delta) = -|\xi|^2$, et la transformée de Fourier de u doit donc vérifier $-|\xi|^2 \hat{u} = 0$.

Il en résulte que la restriction de $\hat{u}$ au complémentaire de l'origine est égale à 0. La distribution $\hat{u}$ dont le support est l'origine est donc une somme de dérivées de la masse de Dirac à l'origine, ce qui signifie que u est un polynôme. Nous avons donc obtenu le résultat suivant.

Les seules distributions harmoniques tempérées sont les polynômes harmoniques. En d'autres termes, toute fonction harmonique majorée par un polynôme est un polynôme harmonique.

En particulier, on retrouve le fait qu'une fonction harmonique qui tend vers 0 à l'infini est nulle. On obtient aussi qu'une fonction harmonique bornée (ou même $o(|x|)$ à l'infini) est une constante, etc.

Nous avons déjà montré que toutes les distributions harmoniques sont des fonctions C^∞ . Il en existe qui ne sont pas tempérées, par exemple $e^x \cos y$ (ou la partie réelle de n'importe quelle fonction holomorphe) dans le plan.

9.5.10. Équation de Poisson. — Pour trouver les solutions tempérées de $\Delta u = f$ avec $f \in \mathcal{S}'$, on doit résoudre l'équation $|\xi|^2 \hat{u} = -\hat{f}$. On peut montrer que c'est toujours possible. Nous nous limiterons ici au cas où $\hat{f}$ est une fonction, bornée dans un voisinage de l'origine, en dimension $n \geq 3$.

Dans ce cas, la fonction $g(\xi) = -\hat{f}(\xi)/|\xi|^2$ est localement sommable, et il n'est pas difficile de montrer (Exercice : le faire) qu'elle est tempérée. On obtient donc une solution en posant $u = \mathcal{F}^{-1}g$.

L'hypothèse contient le cas $f \in \mathcal{E}'$ que nous avons déjà résolu dans la section 8.2, mais aussi les cas $f \in L^1$ et $f \in \mathcal{O}'_C$.

Exercice 9.5.11. — Calculer la transformée de Fourier inverse de la fonction $1/|\xi|^2$ dans $\mathbb{R}^3$, et retrouver ainsi la solution élémentaire du Laplacien considérée dans la section 8.2.

9.5.12. Équation $(\Delta - \lambda)u = f$. — Soit $\lambda > 0$ et cherchons les solutions tempérées de cette équation. On a à résoudre $(-|\xi|^2 - \lambda)\hat{u} = \hat{f}$. En remarquant que $(|\xi|^2 + \lambda)^{-1} \in \mathcal{O}_M$, on voit que l'équation possède une et une seule solution tempérée. On a donc le résultat suivant.

Pour chaque $f \in \mathcal{S}'$, il existe une et une seule solution tempérée u de l'équation $(\Delta - \lambda)u = f$, et celle-ci est donnée par

$$u = (2\pi)^{-n} \overline{\mathcal{F}} \left\{ \frac{-1}{|\xi|^2 + \lambda} \mathcal{F}f \right\}.$$

Lorsque $f \in \mathcal{S}$, il en résulte facilement que $u \in \mathcal{S}$. Nous verrons que les espaces de Sobolev nous permettront de décrire plus finement la régularité de la solution u en fonction de celle du second membre f .

Exercice 9.5.13. — Démontrer que la solution élémentaire de $(\Delta - \lambda)$ ainsi construite est de classe C^∞ en dehors de l'origine. En déduire qu'une distribution u solution de $(\Delta - \lambda)u = 0$ dans un ouvert est nécessairement une fonction C^∞ (voir le corollaire 8.2.3 et la remarque qui le précède).

Remarque 9.5.14 (Considérations sur le prolongement des applications bilinéaires)

Lorsqu'une application linéaire ($\mathcal{F}$ par exemple) est définie sur un espace (L^1 par exemple) que l'on estime trop petit, on peut chercher à obtenir un prolongement par continuité. Il s'agit de trouver un espace $\mathcal{A}$ muni d'une notion de convergence, dans lequel L^1 soit dense, et un prolongement continu (nécessairement unique) de $\mathcal{F}$ comme application de $\mathcal{A}$ dans un espace raisonnable (les distributions par exemple). Même si, posé ainsi, le problème n'a pas toujours de solution optimale (avec $\mathcal{A}$ le plus grand possible), on peut lui apporter des solutions très satisfaisantes, l'extension de $\mathcal{F}$ à $\mathcal{S}'$ étant l'un des meilleurs exemples.

La situation est très différente pour les applications bilinéaires (penser au produit ordinaire, au produit de convolution, à la dualité). On cherche cette fois des couples d'espaces $\mathcal{A}, \mathcal{B}$ tels que l'application bilinéaire se prolonge continûment à $\mathcal{A} \times \mathcal{B}$, mais il est clair que l'on ne peut pas espérer "une" bonne solution. Plus on sera exigeant pour l'appartenance à $\mathcal{A}$, plus on pourra être laxiste pour l'appartenance à $\mathcal{B}$, et on peut avoir une grande quantité de théorèmes intéressants correspondant à ces phénomènes de compensation.

On ne peut pas définir en général le produit de deux distributions. L'obstacle n'est pas la croissance à l'infini (on peut multiplier deux fonctions continues quelles que soient leurs croissances), mais la régularité. Plus l'un des facteurs est régulier, plus on peut admettre de singularité pour l'autre. Par exemple, on peut définir le produit d'une fonction C^∞ et d'une distribution quelconque ; d'une fonction de classe C^m et d'une distribution d'ordre m ; d'un élément de H^m et d'un élément de H^{-m} ; d'un élément de L_{loc}^∞ et d'un élément de L_{loc}^1 ; de deux éléments de L_{loc}^2 , ... La liste est loin d'être limitative, et nous évoquerons des extensions de nature très différentes dans la section 9.7.

Pour le produit de convolution, il n'y a aucun problème de régularité (il est défini pour deux distributions à support compact). Il s'agit d'un problème de décroissance à l'infini, et plus l'un des facteurs est petit à l'infini, plus l'autre peut être grand. Par exemple, on peut définir le produit de convolution d'un élément de $\mathcal{E}'$ (ce qui se fait de plus petit à l'infini) avec une distribution quelconque ; d'un élément de $\mathcal{O}'_C$ avec un élément de $\mathcal{S}'$; d'une fonction localement sommable $\mathcal{O}(|x|^{-p})$ à l'infini par une fonction $\mathcal{O}(|x|^{p-n-1})$; d'un élément de L^1 par un élément de L^∞ ; de deux éléments de L^2 , ... Le mécanisme de compensation peut être même plus subtil, puisqu'une grande croissance de l'un des facteurs dans certaines directions peut être compensée par la nullité (ou une grande décroissance) de l'autre dans les directions opposées : c'est la signification des hypothèses de convolutivité du chapitre 7.

En ce qui concerne la dualité, les deux phénomènes se superposent. Il doit y avoir compensation des régularités et des comportements à l'infini. Les deux cas extrêmes sont les définitions de $\langle u, \varphi \rangle$ pour $u \in \mathcal{D}'$ et $\varphi \in C_0^\infty$ d'une part, et pour $u \in \mathcal{E}'$ et $\varphi \in C^\infty$ d'autre part. Nous avons également rencontré le cas $\varphi \in H^m$ et $u \in H^{-m}$, le cas $\varphi \in \mathcal{S}$ et $u \in \mathcal{S}'$ et le cas du n° 6.1.8. Le lecteur pourra en imaginer bien d'autres.

On aura remarqué, bien sûr, la cohérence entre le fait que la transformation de Fourier échange la régularité et la décroissance à l'infini, le fait qu'elle échange multiplication et convolution, et le fait qu'elle transforme la dualité en elle-même (aux conjugués complexes et au $(2\pi)^{-n}$ près).

9.6. Transformation de Fourier partielle et équations d'évolution

Dans l'espace-temps $\mathbb{R}^{n+1}$, on veut souvent faire jouer à la variable de temps un rôle spécifique (rester local en t , distinguer passé et avenir, ...) qu'une transformation de Fourier globale prendrait mal en compte. Il est souvent préférable de faire une transformation de Fourier en les variables d'espace "à t fixé" (ce qui a un sens clair pour des fonctions, mais doit être défini pour des distributions). Seules les dérivations en x sont alors transformées en multiplications, et l'étude des équations aux dérivées partielles à coefficients constants est ramenée à l'étude d'équations différentielles en t dépendant du paramètre ξ .

Nous nous limiterons à ce cadre, mais il est bien sûr possible de faire une transformation de Fourier partielle en x dans $\mathbb{R}_x^p \times \mathbb{R}_y^q$.

Théorème et Définition 9.6.1. — Soit $\varphi \in \mathcal{S}(\mathbb{R}^{n+1})$. On définit sa transformée de Fourier partielle, notée $\mathcal{F}'\varphi$ ou $\tilde{\varphi}$, par

$$\tilde{\varphi}(t, \xi) = \int e^{-ix \cdot \xi} \varphi(t, x) dx.$$

L'opérateur $\mathcal{F}'$ applique bijectivement $\mathcal{S}(\mathbb{R}^{n+1})$ dans lui-même, et son inverse est $(2\pi)^{-n} \overline{\mathcal{F}'}$.

Nous laissons au lecteur le soin de montrer, en reprenant les arguments du théorème 9.2.4, que $\mathcal{F}'$ applique $\mathcal{S}$ dans lui-même, et de démontrer que l'on a

$$\mathcal{N}_p(\tilde{\varphi}) \leq C_p \mathcal{N}_{p+n+1}(\varphi). \quad (9.17)$$

Quant à la démonstration de la formule d'inversion, il suffit d'appliquer pour chaque valeur de t le théorème 9.1.5.

Théorème et Définition 9.6.2. — Soit $u \in \mathcal{S}'(\mathbb{R}^{n+1})$. La transformée de Fourier partielle de u est la distribution tempérée, notée $\mathcal{F}'u$ ou $\tilde{u}$, définie par

$$\forall \varphi \in \mathcal{S}(\mathbb{R}^{n+1}), \quad \langle \tilde{u}, \varphi \rangle = \langle u, \tilde{\varphi} \rangle.$$

L'opérateur $\mathcal{F}'$ est un isomorphisme de $\mathcal{S}'(\mathbb{R}^{n+1})$ sur lui-même d'inverse $(2\pi)^{-n}\overline{\mathcal{F}'}$. En outre, si $u_j \rightarrow u$ dans $\mathcal{S}'$, on a $\tilde{u}_j \rightarrow \tilde{u}$ dans $\mathcal{S}'$.

Le fait que $\tilde{u}$ appartienne à $\mathcal{S}'$ résulte simplement de (9.17). Les démonstrations des autres propriétés ne font que reprendre mot pour mot celles que nous avons vues dans la sous-section 9.4.A.

On obtient de même facilement les formules suivantes

$$\begin{aligned}\mathcal{F}'(\partial_x^\alpha u) &= i^{|\alpha|} \xi^\alpha \mathcal{F}'u \\ \mathcal{F}'(\partial_t^k u) &= \partial_t^k (\mathcal{F}'u) \\ \mathcal{F}'(x^\alpha u) &= i^{|\alpha|} \partial_\xi^\alpha (\mathcal{F}'u) \\ \mathcal{F}'(t^k u) &= t^k \mathcal{F}'u.\end{aligned}$$

Remarque 9.6.3. — Si I est un intervalle ouvert de $\mathbb{R}$, on peut définir l'espace des distributions partiellement tempérées en x sur $I \times \mathbb{R}^n$ de la manière suivante : c'est l'espace des $u \in \mathcal{D}'(I \times \mathbb{R}^n)$ telles que, pour chaque fonction $\psi(t) \in C_0^\infty(I)$, on ait $\psi(t)u \in \mathcal{S}'(\mathbb{R}^{n+1})$.

Pour une telle distribution u , on voit facilement que, si ψ est égale à 1 dans un intervalle $]a, b[$, la restriction de $\widetilde{\psi u}$ à $]a, b[\times \mathbb{R}^n$ ne dépend pas de ψ . Cela permet de définir une distribution notée encore $\tilde{u}$ partiellement tempérée en ξ dans $I \times \mathbb{R}^n$, caractérisée par le fait que, pour tout $\psi \in C_0^\infty(I)$, on ait $\mathcal{F}'(\psi u) = \psi \tilde{u}$.

9.6.4. Transformée de Fourier partielle de δ . — On doit avoir par définition

$$\langle \tilde{\delta}, \varphi \rangle = \langle \delta, \tilde{\varphi} \rangle = \tilde{\varphi}(0, 0)$$

et donc

$$\langle \tilde{\delta}, \varphi \rangle = \int \varphi(0, x) dx.$$

La transformée de Fourier partielle de δ est donc la distribution de simple couche de densité 1 portée par l'hyperplan $t = 0$.

9.6.5. Solution élémentaire de l'équation des ondes. — En dimension d'espace $n = 3$, il s'agit de déterminer une distribution E nulle dans le passé vérifiant $\square E = \delta$. En espérant que E soit tempérée, il est équivalent de demander que $\tilde{E}$ soit nulle dans le passé et vérifie

$$(\partial_t^2 + |\xi|^2)\tilde{E} = \tilde{\delta}. \quad (9.18)$$

En dehors de $t = 0$, on doit donc avoir $(\partial_t^2 + |\xi|^2)\tilde{E} = 0$, équation différentielle dont on connaît bien les solutions pour chaque ξ fixé. Compte tenu

de la condition de nullité dans le passé, il est donc raisonnable de chercher $\tilde{E}$ sous la forme d'une fonction du type

$$\tilde{E}(t, \xi) = \begin{cases} 0 & \text{pour } t < 0 \\ a(\xi) \cos(|\xi| t) + b(\xi) \sin(|\xi| t) & \text{pour } t > 0 \end{cases},$$

où les fonctions a et b sont à déterminer pour que (9.18) ait lieu. En appliquant deux fois la formule des sauts pour calculer $\partial_t^2 \tilde{E}$, on obtient

$$(\partial_t^2 + |\xi|^2) \tilde{E} = \partial_t \mu + \nu,$$

où μ et ν sont les distributions de simple couche portées par $t = 0$, de densité respectives $a(\xi)$ et $|\xi| b(\xi)$.

En choisissant $a = 0$ et $b = 1/|\xi|$, nous avons donc obtenu une fonction $\tilde{E}$, égale à 0 pour $t < 0$ et à $\sin(t|\xi|)/|\xi|$ pour $t > 0$, qui est solution de l'équation (9.18) et qui est une distribution tempérée (on a en effet $|\tilde{E}(t, \xi)| \leq \max(t, 0)$). On sait donc que $(\mathcal{F}')^{-1} \tilde{E}$ sera une solution élémentaire, nulle pour $t < 0$, et il ne reste plus qu'à l'expliciter.

Cette explicitation sera facilitée par le fait que nous avons calculé, au n°9.4.9 la transformée de Fourier de la mesure de surface $d\sigma_R$ de la sphère de rayon R , qui est la fonction $4\pi R(\sin R|\xi|)/|\xi|$. Pour $\varphi \in \mathcal{S}$, on doit avoir

$$\langle E, \tilde{\varphi} \rangle = \left\langle \tilde{E}, \varphi \right\rangle = \int_0^\infty dt \int \varphi(t, \xi) \frac{\sin(t|\xi|)}{|\xi|} d\xi.$$

Nous avons appliqué le théorème de Fubini, la fonction à intégrer étant majorée en module par $|t\varphi(t, \xi)|$ sommable dans $\mathbb{R}^4$. En utilisant la transformée de Fourier dans $\mathbb{R}^3$ pour chaque t fixé non nul, on a donc

$$\langle E, \tilde{\varphi} \rangle = \int_0^\infty \left\langle \frac{d\sigma_t(x)}{4\pi t}, \tilde{\varphi}(t, x) \right\rangle dt.$$

En posant $\tilde{\varphi} = \psi$, et en utilisant les coordonnées polaires $x = r\theta$, $r > 0$, $\theta \in \mathbb{S}^2$, on obtient

$$\langle E, \psi \rangle = \int_0^\infty \frac{dt}{4\pi t} \int_{r=t} \psi(t, r\theta) r^2 d\sigma_\theta = \iint_{[0, \infty] \times \mathbb{S}^2} \frac{\psi(r, r\theta)}{4\pi r} r^2 dr d\sigma_\theta.$$

On reconnaît là la mesure de Radon $d\sigma/4\pi\rho$ définie par (8.5) et l'expression de la solution élémentaire de $\square$ que nous avons admise jusqu'ici.

Remarque 9.6.6. — Une partie de l'analyse ci-dessus est valable en toute dimension : il existe une solution élémentaire tempérée à support dans $t \geq 0$ donnée par $E = (\mathcal{F}')^{-1}(H(t)\sin(t|\xi|)/|\xi|)$. D'autre part, l'argument du théorème 8.3.2 montre que c'est l'unique solution élémentaire nulle dans le passé.

Par contre la forme explicite de E dépend de la dimension. Pour $n = 2$, on obtient l'expression de l'exercice n°8.3.8. Pour $n \geq 3$ impair, on obtient une couche $\frac{n-1}{2}$ -multiple portée par la surface du cône d'onde, tandis que, pour n pair, on obtient (à une constante près) une partie finie, qui doit être précisée, de $H(t-r)(t^2-r^2)^{-(n-1)/2}$.

9.6.7. Solutions élémentaires des équations de la chaleur et de Schrödinger

Les exercices qui vont suivre permettent d'obtenir, par le même procédé, l'existence de solutions élémentaires et la forme précise de leurs transformées de Fourier. L'écriture de la solution élémentaire elle-même, qui est un peu délicate pour l'équation de Schrödinger, sera donnée ensuite.

Exercice 9.6.8. — Démontrer que l'opérateur de la chaleur

$$\partial_t - \Delta$$

dans l'espace-temps $\mathbb{R}^{n+1}$ possède une unique solution élémentaire E qui soit tempérée et à support dans $t \geq 0$, et que l'on a $E = (\mathcal{F}')^{-1} (H(t)e^{-t|\xi|^2})$.

Exercice 9.6.9. — Démontrer que l'opérateur de Schrödinger

$$i\partial_t + \Delta$$

dans l'espace-temps $\mathbb{R}^{n+1}$ possède une unique solution élémentaire F qui soit tempérée et à support dans $t \geq 0$, et que l'on a $F = (\mathcal{F}')^{-1} (H(t)e^{-it|\xi|^2})$.

Exercice 9.6.10. — Montrer que sur $\mathbb{R}$ la fonction $u(x) = e^{ix^2}$ vérifie $u' = 2ixu$, et que réciproquement, les seules distributions solutions de cette équation différentielle sont les fonctions Ce^{ix^2} .

En déduire qu'il existe une constante C_1 telle que la transformée de Fourier de e^{ix^2} soit égale à $C_1 e^{-i\xi^2/4}$, et que l'on a en dimension n

$$\mathcal{F} \left(e^{i|x|^2} \right) = C_n e^{-i|\xi|^2/4}$$

avec $C_n = (C_1)^n$. On admettra que l'on a $C_n = e^{in\pi/4} \pi^{n/2}$.

D'après le premier de ces exercices, la solution élémentaire E de l'équation de la chaleur vérifie donc, pour toute fonction $\varphi(t, x) \in \mathcal{S}$,

$$\langle E, \tilde{\varphi} \rangle = \left\langle \tilde{E}, \varphi \right\rangle = \int_0^\infty dt \int \varphi(t, \xi) e^{-t|\xi|^2} d\xi.$$

En utilisant, pour chaque $t > 0$ fixé, le fait que nous connaissons bien les transformées de Fourier des gaussiennes dans $\mathbb{R}^n$, on obtient

$$\langle E, \tilde{\varphi} \rangle = \int_0^\infty dt \left\{ (4\pi t)^{-n/2} \int e^{-|x|^2/4t} \tilde{\varphi}(t, x) dx \right\}. \quad (9.19)$$

Il est facile de vérifier que la fonction $H(t)(4\pi t)^{-n/2} e^{-|x|^2/4t}$ est localement sommable, et qu'elle devient sommable après multiplication par un élément

de $\mathcal{S}$. On peut donc appliquer le théorème de Fubini et il en résulte qu'une solution élémentaire de l'équation de la chaleur est la fonction

$$E(t, x) = H(t)(4\pi t)^{-n/2} e^{-|x|^2/4t}.$$

Pour l'équation de Schrödinger, on peut répéter mot pour mot l'argument précédent jusqu'à l'obtention de l'équivalent de (9.19). Il n'y a aucune difficulté à utiliser le théorème de Fubini dans l'espace des (t, ξ) , et l'exercice n°9.6.10 nous donne les valeurs des transformées de Fourier des gaussiennes imaginaires. On obtient

$$\langle F, \tilde{\varphi} \rangle = \int_0^\infty dt \left\{ (4\pi t)^{-n/2} e^{-in\pi/4} \int e^{i|x|^2/4t} \tilde{\varphi}(t, x) dx \right\}.$$

La fonction $H(t)(4\pi t)^{-n/2} e^{-i|x|^2/4t}$, dont le module vaut $H(t)(4\pi t)^{-n/2}$, n'est pas localement sommable (sauf pour $n = 1$) dans $\mathbb{R}^{n+1}$, et il est exclu d'appliquer le théorème de Fubini. Ce que l'on a démontré doit être énoncé avec précision : *l'équation de Schrödinger possède une solution élémentaire tempérée à support dans $t \geq 0$. Il s'agit de la distribution F définie, pour tout $\psi \in \mathcal{S}(\mathbb{R}^{n+1})$, par*

$$\langle F, \psi \rangle = e^{-in\pi/4} \int_0^\infty \frac{dt}{(4\pi t)^{n/2}} \left\{ \int e^{i|x|^2/4t} \psi(t, x) dx \right\},$$

ce qui précède garantissant la convergence de cette succession d'intégrales (Exercice : le prouver directement). Cette distribution n'est pas une fonction (c'est si on veut une sorte de partie finie ou de valeur principale). En particulier, ce n'est pas une distribution d'ordre 0.

9.6.11. Résolution du problème de Cauchy. — Cherchons par exemple une solution partiellement tempérée de l'équation des ondes vérifiant $u(0, x) = h_0(x)$ et $\partial_t u(0, x) = h_1(x)$, où les fonctions h_j sont données et appartiennent à $\mathcal{S}$. La transformée de Fourier partielle doit donc vérifier $(\partial_t^2 + |\xi|^2)\tilde{u} = 0$ avec $\tilde{u}(0, \xi) = \widehat{h_0}(\xi)$ et $\partial_t \tilde{u}(0, \xi) = \widehat{h_1}(\xi)$.

On cherche $\tilde{u}$ sous la forme

$$\tilde{u}(t, \xi) = a(\xi) \cos(|\xi| t) + b(\xi) \sin(|\xi| t),$$

les deux conditions pour $t = 0$ déterminent a et b de façon unique, et on a

$$\tilde{u}(t, \xi) = \widehat{h_0}(\xi) \cos(|\xi| t) + \widehat{h_1}(\xi) \sin(|\xi| t) / |\xi|.$$

On vérifie facilement que la fonction $\tilde{u}$ ainsi obtenue est tempérée, et il suffit de prendre la transformée de Fourier partielle inverse pour obtenir le résultat. Bien entendu, la formule générale donnant u en fonction des h_j sera la même que celle que nous avons obtenue dans la section 8.3. Cela dit, il peut

arriver que pour des valeurs particulières des h_j , le calcul explicite soit très simplifié. D'autre part, pour des h_j tempérées n'appartenant pas à $\mathcal{S}$ mais dont les transformées de Fourier sont des fonctions, on peut plus facilement prouver que la formule ci-dessus fournit effectivement une solution.

Le lecteur pourra montrer que la même méthode permet de trouver des solutions u de l'équation de la chaleur (pour $t \geq 0$ mais pas pour $t < 0$ en général) ou de l'équation de Schrödinger (pour t quelconque) vérifiant $u(0, x) = h(x)$ avec $h \in \mathcal{S}$. On obtient

$$\tilde{u}(t, \xi) = \widehat{h}(\xi) e^{-t|\xi|^2}$$

pour l'équation de la chaleur, et

$$\tilde{u}(t, \xi) = \widehat{h}(\xi) e^{-it|\xi|^2}$$

pour l'équation de Schrödinger.

9.7. Vers l'analyse microlocale

La transformation de Fourier permet de lire la régularité d'une distribution u sur la décroissance à l'infini de $\widehat{u}$, mais il s'agit d'une analyse globale, qui ne permet pas de distinguer entre une u singulière en un point, et une u singulière partout. En outre, on ne peut pas parler de transformée de Fourier pour une distribution définie seulement dans un ouvert.

Une autre rigidité de la transformation de Fourier est sa relation étroite avec la structure vectorielle de $\mathbb{R}^n$. Si elle fournit immédiatement des résultats précis sur les équations aux dérivées partielles à coefficients constants, elle n'est pas directement efficace lorsque les coefficients sont variables. Si on a $\sum a_\alpha(x) \partial^\alpha u = 0$, dans l'équation correspondante pour $\widehat{u}$ les dérivations seront bien transformées en multiplications, mais les multiplications par les a_α seront transformées en des dérivations (si les a_α sont des polynômes) ou en des convolutions, et le problème n'est pas simplifié.

C'est à partir de 1970 que sont apparus des concepts mathématiques permettant de faire une analyse locale à la fois dans l'espace ambiant (de la variable x) et dans l'espace des fréquences (de la variable ξ). Nous ne pourrions ici que décrire l'un de ceux-ci, en essayant de donner une idée de l'efficacité de cette branche récente, connue maintenant sous le nom d'analyse microlocale.

Une notion déjà ancienne est celle de *support singulier* d'une distribution. Si $u \in \mathcal{D}'(\Omega)$, on montre comme dans le théorème 6.1.1 qu'il existe un plus grand sous-ouvert ω de Ω tel que la restriction de u à ω soit de classe C^∞ . Le complémentaire de ω se note $\text{Suppsing}(u)$.

La propriété $x_0 \notin \text{Suppsing}(u)$ est équivalente à la propriété suivante : il existe un voisinage ω de x_0 tel que, pour tout $\varphi \in C_0^\infty(\omega)$, la fonction $\mathcal{F}(\varphi u)$ soit à décroissance rapide (c'est-à-dire $\widehat{\varphi u}(\xi) = O(|\xi|^{-N})$ pour tout N). En effet, pour tout α , la fonction $\partial^\alpha(\varphi u)$ a alors une transformée de Fourier sommable et est donc continue.

La définition suivante, due ainsi que les résultats qui vont suivre à L. Hörmander, associe aux singularités de u un objet géométrique plus riche : un ensemble fermé de l'espace des phases (espace des couples (x, ξ)).

Définition 9.7.1. — On dit que u est de classe C^∞ microlocalement au voisinage de $(x_0, \xi_0) \in \Omega \times (\mathbb{R}^n \setminus \{0\})$ s'il existe un voisinage ω de x_0 et un voisinage conique Γ de ξ_0 tel que, pour $\varphi \in C_0^\infty(\omega)$, la fonction $\mathcal{F}(\varphi u)$ soit à décroissance rapide dans Γ .

On appelle front d'onde de $u^{(1)}$, et on note $\text{WF}(u)$ le complémentaire de l'ensemble des points (x_0, ξ_0) au voisinage desquels u est microlocalement de classe C^∞ . C'est un sous-ensemble fermé de $\Omega \times (\mathbb{R}^n \setminus \{0\})$ dont la projection sur l'espace des x est $\text{Supp} \sin g(u)$.

Cet ensemble contient des renseignements non seulement sur les endroits où u est singulière, mais sur les directions (qu'il faut considérer comme des vecteurs covariants) de ces singularités. Par exemple, si u est une distribution de simple couche portée par $x_n = 0$ de densité $h(x') \in C^\infty(\mathbb{R}^{n-1})$, son front d'onde sera constitué des $(x', x_n; \xi', \xi_n)$ vérifiant $x' \in \text{Supp}(h)$, $x_n = 0$, $\xi' = 0$, $\xi_n \neq 0$. Plus généralement, le front d'onde d'une distribution de simple couche, ou de couche multiple, de densité C^∞ portée par une hypersurface Σ sera contenu dans l'ensemble des (x, ξ) avec $x \in \Sigma$ et ξ orthogonal à Σ en x .

Une propriété remarquable est que l'on peut définir de manière raisonnable le produit de deux distributions u et v sous l'hypothèse suivante

$$\{(x, \xi) \mid (x, \xi) \in \text{WF}(u) \text{ et } (x, -\xi) \in \text{WF}(v)\} = \emptyset.$$

Sans en donner de démonstration, on peut indiquer que pour φ de classe C^∞ à support dans un petit voisinage d'un point x_0 , les fonctions $\widehat{\varphi u}(\xi)$ et $\widehat{\varphi v}(\xi)$ sont à décroissance rapide en dehors respectivement de deux cônes qui sont convolutifs. On conçoit que l'on puisse en déduire le fait que $\widehat{\varphi u} * \widehat{\varphi v}$ existe, et donc une définition raisonnable de $\varphi^2 uv$ puis de uv .

Les applications les plus spectaculaires concernent les équations aux dérivées partielles à coefficients variables. Si u est une solution de

$$\sum_{|\alpha| \leq m} a_\alpha(x) \partial^\alpha u(x) = 0,$$

on a

$$\text{WF}(u) \subset \left\{ (x, \xi) \mid \sum_{|\alpha| = m} a_\alpha(x) \xi^\alpha = 0 \right\}.$$

En outre, pour de nombreux problèmes d'évolution (penser par exemple à des propagations d'ondes dans des milieux anisotropes inhomogènes), la connaissance du front d'onde de u dans le passé détermine complètement et explicitement celui-ci dans l'avenir. Il y a propagation de $\text{WF}(u)$ dans l'espace des phases le long des courbes associées au hamiltonien $\sum_{|\alpha| = m} a_\alpha(x) \xi^\alpha$.

Les opérateurs permettant d'écrire la valeur de u à l'instant t à partir des données de Cauchy à l'instant 0 ne s'écrivent plus ni par des convolutions, ni en termes des transformées de Fourier. Par contre il en existe des écritures voisines sous forme d'opérateurs intégraux, la différence essentielle étant que la phase $x \cdot \xi$ de la transformation de Fourier est remplacée par une phase Φ plus générale, mais conservant le caractère essentiel que, pour ξ grand, la phase Φ est grande, et donc l'exponentielle $e^{i\Phi}$ très oscillante. Les opérateurs de ce type, appelés opérateurs intégraux de Fourier, jouent en quelque sorte le rôle d'une "transformation de Fourier à coefficients variables".

Un point de vue assez différent et se prêtant bien au calcul numérique, mais relevant de la même philosophie, est celui des bases d'ondelettes (Y. Meyer, 1985). Il s'agit, en dimension

⁽¹⁾Wave front en anglais

1, de bases hilbertiennes de L^2 formées des translatées et dilatées dyadiques d'une même fonction ψ

$$\psi_{k,j}(x) = 2^{k/2} \psi(2^k x - j).$$

Une fonction f de carré sommable se décompose donc sous la forme

$$f = \sum_{k,j \in \mathbb{Z}} c_{k,j}(f) \psi_{k,j}.$$

La fonction ψ doit posséder bien entendu des propriétés très particulières. Il nous suffira de dire ici que ψ est oscillante et à décroissance rapide à l'infini, si bien que l'on peut considérer que, pour k grand, $\psi_{k,j}$ a une "fréquence" de l'ordre de 2^k et est "concentrée" autour du point $j/2^k$.

On peut lire la régularité de f sur la décroissance, pour $k \rightarrow +\infty$, de ses coefficients d'ondelettes $c_{k,j}(f)$, mais le point important est que l'on peut y lire la régularité locale de f au voisinage d'un point x_0 sur la décroissance en k des $c_{k,j}(f)$ correspondant aux ondelettes concentrées autour de points voisins de x_0 .

9.8. Transformation de Laplace

L'étude de cette transformation nécessite quelques connaissances sur les fonctions d'une variable complexe. Rappelons qu'une fonction F de la variable $p = \lambda + i\mu$, définie dans un ouvert ω du plan est dite holomorphe si elle est dérivable au sens complexe, c'est-à-dire si $(F(p+h) - F(p))/h$ tend vers une limite qu'on note $F'(p)$ lorsque $h \in \mathbb{C} \setminus \{0\}$ tend vers 0.

On démontre qu'il est équivalent de dire que, pour tout disque ouvert D (de centre p_0 et de rayon r) contenu dans ω , la fonction F est développable en série entière

$$\forall p \in D, \quad F(p) = \sum \frac{F^j(p_0)}{j!} (p - p_0)^j$$

convergente dans ce disque.

Définition 9.8.1. — Soit f une fonction localement sommable, à support dans $[0, \infty[$, telle qu'il existe C et λ_0 avec $|f(t)| \leq C e^{\lambda_0 t}$. La transformée de Laplace de f est la fonction $\mathcal{L}f$ holomorphe dans le demi-plan $\operatorname{Re} p > \lambda_0$, définie par

$$\mathcal{L}f(p) = \int_0^\infty e^{-pt} f(t) dt.$$

Pour montrer que $\mathcal{L}f$ est holomorphe, on applique le théorème de dérivation sous le signe somme de Lebesgue, qui est valable pour les fonctions dérivables d'un paramètre complexe. En effet, dans chaque demi-plan $\operatorname{Re} p > \lambda_1 > \lambda_0$, la fonction $p \mapsto e^{-pt} f(t)$ a pour dérivée $-te^{-pt} f(t)$ qui est majorée en module par la fonction $Cte^{(\lambda_0 - \lambda_1)t}$ sommable dans $[0, \infty[$.

Définition 9.8.2. — Soit $u \in \mathcal{E}'$. Sa transformée de Laplace est la fonction

$$\mathcal{L}f(p) = \langle u(t), e^{-pt} \rangle$$

qui est holomorphe dans tout le plan.

Il suffirait de montrer que le théorème de dérivation sous le crochet s'applique aussi aux dérivations complexes pour obtenir le résultat.

La transformation de Laplace des fonctions est insuffisante pour beaucoup d'applications. On souhaite pouvoir utiliser la formule des sauts pour dériver des fonctions discontinues, et éventuellement utiliser des parties finies. La sous-algèbre suivante de $\mathcal{D}'_+$ est souvent suffisante.

9.8.3. Une algèbre de convolution. — On notera $\mathcal{A}$ l'ensemble des $u \in \mathcal{D}'_+$ qui peuvent s'écrire de la manière suivante : $u = v + f$ où v appartient à $\mathcal{E}'$ et où f est une fonction de classe C^∞ , à support dans $]0, \infty[$, telle qu'il existe $\lambda_0 \in \mathbb{R}$ avec $e^{-\lambda_0 t} f(t) \in \mathcal{S}$.

Cette propriété signifie simplement que, en dehors d'un compact, la distribution u est de classe C^∞ , et que ses dérivées ont une croissance au plus exponentielle (avec un même exposant pour toutes les dérivées). Si $\varphi \in C_0^\infty$ est égale à 1 au voisinage du compact en question, il suffit en effet de poser $v = \varphi u$ et $f = (1 - \varphi)u$.

On peut alors définir la transformée de Laplace $\mathcal{L}u$ de u comme étant égale à $\mathcal{L}v + \mathcal{L}f$, ces deux fonctions étant données par les définitions 9.8.1 et 9.8.2. Il faut bien sûr vérifier que la fonction obtenue, qui est holomorphe pour $\operatorname{Re} p > \lambda_0$, ne dépend pas de la décomposition $u = v + f$ choisie. Cela revient à vérifier que les deux définitions coïncident pour une fonction C^∞ à support compact, ce qui est évident.

On dira que λ_0 est une *abscisse de définition* de la transformée de u .

Exercice 9.8.4. — Démontrer que, pour u_1 et u_2 appartenant à $\mathcal{D}'_+$, et pour λ réel, on a

$$e^{-\lambda t}(u_1 \star u_2) = (e^{-\lambda t}u_1) \star (e^{-\lambda t}u_2). \quad (9.20)$$

En déduire que $\mathcal{A}$ est une algèbre de convolution (on vérifiera que, si $w \in \mathcal{E}'$ et si $e^{-\lambda t}f \in \mathcal{S}$ est comme ci-dessus, on a $e^{-\lambda t}(w \star f) \in \mathcal{S}$).

9.8.5. Transformées de Laplace et de Fourier. — Soit $u = v + f$ un élément de $\mathcal{A}$, d'abscisse de définition λ_0 . Pour $\lambda > \lambda_0$, on a

$$\mathcal{L}v(\lambda + iv) = \left\langle v(t), e^{-\lambda t} e^{-ivt} \right\rangle = \left\langle e^{-\lambda t} v(t), e^{-ivt} \right\rangle$$

et

$$\mathcal{L}f(\lambda + iv) = \int e^{-ivt} \left\{ e^{-\lambda t} f(t) \right\} dt.$$

Cela montre que la valeur au point $p = \lambda + iv$ de la transformée de Laplace de u est égale à la valeur au point v de la transformée de Fourier de $e^{-\lambda t}u$.

Les propriétés bien connues de la transformée de Fourier donnent immédiatement le théorème suivant (dont la démonstration directe n'est pas très difficile). On utilisera (9.20) pour la démonstration de la partie (c).

Théorème 9.8.6. — Soient $u, u_1, u_2 \in \mathcal{A}$.

(a) (injectivité) Si les transformées de Laplace de u_1 et u_2 coïncident pour $\operatorname{Re} p$ assez grand, alors $u_1 = u_2$.

(b) On a

$$\mathcal{L}\left(\frac{du}{dt}\right) = p\mathcal{L}u(p), \quad \mathcal{L}(tu(t)) = -\frac{d}{dp}\mathcal{L}u.$$

(c) Soit λ_0 la plus grande des abscisses de définition des transformées de Laplace des u_j . On a pour $\operatorname{Re} p > \lambda_0$

$$\mathcal{L}(u_1 \star u_2)(p) = \mathcal{L}u_1(p)\mathcal{L}u_2(p).$$

On peut reconstruire u à partir de sa transformée de Laplace en utilisant la formule d'inversion de Fourier sur une droite $\operatorname{Re} p = \lambda$. Grâce aux propriétés des fonctions holomorphes, on a une propriété d'injectivité beaucoup plus forte : il suffit que $\mathcal{L}u_1$ et $\mathcal{L}u_2$ coïncident sur un ensemble ayant un point d'accumulation (un segment de l'axe réel par exemple) pour avoir $u_1 = u_2$. Pour pouvoir utiliser "à l'envers" les tables de transformées de Laplace, il suffit donc de connaître $\mathcal{L}u(p)$ pour p réel assez grand.

9.8.7. Une algèbre de convolution plus grande. — On peut plus généralement considérer l'ensemble $\mathcal{B}$ des $u \in \mathcal{D}'_+$ vérifiant $e^{-\lambda_0 t} u \in \mathcal{S}'$ pour un λ_0 réel. L'ensemble $\mathcal{B}$ est une algèbre de convolution pour laquelle on peut définir la transformation de Laplace. Pour $\operatorname{Re} p > \lambda_0$, on pose

$$\mathcal{L}u(p) = \left\langle e^{-\lambda_0 t} u(t), e^{\lambda_0 t} e^{-pt} \Psi(t) \right\rangle \quad (9.21)$$

où la fonction Ψ est de classe C^∞ , à support dans $[-1, \infty[$ et égale à 1 au voisinage de $[0, \infty[$. La fonction figurant à droite dans le crochet est alors dans $\mathcal{S}(\mathbb{R})$, ce qui donne un sens à (9.21). La fonction ainsi définie ne dépend pas des choix de Ψ et λ_0 , et le théorème 9.8.6 s'étend aux éléments de $\mathcal{B}$.

Exercice 9.8.8. — Montrer que la transformation de Laplace est un isomorphisme du corps $\mathcal{K}$ défini au n° 8.4.7, sur le corps des fractions rationnelles en p , et que cet isomorphisme est précisément celui du calcul symbolique.

Dans l'exercice suivant, on notera p^k , pour k réel et $\operatorname{Re} p > 0$ le nombre complexe de module $|p|^k$ et d'argument $k \arg p$, avec $-\pi/2 < \arg p < \pi/2$.

Exercice 9.8.9 (Dérivations d'ordre non entier). — (a) Démontrer que la transformée de Laplace de $H(t)t^{k-1}$, pour $k > 0$ est égale à $c_k p^{-k}$, où c_k est une constante. On pourra montrer que $c_k = \Gamma(k)$, où Γ est la fonction d'Euler (voir l'exercice 1.2.12).

(b) En utilisant l'exercice 5.3.5 montrer que pour k négatif non entier, la transformée de Laplace de $\operatorname{pf}(t_+^{k-1})$ est égale à $c_k p^{-k}$.

(c) Montrer que, pour tout $k \in \mathbb{R}$, il existe une distribution T_k et une seule appartenant à $\mathcal{D}'_+$ dont la transformée de Laplace est p^k . En notant D^k l'opérateur $u \mapsto T_k \star u$, montrer que l'on a

$$\forall k, m \in \mathbb{R}, D^k(D^m u) = D^{k+m} u,$$

et que pour k entier positif, $D^k u$ est la dérivée d'ordre k de u tandis que $D^{-k} u$ en est l'unique primitive d'ordre k à support dans $[0, \infty[$.

CHAPITRE 10

ESPACES DE SOBOLEV

Nous avons déjà défini les espaces H^m d'exposant entier. Comme nous allons le voir, pour décrire avec précision leurs restrictions à des hypersurfaces, il est indispensable d'introduire des espaces d'exposant fractionnaire. Or, la principale application que nous avons en vue, le problème de Dirichlet pour l'équation de Laplace dans un ouvert borné de $\mathbb{R}^n$, nécessite l'étude des restrictions à la frontière des éléments de l'espace H^1 .

Cela dit, grâce à la transformation de Fourier, l'étude des espaces de Sobolev d'exposant quelconque n'est pas plus difficile. Pour les espaces d'exposant entier, certaines démonstrations sont même grandement simplifiées si on les compare aux démonstrations directes.

10.1. Structure hilbertienne et dualité

Définition 10.1.1 (Espaces de Sobolev). — Soit $s \in \mathbb{R}$. On dit qu'une distribution u dans $\mathbb{R}^n$ appartient à l'espace H^s si u est tempérée, si $\widehat{u}$ est une fonction localement sommable, et si on a

$$\int (1 + |\xi|^2)^s |\widehat{u}(\xi)|^2 d\xi < +\infty.$$

Cette définition comporte deux aspects. D'une part, elle exige une certaine régularité de $\widehat{u}$: être localement sommable (et même localement de carré sommable), ce qui interdit à u d'être "trop grande" à l'infini. D'autre part, elle exige une décroissance de $\widehat{u}$ à l'infini, d'autant plus rigoureuse que s est grand, qui correspond à une régularité de u .

Une formulation équivalente consiste à dire que $(1 + |\xi|^2)^{s/2} \widehat{u}$ appartient à L^2 . Lorsque $s \in \mathbb{N}$, cela équivaut à demander que $\xi^\alpha \widehat{u}$ appartienne à L^2 pour

$|\alpha| \leq s$, et donc à la condition $\partial^\alpha u \in L^2$ pour ces mêmes α . On retrouve bien la définition donnée dans la section 6.2.

Il est évident, au vu de la définition, que les espaces H^s décroissent avec s , et que, pour $u \in H^s$, les dérivées d'ordre m de u appartiennent à H^{s-m} .

Théorème 10.1.2. — *Les espaces H^s sont hilbertisables : munis du produit scalaire*

$$(u | v)_s = \int (1 + |\xi|^2)^s \widehat{u}(\xi) \overline{\widehat{v}(\xi)} d\xi,$$

ou de tout autre produit scalaire donnant une norme équivalente à la norme

$$\|u\|_s = \left\| (1 + |\xi|^2)^{s/2} \widehat{u} \right\|_{L^2},$$

ce sont des espaces de Hilbert.

Il est clair que $(\cdot | \cdot)_s$ est un produit scalaire. D'autre part, l'application $u \mapsto (1 + |\xi|^2)^{s/2} \widehat{u}$ est par définition une bijection isométrique de H^s sur $L^2(\mathbb{R}^n)$. Ce dernier espace étant complet, il en est de même de H^s , pour la norme ci-dessus ou pour toute norme équivalente.

Théorème 10.1.3. — *L'espace C_0^∞ est dense dans H^s pour tout s .*

Il est facile de voir que $\mathcal{S}$ est dense dans H^s . En effet, l'isométrie inverse de $u \mapsto (2\pi)^{-n/2} (1 + |\xi|^2)^{s/2} \widehat{u}$ doit transformer le sous-espace dense $\mathcal{S}$ de L^2 en un sous-espace dense de H^s . Or cette même application est une bijection de $\mathcal{S}$ sur lui-même : elle est composée de la transformation de Fourier et de la multiplication par $(1 + |\xi|^2)^{s/2}$, fonction qui appartient à $\mathcal{O}_M$ ainsi que son inverse.

Montrons maintenant qu'il existe p et C , ne dépendant que de s , tels que

$$\forall \varphi \in \mathcal{S}, \quad \|\varphi\|_s \leq C \mathcal{N}_p(\varphi). \quad (10.1)$$

Pour chaque N , on a en effet

$$\left\| (1 + |\xi|^2)^{s/2} \widehat{\varphi} \right\|_{L^2} \leq \sup_{\xi} ((1 + |\xi|^2)^{s/2+N} |\widehat{\varphi}(\xi)|) \left\| (1 + |\xi|^2)^{-N} \right\|_{L^2}.$$

En choisissant $N > n/2$, le membre de droite est majoré par une constante fois $\mathcal{N}_q(\widehat{\varphi})$, où q est le plus petit entier supérieur ou égal à $s/2 + N$. D'après le théorème 9.2.4, on a l'estimation (10.1) avec $p = q + n + 1$.

Un élément u de H^s et $\varepsilon > 0$ étant donnés, on peut d'abord trouver $\varphi \in \mathcal{S}$ vérifiant $\|u - \varphi\|_s \leq \varepsilon/2$. D'après le théorème 9.2.9, il existe une suite $\varphi_j \in C_0^\infty$ telle que l'on ait $\mathcal{N}_p(\varphi - \varphi_j) \rightarrow 0$. Il résulte de la majoration (10.1) que l'on a $\|\varphi - \varphi_j\|_s \leq \varepsilon/2$ en choisissant j assez grand. On peut donc, pour tout ε , trouver un élément $\varphi_j \in C_0^\infty$ tel que $\|u - \varphi_j\|_s \leq \varepsilon$, ce qui achève la démonstration.

Théorème 10.1.4 (Extension de la dualité). — Soit $s \in \mathbb{R}$ et $u \in H^{-s}$. Alors, l'application $\varphi \mapsto \langle u, \varphi \rangle$ de $\mathcal{S}$ dans $\mathbb{C}$ se prolonge de manière unique en une forme linéaire continue sur H^s . On notera encore $\langle u, v \rangle$, pour $u \in H^{-s}$ et $v \in H^s$, ce prolongement.

Cette extension identifie canoniquement H^{-s} au dual de H^s : pour toute forme linéaire L continue sur H^s , il existe un et un seul $u \in H^{-s}$ vérifiant $L(v) = \langle u, v \rangle$ pour tout $v \in H^s$.

Soit donc $u \in H^{-s}$. On a $\langle \widehat{u}, \psi \rangle = \langle u, \widehat{\psi} \rangle$ pour $\psi \in \mathcal{S}$. En posant $\widehat{\psi} = \varphi$, on obtient pour tout $\varphi \in \mathcal{S}$

$$\langle u, \varphi \rangle = (2\pi)^{-n} \langle \widehat{u}, \mathcal{F}\varphi \rangle = (2\pi)^{-n} \int [(1 + |\xi|^2)^{-s/2} \widehat{u}] [(1 + |\xi|^2)^{s/2} \mathcal{F}\varphi] d\xi.$$

L'inégalité de Cauchy-Schwarz (et un changement de variable $\xi \mapsto -\xi$) donnent immédiatement

$$|\langle u, \varphi \rangle| \leq (2\pi)^{-n} \|u\|_{-s} \|\varphi\|_s.$$

Cela prouve que la forme linéaire $\varphi \mapsto \langle u, \varphi \rangle$, qui est définie sur le sous-espace dense $\mathcal{S}$ de H^s est continue. Elle se prolonge donc de façon unique en une forme linéaire continue sur H^s .

Réciproquement, soit L une forme linéaire continue sur H^s , et formons l'application suivante M de L^2 dans $\mathbb{C}$

$$M(f) = L \left(\mathcal{F}^{-1}[(1 + |\xi|^2)^{-s/2} f] \right).$$

Si on pose $w = \mathcal{F}^{-1}[(1 + |\xi|^2)^{-s/2} f]$, on a $w \in H^s$ et $\|w\|_s = \|f\|_{L^2}$. Par continuité de L , il existe une constante C telle que

$$\forall f \in L^2, |M(f)| \leq C \|w\|_s = C \|f\|_{L^2}.$$

Cela montre que la forme linéaire M est continue sur l'espace de Hilbert L^2 , et il existe donc un élément $g \in L^2$ tel que

$$\forall f \in L^2, M(f) = \int g(\xi) f(\xi) d\xi.$$

Notons u la distribution $\mathcal{F}[(1 + |\xi|^2)^{s/2} g]$. On a $(1 + |\xi|^2)^{-s/2} \mathcal{F}u = (2\pi)^n g \in L^2$ et donc $u \in H^{-s}$. D'autre part, pour $\varphi \in \mathcal{S}$, on a

$$\langle u, \varphi \rangle = \left\langle \mathcal{F}[(1 + |\xi|^2)^{s/2} g], \varphi \right\rangle = \left\langle [(1 + |\xi|^2)^{s/2} g], \widehat{\varphi} \right\rangle,$$

et donc

$$\langle u, \varphi \rangle = \int g(\xi) (1 + |\xi|^2)^{s/2} \widehat{\varphi}(\xi) d\xi = M \left((1 + |\xi|^2)^{s/2} \widehat{\varphi} \right)$$

ou encore, par définition de M

$$\langle u, \varphi \rangle = L \left(\mathcal{F}^{-1}[(1 + |\xi|^2)^{-s/2}(1 + |\xi|^2)^{s/2}\widehat{\varphi}] \right).$$

Nous obtenons donc $\langle u, \varphi \rangle = L(\varphi)$, et ces deux formes linéaires continues sur H^s qui coïncident sur $\mathcal{S}$ sont donc égales. Cela achève la démonstration.

Remarque 10.1.5. — Compte tenu du théorème 6.2.5, le résultat précédent montre que, pour s entier négatif, la nouvelle définition des espaces H^s coïncide avec celle que nous avons donnée dans la section 6.2.

Comme pour les espaces d'indice entier, le dual de H^s s'identifie canoniquement à H^{-s} . Pour chaque choix d'un produit scalaire sur H^s on dispose aussi de l'identification — dépendant de ce choix — de l'espace H^s avec son dual, comme pour tout espace de Hilbert.

10.2. Régularité et caractère local

Les inégalités élémentaires suivantes nous serviront à majorer des produits de convolution en ξ , et donc à contrôler la régularité de produits dans l'espace des x . La seconde est connue sous le nom d'inégalité de Peetre.

Lemme 10.2.1. — *Pour ξ et η dans $\mathbb{R}^n$, on a*

$$\forall s \geq 0, (1 + |\xi|^2)^s \leq 4^s \left\{ (1 + |\xi - \eta|^2)^s + (1 + |\eta|^2)^s \right\} \quad (10.2)$$

et

$$\forall s \in \mathbb{R}, (1 + |\xi|^2)^s \leq 2^{|s|} (1 + |\eta|^2)^s (1 + |\xi - \eta|^2)^{|s|}. \quad (10.3)$$

La première majoration résulte simplement de l'inégalité triangulaire et de la majoration $(a + b)^s \leq 2^s(a^s + b^s)$ pour des nombres positifs. Quant à la seconde, il est équivalent de montrer que

$$(1 + |\xi|^2)^s (1 + |\eta|^2)^{-s} \leq 2^{|s|} (1 + |\xi - \eta|^2)^{|s|},$$

et on voit, en échangeant éventuellement ξ et η , qu'il suffit de démontrer (10.3) lorsque s est positif. Par homogénéité, il suffit alors de le prouver pour $s = 1$, ce que nous pensons pouvoir laisser au lecteur.

Théorème 10.2.2. — *Pour $s > n/2$ les éléments de H^s sont des fonctions continues tendant vers 0 à l'infini. En outre, le produit de deux éléments de H^s est encore dans H^s .*

On peut en effet écrire

$$\widehat{u}(\xi) = [(1 + |\xi|^2)^{s/2}\widehat{u}(\xi)][(1 + |\xi|^2)^{-s/2}].$$

Le premier facteur appartient à L^2 si $u \in H^s$, et le second facteur appartient à L^2 dès que $s > n/2$. On a donc $\widehat{u} \in L^1$, et d'après le théorème 9.1.1, la fonction u est continue et tend vers 0 à l'infini.

Pour démontrer que H^s est une algèbre, on écrit pour deux éléments u et v de cet espace

$$\widehat{uv}(\xi) = (2\pi)^{-n} \int \widehat{u}(\xi - \eta) \widehat{v}(\eta) d\eta.$$

C'est en effet, "lue à l'envers", la partie (a) du théorème 9.5.1, qui assure que la transformée de Fourier du produit de convolution de deux fonctions sommables est le produit des transformées de Fourier. En multipliant les deux membres par $(1 + |\xi|^2)^{s/2}$, et en majorant cette dernière quantité à l'aide de (10.2), on obtient

$$\begin{aligned} (1 + |\xi|^2)^{s/2} |\widehat{uv}(\xi)| &\leq C \int \left[(1 + |\xi - \eta|^2)^{s/2} |\widehat{u}(\xi - \eta)| \right] |\widehat{v}(\eta)| d\eta \\ &\quad + C \int |\widehat{u}(\xi - \eta)| \left[(1 + |\eta|^2)^{s/2} |\widehat{v}(\eta)| \right] d\eta \end{aligned}$$

On reconnaît dans chacun des deux termes du membre de droite le produit de convolution d'une fonction de L^1 et d'une fonction de L^2 qui, d'après l'exercice 3.3.7, appartient à L^2 . On a donc $(1 + |\xi|^2)^{s/2} \widehat{uv}(\xi) \in L^2$ et $uv \in H^s$.

Corollaire 10.2.3. — *Pour m entier positif et $s > n/2 + m$, les éléments de H^s sont des fonctions de classe C^m . En particulier, une fonction appartenant à H^s pour tout s est une fonction de classe C^∞ .*

Il suffit d'appliquer le théorème aux dérivées de u d'ordre $\leq m$.

Théorème 10.2.4. — *Quel que soit $s \in \mathbb{R}$, pour $u \in H^s$ et $\varphi \in \mathcal{S}$, on a $\varphi u \in H^s$.*

Le théorème 10.2.2 entraînerait ce résultat, mais seulement pour $s > n/2$. On écrit encore $\widehat{\varphi u}(\xi)$ sous forme d'un produit de convolution (c'est cette fois-ci le résultat de l'exercice 9.5.5 "lu à l'envers"), et on utilise ici l'inégalité (10.3) pour majorer $(1 + |\xi|^2)^{s/2}$. On obtient

$$\begin{aligned} (1 + |\xi|^2)^{s/2} |\widehat{\varphi u}(\xi)| \\ \leq C \int \left\{ (1 + |\xi - \eta|^2)^{|s|/2} |\widehat{\varphi}(\xi - \eta)| \right\} \left\{ (1 + |\eta|^2)^{s/2} |\widehat{u}(\eta)| \right\} d\eta. \end{aligned}$$

On reconnaît encore là le produit de convolution de la fonction $\xi \mapsto (1 + |\xi|^2)^{|s|/2} |\widehat{\varphi}(\xi)|$ qui appartient à L^1 , et d'une fonction de L^2 . On a donc $(1 + |\xi|^2)^{s/2} \widehat{\varphi u}(\xi) \in L^2$ et la conclusion.

Remarque 10.2.5. — Un tel résultat permet de définir les espaces $H_{\text{loc}}^s(\Omega)$ dans un ouvert $\Omega \subset \mathbb{R}^n$, et de montrer qu'ils jouissent de propriétés raisonnables. Si $u \in \mathcal{D}'(\Omega)$, on dit que $u \in H_{\text{loc}}^s(\Omega)$ si pour toute $\varphi \in C_0^\infty(\Omega)$, on a $\varphi u \in H^s$.

Le théorème précédent assure que la restriction à Ω d'un élément de $H^s(\mathbb{R}^n)$ appartient à $H_{\text{loc}}^s(\Omega)$, ce qui est le minimum que l'on puisse exiger d'une telle notion. Le lecteur pourra montrer à titre d'exercice que, pour que u appartienne à $H_{\text{loc}}^s(\Omega)$, il faut et il suffit que l'on puisse trouver en chaque point $x_0 \in \Omega$ une fonction $\varphi \in C_0^\infty(\Omega)$ non nulle en x_0 telle que $\varphi u \in H^s$ (utiliser les partitions de l'unité et la compacité).

Le corollaire 10.2.3 montre que les éléments de H_{loc}^s appartiennent à C^m pour $s > n/2 + m$. Il est clair d'autre part que les éléments de C^m appartiennent à H_{loc}^m . Il s'agit de deux échelles de régularité qui ne sont comparables qu'à un décalage près.

Localement, les espaces H^s couvrent tout l'éventail des régularités. L'intersection des H_{loc}^s est l'espace des fonctions C^∞ , et le lecteur pourra montrer que la réunion des H_{loc}^s est l'espace des distributions d'ordre fini.

10.2.6. Application : régularité des solutions de l'équation

$$\Delta u - \lambda u = f. \quad (10.4)$$

Nous avons vu au n° 9.5.12 que, pour $f \in \mathcal{S}'$ et $\lambda > 0$, cette équation possède une et une seule solution tempérée, et que l'on a $\widehat{u} = -(|\xi|^2 + \lambda)^{-1} \widehat{f}$. Si maintenant $f \in H^s$, on a

$$(1 + |\xi|^2)^{(s+2)/2} |\widehat{u}(\xi)| = \frac{1 + |\xi|^2}{\lambda + |\xi|^2} (1 + |\xi|^2)^{s/2} |\widehat{f}(\xi)|.$$

La fraction étant uniformément bornée, il en résulte que le membre de gauche appartient à L^2 . Nous avons donc montré que pour $f \in H^s$, l'unique solution tempérée de (10.4) appartient à H^{s+2} .

Il est clair qu'un tel résultat est le meilleur possible, l'opérateur $\Delta - \lambda$ appliquant H^{s+2} dans H^s . Contrairement à ce que l'on pourrait penser, l'analogue dans les espaces C^m est faux. Si $(\Delta - \lambda)u$ est de classe C^m , on peut montrer que la fonction u est de classe C^{m+1} , mais elle n'est pas en général de classe C^{m+2} .

10.3. Traces et prolongements

A. Trace d'une fonction définie dans $\mathbb{R}^n$

Considérons l'hyperplan $x_n = 0$ dans $\mathbb{R}^n$. Nous nous intéressons à l'opérateur de trace γ qui à une "bonne" fonction $\varphi(x', x_n)$ dans $\mathbb{R}^n$ associe la

fonction $\gamma\varphi(x') = \varphi(x', 0)$. Cet opérateur, qui est bien défini pour des fonctions continues, n'a pas de sens a priori pour des (classes de) fonctions localement sommables, l'hyperplan étant de mesure nulle dans $\mathbb{R}^n$. D'autre part, il existe des éléments de $L^2(\mathbb{R}^n)$ qui sont des fonctions continues pour $x_n \neq 0$ et qui tendent vers $+\infty$ lorsque x_n tend vers 0, et il paraît exclu de définir la trace d'une telle fonction.

Nous allons montrer que l'on peut définir raisonnablement γu dès que $u \in H^s$ avec $s > 1/2$. Sauf en dimension 1, une telle condition n'implique pas la continuité de u , qui n'est, a priori, qu'une classe de fonctions définies à un ensemble de mesure nulle près. C'est le prolongement par continuité de l'opérateur de trace usuel qui nous fournira la solution.

Nous noterons systématiquement (x', x_n) le point courant de $\mathbb{R}^n$, avec $x' \in \mathbb{R}^{n-1}$, et nous utiliserons des notations analogues pour la variable de Fourier.

Théorème 10.3.1. — (a) Pour tout $s > 1/2$, l'opérateur γ , de $\mathcal{S}(\mathbb{R}^n)$ dans $\mathcal{S}(\mathbb{R}^{n-1})$ défini par $\gamma\varphi(x') = \varphi(x', 0)$ se prolonge de manière unique en un opérateur linéaire continu, noté encore γ de $H^s(\mathbb{R}^n)$ dans $H^{s-1/2}(\mathbb{R}^{n-1})$.

(b) L'opérateur γ est surjectif, de $H^s(\mathbb{R}^n)$ sur $H^{s-1/2}(\mathbb{R}^{n-1})$.

Pour démontrer la première partie, $\mathcal{S}$ étant dense dans H^s , il nous suffira de prouver qu'il existe une constante C telle que

$$\forall \varphi \in \mathcal{S}(\mathbb{R}^n), \quad \|\gamma\varphi\|_{H^{s-1/2}(\mathbb{R}^{n-1})} \leq C \|\varphi\|_{H^s(\mathbb{R}^n)}. \quad (10.5)$$

Le point important est d'écrire l'opérateur de trace en termes de transformation de Fourier. On a

$$\varphi(x', 0) = (2\pi)^{-n} \iint e^{ix' \cdot \xi'} \widehat{\varphi}(\xi', \xi_n) d\xi' d\xi_n$$

d'après la formule d'inversion dans $\mathbb{R}^n$. On a donc

$$\varphi(x', 0) = (2\pi)^{-(n-1)} \int e^{ix' \cdot \xi'} \left\{ \frac{1}{2\pi} \int \widehat{\varphi}(\xi', \xi_n) d\xi_n \right\} d\xi'$$

et il en résulte que la transformée de Fourier (dans $\mathbb{R}^{n-1}$) de $\gamma\varphi$ est la fonction entre accolades.

Nous désignerons désormais par $f(\xi)$ et $g(\xi')$ les transformées de Fourier respectives de φ et $\gamma\varphi$. On peut écrire

$$g(\xi') = (1/2\pi) \int \left[f(\xi', \xi_n) (1 + |\xi'|^2 + \xi_n^2)^{s/2} \right] \left[(1 + |\xi'|^2 + \xi_n^2)^{-s/2} \right] d\xi_n.$$

L'inégalité de Cauchy-Schwarz donne

$$4\pi^2 |g(\xi')|^2 \leq \left(\int |f(\xi)|^2 (1 + |\xi|^2)^s d\xi_n \right) \left(\int (1 + |\xi'|^2 + \xi_n^2)^{-s} d\xi_n \right) \quad (10.6)$$

en remarquant, c'est là qu'intervient l'hypothèse $s > 1/2$, que la seconde intégrale est finie. Celle-ci se calcule à l'aide du changement de variable $\xi_n = (1 + |\xi'|^2)^{1/2}\lambda$, et on obtient

$$\int (1 + |\xi'|^2 + \xi_n^2)^{-s} d\xi_n = c_s (1 + |\xi'|^2)^{-s+1/2} \quad (10.7)$$

avec

$$c_s = \int (1 + \lambda^2)^{-s} d\lambda.$$

En majorant $|g(\xi')|^2$ par (10.6), et compte tenu du calcul précédent, on obtient

$$\int (1 + |\xi'|^2)^{s-1/2} |g(\xi')|^2 d\xi' \leq \frac{c_s}{4\pi^2} \int (1 + |\xi|^2)^s |f(\xi)|^2 d\xi.$$

Ces intégrales étant égales respectivement aux carrés de $\|\gamma\varphi\|_{H^{s-1/2}(\mathbb{R}^{n-1})}$ et de $\|\varphi\|_{H^s(\mathbb{R}^n)}$, nous avons établi l'estimation (10.5), ce qui achève la démonstration de la partie (a) du théorème.

Pour $u \in H^s(\mathbb{R}^n)$, $s > 1/2$, on peut d'après ce qui précède expliciter l'opérateur γ de la manière suivante :

$$\gamma u = \mathcal{F}_{n-1}^{-1} \left\{ \frac{1}{2\pi} \int \mathcal{F}_n u(\xi', \xi_n) d\xi_n \right\},$$

où on a mis en indice de la transformation de Fourier la dimension de l'espace concerné.

10.3.2. Démonstration du théorème 10.3.1 (b). — Soit maintenant v un élément de $H^{s-1/2}(\mathbb{R}^{n-1})$. En notant $g(\xi')$ la transformée de Fourier de v , nous allons définir une distribution u dans $\mathbb{R}^n$ comme la transformée de Fourier inverse de la fonction f suivante

$$f(\xi) = k_N \frac{(1 + |\xi'|^2)^N}{(1 + |\xi|^2)^{N+1/2}} g(\xi').$$

Notre but est de démontrer, en choisissant judicieusement les constantes N et k_N , que l'on a $\gamma u = v$ et $u \in H^s$. On est donc ramené à prouver que l'on a

$$\int (1 + |\xi|^2)^s |f(\xi)|^2 d\xi < \infty, \quad (10.8)$$

$$(2\pi)^{-1} \int f(\xi', \xi_n) d\xi_n = g(\xi'). \quad (10.9)$$

On a, par définition de f

$$\int (1 + |\xi|^2)^s |f(\xi)|^2 d\xi \leq C \int (1 + |\xi'|^2)^{2N} |g(\xi')|^2 d\xi' \int (1 + |\xi|^2)^{s-2N-1} d\xi_n. \quad (10.10)$$

La seconde intégrale est finie à condition de choisir $N > s/2 - 1/4$, et elle est égale d'après (10.7) à une constante fois $(1 + |\xi'|^2)^{s-2N-1/2}$. Le membre de droite de (10.10) est donc égal à une constante fois $\int (1 + |\xi'|^2)^{s-1/2} |g(\xi')|^2 d\xi'$, dont la finitude exprime précisément l'hypothèse $v \in H^{s-1/2}$. Nous avons donc établi (10.8).

D'après l'expression de f , on a

$$\int f(\xi', \xi_n) d\xi_n = k_N (1 + |\xi'|^2)^N g(\xi') \int (1 + |\xi|^2)^{-N-1/2} d\xi_n.$$

L'intégrale de droite est égale, d'après (10.7) à $c_{N+1/2} (1 + |\xi'|^2)^{-N}$. Il suffit donc de choisir la constante k_N égale à $2\pi/c_{N+1/2}$ pour avoir (10.9). Cela achève la démonstration du théorème 10.3.1.

B. L'espace $H^1(\mathbb{R}_+^n)$

Nous noterons $\mathbb{R}_+^n$ le demi-espace ouvert de $\mathbb{R}^n$ défini par $x_n > 0$. Bien qu'il soit possible, et utile, de définir les espaces de Sobolev d'exposant quelconque dans un ouvert régulier — le demi-espace étant le cas le plus simple — nous nous limiterons au cas de H^1 , qui nous suffira pour l'étude du problème de Dirichlet.

Pour bien comprendre ce qui va suivre, il est bon de considérer d'abord les conséquences de la formule des sauts pour une fonction f appartenant à $C^1(\overline{\mathbb{R}_+^n})$. Si on prolonge cette fonction par 0 pour $x_n < 0$, on obtient une fonction Zf bornée dont les dérivées ne sont pas en général des fonctions. Plus précisément, la dérivée par rapport à x_n fait apparaître une simple couche ayant pour densité la trace de f sur l'hyperplan $x_n = 0$.

Il est donc équivalent dans ce cas de dire que la trace de f est nulle, et de dire que les dérivées de Zf sont des fonctions. Par contre, si on prolonge f par symétrie par rapport à $x_n = 0$, on obtient une fonction de classe C^1 par morceaux dont les sauts sont nuls, et les dérivées de ce prolongement sont des fonctions. Nous retrouverons ci-dessous les équivalents de ces remarques.

Définition 10.3.3. — (a) On appelle $H^1(\mathbb{R}_+^n)$ l'espace des $u \in L^2(\mathbb{R}_+^n)$ dont les dérivées premières au sens des distributions appartiennent également à $L^2(\mathbb{R}_+^n)$.

(b) On appelle $H_0^1(\mathbb{R}_+^n)$ l'espace des fonctions $u \in H^1(\mathbb{R}^n)$ qui sont nulles pour $x_n < 0$. Il s'identifie au sous-espace des $u \in H^1(\mathbb{R}_+^n)$ tels que $Zu \in H^1(\mathbb{R}^n)$,

où le prolongement Z est défini par $Zu(x) = u(x)$ pour $x_n > 0$ et $Zu(x) = 0$ sinon.

Ces deux espaces, munis du produit scalaire

$$(u | v) = \sum \int_{\mathbb{R}_+^n} \partial_i u(x) \overline{\partial_i v(x)} dx + \int_{\mathbb{R}_+^n} u(x) \overline{v(x)} dx$$

sont des espaces de Hilbert.

Il est facile de montrer le caractère hilbertien de ces espaces — et donc que H_0^1 étant complet est fermé dans $H^1(\mathbb{R}_+^n)$ — en reprenant la démonstration du théorème 6.2.2.

Il ne faut pas confondre $H^1(\mathbb{R}_+^n)$ avec l'espace $H_{\text{loc}}^1(\mathbb{R}_+^n)$ défini dans la section précédente. Rien ne limite la croissance d'un élément de ce dernier espace lorsque x tend vers l'infini ou vers l'hyperplan $x_n = 0$, tandis que le fait d'être de carré sommable dans le demi-espace (et la condition sur les dérivées) constituent une limitation sérieuse pour les éléments de $H^1(\mathbb{R}_+^n)$.

Théorème 10.3.4. — (a) L'espace $C_0^\infty(\overline{\mathbb{R}_+^n})$ constitué des restrictions à $\mathbb{R}_+^n$ des éléments de $C_0^\infty(\mathbb{R}^n)$ est dense dans $H^1(\mathbb{R}_+^n)$.

(b) L'espace $C_0^\infty(\mathbb{R}_+^n)$ est dense dans $H_0^1(\mathbb{R}_+^n)$.

Les éléments de $C_0^\infty(\overline{\mathbb{R}_+^n})$ sont des fonctions de classe C^∞ dans $\mathbb{R}_+^n$, nulles en dehors d'un ensemble borné, et dont toutes les dérivées se prolongent continûment à $\overline{\mathbb{R}_+^n}$. Les éléments de $C_0^\infty(\mathbb{R}_+^n)$ sont de plus nuls au voisinage de la frontière $x_n = 0$.

Nous allons utiliser, dans les deux cas, la méthode habituelle : régularisation et troncature. Toutefois, le fait que les éléments de $H^1(\mathbb{R}_+^n)$ ne soient pas définis dans tout l'espace, et le fait qu'on veuille dans le second cas des régularisées nulles près de $x_n = 0$ nécessite un contrôle plus précis des supports.

Soit χ une fonction positive de classe C^∞ , d'intégrale 1, à support dans l'intersection de la boule unité et de $\{x | x_n < -1/2\}$. On forme l'approximation de l'identité $\chi_\varepsilon(x) = \varepsilon^{-n} \chi(x/\varepsilon)$. La propriété importante est que, pour u définie et de carré sommable sur $\mathbb{R}_+^n$, la fonction

$$u_\varepsilon(x) = (u \star \chi_\varepsilon)(x) = \int \chi_\varepsilon(x - y) u(y) dy$$

est parfaitement définie et de classe C^∞ pour $x_n > -\varepsilon/2$, la fonction u étant alors définie sur le domaine d'intégration. La démonstration habituelle montre alors que, pour $u \in H^1(\mathbb{R}_+^n)$ et pour $x_n > -\varepsilon/2$, on a $\partial_i u_\varepsilon = \chi_\varepsilon \star (\partial_i u)$.

Si on note $\tilde{u} \in L^2(\mathbb{R}^n)$ un prolongement quelconque (Zu par exemple) de u , on a $\tilde{u} \star \chi_\varepsilon = u \star \chi_\varepsilon$ dans $\mathbb{R}_+^n$, et donc

$$\|u_\varepsilon - u\|_{L^2(\mathbb{R}_+^n)} \leq \|\tilde{u} \star \chi_\varepsilon - \tilde{u}\|_{L^2(\mathbb{R}^n)} \rightarrow 0.$$

On montre de même (en prolongeant l'élément $\partial_i u$ de $L^2(\mathbb{R}_+^n)$) que $\chi_\varepsilon \star (\partial_i u) - \partial_i u$ tend vers 0 en norme L^2 dans $\mathbb{R}_+^n$. Cela montre que l'on a

$$\|u_\varepsilon - u\|_{H^1(\mathbb{R}_+^n)} \rightarrow 0.$$

Un nombre $\beta > 0$ étant donné, on peut donc fixer un ε tel que la norme ci-dessus soit inférieure à $\beta/2$. Il ne reste plus qu'à prendre une fonction $\psi \in C_0^\infty(\mathbb{R}^n)$, égale à 1 sur la boule unité, et une fonction $\theta \in C^\infty(\mathbb{R})$, à support dans $[-\varepsilon/4, \infty[$, égale à 1 au voisinage de $[0, \infty[$, et à poser

$$\varphi_R(x) = \psi(x/R)\theta(x_n)u_\varepsilon(x).$$

Ces fonctions appartiennent à $C_0^\infty(\mathbb{R}^n)$, et le lecteur montrera facilement que les φ_R et les $\partial_i \varphi_R$ convergent respectivement vers u_ε et $\partial_i u_\varepsilon$ pour la norme L^2 dans $\mathbb{R}_+^n$. Pour R assez grand on a donc $\|u_\varepsilon - \varphi_R\|_{H^1(\mathbb{R}_+^n)} < \beta/2$.

La restriction de φ_R à $\mathbb{R}_+^n$ appartient à $C_0^\infty(\overline{\mathbb{R}_+^n})$, et approche u en norme à β près. Cela achève la démonstration de la partie (a).

La démonstration de la partie (b) est semblable, mais on choisit cette fois la fonction χ à support dans l'intersection de la boule unité et de $\{x | x_n > 1/2\}$. Si $u \in H_0^1(\mathbb{R}_+^n)$, c'est donc un élément de $H^1(\mathbb{R}^n)$ nul pour $x_n < 0$. On sait déjà que les $u_\varepsilon = u \star \chi_\varepsilon$ convergent vers u dans $H^1(\mathbb{R}^n)$. La nouveauté est que les u_ε sont à support dans $x_n \geq \varepsilon/2$. La fin de la démonstration est analogue, on approche ensuite u_ε par les $u_\varepsilon \psi(\cdot/R)$ qui appartiennent à $C_0^\infty(\mathbb{R}_+^n)$.

Théorème 10.3.5. — *L'application γ qui à une fonction de $C_0^\infty(\overline{\mathbb{R}_+^n})$ fait correspondre sa restriction à $x_n = 0$ se prolonge de manière unique en une application linéaire continue surjective, notée encore γ de $H^1(\mathbb{R}_+^n)$ sur $H^{1/2}(\mathbb{R}^{n-1})$.*

Considérons d'abord le prolongement par symétrie S qui à une fonction $u \in H^1(\mathbb{R}_+^n)$ associe la fonction Su définie par $Su(x', x_n) = u(x', |x_n|)$. Il est clair que $Su \in L^2(\mathbb{R}^n)$ et que l'on a $\|Su\|_{L^2(\mathbb{R}^n)} = \sqrt{2} \|u\|_{L^2(\mathbb{R}_+^n)}$.

Soit φ_j une suite d'éléments de $C_0^\infty(\overline{\mathbb{R}_+^n})$ qui converge vers u dans $H^1(\mathbb{R}_+^n)$. Alors les fonctions $S\varphi_j$ sont de classe C^1 par morceaux, et le saut est nul sur l'hyperplan $x_n = 0$. Les dérivées des $S\varphi_j$ sont donc des fonctions définies presque partout, et on a aussi $\|\partial_i Su\|_{L^2(\mathbb{R}^n)} = \sqrt{2} \|\partial_i u\|_{L^2(\mathbb{R}_+^n)}$. Il en résulte que (au $\sqrt{2}$ près) la restriction de S à $C_0^\infty(\overline{\mathbb{R}_+^n})$ est isométrique pour les normes H^1 . L'espace $C_0^\infty(\overline{\mathbb{R}_+^n})$ étant dense dans $H^1(\mathbb{R}_+^n)$, l'application S est une isométrie de ce dernier espace dans $H^1(\mathbb{R}^n)$.

Nous pouvons maintenant définir l'opérateur de trace de $H^1(\mathbb{R}_+^n)$ dans $H^{1/2}(\mathbb{R}^{n-1})$ comme le composé de S , qui envoie le premier espace dans $H^1(\mathbb{R}^n)$, et de l'opérateur de trace défini dans la section précédente, qui applique $H^1(\mathbb{R}^n)$ dans $H^{1/2}(\mathbb{R}^{n-1})$. Cet opérateur est continu comme composé d'opérateurs continus, et donne évidemment la trace usuelle pour les fonctions de $C_0^\infty(\overline{\mathbb{R}_+^n})$.

Il est facile de voir que γ est surjectif. Si $v \in H^{1/2}(\mathbb{R}^{n-1})$, on peut d'après le théorème 10.3.1 trouver $w \in H^1(\mathbb{R}^n)$ dont la trace est égale à v . Il suffit de désigner par u la restriction de w à $\mathbb{R}_+^n$. On a $u \in H^1(\mathbb{R}_+^n)$ et $\gamma u = v$.

Théorème 10.3.6 (Formule des sauts). — Soit $u \in H^1(\mathbb{R}_+^n)$, et soit Zu son prolongement par 0. On a alors $\partial_i(Zu) = Z(\partial_i u)$ pour $i = 1, \dots, n-1$ et

$$\partial_n(Zu) = Z(\partial_n u) + \mu \quad (10.11)$$

où μ est la distribution de simple couche, portée par $x_n = 0$, de densité γu .

Il suffit d'approcher un élément quelconque $u \in H^1(\mathbb{R}_+^n)$ par une suite φ_j d'éléments de $C_0^\infty(\overline{\mathbb{R}_+^n})$. Pour ceux-ci, la formule des sauts habituelle donne

$$\partial_n(Z\varphi_j) = Z(\partial_n \varphi_j) + \mu_j \quad (10.12)$$

où la simple couche μ_j a pour densité $\gamma\varphi_j$. Pour $j \rightarrow \infty$, on a $Z\varphi_j \rightarrow Zu$ dans L^2 et donc au sens des distributions, ce qui entraîne que $\partial_n(Z\varphi_j) \rightarrow \partial_n(Zu)$ au sens des distributions. Il est clair que $Z(\partial_n \varphi_j) \rightarrow Z(\partial_n u)$ dans L^2 et donc au sens des distributions. On a enfin, pour $\psi \in C_0^\infty(\mathbb{R}^n)$

$$\langle \mu_j, \psi \rangle = \int \gamma\varphi_j(x')\psi(x', 0) dx' \rightarrow \int \gamma u(x')\psi(x', 0) dx',$$

les $\gamma\varphi_j$ tendant vers γu dans $H^{1/2}$ et donc au sens des distributions dans $\mathbb{R}^{n-1}$. Le membre de droite étant égal à $\langle \mu, \psi \rangle$, l'égalité (10.12) entraîne donc (10.11) par passage à la limite. Le cas des dérivées tangentielles est plus facile et laissé au lecteur.

Corollaire 10.3.7. — Pour qu'un élément $u \in H^1(\mathbb{R}_+^n)$ appartienne à $H_0^1(\mathbb{R}_+^n)$, il faut et il suffit que $\gamma u = 0$.

Nous savons que $u \in H_0^1(\mathbb{R}_+^n)$ peut être approché en norme H^1 par des éléments $\varphi_j \in C_0^\infty(\mathbb{R}_+^n)$. Les traces des φ_j sont nulles, et convergent vers γu dans $H^{1/2}(\mathbb{R}^{n-1})$ d'après le théorème 10.3.5. On a donc $\gamma u = 0$.

Réciproquement, soit $u \in H^1(\mathbb{R}_+^n)$ tel que $\gamma u = 0$. D'après le théorème précédent, on a $\partial_i(Zu) = Z(\partial_i u) \in L^2$ pour $i = 1, \dots, n$. Cela signifie que $Zu \in H^1(\mathbb{R}^n)$, et donc que $u \in H_0^1(\mathbb{R}_+^n)$.

10.4. Problème de Dirichlet dans un ouvert régulier

Le problème de Dirichlet dans un ouvert consiste à déterminer une fonction harmonique (ou une solution de $\Delta u = f$) prenant des valeurs données à la frontière de l'ouvert. Ce problème se pose dans les multiples domaines de la physique mathématique où intervient le Laplacien. Une autre question

importante est le problème de Neumann, où on impose les valeurs de la dérivée normale (c'est à dire le flux du gradient du potentiel u) à la frontière.

Les méthodes hilbertiennes que nous avons appliquées dans la section 6.2C à l'équation $(\Delta - \lambda)u = f$ dans tout l'espace vont nous permettre de montrer l'existence et l'unicité de la solution du problème de Dirichlet. Cela nécessite ici non seulement l'intervention de l'espace H^1 dans l'ouvert, mais aussi l'étude des traces des éléments de cet espace à la frontière.

A. Traces

Soit Ω un ouvert borné régulier (définition 5.4.5) et soit $\partial\Omega$ sa frontière. Nous n'utiliserons que l'espace de Sobolev $H^1(\Omega)$ dont on démontre le caractère hilbertien en reprenant mot pour mot l'argument du théorème 6.2.2.

Définition 10.4.1. — On note $H^1(\Omega)$ l'espace des u appartenant à $L^2(\Omega)$ dont les dérivées $\partial_i u$ au sens des distributions appartiennent également à $L^2(\Omega)$.

C'est un espace de Hilbert pour le produit scalaire

$$(u | v)_1 = \int_{\Omega} \partial_i u \overline{\partial_i v} dx + \int_{\Omega} u \overline{v} dx$$

ou pour tout autre produit scalaire donnant une norme équivalente.

Il est possible de définir les espaces H^s sur la sous-variété $\partial\Omega$, et nous aurons besoin effectivement de l'espace $H^{1/2}(\partial\Omega)$. Nous admettrons les résultats principaux, et n'indiquons ici que le principe de la définition des $H^s(\partial\Omega)$ pour $s \geq 0$.

Soit f une fonction définie sur $\partial\Omega$, de carré sommable pour la mesure de surface. Soient (voir les définitions 5.4.6 et 5.4.7) $(y^z); B^z; \Phi^z; \chi^z$ respectivement des systèmes de coordonnées orthogonales, des "cylindres" recouvrant $\partial\Omega$, l'équation locale de $\partial\Omega$, une partition de l'unité associée. On peut alors écrire, dans les coordonnées (y^z) , la fonction $\chi^z(y)f(y)$ sous la forme $g^z(y_1^z, \dots, y_{n-1}^z)$. On dit alors que f appartient à $H^s(\partial\Omega)$ si toutes les fonctions g^z appartiennent à $H^s(\mathbb{R}^{n-1})$.

Il faut bien sûr vérifier que cette définition ne dépend pas des choix des $(y^z); B^z; \Phi^z; \chi^z$. L'ingrédient essentiel de cette vérification est un théorème qui, pour s non entier, n'a rien d'évident : si $u \in L^2$ est à support compact dans un ouvert U de $\mathbb{R}^n$, et si Φ est un difféomorphisme de U sur un ouvert U' , on a

$$u \in H^s(\mathbb{R}^n) \Leftrightarrow u \circ \Phi^{-1} \in H^s(\mathbb{R}^n).$$

Nous pouvons maintenant énoncer, sans démonstration, les théorèmes suivants. Ils se prouvent en se ramenant, par partition de l'unité et cartes locales comme ci-dessus, aux théorèmes analogues que nous avons démontrés pour $\Omega = \mathbb{R}_+^n$.

Théorème 10.4.2. — *L'application γ de restriction au bord, de $C^\infty(\overline{\Omega})$ dans $C^\infty(\partial\Omega)$ se prolonge en une application linéaire continue surjective, notée encore γ , de $H^1(\Omega)$ sur $H^{1/2}(\partial\Omega)$.*

Théorème 10.4.3. — *Soit $u \in H^1(\Omega)$. Les trois propriétés suivantes sont équivalentes :*

- (a) $\gamma u = 0$.
- (b) $Zu \in H^1(\mathbb{R}^n)$, où Zu désigne la fonction égale à u dans Ω et à 0 ailleurs,
- (c) il existe une suite φ_j d'éléments de $C_0^\infty(\Omega)$ telle que $\|u - \varphi_j\|_1 \rightarrow 0$.

On note $H_0^1(\Omega)$ l'espace des fonctions u vérifiant les conditions précédentes. C'est un sous-espace fermé, et donc hilbertien, de $H^1(\Omega)$.

Si les théorèmes précédents généralisaient des résultats valables dans le demi-espace, l'inégalité qui va suivre est particulière aux ouverts bornés (ou au moins aux ouverts contenus dans une bande).

Théorème 10.4.4 (Inégalité de Poincaré). — *Il existe une constante C telle que*

$$\forall u \in H_0^1(\Omega), \quad \|u\|_{L^2} \leq C \left(\int_{\Omega} |\nabla u(x)|^2 dx \right)^{1/2}. \quad (10.13)$$

Considérons d'abord le cas d'une fonction $\varphi \in C_0^\infty(\Omega)$. Soit $R > 0$ tel que l'on ait $\Omega \subset \{x \mid -R \leq x_n \leq R\}$. On a alors

$$\varphi(x', x_n) = \int_{-R}^{x_n} \partial_n \varphi(x', t) dt.$$

L'inégalité de Cauchy-Schwarz, en majorant le domaine d'intégration par $[-R, R]$, donne

$$|\varphi(x', x_n)|^2 \leq \left(\int_{-R}^R 1 dt \right) \int_{-R}^R |\partial_n \varphi(x', t)|^2 dt.$$

On obtient donc

$$\begin{aligned} \iint_{\Omega} |\varphi(x', x_n)|^2 dx' dx_n &\leq \iint \int_{|x_n| \leq R: |t| \leq R} 2R |\partial_n \varphi(x', t)|^2 dt dx' dx_n \\ &\leq 4R^2 \iint_{\Omega} |\partial_n \varphi(x', t)|^2 dx' dt, \end{aligned}$$

et l'inégalité voulue, avec $C = 2R$, lorsque $\varphi \in C_0^\infty(\Omega)$.

Si maintenant $u \in H_0^1(\Omega)$, il existe une suite de fonctions φ_j appartenant à $C_0^\infty(\Omega)$ telles que $\|u - \varphi_j\|_1 \rightarrow 0$. L'inégalité (10.13) est valable pour chaque φ_j , et chacun des deux membres converge vers l'expression correspondante pour u , ce qui établit le résultat.

10.4.5. Notations. — Pour u et v appartenant à $H^1(\Omega)$, on posera

$$B(u, v) = \sum_i \int_{\Omega} \partial_i u \overline{\partial_i v} \, dx,$$

et on appellera *énergie* de u la quantité

$$E(u) = B(u, u) = \int_{\Omega} |\nabla u(x)|^2 \, dx.$$

La quantité $E(u)^{1/2}$ n'est pas une norme sur $H^1(\Omega)$, les constantes ayant une énergie nulle. On a par contre le résultat suivant.

Corollaire 10.4.6. — *Sur l'espace $H_0^1(\Omega)$, la forme $B(u, v)$ est un produit scalaire, et la norme $E(u)^{1/2}$ est équivalente à la norme $\|u\|_1$. En particulier, $H_0^1(\Omega)$ muni de B est un espace de Hilbert.*

On a en effet

$$\|u\|_1^2 = E(u) + \|u\|_{L^2}^2,$$

et, pour $u \in H_0^1(\Omega)$, on a d'après l'inégalité de Poincaré

$$E(u) \leq \|u\|_1^2 \leq (1 + C^2)E(u).$$

B. Problème de Dirichlet homogène

Nous ne considérerons plus que des fonctions à valeurs réelles. La forme bilinéaire B est alors symétrique.

Lemme 10.4.7. — *Soit $u \in H^1(\Omega)$. Il existe alors un élément $\tilde{u} \in H_0^1(\Omega)$ tel que*

$$\forall v \in H_0^1(\Omega), \quad B(u, v) = B(\tilde{u}, v). \quad (10.14)$$

Considérons en effet la forme linéaire $L(v) = B(u, v)$ sur $H_0^1(\Omega)$. On a $|L(v)| \leq E(u)^{1/2} E(v)^{1/2}$. C'est donc une forme linéaire continue sur l'espace de Hilbert $H_0^1(\Omega)$ muni du produit scalaire B , et il existe donc un élément $\tilde{u}$ de l'espace H_0^1 tel que $L(v) = B(\tilde{u}, v)$ pour tout $v \in H_0^1$.

Théorème 10.4.8. — *Soit $g \in H^{1/2}(\partial\Omega)$.*

(a) Il existe une et une seule fonction $u \in H^1(\Omega)$ solution du problème de Dirichlet

$$\begin{aligned}\Delta u &= 0, & \text{dans } \Omega \\ \gamma u &= g.\end{aligned}$$

(b) La solution u est, parmi les éléments de $H^1(\Omega)$ dont la trace est égale à g , l'unique fonction dont l'énergie soit minimum.

L'opérateur de trace étant surjectif, il existe $u_1 \in H^1(\Omega)$ telle que $\gamma u_1 = g$. Soit $\widetilde{u}_1$ la fonction fournie par le lemme précédent. Elle appartient à H_0^1 et est donc de trace nulle, et la fonction $u = u_1 - \widetilde{u}_1$ vérifie donc $\gamma u = g$.

L'égalité (10.14) appliquée à u_1 donne immédiatement que $B(u, v) = 0$ pour tout $v \in H_0^1$. On a donc en particulier pour $\varphi \in C_0^\infty(\Omega)$

$$B(u, \varphi) = \sum_i \int \partial_i u \partial_i \varphi \, dx = \sum_i \langle \partial_i u, \partial_i \varphi \rangle = 0,$$

et donc

$$\langle \Delta u, \varphi \rangle = 0.$$

Cela assure que u est une solution du problème de Dirichlet.

Démontrons l'unicité. La différence U de deux solutions doit être un élément de H^1 de trace nulle, et donc appartenir à H_0^1 , et doit vérifier $\Delta U = 0$. On a alors, pour toute fonction $\varphi \in C_0^\infty(\Omega)$

$$\langle \Delta U, \varphi \rangle = - \sum_i \int \partial_i U \partial_i \varphi \, dx = 0.$$

On a $B(U, \varphi) = 0$ pour tout $\varphi \in C_0^\infty(\Omega)$. L'élément U de l'espace de Hilbert H_0^1 , qui doit être orthogonal au sous-espace dense $C_0^\infty(\Omega)$, est donc nécessairement nul.

Démontrons maintenant la partie (b) du théorème. Soit w un élément quelconque de $H^1(\Omega)$ dont la trace est égale à g . La différence $w - u$ a donc une trace nulle, et appartient à H_0^1 . On a

$$E(w) = B(w, w) = B(u, u) + 2B(u, w - u) + B(w - u, w - u).$$

Nous avons vu que u est orthogonal à tous les éléments de H_0^1 et en particulier à $w - u$. On obtient donc

$$E(w) = E(u) + E(w - u),$$

quantité toujours supérieure ou égale à $E(u)$, et qui ne lui est égale que si l'élément $w - u$ de H_0^1 a une énergie nulle, ce qui entraîne $w = u$.

C. Problème de Dirichlet non homogène

Définition 10.4.9. — On dit qu'une distribution f dans Ω appartient à l'espace $H^{-1}(\Omega)$ s'il existe une constante C telle que

$$\forall \varphi \in C_0^\infty(\Omega), \quad |\langle f, \varphi \rangle| \leq C E(\varphi)^{1/2}.$$

On voit facilement qu'un élément f de $L^2(\Omega)$ appartient à $H^{-1}(\Omega)$. On a en effet $|\langle f, \varphi \rangle| \leq \|f\|_{L^2} \|\varphi\|_{L^2} \leq C \|u\|_{L^2} E(\varphi)^{1/2}$ d'après l'inégalité de Poincaré. De même, toujours pour $f \in L^2(\Omega)$, on a $\partial_i f \in H^{-1}(\Omega)$: on a en effet

$$|\langle \partial_i f, \varphi \rangle| = |\langle f, \partial_i \varphi \rangle| \leq \|f\|_{L^2} \|\partial_i \varphi\|_{L^2} \leq \|f\|_{L^2} E(\varphi)^{1/2}.$$

Nous verrons ci-dessous la réciproque de cette propriété.

Théorème 10.4.10. — Pour tout $f \in H^{-1}(\Omega)$, il existe un et un seul u appartenant à $H_0^1(\Omega)$ tel que l'on ait

$$\Delta u = -f.$$

La forme linéaire $\varphi \rightarrow \langle f, \varphi \rangle$ est définie sur le sous-espace dense $C_0^\infty(\Omega)$ de H_0^1 , et est continue pour la norme $E^{1/2}$ de cet espace. Elle se prolonge donc en une forme linéaire continue L sur H_0^1 entier. Il existe donc $u \in H_0^1$ tel que l'on ait $L(v) = B(u, v)$ pour tout $v \in H_0^1$. On a donc en particulier, pour $\varphi \in C_0^\infty(\Omega)$

$$\langle f, \varphi \rangle = B(u, \varphi) = \sum_i \langle \partial_i u, \partial_i \varphi \rangle = -\langle \Delta u, \varphi \rangle.$$

On a donc $\Delta u = -f$. L'unicité a déjà été démontrée dans la preuve du théorème 10.4.8 : la différence U de deux solutions appartiendrait à H_0^1 et vérifierait $\Delta U = 0$, et nous avons vu que cela impliquait $U = 0$.

Corollaire 10.4.11. — Tout élément $f \in H^{-1}(\Omega)$ est somme de dérivées d'éléments de $L^2(\Omega)$.

On sait en effet qu'il existe $u \in H_0^1$ avec $\Delta u = -f$. On peut donc écrire $f = -\sum_i \partial_i (\partial_i u)$, et les fonctions $\partial_i u$ appartiennent à L^2 .

Corollaire 10.4.12. — Soient $f \in H^{-1}(\Omega)$ et $g \in H^{1/2}(\partial\Omega)$. Il existe alors une et une seule solution $u \in H^1(\Omega)$ du problème de Dirichlet

$$\begin{aligned} \Delta u &= -f \\ \gamma u &= g. \end{aligned}$$

Il existe d'une part un élément $v \in H^1(\Omega)$ vérifiant $\Delta v = 0$ et $\gamma v = g$, d'autre part un élément $w \in H_0^1(\Omega)$ vérifiant $\Delta w = -f$. La fonction $u = v + w$ est une solution du problème posé.

Quant à l'unicité, elle résulte encore du théorème 10.4.8 : la différence U de deux solutions serait un élément de H_0^1 vérifiant $\Delta U = 0$.

10.4.13. Un peu d'histoire. — Les démonstrations précédentes reprennent d'assez près des idées de Gauss et de Riemann. Toutefois, ces mathématiciens qui furent des modèles de rigueur dans d'autres branches des mathématiques étaient moins stricts (comme beaucoup de leurs contemporains) en physique mathématique. Riemann, pour démontrer notre théorème 10.4.8, considère une fonction u dont l'énergie est minimum parmi les fonctions égales à g sur $\partial\Omega$, et prouve très correctement que u est alors une solution. Par contre, il ne démontre pas que le minimum est effectivement atteint.

Des démonstrations rigoureuses, utilisant des idées différentes, ont été données par Neumann et Poincaré, et Hilbert prouve vers 1900 que le minimum de l'énergie est effectivement atteint. Ce n'est toutefois que depuis 1950, avec l'introduction systématique par J. Deny des espaces du type H^1 et H^{-1} en théorie du potentiel, que l'on peut utiliser presque telles quelles les idées de Gauss et de Riemann.

L'exercice suivant reprend les idées d'un mémoire de Gauss (existence du potentiel d'équilibre) datant de 1843.

Exercice 10.4.14. — Soient Ω un ouvert borné régulier, et ω un ouvert relativement compact dans Ω .

(a) Soit C l'ensemble des $v \in H_0^1$ vérifiant $v(x) \geq 1$ presque partout sur ω . Démontrer qu'il existe un et un seul élément u de C dont l'énergie soit minimum. Montrer que l'on a $B(u, v) = 0$ pour tout $v \in H_0^1(\Omega)$ s'annulant sur ω et en déduire que $\Delta u = 0$ dans $\Omega \setminus \overline{\omega}$.

(b) On admettra le résultat suivant : si F est une application contractante (c'est-à-dire telle que $|F(s) - F(t)| \leq |s - t|$) de $\mathbb{R}$ dans lui-même s'annulant à l'origine, et si $v \in H_0^1(\Omega)$, alors $F \circ v$ appartient au même espace, et on a $E(F \circ v) \leq E(v)$. Il s'agit d'un résultat assez naturel (si $|w(x) - w(y)| \leq |v(x) - v(y)|$ quels que soient x et y le gradient de w en chaque point doit être plus petit en norme que celui de v), mais dont la démonstration nécessite des résultats fins sur les rapports entre dérivées usuelles et dérivées au sens des distributions.

Démontrer que la fonction u est presque partout comprise entre 0 et 1, et qu'elle est égale (p.p.) à 1 sur ω . On pourra considérer $F(t) = \max(0, \min(1, t))$.

(c) On pose $\mu = -\Delta u$. Démontrer que μ qui appartient à $H^{-1}(\Omega)$ est une mesure de Radon positive dont le support est contenu dans la frontière de ω . On écrira que $E(u + \varepsilon\varphi) \geq E(u)$ pour $\varphi \in C_0^\infty(\Omega)$ positive.

(d) On pose $C(\omega) = E(u)$. Montrer que $C(\omega)$ est aussi égal à la masse de μ (c'est-à-dire à $(\mu, 1)$).

En électrostatique, la fonction u représente le potentiel créé par un conducteur chargé ($\overline{\omega}$) en équilibre (énergie minimum). Ce potentiel est constant (normalisé ici à 1) sur le conducteur et harmonique en dehors. Les charges μ qui l'engendrent se répartissent à la frontière du conducteur. Le rapport $C(\omega)$ entre la masse et le potentiel créé s'appelle la capacité newtonienne (ou coulombienne) de ω . Le concept général de capacité (G. Choquet,

1953) est une extension de la notion de mesure ensembliste. Le lecteur pourra montrer facilement ici que l'application $\omega \mapsto C(\omega)$ est croissante et sous-additive [$C(\omega_1 \cup \omega_2) \leq C(\omega_1) + C(\omega_2)$], et plus difficilement qu'elle est fortement sous-additive [$C(\omega_1 \cup \omega_2) + C(\omega_1 \cap \omega_2) \leq C(\omega_1) + C(\omega_2)$].

D. Vers l'analyse spectrale

En remplaçant l'opérateur Δ par l'opérateur $\Delta - \lambda$, on obtient pour $\lambda > 0$ les mêmes résultats d'existence et d'unicité. Il suffit de remplacer la forme bilinéaire B par $B_\lambda(u, v) = B(u, v) + \lambda \int uv \, dx$.

Pour $\lambda < 0$ la situation est différente, et on peut démontrer le résultat suivant : il existe un ensemble

$$S = \{-\mu_1, \dots, -\mu_j, \dots\}$$

où la suite des μ_j est croissante et tend vers $+\infty$ tel que

- (a) Pour $\lambda \notin S$, l'opérateur $\Delta - \lambda$ est bijectif de $H_0^1(\Omega)$ sur $H^{-1}(\Omega)$.
- (b) Pour $\lambda = -\mu_j \in S$, l'espace E_j des solutions $u \in H_0^1$ de $(\Delta - \lambda)u = 0$ est de dimension finie $d_j \geq 1$.
- (c) Si on choisit, pour chaque j une base orthonormale $\varphi_{j,1} \dots \varphi_{j,d_j}$ (pour le produit scalaire de L^2) de l'espace E_j , alors la collection des $\varphi_{j,k}$ est une base hilbertienne de L^2 .

On appelle S le spectre du Laplacien (ou plus précisément de l'opérateur dans Ω associé au Laplacien et à la condition de nullité au bord). Les $-\mu_j$ sont les valeurs propres de Δ , et les éléments de E_j sont les fonctions propres associées.

La démonstration de ces propriétés, qui évoquent bien entendu les propriétés des matrices hermitiennes, repose sur des résultats d'analyse fonctionnelle abstraite, et sur deux propriétés du Laplacien. La première (caractère autoadjoint ou hermitien du Laplacien) est la formule de Green $\int u \Delta v \, dx = \int v \Delta u \, dx$ étendue au cas où u et v appartiennent à H_0^1 . La seconde est que l'opérateur G qui à $f \in L^2$ associe la solution $u = Gf \in H_0^1$ de $\Delta u = f$ est un opérateur compact de L^2 dans lui-même, c'est-à-dire qu'il transforme les boules de L^2 en des ensembles relativement compacts de L^2 .

Pour de tels opérateurs G , on a une bonne théorie spectrale, très proche de celle des endomorphismes hermitiens en dimension finie : existence de valeurs propres, et d'une base orthonormale de fonctions propres. Les valeurs propres de Δ sont alors les inverses de celles de G , les fonctions propres étant les mêmes.

Si φ est une fonction propre relative à la valeur propre $-\mu$, il est facile de vérifier que la fonction $F(t, x) = \varphi(x)e^{-\mu t}$ est une solution de

$$\partial_t F = \Delta F \quad \text{dans } [0, \infty[\times \Omega,$$

que la fonction $G(t, x) = \varphi(x)e^{i\mu t}$ est une solution de

$$i\partial_t G = \Delta G \quad \text{dans } \mathbb{R} \times \Omega,$$

et que les fonctions $H_\pm(t, x) = \varphi(x)e^{\pm i\sqrt{\mu}t}$ sont des solutions de

$$\partial_t^2 H = \Delta H \quad \text{dans } \mathbb{R} \times \Omega.$$

Cela permet de résoudre le "problème de Cauchy dans Ω avec conditions de Dirichlet", qui se formule ainsi pour l'équation de la chaleur (nous laissons au lecteur le soin d'écrire

l'analogue pour les équations des ondes et de Schrödinger)

$$\begin{aligned}\partial_t u(t, x) - \Delta u(t, x) &= 0 && \text{dans } [0, \infty[\times \Omega \\ u(t, x) &= 0 && \text{dans } [0, \infty[\times \partial\Omega \\ u(0, x) &= u_0(x) && \text{dans } \Omega\end{aligned}$$

On décompose la fonction u_0 donnée sur la base hilbertienne des $\varphi_{j,k}$. On a donc $u_0(x) = \sum \alpha_{j,k} \varphi_{j,k}(x)$, et on obtient la solution u par

$$u(t, x) = \sum \alpha_{j,k} \varphi_{j,k}(x) e^{-\mu_j t},$$

l'appartenance à H_0^1 pour chaque t traduisant la vérification de la seconde propriété. Bien entendu, il faudrait prouver la convergence de la série ci-dessus dans H_0^1 , ce qui n'est d'ailleurs pas difficile.

Le lecteur connaît bien un cas particulier de la méthode ci-dessus. C'est celui de la dimension 1 d'espace. Si Ω est l'intervalle $]0, \pi[$, les fonctions propres (dans H_0^1) de l'opérateur d^2/dx^2 sont les fonctions $C \sin jx$ et les valeurs propres sont égales à $-\mu_j = -j^2$. On retrouve la méthode dite "de séparation des variables".

En toute dimension, la solution du problème de Cauchy pour l'équation des ondes s'écrit comme somme de fonctions périodiques en t , dont les fréquences sont les $\sqrt{\mu_j}$. En dimension 1 d'espace, ces fréquences sont multiples d'une même fréquence fondamentale. Il n'en est plus de même en dimension supérieure (membranes et volumes vibrants) où le spectre n'a plus de structure aussi simple. L'étude des relations entre la géométrie de Ω (ou d'une variété riemannienne) et les valeurs propres du Laplacien fait toujours l'objet de recherches actives.

10.5. Problème de Cauchy et semi-groupes

Nous avons vu dans la section 6.6 comment la transformation de Fourier partielle donne l'expression des solutions $u(t, x)$ des équations des ondes, de la chaleur et de Schrödinger, connaissant les données de Cauchy (dans $\mathcal{S}(\mathbb{R}^n)$) à l'instant 0.

Nous allons nous intéresser ici à la famille, indexée par t , des opérateurs N_t qui, aux traces de la solution u à l'instant 0, font correspondre les traces de u à l'instant t .

Nous verrons d'une part que ces opérateurs N_t se prolongent à des espaces fonctionnels plus gros que $\mathcal{S}$, ce qui permet de résoudre le problème de Cauchy sous des hypothèses plus générales. D'autre part, nous constaterons la conservation de certaines normes au cours de l'évolution, ce qui a une signification physique précise. Enfin, nous verrons que la famille des N_t constitue un (semi)-groupe à un paramètre. Il s'agit là d'un phénomène d'une grande généralité, dont nous expliquerons les raisons.

La plupart des calculs (faciles) seront laissés au lecteur à titre d'exercice.

10.5.1. Équation de Schrödinger. — Nous avons obtenu la solution du problème de Cauchy

$$\begin{aligned}(i\partial_t + \Delta)u &= 0 \\ u(0, x) &= u_0(x)\end{aligned}$$

sous la forme $\tilde{u}(t, \xi) = e^{-it|\xi|^2} \widehat{u_0}(\xi)$, où $\tilde{u}$ désigne la transformée de Fourier partielle de u .

Cela nous amène à définir l'opérateur, de $\mathcal{S}(\mathbb{R}^n)$ dans lui-même

$$P_t = \mathcal{F}^{-1} [e^{-it|\xi|^2} \cdot] \mathcal{F}, \quad (10.15)$$

où on a noté $[e^{-it|\xi|^2} \cdot]$ l'opérateur de multiplication par $e^{-it|\xi|^2}$. Nous avons vu que l'opérateur P_t est l'opérateur de convolution (dans $\mathbb{R}^n$, à t fixé) par $(4\pi t)^{-n/2} e^{-in\pi/4} e^{i|x|^2/4t}$. Le lecteur vérifiera les points suivants.

— L'opérateur P_t défini par (10.15) est une isométrie de $H^s(\mathbb{R}^n)$ sur lui-même pour tout s .

— Pour $u_0 \in H^s(\mathbb{R}^n)$, l'application $t \mapsto P_t u_0$ est continue de $\mathbb{R}$ dans $H^s(\mathbb{R}^n)$ (on déterminera la limite de $\int \left| \widehat{P_{t+h} u_0} - \widehat{P_t u_0} \right|^2 (1 + |\xi|^2)^s d\xi$ à l'aide du théorème de Lebesgue).

— L'opérateur P_0 est l'identité et on a $P_t \circ P_{t'} = P_{t+t'}$ pour $t, t' \in \mathbb{R}$, ce qui est conséquence immédiate de la relation $e^{-it|\xi|^2} e^{-it'|\xi|^2} = e^{-i(t+t')|\xi|^2}$. On dit que (P_t) est un groupe à un paramètre d'opérateurs, c'est-à-dire un homomorphisme de groupes de $\mathbb{R}$ muni de l'addition dans l'ensemble des opérateurs muni de la composition.

La première propriété, pour $s = 0$, a un sens physique bien connu. Si u_0 est normalisée par $\|u_0\|_{L^2} = 1$, alors la probabilité totale de présence dans l'espace entier de la particule à l'instant t , égale à $\int |P_t u_0|^2 dx$, reste égale à 1.

Si $u_0 \in H^s(\mathbb{R}^n)$, on vérifie facilement en utilisant la seconde propriété que, pour $\varphi \in C_0^\infty(\mathbb{R}^{n+1})$, l'application $t \mapsto \langle P_t u_0(x), \varphi(t, x) \rangle$ est continue, ce qui permet de définir la distribution (que nous noterons $u(t, x)$) suivante dans $\mathbb{R}^{n+1}$

$$\langle u(t, x), \varphi(t, x) \rangle = \int \langle P_t u_0(x), \varphi(t, x) \rangle dt,$$

où les crochets désignent respectivement la dualité en dimension $n+1$ et n .

En approchant u_0 par des $\varphi_j \in \mathcal{S}$, on montre facilement que u est limite, au sens des distributions, des solutions u_j correspondantes, ce qui montre que u est solution de $(i\partial_t + \Delta)u = 0$. La trace pour $t = 0$ (dont la définition est immédiate pour une distribution donnée par une fonction continue de t à valeurs dans un espace fonctionnel) est bien égale à u_0 . Pour être complet, il faudrait

dire en quel sens une telle solution est unique, ce que nous n'aborderons pas ici.

10.5.2. Équation de la chaleur. — Pour $t \geq 0$, nous avons déterminé une solution du problème de Cauchy

$$\begin{aligned}(\partial_t - \Delta)u &= 0 \\ u(0, x) &= u_0(x)\end{aligned}$$

par $\tilde{u}(t, \xi) = e^{-t|\xi|^2} \widehat{u_0}(\xi)$.

Cela nous amène à définir la famille d'opérateurs

$$G_t = \mathcal{F}^{-1} [e^{-t|\xi|^2} \cdot] \mathcal{F}.$$

On sait que l'on a

$$G_t u_0 = \left[(4\pi t)^{-n/2} e^{-|x|^2/4t} \right] \star u_0. \quad (10.16)$$

Le lecteur vérifiera les points suivants.

— Pour $t \geq 0$, l'opérateur G_t est continu de $H^s(\mathbb{R}^n)$ dans lui-même. En outre, pour $t > 0$, il est continu de H^s dans H^{s+N} pour tout N (il y a régularisation immédiate de la solution dès le début de l'évolution).

— Pour $u_0 \in H^s$, l'application $t \mapsto G_t u_0$ est continue de $[0, \infty[$ dans H^s .

— Si u_0 est positive et sommable, il en est de même de $G_t u_0$ et on a $\|G_t u_0\|_{L^1} = \|u_0\|_{L^1}$. Cela résulte facilement de (10.16). Le sens physique est clair : l'intégrale de la température est, à une constante (la chaleur volumique) près, la quantité de chaleur totale, qui est conservée au cours de l'évolution.

— On a $G_t \circ G_{t'} = G_{t+t'}$ pour $t, t' \geq 0$, et G_0 est l'identité. On dit que l'on a un semi-groupe d'opérateurs.

Comme pour l'équation de Schrödinger, on montre que, si u_0 appartient à H^s ou à L^1 la famille des $G_t u_0$ définit une solution du problème de Cauchy dans $[0, \infty[\times \mathbb{R}^n$.

10.5.3. Équation des ondes. — Il faut cette fois-ci se donner u_0 et v_0 dans $\mathcal{S}(\mathbb{R}^n)$ pour déterminer une solution du problème de Cauchy

$$\begin{aligned}\Box u(t, x) &= 0 \\ u(0, x) &= u_0(x) \\ \partial_t u(0, x) &= v_0(x).\end{aligned}$$

Nous avons vu que la transformée de Fourier partielle de la solution était donnée par $\widehat{u_0}(\xi) \cos(|\xi|t) + \widehat{v_0}(\xi) \sin(|\xi|t)/|\xi|$. On a donc $\partial_t \tilde{u}(t, \xi) =$

$-\widehat{u_0}(\xi)|\xi|\sin(|\xi|t) + \widehat{v_0}(\xi)\cos(|\xi|t)$. Cela amène à définir la famille des opérateurs suivants, appliquant l'espace $\mathcal{S}(\mathbb{R}^n) \times \mathcal{S}(\mathbb{R}^n)$ dans lui-même, qui aux données de Cauchy à l'instant 0 associent les données de Cauchy à l'instant t .

$$\mathcal{U}_t = \begin{pmatrix} \mathcal{F}^{-1} & 0 \\ 0 & \mathcal{F}^{-1} \end{pmatrix} \begin{pmatrix} \cos(|\xi|t) & \sin(|\xi|t)/|\xi| \\ -|\xi|\sin(|\xi|t) & \cos(|\xi|t) \end{pmatrix} \begin{pmatrix} \mathcal{F} & 0 \\ 0 & \mathcal{F} \end{pmatrix}$$

Le lecteur vérifiera facilement les propriétés suivantes.

- $\mathcal{U}_t$ applique continûment $H^{s+1}(\mathbb{R}^n) \times H^s(\mathbb{R}^n)$ dans lui-même pour tout s .
- Si $W = (u_0, v_0) \in H^{s+1}(\mathbb{R}^n) \times H^s(\mathbb{R}^n)$, alors l'application $t \mapsto \mathcal{U}_t W$ est continue de $\mathbb{R}$ dans ce même espace.
- Posons, pour $W_0 = (u_0, v_0) \in H^1(\mathbb{R}^n) \times L^2(\mathbb{R}^n)$

$$E(W_0) = \int |\nabla u_0|^2 dx + \int |v_0|^2 dx.$$

On a alors $E(\mathcal{U}_t W_0) = E(W_0)$. Au facteur $1/2$ près, la première intégrale représente l'énergie potentielle, et la seconde l'énergie cinétique, et la relation exprime la conservation de l'énergie E au cours de l'évolution.

- On a $\mathcal{U}_0 = I$ et $\mathcal{U}_t \circ \mathcal{U}_{t'} = \mathcal{U}_{t+t'}$, et donc encore un groupe à un paramètre.

Remarque 10.5.4. — Pour des opérateurs analogues, mais à coefficients variables, on ne peut utiliser les démonstrations ci-dessus à base de transformation de Fourier. Par contre, il est possible de démontrer directement les estimations de conservation de l'énergie, ce qui est une étape importante pour prouver l'existence de solutions.

10.5.5. Pourquoi des (semi-)groupes. — Le lecteur sera sans doute — du moins nous l'espérons — insatisfait des démonstrations précédentes conduisant à l'existence de semi-groupes. Il préférera certainement, à ces calculs simples mais aveugles, une raison générale non moins simple pourvue d'une vertu explicative.

Le cadre général est le suivant. On se donne un ensemble E , représentant l'ensemble des états possibles d'un système, et une loi d'évolution qui, parmi toutes les applications d'un intervalle de $\mathbb{R}$ dans E , sélectionne certaines d'entre elles appelées évolutions (possibles) du système. On suppose que, si $J \subset I$ sont deux intervalles de $\mathbb{R}$, et si $t \mapsto e(t)$ est une évolution dans I , alors la restriction de $e(\cdot)$ à J est une évolution dans J .

Les exemples rentrant dans ce cadre sont multiples, citons :

- E est un ouvert $\Omega \subset \mathbb{R}^n$ (ou plus généralement une variété différentielle), et les évolutions du système sont les solutions d'une équation différentielle du

premier ordre, s'écrivant $dX/dt = F(t, X(t))$.

— E est de la forme $\Omega \times \mathbb{R}^n$ où Ω est un ouvert de $\mathbb{R}^n$ (ou plus généralement, E est le fibré tangent $T\Omega$ d'une variété différentielle Ω). Les évolutions sont les solutions d'une équation différentielle du second ordre sur Ω qui s'écrit $d^2X/dt^2 = G(t, X, dX/dt)$, et que l'on peut réécrire $dX/dt = V$, $dV/dt = G(t, X, V)$.

— E est un espace fonctionnel sur $\mathbb{R}^n$ ($\mathcal{S}, \mathcal{S}', H^s, L^1, \dots$) et les évolutions du système sont les solutions d'une équation aux dérivées partielles du premier ordre par rapport à t (ou plus généralement d'une équation du type $\partial_t u(t, x) = \mathcal{A}_t u(t, x)$ où $\mathcal{A}_t$ est pour chaque t une application entre espaces fonctionnels dans $\mathbb{R}^n$).

— E est un produit d'espaces fonctionnels dans $\mathbb{R}^n$ ($H^{s+1} \times H^s$ par exemple), et la loi d'évolution est une équation aux dérivées partielles du second ordre par rapport à t .

Nous serons amenés à examiner les conséquences des deux propriétés suivantes.

(a) *Déterminisme*. — Pour tout $e_0 \in E$ et pour tout $t_0 \in \mathbb{R}$, il existe une et une seule évolution $e(\cdot)$ du système, définie dans $\mathbb{R}$ entier (variante : dans $[t_0, \infty[$) telle que $e(t_0) = e_0$.

Bien entendu, une telle propriété n'a rien d'automatique. Lorsqu'elle est vraie, il faut le plus souvent prouver un théorème d'existence et d'unicité locale puis, par d'autres méthodes en général, prouver l'existence globale.

(b) *Homogénéité dans le temps*. — Si $e(t)$ est une évolution dans un intervalle, alors $e(t - T)$ est une évolution dans l'intervalle translaté.

Techniquement, cette hypothèse correspond au cas où les équations, comme on dit, "ne dépendent pas explicitement de t ". Dans le premier exemple ci-dessus, elle signifie que la fonction F ne dépend que de X ; dans le second, que G ne dépend que de X et V ; dans le troisième, que pour chaque t , l'opérateur $\mathcal{A}_t$ est égal à un opérateur $\mathcal{A}$ fixe.

Sous la seule hypothèse (a), on peut introduire les opérateurs R_s^t , pour $s, t \in \mathbb{R}$ (variante : pour $s \leq t$) définis comme suit : on considère l'unique évolution $e(\cdot)$ qui vérifie $e(s) = e_0$, et on pose $R_s^t(e_0) = e(t)$. C'est l'opérateur qui fait passer de l'état du système à l'instant s à l'état à l'instant t .

Il est clair que R_t^t est l'identité, et que l'on a

$$R_t^u \circ R_s^t = R_s^u$$

pour $s, t, u \in \mathbb{R}$ (variante : pour $s \leq t \leq u$). En effet, l'unique évolution $e(\cdot)$ du système vérifiant $e(s) = e_0$ est aussi l'unique évolution du système vérifiant $e(t) = R_s^t e_0$, et l'état $e(u)$ peut s'écrire indifféremment $R_s^u e_0$ ou $R_t^u (R_s^t e_0)$.

Si on ajoute l'hypothèse (b) on a de plus $R_{s+T}^{t+T} = R_s^t$ pour tout T , et en particulier $R_s^t = R_0^{t-s}$ pour s et t réels (variante : pour $s \leq t$). Si on pose $N_t = R_0^t$, on obtient

$$N_t \circ N_s = R_s^{s+t} \circ R_0^s = R_0^{s+t} = N_{s+t}$$

quels que soient s et t (variante : pour $s, t \geq 0$), et les N_t forment un groupe (variante : un semi-groupe) à un paramètre.

Dans les deux premiers exemples donnés, on obtient des groupes de difféomorphismes de Ω et de $\Omega \times \mathbb{R}^n$ respectivement. Dans les deux suivants, rien n'oblige l'équation aux dérivées partielles à être linéaire (il est toutefois plus rare que l'hypothèse (a) soit satisfaite dans le cas non linéaire). Lorsqu'elle l'est, les opérateurs N_t sont linéaires, et lorsqu'elle est à coefficients constants, les N_t qui commutent avec les translations dans l'espace sont des opérateurs de convolution dans $\mathbb{R}^n$. C'est ce que nous avons constaté pour les trois équations d'évolution que nous avons étudiées.

Pour les équations du type de l'équation des ondes, mais à coefficients variables, on obtient des groupes à un paramètre d'opérateurs intégraux de Fourier, opérateurs dont nous avons dit quelques mots dans la section 9.7.

APPENDICE A

COMPLÉMENTS DE CALCUL DIFFÉRENTIEL

A.1. Applications différentiables

Dans toute cette section, E et F désigneront des espaces vectoriels de dimension finie sur $\mathbb{R}$. On note $\mathcal{L}(E, F)$ l'espace des applications linéaires de E dans F .

Définition A.1.1. — Soient Ω un ouvert de E , x un point de Ω , et φ une application de Ω dans F . On dit que φ est différentiable en x s'il existe une application linéaire $L \in \mathcal{L}(E, F)$ telle que l'on ait, pour $h \in E$ assez voisin de 0,

$$\varphi(x + h) = \varphi(x) + L(h) + |h|_E \omega(h) \quad \text{avec} \quad |\omega(h)|_F \xrightarrow{h \rightarrow 0} 0. \quad (\text{A.1})$$

L'application L est alors unique et est appelée différentielle de φ au point x . Nous la noterons⁽¹⁾ $h \mapsto D\varphi(x) \cdot h$.

Cette définition fait apparemment appel au choix de normes sur E et F , mais il est clair que si on remplace une de ces normes par une autre norme (nécessairement équivalente), la propriété est inchangée. On dit aussi que $Df(x)$ est l'application linéaire tangente à φ au point x .

Définition A.1.2. — Soient Ω un ouvert de E et φ une application de Ω dans F . On dit que φ est continûment différentiable dans Ω , ou que φ est de classe C^1 dans Ω si φ est différentiable en tout point de Ω et si l'application $x \mapsto D\varphi(x)$ de Ω dans $\mathcal{L}(E, F)$ est continue.

⁽¹⁾Elle est notée systématiquement $h \mapsto h\varphi'(x)$ lorsque l'espace de départ est $\mathbb{R}$, et $h \mapsto d\varphi(x) \cdot h$ ou $\langle d\varphi(x), h \rangle$ lorsque l'espace d'arrivée est $\mathbb{R}$ ou $\mathbb{C}$. Dans le cas général, les notations varient selon les auteurs.

Théorème A.1.3. — Soient E, F et G des espaces vectoriels de dimension finie, Ω un ouvert de E , Ω' un ouvert de F , φ une application de Ω dans Ω' et ψ une application de Ω' dans G .

(a) Si φ est différentiable en un point $x_0 \in \Omega$ et si ψ est différentiable au point $y_0 = \varphi(x_0)$, la fonction $h = \psi \circ \varphi$ est différentiable en x_0 et on a $Dh(x_0) = D\psi(y_0) \circ D\varphi(x_0)$.

(b) Si φ et ψ sont de classe C^1 respectivement dans Ω et Ω' , la fonction $\psi \circ \varphi$ est de classe C^1 dans Ω .

Étant donnés deux ouverts Ω et Ω' de E et F respectivement, on appelle *difféomorphisme* (de classe C^1) une application φ bijective de Ω sur Ω' qui est de classe C^1 ainsi que son inverse $\psi = \varphi^{-1}$.

Les applications $\psi \circ \varphi$ et $\varphi \circ \psi$ sont alors les applications identiques de Ω et Ω' , dont la différentielle est l'identité en tout point, et on obtient donc, pour $x_0 \in \Omega$ et en posant $y_0 = \varphi(x_0)$

$$D\psi(y_0) \circ D\varphi(x_0) = I_E \quad D\varphi(x_0) \circ D\psi(y_0) = I_F.$$

Il en résulte que les espaces E et F ont nécessairement la même dimension, que $D\varphi(x_0)$ est une application linéaire inversible de E dans F , et que l'on a $D\psi(y_0) = (D\varphi(x_0))^{-1}$.

Le théorème suivant, d'une grande importance, fournit une réciproque locale aux énoncés qui précèdent.

Théorème A.1.4 (d'inversion locale). — Supposons E et F de même dimension et soit φ une application de classe C^1 d'un ouvert $\Omega \subset E$ dans F . Supposons que, en un point $x_0 \in \Omega$, la différentielle $D\varphi(x_0)$ soit inversible. Il existe alors un voisinage ouvert V de x_0 contenu dans Ω et un voisinage ouvert W de $\varphi(x_0)$ tels que φ soit un difféomorphisme de V sur W .

Remarque A.1.5. — Il est bien sûr nécessaire de se restreindre à un voisinage de x_0 pour avoir la garantie que $D\varphi(x)$ reste inversible, mais ce n'est pas la seule raison qui rende cette restriction nécessaire. Une application surjective de Ω sur Ω' dont la différentielle est inversible en tout point peut ne pas être injective. Un exemple classique est l'application $(r, \theta) \mapsto (r \cos \theta, r \sin \theta)$ de $]0, \infty[\times \mathbb{R}$ dans le plan privé de l'origine.

Exercice A.1.6. — Soit φ de classe C^1 et bijective de Ω sur Ω' dont la différentielle est inversible en tout point. Démontrer que φ est un difféomorphisme de Ω sur Ω' .

Exercice A.1.7. — Soit φ une application de classe C^1 de Ω dans F dont la différentielle est inversible en tout point. Démontrer que pour tout ouvert $U \subset \Omega$, son image $\varphi(U)$ est un ouvert.

A.1.8. Différentielles et coordonnées locales. — Soit $(e_1, \dots, e_n)$ une base de E et $(f_1, \dots, f_p)$ une base de F . On peut alors identifier E et $\mathbb{R}^n$ par l'application $(x^1, \dots, x^n) \mapsto \sum x^i e_i$ et identifier de même F et $\mathbb{R}^p$. Une application φ d'un ouvert de E dans F s'identifie alors à une famille de p fonctions de n variables $\varphi^j(x^1, \dots, x^n)$, $j = 1, \dots, p$ c'est-à-dire que l'on a $\varphi(\sum x^i e_i) = \sum \varphi^j(x^1, \dots, x^n) f_j$. L'application φ est différentiable en un point ou de classe C^1 dans un ouvert si et seulement s'il en est de même des applications composantes φ^j . En particulier, φ est de classe C^1 si et seulement si les dérivées partielles $\partial \varphi^j / \partial x^i$ existent et sont continues.

La définition A.1 devient alors

$$\varphi^j(x^1 + h^1, \dots, x^n + h^n) = \varphi^j(x^1, \dots, x^n) + \sum_{i=1}^n \frac{\partial \varphi^j}{\partial x^i}(x) h^i + o(h), \quad (\text{A.2})$$

pour $j = 1, \dots, p$. La matrice $(\partial \varphi^j / \partial x^i(x))$, qui est la matrice de l'application linéaire $D\varphi(x)$ dans les bases considérées est appelée *matrice jacobienne* de φ , son déterminant étant appelé le *jacobien* de φ au point x .

A.1.9. Applications de classe C^k . — On dit que φ est de classe C^k si, des bases étant choisies, les fonctions composantes admettent des dérivées partielles continues jusqu'à l'ordre k . Il faut bien sûr vérifier que, si on change les bases de E et F , les nouvelles composantes $\tilde{\varphi}^j$ possèdent les mêmes propriétés. Les formules de changement de base fournissent immédiatement des relations

$$\tilde{\varphi}^j(x^1, \dots, x^n) = \sum_{l=1}^p \beta_l^j \varphi^l \left(\sum_{k=1}^n \alpha_k^1 x^k, \dots, \sum_{k=1}^n \alpha_k^n x^k \right)$$

et il suffit d'appliquer par récurrence le théorème A.1.3 dans un cas très simple (composition avec des applications linéaires).

On appelle *difféomorphisme* de classe C^k une bijection de classe C^k ainsi que son inverse. Dans l'énoncé du théorème d'inversion locale, si on suppose l'application φ de classe C^k , on obtient un difféomorphisme de classe C^k .

On peut en fait donner une présentation intrinsèque, indépendante des bases, des différentielles d'ordre supérieur. Si φ est de classe C^1 de Ω dans F , l'application $D\varphi : x \mapsto D\varphi(x)$ est une application de Ω dans $\mathcal{L}(E, F)$. On dit que φ est de classe C^2 si l'application $D\varphi$ est elle-même de classe C^1 , sa différentielle en un point x , notée $D(D\varphi)(x)$ ou $D^2\varphi(x)$ appartenant à $\mathcal{L}(E, \mathcal{L}(E, F))$. Si $D^2\varphi$ est continûment différentiable, on dit que φ est de classe C^3 , et on définit $D^3\varphi(x) \in \mathcal{L}(E, \mathcal{L}(E, \mathcal{L}(E, F)))$, etc.

Un élément de $\mathcal{L}(E, \mathcal{L}(E, F))$ s'identifie à une application bilinéaire de $E \times E$ dans F . D'après le lemme de Schwarz, pour φ de classe C^2 , l'application bilinéaire $D^2\varphi(x)$ est une application bilinéaire symétrique. De même, $D^3\varphi(x)$ s'identifie à une application trilinéaire symétrique, ... La formule de Taylor sous forme intrinsèque devient $\varphi(x + h) = \varphi(x) + D\varphi(x) \cdot h + 1/2! D^2\varphi(x)(h, h) + \dots$

A.2. Hypersurfaces

Un concept élémentaire, mais insuffisant pour décrire ne serait-ce qu'une sphère, est celui de graphe d'une application différentiable de $\mathbb{R}^{n-1}$ dans $\mathbb{R}$, une base étant choisie. Par contre, une des bonnes définitions des hypersurfaces les présentera comme des "recolllements" de graphes, ce qui ne peut avoir de sens qu'après avoir établi des propriétés d'invariance par changement de base. La clef de toutes les questions d'invariance sera le théorème d'inversion locale.

Dans toute cette section, E désigne un espace vectoriel sur $\mathbb{R}$ de dimension finie n .

Pour $x \in \mathbb{R}^n$, on notera souvent x' la projection $(x^1, \dots, x^{n-1})$ de x sur $\mathbb{R}^{n-1}$. Sauf indication contraire, une expression du type $\sum a_j b^j$ désignera une somme où l'indice répété j varie de 1 à n , tandis que $\sum' a_j b^j$ désigne la somme de 1 à $n-1$.

Définition A.2.1. — On dit qu'un sous-ensemble Γ de E est un graphe (de dimension $n-1$ et de classe C^1) s'il existe : 1. une base $(e_1, \dots, e_n)$ de E ; 2. un ouvert V de $\mathbb{R}^{n-1}$; 3. une application φ de classe C^1 de V dans $\mathbb{R}$; tels que l'on ait

$$\Gamma = \{x^1 e_1 + \dots + x^{n-1} e_{n-1} + \varphi(x') e_n \mid x' = (x^1, \dots, x^{n-1}) \in V\}.$$

On dira que le triplet $((e_j), V, \varphi)$ est une présentation du graphe Γ .

Théorème et Définition A.2.2 (Hyperplan tangent)

Soient Γ un graphe et m_0 un point de Γ . Il existe un unique hyperplan affine H de E passant par m_0 tel que l'on ait pour $p \in \Gamma$, et en notant d la distance pour une norme de E ,

$$d(p, H) = |p - m_0| \omega(p) \quad \text{avec} \quad \omega(p) \xrightarrow[p \rightarrow m_0, p \in \Gamma]{} 0. \quad (\text{A.3})$$

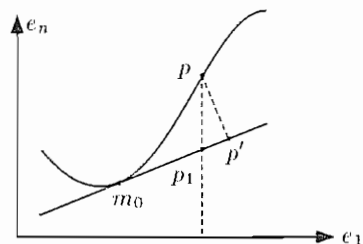
Si $((e_j), V, \varphi)$ est une présentation de Γ et si $x_0^1, \dots, x_0^n$ sont les coordonnées de m_0 dans la base (e_j) , l'hyperplan H a pour équation

$$x^n - x_0^n = \sum' \left(\frac{\partial \varphi}{\partial x^j}(x_0') \right) (x^j - x_0^j) \quad (\text{A.4})$$

dans cette même base.

Rappelons le principe de la démonstration que le lecteur connaît certainement, au moins en dimension 3. La propriété ne dépendant pas du choix de la norme, on peut choisir celle-ci euclidienne de telle sorte que la base (e_j) soit ortho-normale. Soit H l'hyperplan défini par (A.4). La formule (A.1), appliquée au point p se projetant en $x_0 + h$ assure que, en appelant p_1 le point de H de même projection, on a $|p - p_1| \leq |h| \omega(h)$ avec $\omega(h) \rightarrow 0$. Il est facile de voir que l'on a $|p - m_0| \geq C^{\text{te}} |h|$ et $|p - p'| = d(p, H) \leq |p - p_1|$, ce qui entraîne (A.3).

Réciproquement, si un hyperplan d'équation $x^n - x_0^n = \sum' \alpha_j (x^j - x_0^j)$ vérifie (A.3), le même calcul montre que l'on a $\varphi(x_0 + h) = \sum' \alpha_j h^j + o(h)$, et par unicité de la différentielle, on a nécessairement $\alpha^j = \partial\varphi/\partial x^j(x_0)$. Il reste à montrer qu'aucun hyperplan parallèle à e_n ne peut vérifier (A.3), ce qui est facile.



Remarque A.2.3. — On vérifie immédiatement, toujours dans la présentation précédente du graphe, que les $n - 1$ vecteurs $V_j = e_j + (\partial\varphi/\partial x^j(x'_0))e_n$ sont linéairement indépendants et sont parallèles à l'hyperplan tangent H .

Un graphe donné peut avoir de nombreuses présentations comme le montrent l'exercice et surtout le théorème suivants. La définition que nous en avons donnée montre que l'hyperplan tangent en un point ne dépend pas de la présentation.

Exercice A.2.4. — Soit Γ un graphe de présentation $((e_j), V, \varphi)$, et soit (f_j) une base telle que f_n soit proportionnel à e_n . Démontrer que Γ admet une présentation dans la base (f_j) . En déduire que, dans un espace E euclidien, un graphe admet toujours une présentation dans une base orthonormale.

Théorème A.2.5. — Soient Γ un graphe de présentation $((e_j), V, \varphi)$, m_0 un point de Γ et H l'hyperplan tangent en ce point. Soit $(f_1, \dots, f_n)$ une base de E telle que f_n ne soit pas parallèle à H et notons $y_0^1, \dots, y_0^n$ les coordonnées de m_0 dans cette base. Il existe alors un voisinage ω de m_0 dans E , un voisinage W de y_0^1 dans $\mathbb{R}^{n-1}$ et une application ψ de classe C^1 de W dans $\mathbb{R}$ tels que $\Gamma \cap \omega$ soit un graphe de présentation $((f_j), W, \psi)$.

La démonstration se fera en plusieurs étapes.

A.2.6. L'application θ de changement de coordonnées. — Étant donné un point $x' \in V$, on peut lui faire correspondre le point $P(x') = \sum' x^j e_j + \varphi(x') e_n$ qui appartient à Γ . Ce même point possède des coordonnées $(y^1, \dots, y^n)$ dans la base (f_j) qu'il est aisé de calculer. Soit en effet (α_j^k) la matrice de changement de base définie par $e_j = \sum \alpha_j^k f_k$, on a alors

$$P(x') = \sum_j' \sum_k x^j \alpha_j^k f_k + \sum_k \varphi(x') \alpha_n^k f_k,$$

et donc

$$y^k = \sum_j' \alpha_j^k x^j + \alpha_n^k \varphi(x') \quad , \quad k = 1, \dots, n. \quad (\text{A.5})$$

Nous appellerons θ l'application de V dans $\mathbb{R}^{n-1}$ qui à x' fait correspondre le point $y' = (y^1, \dots, y^{n-1})$ où les y^j sont données par (A.5).

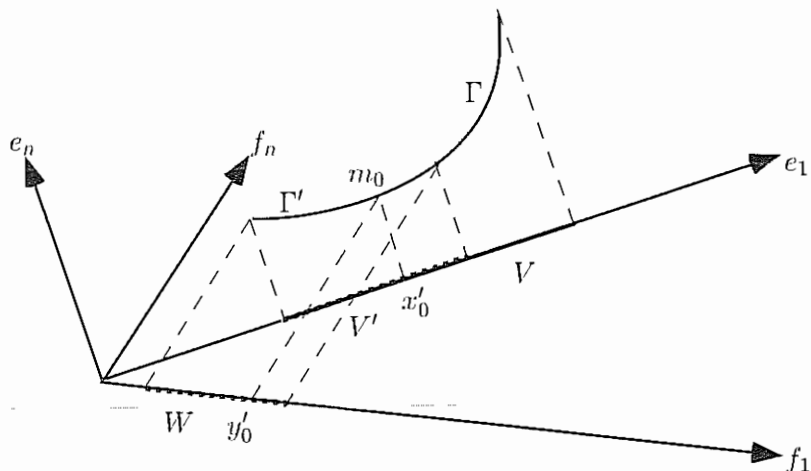


FIGURE 1. On a représenté, à la place des ouverts $V' \subset V$ et W qui vivent dans $\mathbb{R}^{n-1}$, leurs images dans E par $x' \mapsto \sum' x^j e_j$ et $y' \mapsto \sum' y^j f_j$.

A.2.7. *Calcul du jacobien de θ .* — Il est facile d'écrire le déterminant $J_\theta(x')$ de la matrice jacobienne de θ . On a

$$J_\theta(x') = \begin{vmatrix} \alpha_1^1 + \alpha_n^1 \partial\varphi/\partial x^1 & \cdots & \alpha_{n-1}^1 + \alpha_n^1 \partial\varphi/\partial x^{n-1} \\ \vdots & \ddots & \vdots \\ \alpha_1^{n-1} + \alpha_n^{n-1} \partial\varphi/\partial x^1 & \cdots & \alpha_{n-1}^{n-1} + \alpha_n^{n-1} \partial\varphi/\partial x^{n-1} \end{vmatrix},$$

ou encore, en développant par rapport à la dernière colonne,

$$J_\theta(x') = \begin{vmatrix} \alpha_1^1 + \alpha_n^1 \partial\varphi/\partial x^1 & \cdots & \alpha_{n-1}^1 + \alpha_n^1 \partial\varphi/\partial x^{n-1} & 0 \\ \vdots & \ddots & \vdots & 0 \\ \alpha_1^{n-1} + \alpha_n^{n-1} \partial\varphi/\partial x^1 & \cdots & \alpha_{n-1}^{n-1} + \alpha_n^{n-1} \partial\varphi/\partial x^{n-1} & 0 \\ \alpha_1^n + \alpha_n^n \partial\varphi/\partial x^1 & \cdots & \alpha_{n-1}^n + \alpha_n^n \partial\varphi/\partial x^{n-1} & 1 \end{vmatrix}.$$

L'intérêt de la seconde écriture est qu'elle fait apparaître en colonne les composantes, dans la base (f_j) , des vecteurs $e_j + (\partial\varphi/\partial x_j)e_n$, $j = 1, \dots, n-1$ et f_n . Avec les notations de la remarque A.2.3, on a donc

$$J_\theta(x'_0) = \det_{(f_j)}(V_1, \dots, V_{n-1}, f_n),$$

en notant ainsi le déterminant d'un système de vecteurs dans une base. En particulier, les V_j étant indépendants et le vecteur f_n n'étant pas, par hypothèse, combinaison linéaire des V_j , on a $J_\theta(x'_0) \neq 0$.

A.2.8. Fin de la démonstration du théorème A.2.5. — L'application θ de V dans $\mathbb{R}^{n-1}$ est de classe C^1 et a une différentielle inversible au point x'_0 . D'après le théorème d'inversion locale, il existe donc un voisinage ouvert $V' \subset V$ de x'_0 et un voisinage ouvert W de y'_0 tel que θ soit un difféomorphisme de V' sur W . On notera $\chi = \theta^{-1}$ le difféomorphisme inverse, et χ^j ses composantes. Posons

$$\Gamma' = \Gamma \cap \left\{ \sum x^j e_j \mid x' \in V' \right\}.$$

Dire que $m \in \Gamma'$ équivaut à dire qu'il existe $x' \in V'$ tel que $m = P(x')$. Il est donc équivalent de dire qu'il existe $y' \in W$ tel que $m = P \circ \chi(y')$. D'après les formules (A.5), il est encore équivalent d'affirmer que $m = \sum z^j f_j$ avec

$$z' = \theta \circ \chi(y') \quad z^n = \sum' \alpha_j^n \chi^j(y') + \alpha_n^n \varphi(\chi(y')) \quad y' \in W.$$

Il reste à poser $\psi(y') = \sum' \alpha_j^n \chi^j(y') + \alpha_n^n \varphi(\chi(y'))$ et, compte tenu de $\theta \circ \chi(y') = y'$, l'ensemble Γ' est constitué des points de la forme $\sum y^j f_j$ avec $y' \in W$ et $y^n = \psi(y')$, ce qui exprime exactement que Γ' est un graphe de présentation $((f_j), W, \psi)$.

Remarque A.2.9. — La démonstration montre que l'on peut prendre comme ouvert ω l'ensemble des $\sum x^j e_j$ avec $x' \in V'$. Il y a bien d'autre choix possibles, mais (voir figure) l'ensemble des points de Γ dont les composantes y' appartiennent à un voisinage, même très petit, de y'_0 ne forment pas nécessairement un graphe.

Remarque A.2.10. — On peut remplacer partout l'expression "de classe C^1 " par "de classe C^k ", $k = 1, 2, \dots, \infty$, dans la définition d'un graphe, dans le théorème précédent et dans la définition suivante.

Définition A.2.11. — On appelle hypersurface (de classe C^1) un sous-ensemble Σ de E tel que, pour tout $m \in \Sigma$, il existe un voisinage ouvert ω de m dans E tel que $\Sigma \cap \omega$ soit un graphe.

Les théorèmes (A.2.2) et (A.2.5) entraînent immédiatement l'existence de l'hyperplan tangent (avec exactement la même définition) en chaque point m de Σ , et le fait que, si une base (f_j) est telle que f_n n'est pas parallèle à cet hyperplan tangent, il existe un voisinage ω de m tel que $\Sigma \cap \omega$ est un graphe admettant une présentation dans (f_j) .

Exercice A.2.12 (Hypersurface définie par une équation implicite)

Soient Ω un ouvert de E et F une application de classe C^1 de Ω dans $\mathbb{R}$. On suppose qu'en tout point de Ω où F s'annule, sa différentielle dF ne soit pas nulle. Démontrer que $\Sigma = \{m \in \Omega \mid F(m) = 0\}$ est une hypersurface. En chaque point $m_0 \in \Sigma$, on choisira une base (e_j) telle que $\langle dF(m_0), e_n \rangle \neq 0$, et on appliquera le théorème d'inversion locale à l'application de Ω dans $\mathbb{R}^n$ qui

à m fait correspondre $(x^1, \dots, x^{n-1}, F(m))$, où les x^j sont les composantes de m dans la base (e_j) .

Démontrer que l'hyperplan tangent à Σ au point m est l'ensemble des $p \in E$ vérifiant $\langle dF(m), p - m \rangle = 0$.

A.3. Intégrale de surface

Dans toute cette section, E désignera un espace vectoriel *eucldien* de dimension n , dont le produit scalaire sera noté $(p, q) \mapsto p \cdot q$.

Il est facile de donner une définition d'apparence raisonnable de l'intégrale de surface sur un graphe muni d'une présentation. Cela dit, ce n'est qu'après avoir montré que la définition ne dépend pas de la présentation choisie que l'on est assuré d'avoir défini un concept géométrique intrinsèque. D'autre part, une fois prouvée cette invariance, les partitions de l'unité permettent facilement de définir l'intégrale de surface sur une hypersurface générale.

Nous conserverons les notations de la section précédente et notamment, pour un graphe muni de deux présentations, celles de la démonstration du théorème A.2.5 (présentations $((e_j), V, \varphi)$ et $((f_j), W, \psi)$, vecteurs V_j , applications P et $\theta, \dots$).

A.3.1. Considérations heuristiques. — Soit Γ un graphe ayant une présentation dans la base orthonormale (e_j) , le point $m_0 \in \Gamma$ ayant comme coordonnées les x_0^j . Pour faire image, nous utiliserons le vocabulaire géométrique (carré, parallélogramme, ...) correspondant au cas $n = 3$. Lorsque x' parcourt un petit "carré" K de mesure $((n-1)$ -dimensionnelle) v centré en x_0' , l'ensemble $P(K)$ est "proche" du parallélogramme que l'on obtiendrait en remplaçant Γ par son hyperplan tangent H en m_0 . La mesure de ce parallélogramme est égale à $v/|\alpha|$, où le rapport α est le cosinus de l'angle de l'hyperplan tangent avec l'hyperplan engendré par les e_j . $j = 1, \dots, n-1$, c'est-à-dire $\alpha(m_0) = \nu_0 \cdot e_n =$ où ν_0 est un vecteur unitaire normal à H .

Le vecteur $\sum' \partial\varphi/\partial x^j e_j - e_n$ étant normal à H , on obtient donc $|\alpha| = (1 + \sum' |\partial\varphi/\partial x^j(x_0')|^2)^{-1/2}$, et il paraît raisonnable de définir l'intégrale d'une fonction h définie sur Γ par

$$\begin{aligned} \int_{\Gamma} h(m) d\sigma_m &= \\ &= \int_{\mathbb{R}^{n-1}} h(\sum' x^j e_j + \varphi(x') e_n) \sqrt{1 + \sum' |\partial\varphi/\partial x^j(x')|^2} dx^1 \dots dx^{n-1}. \end{aligned}$$

Si $((f_j), W, \psi)$ est une autre présentation de Γ dans une base orthonormale, le presque parallélogramme $P(K)$ devrait correspondre, pour les mêmes raisons, à un presque parallélogramme K' de l'espace des y' , de mesure voisine de $v|\beta|/|\alpha|$, où $\beta = \nu_0 \cdot f_n$. Mais d'autre part, K' est l'image de K par l'application θ et doit être proche du parallélogramme image de K par l'application linéaire tangente, dont le volume est $v|J_\theta(x'_0)|$. Ce jacobien, qui intervient dans la formule de changement de variable dans les intégrales, devrait donc être égal au rapport β/α .

Il n'est pas impossible de faire une démonstration rigoureuse en développant les arguments ci-dessus. La démonstration qui va suivre est nettement plus courte, et nous espérons que le lecteur reconnaîtra, au milieu de ces calculs aveugles, les équivalents des considérations plus intuitives qui précèdent.

Théorème et Définition A.3.2. — Soit Γ un graphe pourvu de deux présentations $((e_j), V, \varphi)$ et $((f_j), W, \psi)$ dans des bases orthonormales.

(a) Si h est une fonction positive définie sur Γ , on a

$$\begin{aligned} \int_V h(\sum' x^j e_j + \varphi(x') e_n) \sqrt{1 + \sum' |\partial\varphi/\partial x^j|^2} dx^1 \dots dx^{n-1} \\ = \int_W h(\sum' y^j f_j + \psi(y') f_n) \sqrt{1 + \sum' |\partial\psi/\partial y^j|^2} dy^1 \dots dy^{n-1}. \end{aligned} \quad (\text{A.6})$$

les deux membres ayant un sens dans $\overline{\mathbb{R}_+}$.

(b) Si h est à valeurs complexes, et si l'intégrale précédente relative à $|h|$ est finie pour une (et donc toute) présentation, on dit que h est sommable sur Γ . Les deux membres de (A.6) sont alors égaux et leur valeur est notée $\int_\Gamma h(m) d\sigma_m$.

En remplaçant m_0 par un point quelconque $m \in \Gamma$, le calcul du n° A.2.7 nous fournit la valeur du jacobien $J_\theta(x')$. On a

$$J_\theta(x') = \pm \det(V_1, \dots, V_{n-1}, f_n),$$

avec $V_j = e_j + (\partial\varphi/\partial x^j) e_n$. Nous supprimons la référence aux (f_j) , le déterminant par rapport à une base orthonormale ne dépendant pas, au signe près, de la base choisie.

Si ν est un vecteur normal unitaire en m , posons $\alpha(m) = \nu \cdot e_n$, $\beta(m) = \nu \cdot f_n$. On a $f_n = \beta(m)\nu + \widetilde{f_n}$, où le vecteur $\widetilde{f_n}$ est orthogonal à ν et est donc combinaison linéaire des V_j . En se remémorant les propriétés des applications multilinéaires alternées on obtient donc

$$J_\theta(x') = \pm \beta(m) \det(V_1, \dots, V_{n-1}, \nu),$$

et de même

$$\det(V_1, \dots, V_{n-1}, e_n) = \pm \alpha(m) \det(V_1, \dots, V_{n-1}, \nu).$$

Le premier membre de l'égalité ci-dessus étant égal à $\det(e_1, \dots, e_{n-1}, e_n)$ et donc à ± 1 , toujours d'après le caractère multilinéaire et alterné, on en déduit que $|J_\theta(x')| = |\beta(m)/\alpha(m)|$.

Il ne reste plus, en supposant la fonction à intégrer positive ou sommable, qu'à faire le changement de variable $y' = \theta(x')$ (théorème 1.4.5) au membre de droite de (A.6). En notant $m = P(x') = \sum' y^j f_j + \psi(y') f_n = \sum' x^j e_j + \varphi(x') e_n$, on obtient

$$\int_W h(\sum' y^j f_j + \psi(y') f_n) \sqrt{1 + \sum' |\partial \psi / \partial y^j|^2} dy^1 \dots dy^{n-1} = \int_V h(\sum' x^j e_j + \varphi(x') e_n) |\beta(m)|^{-1} |J_\theta(x')| dx^1 \dots dx^{n-1}.$$

Il reste à emplacer $|J_\theta/\beta(m)|$ par $1/|\alpha(m)|$ et à exprimer cette dernière quantité en fonction de φ pour obtenir le membre de gauche de (A.6).

Remarque A.3.3. — Le lecteur pourra vérifier que le signe de J_θ est positif si les bases (e_j) et (f_j) ont la même orientation, et si les vecteurs e_n et f_n sont dirigés du même côté de Γ (c'est-à-dire que $\nu_0 \cdot e_n$ et $\nu_0 \cdot f_n$ sont de même signe). Par contre il est négatif si une et une seule de ces conditions n'est pas remplie. Ce type d'argument permet de montrer l'invariance du concept de normale extérieure dans la définition 5.4.6.

Théorème et Définition A.3.4. — Soit Σ une hypersurface compacte de E . Soient ω_i , $i = 1, \dots, N$ des ouverts de E qui recouvrent Σ tels que les ensembles $\Gamma_i = \Sigma \cap \omega_i$ soient des graphes. Soient enfin $\chi_i \in C_0^\infty(\omega_i)$ vérifiant $\sum \chi_i = 1$ au voisinage de Σ . On pose

$$\int_\Sigma h(m) d\sigma_m = \sum_{i=1}^n \int_{\Gamma_i} h(m) \chi_i(m) d\sigma_m \quad (\text{A.7})$$

dans les deux cas suivants : (a) h est positive ; (b) $\int_\Sigma |h(m)| d\sigma_m < \infty$ (on dit alors que h est sommable sur Σ).

Le résultat ne dépend pas du choix des ω_i et χ_i . Il est déterminé uniquement par Σ , h et la structure euclidienne de E .

Il faut donc vérifier que, si (U_λ) est un autre recouvrement ouvert fini du même type et si ψ_λ est une partition de l'unité associée, les deux modes de calcul donnent le même résultat. En remarquant que l'on a $\psi_\lambda = \sum_i \psi_\lambda \chi_i$ sur Σ , on obtient

$$\sum_\lambda \int_{\Sigma \cap U_\lambda} \psi_\lambda(m) h(m) d\sigma_m = \sum_\lambda \sum_i \int_{\Sigma \cap U_\lambda} \psi_\lambda(m) \chi_i(m) h(m) d\sigma_m. \quad (\text{A.8})$$

Il est évident que $\Sigma \cap U_\lambda \cap \omega_i$, intersection d'un graphe et d'un ouvert, est encore un graphe. En outre, la fonction $\psi_\lambda(m) \chi_i(m) h(m)$ étant nulle hors de

$\Sigma \cap U_\lambda \cap \omega_i$, son intégrale sur ce dernier ensemble est égale à son intégrale sur $\Sigma \cap U_\lambda$. Le premier membre de (A.8) est donc égal à

$$\sum_{\lambda} \sum_i \int_{\Sigma \cap U_\lambda \cap \omega_i} \psi_\lambda(m) \chi_i(m) h(m) d\sigma_m,$$

expression dont la symétrie nous garantit qu'elle est aussi égale au second membre de (A.7), ce qui achève la démonstration.

Remarque A.3.5. — Le résultat du théorème précédent, et la définition de l'intégrale sont valables lorsque l'hypersurface Σ n'est pas compacte, si la fonction h est nulle hors d'un compact de Σ : il suffit de prendre des partitions de l'unité au voisinage du compact en question, et la démonstration précédente s'applique mot pour mot.

En fait, en utilisant les partitions de l'unité localement finies que nous avons évoquées après le théorème 3.2.9, on peut définir l'intégrale de surface sans aucune hypothèse de compacité.

A.4. Cartes et sous-variétés

En notant E un espace vectoriel sur $\mathbb{R}$, non nécessairement euclidien, de dimension n , la présentation que nous avons donnée des hypersurfaces dans la section A.2 s'étend facilement au cas de sous-variétés de dimension quelconque $d < n$ de E .

Un graphe d dimensionnel est défini par une base (e_j) de E , un ouvert V de $\mathbb{R}^d$, et une application φ de classe C^1 de V dans $\mathbb{R}^{n-d}$ de composantes $\varphi^{d+1}, \dots, \varphi^n$, en posant

$$\Gamma = \left\{ \sum x^i e_i \mid x' = (x^1, \dots, x^d) \in V \text{ et } x^j = \varphi^j(x'), j = d+1, \dots, n \right\}.$$

L'équivalent de la définition A.2.2 permet de définir en tout point $m \in \Gamma$ la variété affine d -dimensionnelle tangente en ce point. Une démonstration tout à fait analogue à celle du théorème A.2.5, et reposant essentiellement sur le théorème d'inversion locale, assure que l'intersection de Γ et d'un voisinage de m admet une présentation dans une base (f_j) pourvu que le sous-espace vectoriel engendré par les $f_{d+1}, \dots, f_n$ soit supplémentaire du sous-espace de E parallèle à la variété affine tangente en m .

On définit alors une *sous-variété de dimension d* de E comme un sous-ensemble $\Sigma \subset E$ tel que, pour tout $m \in \Sigma$, il existe un voisinage ouvert ω de m dans E tel que $\Sigma \cap \omega$ soit un graphe d -dimensionnel.

L'équivalent de l'exercice A.2.12 fournit l'énoncé suivant. Soient Ω un ouvert de E et $F_1, \dots, F_{n-d}$ des fonctions de classe C^1 dans Ω telles que, en tout point où les F_j s'annulent simultanément, les dF_j soient linéairement indépendantes. Alors l'ensemble $\Sigma = \{m \in \Omega \mid F_1(m) = \dots = F_{n-d}(m) = 0\}$ est une sous-variété de dimension d .

On peut de manière équivalente définir une sous-variété de dimension d de E comme un ensemble Σ tel que pour tout $m \in \Sigma$, on peut trouver un voisinage Ω de m et des F_j comme ci-dessus tels que l'on ait $\Sigma \cap \Omega = \{m \in \Omega \mid F_j(m) = 0\}, j = 1, \dots, n-d$.

Un troisième type de définition repose sur le concept suivant.

Définition A.4.1. — Soit $U \subset E$. On appelle carte d -dimensionnelle de domaine U un homéomorphisme χ de U sur un ouvert V de $\mathbb{R}^d$ tel que l'application réciproque χ^{-1} de V dans E soit de classe C^1 et ait une différentielle de rang d en tout point.

L'application χ^{-1} s'appelle souvent une représentation paramétrique régulière de U .

Le lecteur pourra, en s'appuyant bien sûr sur le théorème d'inversion locale, démontrer le théorème suivant. On rappelle que pour $U \subset \Sigma \subset E$, l'ensemble U est un ouvert du sous-espace Σ si et seulement s'il est de la forme $\Sigma \cap \Omega$ avec Ω ouvert dans E .

Théorème A.4.2. — (a) Soit $\Sigma \subset E$. Pour que Σ soit une sous-variété de dimension d , il faut et il suffit que, pour tout $m \in \Sigma$, il existe un voisinage ouvert U de m dans Σ et une carte d -dimensionnelle χ de U .

(b) Soient χ_1 et χ_2 deux cartes envoyant respectivement un même ouvert U de Σ sur des ouverts V_1 et V_2 de $\mathbb{R}^d$. Alors l'application $\chi_2 \circ \chi_1^{-1}$ est un difféomorphisme de V_1 sur V_2 .

A.4.3. Sous-variétés et variétés. — Il est rarement possible de représenter une sous-variété entière à l'aide d'une seule carte, mais les cartes permettent en général d'en représenter des portions plus grandes que les graphes. Par exemple, pour la sphère unité de $\mathbb{R}^3$, la projection stéréographique (à partir du pôle nord) fournit une carte définie sur la sphère privée d'un point. Il suffit donc de deux telles cartes pour représenter toute la sphère. Au contraire, un graphe ne peut représenter au mieux qu'un hémisphère ouvert, et il en faut au moins quatre pour recouvrir la sphère.

Il existe une raison plus profonde de préférer les représentations par cartes. On souhaite souvent — pour des raisons relevant comme nous le verrons des mathématiques mais aussi de la physique et de la mécanique — étudier une sous-variété Σ "en elle-même", indépendamment de l'espace dans lequel elle est plongée. Alors que les représentations par graphes ou par équations implicites font appel directement à E , une carte χ n'est définie que sur un ouvert de Σ . Cela dit, l'indépendance par rapport à E n'est pas complète puisque l'on demande que χ^{-1} , considérée comme application à valeurs dans E , soit de classe C^1 et ait une différentielle de rang maximum.

La véritable notion intrinsèque, indépendante de tout plongement, est celle de *variété différentielle* qui constitue le cadre général dans lequel on peut développer le calcul différentiel. Une telle variété M est par définition un espace topologique séparé⁽²⁾, muni d'un ensemble d'homéomorphismes appelés cartes allant d'un ouvert U de M (le domaine de la carte) sur un ouvert V de $\mathbb{R}^d$. On demande que les domaines des cartes recouvrent M , ainsi qu'une condition de cohérence entre les cartes. Celle-ci est directement inspirée de la partie (b) du théorème A.4.2 et s'énonce ainsi : si les domaines U_1 et U_2 de deux cartes χ_1 et χ_2 ont une intersection non vide, en notant $\widetilde{\chi}_1$ et $\widetilde{\chi}_2$ les restrictions de χ_1 et χ_2 à $U_1 \cap U_2$, l'application $\widetilde{\chi}_1^{-1} \circ \widetilde{\chi}_2$ est un difféomorphisme de $\chi_1(U_1 \cap U_2)$ sur $\chi_2(U_1 \cap U_2)$.

Remarque A.4.4. — Il existe de nombreux cas où un même concept mathématique peut se représenter naturellement de plusieurs manières comme une sous-variété. Par exemple, dans le plan, identifié à $\mathbb{C}$, le groupe des rotations peut se décrire comme l'ensemble des nombres complexes de module 1 (sous-variété de codimension 1 dans un espace de dimension 2) ou comme l'ensemble des matrices $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ vérifiant $a = d$, $b = -c$, $a^2 + b^2 = 1$ (sous-variété de codimension 3 dans un espace de dimension 4).

⁽²⁾On ajoute le plus souvent la condition que M est réunion dénombrable de compacts.

De même, en mécanique, l'ensemble des positions possibles dans l'espace d'un parallélogramme articulé peut être décrit comme l'ensemble des points de $\mathbb{R}^{12}$ (coordonnées des 4 sommets) assujettis à 5 liaisons mais aussi comme l'ensemble des points de $\mathbb{R}^9$ (coordonnées de trois sommets) assujettis à deux liaisons, etc. Il est plus satisfaisant de penser l'espace des configurations comme une variété différentielle de dimension 7 sur laquelle on a des coordonnées locales, que de choisir arbitrairement une de ses réalisations dans un espace $\mathbb{R}^n$.

Il existe également des cas où des êtres mathématiques ou physiques sur lesquels on veut faire du calcul différentiel n'ont pas de plongement naturel dans un espace vectoriel (même si un théorème abstrait assure que de tels plongements sont toujours possibles). Par exemple, le plan projectif réel se définit comme le quotient de $\mathbb{R}^3 \setminus \{0\}$ par la relation d'équivalence : x et y sont alignés avec l'origine. On peut aussi le définir comme la réunion du plan affine et d'une droite "à l'infini" (en précisant la structure de cette réunion). Par contre, aucune définition comme sous-espace d'un espace vectoriel n'apparaît naturelle.

Quant à l'espace-temps de la relativité générale, il est de dimension 4 et n'a pas de structure vectorielle. On veut néanmoins y définir des champs de tenseurs et y écrire des équations aux dérivées partielles. La représentation comme variété différentielle s'impose, tout plongement introduirait des dimensions artificielles ne correspondant à aucune réalité physique.

APPENDICE B

ESPACES DE BAIRE

B.1. Résultats fondamentaux

Définition B.1.1. — On dit qu'un espace topologique X est un espace de Baire si toute intersection dénombrable d'ouverts denses dans X est un sous-ensemble partout dense. Il est équivalent de dire qu'une réunion dénombrable de fermés d'intérieur vide est un ensemble d'intérieur vide.

Définition B.1.2. — On dit qu'une partie A d'un espace de Baire X est un résiduel si A contient une intersection dénombrable d'ouverts partout denses. On dit que A est un ensemble maigre (ou un ensemble de première catégorie de Baire) si son complémentaire est un résiduel, ce qui signifie que A est contenu dans une réunion dénombrable de fermés d'intérieur vide.

Il est clair que les résiduels sont stables par intersection dénombrable, et que les ensembles maigres sont stables par réunion dénombrable.

Théorème B.1.3. — (a) Tout espace métrique complet est un espace de Baire.

(b) Tout espace topologique localement compact est un espace de Baire.

Soit donc (U_j) une suite d'ouverts partout denses. Pour prouver que l'intersection est partout dense il suffit de montrer que, si V est un ouvert non vide quelconque, il existe un point commun à V et à tous les U_j . Nous allons dans les deux cas construire par récurrence une suite d'ensembles fermés B_j vérifiant $B_1 \subset U_1 \cap V$ et $B_{j+1} \subset U_{j+1} \cap \overset{\circ}{B_j}$. Il nous suffira alors de montrer que l'intersection des B_j est non vide pour avoir le résultat.

Dans le cas (a) nous allons choisir comme ensembles B_j des boules fermées, centrées en un point x_j , et de rayon $\leq 1/j$ strictement positif. L'existence de B_1 ne présente pas de difficulté. La boule B_j étant construite, l'ouvert $U_{j+1} \cap \overset{\circ}{B_j}$ est alors non vide et contient donc un point x_{j+1} . Il contient par

conséquent une boule centrée en ce point, que l'on peut supposer fermée et de rayon $\leq 1/(j+1)$. On peut ainsi construire par récurrence la suite des B_j .

A partir du rang j_0 , tous les points x_j appartiennent à la boule B_{j_0} et ont une distance mutuelle $\leq 2/j_0$. La suite x_j est donc une suite de Cauchy. L'espace étant complet, elle converge vers un certain point x qui appartient à la boule B_{j_0} , celle-ci étant fermée. Cela étant valable pour tout j_0 , nous avons prouvé que l'intersection des B_j contient le point x et est donc non vide.

Dans le cas (b) rappelons d'abord qu'un espace (métrique ou plus généralement topologique séparé) est localement compact si pour tout point x et pour tout voisinage W de x , il existe un voisinage compact de x contenu dans W . Nous allons cette fois exiger que les ensembles B_j soient des compacts d'intérieur non vide.

L'ouvert $U_1 \cap V$ est non vide et est donc un voisinage de l'un quelconque de ses points x . Nous prendrons comme ensemble B_1 un voisinage compact de x contenu dans $U_1 \cap V$. De même, B_j étant supposé construit, nous choisirons un point x de l'ouvert non vide $U_{j+1} \cap \overset{\circ}{B}_j$ et B_{j+1} sera un voisinage compact de x contenu dans cet ouvert.

Une suite décroissante de compacts non vides ayant toujours, d'après le théorème 2.3.2, une intersection non vide, cela achève la preuve du théorème.

Théorème B.1.4. — *Soit X un espace de Baire. et soit F_j une suite de fermés qui recouvre X . Alors la réunion des $\overset{\circ}{F}_j$ est un ouvert partout dense.*

Soit G le complémentaire de la réunion des $\overset{\circ}{F}_j$. C'est un ensemble fermé, et il nous faut prouver qu'il est d'intérieur vide. Chacun des $F_j \cap G$ étant un fermé d'intérieur vide, et leur réunion étant égale à G , cela résulte du fait que X est un espace de Baire.

B.2. Quelques applications

Nous donnons d'abord l'exemple historique, dû à Baire lui-même, de l'utilisation des concepts précédents.

Théorème B.2.1. — *Soit f_j une suite de fonctions continues sur $[0, 1]$ (ou sur un espace de Baire quelconque) qui converge en chaque point vers une fonction $f(x)$. Alors, l'ensemble des points où f est continue est un résiduel, et est en particulier non vide.*

Pour $\delta > 0$ et $n \in \mathbb{N}$, considérons l'ensemble

$$F_{n,\delta} = \{x \mid \forall j \geq n, \forall k \geq n, |f_j(x) - f_k(x)| \leq \delta\}.$$

Pour $x \in F_{n,\delta}$ et $j \geq n$, on a $|f_j(x) - f(x)| \leq \delta$. Pour δ fixé, la réunion des ensembles fermés $F_{n,\delta}$ est égale à tout l'espace. D'après le théorème précédent, il en résulte que $U_\delta = \bigcup_n \overset{\circ}{F}_{n,\delta}$ est un ouvert partout dense.

L'ensemble $R = \bigcap_k U_{1/k}$ est donc un résiduel, et il nous reste à montrer que f est continue en un point quelconque x_0 de R . Donnons nous $\varepsilon > 0$. Le point x_0 appartenant à $U_{\varepsilon/3}$, il existe donc un voisinage V_1 de ce point, et un entier n tel que l'on ait $V_1 \subset F_{n,\varepsilon/3}$.

D'autre part, la fonction f_n étant continue, il existe un voisinage V_2 de x_0 tel que l'on ait $|f_n(x) - f_n(x_0)| \leq \varepsilon/3$ pour x dans ce voisinage. Pour tout x appartenant au voisinage $V_1 \cap V_2$ de x_0 , on a donc

$$|f(x) - f(x_0)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)| \leq \varepsilon,$$

ce qui achève la démonstration.

Remarque B.2.2. — On dit qu'une fonction est de première classe de Baire si elle est limite simple d'une suite de fonctions continues. On dit qu'une fonction est de deuxième classe de Baire si elle est limite simple d'une suite de fonctions de première classe de Baire, et ainsi de suite (cette expression anodine cachant une récurrence transfinie).

Exercice B.2.3. — Démontrer que, sur $[0, 1]$, la fonction f égale à $1/q$ en tout point rationnel mis sous forme irréductible p/q et à 0 en tout point irrationnel est de première classe de Baire. Vérifier que ses points de continuité forment un résiduel.

Exercice B.2.4. — Démontrer que la fonction g égale à 1 en tout point rationnel et à 0 en tout point irrationnel est de deuxième classe de Baire mais n'est pas de première classe de Baire.

La topologie de la convergence simple sur l'espace $\mathcal{F}$ de toutes les applications de $[0, 1]$ dans $\mathbb{R}$ ne peut pas être définie par une métrique. À partir de l'espace C^0 des fonctions continues, on est amené à considérer au moins les trois espaces suivants.

— L'espace des fonctions de première classe de Baire, c'est à dire l'espace des éléments de $\mathcal{F}$ qui sont limite simple d'une suite d'éléments de C^0 .

— L'espace des fonctions boréliennes, c'est-à-dire le plus petit sous-espace de $\mathcal{F}$ contenant C^0 qui soit stable pour la convergence simple des suites. C'est un sous-espace beaucoup plus gros : il contient aussi l'espace (strictement plus grand) des fonctions de deuxième classe de Baire et ainsi de suite.

— L'adhérence de C^0 dans $\mathcal{F}$. C'est un espace encore plus gros : il est égal à $\mathcal{F}$ entier.

Le lecteur pourra maintenant apprécier le confort que lui procure le théorème 7.2.5 : bien que la topologie de l'espace des distributions ne puisse pas être définie par une métrique, chaque élément de cet espace est limite d'une suite d'éléments de C_0^∞ .

B.2.5. Existence de fonctions continues nulle part dérivables. — Cette "plaie lamentable" dont se détournait Hermite est plus étendue qu'on ne pourrait le penser : les fonctions qui sont dérivables en au moins un point ne constituent qu'un ensemble maigre dans l'espace des fonctions continues.

Exercice B.2.6. — On considère l'espace de Banach C^0 des fonctions continues sur $[0, 1]$ muni de la norme uniforme. Pour $\varepsilon > 0$ et $n \in \mathbb{N}$, on considère l'ensemble $U_{\varepsilon, n}$ constitué des $f \in C^0$ vérifiant la propriété suivante

$$\forall x \in [0, 1], \exists y \in [0, 1], \quad 0 < |y - x| \leq \varepsilon \text{ et } \left| \frac{f(y) - f(x)}{y - x} \right| > n.$$

(a) Démontrer que, pour $f \in U_{\varepsilon, n}$, on peut trouver un nombre fini de compacts K_j , $j = 1 \dots p$ recouvrant $[0, 1]$, et des points $y_j \notin K_j$ tels que, pour $x \in K_j$ on ait $|y_j - x| \leq \varepsilon$ et que

$$\inf_j \inf_{x \in K_j} \left| \frac{f(y_j) - f(x)}{y_j - x} \right| > n.$$

En déduire que $U_{\varepsilon, n}$ est un ouvert de C^0 .

(b) Démontrer que $U_{\varepsilon, n}$ est partout dense dans C^0 . On montrera que, pour f quelconque dans C^0 , et pour $\delta > 0$, la fonction $f(x) + \delta \sin Nx$ appartient à $U_{\varepsilon, n}$ pourvu que N soit choisi assez grand en fonction de ε, n, δ et de l'expression de la continuité uniforme de f .

(c) On appelle R l'intersection des $U_{1/n, n}$. Montrer que R est un résiduel, et que les éléments de R ne sont dérivables en aucun point.

APPENDICE C

ESPACES DE FRÉCHET

Dans cet appendice, le corps des scalaires des espaces vectoriels sera toujours supposé être $\mathbb{R}$ ou $\mathbb{C}$.

C.1. Espaces localement convexe métrisables

Définition C.1.1. — Soit E un espace vectoriel. On appelle *semi-norme* sur E une application $\mathcal{P}$ de E dans $[0, \infty[$ vérifiant $\mathcal{P}(f + g) \leq \mathcal{P}(f) + \mathcal{P}(g)$ et $\mathcal{P}(\lambda f) = |\lambda| \mathcal{P}(f)$ pour $f, g \in E$ et λ scalaire.

Dire que $\mathcal{P}$ est une norme équivaut à dire que $\mathcal{P}$ est une semi-norme et que le seul élément f vérifiant $\mathcal{P}(f) = 0$ est l'élément 0.

Définition C.1.2. — On appelle *espace localement convexe métrisable* un espace vectoriel muni d'une suite croissante de semi-normes $\mathcal{P}_j$ telles que l'on ait

$$\{\forall j, \mathcal{P}_j(f) = 0\} \iff f = 0$$

On dira que deux suites croissantes de semi-normes sont équivalentes et définissent la même structure d'espace localement convexe métrisable sur E si on a

$$\begin{aligned} \forall j, \exists k, \exists C, \quad \mathcal{P}_j &\leq C \mathcal{Q}_k \\ \forall j, \exists k, \exists C, \quad \mathcal{Q}_j &\leq C \mathcal{P}_k \end{aligned}$$

Remarque C.1.3. — On définit parfois la structure en se donnant une suite quelconque $\mathcal{P}_j$ de semi-normes. On se ramène au cas d'une suite croissante en posant $\mathcal{Q}_j = \sum_{k \leq j} \mathcal{P}_k$.

Bien entendu, les espaces normés sont un cas (très) particulier des espaces du type précédent. Il suffit de prendre toutes les semi-normes égales à la norme.

C.1.4. Structures d'espace métrique d'un espace localement convexe métrisable. — Soit α_j une suite de nombres réels > 0 tels que l'on ait $\sum \alpha_j < \infty$. On pose

$$d(f, g) = \sum \alpha_j \min \{1, \mathcal{P}_j(f - g)\}.$$

Le lecteur vérifiera que d est effectivement une distance sur E . Il vérifiera également que, si on modifie les α_j ou si on remplace la suite des $\mathcal{P}_j$ par une suite équivalente de semi-normes, on obtient une distance $\tilde{d}$ uniformément équivalente (voir (2.2)) à d .

Définition C.1.5. — On appelle espace de Fréchet un espace localement convexe métrisable qui est complet pour la structure d'espace métrique ci-dessus.

C.1.6. Utilité et nocivité de la métrique. — Il est bien sûr important de pouvoir appliquer aux espaces localement convexe métrisables toutes les définitions et propriétés connues des espaces métriques : voisinages, suites convergentes, de Cauchy, ..., et notamment le fait qu'une application Φ de E dans un espace topologique est continue si pour toute suite f_n tendant vers f la suite $\Phi(f_n)$ converge vers $\Phi(f)$.

Par contre, il ne faut jamais utiliser la distance d définie ci-dessus, qui n'a pas de propriété d'homogénéité et où la manipulation des α_j et des minimums est très désagréable. Il est donc important d'exprimer, uniquement en termes de semi-normes, les concepts précédents. Il s'agit d'exercices faciles que nous laissons au lecteur.

— Pour qu'une suite f_n converge vers f , il faut et il suffit que pour tout j , on ait $\mathcal{P}_j(f_n - f) \rightarrow 0$. Pour que f_n soit de Cauchy, il faut et il suffit que pour tout j on ait $\lim_{m, n \rightarrow \infty} \mathcal{P}_j(f_n - f_m) = 0$.

— Pour que V soit un voisinage de f , il faut et il suffit qu'il contienne une semi-boule centrée en ce point, c'est-à-dire qu'il existe j et $\varepsilon > 0$ tels que

$$\{g \mid \mathcal{P}_j(g - f) < \varepsilon\} \subset V.$$

— Les semi-normes $\mathcal{P}_j$ sont continues de E dans $\mathbb{R}$. Les translations et les homothéties sont continues de E dans lui-même.

Théorème C.1.7. — Soit L une application linéaire de E dans F , ceux-ci étant des espaces localement convexe métrisables munis respectivement des suites de semi-normes $\mathcal{P}_j$ et $\mathcal{Q}_k$. Les trois propriétés suivantes sont équivalentes.

(a) L est continue en 0.

(b) L est continue en tout point.

(c) $\forall k, \exists j, \exists C, \forall f$

$$Q_k(L(f)) \leq C P_j(f). \quad (C.1)$$

Montrons que la continuité de L en 0 entraîne (c). L'image réciproque de la semi-boule $\{g \mid Q_k(g) \leq 1\}$ doit être un voisinage de 0, et donc contenir une certaine semi-boule $\{f \mid P_j(f) \leq \varepsilon\}$. On a l'implication $P_j(f) \leq \varepsilon \Rightarrow Q_k(L(f)) \leq 1$ et, par homogénéité, la relation (C.1) avec $C = 1/\varepsilon$.

Montrons que (c) entraîne (b). Si une suite f_n converge vers f dans E , on a $P_j(f_n - f) \rightarrow 0$ pour tout j , et il résulte des estimations (C.1) que l'on a $Q_k(L(f_n) - L(f)) \rightarrow 0$ pour tout k , ce qui exprime que $L(f_n) \rightarrow L(f)$. Enfin, il est trivial que (b) implique (a).

Remarque C.1.8. — On définit plus généralement la structure d'espace localement convexe sur un espace vectoriel E par la donnée d'un ensemble d'indices I ordonné filtrant croissant (c'est-à-dire tel que deux éléments de I ont toujours un majorant commun), et la donnée d'une famille croissante $\mathcal{P}_I$ de semi-normes indexées par I .

On en déduit une structure d'espace topologique sur E , les voisinages d'un point étant les ensembles contenant une semi-boule centrée en ce point. On définit les suites de Cauchy et les suites convergentes par des expressions analogues à celles que nous avons vues ci-dessus, mais ces concepts sont insuffisants pour rendre compte de toute la structure, qui ne s'exprime complètement qu'à l'aide du concept de filtre.

Un exemple important est la topologie de la convergence simple sur l'espace $\mathcal{F}$ de toutes les fonctions définies sur un ensemble A . On prend comme ensemble d'indice l'ensemble B des sous-ensembles finis de A , et on définit les semi-normes $\mathcal{P}_B(f) = \sum_{x \in B} |f(x)|$ pour $B \in B$.

Un autre exemple important est la topologie faible sur l'espace des distributions. L'ensemble d'indices est l'ensemble des sous-ensembles finis de C_0^∞ , et les semi-normes sont définies par $\mathcal{P}_B(u) = \sum_{\varphi \in B} |\langle u, \varphi \rangle|$. Cette topologie donne la définition des suites convergentes de distributions que nous avons utilisée dans le cours. Il existe une autre topologie, dite forte, sur $\mathcal{D}'$ qui est différente mais pour qui les suites convergentes sont les mêmes. Comme nous n'avons utilisé que la convergence des suites, nous n'avons pas eu à choisir.

C.2. Exemples d'espaces de Fréchet

Les espaces fonctionnels sont souvent définis comme ensemble de fonctions f telles que certaines quantités (intégrales, maximums...) dépendant de f soient finies. Ces quantités ont souvent les propriétés de convexité qui en font des (semi-)normes. Lorsqu'il n'y en a qu'une (ou un nombre fini ce qui revient au même), on munit naturellement l'espace d'une structure d'espace normé. Par contre, lorsqu'il y en a une infinité dénombrable, et notamment lorsque "pour chaque compact" ou "pour chaque ordre de dérivation" une quantité doit être finie, c'est le cadre des espaces localement convexe métrisables qui s'impose.

Tous les espaces que nous allons considérer seront des espaces de Fréchet, ce qui leur confère beaucoup de propriétés intéressantes. S'il faut s'en réjouir, il ne faut pas s'en étonner : ceux qui n'étaient pas complets au départ (l'espace des fonctions intégrables au sens de Riemann par exemple) ont été dûment complétés.

C.2.1. L'espace C_K^m . — L'espace des fonctions m fois continûment dérivables à support dans un compact K de $\mathbb{R}^n$ est un espace de Banach pour la norme

$$\mathcal{P}_m(f) = \sup_{x \in K, |\alpha| \leq m} |\partial^\alpha f(x)|. \quad (\text{C.2})$$

Le fait que l'espace soit complet est une partie de la démonstration qui va suivre.

C.2.2. L'espace C_K^∞ . — Il sera muni de la suite des normes $\mathcal{P}_m$ définies encore par la formule (C.2).

Montrons que cet espace est complet. Soit f_n une suite de Cauchy. On a alors $\lim_{m,n \rightarrow \infty} \mathcal{P}_j(f_m - f_n) = 0$ pour chaque j , ce qui entraîne que pour chaque α , on a $\sup_x |\partial^\alpha f_m(x) - \partial^\alpha f_n(x)| \rightarrow 0$. La suite des $\partial^\alpha f_n$ est donc une suite de Cauchy pour la convergence uniforme, et elle tend donc uniformément vers une limite v_α qui est une fonction continue à support dans K .

Il suffit maintenant d'appliquer le théorème classique assurant que si f_n converge vers v_0 uniformément, et si les dérivées $\partial_i f_n$ sont continues et convergent vers des limites v_i uniformément, alors la fonction v_0 est de classe C^1 et les dérivées de la limite sont les limites des dérivées. En appliquant par récurrence ce résultat aux dérivées successives de v_0 , on obtient que $\partial^\alpha v_0 = v_\alpha$ et la convergence uniforme des $\partial^\alpha f_n$ vers les $\partial^\alpha v_0$ exprime précisément que $\mathcal{P}_j(f_n - v_0) \rightarrow 0$ pour tout j .

Remarque C.2.3. — L'espace $C_0^\infty(\Omega)$ n'a pas de topologie aussi simple, mais il est réunion des C_K^∞ pour K compact contenu dans Ω . La définition que nous avons donnée des distributions (voir définition 4.2.1) est précisément la suivante : forme linéaire sur $C_0^\infty(\Omega)$ dont la restriction à chaque C_K^∞ est continue.

C.2.4. L'espace $L_{loc}^1(\Omega)$. — C'est l'espace des classes de fonctions sommables sur chaque compact de Ω . Si K_j est une suite exhaustive de compacts de Ω , on munit l'espace des semi-normes suivantes

$$\mathcal{P}_j(f) = \int_{K_j} |f(x)| \, dx.$$

La notion de convergence associée est la convergence en norme L^1 sur chaque compact.

Si on se donne une suite de Cauchy f_n , chacun des espaces $L^1(K)$ étant complet, la restriction à K_j des f_n converge vers un élément g_j de $L^1(K_j)$. Mais la restriction de f_n à K_{j+1} convergeant vers g_{j+1} dans $L^1(K_{j+1})$, en en prenant les restrictions à K_j , on obtient que la restriction de g_{j+1} à K_j est égale à g_j (en tant que classe de fonctions, c'est-à-dire presque partout si on prend des représentants).

Les (classes de) fonctions g_j "se recollent" et constituent les restrictions aux K_j d'une même (classe de) fonction $g \in L^1_{\text{loc}}(\Omega)$. La convergence des restrictions des f_n aux K_j vers les g_j exprime précisément que f_n tend vers g en norme L^1 sur chaque compact.

C.2.5. Les espaces $L^2_{\text{loc}}(\Omega)$ et $L^\infty_{\text{loc}}(\Omega)$. — Ils se définissent de même, en prenant comme semi-normes les normes L^2 ou L^∞ des restrictions à une suite exhaustive de compacts. On montre de même que ce sont des espaces de Fréchet.

C.2.6. L'espace $C^\infty(\Omega)$. — Il sera muni de la suite de semi-normes suivantes

$$\mathcal{P}_j(f) = \sum_{|\alpha| \leq j} \sup_{x \in K_j} |\partial^\alpha f(x)|,$$

où les K_j forment encore une suite exhaustive de compacts. La convergence de f_n vers f dans cet espace exprime que chaque dérivée de f_n converge vers la dérivée correspondante de f uniformément sur chaque compact de Ω . On montre que cet espace est complet à l'aide des deux arguments utilisés précédemment : égalité entre dérivée de la limite et limite des dérivées d'une part, et recollement des limites définies sur les K_j d'autre part.

La remarque 6.1.7 exprime précisément que $\mathcal{E}'(\Omega)$ est l'espace des formes linéaires continues sur $C^\infty(\Omega)$.

C.2.7. L'espace $\mathcal{S}(\mathbb{R}^n)$. — Il est muni des semi-normes notées $\mathcal{N}_p$ au chapitre 9. Nous laissons encore au lecteur le soin de prouver, par les méthodes ci-dessus, qu'il est complet. Les théorèmes et définitions 9.2.4, 9.2.9 et 9.3.2 expriment respectivement que la transformation de Fourier est continue de $\mathcal{S}$ dans lui-même, que C^∞_0 est dense, et que $\mathcal{S}'$ est l'espace des formes linéaires continues sur $\mathcal{S}$.

C.3. Le théorème de Banach-Steinhaus

Nous donnerons deux formes de ce théorème. La première, plus simple, est la plus fréquemment employée. La seconde, qui entraîne la première, est plus technique mais contient plus d'information.

Théorème C.3.1. — Soient E un espace de Fréchet, F un espace localement convexe métrisable et L_n une suite d'applications linéaires continues de E dans F . On suppose que, pour tout $f \in E$ la suite $L_n(f)$ converge dans F vers une limite que l'on note $L(f)$. On a alors :

- (a) l'application linéaire L est continue de E dans F ,
- (b) pour toute suite f_n qui converge dans E vers un élément f , la suite $L_n(f_n)$ converge vers $L(f)$ dans F .

Théorème C.3.2. — Sous les hypothèses du théorème précédent, on a :

- (a) la famille des applications L_n est équicontinue, c'est-à-dire que, en notant $(\mathcal{P}_j)$ et $(\mathcal{Q}_k)$ respectivement les semi-normes de E et F , on peut trouver pour chaque k une constante C et un indice j tels que

$$\forall n, \forall f, \quad \mathcal{Q}_k(L_n(f)) \leq C\mathcal{P}_j(f). \quad (\text{C.3})$$

- (b) pour tout compact K de E , la convergence de L_n vers L est uniforme sur K , c'est-à-dire que

$$\forall k, \quad \lim_{n \rightarrow \infty} \sup_{f \in K} \mathcal{Q}_k(L_n(f) - L(f)) = 0. \quad (\text{C.4})$$

Démonstration : le théorème C.3.2 entraîne le théorème C.3.1. — La première partie est facile. Il suffit de faire tendre n vers l'infini dans la relation (C.3), et d'utiliser la continuité des semi-normes. On obtient pour tout k l'existence de C et j tels que $\mathcal{Q}_k(L(f)) \leq C\mathcal{P}_j(f)$ ce qui exprime précisément la continuité de L .

L'ensemble constitué des f_n et de leur limite f est un compact de E . La relation (C.4) nous assure que

$$\lim_{n \rightarrow \infty} \sup_m \mathcal{Q}_k(L_n(f_m) - L(f_m)) = 0.$$

Il suffit de majorer $\mathcal{Q}_k(L_n(f_n) - L(f))$ par la somme de $\mathcal{Q}_k(L(f_n) - L(f))$ et de $\mathcal{Q}_k(L_n(f_n) - L(f_n))$ qui tendent tous deux vers 0 pour conclure.

Nous aurons besoin du résultat suivant.

Lemme C.3.3. — Soit C un ensemble convexe symétrique (c'est-à-dire tel que $f \in C$ implique $-f \in C$) dans un espace localement convexe métrisable. Si C est d'intérieur non vide, c'est un voisinage de 0.

L'ensemble C étant d'intérieur non vide, il contient une certaine semi-boule $\{g \mid \mathcal{P}_j(g - f) < \varepsilon\}$ centrée en un point f . Il en résulte que C contient la semi-boule centrée à l'origine $\{h \mid \mathcal{P}_j(h) < \varepsilon\}$. En effet, un point quelconque h de cette semi-boule est le milieu des points $-(f - h)$ et $f + h$ qui appartiennent tous deux à C .

Démonstration du théorème C.3.2(a). — Soit $\mathcal{Q}_k$ une semi-norme sur F et considérons, pour $p \in \mathbb{N}$, les ensembles

$$C_p = \{f \in E \mid \forall n, \mathcal{Q}_k(L_n(f)) \leq p\}.$$

Les ensembles C_p sont fermés, et leur réunion est égale à E entier. En effet, pour chaque $f \in E$, la suite $L_n(f)$ étant convergente, la suite $\mathcal{Q}_k(L_n(f))$ est aussi convergente et donc bornée.

L'espace E étant métrique complet, c'est un espace de Baire d'après le théorème B.1.3, et il est donc impossible que tous les C_p soient d'intérieur vide. Comme ils sont symétriques, il résulte du lemme précédent qu'il existe un indice p tel que C_p soit un voisinage de 0 dans E . Il contient donc une semi-boule, relative à une des semi-normes $\mathcal{P}_j$ et de rayon $\varepsilon > 0$.

On a donc l'implication

$$\mathcal{P}_j(f) \leq \varepsilon \implies \forall n, \mathcal{Q}_k(L_n(f)) \leq p$$

et par homogénéité

$$\forall n, \mathcal{Q}_k(L_n(f)) \leq \frac{p}{\varepsilon} \mathcal{P}_j(f),$$

ce qui est le résultat voulu.

Démonstration du théorème C.3.2(b). — Il s'agit d'une propriété générale des familles équicontinues de fonctions sur un espace compact : la convergence simple entraîne la convergence uniforme.

Fixons $\varepsilon > 0$ et une semi-norme $\mathcal{Q}_k$, notre but est de montrer que, pour n assez grand, on a $\sup_{f \in K} \mathcal{Q}_k(L_n(f) - L(f)) \leq \varepsilon$.

Soient C et $\mathcal{P}_j$ tels que l'estimation (C.3) soit valable. On peut recouvrir le compact K par un nombre fini de semi-boules relatives à $\mathcal{P}_j$ et de rayon $\varepsilon/3C$. En notant g_λ leurs centres, il résulte de C.3 que l'on a

$$\forall f \in K, \exists g_\lambda, \forall n, \mathcal{Q}_k(L_n(f) - L_n(g_\lambda)) \leq \varepsilon/3,$$

et cette relation est également valable en remplaçant L_n par L . L'ensemble des g_λ étant fini, on peut trouver un entier N tel que, pour $n \geq N$, on ait $\mathcal{Q}_k(L(g_\lambda) - L_n(g_\lambda)) \leq \varepsilon/3$ en tous ces points. L'inégalité triangulaire montre alors que $\mathcal{Q}_k(L_n(f) - L(f)) \leq \varepsilon$ pour tout $f \in K$ dès que $n \geq N$, ce qui achève la démonstration.

C.3.4. Application : démonstration du théorème 4.2.3. — Il s'agit de prouver que, si (u_j) est une suite de distributions dans un ouvert Ω telle que pour tout $\varphi \in C_0^\infty(\Omega)$ la suite $\langle u_j, \varphi \rangle$ converge vers une limite que l'on note $\langle u, \varphi \rangle$, la forme linéaire u est une distribution. On sait (voir remarque C.2.3) que dire que u_j est une distribution signifie que la restriction de u_j à chaque C_K^∞ est

une forme linéaire continue sur cet espace de Fréchet. Le théorème C.3.1(a) assure qu'il en est de même pour u , ce qui est le résultat voulu.

C.4. Continuité des applications bilinéaires

Pour beaucoup d'applications bilinéaires (produit, convolution, dualité), il est facile de prouver la continuité séparée. Le corollaire suivant du théorème de Banach-Steinhaus, et des variantes de celui-ci, permettent d'en déduire la continuité (séquentielle) par rapport au couple des variables.

Corollaire C.4.1. — Soient E un espace de Fréchet, F un espace localement convexe métrisable, et D un espace vectoriel sur lequel on a défini une notion de suite convergente. Soit B une application bilinéaire séparément continue de $D \times E$ dans F , c'est à dire telle que

(a) Pour toute suite f_n d'éléments de D convergeant vers un élément f_0 , et pour tout $g \in E$, la suite $B(f_n, g)$ converge vers $B(f_0, g)$ dans F .

(b) Pour toute suite g_n d'éléments de E convergeant vers un élément g_0 , et pour tout $f \in D$, la suite $B(f, g_n)$ converge vers $B(f, g_0)$ dans F .

Alors, si f_n et g_n sont des suites convergentes dans D et E respectivement, on a

$$\lim_{n \rightarrow \infty} B(f_n, g_n) = B\left(\lim_{n \rightarrow \infty} f_n, \lim_{n \rightarrow \infty} g_n\right).$$

Soient donc f_n et g_n qui convergent vers f_0 et g_0 , et considérons les applications linéaires L_n de E dans F définies par $L_n(g) = B(f_n, g)$. La propriété (b) affirme que ces applications sont continues, tandis que la propriété (a) assure que pour tout g la suite $L_n(g)$ converge vers $B(f_0, g)$. D'après la partie (b) du théorème C.3.1, la suite $L_n(g_n)$ converge vers $B(f_0, g_0)$, ce qui est la propriété voulue.

Lorsque la convergence dans D est la convergence pour une structure d'espace métrique, la propriété ci-dessus exprime que B est continue de l'espace produit $D \times E$ dans F . Si par contre cette convergence est associée à une structure d'espace topologique non métrisable, le théorème prouve ce que l'on appelle la "continuité séquentielle" de B , mais il se peut que B ne soit pas continue sur $D \times E$.

C.4.2. Applications. — (a) Si $u_n \rightarrow u$ dans $\mathcal{D}'(\Omega)$, et si $\varphi_n \rightarrow \varphi$ dans C_K^∞ , où K est un compact de Ω , alors $\langle u_n, \varphi_n \rangle \rightarrow \langle u, \varphi \rangle$.

C'est une application immédiate du corollaire : la continuité en u n'est rien d'autre que la définition des suites convergentes dans $\mathcal{D}'$, tandis que la continuité en φ est la définition des distributions comme formes linéaires continues sur chaque C_K^∞ .

On démontre de même le résultat analogue lorsque $u_n \rightarrow u$ dans $\mathcal{S}'$ et $\varphi_n \rightarrow \varphi$ dans $\mathcal{S}$.

(b) Si $f_n \rightarrow f$ dans $C^\infty(\Omega)$, et si $u_n \rightarrow u$ dans $\mathcal{D}'(\Omega)$, alors $f_n u_n \rightarrow fu$ dans $\mathcal{D}'(\Omega)$.

Il s'agit de prouver que, pour tout $\varphi \in C_0^\infty(\Omega)$, on a $\langle f_n u_n, \varphi \rangle \rightarrow \langle fu, \varphi \rangle$. Le corollaire appliqué à la forme bilinéaire $B(u, f) = \langle fu, \varphi \rangle = \langle u, f\varphi \rangle$ fournit immédiatement le résultat.

C.4.3. Continuité séquentielle du produit de convolution. — Nous allons prouver le résultat énoncé au théorème 7.4.9 : Si F et G sont deux fermés convolutifs, et si deux suites de distributions u_j et v_j respectivement à support dans F et G sont convergentes, alors $\lim(u_j \star v_j) = (\lim u_j) \star (\lim v_j)$.

Dans la démonstration de la continuité séparée, nous avons montré qu'il suffit de démontrer le théorème dans le cas où F et G sont compacts, *ce que nous supposons désormais*. Nous avons aussi établi la relation $\langle u \star v, \varphi \rangle = \langle u, \tilde{v} \star \varphi \rangle$ que nous aurons à utiliser. La démonstration se fera en deux étapes.

Montrons d'abord la continuité séquentielle de l'application $(u, \varphi) \mapsto u \star \varphi$ de $\mathcal{D}'_F \times C_K^\infty$ dans C_{F+K}^∞ , où K est un compact, et où on note $\mathcal{D}'_F$ l'espace des distributions à support dans F .

Il faut donc prouver que si $u_j \rightarrow u_0$ dans $\mathcal{D}'_F$ et si $\varphi_j \rightarrow \varphi_0$ dans C_K^∞ , alors les fonctions $\partial^\alpha(u_j \star \varphi_j)$ convergent uniformément vers $\partial^\alpha(u_0 \star \varphi_0)$. Raisonnons par l'absurde en supposant qu'il n'en soit pas ainsi. On pourrait alors, quitte à extraire une sous-suite, trouver une suite x_j de points de $F + K$ qui converge vers un point x_0 telle que la suite $(\partial^\alpha u_j \star \varphi_j)(x_j)$ ne converge pas vers $(\partial^\alpha u_0 \star \varphi_0)(x_0)$.

Or, en notant ψ_j la fonction $y \mapsto \varphi_j(x_j - y)$, on voit facilement que ψ_j converge vers ψ_0 dans C_L^∞ , où L est le compact $K + F + (-K)$. On a donc d'après le résultat (a) du n° C.4.2

$$(\partial^\alpha u_j \star \varphi_j)(x_j) = \langle \partial^\alpha u_j, \psi_j \rangle \rightarrow \langle \partial^\alpha u_0, \psi_0 \rangle = (\partial^\alpha u_0 \star \varphi_0)(x_0),$$

ce qui est la contradiction cherchée.

Revenons à la démonstration du théorème 7.4.9. Nous devons prouver que, pour $\varphi \in C_0^\infty$, on a

$$\langle u_j, \tilde{v}_j \star \varphi \rangle = \langle u_j \star v_j, \varphi \rangle \rightarrow \langle u_0 \star v_0, \varphi \rangle. \quad (\text{C.5})$$

D'après le résultat que nous venons de montrer, la fonction $\tilde{v}_j \star \varphi$ converge dans C_L^∞ vers la fonction $\tilde{v}_0 \star \varphi$, où L est le compact $\text{Supp}(\varphi) + (-G)$. Une nouvelle application du (a) du n° C.4.2 montre que le membre de gauche de (C.5)

converge vers $\langle u_0, \check{v}_0 \star \varphi \rangle$ qui est égal au membre de droite de (C.5). Cela achève la démonstration.

BIBLIOGRAPHIE

I Ouvrages d'une lecture relativement facile.

1. L. Schwartz, Méthodes mathématiques de la Physique. Hermann.
2. Y. Choquet-Bruhat, Distributions, Théorie et Problèmes. Masson (1973).
3. H. Brezis, Analyse fonctionnelle, Théorie et Applications. Masson (1983).
4. J. Arsac, Transformation de Fourier et Théorie des Distributions. Dunod (1961).
5. C. Zuily, Distributions et Équations aux Dérivées partielles : Exercices corrigés. Hermann.

II Pour le lecteur qui souhaite en savoir plus.

1. L. Schwartz, Théorie des Distributions. Hermann.
2. I.M. Gelfand et G.E. Šilov, Les Distributions, tomes 1,2 et 3. Dunod (1962).
3. W. Rudin, Real and Complex Analysis. Mc Graw-Hill (1966), traduction française Masson (1980).
4. M. Reed et B. Simon, Methods of Modern Mathematical Physics, tome 1 : Fonctionnal Analysis. Academic Press.
5. J. Dieudonné, Éléments d'Analyse, tomes 1 et 2. Gauthier-Villars.

III *Enfin le lecteur pourra aborder la lecture des 8 premiers chapitres du tome 1 du traité de L. Hörmander, qui recouvrent largement le contenu de ce cours. Il est préférable de demander conseil sur ce qui peut être sauté en première lecture. La seconde édition contient des exercices.*

L. Hörmander, The Analysis of Linear Partial Differential Operators, Springer-Verlag (1983).

INDEX

A

absolue convergence (d'une intégrale)	17
additivité dénombrable	8
adhérence	30
approximation de l'identité	79
Ascoli (théorème d')	39
Associativité (du produit de convolution)	135, 143

B

Baire (ensemble de première catégorie de)	235
Baire (espace de)	235
Banach (espace de)	46
Banach-Steinhaus (théorème de)	244
base hilbertienne	53
Beppo Levi (théorème de)	8
Bessel-Parseval (égalité de)	54, 66
binôme (formule du)	71
borélienne (mesure)	94
Borel-Lebesgue	35
borne supérieure essentielle	21

C

calcul symbolique de Heaviside	161
--	-----

Cauchy (problème de)	156, 188
Cauchy (suite de)	32
Cauchy-Schwarz (inégalité de)	49
causalité	156
chaleur (équation de la)	187, 216
changement de variable	19
commutativité (du produit de convolution)	144
compact (espace)	35
compacts (suite exhaustive de)	38
complet (espace métrique)	32
condition de Rankine-Hugoniot	112
conjugaison complexe (d'une distribution)	99
conjugaison complexe (et parité)	175
continuité (du produit de convolution)	146
continuité uniforme	33, 45
continûment différentiable (fonction)	69, 221
contractante (application)	34
convergence (au sens de $\mathcal{D}'$)	88
convergence (au sens de $\mathcal{S}'$)	171
convergence dominée	13, 93
convergence en moyenne	20
convergence faible	55
convergence monotone	8
convolutifs (ensembles)	132
convolution (associativité)	135, 143
convolution (commutativité)	144
convolution (continuité)	146
convolution (définition)	76, 135, 142, 143
convolution (dérivation)	78, 144
convolution (équations de)	149, 181
convolution (propriété caractéristique)	138
convolution (translation)	144

couche (distribution de double)	96
couche (distribution de simple)	95
couche (multiple)	96
crochet (dérivation sous le)	131
crochet (intégration sous le)	132
croissance lente (fonction à)	172
croissance lente (suite à)	128, 171

D

d'Alembertien	154
dérivation (sous le crochet)	131
dérivation (sous le signe f)	14
dérivée (d'une distribution)	90
dérivée (d'un produit de convolution)	78, 144
déterminisme	218
différentielle	69, 221
dilatation (d'une distribution)	100
Dirac (distribution de)	87
Dirichlet (problème de)	209
Dirichlet (problème non homogène)	211
distance	29
distance (à un sous-ensemble)	31
distribution (conjuguée complexe d'une)	99
distribution (définition)	88
distribution (dérivée)	90
distribution (dilatation d'une)	100
distribution (de Dirac)	87
distribution (de double couche)	96
distribution (de simple couche)	95
distribution (homogène de degré k)	100
distribution (nulle dans le passé)	155
distribution (ordre d'une)	88
distribution (périodique)	128

distribution (primitives d'une) 105

distribution (produit par une fonction C^∞) 101

distribution (restriction d'une) 99

distribution (série de Fourier d'une d. périodique) 129

distribution (support d'une) 115

distribution (à support compact) 116

distribution (à support réduit à un point) 118

distribution (tempérée) 170

distribution (translatée d'une) 100

divergence 107

dominée (convergence) 13, 93

double couche (distribution de) 96

dual d'un espace de Hilbert 52

dualité (extension à $\mathcal{E}', C^\infty$) 118

dualité (extension à H^{-m}, H^m) 122

dualité (extension à H^{-s}, H^s) 197

dualité (extension à $\mathcal{S}', \mathcal{S}$) 170

E

écart qudratique moyen 21

égalité de Bessel-Parseval 54, 66

élémentaire (solution) 150

ensembles convolutifs 132

ensembles maigres 235

ensembles (mesure des) 8

ensembles négligeables 9

ensembles de première catégorie de Baire 235

équation de la chaleur 187, 216

équation de convolution 149, 181

équation intégrale de première espèce 161

équation intégrale de seconde espèce 160

équation de Laplace 151, 182

équation des ondes 154, 185, 216

équation de Poisson	151, 182
équation de Schrödinger	187, 215
équicontinu (ensemble de fonctions)	39, 244
espace de Baire	235
espace de Banach	46
espace de Fréchet	240
espace de Hilbert	49
espace de Sobolev	120, 195
espace localement convexe métrisable	239
espace vectoriel normé	44
essai (fonction d')	88
exhaustive (suite e. de compacts)	38
extension de la dualité (à $\mathcal{E}', C^\infty$)	118
extension de la dualité (à H^{-m}, H^m)	122
extension de la dualité (à H^{-s}, H^s)	197
extension de la dualité (à $\mathcal{S}', \mathcal{S}$)	170

F

Fatou (théorème de)	10
fermé (sous-ensemble)	29
finies (parties)	98
Fischer-Riesz (théorème de)	58
fonctions à croissance lente	172
fonctions d'essai	88
fonction de Heaviside	92
fondamentale (solution)	150
formule de Green	110
formule de Leibniz	71, 102
formule des sauts	103
formule des sauts dans l'espace	111, 111, 206
formule de Stokes	107, 110
formule de Taylor	71
formule du binôme	71

formule sommatoire de Poisson	129
Fourier (formule d'inversion de)	164, 174
Fourier (série de)	66, 129
Fourier (transformée de , $u \in L^1$)	163
Fourier (transformée de , $u \in \mathcal{S}'$)	173
Fourier (transformée de , $u \in \mathcal{E}'$)	175
Fourier (transformée de , $u \in L^2$)	177
Fourier (transformée de F. partielle)	184
Fréchet (espace de)	240
frontière	30
Fubini (théorème de)	18

G, H

gaussiennes (transformée de Fourier des)	164
Green (formule de)	110
Hadamard (lemme de)	72
harmonique (fonction)	151
Heaviside (calcul symbolique de)	161
Heaviside (fonction de)	92
hermitienne (symétrie)	49, 175
Hilbert (espace de)	49
hilbertienne (base)	53
homogène (distribution)	100
homogénéité dans le temps	218
Hugoniot (condition de Rankine-H.)	112

I, J

identité (approximation de l')	79
inégalité de Cauchy-Schwarz	49
inégalité de Peetre	198
inégalité de Poincaré	208
intégrabilité (au sens de Lebesgue)	11
intégrales (absolument convergentes)	17

intégrales (équations i. de première espèce)	161
intégrales (équations i. de seconde espèce)	34, 160
intégrales (semi-convergentes)	18
intégration (par parties)	110
intégration (sous le crochet)	132
intérieur	30
inversion de Fourier	164, 174
jacobien	19, 223
jacobienne (matrice)	19, 223

L

Laplace (équation de)	151, 182
Lebesgue (intégrabilité au sens de)	11
Lebesgue (mesure de)	8
Lebesgue (théorème de)	13
Leibniz (formule de)	71, 102
lemme de Hadamard	72
lente (fonction à croissance)	172
lente (suite à croissance)	128, 171
Levi (théorème de Beppo-L.)	8
lipschitzienne (application)	34
localement convexe métrisable	239
longueur d'un multiindice	70

M

maigre (ensemble)	235
mesure borélienne	94
mesure d'un ensemble	8
mesure de Lebesgue	8
mesure de Radon	94
mesure de surface	106, 109, 229
monotone (théorème de convergence)	8
moyen (écart quadratique)	21

moyenne (propriété de)	153
multiindice	70
multiple (couche)	96
multipôle	96

N

négligeable	9
Neumann (problème de)	206
normale extérieure unitaire	107, 108
norme	44
nulle dans le passé (distribution)	155

O

ondes (équation des ondes)	154, 185, 216
opérateur de trace	201, 208
ordre d'une distribution	88
orthogonal (supplémentaire)	51
orthonormalisation de Schmidt	54
ouvert	29

P

parité et conjugaison complexe	175
Parseval (égalité de Bessel-)	54, 66
partielle (transformée de Fourier)	184
parties finies	98
partitions de l'unité	43, 75
passé (distribution nulle dans le)	155
Peetre (inégalité de)	198
périodique (distribution)	128
Picard (théorème du point fixe de)	34
Poincaré (inégalité de)	208
Poisson (équation de)	151, 182
Poisson (formule sommatoire de)	129
presque partout (p.p.)	9

primitives d'une distribution	105
principale (valeur)	97
problème de Cauchy	156, 188
problème de Dirichlet	209
problème de Dirichlet non homogène	211
problème de Neumann	206
produit (d'une distribution par une fonction C^∞)	101
projection (sur un convexe fermé)	50
produit (support d'un)	116
propriété caractéristique de la convolution	138
propriété de moyenne	153
Pythagore (théorème de)	49

Q, R

quadratique (écart q. moyen)	21
Radon (mesure de)	94
Rankine-Hugoniot (condition de)	112
Riesz (théorème de)	52
régularisation	80
résiduel	235
restriction d'une distribution	99

S

sauts (formule des)	103
sauts (formule des s. dans l'espace)	111, 111, 206
Schmidt (orthonormalisation de)	54
Schrödinger (équation de)	187, 215
Schwartz (espace $\mathcal{S}$ de)	167
Schwartz (théorème des noyaux de)	141
Schwarz (inégalité de Cauchy-)	49
Schwarz (lemme de)	70
semi-convergente (intégrale)	18
semi-norme	239

séparable (espace vectoriel normé)	46
série de Fourier	66, 129
sesquilinéaire (application)	49
simple couche (distribution de)	95
Sobolev (espace de)	120, 195
solution élémentaire	150
sommatoire (formule s. de Poisson)	129
somme algébrique de deux ensembles	39, 133
Steinhaus (théorème de Banach-)	244
Stone-Weierstrass (théorème de)	40
Stokes (formule de)	107, 110
suite à croissance lente	128, 171
suite de Cauchy	32
suite exhaustive de compacts	38
supplémentaire orthogonal	51
support (distribution à s. compact)	116
support (d'une distribution)	115
support (réduit à un point)	118
surface (mesure de)	106, 109, 229
surgraphe	106
symbolique (calcul s. de Heaviside)	161
symétrie hermitienne	49, 175

T

Taylor (formule de)	71
tempérée (distribution)	170
théorème de Banach-Steinhaus	244
théorème de Beppo Levi	8
théorème de la convergence dominée	13
théorème de la convergence monotone	8
théorème de Fubini	18
théorème de Lebesgue	13
total (sous-ensemble)	46

totalité (critère de) 52

trace 201, 208

transformation de Fourier, $(u \in L^1)$ 163

transformation de Fourier, $(u \in \mathcal{S}')$ 173

transformation de Fourier, $(u \in \mathcal{E}')$ 175

transformation de Fourier, $(u \in L^2)$ 177

transformation de Fourier partielle 184

translation d'une distribution 100

translation d'un produit de convolution 144

U, V

uniforme (continuité) 33, 45

unité (partitions de l')

valeur principale 97

variable (changement de) 19

Index des notations

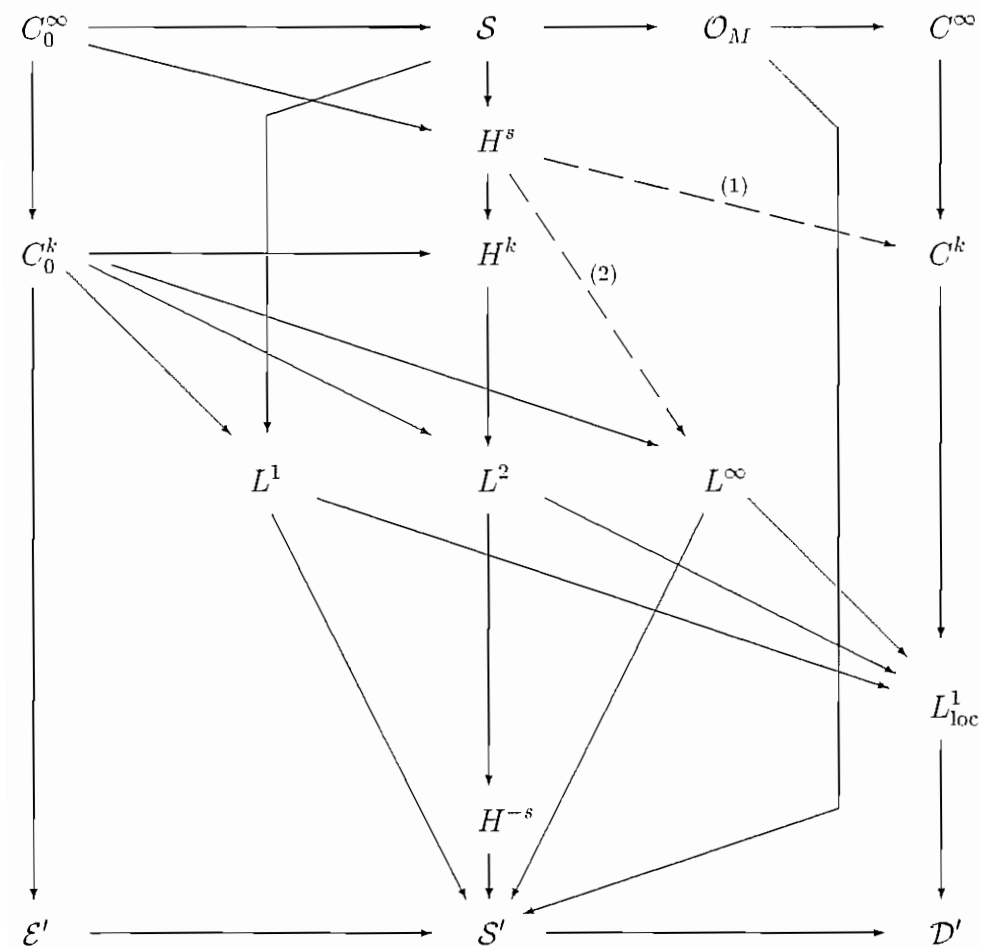
$\langle u, \varphi \rangle$	88	$\mathcal{D}'(\Omega)$	88
$\langle u(x), \varphi(x) \rangle$	88	$\mathcal{D}'_+$	159
$\alpha \leq \beta$	70	$\operatorname{div} X$	107
$\alpha!$	71	$d\sigma$	106, 229
$\binom{\alpha}{\beta}$	71	$\mathcal{E}(\Omega)$	116
$\star, f \star g$	76	$\mathcal{E}'(\Omega)$	116
$\star, u \star \varphi$	135	$\mathcal{F}f$	163
$\star, u \star v$	142	$\mathcal{F}u$	173
$\star, u^{\star-1}$	159	$\mathcal{F}'\varphi, \mathcal{F}'u$	184
$\star, u^{\star n}$	159	H	92
$\tilde{f}, \tilde{u}$	100	H^m	120
$\{f'\}$	103	H^{-m}	121
$\tilde{\Omega}$	106	H^s	195
$C^m(\tilde{\Omega})$	107	J_Φ	19
$\tilde{\varphi}, \tilde{u}$	184	$\mathcal{L}^1(\mathbf{R}^n)$	11
$\nabla'\varphi$	106	$L^1(\mathbf{R}^n)$	58
$C^m(\overline{\Omega})$	109	$\mathcal{L}^2(\mathbf{R}^n)$	21
$\overline{\mathbf{R}}_+$	7	$L^2(\mathbf{R}^n)$	60
$\square u$	154	$\mathcal{L}^\infty(\mathbf{R}^n)$	21
$\widehat{f}$	163	$L^\infty(\mathbf{R}^n)$	62
$\widehat{u}$	173	$L^1_{loc}(\Omega)$	64
γ	201	$L^2_{loc}(\Omega)$	64
δ	87	$L^\infty_{loc}(\Omega)$	64
$u_\lambda, u(x/\lambda)$	100	$\mathcal{N}_p(\varphi)$	167
$\tau_a f, \tau_a u$	100	$\mathcal{O}_M(\mathbf{R}^n)$	172
χ_ε	79	$\operatorname{pf} x_+^\alpha$	98
$C_0^m(\Omega), C_K^m(\Omega), C^m(\Omega)$	73	$\mathbf{R}_+^n$	203
$C_0^\infty(\Omega), C_K^\infty(\Omega), C^\infty(\Omega)$	73	$\mathcal{S}(\mathbf{R}^n)$	167
$\partial^\alpha f, \partial_i f$	70	$\mathcal{S}'(\mathbf{R}^n)$	170
$\partial\Omega$	106, 108	$\operatorname{Supp}(u)$	115
$\mathcal{D}(\Omega)$	88	$\operatorname{vp}(1/x)$	97

Principaux espaces fonctionnels

On a représenté ci-dessous la plus grande partie des espaces de fonctions ou de distributions introduits dans ce cours, k désignant un entier ≥ 0 et s un réel $\geq k$. Une flèche $\mathcal{A} \rightarrow \mathcal{B}$ signifie que l'espace $\mathcal{A}$ est inclus dans $\mathcal{B}$ et que l'inclusion est continue : si $u_j \rightarrow u$ dans $\mathcal{A}$, alors $u_j \rightarrow u$ dans $\mathcal{B}$. Bien entendu, la composée de deux flèches est une flèche (non représentée).

La régularité décroît en lisant le tableau du haut vers le bas, et la "petitesse à l'infini" décroît en lisant de gauche à droite. Une représentation plus fidèle aurait décalé, comme dans la figure 9.3, l'espace $\mathcal{S}$ vers la gauche et l'espace $\mathcal{S}'$ vers la droite.

Il n'y a pas d'inclusion entre les espaces L^1 , L^2 et L^∞ , mais on a $L_{\text{loc}}^\infty \subset L_{\text{loc}}^2 \subset L_{\text{loc}}^1$, espaces qui prendraient place sur la verticale de droite.



(1) Pour $s > k + n/2$.

(2) Pour $s > n/2$.

Table des matières

Introduction	1
1. L'intégrale de Lebesgue	7
1.1. Intégrale des fonctions positives	7
1.2. Fonctions sommables	11
1.3. Cas de la dimension 1	16
1.4. Intégrales multiples	18
1.5. Espaces $\mathcal{L}^1$, $\mathcal{L}^2$, $\mathcal{L}^\infty$	20
1.6. Sur la construction de l'intégrale	22
1.7. Les quatre opérations	25
2. Topologie générale et espaces fonctionnels	29
2.1. Espaces métriques (propriétés topologiques)	29
2.2. Espaces métriques (propriétés uniformes)	32
2.3. Espaces métriques compacts	35
A. Généralités	35
B. Exemples et applications	38
C. Partitions de l'unité	42
2.4. Espaces vectoriels normés	44
2.5. Espaces de Hilbert	48
2.6. Espaces fonctionnels classiques	56
2.7. Séries de Fourier	64

3. Fonctions différentiables et approximation	69
3.1. Espaces de fonctions différentiables	69
3.2. Partitions de l'unité C^∞	73
3.3. Convolution	76
3.4. Régularisation	79
3.5. Approximation dans un ouvert	82
4. Les distributions	85
4.1. Introduction	85
4.2. Définition et convergence	88
4.3. Dérivées	90
4.4. Exemples de distributions	92
A. Fonctions localement sommables	92
B. Mesures de Radon	94
C. Multipôles, couches multiples	96
D. Valeurs principales et parties finies	97
5. Opérations sur les distributions	99
5.1. Opérations élémentaires	99
5.2. Multiplication par les fonctions C^∞	101
5.3. Dérivation (dimension 1)	102
5.4. Dérivation (dimension quelconque)	106
A. Formule de Stokes (cas d'un surgraphe)	106
B. Formule de Stokes (cas d'un ouvert régulier)	108
C. Formule des sauts dans l'espace	111
D. Applications	112
6. Espaces particuliers de distributions	115
6.1. Distributions à support compact	115
6.2. Espaces de Sobolev d'ordre entier	119
A. Notions de régularité	119
B. Définition et propriétés	120
C. Applications	124
6.3. Distributions périodiques	128

7. Convolution	131
7.1. Préliminaires	131
7.2. Convolution d'une distribution et d'une fonction C^∞	135
7.3. Convolution et translations	138
A. Propriété caractéristique de la convolution	138
B. Interprétation physique	140
7.4. Convolution des distributions	142
7.5. Mode d'emploi	146
A. Conditions de Définition	146
B. Propriétés fondamentales	147
C. Modes de calcul	147
8. Quelques équations de la physique mathématique	149
8.1. Généralités sur les équations de convolution	149
8.2. Équations de Laplace et de Poisson	151
8.3. Équation des ondes	154
8.4. Équations différentielles et intégrales	159
9. Transformation de Fourier	163
9.1. Transformation de Fourier des fonctions sommables	163
9.2. L'espace $\mathcal{S}$ de Schwartz	167
9.3. L'espace $\mathcal{S}'$ des distributions tempérées	170
9.4. Transformation de Fourier des distributions tempérées	173
A. Résultats généraux	173
B. Transformation de Fourier dans $\mathcal{E}'$	175
C. Transformation de Fourier dans L^2	177
9.5. Les propriétés fondamentales	178
A. L'échange de la convolution et de la multiplication	178
B. Équations de convolution	181
9.6. Transformation de Fourier partielle et équations d'évolution	184
9.7. Vers l'analyse microlocale	189
9.8. Transformation de Laplace	191

10. Espaces de Sobolev 195

10.1. Structure hilbertienne et dualité 195

10.2. Régularité et caractère local 198

10.3. Traces et prolongements 200

 A. Trace d’une fonction définie dans $\mathbb{R}^n$ 200

 B. L’espace $H^1(\mathbb{R}_+^n)$ 203

10.4. Problème de Dirichlet dans un ouvert régulier 206

 A. Traces 207

 B. Problème de Dirichlet homogène 209

 C. Problème de Dirichlet non homogène 211

 D. Vers l’analyse spectrale 213

10.5. Problème de Cauchy et semi-groupes 214

A. Compléments de calcul différentiel 221

 A.1. Applications différentiables 221

 A.2. Hypersurfaces 224

 A.3. Intégrale de surface 228

 A.4. Cartes et sous-variétés 231

B. Espaces de Baire 235

 B.1. Résultats fondamentaux 235

 B.2. Quelques applications 236

C. Espaces de Fréchet 239

 C.1. Espaces localement convexe métrisables 239

 C.2. Exemples d’espaces de Fréchet 241

 C.3. Le théorème de Banach-Steinhaus 243

 C.4. Continuité des applications bilinéaires 246

Bibliographie 249

Index 251

 Index des notations 262

 Principaux espaces fonctionnels 263

