ANALYSE

Eric CANCES et Alexandre ERN

Septembre 2011

Table des matières

Avant-Propos

Ce document reprend, en le complétant, le contenu du cours d'analyse de première année de l'Ecole des Ponts.

Nous sommes redevables aux Professeurs Mikhael Balabane et Jean-Michel Bony dont les cours d'analyse à l'Ecole des Ponts et à l'Ecole Polytechnique ont inspiré le contenu et la structure de certains chapitres. Nous remercions chaleureusement Guy Bencteux, Virginie Ehrlacher, Jean-Frédéric Gerbeau, Frédéric Legoll, Tony Lelièvre, Régis Monneau, Julien Sabin, Gabriel Stoltz et Florian Thomines pour leurs précieux commentaires sur les versions successives de ce texte.

Introduction

De la mécanique à la finance, en passant par l'environnement et les transports, les domaines scientifiques et techniques couverts par la formation dispensée dans les Ecoles d'ingénieurs font intervenir massivement des modèles reposant sur des équations aux dérivées partielles (EDP).

L'objectif de ce cours est de présenter certains outils mathématiques de base permettant d'appréhender ce type d'équations. Il ne s'agit pas de réaliser une étude exhaustive des principales techniques d'analyse des EDP (ce qui serait impossible dans un cours introductif), mais de traiter complètement un cas important, celui du problème de Poisson sur un ouvert borné avec conditions aux bords de Dirichlet, dont l'énoncé formel s'écrit

$$(I) \qquad \left\{ \begin{array}{l} \text{Chercher une fonction } u\,:\,\Omega \longrightarrow \mathbb{R} \text{ vérifiant} \\ -\Delta u = f \quad \text{ dans } \Omega \\ u = 0 \quad \text{ sur } \partial\Omega \end{array} \right.$$

où Ω désigne un ouvert de $\mathbb{R}^d$, $\partial\Omega$ le bord de Ω et f une fonction donnée définie sur Ω et à valeurs dans $\mathbb{R}$. Les questions auxquelles nous allons tenter de répondre sont les suivantes :

- Ce problème a-t-il une solution (dans une classe à préciser) ?
- Si oui est-elle unique ?
- Dans ce cas, l'application $f \mapsto u$ est-elle continue (là encore, en un sens à préciser) ?

Le problème (I) est un *problème aux limites* en ce sens qu'il fait intervenir une EDP et une condition au bord ($u = 0$ sur $\partial\Omega$). Pour $d = 2$, on pourra se représenter la solution de (I) comme la déformation verticale sous le chargement f d'une membrane élastique à réponse linéaire tendue et dont le bord est maintenu fixe (cf. figure 1).

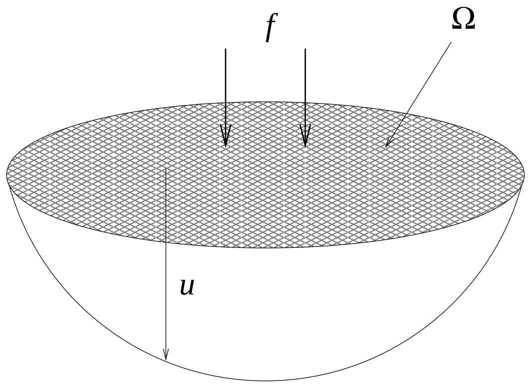


FIG. 1 – Déformation sous un chargement d'une membrane élastique

Le problème (I) présente le triple avantage d'être

1. un problème standard en analyse des EDP,

2. un problème générique en sciences de l'ingénieur (on le retrouve en mécanique, en thermique, en électromagnétisme, en biologie, en finance, ...),
3. le problème aux limites pris comme exemple dans le cours de Calcul Scientifique de l'Ecole des Ponts pour mettre en œuvre la méthode des éléments finis.

L'étude du problème (I) permet également d'illustrer très clairement l'une des idées phares de l'analyse moderne : l'utilisation d'outils de *géométrie* dans des espaces vectoriels de dimension infinie permet d'obtenir simplement et élégamment des résultats d'*analyse*.

Pour mener à bien ce programme, nous devrons tout d'abord nous familiariser avec les espaces vectoriels normés de dimension infinie (chapitre 1) et étudier de plus près les propriétés géométriques

- des espaces de Banach, qui sont les espaces vectoriels normés complets pour la topologie définie par la norme (chapitre 2) ;
- des espaces de Hilbert qui sont les espaces de Banach dont la norme est définie par un produit scalaire (chapitre 3).

Afin d'utiliser les résultats de géométrie démontrés aux chapitres 2 et 3 pour résoudre des problèmes d'analyse, et en particulier le problème (I), il nous faudra construire des espaces vectoriels de fonctions (on parle d'*espaces fonctionnels*) ayant une structure d'espace de Banach ou d'espace de Hilbert. Pour cela, nous introduirons dans un premier temps l'intégrale de Lebesgue (chapitre 4) qui, contrairement à l'intégrale de Riemann, permet de construire des espaces fonctionnels complets.

Le chapitre 5 porte sur la construction et les propriétés des espaces L^p, qui jouent un rôle central en analyse fonctionnelle. Les chapitres 6 et 7 sont consacrés à la théorie des distributions, qui est un outil fondamental de l'analyse moderne ; nous l'utiliserons notamment pour établir le cadre fonctionnel adapté à l'analyse du problème (I) (et plus généralement d'une très large classe de problèmes aux limites), celui des *espaces de Sobolev* (chapitre 8).

Nous nous baserons ensuite sur les résultats “abstraits” de géométrie en dimension infinie établis dans les chapitres 2 et 3 ainsi que sur les espaces fonctionnels construits aux chapitres 5 et 8 pour résoudre le problème aux limites (I) et plusieurs variantes de ce problème ; nous introduirons à cette occasion la notion de *formulation variationnelle* (ou *formulation faible*) d'un problème au limites, et nous démontrerons (et appliquerons) le théorème de Lax-Milgram.

Enfin, nous présenterons la transformation de Fourier, qui est un outil incontournable de l'analyse des phénomènes linéaires, constamment utilisé dans de nombreuses applications, et qui sera mis en pratique dans les cours de Traitement du Signal et d'Analyse en Fréquences de l'Ecole des Ponts.

Chapitre 1

Espaces vectoriels normés

Ce chapitre contient essentiellement des rappels sur les normes, les applications linéaires continues et la dualité. À travers l'étude des exemples présentés dans ce chapitre, on retiendra avant tout le fait que certains résultats bien connus en dimension finie ne sont plus valable en dimension infinie, ce qui justifie l'emploi de nouveaux concepts. Dans un premier temps, nous ne considérerons que deux exemples d'espaces vectoriels de dimension infinie : les espaces C^k dont les éléments sont des fonctions suffisamment régulières et les espaces l^p dont les éléments sont des suites. Notre "collection" d'espaces vectoriels normés de dimension infinie s'enrichira au fil de la lecture de cet ouvrage, notamment après avoir introduit les notions d'intégrale de Lebesgue (chapitres 4 et 5) et de distribution (chapitres 6 à 8).

1.1 Normes

Soit K un corps qu'on prendra égal à $\mathbb{R}$ ou $\mathbb{C}$. Pour $x \in K$, on notera $|x|$ sa valeur absolue si $x \in \mathbb{R}$ et son module si $x \in \mathbb{C}$. Par ailleurs, on utilisera l'abréviation K-ev pour désigner un espace vectoriel sur K.

Définition 1.1. *Soit V un K-ev. Une* norme *est une application définie sur V à valeurs dans $\mathbb{R}_+$, notée $\|\cdot\|_V$, et satisfaisant les trois propriétés suivantes :*
(i) $\|v\|_V = 0 \iff v = 0$,
(ii) $\forall \lambda \in K$, $\forall v \in V$, $\|\lambda v\|_V = |\lambda|\,\|v\|_V$,
(iii) inégalité triangulaire : $\forall (v, w) \in V \times V$, $\|v + w\|_V \leq \|v\|_V + \|w\|_V$.

Exemples en dimension finie. Soit un entier $d \geq 1$ et considérons le K-ev K^d de dimension finie d. On pose pour tout réel p avec $1 \leq p < \infty$

$$\forall v = (v_1, \ldots, v_d) \in K^d, \qquad \|x\|_p = \left(\sum_{i=1}^{d} |v_i|^p\right)^{1/p}, \tag{1.1}$$

et pour $p = \infty$,

$$\forall v = (v_1, \ldots, v_d) \in K^d, \qquad \|v\|_\infty = \max_{1 \leq i \leq d} |v_i|. \tag{1.2}$$

On vérifie facilement que $\|\cdot\|_p$, $1 \leq p \leq \infty$, définit une norme sur K^d. Le seul point technique consiste à montrer l'inégalité triangulaire, qui résulte de l'inégalité de Minkowski (voir exercice 1.1)

$$\forall u, v \in \mathbb{R}_+^d,\ 1 \leq p < \infty, \quad \left(\sum_{i=1}^{d} (u_i + v_i)^p\right)^{1/p} \leq \left(\sum_{i=1}^{d} u_i^p\right)^{1/p} + \left(\sum_{i=1}^{d} v_i^p\right)^{1/p}. \tag{1.3}$$

Les exemples présentés ci-dessus s'étendent de façon naturelle pour construire des normes sur des produits d'espaces vectoriels normés. En effet, nous avons la

Proposition 1.2. *Soit un entier $d \geq 1$. Considérons la famille de K-ev $(V_i)_{1\leq i\leq d}$, chacun des V_i étant équipé d'une norme notée $\|\cdot\|_{V_i}$. Soit un réel $p \in [1, +\infty]$. Alors, l'application de $V = V_1 \times \ldots \times V_d$ dans $\mathbb{R}_+$ définie pour tout $v = (v_1, \ldots, v_d) \in V$ par*

$$\|v\|_p = \left(\sum_{i=1}^{d} \|v_i\|_{V_i}^p\right)^{\frac{1}{p}}, \qquad \text{si } p \in [1, +\infty[,$$
$$\|v\|_\infty = \max_{1\leq i\leq d} \|v_i\|_{V_i}, \qquad \text{si } p = +\infty,$$

est une norme sur V.

Preuve. Laissée en exercice. □

Exemples en dimension infinie.

- Soit $[a, b]$ un intervalle de $\mathbb{R}$ et $\mathcal{C}^0([a, b])$ l'espace vectoriel constitué des fonctions continues sur $[a, b]$ à valeurs dans K. L'application $\|\cdot\|_{\mathcal{C}^0([a,b])}$ qui à $f \in \mathcal{C}^0([a, b])$ associe
$$\|f\|_{\mathcal{C}^0([a,b])} = \sup_{t\in[a,b]} |f(t)|,$$
définit une norme sur $\mathcal{C}^0([a, b])$. Plus généralement, pour un entier $k \geq 0$, on note $C^k([a, b])$ l'espace vectoriel constitué des fonctions k fois continûment dérivables sur $[a, b]$ à valeurs dans K. L'application $\|\cdot\|_{\mathcal{C}^k([a,b])}$ qui à $f \in \mathcal{C}^k([a, b])$ associe
$$\|f\|_{\mathcal{C}^k([a,b])} = \max_{0\leq l\leq k} \sup_{t\in[a,b]} |f^{(l)}(t)|,$$
définit une norme sur $\mathcal{C}^k([a, b])$ (le vérifier en exercice) ;
- Un deuxième exemple d'espace vectoriel normé de dimension infinie est un espace dont les éléments sont des suites de K. Soit $\mathrm{u} = (u_i)_{i\geq 0} \in K^{\mathbb{N}}$ une suite à valeurs dans K. Pour $1 \leq p \leq \infty$, on considère l'ensemble
$$l^p = \left\{ \mathrm{u} \in K^{\mathbb{N}},\ \|\mathrm{u}\|_{l^p} < +\infty \right\},$$
où on a posé
$$\|\mathrm{u}\|_{l^p} = \left(\sum_{i\geq 0} |u_i|^p\right)^{1/p}, \qquad 1 \leq p < \infty$$
et
$$\|\mathrm{u}\|_{l^\infty} = \max_{i\geq 0} |u_i|.$$
On vérifie que l^p est un K-ev et que $\|\cdot\|_{l^p}$ définit bien une norme sur l^p (voir exercice 1.3) ;
- Nous introduirons d'autres espaces vectoriels normés de dimension infinie, plus intéressants pour l'analyse des problèmes posés en sciences de l'ingénieur (EDP, analyse harmonique, ...) lorsque nous disposerons des notions d'intégrale de Lebesgue (chapitres 4 et 5) et de distribution (chapitres 6 à 8).

1.2 Topologie des espaces vectoriels normés

Cette section rappelle brièvement diverses notions d'analyse vues en premier cycle universitaire. Le lecteur désireux d'approfondir ces notions pourra consulter [AF 88, chapitres X et XI].

1.2.1 Ouverts et fermés

Définition 1.3. *Soit V un K-ev normé équipé d'une norme $\|\cdot\|_V$.*

– *On dit qu'un ensemble $O \subset V$ est* ouvert dans V *si*

$$\forall x \in O, \quad \exists r > 0 \text{ tel que } B_x(r) \subset O,$$

où $B_x(r) = \{y \in V, \ \ \|y - x\|_V < r\}$ est la boule ouverte de centre x et de rayon r ;

– *On dit qu'un ensemble $F \subset V$ est* fermé dans V *si l'ensemble $V \setminus F$ est ouvert dans V.*

L'ensemble des ouverts de V selon la définition ci-dessus est appelé la topologie de V induite par la norme $\|\cdot\|_V$.

Remarque 1.4. Lorsqu'il n'y a pas d'ambiguïté, on dira simplement qu'un ensemble est ouvert ou fermé, sans préciser "dans V".

Proposition 1.5. *Soit V un K-ev normé équipé d'une norme $\|\cdot\|_V$. Soit F un fermé dans V et soit $(u_n)_{n\geq 0}$ une suite d'éléments de F. Alors, si $(u_n)_{n\geq 0}$ converge dans V, sa limite est dans F.*

Preuve. Le complémentaire de F étant ouvert, la limite de la suite $(u_n)_{n\geq 0}$ ne peut s'y trouver. □

Définition 1.6. *Soit V un K-ev normé équipé d'une norme $\|\cdot\|_V$ et A un ensemble inclus dans V. On définit* l'adhérence de l'ensemble A dans V *et on note $\overline{A}^V$ l'ensemble des points de V qui sont limites de points de A.*

Proposition 1.7. *F est fermé dans V si et seulement si $F = \overline{F}^V$.*

Preuve. Laissée en exercice. □

Définition 1.8. *On dit qu'un ensemble $A \subset V$ est* dense *dans V si $\overline{A}^V = V$.*

1.2.2 Équivalence de normes

Sur un espace vectoriel, on peut considérer plusieurs normes et il est légitime de se demander si les topologies induites sont les mêmes. Pour cela, nous introduisons la notion d'équivalence de normes.

Définition 1.9. *Soit V un K-ev. On dit que deux normes $\|\cdot\|_{V,1}$ et $\|\cdot\|_{V,2}$ sont* équivalentes *sur V s'il existe deux constantes c_1 et c_2 strictement positives telles que*

$$\forall v \in V, \qquad c_1 \|v\|_{V,2} \leq \|v\|_{V,1} \leq c_2 \|v\|_{V,2}. \tag{1.4}$$

L'importance de la définition 1.9 provient du fait que si deux normes sont équivalentes, elles induisent la même topologie.

Proposition 1.10. *Soit V un K-ev équipé de deux normes équivalentes $\|\cdot\|_{V,1}$ et $\|\cdot\|_{V,2}$. Soit $O \subset V$. Alors, O est ouvert pour $\|\cdot\|_{V,1}$ si et seulement si O est ouvert pour $\|\cdot\|_{V,2}$.*

Preuve. Laissée en exercice. □

Proposition 1.11. *Si V est un K-ev de dimension finie, toutes les normes sont équivalentes.*

Preuve. Voir exercice 1.4. □

Contre-exemples en dimension infinie. La proposition 1.11 ne s'étend pas aux espaces vectoriels normés de dimension infinie. Donnons quelques contre-exemples.

– Sur $C^0([0,1])$, l'application $f \mapsto \|f\|_{L^1} = \int_0^1 |f(s)|\,ds$ définit bien une norme, mais celle-ci n'est pas équivalente à la norme $\|\cdot\|_{C^0([0,1])}$. Plus précisément, on a clairement la majoration $\|f\|_{L^1} \leq \|f\|_{C^0([0,1])}$ pour tout $f \in C^0([0,1])$. En revanche, il n'existe pas de constante c telle que $\|f\|_{C^0([0,1])} \leq c\|f\|_{L^1}$ (on pourra considérer pour tout $\varepsilon > 0$, les fonctions f_ε définies par $f_\varepsilon(t) = 1 - t/\varepsilon$ pour $t \in [0,\varepsilon]$ et $f_\varepsilon(t) = 0$ sinon (voir figure 1.1) ; ces fonctions satisfont $\|f_\varepsilon\|_{C^0([0,1])} = 1$ et $\|f_\varepsilon\|_{L^1} = \varepsilon/2$) ;

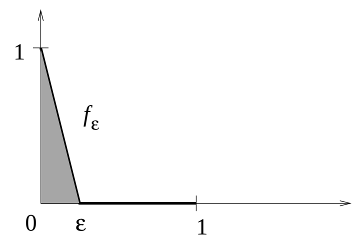


FIG. 1.1 – Contre-exemple pour l'équivalence des normes en dimension infinie.

– $\|\cdot\|_{C^0([a,b])}$ définit bien une norme sur $C^1([a,b])$ mais elle n'est pas équivalente à la norme $\|\cdot\|_{C^1([a,b])}$. La construction d'un contre-exemple est laissée en exercice ;
– Sur l^1, l'application

$$\|\cdot\|_h \,:\, \mathtt{u} \mapsto \|\mathtt{u}\|_h = \sum_{i\geq 0} \tfrac{1}{i+1}|u_i|,$$

définit une norme et on a clairement $\|\mathtt{u}\|_h \leq \|\mathtt{u}\|_{l^1}$, $\forall \mathtt{u} \in l^1$. En revanche, il n'existe pas de constante c telle que pour tout $\mathtt{u} \in l^1$, on ait $\|\mathtt{u}\|_{l^1} \leq c\,\|\mathtt{u}\|_h$. En effet, en considérant pour tout entier $N > 0$, la suite $u^N \in l^1$ telle que $u_i^N = 1$ si $i \leq N$ et 0 sinon, on aurait une inégalité de la forme $N \leq c \log N$ avec c indépendant de N, ce qui est absurde.

Remarque 1.12. Si V est un K-ev normé et F un sous-espace *de dimension finie*, alors F est nécessairement fermé. Ce n'est plus le cas si F est de dimension infinie. À titre de contre-exemple, on pourra considérer $C^1([-1,1])$ comme sous-espace de $C^0([-1,1])$ équipé de la norme $\|\cdot\|_{C^0([-1,1])}$ et la famille de fonctions $f_\varepsilon \in C^1([-1,1])$ définies par $f_\varepsilon(t) = |t|$ si $|t| \geq \varepsilon$ et $f_\varepsilon(t) = \frac{\varepsilon}{2}(1 + (\frac{t}{\varepsilon})^2)$ sinon (voir figure 1.2).

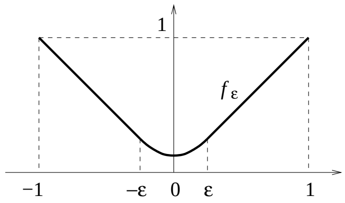


FIG. 1.2 – Illustration de la remarque 1.12.

1.2.3 Compacité

Définition 1.13. *Soit V un K-ev normé. Un ensemble $A \subset V$ est* compact *si de toute suite d'éléments de A on peut extraire une sous-suite convergente.*

Proposition 1.14. *Si A est compact, alors A est fermé et borné.*

Preuve. Vérification immédiate. □

Proposition 1.15. *Si V est de dimension finie, alors A est compact si et seulement si A est fermé et borné.*

Preuve. Voir exercice 1.5. □

On peut en fait énoncer un résultat plus général qui souligne bien la différence entre la dimension finie et la dimension infinie (voir par exemple [AF 88, p. 582] ou [Brézis 83, p. 92]).

Théorème 1.16. *Soit V un K-ev. La boule unité (fermée) de V est compacte si et seulement si V est de dimension finie.*

1.3 Applications linéaires continues

1.3.1 Définition et premières propriétés

Dans cette section, V et W désignent deux K-ev normés équipés respectivement des normes $\|\cdot\|_V$ et $\|\cdot\|_W$.

Proposition 1.17. *Soit A une application linéaire de V dans W. Les propositions suivantes sont équivalentes :*

(1) A est continue.
(2) A est continue en 0.
(3) Il existe une constante $c \in \mathbb{R}_+$ telle que

$$\forall u \in V, \qquad \|Au\|_W \leq c\,\|u\|_V.$$

Preuve. Les implications (1) $\Longrightarrow$ (2) et (3) $\Longrightarrow$ (1) sont évidentes. Il nous reste donc à prouver l'implication (2) $\Longrightarrow$ (3). Soit A une application linéaire continue en 0. Il existe $\eta > 0$ tel que pour tout $u \in B(0, \eta)$, la boule fermée de centre 0 et de rayon η, on ait

$$\|Au - \underbrace{A0}_{=0}\|_W \leq 1.$$

Soit maintenant $u \in V$ avec $u \neq 0$. Le vecteur $v = \frac{\eta}{\|u\|_V} u$ appartient à $B_0(\eta)$ si bien que $\|Av\|_W \leq 1$, ce qui, par linéarité, implique (3) avec $c = \frac{1}{\eta}$. Enfin, l'inégalité dans (3) est trivialement satisfaite si $u = 0$. $\square$

Proposition 1.18. *Soit A une application linéaire de V dans W. Supposons V de dimension finie. Alors, A est nécessairement continue.*

Preuve. Soit $(e_1, \dots, e_d)$ une base de V. Pour tout $u = \sum_{i=1}^{d} u_i e_i \in V$, on a

$$\|Au\|_W \leq \sum_{i=1}^{d} |u_i| \, \|Ae_i\|_W \leq \left(\sum_{i=1}^{d} \|Ae_i\|_W \right) \|u\|_\infty,$$

et on conclut en utilisant l'équivalence des normes sur V. $\square$

Exemples et contre-exemples en dimension infinie.

- L'application qui à $f \in C^1([a,b])$ associe $f' \in C^0([a,b])$ est linéaire continue lorsque ces espaces sont équipés des normes $\|\cdot\|_{C^1([a,b])}$ et $\|\cdot\|_{C^0([a,b])}$.
- Cette application n'est pas continue si $C^1([a,b])$ est équipé de la norme $|f(a)| + \int_a^b |f'(s)|\, ds$ (voir exercice 1.7).
- Plus généralement, d'autres contre-exemples peuvent être construits en considérant deux normes $\|\cdot\|_{V,1}$ et $\|\cdot\|_{V,2}$ qui ne sont pas équivalentes. En effet, la continuité de l'injection de V équipé de $\|\cdot\|_{V,1}$ dans V équipé de $\|\cdot\|_{V,2}$ équivaut à l'existence d'une constante $c > 0$ telle que $\forall u \in V$, $\|u\|_{V,2} \leq c\|u\|_{V,1}$.

Définition 1.19. *L'espace vectoriel des applications linéaires continues de V dans W est noté $\mathcal{L}(V, W)$.*

Proposition 1.20. *L'application*

$$\begin{array}{rcl} \|\cdot\|_{\mathcal{L}(V,W)} : \mathcal{L}(V,W) & \longrightarrow & \mathbb{R}_+ \\ A & \longmapsto & \|A\|_{\mathcal{L}(V,W)} = \sup_{\substack{u \in V \\ u \neq 0}} \frac{\|Au\|_W}{\|u\|_V} \end{array}$$

est une norme sur $\mathcal{L}(V, W)$.

Preuve. (i) $\|A\|_{\mathcal{L}(V,W)} = 0$ implique $Au = 0$ pour tout $u \in V$ avec $u \neq 0$. Par ailleurs, $Au = 0$ pour $u = 0$. D'où $A = 0$.
(ii) Soit $\lambda \in K$. On a

$$\|(\lambda A)u\|_W = |\lambda| \, \|Au\|_W \leq |\lambda| \, \|A\|_{\mathcal{L}(V,W)} \, \|u\|_V,$$

d'où on déduit que $\|\lambda A\|_{\mathcal{L}(V,W)} \leq |\lambda| \, \|A\|_{\mathcal{L}(V,W)}$. Pour $\lambda \neq 0$, on a donc

$$\|A\|_{\mathcal{L}(V,W)} = \|\tfrac{1}{\lambda} \lambda A\|_{\mathcal{L}(V,W)} \leq \tfrac{1}{|\lambda|} \|\lambda A\|_{\mathcal{L}(V,W)}$$

si bien que $\|\lambda A\|_{\mathcal{L}(V,W)} = |\lambda| \, \|A\|_{\mathcal{L}(V,W)}$ et l'égalité est trivialement satisfaite si $\lambda = 0$.
(iii) Pour A et $B \in \mathcal{L}(V, W)$, on a

$$\|(A+B)u\|_W \leq \|Au\|_W + \|Bu\|_W \leq (\|A\|_{\mathcal{L}(V,W)} + \|B\|_{\mathcal{L}(V,W)})\|u\|_V,$$

si bien que $\|A+B\|_{\mathcal{L}(V,W)} \leq \|A\|_{\mathcal{L}(V,W)} + \|B\|_{\mathcal{L}(V,W)}$. □

1.3.2 Dualité

Un cas particulier important d'applications linéaires continues entre K-ev normés est celui où l'espace d'arrivée est $W = K$.

Définition 1.21. *Soit V un K-ev normé. $\mathcal{L}(V,K)$ est appelé l'*espace dual *de V et est noté V'. Un élément $A \in V'$ est appelé une* forme linéaire continue *et son action sur un élément $v \in V$ est notée à l'aide du crochet de dualité $\langle \cdot, \cdot \rangle_{V',V}$. V' est équipé de la norme canonique*

$$\|A\|_{V'} = \sup_{\substack{u \in V \\ u \neq 0}} \frac{|\langle A, u \rangle_{V',V}|}{\|u\|_V}.$$

Remarque 1.22. La notion introduite dans la définition 1.21 est celle de dual *topologique* que l'on distingue du dual *algébrique* dans lequel la propriété de continuité n'est pas requise. En dimension finie, le dual topologique coïncide avec le dual algébrique mais en dimension infinie, ces deux notions sont différentes. Dans ce cours, seule la notion de dual topologique sera utilisée.

Exemples.
- Soit $[a,b] \subset \mathbb{R}$ et $c \in [a,b]$. L'application qui à $f \in C^0([a,b])$ associe $f(c)$ est une forme linéaire continue sur $C^0([a,b])$ de norme 1 ;
- L'application qui à $\mathrm{u} \in l^p$, $1 \leq p \leq \infty$, associe u_0 est une forme linéaire continue sur l^p de norme 1 ;
- Soit $1 \leq p < \infty$ et p' tel que $\frac{1}{p} + \frac{1}{p'} = 1$ ($p' = \infty$ si $p = 1$). Soit $\mathrm{v} \in l^{p'}$. Alors, en posant pour tout $u \in l^p$, $\langle \tilde{\mathrm{v}}, \mathrm{u} \rangle_{(l^p)', l^p} = \sum_{i \geq 0} u_i v_i$, on définit une forme linéaire continue sur l^p, i.e. $\tilde{\mathrm{v}} \in (l^p)'$. De plus, l'application

  $$\mathrm{v} \in l^{p'} \mapsto \tilde{\mathrm{v}} \in (l^p)',$$

  est linéaire, bijective et isométrique. Ce résultat, dont la preuve fait l'objet de l'exercice 1.8, permet de faire l'identification $(l^p)' = l^{p'}$ pour $1 \leq p < \infty$. En revanche, le dual de l^∞ ne coïncide pas avec l^1 mais est strictement plus grand que l^1. En fait, on peut identifier l^1 au dual topologique du K-ev des suites de K à support fini, cet espace étant strictement inclus dans l^∞.

Proposition 1.23. *Soit Z et V deux K-ev équipés respectivement des normes $\|\cdot\|_Z$ et $\|\cdot\|_V$. Supposons que $Z \subset V$ avec injection continue, i.e., qu'il existe $c_i \in \mathbb{R}_+$ telle que $\forall z \in Z$, $\|z\|_V \leq c_i \|z\|_Z$. Alors, $V' \subset Z'$ avec injection continue.*

Preuve. Soit $f \in V'$. Considérons l'application $Tf : Z \to \mathbb{R}$ qui à $z \in Z$ associe $Tf(z) = \langle f, z \rangle_{V',V}$. Il est clair que Tf est linéaire. De plus, Tf est continue sur Z puisque

$$\forall z \in Z, \quad Tf(z) = \langle f, z \rangle_{V',V} \leq \|f\|_{V'} \|z\|_V \leq c_i \|f\|_{V'} \|z\|_Z.$$

Par conséquent, $Tf \in Z'$ et $\|Tf\|_{Z'} \leq c_i \|f\|_{V'}$, ce qui complète la preuve. □

Nous concluons cette section avec deux résultats qui nous serviront par la suite : le théorème de prolongement d'une forme linéaire et une caractérisation des sous-espaces denses. Ces résultats, dont la preuve repose sur un théorème profond d'analyse, le théorème de Hahn-Banach, sont admis (on pourra consulter [Brézis 83, p. 3 et 7]).

Théorème 1.24. *Soit V un K-ev normé et Z un sous-espace vectoriel de V. Soit $z \in Z'$ une forme linéaire continue de norme $\|z\|_{Z'}$. Alors, il existe $A \in V'$ qui prolonge z, i.e.*

$$\forall u \in Z, \quad \langle z, u\rangle_{Z',Z} = \langle A, u\rangle_{V',V},$$

et telle que $\|A\|_{V'} = \|z\|_{Z'}$.

Théorème 1.25. *Soit V un K-ev normé et Z un sous-espace vectoriel de V. Si toute forme linéaire continue $f \in V'$ qui s'annule sur Z est identiquement nulle sur V, alors Z est dense dans V.*

1.3.3 Formes bilinéaires continues

Définition 1.26. *Soit V et W deux K-ev. Une* forme bilinéaire *sur $V \times W$ est une application $a : V \times W \to K$ telle que*
- *$\forall v \in V$, l'application $a(v, \cdot) : W \to K$ est linéaire ;*
- *$\forall w \in W$, l'application $a(\cdot, w) : V \to K$ est linéaire.*

Définition 1.27. *Soit V et W deux K-ev. On dit qu'une forme bilinéaire $a : V \times W \to K$ est* continue *sur $V \times W$ s'il existe une constante $c \in \mathbb{R}_+$ telle que*

$$\forall (v, w) \in V \times W, \qquad |a(v, w)| \leq c\, \|v\|_V\, \|w\|_W.$$

On note

$$\|a\| = \sup_{\substack{(v,w)\in V\times W \\ v\neq 0, w\neq 0}} \frac{a(v,w)}{\|v\|_V\, \|w\|_W}.$$

L'intérêt des formes bilinéaires continues est qu'elles interviennent fréquemment dans la formulation faible des EDP (cf. chapitre 9). Dans ce cadre, il est intéressant de garder à l'esprit qu'une forme bilinéaire continue de $V \times W$ à valeurs dans K peut s'interpréter comme un élément A de $\mathcal{L}(V, W')$ en posant

$$\forall w \in W, \qquad \langle Av, w\rangle_{W',W} = a(v, w).$$

On vérifie aisément que $Av \in W'$ et que

$$\|A\|_{\mathcal{L}(V,W')} = \|a\|.$$

1.4 Exercices

$\triangleright$ **1.1.** Montrer l'inégalité de Minkowski (1.3). Pour cela, on utilisera la concavité de la fonction $f : \mathbb{R}_+ \to \mathbb{R}$, $t \mapsto (1 + t^{1/p})^p$ qui implique que pour des $\lambda_i \geq 0$ tels que $\sum_{i=1}^d \lambda_i = 1$ et des $t_i \geq 0$, on a

$$f\left(\sum_{i=1}^d \lambda_i t_i\right) \geq \sum_{i=1}^d \lambda_i f(t_i).$$

▷ **1.2.** Vérifier que l'application $\|\cdot\|_{\mathcal{C}^k([a,b])}$ définit une norme sur $\mathcal{C}^k([a,b])$.

▷ **1.3.** L'objet de cet exercice est de montrer que les l^p sont bien des espaces vectoriels normés.

1. Vérifier que l^p est un espace vectoriel ;
2. Soient $1 \leq p, p' \leq \infty$ avec $\frac{1}{p} + \frac{1}{p'} = 1$ et N un entier positif. Montrer l'inégalité de Hölder
$$\forall u, v \in \mathbb{R}_+^N, \quad \sum_{i=1}^N u_i v_i \leq \left(\sum_{i=1}^N u_i^p\right)^{1/p} \left(\sum_{i=1}^N v_i^{p'}\right)^{1/p'}. \tag{1.5}$$
Indication : en utilisant la concavité de la fonction $f \ : \ \mathbb{R}_+ \to \mathbb{R}$, $t \mapsto \log t$, on montrera d'abord que pour u et v réels positifs, on a l'inégalité de Young (sur $\mathbb{R}$)
$$uv \leq \frac{u^p}{p} + \frac{v^{p'}}{p'}. \tag{1.6}$$
Puis, considérant x et $y \in \mathbb{R}^N$, on appliquera l'inégalité ci-dessus à $u = x_i/\|x\|_p$ et $v = y_i/\|y\|_{p'}$;
3. Déduire de (1.5) que pour $\mathtt{u} \in l^p$ et $\mathtt{v} \in l^{p'}$, la suite produit $\mathtt{uv}$ définie par $\mathtt{uv} = (u_i v_i)_{i \geq 0}$ est dans l^1 et que l'on a
$$\|\mathtt{uv}\|_{l^1} \leq \|\mathtt{u}\|_{l^p} \, \|\mathtt{v}\|_{l^{p'}} \ ; \tag{1.7}$$
4. Montrer que $\|\cdot\|$ satisfait bien l'inégalité triangulaire. Indication : on pourra utiliser l'inégalité de Hölder (1.7) en remarquant que pour $\mathtt{u}, \mathtt{v} \in l^p$, $|\mathtt{u} + \mathtt{v}|^{p-1} \in l^{p'}$;
5. Déduire l'inégalité de Minkowski (1.3) de l'inégalité de Hölder (1.5).

▷ **1.4.** L'objectif de cet exercice est de montrer que si V est un K-ev de dimension finie, alors toutes les normes sur V sont équivalentes. Soit $(e_1, \ldots, e_d)$ une base de V. Pour $v \in V$, on note $(v_1, \ldots, v_d)$ les coordonnées de v dans cette base.

1. Vérifier qu'il suffit de montrer que toutes les normes de V sont équivalentes à la norme $\|\cdot\|_\infty$ donnée par (1.2) ;
2. Soit $\|\cdot\|_V$ une norme de V. De l'inégalité triangulaire, déduire qu'il existe une constante $c_2 > 0$ telle que $\forall v \in V$, $\|v\|_V \leq c_2 \, \|v\|_\infty$;
3. Soit B_∞ la boule unité (fermée) de V pour la norme $\|\cdot\|_\infty$. Déduire de la question précédente que l'application
$$\begin{array}{rccc} id \ : & (B_\infty, \|\cdot\|_\infty) & \to & (B_\infty, \|\cdot\|_V) \\ & v & \mapsto & v, \end{array}$$
est continue ;
4. En utilisant un argument de compacité, montrer qu'il existe une constante $c_1 > 0$ telle que $\forall v \in V$, $\|v\|_\infty \leq c_1 \, \|v\|_V$.

▷ **1.5.** L'objet de cet exercice est de montrer qu'en dimension finie, tout fermé borné est compact.

1. La première étape fait appel à la notion de complétude qui est introduite dans le chapitre 2. En admettant que $\mathbb{R}$ est complet, montrer que sur $\mathbb{R}$, tout fermé borné est compact ;
2. En déduire que sur $\mathbb{R}^d$ équipé de la norme $\|\cdot\|_\infty$, tout fermé borné est compact ;
3. Conclure en utilisant l'équivalence des normes en dimension finie.

▷ **1.6.** Construire un contre-exemple montrant que les normes $\|\cdot\|_{C^0([a,b])}$ et $\|\cdot\|_{C^1([a,b])}$ ne sont pas équivalentes sur $C^1([a,b])$.

▷ **1.7.** Vérifier que l'application $f \mapsto |f(0)| + \int_a^b |f'(s)|\,ds$ définit une norme sur $C^1([a,b])$. En s'inspirant de la figure 1.1, construire un contre-exemple montrant que l'application qui à $f \in C^1([a,b])$ associe $f' \in C^0([a,b])$ n'est pas continue si $C^1([a,b])$ est équipé de la norme ci-dessus et $C^0([a,b])$ de la norme canonique $\|\cdot\|_{C^0([a,b])}$.

▷ **1.8.** L'objet de cet exercice est de montrer que pour $1 \leq p < \infty$, on peut faire l'identification $(l^p)' = l^{p'}$ où $\frac{1}{p} + \frac{1}{p'} = 1$ (on pose $p' = \infty$ si $p = 1$). L'identification se fait par le biais de l'application

$$\begin{array}{rcl} T\ :\ l^{p'} & \to & (l^p)' \\ \mathtt{v} & \mapsto & \tilde{\mathtt{v}} \quad \text{avec} \quad \forall \mathtt{u} \in l^p,\ \ \langle \tilde{\mathtt{v}}, \mathtt{u}\rangle_{(l^p)',l^p} = \sum_{i\geq 0} u_i v_i. \end{array}$$

1. On commence par traiter le cas $1 < p < \infty$. Montrer que l'application T est bien définie. Indication : en utilisant l'inégalité de Hölder (1.7), montrer que si $\mathtt{u} \in l^p$ et $\mathtt{v} \in l^{p'}$, la série $\sum_{i\geq 0} u_i v_i$ est absolument convergente. En déduire que $T(\mathtt{v}) \in (l^p)'$ avec $\|T(\mathtt{v})\|_{(l^p)'} \leq \|\mathtt{v}\|_{l^{p'}}$;
2. Montrer que l'application T est linéaire et injective.
3. Montrer que l'application T est surjective. Indication : soit $\varphi \in (l^p)'$ donné. Pour tout entier $i \geq 0$, poser $\mathtt{e}_i = (\delta_{ij})_{j\geq 0}$ où δ_{ij} est le symbole de Kronecker et introduire la suite $\mathtt{v}$ de terme général $v_i = \varphi(\mathtt{e}_i)$. Montrer que $\mathtt{v} \in l^{p'}$ et que $T(\mathtt{v}) = \varphi$;
4. Montrer que T est une isométrie. Indication : il suffit de montrer que $\|\mathtt{v}\|_{l^{p'}} \leq \|T(\mathtt{v})\|_{(l^p)'}$. Pour cela, se donner une suite $\mathtt{v} \in l^{p'}$ et considérer la suite $\mathtt{u}$ de terme général $u_i = |v_i|^{p'-1}\mathrm{sgn}(v_i)$;
5. Reprendre les étapes précédentes dans le cas où $p = 1$ et $p' = \infty$.

▷ **1.9.** Soit $\mathtt{u} \in \mathbb{R}^{\mathbb{N}}$. On note A l'opérateur qui à $\mathtt{u}$ associe la suite $\mathtt{v}$ de terme général $v_i = (-1)^i u_i$. Montrer que pour $1 \leq p \leq \infty$, $A \in \mathcal{L}(l^p, l^p)$ et que A est de norme 1. Reprendre l'exercice avec l'opérateur qui à $\mathtt{u}$ associe la suite $\mathtt{v}$ de terme général $v_i = u_{i+1}$.

▷ **1.10.** Soit γ un réel non nul. On considère l'opérateur A qui à $f \in C^0(\mathbb{R})$ associe $g = Af \in C^0(\mathbb{R})$ définie par $g(t) = f(\gamma t)$ pour $t \in \mathbb{R}$. Montrer que $A \in \mathcal{L}(C^0(\mathbb{R}), C^0(\mathbb{R}))$ et que A est de norme 1.

Chapitre 2

Espaces de Banach

Ce chapitre introduit la notion de complétude : un espace vectoriel normé est dit complet si toute suite de Cauchy y est convergente. De tels espaces, appelés espaces de Banach, interviendront fréquemment dans la suite de ce cours car ils jouissent de plusieurs propriétés remarquables. Nous démontrerons notamment le théorème du point fixe de Picard (qui ne s'applique que dans les espaces de Banach) et nous donnerons un premier exemple d'application de ce théorème en prouvant l'existence et l'unicité de la solution pour une certaine classe d'équations différentielles ordinaires.

2.1 Définitions et premiers exemples

Dans cette section, nous introduisons les notions de suite de Cauchy, d'espace complet et d'espace de Banach.

2.1.1 Suites de Cauchy

Définition 2.1. *Soit V un K-ev normé. On dit qu'une suite $(u_n)_{n\geq 0}$ d'éléments de V est* de Cauchy *si on a la propriété suivante*

$$\forall \varepsilon > 0, \quad \exists N \geq 0, \quad \forall n \geq N, \, \forall p \geq 0, \quad \|u_{n+p} - u_n\|_V \leq \varepsilon. \tag{2.1}$$

Proposition 2.2. *Toute suite convergente est de Cauchy.*

Preuve. Conséquence immédiate de (2.1). □

Proposition 2.3. *Toute suite de Cauchy est bornée.*

Preuve. Immédiate. □

Proposition 2.4. *Soit V_1 et V_2 deux K-ev normés et $(u_n)_{n\geq 0}$ une suite de Cauchy dans V_1. Soit $f : V_1 \to V_2$ une application uniformément continue. Alors, la suite $(f(u_n))_{n\geq 0}$ est de Cauchy dans V_2.*

Preuve. Laissée en exercice. □

2.1.2 Espaces complets

La réciproque de la proposition 2.2 n'étant pas vraie d'une façon générale, cela motive la définition suivante.

Définition 2.5. *Un K-ev V est dit* complet pour la norme $\|\cdot\|_V$ *si toute suite de Cauchy est convergente. Un K-ev complet pour sa norme est appelé un* espace de Banach.

Proposition 2.6. *Soit V un K-ev équipé de deux normes équivalentes $\|\cdot\|_{V,1}$ et $\|\cdot\|_{V,2}$. Alors, V est complet pour la norme $\|\cdot\|_{V,1}$ si et seulement si il est complet pour la norme $\|\cdot\|_{V,2}$.*

Preuve. Supposons V complet pour la norme $\|\cdot\|_{V,1}$. Soit $(u_n)_{n\geq 0}$ une suite de Cauchy dans V pour la norme $\|\cdot\|_{V,2}$. De l'inégalité $\|v\|_{V,1} \leq c_2 \|v\|_{V,2}$, nous déduisons que $(u_n)_{n\geq 0}$ est une suite de Cauchy pour la norme $\|\cdot\|_{V,1}$. Elle est donc convergente par hypothèse. De l'inégalité $\|v\|_{V,2} \leq c_1 \|v\|_{V,1}$, nous déduisons qu'elle converge également pour la norme $\|\cdot\|_{V,2}$. □

L'étude des espaces complets est très simple en dimension finie, puisque nous avons le résultat suivant.

Proposition 2.7. *Tout espace vectoriel normé de dimension finie est complet.*

Preuve. (i) Nous admettons que $\mathbb{R}$ est complet, propriété qui résulte de la construction de $\mathbb{R}$ (voir par exemple [AF 88, chapitre I]).
(ii) Considérons maintenant un $\mathbb{R}$-ev de dimension finie d. Soit $(u_n)_{n\geq 0}$ une suite de Cauchy dans cet espace. Notons $(u_{n,1}, \ldots, u_{n,d})$ les composantes de u_n dans une base donnée. Toutes les normes étant équivalentes en dimension finie d'après la proposition 1.11, nous pouvons utiliser la norme $\|\cdot\|_\infty$ et en déduire que pour tout i, $1 \leq i \leq d$, les $(u_{n,i})_{n\geq 0}$ sont des suites de Cauchy dans $\mathbb{R}$. Ces suites sont donc convergentes vers une limite l_i, $1 \leq i \leq d$, et il est immédiat de prouver que $(u_n)_{n\geq 0}$ converge vers $(l_1, \ldots, l_d)$.
(iii) Le cas d'un $\mathbb{C}$-ev se traite de manière identique grâce au fait que $\mathbb{C}$ est un $\mathbb{R}$-ev de dimension deux. □

Remarque 2.8. Du fait du résultat précédent, l'expression "espace de Banach" ne s'emploie guère en dimension finie. Elle est plutôt réservée à la dimension infinie où la notion de complétude est plus riche.

Donnons deux exemples puis un contre-exemple d'espaces de Banach.

Proposition 2.9. *Pour $1 \leq p \leq \infty$, l^p équipé de la norme $\|\cdot\|_{l^p}$ est un espace de Banach.*

Preuve. Voir exercice 2.1. □

Proposition 2.10. *$C^0([a,b])$ équipé de la norme $\|\cdot\|_{C^0([a,b])}$ est un espace de Banach.*

Preuve. Soit $(u_n)_{n\geq 0}$ une suite de Cauchy dans $C^0([a,b])$, i.e.,

$$\forall\varepsilon>0,\quad \exists N\geq 0,\quad \forall n\geq N,\ \forall p\geq 0,\quad \|u_{n+p}-u_n\|_{C^0([a,b])}=\sup_{x\in[a,b]}|u_{n+p}(x)-u_n(x)|\leq\varepsilon.$$

Pour tout $x\in[a,b]$, $(u_n(x))_{n\geq 0}$ est de Cauchy dans $\mathbb{R}$ et converge donc vers une limite notée $u(x)$. En faisant tendre p vers l'infini, on obtient

$$\forall\varepsilon>0,\quad \exists N\geq 0,\quad \forall n\geq N,\quad \sup_{x\in[a,b]}|u(x)-u_n(x)|\leq\varepsilon. \tag{2.2}$$

Montrons que u est une fonction continue sur $[a,b]$. Soit $x\in[a,b]$ et $\varepsilon>0$. On sait qu'il existe $N\in\mathbb{N}$ tel que

$$\sup_{z\in[a,b]}|u(z)-u_N(z)|\leq\varepsilon/3.$$

Par ailleurs, u_N étant continue, il existe V_x, voisinage de x dans $[a,b]$ tel que $|u_N(y)-u_N(x)|\leq\varepsilon/3$ pour $y\in V_x$. L'inégalité triangulaire permet de conclure : pour tout $y\in V_x$, on a

$$|u(y)-u(x)|\leq|u(y)-u_N(y)|+|u_N(y)-u_N(x)|+|u(x)-u_N(x)|\leq\varepsilon.$$

Cela signifie que $u\in C^0([a,b])$. On conclut en remarquant que (2.2) signifie que $(u_n)_{n\geq 0}$ converge vers u dans $C^0([a,b])$. $\square$

Contre-exemple. $C^0([0,1])$ équipé de la norme $\|f\|_{L^1}=\int_0^1|f(s)|\,ds$ n'est pas un espace de Banach. En effet, considérons la suite de fonctions $(f_n)_{n\geq 1}$ définie par (voir figure 2.1)

$$f_n(t)=\begin{cases}0 & 0\leq t\leq\frac{1}{2}-\frac{1}{n}\\ \frac{n}{2}(t-\frac{1}{2}+\frac{1}{n}) & \frac{1}{2}-\frac{1}{n}\leq t\leq\frac{1}{2}+\frac{1}{n}\\ 1 & \frac{1}{2}+\frac{1}{n}\leq t\leq 1\end{cases}$$

La suite $(f_n)_{n\geq 1}$ converge en norme $\|\cdot\|_{L^1}$ vers la fonction f_* qui vaut 0 si $0\leq t\leq\frac{1}{2}$ et 1 sinon. Elle est donc de Cauchy pour cette norme. Cependant, la fonction f_* n'appartient pas à $C^0([0,1])$.

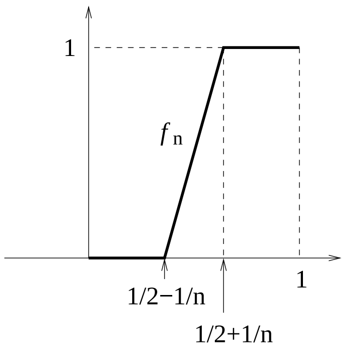


FIG. 2.1 – Graphe de la fonction f_n.

Proposition 2.11. *Soit V un espace de Banach et F un sous-espace* fermé *de V. Alors, F équipé de la norme induite par V est un espace de Banach.*

Preuve. Toute suite de Cauchy dans F est de Cauchy dans V donc convergente dans V. Comme F est fermé, la limite est dans F. $\square$

Proposition 2.12. *Soit un entier $d \geq 1$. Considérons une famille d'espaces de Banach $(V_i)_{1 \leq i \leq d}$, chacun des V_i étant équipé d'une norme notée $\|\cdot\|_{V_i}$. Soit V le K-ev $V = V_1 \times \ldots \times V_d$ équipé de la norme $\|\cdot\|_p$ définie dans la proposition 1.2. Alors, V est un espace de Banach.*

Preuve. Laissée en exercice. □

2.2 Applications linéaires continues dans les espaces de Banach

Théorème 2.13. *Soit V un espace vectoriel normé et W un espace de Banach. Alors, $\mathcal{L}(V, W)$ équipé de la norme $\|\cdot\|_{\mathcal{L}(V,W)}$ est un espace de Banach.*

Preuve. Soit $(A_n)_{n\geq 0}$ une suite de Cauchy dans $\mathcal{L}(V, W)$, i.e.,

$$\forall \varepsilon > 0, \quad \exists N \geq 0, \quad \forall n \geq N, \ \forall p \geq 0, \quad \|A_{n+p} - A_n\|_{\mathcal{L}(V,W)} = \sup_{\substack{u \in V \\ u \neq 0}} \frac{\|A_{n+p}u - A_n u\|_W}{\|u\|_V} \leq \varepsilon.$$

Pour tout $u \in V$, $u \neq 0$, $(A_n u)_{n\geq 0}$ est donc une suite de Cauchy dans W. En notant $l \in W$ sa limite, nous définissons une application A de V dans W qui à $u \in V$ associe $Au = l \in W$. Il est clair que A est linéaire. En faisant tendre $p \to \infty$, on obtient

$$\forall \varepsilon > 0, \quad \exists N \geq 0, \quad \forall n \geq N, \quad \forall u \in V, \ \|Au - A_n u\|_W \leq \varepsilon \, \|u\|_V. \tag{2.3}$$

Pour $\varepsilon = 1$, on en déduit en particulier que

$$\forall u \in V, \quad \|Au\|_W \leq \|Au - A_N u\|_W + \|A_N u\|_W \leq (1 + \|A_N\|_{\mathcal{L}(V,W)}) \, \|u\|_V,$$

ce qui montre que A est continue. Enfin, (2.3) exprime le fait que A_n tend vers A dans $\mathcal{L}(V, W)$. □

Corollaire 2.14. *Soit V un espace vectoriel normé. Alors, son espace dual V' est un espace de Banach.*

Preuve. Conséquence directe du théorème 2.13 puisque K, de dimension finie, est complet. □

2.3 Séries normalement convergentes

Un des intérêts des espaces de Banach est que l'étude de la convergence des séries peut se faire de façon particulièrement simple en étudiant leur convergence normale.

Définition 2.15. *Soit V un espace vectoriel normé. Soit $(u_n)_{n\geq 0}$ une suite d'éléments de V. On dit que la série de terme général u_n est* normalement convergente *si la série $\sum_{n\geq 0} \|u_n\|_V$ est convergente.*

Théorème 2.16. *Soit V un espace vectoriel normé. V est un espace de Banach si et seulement si toute série normalement convergente est convergente.*

Preuve. (i) Supposons que V soit un espace de Banach. Soit $S_n = \sum_{m=0}^{n} u_m$ une série normalement convergente de terme général u_n. On constate que S_n est une suite de Cauchy puisque

$$\|S_{n+p} - S_n\|_V \leq \|u_{n+1}\|_V + \ldots + \|u_{n+p}\|_V \leq \sum_{m\geq n+1}^{\infty} \|u_m\|_V \underset{n\to\infty}{\longrightarrow} 0.$$

V étant complet, S_n est convergente.
(ii) Réciproquement, supposons que toute série normalement convergente soit convergente. Soit $(u_n)_{n\geq 0}$ une suite de Cauchy. Nous pouvons extraire une sous-suite $v_n = u_{\varphi(n)}$ de sorte que

$$\|v_{n+1} - v_n\|_V \leq \frac{1}{2^n}.$$

La série de terme général $v_{n+1} - v_n$ est donc normalement convergente. Comme

$$S_n = \sum_{m=0}^{n} v_{m+1} - v_m = v_{n+1} - v_0,$$

nous en déduisons que la suite extraite $(v_n)_{n\geq 0}$ est convergente. Comme $(u_n)_{n\geq 0}$ est une suite de Cauchy, il résulte de (2.1) que toute la suite $(u_n)_{n\geq 0}$ est convergente. □

2.4 Théorème du point fixe de Picard

Définition 2.17. *Soit V un espace vectoriel normé et A une application de V dans V. On dit que A est* contractante *s'il existe un réel $\alpha < 1$ tel que*

$$\forall u, v \in V, \quad \|Au - Av\|_V \leq \alpha \|u - v\|_V. \tag{2.4}$$

Remarque 2.18. Une application contractante est continue.

Théorème 2.19. *Soit V un espace de Banach et A une application contractante de V dans V. Alors, l'équation*

$$Au = u,$$

admet une solution unique $u_* \in V$. Cette solution est appelée le point fixe *de A.*

Preuve. (i) Existence : soit $u_0 \in V$ quelconque et $(u_n)_{n\geq 0}$ la suite définie par $u_{n+1} = Au_n$. On constate que

$$\|u_{n+1} - u_n\|_V = \|Au_n - Au_{n-1}\|_V \leq \alpha \|u_n - u_{n-1}\|_V \leq \alpha^n \|u_1 - u_0\|_V,$$

d'où on déduit que la série de terme général $u_{n+1} - u_n$ est normalement convergente. L'espace V étant complet, nous déduisons du théorème 2.16 que la série de terme général $u_{n+1} - u_n$ est convergente, i.e., que la suite $(u_n)_{n\geq 0}$ est convergente. En notant u_* sa limite et en passant à la limite dans la relation $u_{n+1} = Au_n$, il vient $Au_* = u_*$ puisque (2.4) implique la continuité de A.
(ii) L'unicité se montre par l'absurde : si on a 2 solutions u_1 et u_2 avec $u_1 \neq u_2$, on obtient

$$\|u_1 - u_2\|_V = \|Au_1 - Au_2\|_V \leq \alpha \|u_1 - u_2\|_V < \|u_1 - u_2\|_V,$$

ce qui est absurde. □

2.5 Application aux équations différentielles ordinaires

Le théorème du point fixe de Picard sert par exemple à montrer l'existence et unicité globale de la solution pour une certaine classe d'équations différentielles ordinaires. Nous considérons le problème suivant.

Définition 2.20. *(Problème de Cauchy). Soit $d \geq 1$ un entier, $u_0 \in \mathbb{R}^d$ et f une fonction définie sur $\mathbb{R}^d$ à valeurs dans $\mathbb{R}^d$. On appelle* problème de Cauchy *le problème suivant : trouver $u \in C^1([0,+\infty);\mathbb{R}^d)$ telle que*

$$\begin{cases} \frac{du}{dt}(t) = f(u(t)), & t \geq 0, \\ u(0) = u_0. \end{cases} \tag{2.5}$$

Théorème 2.21. *(Théorème de Cauchy-Lipschitz). On suppose que f est lipschitzienne sur $\mathbb{R}^d$, i.e., qu'il existe une constante $L \in \mathbb{R}_+$ telle que*

$$\forall u_1, u_2 \in \mathbb{R}^d, \quad \|f(u_1) - f(u_2)\| \leq L\, \|u_1 - u_2\|,$$

où $\|\cdot\|$ désigne une norme quelconque sur $\mathbb{R}^d$. Alors, le problème (2.5) admet une solution u et une seule dans $C^1([0,+\infty[;\mathbb{R}^d)$.

Remarque 2.22. En changeant t en $-t$, il est clair que dans le cadre des hypothèses du théorème 2.21, le problème (2.5) admet une solution unique $u \in C^1(]-\infty,,+\infty);\mathbb{R}^d[$.

Preuve du théorème 2.21. (i) Nous allons démontrer d'abord l'existence et l'unicité locale en temps en utilisant le théorème du point fixe. Posons $T = 1/(2L)$ et notons $E = C^0([0,T];\mathbb{R}^d)$ l'espace des fonctions continues sur $[0,T]$ à valeurs dans $\mathbb{R}^d$ équipé de la norme

$$\forall u \in E, \quad \|u\|_E = \sup_{t\in[0,T]} \|u(t)\|.$$

On vérifie facilement que E muni de cette norme est un espace de Banach. Il est clair par ailleurs que la fonction

$$t \mapsto u_0 + \int_0^t f(u(s))\,ds,$$

est de classe C^1, sa dérivée valant $f(u(t))$. On peut donc définir l'application

$$\begin{array}{rcl} A : E & \to & E \\ u & \mapsto & Au \quad \text{avec} \quad Au(t) = u_0 + \int_0^t f(u(s))\,ds, \end{array}$$

qui vérifie la propriété suivante : $u \in E$ est solution de (2.5) sur $[0,T]$ si et seulement si $Au = u$. Or, A est contractante puisque pour u_1 et $u_2 \in E$, on a

$$\|Au_2 - Au_1\|_E = \sup_{t\in[0,T]} \left\| \int_0^t (f(u_2(s)) - f(u_1(s)))\,ds \right\| \leq TL\|u_2 - u_1\|_E,$$

et $TL = \frac{1}{2}$. Nous déduisons du théorème du point fixe que le système (2.5) admet une solution unique sur $[0,T]$.
(ii) En appliquant le résultat d'existence et unicité locale au problème de Cauchy

$$\begin{cases} \frac{d}{dt}v(t) = f(v), \\ v(0) = u(T), \end{cases}$$

et en posant pour $t \in [T, 2T]$, $u(t) = v(t-T)$, on obtient l'existence et unicité de la solution du système (2.5) sur $[0,2T]$. On recouvre ainsi de proche en proche l'intervalle $[0,\infty[$. Enfin, en changeant t en $-t$ et f en $-f$, on obtient l'existence et l'unicité de la solution du système (2.5) sur $]-\infty,+\infty[$. $\square$

Remarque 2.23. Le caractère lipschitzien de f est essentiel pour assurer l'unicité de la solution. Ainsi, pour $d = 1$ et $f(u) = \sqrt{u}$, (2.5) avec $u_0 = 0$ admet la famille de solutions

$$u_c(t) = \begin{cases} 0 & t \leq c \\ \frac{1}{4}(t-c)^2 & t \geq c \end{cases}$$

pour tout réel $c > 0$.

Remarque 2.24. On peut cependant affaiblir les hypothèses du théorème 2.21. Soit Ω un ouvert non vide de $\mathbb{R} \times \mathbb{R}^d$, $(t_0, u_0) \in \Omega$ et f une application de Ω dans $\mathbb{R}^d$ continue sur Ω et localement lipschitzienne en la deuxième variable, i.e. tout point $(t, u) \in \Omega$ admet un voisinage $V_{(t,u)} \subset \Omega$ tel que pour un réel L (pouvant dépendre de $V_{(t,u)}$), on ait

$$\forall (t, u_1), (t, u_2) \in V_{(t,u)}, \quad \|f(t, u_1) - f(t, u_2)\| \leq L \|u_1 - u_2\|.$$

Dans ces conditions, le problème de Cauchy

$$\begin{cases} \frac{d}{dt}u(t) = f(t, u), \\ u(t_0) = u_0, \end{cases}$$

admet une unique solution maximale (i.e. une solution qu'on ne peut prolonger ni à droite ni à gauche). Celle-ci est définie sur un intervalle ouvert de $\mathbb{R}$. Par exemple, $f(t, u) = 1 + u^2$ est continue et localement lipschitzienne en u ; le problème de Cauchy avec la condition initiale $u(0) = 0$ admet pour unique solution maximale la fonction $u(t) = \tan t$ définie sur l'intervalle ouvert $]-\pi/2, +\pi/2[$.

2.6 Exercices

$\triangleright$ **2.1.** L'objet de cet exercice est de montrer que pour $1 \leq p \leq \infty$, l^p équipé de la norme $\|\cdot\|_{l^p}$ est un espace de Banach.

1. On traitera d'abord le cas $1 \leq p < \infty$. Soit $(\mathtt{u}_n)_{n\geq 0}$ une suite de Cauchy dans l^p. On note $u_{n,m}$ le m-ième terme de la suite $\mathtt{u}_n$. Montrer que pour tout $m \geq 0$, la suite $(u_{n,m})_{n\geq 0}$ est de Cauchy dans K. On notera l_m sa limite et $\mathtt{l}$ la suite de terme général l_m ;
2. Montrer que $\mathtt{l} \in l^p$. Indication : montrer qu'il existe une constante C telle que pour tout $n \geq 0$ et pour tout $I \geq 0$,
$$\left(\sum_{i=0}^{I} |u_{n,i}|^p\right)^{1/p} \leq \|\mathtt{u}_n\|_{l^p} \leq C \ ;$$
3. Montrer que $(\mathtt{u}_n)_{n\geq 0}$ converge vers $\mathtt{l}$ dans l^p. Indication : en se donnant $\varepsilon > 0$, montrer qu'il existe $N \geq 0$ tel que pour $n \geq N$ et $q \geq 0$ et pour tout $I \geq 0$, on ait
$$\left(\sum_{i=0}^{I} |u_{n+q,i} - u_{n,i}|^p\right)^{1/p} \leq \|\mathtt{u}_{n+q} - \mathtt{u}_n\|_{l^p} \leq \varepsilon \ ;$$
4. Reprendre les étapes ci-dessus dans le cas $p = \infty$.

$\triangleright$ **2.2.** Soit E un espace de Banach. On note $F = C^0([a, b]; E)$ l'espace vectoriel des fonctions définies et continues sur $[a, b]$ à valeurs dans E. On considère sur F l'application $\|\cdot\|_F : F \to \mathbb{R}$ qui à $f \in F$ associe $\|f\|_F = \max_{t\in[a,b]} \|f(t)\|_E$.

– Vérifier que $\|\cdot\|_F$ définit bien une norme sur F.

– En reprenant la preuve de la proposition 2.10, montrer que F équipé de la norme $\|\cdot\|_F$ est un espace de Banach.

▷ **2.3.** Soit un réel p avec $1 < p < \infty$. Montrer par un contre-exemple que $C^0([0,1])$ équipé de la norme $\|f\|_{L^p} = (\int_0^1 |f(s)|^p \, ds)^{1/p}$ n'est pas un espace de Banach.

▷ **2.4.** Soit V un espace de Banach et A une application de V dans V. Pour un entier $p > 0$, on note $A^{(p)}$ l'application $A \circ \ldots \circ A$ obtenue en composant l'application A p fois avec elle-même. On suppose trouvé un entier $p > 0$ tel que $A^{(p)}$ soit contractante. Montrer que A admet un point fixe unique.

▷ **2.5.** On note E l'espace de Banach $C^0([a,b])$ équipé de la norme $\|\cdot\|_{C^0([a,b])}$. Soit $k : [a,b]^2 \to \mathbb{R}$ une fonction continue sur $[a,b]^2$ et $\rho = \max_{(x,y)\in[a,b]^2} |k(x,y)|$. Pour $f \in E$, on note Af l'application de $[a,b]$ dans $\mathbb{R}$ définie par $Af(x) = \int_a^b k(x,y)f(y)\,dy$.

1. Montrer que A est une application linéaire continue de E dans E.
2. Montrer que si $\rho|b-a| < 1$, A est contractante. On conservera cette hypothèse pour la suite de l'exercice.
3. Soit g donnée dans E. Montrer que l'équation intégrale
$$h(x) - \int_a^b k(x,y)h(y)\,dy = g(x), \quad x \in [a,b],$$
admet une solution unique h dans E.
4. Montrer que la série de terme général $A^{(n)}g$ est convergente puis que $h = \sum_{n\geq 0} A^{(n)}g$.

▷ **2.6.** On note c_0 l'espace vectoriel formé par les suites de nombres (réels ou complexes) qui tendent vers 0. On munit c_0 de la norme définie par
$$\forall u = (u_n)_{n\geq 0} \in c_0, \quad \|u\|_{c_0} = \max_{n\in\mathbb{N}} |u_n|.$$
Pour tout $k \in \mathbb{N}$, on note e^k la suite $(e^k_n)_{n\geq 0} \in c_0$ définie par $e^k_n = \delta_{kn}$ pour tout $n \in \mathbb{N}$.

1. Soit $a = (a_n)_{n\geq 0} \in l^1$. Vérifier que l'application linéaire
$$L_a(u) = \sum_{n\in\mathbb{N}} a_n\, u_n$$
définit une forme linéaire continue sur c_0, puis montrer que
$$\|L_a\|_{c_0'} = \|a\|_{l^1}.$$
2. Soit $\Lambda \in c_0'$, $\lambda_k = \Lambda(e_k)$ et
$$u^l = \sum_{0\leq k\leq l,\ \lambda_k\neq 0} \frac{|\lambda_k|}{\lambda_k} e^k.$$
Montrer que pour tout $l \in \mathbb{N}$, $u^l \in c_0$, $\|u^l\|_{c_0} = 1$ et
$$\Lambda(u^l) = \sum_{k=1}^{l} |\lambda_k|.$$
En déduire que $\lambda = (\lambda_k)_{k\geq 0} \in l^1$.
3. Montrer que $\Lambda = L_\lambda$.
4. En déduire que l^1 s'identifie au dual topologique de c_0.

Chapitre 3

Espaces de Hilbert

Les espaces de Hilbert jouent un rôle fondamental en analyse car ils fournissent un cadre mathématique général pour "faire de la géométrie en dimension infinie". Dans ce chapitre, ces considérations seront illustrées par plusieurs résultats importants, notamment le théorème de projection orthogonale, le théorème de Riesz et la notion de base hilbertienne. Nous évoquerons aussi le calcul différentiel et les problèmes d'optimisation (avec ou sans contrainte) dans les espaces de Hilbert.

3.1 Définitions

Dans cette section, nous introduisons la notion de produit scalaire et celle d'espace de Hilbert. Pour plus de clarté, nous distinguons les espaces vectoriels réels et complexes.

3.1.1 Espaces euclidiens

Définition 3.1. *Soit V un $\mathbb{R}$-ev. Un* produit scalaire *sur V est une application bilinéaire sur V, notée $(\cdot,\cdot)_V : V \times V \to \mathbb{R}$, satisfaisant les propriétés suivantes :*

(i) symétrie : $\forall (v,w) \in V \times V$, $(v,w)_V = (w,v)_V$;

(ii) positivité : $\forall v \in V$, $(v,v)_V \geq 0$;

(iii) $(v,v)_V = 0 \iff v = 0$.

Un $\mathbb{R}$-ev équipé d'un produit scalaire est appelé un espace euclidien.

3.1.2 Espaces hermitiens

Définition 3.2. *Soit V un $\mathbb{C}$-ev. Un* produit scalaire *sur V est une application sesquilinéaire à gauche sur V, notée $(\cdot,\cdot)_V : V \times V \to \mathbb{C}$, possédant les propriétés suivantes :*

(i) hermiticité : $\forall v,w \in V$, $(v,w)_V = \overline{(w,v)_V}$;

(ii) positivité : $\forall v \in V$, $(v,v)_V \geq 0$;

(iii) $(v,v)_V = 0 \iff v = 0$.

On rappelle que la sesquilinéarité à gauche signifie que

$$\forall (\lambda_1,\lambda_2) \in \mathbb{C} \times \mathbb{C},\ \forall (v_1,v_2,w) \in V \times V \times V, \qquad \begin{cases} (\lambda_1 v_1 + \lambda_2 v_2, w)_V = \overline{\lambda}_1 (v_1,w)_V + \overline{\lambda}_2 (v_2,w)_V, \\ (w, \lambda_1 v_1 + \lambda_2 v_2)_V = \lambda_1 (w,v_1)_V + \lambda_2 (w,v_2)_V. \end{cases}$$

Un $\mathbb{C}$-ev équipé d'un produit scalaire est appelé un espace hermitien.

Remarque 3.3. On peut définir de façon analogue les formes sesquilinéaires à droite et convenir de considérer les formes sesquilinéaires à droite pour définir les produits scalaires.

3.1.3 Inégalité de Cauchy-Schwarz

Proposition 3.4. *Soit V un K-ev et $(\cdot,\cdot)_V$ un produit scalaire sur V. L'application $\|\cdot\|_V : V \to \mathbb{R}$ définie par*

$$\forall v \in V, \quad \|v\|_V = (v,v)_V^{1/2}, \tag{3.1}$$

est une norme sur V appelée norme induite par le produit scalaire. *De plus, on a l'inégalité de Cauchy-Schwarz*

$$\forall v, w \in V, \quad |(v,w)_V| \le \|v\|_V \, \|w\|_V. \tag{3.2}$$

Preuve. Prouvons le théorème ci-dessus pour $K = \mathbb{R}$ (le cas $K = \mathbb{C}$ se traite de façon analogue). Commençons par établir l'inégalité de Cauchy–Schwarz (3.2). Soit $(v,w) \in V \times V$. Supposons $w \neq 0$ (sinon, l'inégalité (3.2) devient une égalité triviale). En développant le membre de droite, on constate que

$$\|v\|_V^2 - \frac{|(v,w)_V|^2}{\|w\|_V^2} = \left\| v - \frac{(v,w)_V}{\|w\|_V^2} w \right\|_V^2 \ge 0.$$

Ceci montre l'inégalité (3.2). De plus, il est clair qu'il y a égalité si et seulement si le membre de droite est nul, i.e. si et seulement si v et w sont colinéaires.

Vérifions maintenant que l'application $\|\cdot\|_V$ définie par (3.1) est bien une norme. Les assertions (i) et (ii) de la définition 1.1 résultent, respectivement, de l'assertion (iii) de la définition 3.1 et de la bilinéarité (ou sesquilinéarité) du produit scalaire. Enfin, l'inégalité triangulaire se déduit de l'inégalité de Cauchy–Schwarz : pour tout $(v,w) \in V \times V$,

$$\begin{aligned}
\|v+w\|_V^2 &= \|v\|_V^2 + 2(v,w)_V + \|w\|_V^2 \\
&\le \|v\|_V^2 + 2|(v,w)_V| + \|w\|_V^2 \\
&\le \|v\|_V^2 + 2\|v\|_V \, \|w\|_V + \|w\|_V^2 \\
&= (\|v\|_V + \|w\|_V)^2,
\end{aligned}$$

ce qui complète la preuve. □

Remarque 3.5. L'inégalité de Cauchy-Schwarz (3.2) exprime la continuité du produit scalaire dans la topologie induite.

3.1.4 Espaces de Hilbert

Définition 3.6. *Un* espace de Hilbert *est un espace euclidien ou hermitien complet pour la norme induite par le produit scalaire.*

Exemples en dimension finie. Soit un entier $d \ge 1$.

– on munit $\mathbb{R}^d$ d'une structure hilbertienne en considérant le produit scalaire euclidien

$$\forall x, y \in \mathbb{R}^d, \quad (x,y)_{\mathbb{R}^d} = \sum_{i=1}^{d} x_i \, y_i.$$

– De façon analogue, $\mathbb{C}^d$ est muni de la structure hilbertienne définie par le produit scalaire canonique

$$\forall x, y \in \mathbb{C}^d, \quad (x,y)_{\mathbb{C}^d} = \sum_{i=1}^{d} \overline{x_i} \, y_i.$$

Exemples en dimension infinie. On munit l'espace de Banach l^2 des suites réelles de carré sommable d'une structure hilbertienne en considérant le produit scalaire

$$\forall \mathtt{u}, \mathtt{v} \in l^2, \quad (\mathtt{u}, \mathtt{v})_{l^2} = \sum_{i \geq 0} u_i \, v_i.$$

Lorsque les suites considérées sont complexes, le produit scalaire est

$$\forall \mathtt{u}, \mathtt{v} \in l^2, \quad (\mathtt{u}, \mathtt{v})_{l^2} = \sum_{i \geq 0} \overline{u_i} \, v_i.$$

Proposition 3.7. *Soit V un espace de Hilbert et F un sous-espace* fermé *de V. Alors, F équipé du produit scalaire induit par V est un espace de Hilbert.*

Preuve. Laissée en exercice □

3.2 Théorème de projection orthogonale

Théorème 3.8. *Soit V un espace de Hilbert et K un sous-espace vectoriel fermé de V. Pour tout $u \in V$, il existe un unique $v = P_K u \in K$, appelé* projection orthogonale *de u sur K (voir figure 3.1), tel que*

$$\|P_K u - u\|_V = \inf_{w \in K} \|w - u\|_V. \tag{3.3}$$

$P_K u$ est caractérisé par

$$P_K u \in K \quad \text{et} \quad (P_K u - u, w)_V = 0, \, \forall w \in K. \tag{3.4}$$

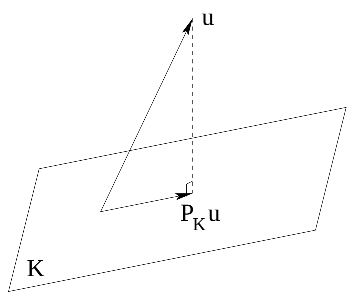


FIG. 3.1 – Illustration du théorème de projection orthogonale.

Preuve. (i) On commence par prouver la formule de la médiane

$$\forall s, t \in V, \quad \tfrac{1}{2}\|s+t\|_V^2 + \tfrac{1}{2}\|s-t\|_V^2 = \|s\|_V^2 + \|t\|_V^2,$$

qui résulte simplement de l'identité $\|s \pm t\|_V^2 = \|s\|_V^2 \pm 2(s,t)_V + \|t\|_V^2$.
(ii) Soit $(v_n)_{n \geq 0}$ une suite minimisante, i.e., une suite d'éléments $v_n \in K$ telle que

$$\lim_{n \to \infty} \|v_n - u\|_V = \inf_{w \in K} \|w - u\|_V.$$

Montrons que $(v_n)_{n\geq 0}$ est de Cauchy. En appliquant la formule de la médiane à $s = v_n - u$ et $t = v_{n+p} - u$, on obtient

$$\begin{aligned}\|v_{n+p} - v_n\|_V^2 &= 2\left(\|v_{n+p} - u\|_V^2 + \|v_n - u\|_V^2 - 2\|\tfrac{1}{2}(v_{n+p} + v_n) - u\|_V^2\right)\\ &\leq 2\left(\|v_{n+p} - u\|_V^2 + \|v_n - u\|_V^2 - 2(\inf_{w\in K} \|w - u\|_V)^2\right),\end{aligned}$$

d'où on déduit aisément que $\lim_{n\to\infty,p>0} \|v_{n+p} - v_n\|_V = 0$. L'espace V étant complet, $\lim_{n\to\infty} v_n = v$ et comme K est fermé, $v \in K$. De plus, $\|v - u\|_V = \inf_{w\in K} \|w - u\|_V$ par continuité de la norme sur V. Ceci montre l'existence d'une solution pour (3.3).
(iii) L'unicité de la solution de (3.3) découle directement de la formule de la médiane : si v_1 et v_2 sont deux solutions de (3.3), en appliquant la formule de la médiane avec $s = u - v_1$ et $t = u - v_2$, on obtient $\|v_1 - v_2\|_V^2 \leq 0$, i.e., $v_1 = v_2$.
(iv) Il nous reste à prouver l'équivalence entre (3.3) et (3.4). Soit $P_K u$ la solution du problème de minimisation (3.3). Soit $w \in K$. Pour $t \in \mathbb{R}$, la fonction

$$\varphi(t) = \|P_K u + tw - u\|_V^2 = \|P_K u - u\|_V^2 + 2t(P_K u - u, w)_V + t^2\|w\|_V^2,$$

est un polynôme de degré 2 en t. De plus, φ est minimale en $t = 0$ puisque $P_K u + tw \in K$ pour tout $t \in \mathbb{R}$. Par conséquent, on a $\varphi'(0) = 0$, i.e., $(P_K u - u, w)_V = 0$. w étant arbitraire, nous avons prouvé que $P_K u$ satisfaisait la caractérisation (3.4). Réciproquement, si $P_K u$ est solution de (3.4), pour tout $w \in K$, on a

$$\begin{aligned}\|u - w\|_V^2 &= \|(u - P_K u) + (P_K u - w)\|_V^2\\ &= \|u - P_K u\|_V^2 + 2(u - P_K u, P_K u - w)_V + \|P_K u - w\|_V^2\\ &= \|u - P_K u\|_V^2 + \|P_K u - w\|_V^2\\ &\geq \|u - P_K u\|_V^2,\end{aligned}$$

puisque $P_K u - w \in K$ si bien que $(u - P_K u, P_K u - w)_V = 0$ par hypothèse. On a donc bien montré que $P_K u$ était solution de (3.3). □

Corollaire 3.9. *$P_K \in \mathcal{L}(V, K)$ et $\|P_K\|_{\mathcal{L}(V,K)} = 1$ si $K \neq \{0\}$.*

Preuve. La linéarité de P_K résulte de la caractérisation (3.4). En effet, pour tout $\alpha, \beta \in \mathbb{R}$ et $u_1, u_2 \in V$, on a

$$(\alpha u_1 + \beta u_2 - (\alpha P_K u_1 + \beta P_K u_2), w)_V = 0, \quad \forall w \in K,$$

par définition de $P_K u_1$ et $P_K u_2$ et par linéarité du produit scalaire. Ceci montre que

$$P_K(\alpha u_1 + \beta u_2) = \alpha P_K u_1 + \beta P_K u_2.$$

Par ailleurs, comme $P_K u \in K$, nous déduisons de la caractérisation (3.4) que $(u - P_K u, P_K u)_V = 0$ si bien que

$$\|P_K u\|_V^2 = \|u - P_K u\|_V^2 + \|u\|_V^2 \leq \|u\|_V^2,$$

ce qui prouve la continuité de P_K avec $\|P_K\|_{\mathcal{L}(V,K)} \leq 1$. Enfin, si $K \neq \{0\}$, on a pour tout $w \in K \setminus \{0\}$, $P_K w = w$ ce qui implique que $\|P_K\|_{\mathcal{L}(V,K)} \geq 1$. D'où finalement $\|P_K\|_{\mathcal{L}(V,K)} = 1$. □

Proposition 3.10. *Pour tout $u_1, u_2 \in V$, on a*

$$\|P_K u_1 - P_K u_2\|_V \leq \|u_1 - u_2\|_V.$$

Preuve. Laissée en exercice. □

Exemple (simple). Soit K le sous-espace vectoriel fermé de l^2 engendré par toutes les suites dont le premier terme est nul et on note P_K la projection orthogonale de l^2 sur K. On vérifie facilement que pour $\mathtt{u} \in l^2$, $P_K\mathtt{u} = \mathtt{v}$ avec $v_0 = 0$ et $v_i = u_i$ pour $i \geq 1$. Un exemple plus sophistiqué de projection orthogonale, mais qui anticipe sur des notions introduites au chapitre 5, est proposé dans l'exercice 3.4.

3.3 Théorème de Riesz

Théorème 3.11. *Soit V un espace de Hilbert. Étant donné $\varphi \in V'$, il existe $u \in V$ unique tel que*

$$\forall w \in V, \qquad \langle \varphi, w \rangle_{V',V} = (u, w)_V.$$

De plus, on a $\|u\|_V = \|\varphi\|_{V'}$. En d'autres termes, l'application de V' dans V qui à φ associe u est un isomorphisme isométrique permettant d'identifier l'espace de Hilbert V et son dual.

Preuve. Soit $\varphi \in V'$ et $K = \varphi^{-1}(0) = \{\, w \in V,\, \langle \varphi, w \rangle_{V',V} = 0 \,\}$. K est par construction un sous-espace fermé de V. Si $K = V$, alors $\varphi = 0$ et il est clair que $\varphi = 0$ peut être trivialement représentée par $u = 0$. Si $K \neq V$, on peut trouver $v_0 \in V$ avec $v_0 \notin K$. Posons

$$v_1 = P_K v_0 \quad \text{et} \quad v = \frac{v_1 - v_0}{\|v_1 - v_0\|_V}.$$

On a $\|v\|_V = 1$, $\langle \varphi, v \rangle_{V',V} \neq 0$ et $(v, w)_V = 0$ pour tout $w \in K$. Posons

$$u = (\langle \varphi, v \rangle_{V',V})\, v.$$

Montrons que u est un représentant de φ. Pour tout $w \in V$, on peut écrire $w = \lambda v + z$ avec $\lambda = \langle \varphi, w \rangle_{V',V} / \langle \varphi, v \rangle_{V',V}$ et $z \in K$. On obtient donc

$$(u, w)_V = \lambda (u, v)_V + (u, z)_V = \tfrac{\langle \varphi, w \rangle_{V',V}}{\langle \varphi, v \rangle_{V',V}} \langle \varphi, v \rangle_{V',V} (v, v)_V + 0 = \varphi(w).$$

L'unicité du représentant est immédiate : si u_1 et $u_2 \in V$ sont des représentants de φ, on a $(u_2 - u_1, w)_V = 0$ pour tout $w \in V$ si bien que $u_1 = u_2$. Il nous reste à montrer l'égalité des normes. On a

$$\|u\|_V = |\langle \varphi, v \rangle_{V',V}| \leq \|\varphi\|_{V',V},$$

puisque $\|v\|_V = 1$. Par ailleurs, pour tout $w \in V$, on a

$$|\langle \varphi, w \rangle_{V',V}| = |(u, w)_V| \leq \|u\|_V \, \|w\|_V,$$

si bien que $\|\varphi\|_{V',V} \leq \|u\|_V$. $\square$

Exemple. La forme linéaire continue sur l^2 qui à $\mathtt{v} \in l^2$ associe v_0 est représentée par $\mathtt{u} \in l$ donné par $u_0 = 1$ et $u_i = 0$ pour $i \geq 1$.

Remarque 3.12. Lorsque plusieurs espaces de Hilbert sont en jeu, on ne peut pas faire plusieurs identifications en même temps. Considérons par exemple une situation où interviennent deux espaces de Hilbert, V et W, chacun équipé de son propre produit scalaire notés $(\cdot,\cdot)_V$ et $(\cdot,\cdot)_W$ respectivement. Supposons que $W \subset V$ avec injection continue et dense.

- Identifions V et V'. On peut alors injecter V dans W' en considérant l'application $T : V \to W'$ qui à $v \in V$ associe $Tv \in W'$ défini par

$$\forall w \in W, \quad \langle Tv, w\rangle_{W',W} = (v, w)_V.$$

  On vérifie aisément (voir exercice 3.5) que T est linéaire, continue, injective et que $T(V)$ est dense dans W' (mais en général, T n'est pas surjective). On a donc le schéma

$$W \subset V \equiv V' \subset W',$$

  avec injections denses continues. On dit que V joue le rôle d'espace pivot.
- On peut aussi choisir d'identifier W et W' à l'aide du produit scalaire $(\cdot, \cdot)_W$. Dans ce cas, les injections ci-dessus deviennent absurdes ce qui montre qu'on ne peut pas faire *simultanément* les deux identifications. Dans ce deuxième cas, on a le schéma

$$V' \subset W' \equiv W \subset V,$$

  avec injections denses continues.

Un exemple utilisant l'espace l^2 est proposé dans l'exercice 3.6.

3.4 Bases hilbertiennes

Définition 3.13. *Soit V un espace de Hilbert. On appelle* base hilbertienne *de V une suite $(e_n)_{n\geq 0}$ d'éléments de V tels que*
- $\forall n$, $\|e_n\|_V = 1$ *et* $\forall m, n$, $m \neq n$, $(e_m, e_n)_V = 0$ *;*
- *l'espace vectoriel engendré par la famille $(e_n)_{n\geq 0}$ est dense dans V.*

Remarque 3.14. La notion de base hilbertienne généralise la notion de base orthonormée en dimension finie.

Remarque 3.15. En dimension infinie, un espace de Hilbert n'admet pas nécessairement de base hilbertienne. Une condition suffisante pour qu'un espace de Hilbert admette une base hilbertienne est que celui-ci soit séparable, i.e. qu'il existe un sous-ensemble $D \subset V$ dénombrable et dense dans V. Tous les espaces de Hilbert rencontrés dans le cadre de ce cours seront séparables, donc admettront une base hilbertienne.

Exemple. Une base hilbertienne de l^2 est donnée par la famille des suites $(\mathsf{e}_n)_{n\geq 0}$ telles que $\forall m \geq 0$, $e_{n,m} = \delta_{nm}$, le symbole de Kronecker.

Proposition 3.16. *Soit V un espace de Hilbert admettant une base hilbertienne $(e_n)_{n\geq 0}$. Soit $u \in V$. Posons $u_n = (u, e_n)_V$ pour tout $n \geq 0$. Alors, les séries $\sum_{n\geq 0} u_n\, e_n$ et $\sum_{n\geq 0} |u_n|^2$ sont convergentes dans V et $\mathbb{R}$ respectivement et on a*

$$u = \sum_{n\geq 0} u_n\, e_n \quad \text{et} \quad \|u\|_V^2 = \sum_{n\geq 0} |u_n|^2, \tag{3.5}$$

*la deuxième égalité portant le nom d'*égalité de Bessel-Parseval.

Preuve. Pour tout entier $n \geq 0$, on note V_n l'espace vectoriel engendré par $(e_0, \dots, e_n)$. Soit P_n la projection orthogonale de V sur V_n. Il résulte de la caractérisation (3.4) que pour tout $u \in V$ on a

$$P_n u = \sum_{m=0}^{n} u_m \, e_m,$$

si bien que

$$\|P_n u\|_V^2 = \sum_{m=0}^{n} |u_m|^2. \tag{3.6}$$

Du corollaire 3.9, nous déduisons que $\|P_n u\|_V \leq \|u\|_V$ si bien que la série $\sum_{m \geq 0} |u_m|^2$ est convergente. De la majoration

$$\|P_{n+p} u - P_n u\|_V^2 \leq \sum_{m=n+1}^{n+p} |u_m|^2,$$

nous déduisons alors que la suite $(P_n u)_{n \geq 0}$ est de Cauchy dans V et donc convergente vers une limite $v \in V$. Comme nous avons $(u - v, e_m)_V = 0$ pour tout $m \geq 0$ et que la famille $(e_m)_{m \geq 0}$ est dense dans V, il en résulte que $u = v$, i.e. $u = \sum_{n \geq 0} u_n \, e_n$. Finalement, l'égalité de Bessel-Parseval résulte de (3.6) en faisant tendre n vers l'infini. $\square$

3.5 Calcul différentiel

Nous montrons dans cette section comment les notions de différentielle et de gradient (supposées connues du lecteur lorsqu'on travaille dans $\mathbb{R}^d$) s'étendent au cadre des espaces de Hilbert.

Définition 3.17. *Soit U un ouvert d'un espace de Hilbert V et $J : U \to \mathbb{R}$ une fonctionnelle définie sur U. On dit que J est* différentiable *en $v \in U$, s'il existe une forme linéaire continue $L \in V'$ telle qu'au voisinage de v,*

$$J(v+h) = J(v) + \langle L, h \rangle_{V',V} + o(h),$$

ce qui signifie

$$\forall \epsilon > 0, \ \exists \eta > 0 \ \text{t.q.} \ \forall h \in V \ \text{t.q.} \ \|h\|_V \leq \eta, \ \text{ on a } v+h \in U \ \text{ et } |J(v+h) - J(v) - \langle L, h \rangle_{V',V}| \leq \epsilon \|h\|_V.$$

Si une telle forme linéaire L existe, elle est unique ; c'est par définition la différentielle *de J au point v. Elle est notée $J'(v)$.*

Si J est différentiable au point v, on a donc

$$J(v+h) = J(v) + \langle J'(v), h \rangle_{V',V} + o(h).$$

Pour montrer en pratique qu'une fonctionnelle J est différentiable en $v \in V$, on se donne $h \in V$ et on développe $J(v+h)$ sous la forme

$$J(v+h) = J(v) + L_v(h) + R(v,h),$$

où $L_v(h)$ regroupe les termes d'ordre 1 en h et où $R(v,h)$ regroupe les termes d'ordre supérieur ou égal à 2 en h. On vérifie ensuite que (i) $h \mapsto L_v(h)$ est une forme linéaire continue (la continuité signifiant qu'il existe $C \in \mathbb{R}_+$ tel que pour tout $h \in V$, $|L_v(h)| \leq C \|h\|_V$, la constante C pouvant dépendre de v) et (ii) $\lim\limits_{h \to 0} |R(v,h)| / \|h\|_V = 0$.

Définition 3.18. *On dit que J est différentiable sur U si J est différentiable en tout point de U. L'application*

$$\begin{array}{rcl} J' : U & \longrightarrow & V' \\ v & \mapsto & J'(v) \end{array}$$

est alors appelée la différentielle de J. On dit que J est de classe C^1 *sur U si J' est continue.*

Définition 3.19. *Soit V un espace de Hilbert (muni d'un produit scalaire noté $(\cdot,\cdot)_V$), U un ouvert de V et $J : U \to \mathbb{R}$ une fonctionnelle différentiable en $v \in U$. L'unique élément de V noté $\nabla J(v)$ et défini par*

$$\forall h \in V, \qquad \langle J'(v), h\rangle_{V',V} = (\nabla J(v), h)_V \tag{3.7}$$

est appelé le gradient *de J au point v.*

Notons que l'existence et l'unicité du gradient découlent du théorème de Riesz.

Remarque 3.20. Si on remplace le produit scalaire $(\cdot,\cdot)_V$ de V, par un produit scalaire $\langle\cdot,\cdot\rangle_V$ définissant une norme équivalente à la norme définie par $(\cdot,\cdot)_V$, une fonctionelle J différentiable en un point $v \in U$ reste différentiable en v, et la différentielle $J'(v)$ demeure inchangée. En revanche, le gradient de J au point v, qui est défini par (3.7), dépend bien évidemment du choix du produit scalaire.

Considérons le cas où $V = \mathbb{R}^d$. Soit $U \subset \mathbb{R}^d$, $x \in U$ et $J : U \to \mathbb{R}$ une fonctionnelle différentiable en x. La différentielle $J'(x)$ de J au point x est une forme linéaire sur $\mathbb{R}^d$. Si on note dx_i la i-ième forme coordonnée, définie par

$$\forall h = \begin{pmatrix} h_1 \\ \cdot \\ \cdot \\ \cdot \\ h_d \end{pmatrix} \in \mathbb{R}^d, \quad dx_i(h) = h_i,$$

on peut développer $J'(x)$ dans la base $(dx_1, \cdots, dx_d)$ de V'. On vérifie facilement que les coordonnées de $J'(x)$ dans cette base ne sont autres que les dérivées partielles de J au point x :

$$J'(x) = \sum_{i=1}^{d} \frac{\partial J}{\partial x_i}(x)\, dx_i,$$

soit encore

$$\forall h \in \mathbb{R}^d, \quad \langle J'(x), h\rangle = \sum_{i=1}^{d} \frac{\partial J}{\partial x_i}(x)\, h_i.$$

Si on munit $\mathbb{R}^d$ du produit scalaire euclidien (ce qu'on fait presque toujours implicitement), on retrouve le fait bien connu que le gradient de J au point x est le vecteur formé par les dérivées partielles de J au point x :

$$\nabla J(x) = \begin{pmatrix} \dfrac{\partial J}{\partial x_1}(x) \\ \cdot \\ \cdot \\ \cdot \\ \dfrac{\partial J}{\partial x_d}(x) \end{pmatrix}$$

3.6 Optimisation

Les problèmes d'optimisation sont légion en sciences de l'ingénieur. Nous nous intéressons ici à des problèmes d'optimisation continue[1]

– *sans contrainte*, du type
$$\inf_{v \in V} J(v), \tag{3.8}$$
– *avec contraintes égalités*, du type
$$\inf_{v \in V \mid F(v)=0} J(v), \tag{3.9}$$
– *avec contraintes inégalités*, du type
$$\inf_{v \in V \mid F(v)\leq 0} J(v). \tag{3.10}$$

Dans les problèmes ci-dessus, V est un espace de Hilbert (de dimension finie ou infinie), $J : V \to \mathbb{R}$ est une fonctionnelle définie sur V à valeurs dans $\mathbb{R}$, et $F : V \to \mathbb{R}^m$ est une fonctionnelle définie sur V à valeurs dans $\mathbb{R}^m$. Dans le langage de l'optimisation, la fonctionnelle J à minimiser est appelée *critère* ou *fonction coût*. On note $F_1, \cdots, F_m$ les coordonnées de F :
$$\forall v \in V, \quad F(v) = \begin{pmatrix} F_1(v) \\ \cdot \\ \cdot \\ \cdot \\ F_m(v) \end{pmatrix} \in \mathbb{R}^m.$$

Les fonctions F_i sont appelées contraintes (scalaires). La notation $F(v) \leq 0$ signifie $F_i(v) \leq 0$ pour tout $1 \leq i \leq m$.

En introduisant les ensembles
$$K = V \quad \text{pour (3.8)},$$
$$K = \{v \in V \mid F(v) = 0\} \quad \text{pour (3.9)},$$
et
$$K = \{v \in V \mid F(v) \leq 0\} \quad \text{pour (3.10)},$$
les problèmes (3.8), (3.9) et (3.10) s'écrivent sous la forme
$$\inf_{v \in K} J(v). \tag{3.11}$$

Le nombre réel $\inf_{v\in K} J(v)$ est appelé l'*infimum* de J sur K. On dit que u est un *minimiseur global* du problème (3.11) si
$$u \in K \quad \text{et} \quad J(u) = \inf_{v \in K} J(v).$$

On dit que u est un *minimiseur local* du problème (3.11) s'il existe $\epsilon > 0$ tel que
$$u \in K \quad \text{et} \quad J(u) = \inf_{v \in K \mid \|v-u\|_V \leq \epsilon} J(v).$$

Les questions que nous allons aborder dans cette introduction à l'optimisation concernent

– l'existence et l'unicité de minimiseurs globaux pour les problèmes (3.8), (3.9) et (3.10) ;

[1] On parle de problèmes d'optimisation continue par opposition aux problèmes d'optimisation discrète dans lesquels l'ensemble sur lequel on optimise est un ensemble discret.

– la dérivation de conditions d'optimalité du premier ordre pour les minimiseurs locaux, débouchant, dans le cadre des problèmes (3.9) et (3.10), sur la notion de *multiplicateur de Lagrange*.

Dans les applications de l'optimisation, on cherche soit à minimiser le critère (lorsque celui-ci est un coût, un temps de trajet, ...) soit à le maximiser (lorsqu'il s'agit d'une rentabilité, d'une utilité, ...). Bien évidemment, nous ne restreignons pas le champ d'application de la théorie en ne considérant que des problèmes de minimisation, puisque maximiser la fonction J revient à minimiser la fonction $-J$.

Certains des résultats que nous allons établir ne sont valables que pour des fonctionnelles *convexes*. Rappelons quelques définitions de base.

Définition 3.21. *Soit V un $\mathbb{R}$-ev et $K \subset V$. On dit que l'ensemble K est* convexe *si*

$$\forall (v,w) \in K \times K, \quad \forall \theta \in [0,1], \quad \theta v + (1-\theta) w \in K.$$

Définition 3.22. *Soit V un $\mathbb{R}$-ev, $K \subset V$ un sous-ensemble convexe de V et $J : K \to \mathbb{R}$.*

– *On dit que J est* convexe *si*

$$\forall (v,w) \in K \times K, \quad \forall \theta \in [0,1], \quad J(\theta v + (1-\theta) w) \leq \theta J(v) + (1-\theta) J(w).$$

– *On dit que J est* strictement convexe *si l'inégalité ci-dessus est stricte lorsque $\theta \in]0,1[$ et $v \neq w$.*

Nous aurons besoin dans certains cas d'une notion plus forte que la stricte convexité.

Définition 3.23. *Soit V un $\mathbb{R}$-ev normé, $K \subset V$ un sous-ensemble convexe de V et $J : K \to \mathbb{R}$. On dit que J est* fortement convexe *de paramètre $\alpha > 0$ (on dit également que J est α-convexe) si*

$$\forall (v,w) \in K \times K, \quad \forall \theta \in [0,1], \quad J(\theta v + (1-\theta) w) \leq \theta J(v) + (1-\theta) J(w) - \alpha \frac{\theta(1-\theta)}{2} \|v-w\|_V^2.$$

Si V est un espace de Hilbert et si $J : V \to \mathbb{R}$ est différentiable, on dispose des caractérisations suivantes (les preuves, élémentaires, sont laissées en exercice).

Proposition 3.24. *Soit V un espace de Hilbert et $J : V \to \mathbb{R}$ une fonctionnelle différentiable sur V. Les assertions suivantes sont équivalentes :*

(i) *J est* convexe *sur V ;*

(ii) *pour tout $(v,w) \in V \times V$, $J(w) \geq J(v) + (\nabla J(v), w-v)_V$;*

(iii) *pour tout $(v,w) \in V \times V$, $(\nabla J(w) - \nabla J(v), w-v)_V \geq 0$.*

La condition (ii) signifie que le graphe de J est au-dessus de ses hyperplans tangents en tout point v. De même, la condition (iii) a une interprétation simple en dimension 1 : elle est équivalente au fait que la dérivé de J est croissante.

Proposition 3.25. *Soit $J : V \to \mathbb{R}$ une fonctionnelle différentiable sur V. Les assertions suivantes sont équivalentes :*

(i) *J est* fortement convexe *sur V (de paramètre α) ;*

(ii) *pour tout $(v,w) \in V \times V$, $J(w) \geq J(v) + (\nabla J(v), w-v)_V + \frac{\alpha}{2}\|w-v\|_V^2$;*

(iii) *pour tout $(v,w) \in V \times V$, $(\nabla J(w) - \nabla J(v), w-v)_V \geq \alpha \|w-v\|_V^2$.*

3.6.1 Optimisation sans contrainte

Commençons par un résultat élémentaire.

Théorème 3.26. *On suppose que l'espace vectoriel V est de dimension* finie. *On suppose en outre que J est continue sur V et infinie à l'infinie au sens où*

$$\lim_{\|v\|_V \to +\infty} J(v) = +\infty. \tag{3.12}$$

Alors, J admet au moins un minimiseur global sur V.

Preuve. Il résulte de la propriété (3.12) que

$$\forall M \in \mathbb{R}, \quad \exists R \in \mathbb{R}_+ \quad \text{tel que} \quad (\|v\|_V \geq R) \Rightarrow (J(v) \geq M).$$

En considérant cette propriété pour $M = J(0)$, il vient

$$\exists R \in \mathbb{R}_+ \quad \text{tel que} \quad (\|v\|_V \geq R) \Rightarrow (J(v) \geq J(0)).$$

Par suite,

$$\inf_{v \in V} J(v) = \inf_{v \in \overline{B_R(0)}} J(v),$$

où $\overline{B_R(0)} = \{v \in V;\ \|v\|_V \leq R\}$. Or, en dimension finie, la boule $B_R(0)$ est compacte. La fonctionnelle J étant continue, elle y atteint son infimum. □

La situation est plus complexe dans un espace de Hilbert de dimension infinie. Ainsi par exemple, le problème

$$\inf_{v \in V} J(v) \quad \text{où} \quad V = l^2 \quad \text{et} \quad J(v) = \left(\|v\|_{l^2}^2 - 1\right)^2 + \sum_{i \in \mathbb{N}} \frac{v_i^2}{i+1},$$

n'a pas de minimiseur global. En effet, pour tout $v \in V$, $J(v) > 0$, tandis que $\inf\limits_{v \in V} J(v) = 0$ (on pourra le vérifier en exercice).

En revanche, l'existence (et l'unicité) d'un minimum global pour (3.8) est assurée dans le cas suivant.

Théorème 3.27. *Soient V un espace de Hilbert et J une fonctionnelle fortement convexe et continue sur V. Alors, J admet un et un seul minimiseur global sur V.*

Preuve. Soit $(v_n)_{n \in \mathbb{N}}$ une suite minimisante de J sur V, c'est-à-dire une suite $(v_n)_{n \in \mathbb{N}}$ d'éléments de V telle que

$$J(v_n) \underset{n \to \infty}{\longrightarrow} I = \inf_{v \in V} J(v).$$

Notons que $I > -\infty$ puisque J est fortement convexe. Soit α le paramètre de forte convexité de J sur V. Il vient pour tout $m, n \geq 0$,

$$\frac{\alpha}{8}\|v_n - v_m\|_V^2 + \underbrace{J\left(\frac{v_n + v_m}{2}\right) - \inf_{v \in V} J(v)}_{\geq 0} \leq \frac{1}{2}\left(J(v_n) - \inf_{v \in V} J(v)\right) + \frac{1}{2}\left(J(v_m) - \inf_{v \in V} J(v)\right),$$

ce qui montre que la suite $(v_n)_{n \in \mathbb{N}}$ est Cauchy dans V, donc converge vers une limite $u \in V$. La fonctionnelle J étant continue, $J(u) = \inf_{v \in V} J(v)$. D'où l'existence du minimiseur global de J sur V. Comme par ailleurs J est fortement, donc strictement convexe sur V, le minimiseur global de J sur V est unique. □

En fait, on peut montrer l'existence d'un minimiseur sous des hypothèses moins contraignantes.

Théorème 3.28. *Soit V un espace de Hilbert. On suppose que $J : V \to \mathbb{R}$ est convexe et continue sur V et infinie à l'infini au sens où*

$$\lim_{\|v\|_V \to +\infty} J(v) = +\infty. \tag{3.13}$$

Alors, J admet au moins un minimiseur global sur V. De plus, tout minimiseur local de J est un minimiseur global et l'ensemble des minimiseurs globaux de J forme un ensemble convexe.

Preuve. La preuve de l'existence d'un minimiseur global fait appel à la notion de convergence faible ; elle fait l'objet de l'exercice 3.7.

Montrons que tout minimiseur local est un minimiseur global. Soit $I = \inf_{v\in V} J(v)$, u un minimiseur local de J sur V et $\epsilon > 0$ tel que

$$\forall h \in V \text{ tel que } \|h\|_V \leq \epsilon, \quad J(u+h) \geq J(u). \tag{3.14}$$

Supposons que u ne soit pas un minimiseur global de J. Il existerait alors $w \in V$ tel que $J(w) < J(u)$. La fonctionnelle J étant convexe, on aurait

$$\forall \theta \in]0,1[, \quad J(u+\theta(w-u)) = J(\theta w + (1-\theta)u) \leq \theta J(w) + (1-\theta)J(u) < J(u),$$

ce qui contredit (3.14) pour θ petit. Le point u est donc un minimiseur global de J.

Enfin, soit C l'ensemble des minimiseurs locaux (donc globaux d'après ce qui précède) de J. Soit u_1 et u_2 deux vecteurs de C. On a $J(u_1) = J(u_2) = I$ et donc, pour tout $\theta \in [0,1]$,

$$J(\theta u_1 + (1-\theta)u_2) \leq \theta J(u_1) + (1-\theta)J(u_2) = I.$$

Le vecteur $\theta u_1 + (1-\theta)u_2$ est donc dans C pour tout $\theta \in [0,1]$, ce qui prouve que C est convexe. $\square$

Il est bien connu qu'une *condition nécessaire* pour qu'un point $x_0 \in \mathbb{R}$ soit un minimiseur local d'une fonction dérivable $f : \mathbb{R} \to \mathbb{R}$ est que la dérivée f' s'annule en x_0. Ce critère peut se généraliser aux fonctionnelles *différentiables* de V dans $\mathbb{R}$.

Théorème 3.29. *Soient V un espace de Hilbert et J une fonctionnelle de V dans $\mathbb{R}$. On suppose que la fonctionnelle J admet un minimum local en $u \in V$ et que J est différentiable en u. Alors,*

$$J'(u) = 0 \quad (\textit{dans } V') \quad \textit{ou de façon équivalente} \quad \nabla J(u) = 0 \quad (\textit{dans } V). \tag{3.15}$$

Un vecteur $u \in V$ vérifiant (3.15) est appelé un point critique *de J.*

Preuve. Soit $v \in V$. Pour tout $t \in \mathbb{R}$ sufisamment petit,

$$J(u) \leq J(u+tv) = J(u) + t(\nabla J(u), v)_V + o(t).$$

En faisant tendre t vers 0 d'abord par valeurs supérieures, puis par valeurs inférieures, on obtient $(\nabla J(u), v)_V = 0$ et comme v est arbitraire dans V, il vient $\nabla J(u) = 0$. $\square$

3.6.2 Optimisation avec contraintes égalités

Hormis en dimension finie, il n'existe pas de résultats très généraux similaires au théorème 3.28 assurant l'existence (et *a fortiori* l'unicité) d'un minimiseur pour un problème d'optimisation sous contraintes égalités. Ceci vient notamment du fait que, sauf cas exceptionnel, l'ensemble K n'est pas convexe même si les contraintes F_i sont convexes. En dimension infinie, l'existence d'un minimiseur pour le problème (3.9) donc être établie au cas par cas.

On connaît en revanche une condition nécessaire vérifié par les minimiseurs locaux du problème

$$\inf_{v\in K} J(v) \quad \text{avec} \quad K = \{v \in V \mid F(v) = 0\}. \tag{3.16}$$

Définition 3.30. *On dit que les contraintes égalités $F(v) = 0$ sont* qualifiées *en un point $u \in K$ en lequel F est différentiable si la famille $\{\nabla F_i(u)\}_{1\le i\le m}$ est libre (dans V).*

Théorème 3.31. *Soit $u \in K$ un minimum local de J sur*

$$K = \{v \in V \mid F(v) = 0\}.$$

On suppose que J est différentiable en u, que F est de classe C^1 au voisinage de u, et que les contraintes égalités $F(v) = 0$ sont qualifiées en u. Si u est un minimiseur local de J sur K, il existe un unique m-uplet de réels $(p_1, \dots, p_m)$, appelés multiplicateurs de Lagrange, *tel que*

$$J'(u) + \sum_{j=1}^{m} p_j F_j'(u) = 0 \quad \textit{ou de manière équivalente} \quad \nabla J(u) + \sum_{i=1}^{m} p_i \nabla F_i(u) = 0. \tag{3.17}$$

Notons que dans le cas où V est un espace de dimension finie d, le système de d équations scalaires (3.17) comporte $d+m$ inconnues (le vecteur u et les m multiplicateurs de Lagrange p_i). Pour obtenir un système comportant autant d'équations que d'inconnues, il faut compléter (3.17) par la contrainte $F(u) = 0$ (qui comporte m équations scalaires). Il est donc préférable d'écrire les équations d'optimalité du premier ordre associées au problème (3.16) sous la forme

$$\left\{\begin{array}{l} J'(u) + p\cdot F'(u) = 0 \\ F(u) = 0 \end{array}\right. \quad \text{ou de manière équivalente} \quad \left\{\begin{array}{l} \nabla J(u) + p\cdot \nabla F(u) = 0 \\ F(u) = 0, \end{array}\right. \tag{3.18}$$

où $\cdot$ désigne le produit scalaire euclidien de $\mathbb{R}^m$.

Il est d'usage en mécanique, en physique, ou en économie d'introduire le *Lagrangien* $\mathcal{L}(v,q)$ associé au problème (3.16), défini sur l'espace $V\times\mathbb{R}^m$ par

$$\forall (v,q) \in V\times\mathbb{R}^m, \quad \mathcal{L}(v,q) = J(v) + q\cdot F(v),$$

et de remarquer que les équations (3.18) s'écrivent aussi

$$\left\{\begin{array}{l} \mathcal{L}_v'(u,p) = 0 \\ \mathcal{L}_q'(u,p) = 0 \end{array}\right. \quad \text{ou de manière équivalente} \quad \left\{\begin{array}{l} \nabla_v\mathcal{L}(u,p) = 0 \\ \nabla_q\mathcal{L}(u,p) = 0, \end{array}\right.$$

où $\mathcal{L}_v'(u,p) = J'(u) + p\cdot F'(u)$ et $\mathcal{L}_q'(u,p) = (\cdot, F(u))_{\mathbb{R}^m}$ représentent respectivement les différentielles des fonctions $V \ni v \mapsto \mathcal{L}(v,q) \in \mathbb{R}$ et $\mathbb{R}^m \ni q \mapsto \mathcal{L}(v,q) \in \mathbb{R}$ au point (u,p), et $\nabla_v\mathcal{L}(u,p)$ et $\nabla_q\mathcal{L}(u,p)$ les gradients associés. Nous entreverrons dans la section suivante, consacrée aux problèmes d'optimisation sous contraintes inégaliés, certaines des propriétés qui font l'intérêt de ce formalisme.

Preuve du théorème 3.31. On se limite pour simplifier au cas où $m = 1$ (une seule contrainte scalaire). Dans ce cas, F est une fonctionnelle de V dans $\mathbb{R}$ et la condition de qualification s'écrit $F'(u) \neq 0$ (soit encore $\nabla F(u) = 0$).

– Commençons par prouver le résultat pour un espace V de dimension finie égale à 2, en nous plaçant dans un repère orthonormé de V. Le problème (3.16) s'écrit alors

$$\inf_{(v_1,v_2)\in\mathbb{R}^2 \mid F(v_1,v_2)=0} J(v_1, v_2),$$

et l'hypothèse de qualification des contraintes $(\partial F/\partial v_1(u), \partial F/\partial v_2(u)) \neq (0, 0)$. Supposons par exemple que $\partial F/\partial v_1(u) \neq 0$ (le cas où $\partial F/\partial v_2(u) \neq 0$ se traite de manière similaire). D'après le théorème des fonctions implicites, il existe un voisinage de $\mathcal{U} =]u_1 - \epsilon, u_1 + \epsilon[\times]u_2 - \eta, u_2 + \eta[$ ($\epsilon > 0$, $\eta > 0$) de $u = (u_1, u_2)$ dans $\mathbb{R}^2$ et une fonction $\phi :]u_1 - \epsilon, u_1 + \epsilon[\to]u_2 - \eta, u_2 + \eta[$ tel que

$$\{v \in \mathcal{U} \mid F(v) = 0\} = \{v = (v_1, \phi(v_1)),\ v_1 \in]u_1 - \epsilon, u_1 + \epsilon[\}\,.$$

De plus ϕ est dérivable en u_1 et

$$\phi'(u_1) = -\frac{\dfrac{\partial F}{\partial v_2}(u)}{\dfrac{\partial F}{\partial v_1}(u)}. \tag{3.19}$$

Or u étant un minimum local du problème (3.16), on a (quitte à diminuer la valeur du réel strictement positif ϵ)

$$J(u_1, u_2) = \inf_{v_1\in]u_1-\epsilon,u_1+\epsilon[,\ v_2=\phi(v_1)} J(v_1, v_2) = \inf_{x_1\in]u_1-\epsilon,u_1+\epsilon[} J(v_1, \phi(v_1)).$$

On en déduit

$$\frac{\partial J}{\partial v_1}(u_1, u_2) + \frac{\partial J}{\partial v_2}(u_1, u_2)\phi'(u_1) = 0.$$

En utilisant (3.19) et en multipliant par $\partial F/\partial v_1(u_1, u_2)$ l'équation ci-dessus, on obtient

$$\frac{\partial J}{\partial v_1}(u_1, u_2)\frac{\partial F}{\partial v_2}(u_1, u_2) - \frac{\partial J}{\partial v_2}(u_1, u_2)\frac{\partial F}{\partial v_1}(u_1, u_2) = 0.$$

Ceci prouve que les vecteurs $\nabla J(u)$ et $\nabla F(u)$ sont colinéaires, autrement dit, comme $\nabla F(u) \neq 0$, qu'il existe $p \in \mathbb{R}$ tel que

$$\nabla J(u) + p\nabla F(u) = 0.$$

– Prouvons maintenant le cas général (V de dimension quelconque, finie ou infinie). Par hypothèse, $F'(u) \neq 0$. Donc $\mathrm{Ker}(F'(u))$ est un hyperplan fermé de V. Soit $v_0 \in V$ tel que $\langle F'(u), v_0\rangle_{V',V} = 1$ et $p := -\langle J'(u), v_0\rangle_{V',V}$. Soit maintenant $h \in V$. Posons

$$\widetilde{J}(x_1, x_2) = J(u + x_1v_0 + x_2h) \quad \text{et} \quad \widetilde{F}(x_1, x_2) = F(u + x_1v_0 + x_2h).$$

Comme u est un minimiseur local de J sur K, il existe $\epsilon > 0$ tel que

$$\widetilde{J}(0, 0) = J(u) = \inf_{v\in V \mid F(v)=0,\ \|v-u\|_V\leq\epsilon} J(v) \leq \inf_{(x_1,x_2)\in\mathbb{R}^2 \mid \|x_1v_0+x_2h\|_V\leq\epsilon,\ \widetilde{F}(x_1,x_2)=0} \widetilde{J}(x_1, x_2) \leq \widetilde{J}(0, 0).$$

Ceci montre que $(0, 0)$ est un minimiseur local du problème

$$\inf_{(x_1,x_2)\in\mathbb{R}^2 \mid \widetilde{F}(x_1,x_2)=0} \widetilde{J}(x_1, x_2). \tag{3.20}$$

Or un calcul simple montre que $\widetilde{J}$ est différentiable en $(0,0)$, que $\widetilde{F}$ est de classe C^1 au voisinage de ce point, et que

$$\langle \widetilde{F}'(0,0),(1,0)\rangle_{\mathbb{R}^2} = \langle F'(u), v_0\rangle_{V',V} = 1, \quad \langle \widetilde{F}'(0,0),(0,1)\rangle_{\mathbb{R}^2} = \langle F'(u), h\rangle_{V',V},$$

et

$$\langle \widetilde{J}'(0,0),(1,0)\rangle_{R^2} = \langle J'(u), v_0\rangle_{V',V} = -p, \quad \langle \widetilde{J}'(0,0),(0,1)\rangle_{\mathbb{R}^2} = \langle J'(u), h\rangle_{V',V}.$$

En appliquant au problème d'optimisation (3.20), posé en dimension 2, le résultat établi dans la première partie de la preuve, on obtient qu'il existe $q \in \mathbb{R}$ tel que

$$\widetilde{J}'(0,0) + q\,\widetilde{F}'(0,0) = 0.$$

Il en résulte que

$$0 = \langle \widetilde{J}'(0,0) + q\,\widetilde{F}'(0,0),(1,0)\rangle_{\mathbb{R}^2} = -p + q,$$

ce qui montre que $q = p$, et par suite que

$$\begin{aligned} 0 &= \langle \widetilde{J}'(0,0) + q\,\widetilde{F}'(0,0),(0,1)\rangle_{\mathbb{R}^2} \\ &= \langle J'(u), h\rangle_{V',V} + q\langle F'(u), h\rangle_{V',V} \\ &= \langle J'(u), h\rangle_{V',V} + p\langle F'(u), h\rangle_{V',V} \\ &= \langle J'(u) + pF'(u), h\rangle_{V',V}. \end{aligned}$$

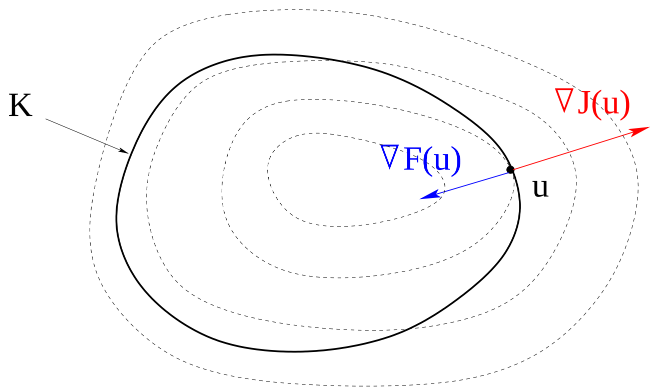
Cette relation étant vraie pour tout $h \in V$, et $p = \langle J'(u), v_0\rangle_{V',V}$ étant indépendant de h, il en résulte que

$$J'(u) + pF'(u) = 0.$$

On obtient ainsi le résultat escompté.

□

Remarque 3.32. *Interprétation géométrique du théorème 3.31 (pour $m = 1$). Si $\nabla J(u) = 0$, l'équation d'Euler-Lagrange est vérifiée avec $p = 0$. Il suffit donc de considérer le cas où $\nabla J(u) \neq 0$. Pour que u soit un point de minimum local de J sur K, il faut que la surface de niveau $S := J^{-1}(J(u))$ soit tangente à K au point u. Sinon, on pourrait faire décroître J en restant sur K. Or K est la surface de niveau 0 de la fonction F. Par ailleurs, par définition du gradient, le gradient d'une fonction en un point est toujours orthogonal à la ligne de niveau de cette fonction contenant ce point. Comme les surfaces de niveau S et K sont tangentes au point u, les gradients $\nabla J(u)$ et $\nabla F(u)$, qui sont respectivement orthogonaux aux surfaces de niveau S et K, sont colinéaires. Le vecteur $\nabla F(u)$ étant non nul par hypothèse, il existe donc un certain $p \in \mathbb{R}$ tel que $\nabla J(u) + p\nabla F(u) = 0$.*



Remarque 3.33. *Interprétation de la valeur du multiplicateur de Lagrange (pour $m = 1$).* On suppose toujours que $m = 1$ et on s'intéresse maintenant à la famille de problèmes d'optimisation

$$I(\alpha) = \inf_{v \in V \mid F(v)=\alpha} J(v).$$

On suppose qu'il existe $\epsilon > 0$ tel que pour tout $\alpha \in]-\epsilon, \epsilon[$, le problème d'optimisation ci-dessus admette un unique minimiseur global $u(\alpha)$ et que la fonction $\alpha \mapsto u(\alpha)$ soit dérivable sur $]-\epsilon, \epsilon[$. On suppose également que la contraite $F(v) = 0$ est qualifiée en $u := u(0)$. Il résulte du Théorème 3.31 qu'il existe $p \in \mathbb{R}$ tel que

$$J'(u) + pF'(u) = 0.$$

Par ailleurs, on a pour tout $\alpha \in]-\epsilon, \epsilon[$,

$$\begin{aligned} I(\alpha) &= J(u(\alpha)) = J(u(0) + \alpha u'(0) + o(\alpha)) \\ &= J(u) + \alpha(\nabla J(u), u'(0))_V + o(\alpha) \\ &= I(0) - \alpha p(\nabla F(u), u'(0))_V + o(\alpha), \end{aligned}$$

et

$$\begin{aligned} \alpha &= F(u(\alpha)) = F(u(0) + \alpha u'(0) + o(\alpha)) \\ &= F(u) + \alpha(\nabla F(u), u'(0))_V + o(\alpha) \\ &= \alpha(\nabla F(u), u'(0))_V + o(\alpha). \end{aligned}$$

On déduit de la première égalité que I est dérivable en 0 et que $I'(0) = -p(\nabla F(u), u'(0))_V$. La deuxième inégalité montre que $(\nabla F(u), u'(0))_V = 1$. Il en résulte que

$$I'(0) = -p.$$

En économie, le multiplicateur de Lagrange p est souvent appelé l'élasticité de la fonction J par rapport à la contrainte $F(v) = 0$. En thermodynamique, les variables intensives (température, pression, potentiels chimiques, ...) sont des multiplicateurs de Lagrange (ou pour certaines des fonctions des multiplicateurs de Lagrange) qui interviennent lors de la maximisation de l'entropie.

3.6.3 Optimisation avec contraintes inégalités

Considérons pour finir le problème d'optimisation sous contraintes inégalités

$$\inf_{v \in K} J(v) \quad \text{avec} \quad K = \{v \in V \mid F(v) \leq 0\}. \tag{3.21}$$

Dans le cas de contraintes inégalités, la notion de qualification des contraintes est plus délicate. Afin d'éviter d'entrer dans des détails trop techniques, nous allons nous restreindre à des situations relativement simples.

Définition 3.34. *On dit que la contrainte F_i est* active *(ou* saturée*) en un point $u \in K$ si $F_i(u) = 0$. On note*

$$\mathcal{A}(u) = \{i \in \{1, \ldots, m\} \mid F_i(u) = 0\}.$$

Définition 3.35. *On suppose que les fonctions F_i sont différentiables. On dit que les contraintes inégalités $F_i \leq 0$ sont* qualifiées *en un point $u \in K$ si l'une des deux conditions suivantes est vérifiée :*

1. *toutes les contraintes actives en u sont affines ;*

2. la famille des gradients des contraintes actives en u, $\{\nabla F_i(u)\}_{i\in\mathcal{A}(u)}$, est libre.

Théorème 3.36. *Soit $u \in K$. On suppose que les fonctions J et F_i $(1 \le i \le m)$ sont différentiables en u et que les contraintes inégalité sont qualifiées en u. Alors, une* condition nécessaire *pour que u soit un minimiseur local de J sur K est qu'il existe un vecteur $p \in \mathbb{R}^m_+$ de composantes* positives *$(p_1, \dots, p_m)$ appelées* multiplicateurs de Lagrange*, tel que*

$$\nabla J(u) + \sum_{j=1}^{m} p_j \nabla F_j(u) = 0 \tag{3.22}$$

$$p \cdot F(u) = 0. \tag{3.23}$$

On rappelle que le symbole $\cdot$ fait ici référence au produit scalaire euclidien de $\mathbb{R}^m$.

La condition (3.23) est appelée *condition des écarts complémentaires* (on parle aussi de *relations d'exclusion*). Puisque toutes les composantes p_i sont positives ($p \in \mathbb{R}^m_+$) et que tous les réels $F_i(u)$ sont négatifs ($u \in K$), on observe que tous les termes dans le produit scalaire $p \cdot F(u)$ ont le même signe. Chacun de ces termes doit donc être nul d'après (3.23). Cette condition s'écrit donc de façon plus explicite sous la forme

$$\forall 1 \le i \le m, \qquad p_i F_i(u) = 0.$$

Elle implique que
- si $p_i > 0$, la i-ème contrainte est active en u ;
- si la i-ème contrainte n'est pas active en u, $p_i = 0$.

Preuve du théorème 3.36. On se limite pour simplifier au cas où $m = 1$ (une seule contrainte scalaire). Deux cas peuvent se présenter : soit la contrainte est active ($F(u) = 0$), soit elle ne l'est pas ($F(u) < 0$).

Si la contrainte n'est pas active ($F(u) < 0$), il résulte de la continuité de F qu'il existe $\epsilon > 0$ tel que la boule fermée de centre u et de rayon ϵ soit incluse dans K. Le point u est donc un minimum global (sans contrainte !) de J sur cette boule. En conséquence, $J'(u) = 0$ et $p = 0$, ce qui fait que les équations (3.22) et (3.22) sont bien satisfaites.

Si la contrainte est active ($F(u) = 0$), alors u est aussi un minimiseur local du problème sous contrainte égalité

$$\inf_{v\in V \,|\, F(v)=0} J(v).$$

Si $F'(u) \neq 0$, on déduit du théorème 3.31 qu'il existe $p \in \mathbb{R}$ tel que $J'(u) + pF'(u) = 0$. Soit $\phi(t) = u - t\nabla F(u)$. La fonction F étant diffŕentiable en u, il vient

$$F(\phi(t)) = F(u - t\nabla F(u)) = F(u) - t|\nabla F(u)|^2 + o(t).$$

Comme $\nabla F(u) \neq 0$, il existe $\eta > 0$ tel que pour tout $t \in [0, \eta]$, $\phi(t) \in K$. On a donc pour $t \ge 0$ assez petit,

$$J(u) \le J(\phi(t)) = J(u - t\nabla F(u)) = J(u) - t\nabla J(u) \cdot \nabla F(u) + o(t) = J(u) + tp|\nabla F(u)|^2 + o(t).$$

Il en résulte que le réel p est positif ou nul. Enfin, si F est affine, ou bien $F'(u) \neq 0$, ce qui nous ramène au cas précédent, ou bien F est constante, auquel cas ou bien $K = \emptyset$, ce qui est absurde puisque $u \in K$, ou bien $K = V$, ce qui nous ramène au cadre de l'optimisation sans contrainte. □

Une façon élégante et utile en pratique de reformuler le problème d'optimisation sous contraintes (3.21) consiste à introduire le *Lagrangien* $\mathcal{L}$ défini comme suit :

$$\mathcal{L} : V \times \mathbb{R}^m_+ \ni (v, q) \longmapsto J(v) + q \cdot F(v) \in \mathbb{R}. \tag{3.24}$$

Définition 3.37. *Soit U une partie de V contenant K. On dit que le couple (u,p) est un* point selle *du Lagrangien $\mathcal{L}$ sur $U \times \mathbb{R}_+^m$ si on a*

$$\forall (v,q) \in U \times \mathbb{R}_+^m, \qquad \mathcal{L}(u,q) \leq \mathcal{L}(u,p) \leq \mathcal{L}(v,p). \tag{3.25}$$

En d'autres termes,

$$\sup_{q \in \mathbb{R}_+^m} \mathcal{L}(u,q) = \mathcal{L}(u,p) = \inf_{v \in U} \mathcal{L}(v,p). \tag{3.26}$$

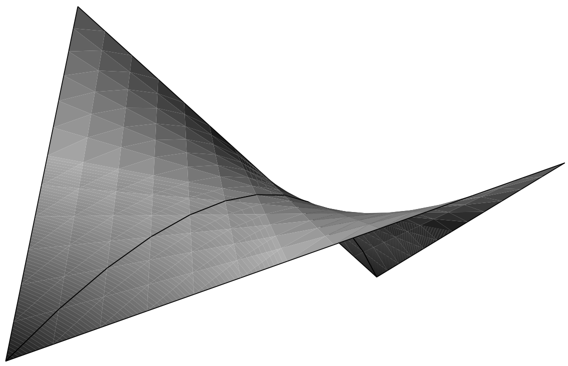


FIG. 3.2 – Graphe d'un Lagrangien présentant un point selle.

La figure 3.2 illustre la notion de point selle. Un point selle satisfait une condition d'optimalité plus générale que (3.26). Celle-ci est précisée dans la proposition suivante.

Proposition 3.38. *Soit (u,p) un point selle du Lagrangien $\mathcal{L}$ sur $U \times \mathbb{R}_+^m$. On a*

$$\sup_{q \in \mathbb{R}_+^m} \inf_{v \in U} \mathcal{L}(v,q) = \mathcal{L}(u,p) = \inf_{v \in U} \sup_{q \in \mathbb{R}_+^m} \mathcal{L}(v,q). \tag{3.27}$$

Preuve. Montrons la première égalité. On introduit la fonctionnelle

$$G : \mathbb{R}_+^m \ni q \longmapsto G(q) := \inf_{v \in U} \mathcal{L}(v,q) \in \mathbb{R} \cup \{-\infty\}. \tag{3.28}$$

Puisque (u,p) est point selle de $\mathcal{L}$, on a pour tout $q \in \mathbb{R}_+^m$,

$$G(q) = \inf_{v \in U} \mathcal{L}(v,q) \leq \mathcal{L}(u,q) \leq \mathcal{L}(u,p).$$

De plus, $G(p) = \mathcal{L}(u,p)$ toujours par la propriété du point selle. D'où

$$\sup_{q \in \mathbb{R}_+^m} G(q) = G(p) = \mathcal{L}(u,p).$$

La preuve de la deuxième inégalité est analogue. On introduit la fonctionnelle

$$H : U \ni v \longmapsto H(v) := \sup_{q \in \mathbb{R}_+^m} \mathcal{L}(v,q) \in \mathbb{R} \cup \{+\infty\}, \tag{3.29}$$

et on vérifie que

$$H(v) \geq \mathcal{L}(v,p) \geq \mathcal{L}(u,p) = H(u),$$

si bien que $\inf_{v \in U} H(v) = H(u) = \mathcal{L}(u,p)$. □

Voyons maintenant quels sont les liens entre la recherche d'un minimiseur du problème d'optimisation sous contraintes et la recherche d'un point selle du Lagrangien.

Proposition 3.39. *Si le couple $(u,p) \in U \times \mathbb{R}_+^m$ est un point selle du Lagrangien $\mathcal{L}$ sur $U \times \mathbb{R}_+^m$ où U est une partie de V contenant K, alors $u \in K$, u est un minimiseur global de J sur K et le couple (u,p) vérifie la condition des écarts complémentaires $p \cdot F(u) = 0$.*

Preuve. Considérons tout d'abord le fait que $\mathcal{L}(u,p) = \sup_{q \in \mathbb{R}_+^m} \mathcal{L}(u,q)$ si bien que

$$\forall q \in \mathbb{R}_+^m, \qquad q \cdot F(u) \le p \cdot F(u).$$

S'il existait $i \in \{1,\dots,m\}$ tel que $F_i(u) > 0$, on pourrait faire exploser le minorant en faisant tendre q_i vers $+\infty$, ce qui est absurde. On a donc $F(u) \le 0$, c'est-à-dire $u \in K$. De plus, pour tout $i \notin \mathcal{A}(u)$, on a $F_i(u) < 0$ et dans ce cas, nécessairement $p_i = 0$ car si p_i était strictement positif, on pourrait augmenter strictement la quantité $p \cdot F(u)$ en diminuant la composante p_i. Par suite, la condition des écarts complémentaires $p \cdot F(u) = 0$ est satisfaite. Enfin, en considérant le fait que $\mathcal{L}(u,p) = \inf_{v \in U} \mathcal{L}(v,p)$ et le fait que $K \subset U$ par hypothèse, il vient pour tout $v \in K$,

$$J(u) = \mathcal{L}(u,p) \le \mathcal{L}(v,p) = J(v) + p \cdot F(v) \le J(v),$$

puisque $p \cdot F(u) = 0$ et $p \cdot F(v) \le 0$. En conséquence, u est un minimiseur global de J sur K. □

Lorsque les fonctionnelles J et F_i sont *convexes* et différentiables, nous disposons d'un résultat *d'équivalence* entre la résolution du problème d'optimisation sous contraintes et la recherche d'un point selle pour le Lagrangien (on rappelle que dans cette section, nous avons supposé que les applications composantes F_i, $1 \le i \le m$, sont convexes). Le résultat ci-dessous est connu sous le nom de théorème de Kuhn et Tucker.

Théorème 3.40. *On suppose que les fonctionnelles J et F_i $(1 \le i \le m)$ sont* convexe *et différentiables en $u \in K$. On suppose en outre que les contraintes inégalités $F_i \le 0$ sont qualifiées en u. Alors, les assertions suivantes sont équivalentes :*
(i) *u est un minimiseur global de J sur K ;*
(ii) *il existe $p \in \mathbb{R}_+^m$ tel que le couple (u,p) est un point selle du Lagrangien $\mathcal{L}$ sur $V \times \mathbb{R}_+^m$;*
(iii) *il existe $p \in \mathbb{R}_+^m$ tel que les relations (3.22) et (3.23) sont satisfaites.*

Preuve. L'implication (i) $\Rightarrow$ (iii) est l'assertion du théorème 3.36. L'implication (ii) $\Rightarrow$ (i) résulte de la proposition 3.39. Il reste donc à prouver l'implication (iii) $\Rightarrow$ (ii). Pour tout $q \in \mathbb{R}_+^m$, il est clair que $q \cdot F(u) \le 0 = p \cdot F(u)$ de par la relation des écarts complémentaires (3.23). D'où

$$\mathcal{L}(u,p) = \sup_{q \in \mathbb{R}_+^m} \mathcal{L}(u,q).$$

Par ailleurs, à p fixé, la fonctionnelle $J_p : V \ni v \mapsto J(v) + p \cdot F(v)$ étant convexe par hypothèse, la relation (3.22) implique que le point u est un point critique, et donc un minimiseur global de J_p. Il en résulte que pour tout $v \in V$, $J_p(u) \le J_p(v)$, ou encore

$$\mathcal{L}(u,p) = \inf_{v \in U} \mathcal{L}(v,p),$$

ce qui complète la preuve. □

Remarque 3.41. *Attention ! Le théorème 3.36 ne dit rien sur l'unicité des multiplicateurs de Lagrange. De même, le théorème 3.40 ne dit rien sur l'unicité du vecteur $p \in \mathbb{R}_+^m$ intervenant dans le point selle. Bien sûr, si la famille $\{\nabla F_i(u)\}_{1 \le i \le m}$ est libre (dans V), ces multiplicateurs sont uniques.*

Remarque 3.42. *On vérifie facilement en considérant la fonctionnelle H définie par (3.29) que*

$$H(v) = \sup_{q \in \mathbb{R}_+^m} \mathcal{L}(v,q) = \begin{cases} J(v) & si\ v \in K, \\ +\infty & sinon. \end{cases}$$

Ainsi, minimiser la fonctionnelle H sur l'ensemble U (qui contient K) revient à chercher le minimiseur u de J sur K. Par ailleurs, maximiser la fonctionnelle G définie par (3.28) fournit un multiplicateur de Lagrange $p \in \mathbb{R}_+^m$. L'équivalence entre les deux approches est exprimée par le biais de la proposition 3.38. Lorsque des méthodes d'optimisation numérique sont conçues à partir de cette équivalence, on parle de méthodes de dualité.

3.7 Exercices

▷ **3.1.** Soit m et n deux entiers positifs et A une matrice de taille $m \times n$. Soit $b \in \mathbb{R}^m$. On appelle solution du système linéaire $Ax = b$ au sens des moindres carrés tout point $x \in \mathbb{R}^n$ minimum de $\|Ax - b\|_m^2$. En utilisant la projection orthogonale de $\mathbb{R}^m$ sur l'image de A, montrer que le problème

$$\inf_{x \in \mathbb{R}^n} \|Ax - b\|_m^2,$$

admet au moins une solution et que toute solution est caractérisée par les équations normales $A^t Ax = A^t b$.

▷ **3.2.** Soit V un espace de Hilbert, K un sous-espace fermé de V et $u_0 \notin K$. Montrer qu'il existe un hyperplan fermé de V qui sépare u_0 de K, i.e., qu'il existe une forme linéaire continue $f \in V'$ et un réel α tels que

$$\forall v \in K, \quad \langle f, v \rangle_{V',V} > \alpha > \langle f, u_0 \rangle_{V',V}.$$

On considérera la forme linéaire $\langle f, v \rangle_{V',V} = (P_K u_0 - u_0, v)_V$.

▷ **3.3.** Soit V un espace de Hilbert, α et δ deux réels avec $\alpha > 0$ et $v \in V$. On considère la fonctionnelle (quadratique) de V dans $\mathbb{R}$ donnée par

$$J(u) = \delta + (v, u)_V + \alpha \|u\|_V^2.$$

Montrer que J est minorée sur V puis qu'il existe un unique minimum de J sur V.

▷ **3.4.** Cet exercice nécessite la connaissance des espaces L^2 introduits au chapitre 5. Soit $I =]a, b[$ un intervalle borné de $\mathbb{R}$. On se donne un entier $N \geq 2$ et on construit un maillage uniforme de I en posant $x_i = a + ih$, $0 \leq i \leq N$, $h = \frac{b-a}{N}$. On note K l'espace des fonctions continues sur $[0, 1]$ et affines par morceaux sur chaque maille $]x_i, x_{i+1}[$.

- Pour $0 \leq i \leq N$, on définit la fonction $\varphi_i \in K$ par les relations $\varphi_i(x_j) = \delta_{ij}$ où δ_{ij} est le symbole de Kronecker. Représenter graphiquement les fonctions φ_i, $0 \leq i \leq N$, et montrer que $(\varphi_0, \ldots, \varphi_N)$ forme une base de K.
- Soit $u \in L^2(I)$ et $(U_0, \ldots, U_N)$ les composantes de $P_K u$ dans la base $(\varphi_0, \ldots, \varphi_N)$. Préciser le système linéaire satisfait par les U_i.

▷ **3.5.** Soit V et W deux espaces de Hilbert équipés respectivement des produits scalaires $(\cdot, \cdot)_V$ et $(\cdot, \cdot)_W$. Supposons que $W \subset V$ avec injection continue et dense. Soit $T : V \to W'$ l'application qui à $v \in V$ associe $Tv \in W'$ défini par

$$\forall w \in W, \quad \langle Tv, w \rangle_{W',W} = (v, w)_V.$$

– Montrer que T est linéaire, continue et injective.
– Montrer que $T(V)$ est dense dans W' (on utilisera le théorème 1.25).

▷ **3.6.** On pose $V = l^2$ et on introduit

$$W = \{\, \mathtt{u} = (u_n)_{n\geq 0},\ \sum_{n\geq 0} n^2 u_n^2 < +\infty \,\}.$$

On équipe W du produit scalaire $(\mathtt{u}, \mathtt{v})_W = \sum_{n\geq 0} n^2 u_n v_n$.
– Montrer que $W \subset V$ avec injection continue et dense.
– On utilise le produit scalaire de V pour identifier V et V'. On note $\mathtt{1} = (1)_{n\geq 0}$ la suite constante égale à un. Montrer que $\mathtt{1} \in W'$ mais que $\mathtt{1} \notin V'$ (et a fortiori $\mathtt{1} \notin W$).
– On utilise le produit scalaire de W pour identifier W et W'. On note $\mathtt{h} = (1/n)_{n\geq 0}$ la suite harmonique. Montrer que $\mathtt{h} \in V$ mais que $\mathtt{h} \notin V'$.

▷ **3.7.** Convergence faible (exercice difficile)

Soit V un espace de Hilbert de dimension infinie. On note $(\cdot,\cdot)$ le produit scalaire de V et $\|\cdot\|$ la norme associée.

1. On dit qu'une suite $(u_n)_{n\in\mathbb{N}}$ d'éléments de V converge faiblement vers un élément u de V, ce qu'on note $u_n \rightharpoonup u$ si pour tout $v \in V$, $(u_n, v) \to (u, v)$. On dit alors que u est la limite faible de la suite $(u_n)_{n\in\mathbb{N}}$.

   (a) Montrer que si une suite $(u_n)_{n\in\mathbb{N}}$ admet une limite faible, celle-ci est unique.

   (b) Montrer que si la suite $(u_n)_{n\in\mathbb{N}}$ converge fortement (ce qui signifie “pour la topologie de la norme”) vers u, alors $(u_n)_{n\in\mathbb{N}}$ converge faiblement vers u.

   (c) Montrer que si $(u_n)_{n\in\mathbb{N}}$ converge faiblement vers u et si $\|u_n\| \to \|u\|$, alors $(u_n)_{n\in\mathbb{N}}$ converge fortement vers u.

   (d) Montrer que si $(u_n)_{n\in\mathbb{N}}$ converge faiblement vers u, alors

   $$\|u\| \leq \liminf_{n\to+\infty} \|u_n\|.$$

2. On suppose dans cette question (et dans cette question seulement) que V séparable.

   (a) Soit $(e_n)_{n\in\mathbb{N}}$ une base hilbertienne de V. Montrer que la suite $(e_n)_{n\in\mathbb{N}}$ n'admet aucune sous-suite fortement convergente, puis que la suite $(e_n)_{n\in\mathbb{N}}$ converge faiblement vers 0 dans V.

   (b) Montrer, en utilisant un procédé diagonal, que toute suite bornée $(u_n)_{n\in\mathbb{N}}$ de V admet une sous-suite faiblement convergente (ce résultat signifie que la boule unité, qui n'est pas compacte pour la topologie forte - la topologie de la norme -, est compacte pour la topologie faible).

3. Montrer que même si V n'est pas séparable, toute suite bornée $(u_n)_{n\in\mathbb{N}}$ de V admet une sous-suite faiblement convergente. *Indication : on pourra considérer l'espace de Hilbert* $V = \overline{\mathrm{Vect}(u_n)}$.

4. Soit $J : V \to \mathbb{R}$ une fonctionnelle convexe et continue sur V et $(u_n)_{n\in\mathbb{N}}$ une suite d'élements de V qui converge faiblement vers un certain $u \in V$. En supposant que J est différentiable en u, montrer que
$$J(u) \le \liminf_{n\to\infty} J(u_n).$$
En fait la propriété ci-dessus reste vraie même si J n'est pas différentiable en u.

5. Utiliser résultats établis dans les questions précédentes pour prouver le Théorème 3.28.

$\triangleright$ **3.8.** L'objectif de cet exercice est de montrer que si V est un espace de Hilbert de dimension infinie, alors la boule unité de V n'est pas compacte. Soit $(x_1, x_2, x_3, \cdots)$ une suite de vecteurs de V linéairement indépendants (i.e. telle que pour tout $N \in \mathbb{N}^*$, $\dim(\mathrm{Vect}(x_1, x_2, \cdots, x_N)) = N$). Le *procédé d'orthogonalisation de Gram–Schmidt* consiste à construire à partir de la suite $(x_1, x_2, x_3, \cdots)$ la suite $(e_1, e_2, e_3, \cdots)$ définie par
$$e_1 = \frac{x_1}{\|x_1\|_V}$$
et par récurrence,
$$y_k = x_k - \sum_{j=1}^{k-1} (e_j, y_k)_V \, e_j, \quad e_k = \frac{y_k}{\|y_k\|_V}.$$

- Montrer que la suite $(e_1, e_2, e_3, \cdots)$ est orthonormée (i.e. que $(e_k, e_l)_V = \delta_{kl}$).
- Montrer qu'on ne peut pas extraire de la suite $(e_k)_{k\in\mathbb{N}^*}$ une sous-suite convergente.

$\triangleright$ **3.9.** Prolongement des formes linéaires.

Soit V un espace de Hilbert et Y un sous-espace vectoriel de V. Soit A une forme linéaire continue sur Y de norme $\|A\|_{Y'}$. Montrer qu'il existe une forme linéaire $\breve{A} \in V'$ qui prolonge A au sens où $\langle \breve{A}, y\rangle_{V',V} = \langle A, y\rangle_{Y',Y}$ pour tout $y \in Y$ et $\|\breve{A}\|_{V'} = \|A\|_{Y'}$. *Indication : prolonger A à $\overline{Y}^V$ par continuité puis utiliser le théorème de projection orthogonale.*

$\triangleright$ **3.10.** Soit V un espace vectoriel normé dont la norme vérifie la formule de la médiane
$$\forall (v, w) \in V \times V, \qquad \|v\|_V^2 + \|w\|_V^2 = \frac{1}{2}\|v + w\|_V^2 + \frac{1}{2}\|v - w\|_V^2.$$
Montrer que cette norme dérive d'un produit scalaire. *Indication : pour $(x, y) \in V \times V$, on posera*
$$(x, y)_V = \frac{1}{4}\left(\|x + y\|_V^2 - \|x - y\|_V^2\right),$$
en on montrera que $(\cdot, \cdot)_V$ vérifie toutes les propriétés requises pour être un produit scalaire sur V.

$\triangleright$ **3.11.** Optimisation dans des espaces de Banach uniformément convexes (exercice difficile)

On appelle module de convexité d'un espace de Banach X la fonction $\delta : [0, 2] \to \mathbb{R}_+$ définie par
$$\delta(\epsilon) = \inf\left\{1 - \left\|\frac{v + w}{2}\right\|, \ (v, w) \in X \times X, \ \|v\| \le 1, \ \|w\| \le 1, \ \|v - w\| \ge \epsilon\right\}.$$
On dit qu'un espace de Banach est *uniformément convexe* si $\delta(\epsilon) > 0$ pour tout $0 < \epsilon \le 2$.

On dit par ailleurs qu'une fonction $J : X \to \overline{\mathbb{R}}$ est semi-continue inférieurement (s.c.i.) si, pour tout $v \in X$ et toute suite $(v_n)_{n\in\mathbb{N}}$ d'éléments de X qui converge vers v dans X, on a

$$J(v) \leq \liminf_{n\to+\infty} J(v_n).$$

Le but de cet exercice est de montrer le résultat suivant, qui généralise le Théorème 3.28 (nous verrons en effet que tout espace de Hilbert est un espace de Banach uniformément convexe).

Théorème 3.28 bis. *Soit X un espace de Banach uniformément convexe, K un sous-ensemble de X convexe, fermé, non vide, et J une fonction convexe et s.c.i. sur K, infinie à l'infini. Alors le problème*

$$\inf_{v\in K} J(v) \tag{3.30}$$

admet un minimiseur.

1. Montrer que pour tout $r \geq 0$, $0 \leq \epsilon \leq 2r$ et $(u, v, w) \in X \times X \times X$,
$$(\|u-v\| \leq r,\ \|u-w\| \leq r,\ \|v-w\| \geq \epsilon) \quad \Rightarrow \quad \left\| u - \frac{v+w}{2} \right\| \leq (1-\delta(\epsilon))\, r.$$

2. Montrer que si X est un espace de Hilbert,
$$\forall \epsilon \in [0,2], \quad \delta(\epsilon) = 1 - \left(1 - \frac{\epsilon^2}{4}\right)^{1/2},$$
et en déduire que tout espace de Hilbert est un espace de Banach uniformément convexe.

3. Soit X un espace de Banach uniformément convexe et $(C_n)_{n\in\mathbb{N}}$ une suite de sous-ensemble convexes, fermés, bornés, non vides de X décroissante pour l'inclusion. Soit $(v_n)_{n\in\mathbb{N}}$ une suite d'éléments de V telle que
$$\forall n \in \mathbb{N}, \quad v_n \in C_n \quad \text{et} \quad \|v_n\| \leq \inf_{w_n \in C_n} \|w_n\| + \frac{1}{1+n}.$$
Montrer que la suite $(v_n)_{n\in\mathbb{N}}$ est de Cauchy. En déduire que
$$\bigcap_{n\in\mathbb{N}} C_n \neq \emptyset.$$

4. Soit X un espace de Banach, K un sous-ensemble convexe, fermé, non vide de X et $J : K \to \mathbb{R}$. Pour tout $\lambda \in \mathbb{R}$, on note
$$S_\lambda = J^{-1}(]-\infty, \lambda]) \subset K \subset X,$$
l'ensemble de sous-niveau λ de J.
   (a) Montrer que si J est convexe, alors S_λ est convexe.
   (b) Montrer que si J est s.c.i., alors S_λ est fermé.
   (c) Montrer que si J est infinie à l'infini, alors S_λ est borné.

5. Prouver le Théorème 3.28 bis. *Indication : on montrera d'abord que $I = \inf_{v\in K} J(v) > -\infty$ en appliquant le résultat de la question 3 aux ensembles $C_n = J^{-1}(]-\infty, -n])$, puis on prouvera que I est atteint en appliquant ce même résultat aux ensembles $C_n = S_{I+1/(n+1)} = J^{-1}(]-\infty, I + 1/(n+1)])$.*

6. Application : soit V un espace de Banach uniformément convexe, K un sous-ensemble convexe, fermé et non vide de V et u un élément de V. Montrer qu'il existe un unique $w \in K$ tel que
$$\|u-w\|_V = d(u,K) = \inf_{v\in K} \|u-v\|_V.$$
Le point v est appelé le *projeté* de u sur C.

Chapitre 4

Théorie de la mesure et de l'intégration

Ce chapitre présente les bases de la théorie de la mesure et de l'intégration. Partant du constat que la notion d'intégrale de Riemann souffre de plusieurs insuffisances, et notamment celle de réquerir des hypothèses très fortes pour permettre le passage à la limite sous le signe somme, l'objectif de ce chapitre est d'introduire le concept de mesure positive puis de poser les bases de la construction de l'intégrale de Lebesgue. Le point d'orgue de ce chapitre est la section 4.4 qui passe en revue les principaux théorèmes de convergence valables pour l'intégrale de Lebesgue, et notamment le théorème 4.42, dit de «convergence dominée», dont nous ferons un usage fréquent dans la suite de cet ouvrage. Un autre fruit de la notion d'intégrale de Lebesgue, que nous récolterons au chapitre suivant, est celui de la construction des espaces fonctionnels L^p qui jouent un rôle fondamental dans de nombreux domaines de l'analyse dont la théorie des EDP et les probabilités.

4.1 Pourquoi aller au-delà de l'intégrale de Riemann ?

Rappelons qu'une fonction $f : [a,b] \longrightarrow \mathbb{R}$ est dite *intégrable au sens de Riemann* si les sommes de Riemann

$$\sum_{i=0}^{N-1} f(\xi_i)\,(x_{i+1}-x_i), \qquad a = x_0 \leq x_1 \leq x_2 \leq \cdots \leq x_N = b, \quad \xi_i \in [x_i, x_{i+1}]$$

convergent lorsque le pas de discrétisation $h = \sup\limits_{0 \leq i \leq N-1} |x_{i+1}-x_i|$ tend vers 0 ; si tel est le cas, l'intégrale de Riemann de f sur $[a,b]$ est par définition la limite des sommes de Riemann.

Graphiquement, l'intégrale d'une fonction positive sur l'intervalle $[a,b]$ représente l'aire de la surface S située entre l'axe des abscisses et la courbe $y = f(x)$; les sommes de Riemann approchent donc cette grandeur selon une méthode de découpage de l'axe des abscisses représentée sur la figure 4.1.

Quoique suffisamment riche pour permettre d'intégrer beaucoup de fonctions usuelles, notamment les fonctions continues par morceaux, l'intégrale de Riemann souffre de plusieurs insuffisances, dont la principale est de réquerir des hypothèses très fortes (du type convergence uniforme) pour permettre des passages à la limite sous le signe somme

$$\int \lim_{n\to+\infty} f_n \overset{?}{=} \lim_{n\to+\infty} \int f_n.$$

Cette difficulté technique n'est pas anodine : elle révèle un manque de complétude de l'ensemble des fonctions Riemann intégrables.

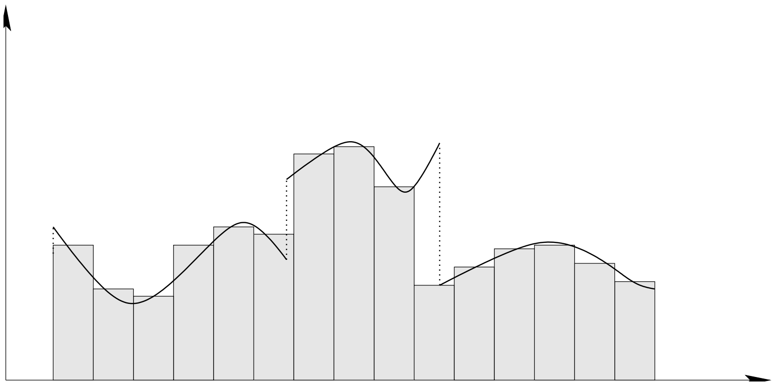


Fig. 4.1 – Approximation d'une intégrale par une somme de Riemann

La méthode de Lebesgue consiste à approcher par valeurs inférieures l'aire de la surface S (pour une fonction f supposée ici positive bornée) à l'aide d'un découpage de l'axe des *ordonnées* :

$$\text{aire}(S) \simeq \sum_{i=0}^{N-1} y_i \text{mes}\,(A_i)\,,$$

où mes (A_i) désigne la mesure de l'ensemble

$$A_i = f^{-1}\left(]y_i, y_{i+1}]\right) = \{x \in [a,b], \quad y_i < f(x) \leq y_{i+1}\}$$

et où $0 = y_0 < y_1 < \cdots < y_N = \max\limits_{[a,b]} f$.

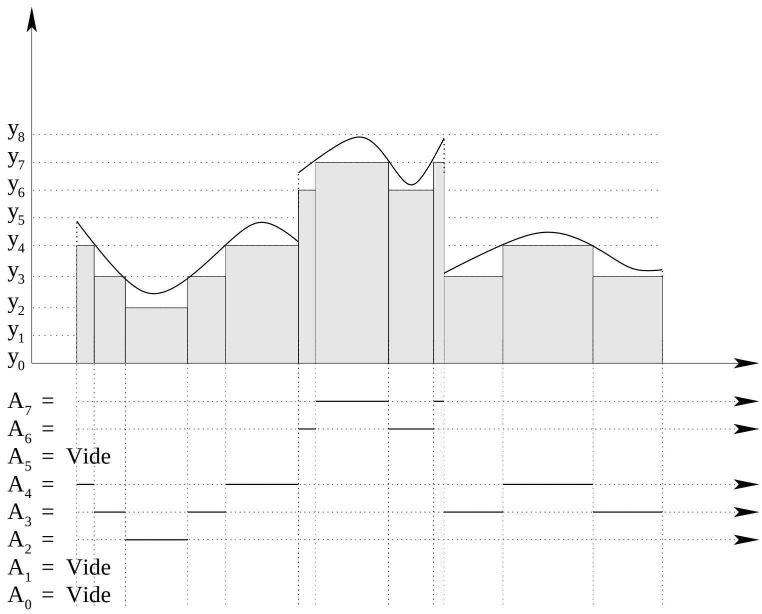


Fig. 4.2 – Approximation d'une intégrale par la méthode de Lebesgue

Remarquons que l'erreur commise en remplaçant $f(x)$ par y_i sur A_i est contrôlée par $|y_{i+1} - y_i|$ (alors qu'il n'y a aucun contrôle de ce type sur les sommes de Riemann). L'intégrale de Lebesgue s'avère en fait bien plus puissante que l'intégrale de Riemann. On verra en effet

– que les résultats de convergence dont on dispose (notamment le théorème de convergence dominée) sont bien plus souples que ceux auxquels on aboutit avec l'intégrale de Riemann (cf. section 4.4) ;
– que les espaces fonctionnels qui découlent de cette construction ont une structure d'espace de Banach voire d'espace de Hilbert (cf. chapitre 5), ce qui permettra par la suite de mettre en œuvre les théorèmes de géométrie établis dans les deux chapitres précédents pour résoudre des problèmes d'analyse (cf. chapitre 8).

Mais auparavant, il nous faut construire rigoureusement l'intégrale de Lebesgue (c'est l'objet de la section 4.3) et pour cela définir précisément ce que recouvre le concept de mesure (ce point est examiné à la section 4.2).

Remarque 4.1. Dans ce chapitre on va être amené à manipuler des fonctions à valeurs dans $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$. On étend les opérations algébriques sur $\mathbb{R}$ par les conventions
– $\forall \alpha \in \mathbb{R}$, $\alpha \pm \infty = \pm\infty$;
– $\forall \alpha > 0$, $\alpha \times (\pm\infty) = \pm\infty$;
– $0 \times (\pm\infty) = 0$.

On s'abstient en revanche, sous peine d'aboutir à une impasse, de définir les opérations $+\infty + (-\infty)$ et $+\infty - (+\infty)$. Remarquons que $\overline{\mathbb{R}}$ n'est pas un espace vectoriel puisque les deux opérations ci-dessus sont interdites. En conséquence, on ne peut pas définir la somme de deux fonctions valant $+\infty$ et $-\infty$ respectivement en un même point.

4.2 Eléments de théorie de la mesure

Quelle est la mesure de l'intervalle $[a, b]$? Spontanément, on a envie de répondre que c'est $b - a$, ce qui est à la fois
– une bonne réponse en ce sens que le réel positif $b - a$ représente ce qu'on appelle la *mesure de Lebesgue* de l'intervalle $[a, b]$;
– une mauvaise réponse car il y a d'autres façons de mesurer l'intervalle $[a, b]$.

Il ne faut pas penser qu'une mesure qui n'associerait pas au segment $[a, b]$ le nombre $b - a$ serait une curiosité mathématique sans intérêt pratique ; c'est notamment dans le domaine des probabilités qu'on est amené à manipuler constamment ce genre d'objets. La mesure de Lebesgue étant cependant la mesure sur $\mathbb{R}$ la plus naturelle (nous verrons plus loin ce que cela signifie précisément) et la plus utile en analyse des EDP, ce sera sur elle que nous mettrons l'accent dans ce cours ; c'est également à elle qu'on se référera dans la suite lorsqu'on parlera de la mesure (sans autre précision) d'un intervalle ou d'un ensemble plus complexe de $\mathbb{R}$.

4.2.1 Tribus et mesures

Soit X un ensemble et $\mathcal{P}(X)$ l'ensemble des parties de X (autrement dit l'ensemble des sous-ensembles de X). Une mesure positive sur X est une application μ d'un sous-ensemble $\mathcal{S}$ de $\mathcal{P}(X)$ dans $\overline{\mathbb{R}}_+$; un élément de $\mathcal{S}$ est appelé un ensemble mesurable. Commençons par recenser les propriétés qu'il est raisonnable d'imposer à une application μ pour qu'elle mérite le label de mesure :
– l'ensemble vide doit être mesurable (i.e. appartenir à $\mathcal{S}$) et sa mesure doit être nulle ; l'ensemble X doit aussi être mesurable mais sa mesure peut être finie ou infinie ;
– si A et B sont mesurables et disjoints alors $A \cup B$ doit être mesurable et on doit avoir

$$\mu(A \cup B) = \mu(A) + \mu(B).$$

– si A et B sont mesurables et tels que $A \subset B$, $B \setminus A$ doit être mesurable et on doit avoir

$$\mu(B \setminus A) = \mu(B) - \mu(A)$$

si $\mu(B) < +\infty$;

On vérifie facilement que pour satisfaire ces conditions, l'ensemble $\mathcal{S}$ des ensembles mesurables doit être une *algèbre*, au sens suivant :

Définition 4.2. *Soit X un ensemble. Une* algèbre *est une famille $\mathcal{A}$ de parties de X ($\mathcal{A} \subset \mathcal{P}(X)$) vérifiant les propriétés suivantes :*

1. *$\emptyset$ et X sont dans $\mathcal{A}$;*
2. *si $A \in \mathcal{A}$, alors $X \setminus A \in \mathcal{A}$;*
3. *si A_1 et A_2 sont dans $\mathcal{A}$, alors $A_1 \cup A_2$ est aussi dans $\mathcal{A}$.*

À titre d'exemple, l'ensemble des réunions finies d'intervalles de $\mathbb{R}$ constitue une algèbre (le vérifier en exercice).

Proposition 4.3. *Soit X un ensemble et $\mathcal{A}$ une algèbre sur X.*

1. *Si A_1, A_2, ..., A_N sont dans $\mathcal{A}$ alors $\displaystyle\bigcup_{1\le k\le N} A_k$ et $\displaystyle\bigcap_{1\le k\le N} A_k$ sont aussi dans $\mathcal{A}$.*
2. *Si A et B sont dans $\mathcal{A}$ et tels que $A \subset B$, alors $B \setminus A$ est aussi dans $\mathcal{A}$.*

Preuve. Laissée en exercice. $\square$

Sur une algèbre, on peut donner la définition suivante d'une mesure :

Définition 4.4. *Soit $\mathcal{A}$ une algèbre sur X. Une mesure (positive) sur $\mathcal{A}$ est une application $\mu : \mathcal{A} \longrightarrow \overline{\mathbb{R}}_+$ possédant les propriétés suivantes :*

1. *$\mu(\emptyset) = 0$;*
2. *Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'ensembles de $\mathcal{A}$ deux à deux disjoints telle que $\displaystyle\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$, alors*

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

On vérifie facilement qu'une mesure sur une algèbre vérifie bien les propriétés suivantes :

Proposition 4.5. *Soit $\mathcal{A}$ une algèbre sur X et μ une mesure sur $\mathcal{A}$. Soit A et B deux ensembles appartenant à l'algèbre $\mathcal{A}$. Les ensembles $A \cup B$ et $A \cap B$ sont dans $\mathcal{A}$ et on a*

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B).$$

Si $A \subset B$, l'ensemble $B \setminus A$ est aussi dans $\mathcal{A}$ et

$$\mu(B \setminus A) = \mu(B) - \mu(A).$$

Preuve. Laissée en exercice. $\square$

Le point faible de la structure d'algèbre réside dans le fait qu'une union dénombrable d'ensembles appartenant tous à l'algèbre n'est pas nécessairement dans l'algèbre (c'est clair si on considère à nouveau l'exemble de l'algèbre des réunions finies d'intervalles de $\mathbb{R}$). Ce manque de "complétude" rend difficile beaucoup de manipulations d'analyse (notamment le passage à la limite dans les suites). Pour cette raison, on va introduire une nouvelle structure, dite de tribu, qui remédie à ce problème :

Définition 4.6. *Soit X un ensemble et $\mathcal{P}(X)$ l'ensemble des parties de X. Une* tribu *est une famille $\mathcal{T}$ de parties de X ($\mathcal{T} \subset \mathcal{P}(X)$) vérifiant les propriétés suivantes :*

1. *$\emptyset$ et X sont dans $\mathcal{T}$;*
2. *si $A \in \mathcal{T}$, alors $X \setminus A \in \mathcal{T}$;*
3. *si $(A_n)_{n\in\mathbb{N}}$ est une suite de parties de X qui appartiennent toutes à $\mathcal{T}$, alors $\bigcup_{n\in\mathbb{N}} A_n$ est aussi dans $\mathcal{T}$.*

Un élément de $\mathcal{T}$ est appelé un ensemble mesurable.

Remarque 4.7. À un ensemble X quelconque, on peut toujours associer les deux tribus suivantes :
- la tribu grossière $\mathcal{T}_0 = \{\emptyset, X\}$,
- la tribu discrète $\mathcal{T}_\infty = \mathcal{P}(X)$.

Malheureusement, la tribu $\mathcal{T}_0$ ne présente aucun intérêt, et la tribu $\mathcal{T}_\infty$ (qui est très utilisée pour les ensembles finis) est trop grosse pour qu'on puisse l'employer pour définir la mesure de Lebesgue sur $\mathbb{R}$ (sauf à renoncer à l'axiome du choix, cf. remarque 4.21 et exercice 4.5). C'est cela qui justifie l'introduction du concept de tribu.

Définition 4.8. *Soit $\mathcal{T}$ une tribu sur X. Une* mesure *(positive) sur $\mathcal{T}$ est une application $\mu : \mathcal{T} \longrightarrow \overline{\mathbb{R}}_+$ possédant les propriétés suivantes :*

1. *$\mu(\emptyset) = 0$;*
2. *Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'ensembles de $\mathcal{T}$ deux à deux disjoints, alors*

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

Définition 4.9.
- *Un espace X muni d'une tribu $\mathcal{T}$ s'appelle un* espace mesurable.
- *Un espace X muni d'une tribu $\mathcal{T}$ et d'une mesure μ sur cette tribu s'appelle un* espace mesuré.

Remarque 4.10. Il est évident que toute tribu est une algèbre ; les propriétés énoncées dans la proposition 4.5 restent donc valables pour une mesure définie sur une tribu.

Enonçons maintenant une proposition dont on fera usage ultérieurement, qui découle directement de la propriété de σ-additivité.

Proposition 4.11. *Soit $(X, \mathcal{T}, \mu)$ un espace mesuré et $(A_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{T}$ croissante pour l'inclusion (i.e. $A_0 \subset A_1 \subset A_2 \subset \cdots$). Soit $B = \bigcup_{n\in\mathbb{N}} A_n$. Alors $B \in \mathcal{T}$ et*

$$\mu(B) = \lim_{n\to+\infty} \mu(A_n).$$

Preuve. En utilisant la convention $A_{-1} = \emptyset$, on peut écrire

$$B = \bigcup_{n\in\mathbb{N}} (A_n \setminus A_{n-1}),$$

la réunion ci-dessus étant cette fois-ci formée d'ensemble deux à deux disjoints. Il découle alors de la propriété de σ-additivité que

$$\mu(B) = \sum_{n\in\mathbb{N}} \mu(A_n \setminus A_{n-1}).$$

Ceci implique que

$$\mu(A_N) = \mu\left(\bigcup_{n=0}^{N} (A_n \setminus A_{n-1})\right) = \sum_{n=0}^{N} \mu(A_n \setminus A_{n-1}) \underset{N\to+\infty}{\longrightarrow} \sum_{n\in\mathbb{N}} \mu(A_n \setminus A_{n-1}) = \mu(B).$$

□

Nous verrons par la suite sur l'exemple de la mesure de Lebesgue qu'il est facile de construire des mesures sur des algèbres. Il en en revanche beaucoup plus difficile de construire "à la main" des mesures (utiles en pratique) sur des tribus ; on utilise pour cela un théorème puissant, le théorème de Carathéodory, qui permet de prolonger une mesure définie sur une algèbre à la tribu engendrée par cette algèbre, le prolongement étant unique sous certaines hypothèses :

Définition 4.12. *Soit $\mathcal{S} \subset \mathcal{P}(X)$. On appelle* tribu engendrée par $\mathcal{S}$ *la plus petite tribu contenant $\mathcal{S}$ (i.e. l'intersection de toutes les tribus contenant $\mathcal{S}$ qui, on le vérifiera en exercice, est une tribu contenant $\mathcal{S}$).*

Théorème 4.13. *(Théorème de Carathéodory). Soit X un ensemble, $\mathcal{A}$ une algèbre sur X et μ une mesure sur $\mathcal{A}$. On suppose qu'il existe une suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{A}$ telle que*

$$X = \bigcup_{n\in\mathbb{N}} A_n \qquad \textit{et} \qquad \mu(A_n) < +\infty, \quad \forall n \in \mathbb{N}$$

(on dit alors que μ est σ-finie). Alors il existe une unique mesure, encore notée μ, qui prolonge μ à la tribu engendrée par $\mathcal{A}$.

Une preuve du théorème de Carathéodory peut être lue dans la référence [Revuz 94], pages 54 et 154.

Concluons cette section en remarquant qu'on peut associer à un espace vectoriel normé (et plus généralement à un espace topologique, voir [AF 88] page 529), une tribu "naturelle" qui jouera un rôle essentiel dans la suite :

Définition 4.14. *Soit X un espace topologique. La* tribu borélienne *sur X est la tribu $\mathcal{B}(X)$ engendrée par l'ensemble $\mathcal{O}$ des ouverts de X. Un élément B de la tribu borélienne $\mathcal{B}(X)$ est appelé un* borélien *de X.*

4.2.2 Ensembles négligeables

Les ensembles de mesure nulle jouent un rôle important dans la théorie de la mesure et de l'intégration. Nous donnons ici deux définitions à retenir.

Définition 4.15. *Soit $(X, \mathcal{T}, \mu)$ un espace mesuré. Un ensemble $A \in \mathcal{T}$ est dit* négligeable *si $\mu(A) = 0$.*

Définition 4.16. *On dit qu'une propriété est vraie* presque partout *si elle est vraie en dehors d'un ensemble inclus dans un ensemble négligeable.*

4.2.3 Mesure de Lebesgue sur $\mathbb{R}^d$

Nous détaillons dans cette section la construction de la mesure de Lebesgue[1] sur (la tribu borélienne de) $\mathbb{R}$; nous indiquons ensuite comment procéder pour construire la mesure de Lebesgue sur (la tribu borélienne de) $\mathbb{R}^d$.

Soit a et b deux réels tels que $a \leq b$. Dans cette section, la notation (a, b) désignera l'un quelconque des quatre intervalles $]a, b[$, $]a, b]$, $[a, b[$ ou $[a, b]$.

Proposition 4.17. *L'ensemble $\mathcal{A}$ des réunions finies d'intervalles forme une algèbre sur $\mathbb{R}$ et la fonction μ définie sur $\mathcal{A}$ par*

$$\forall A = \bigcup_{i=1}^{N} (a_i, b_i) \in \mathcal{A}, \quad \textit{avec } a_1 \leq b_1 \leq a_2 \leq b_2 \leq \cdots \leq a_N \leq b_N,$$

$$\mu(A) = \sum_{i=1}^{N} (b_i - a_i),$$

est une mesure sur $\mathcal{A}$.

Preuve. Laissée en exercice. □

Définition - Théorème 4.18. *La mesure μ définie dans la proposition 4.17 se prolonge de manière unique en une mesure sur la tribu borélienne de $\mathbb{R}$, appelée* mesure de Lebesgue.

Preuve. Le caractère σ-fini de la mesure sur l'algèbre des réunions finies d'intervalles définie par la proposition 4.17 se démontre en remarquant que

$$\mathbb{R} = \bigcup_{N \in \mathbb{N}} [-N, N] \qquad \text{et} \qquad \mu([-N, N]) = 2N < +\infty, \quad \forall N \in \mathbb{N}.$$

La tribu borélienne étant engendrée par l'algèbre $\mathcal{A}$ (cf. exercice 4.3), on peut donc conclure en utilisant le théorème de Carathéodory. □

Remarque 4.19. Pour construire la mesure de Lebesgue sur $\mathbb{R}^d$, on s'y prend de même :

1. on commence par définir la mesure d'un pavé $P = \prod_{i=1}^{n} (a_i, b_i)$ par $\mu(P) = \prod_{i=1}^{n} (b_i - a_i)$ (par convention, ce produit vaut 0 si l'un des facteurs $(b_i - a_i)$ est nul, même si d'autres facteurs valent $+\infty$) ; cette mesure est la mesure "naturelle" des surfaces dans $\mathbb{R}^2$, des volumes dans $\mathbb{R}^3$, ...
2. on peut ainsi définir une mesure sur l'algèbre $\mathcal{A}$ des réunions finies de pavés ;
3. cette mesure étant σ-finie, on peut la prolonger en une mesure sur la tribu engendrée par $\mathcal{A}$, qui est en fait la tribu borélienne de $\mathbb{R}^d$ (exercice 4.4).

La spécificité de la mesure de Lebesgue tient au fait que c'est la seule (à une constante multiplicative près) qui soit invariante par translation :

[1] Une remarque concernant la terminologie : il est souvent d'usage d'appeler *mesure de Borel* la mesure sur la tribu borélienne que nous allons construire, de réserver le terme *mesure de Lebesgue* à la mesure qui prolonge la mesure de Borel à la tribu de Lebesgue (voir remarque 4.19). Comme la distinction entre mesure de Borel et mesure de Lebesgue n'est pas fondamentale pour notre propos, nous parlerons de mesure de Lebesgue dans les deux cas.

Proposition 4.20. *La tribu borélienne $\mathcal{B}(\mathbb{R}^d)$ et la mesure de Lebesgue λ sur $\mathbb{R}^d$ sont invariantes par translation : pour tout borélien $B \in \mathcal{B}(\mathbb{R}^d)$, et pour tout $a \in \mathbb{R}^d$, le translaté de B*

$$\tau_a(B) = \{x \in \mathbb{R}^d, \quad x - a \in B\}$$

est un borélien et $\lambda(\tau_a(B)) = \lambda(B)$.

Preuve. L'algèbre $\mathcal{A}$ des réunions finies de pavés de $\mathbb{R}^d$ ainsi que la mesure μ définie dans la remarque 4.19 sont clairement invariants par translation. Il en résulte d'une part que la tribu engendrée par $\mathcal{A}$ est invariante par translation ($\mathcal{A}$ et $\tau_a(\mathcal{A})$ sont toutes deux la "plus petite tribu" contenant les réunions finies de pavés et sont donc égales), et d'autre part que λ est invariante par translation (λ et $\tau_a\lambda$ définie par $\tau_a\lambda(B) = \lambda(\tau_a(B))$ pour tout $B \in \mathcal{B}(\mathbb{R}^d)$) prolongent toutes deux la mesure définie dans la remarque 4.19, et sont donc égales en vertu du théorème de Carathéodory). □

Remarque 4.21. La tribu borélienne de $\mathbb{R}^d$ n'est pas égale à la tribu discrète $\mathcal{T}_\infty = \mathcal{P}(\mathbb{R}^d)$; elle est cependant suffisamment grande pour contenir les ensembles communément rencontrés en pratique (et particulier les ouverts, les fermés et les réunions dénombrables d'ouverts et de fermés) si bien qu'il est rarissime qu'on ait à se demander si telle ou telle partie de $\mathbb{R}^d$ est mesurable ou non pour cette tribu. Il est cependant naturel de se demander si on peut étendre la mesure de Lebesgue sur les boréliens à une tribu plus grande. La réponse à cette question est subtile :

- on peut, de manière unique, étendre la mesure de Lebesgue à la plus petite tribu complète contenant la tribu borélienne (on dit qu'une tribu est complète si tout ensemble inclus dans un ensemble négligeable est lui-même négligeable). La tribu en question, appelée tribu de Lebesgue, est elle-aussi plus petite que la tribu discrète ;
- si on veut conserver la possibilité d'avoir recours à l'axiome du choix (voir [AF 87], page 21), on ne peut pas prolonger la mesure de Lebesgue en une mesure sur $\mathcal{T}_\infty$; il faut alors accepter qu'il y ait des ensembles non mesurables (l'exercice 4.5 propose d'en construire un) ;
- si on renonce à l'axiome du choix, il n'y a plus de contradiction à supposer que tout ensemble est mesurable pour la mesure de Lebesgue.

4.3 Construction de l'intégrale

4.3.1 Fonctions mesurables

Définition 4.22. *Soit $(X, \mathcal{T})$ un espace mesurable. Une fonction $f : X \longrightarrow \overline{\mathbb{R}}$ est dite* mesurable *si*

$$\forall \alpha \in \mathbb{R}, \qquad f^{-1}(]\alpha, +\infty]) = \{x \in X, \quad f(x) > \alpha\} \in \mathcal{T}.$$

Remarque 4.23. Si f est à valeurs réelles, f est mesurable si et seulement si l'image réciproque de tout ensemble mesurable de la tribu borélienne de $\mathbb{R}$ est dans $\mathcal{T}$.

La notion de fonction mesurable est stable par les opérations élémentaires de l'analyse :

Proposition 4.24. *Soit $(X, \mathcal{T})$ un espace mesurable. Soit f et g deux fonctions mesurables de X à valeurs dans $\overline{\mathbb{R}}$, et α et β deux réels. Alors les fonctions $|f|$, fg, $\max(f, g)$ et $\min(f, g)$ sont mesurables. Il en est de même pour les fonctions $\alpha f + \beta g$ et f/g sous réserve que celles-ci soient bien définies.*

Preuve. Laissée en exercice (l'énoncé de l'exercice 4.6 fournit des indications). □

La notion de fonction mesurable a en outre le bon goût d'être stable par passage à la limite dénombrable :

Proposition 4.25. *Soit $(X, \mathcal{T})$ un espace mesurable. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables de X dans $\overline{\mathbb{R}}$. Alors*

1. *les fonctions $\underline{f}(x) = \inf_{n\in\mathbb{N}} f_n(x)$ et $\overline{f}(x) = \sup_{n\in\mathbb{N}} f_n(x)$ sont mesurables ;*
2. *les fonctions $\underline{l}(x) = \liminf_{n\in\mathbb{N}} f_n(x)$ et $\overline{l}(x) = \limsup_{n\in\mathbb{N}} f_n(x)$ sont mesurables ;*
3. *si la suite de fonctions $(f_n)_{n\in\mathbb{N}}$ converge simplement vers une fonction f, alors f est mesurable.*

On rappelle qu'on dit que $(f_n)_{n\in\mathbb{N}}$ *converge simplement* vers f si $\forall x \in X$, $f_n(x) \underset{n\to+\infty}{\longrightarrow} f(x)$.

Preuve.

1. Pour prouver la mesurabilité de $\overline{f}$, il suffit de remarquer que
$$\overline{f}^{-1}(]\alpha, +\infty]) = \bigcup_{n\in\mathbb{N}} f_n^{-1}(]\alpha, +\infty]) \in \mathcal{T}$$
(une union dénombrable d'ensembles mesurables est mesurable). Ce résultat permet de conclure aussi à la mesurabilité de $\underline{f}$ puisque $\underline{f} = -\sup_{n\in\mathbb{N}}(-f_n(x))$.
2. Il résulte de ce qui précède que les fonctions
$$\underline{l}(x) = \liminf_{n\in\mathbb{N}} f_n(x) = \sup_{N\in\mathbb{N}} \left(\inf_{n\geq N} f_n(x) \right)$$
et
$$\overline{l}(x) = \limsup_{n\in\mathbb{N}} f_n(x) = \inf_{N\in\mathbb{N}} \left(\sup_{n\geq N} f_n(x) \right)$$
sont aussi mesurables.
3. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions qui converge simplement vers une fonction f. On a alors
$$f(x) = \liminf_{n\in\mathbb{N}} f_n(x) = \limsup_{n\in\mathbb{N}} f_n(x) \ ;$$
la fonction f est donc mesurable de par le point 2.

□

Comme exemples de fonctions mesurables, on peut citer :

1. les fonctions caractéristiques d'ensemble mesurables. Soit en effet $A \in \mathcal{T}$ et χ_A sa *fonction caractéristique* (on dit encore *indicatrice*) définie par
$$\chi_A(x) = \left| \begin{array}{ll} 1 & \text{si } x \in A, \\ 0 & \text{si } x \notin A. \end{array} \right.$$
On a
$$\chi_A^{-1}(]\alpha, +\infty]) = \left| \begin{array}{ll} \emptyset & \text{si } \alpha \geq 1, \\ A & \text{si } 0 \leq \alpha < 1, \\ X & \text{si } \alpha < 0. \end{array} \right.$$
Comme les ensembles $\emptyset$, A et X appartiennent à la tribu $\mathcal{T}$, la fonction χ_A est mesurable ;

2. les *fonctions étagées*, qui sont par définition les combinaisons linéaires *finies* de fonctions caractéristiques d'ensembles mesurables ; une fonction étagée e peut donc s'écrire sous la forme
$$e(x) = \sum_{i=1}^{N} \alpha_i \, \chi_{A_i}(x)$$
où les fonctions χ_{A_i} sont mesurables en vertu du point 1 ci-dessus ; il résulte de la proposition 4.24 que la fonction étagée e est bien mesurable ;

3. les limites simples de fonctions étagées sont donc aussi mesurables ; ceci découle de la proposition 4.25.

Nous venons en fait d'exhiber *toutes* les fonctions mesurables, comme le prouve le

Théorème 4.26. *Toute fonction mesurable est limite simple d'une suite de fonctions étagées. En outre*

1. *si f est bornée, la convergence est uniforme ;*
2. *si f est positive, la suite e_n peut être choisie croissante et composée de fonctions positives.*

Preuve. Supposons pour commencer f positive. Pour tout $n \in \mathbb{N}$ et tout entier $1 \leq k \leq 2^{2n}$, on pose
$$A_k^n = f^{-1}\left(\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right[\right).$$
On pose également
$$A_\infty^n = f^{-1}\left([2^n, +\infty]\right).$$
Ces ensembles étant mesurables,
$$e_n(x) = \sum_{k=1}^{2^{2n}} \frac{k-1}{2^n} \, \chi_{A_k^n}(x) + 2^n \, \chi_{A_\infty^n}(x)$$
définit une fonction étagée positive sur l'espace mesuré $(X, \mathcal{T})$. On vérifie facilement que

– la suite de fonctions étagées positives $(e_n)_{n\in\mathbb{N}}$ est croissante ;
– pour tout $x \in X$,
$$\begin{cases} 0 \leq f(x) - e_n(x) \leq \dfrac{1}{2^n} & \text{si } f(x) < 2^n, \\ e_n(x) = 2^n & \text{si } f(x) \geq 2^n. \end{cases}$$

Il est donc clair que la suite $(e_n)_{n\in\mathbb{N}}$ converge simplement vers f, et uniformément sur X si f est bornée.

Si f n'est pas positive, on construit comme précédemment deux suites croissantes $(e_n^+)_{n\in\mathbb{N}}$ et $(e_n^-)_{n\in\mathbb{N}}$ de fonctions étagées positives qui convergent simplement vers $f_+ = \max(f, 0)$ et $f_- = -\min(f, 0)$ respectivement. La suite de fonctions étagées $(e_n)_{n\in\mathbb{N}}$ définie par $\forall n \in \mathbb{N}$, $e_n = e_n^+ - e_n^-$ converge simplement vers f. □

Remarque 4.27. Nous avons vu précédemment (cf. remarque 4.21) que tous les ensembles que l'on rencontre en pratique en analyse sont mesurables. De même, toutes les fonctions qu'on manipule usuellement sont mesurables. En particulier, les fonctions continues définies sur un espace mesurable à valeurs dans $\overline{\mathbb{R}}$ sont mesurables puisque l'image réciproque de l'ouvert $]\alpha, +\infty]$ de $\overline{\mathbb{R}}$ par une application continue est un ouvert, donc un borélien.

4.3.2 Intégrale d'une fonction mesurable

La construction de l'intégrale se fait en trois étapes :

1. on commence par définir l'intégrale des fonctions étagées positives ;
2. on construit ensuite l'intégrale d'une fonction mesurable positive ;
3. on définit ensuite l'intégrale (si elle existe) d'une fonction mesurable de signe quelconque.

Définition 4.28. *Soit $(X, \mathcal{T}, \mu)$ un espace mesuré et $e(x) = \sum_{i=1}^{N} \alpha_i \chi_{A_i}(x)$ une fonction étagée positive (on peut toujours supposer les A_i deux à deux disjoints, ce qui implique alors que tous les α_i sont positifs). On appelle* intégrale de e sur X par rapport à la mesure μ *et on note* $\int_X e\, d\mu$ *l'élément de $\mathbb{R}_+ \cup \{+\infty\}$ défini par*

$$\int_X e\, d\mu = \sum_{i=1}^{N} \alpha_i\, \mu(A_i).$$

On dit que la fonction e est intégrable *si* $\int_X e\, d\mu < +\infty$.

On vérifiera à titre d'exercice que la valeur de l'intégrale ne dépend pas de la décomposition de la fonction e en combinaison linéaire finie de fonctions caractéristiques d'ensembles mesurables.

Définition 4.29. *Soit $(X, \mathcal{T}, \mu)$ un espace mesuré et f une fonction positive mesurable. On appelle* intégrale de f sur X par rapport à la mesure μ *et on note* $\int_X f\, d\mu$ *l'élément de $\mathbb{R}_+ \cup \{+\infty\}$ défini par*

$$\int_X f\, d\mu = \sup\left\{\int_X e\, d\mu, \quad e \text{ étagée vérifiant } 0 \le e \le f\right\}.$$

On dit que la fonction f est intégrable *si* $\int_X f\, d\mu < +\infty$.

Définition 4.30. *Soit $(X, \mathcal{T}, \mu)$ un espace mesuré et f une fonction mesurable. Soit $f = f_+ - f_-$ la décomposition de f en la différence de deux fonctions positives ($f_+ = max(f, 0)$, $f_- = max(-f, 0)$). On dit que f est intégrable si f_+ et f_- sont intégrables ; on appelle alors* intégrale de f sur X par rapport à la mesure μ *le réel*

$$\int_X f\, d\mu = \int_X f_+\, d\mu - \int_X f_-\, d\mu.$$

On note $\mathcal{L}^1(X, \mathcal{T}, \mu)$ l'ensemble des fonctions intégrables.

Définition 4.31. *Soit $(X, \mathcal{T}, \mu)$ un espace mesuré, $Y \subset X$ un ensemble mesurable et f une fonction mesurable. On dit que f est intégrable sur Y (par rapport à la mesure μ) si la fonction $f\chi_Y$ est intégrable sur X. L'*intégrale de f sur Y *(par rapport à la mesure μ) est alors le réel noté* $\int_Y f\, d\mu$ *défini par*

$$\int_Y f\, d\mu = \int_X f\, \chi_Y\, d\mu.$$

Proposition 4.32. *Si Y est un ensemble négligeable, alors toute fonction mesurable f est intégrable sur Y et*

$$\int_Y f\,d\mu = 0.$$

Preuve. Laissée en exercice. □

Proposition 4.33. *Soit $f \in \mathcal{L}^1(X, \mathcal{T}, \mu)$ une fonction intégrable et $A = \{x \in X, \quad f(x) = \pm\infty\}$. L'ensemble A est de mesure nulle.*

Preuve. Laissée en exercice. □

4.3.3 Propriétés fondamentales de l'intégrale

Proposition 4.34. *Soit $(X, \mathcal{T}, \mu)$ un espace mesuré. L'intégrale possède les deux propriétés suivantes :*

1. Linéarité : *soit f et g deux fonctions intégrables, α et β deux réels. Si elle est définie (cf. remarque 4.1), la fonction $\alpha f + \beta g$ est intégrable et*

$$\int_X (\alpha f + \beta g)\,d\mu = \alpha\left(\int_X f\,d\mu\right) + \beta\left(\int_X g\,d\mu\right) \;;$$

2. Monotonie : *soit f et g deux fonctions intégrables telles que $f \leq g$ sur X. Alors*

$$\int_X f\,d\mu \leq \int_X g\,d\mu.$$

Preuve.

1. Si f, g, α et β sont positifs, il découle directement de la définition 4.29 de l'intégrale des fonctions positives que

$$\int_X (\alpha f + \beta g)\;d\mu \geq \alpha \int_X f\,d\mu + \beta \int_X g\,d\mu.$$

   Le fait que cette inégalité soit en fait une égalité résulte du théorème 4.39 (de convergence monotone) qui sera énoncé plus loin. Pour des fonctions et des coefficients de signe quelconque, on démontre la linéarité de l'intégrale en décomposant les fonctions f et g en la somme de leur partie positive et de leur partie négative.

2. De par la linéarité de l'intégrale

$$\int_X g\,d\mu - \int_X f\,d\mu = \int_X (f-g)\,d\mu.$$

   La fonction $f - g$ étant positive, son intégrale l'est aussi (cela résulte de la définition 4.29). D'où le résultat de monotonie annoncé.

□

4.3.4 Intégrale de Lebesgue sur $\mathbb{R}^d$

L'intégrale de Lebesgue sur $\mathbb{R}^d$ (resp. sur un borélien B de $\mathbb{R}^d$) est l'intégrale sur $\mathbb{R}^d$ (resp. sur B) par rapport à la mesure de Lebesgue construite dans la section 4.2.3. On utilise généralement les notations abrégées $\mathcal{L}^1(\mathbb{R}^d)$ (resp. $\mathcal{L}^1(B)$) pour désigner l'ensemble des fonctions intégrables sur $\mathbb{R}^d$ (resp. sur un borélien B de $\mathbb{R}^d$) par rapport à la mesure de Lebesgue ; en outre, pour $f \in \mathcal{L}^1(\mathbb{R}^d)$ (resp. $f \in \mathcal{L}^1(B)$), on note souvent

$$\int_{\mathbb{R}^d} f \qquad (\text{resp. } \int_B f),$$

sans préciser la mesure, ou encore

$$\int_{\mathbb{R}^d} f(x)\,dx \qquad (\text{resp. } \int_B f(x)\,dx),$$

l'intégrale de f sur $\mathbb{R}^d$ (resp. sur B) par rapport à la mesure de Lebesgue.

Enonçons tout de suite un résultat important garantissant que ce nouveau concept d'intégrale colle avec l'ancien :

Théorème 4.35. *Soit $[a,b]$ un intervalle borné de $\mathbb{R}$ et $f\,:\,[a,b] \longrightarrow \mathbb{R}$ une fonction bornée. Si f est Riemann intégrable, alors f est Lebesgue intégrable et les deux intégrales ont même valeur.*

Ce résultat (qui s'adapte aux dimensions supérieures) a une conséquence pratique importante : on pourra calculer numériquement l'intégrale de Lebesgue d'une fonction continue par morceaux en calculant son intégrale de Riemann. Les méthodes d'intégration numérique usuelles (formule du trapèze, formule de Simpson, formules de Gauss, ...) sont en effet basées sur le principe de découpage de l'axe des abscisses sur lequel repose l'intégrale de Riemann. Et bien sûr, on dispose donc toujours du théorème suivant, qui est bien commode pour calculer analytiquement les intégrales des fonctions dont on connaît une primitive.

Théorème 4.36. *Soit $[a,b]$ un intervalle borné de $\mathbb{R}$, $f\,:\,[a,b] \longrightarrow \mathbb{R}$ une fonction continue et F une primitive de f. On a*

$$\int_a^b f = F(b) - F(a).$$

Avant de donner la preuve du théorème 4.35, rappelons que la définition de l'intégrale de Riemann est fondée sur la notion de fonction en escalier (alors que l'intégrale de Lebesgue est fondée sur la notion de fonction étagée). Une fonction en escalier est une combinaison linéaire finie de fonctions caractéristiques d'intervalles de $\mathbb{R}$, et l'intégrale de Riemann sur $[a,b]$ d'une fonction en escalier $e = \displaystyle\sum_{i=0}^{N-1} \alpha_i \chi_{(x_i,x_{i+1})}$ avec $a = x_0 \leq x_1 \leq \cdots \leq x_N = b$ est définie par $\displaystyle\int_a^b e(x)\,dx = \sum_{i=0}^{N-1} \alpha_i\ (x_{i+1} - x_i)$. Il est donc clair que toute fonction en escalier sur $[a,b]$ est une fonction étagée intégrable, et que sur l'ensemble $\mathrm{Esc}([a,b])$ des fonctions en escalier sur $[a,b]$, intégrales de Riemann et de Lebesgue coïncident. Rappelons enfin que la fonction f est dite Riemann intégrable si les deux réels

$$\sup_{e\in \mathrm{Esc}([a,b]),\, e\leq f} \int_a^b e(x)\,dx \tag{4.1}$$

et

$$\inf_{e\in \mathrm{Esc}([a,b]),\, e\geq f} \int_a^b e(x)\,dx \tag{4.2}$$

sont égaux, l'intégrale de Riemann de f sur $[a,b]$ étant alors la valeur commune de ces deux réels.

Preuve du théorème 4.35. Soit f une fonction bornée positive Riemann intégrable sur $[a,b]$. Notons R la valeur de l'intégrale de Riemann de f et L celle de l'intégrale de Lebesgue de f. Comme l'ensemble des fonctions en escalier sur $[a,b]$ est inclus dans l'ensemble des fonctions étagées sur $[a,b]$, on vérifie sans peine en utilisant (4.2) et la définition 4.29 que $R \leq L$. Soit maintenant $\epsilon > 0$ et une fonction en escalier e telle que $e \geq f$ et $\displaystyle\int_a^b e(x)\,dx \leq R - \epsilon$. De par la monotonie de l'intégrale de Lebesgue, $L \leq \displaystyle\int_a^b e(x)\,dx$ (rappelons qu'intégrales de Riemann et de Lebesgue coïncident sur l'ensemble des fonctions en escalier). Donc $L \leq R - \epsilon$. Cette inégalité est vraie pour tout $\epsilon > 0$. Donc $L \leq R$. Finalement $L = R$. Pour une fonction f de signe quelconque, on prouve cette égalité en décomposant f en ses parties positive et négative et en utilisant la linéarité de l'intégrale. □

Remarque 4.37. La réciproque du théorème 4.35 est fausse : il existe des fonctions bornées Lebesgue intégrables qui ne sont pas Riemann intégrables (cf. exercice 4.7).

La remarque ci-dessus prouve la supériorité de l'intégrale de Lebesgue. Cette dernière se comporte également mieux que l'intégrale de Riemann vis-à-vis des singularités. En effet, une fonction définie sur un intervalle borné $[a,b]$ présentant une singularité en $c \in]a,b[$, par exemple la fonction $f(x) = \dfrac{1}{\sqrt{|x|}}$ sur $[-1,1]$, n'est jamais Riemann intégrable au sens usuel (les séries de Riemann ne convergent pas) ; pour définir son intégrale au sens de Riemann généralisé, il faut considérer les quantités

$$I_{\epsilon_-,\epsilon_+} = \int_a^{c-\epsilon_-} f(x)\,dx + \int_{c+\epsilon_+}^b f(x)\,dx$$

et étudier leur comportement lorsque ϵ_- et ϵ_+ tendent vers 0^+. Si les $I_{\epsilon_-,\epsilon_+}$ convergent, on définit l'intégrale de Riemann (généralisée) de f sur $[a,b]$ par

$$\int_a^b f(x)\,dx = \lim_{\epsilon_-\to 0^+,\ \epsilon_+\to 0^+} I_{\epsilon_-,\epsilon_+}.$$

Définir l'intégrale de Riemann (généralisée) d'une fonction non bornée requiert donc un traitement spécifique des singularités, et ce traitement n'est pas pratiquable si l'ensemble des singularités a un point d'accumulation intérieur à $[a,b]$. À l'inverse, les définitions 4.29-4.30 de l'intégrale de Lebesgue sont valables pour toute fonction possédant ou non des singularités. Là encore, quand une fonction définie sur un borné et présentant un nombre fini de singularités est intégrable au sens de Riemann généralisé, elle est aussi intégrable au sens de Lebesgue et les valeurs des deux intégrales coïncident.

Remarque 4.38. Il arrive en revanche que sur un intervalle *non borné*, une fonction intégrable au sens de Riemann généralisé ne soit pas intégrable au sens de Lebesgue ; ceci est dû au fait que la mesure de Lebesgue d'un intervalle non borné est infinie. Ainsi, par exemple la fonction $f(x) = \dfrac{\sin x}{1+|x|}$ est intégrable sur $\mathbb{R}$ au sens de Riemann généralisé, en ce sens que

$$\lim_{R_-\to+\infty,\ R_+\to+\infty} \int_{-R_-}^{R_+} f(x)\,dx$$

existe, mais n'est pas Lebesgue intégrable.

4.4 Théorèmes de convergence

Les théorèmes présentés ci-dessous permettent de “passer à la limite sous le signe somme” sous des hypothèses bien moins restrictives que celles imposées pour les fonctions intégrables au sens de Riemann.

4.4.1 Convergence monotone

Le théorème suivant, dit de convergence monotone, s'utilise rarement tel quel en pratique ; il est en revanche tout à fait fondamental d'un point de vue théorique car c'est sur lui que reposent les preuves de plusieurs résultats centraux de la théorie de l'intégration.

Théorème 4.39. *(Théorème de convergence monotone). Soit $(X, \mathcal{T}, \mu)$ un espace mesuré et $(f_n)_{n\in\mathbb{N}}$ une suite croissante de fonctions positives mesurables. On note f la limite simple de la suite $(f_n)_{n\in\mathbb{N}}$. On a*

$$\int_X f\,d\mu = \lim_{n\to+\infty} \int_X f_n\,d\mu.$$

Preuve. Comme pour tout $n \in \mathbb{N}$, $f_n \leq f_{n+1}$ (la suite $(f_n)_{n\in\mathbb{N}}$ est croissante), il découle de l'alinea 2 de la proposition 4.34 que $\displaystyle\int_X f_n\,d\mu \leq \int_X f_{n+1}\,d\mu$, ce qui montre que la suite réelle positive $\displaystyle\left(\int_X f_n\,d\mu\right)_{n\in\mathbb{N}}$ est croissante, donc converge vers un certain $\alpha \in \mathbb{R}^+ \cup \{+\infty\}$. Puisque la fonction $f = \lim\limits_{n\to+\infty} f_n$ est mesurable (cf. proposition 4.25) et puisque pour tout $n \in \mathbb{N}$, $f_n \leq f$, on a $\displaystyle\int_X f_n\,d\mu \leq \int_X f\,d\mu$, et on obtient donc, en faisant tendre n vers l'infini,

$$\alpha \leq \int_X f\,d\mu. \tag{4.3}$$

Considérons maintenant une fonction étagée e vérifiant $0 \leq e \leq f$ et un réel $0 < c < 1$. Posons

$$A_n = \{x \in X, \quad f_n(x) \geq ce(x)\}.$$

Il est clair que $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{T}$ croissante pour l'inclusion (i.e. $A_0 \subset A_1 \subset A_2 \subset \cdots$) ; en outre, $\displaystyle\bigcup_{n\in\mathbb{N}} A_n = X$. Pour voir cela, il suffit de remarquer que $x \in A_0$ si $f(x) = 0$ et que $ce(x) < f(x)$ dans le cas contraire, ce qui implique que $x \in A_n$ pour n assez grand puisque $(f_n(x))_{n\in\mathbb{N}}$ tend vers $f(x)$. De plus

$$c\int_{A_n} e\,d\mu \leq \int_{A_n} f_n\,d\mu \leq \int_X f_n\,d\mu,$$

et donc

$$c\int_{A_n} e\,d\mu \leq \alpha. \tag{4.4}$$

En décomposant e sous la forme $e = \displaystyle\sum_{i=1}^N \beta_i \chi_{B_i}$ ou les β_i sont des réels positifs et les B_i des ensembles mesurables, on peut écrire

$$\int_{A_n} e\,d\mu = \int_X \sum_{i=1}^N \beta_i \chi_{B_i}\chi_{A_n} = \int_X \sum_{i=1}^N \beta_i \chi_{B_i\cap A_n} = \sum_{i=1}^N \beta_i \mu(B_i \cap A_n).$$

La suite $(B_i \cap A_n)_{n\in\mathbb{N}}$ étant une suite croissante d'ensembles mesurables telle que $\bigcup_{n\in\mathbb{N}} (B_i \cap A_n) = B_i$, on déduit de la proposition 4.11 que

$$\lim_{n\to+\infty} \mu(B_i \cap A_n) = \mu(B_i).$$

En faisant tendre n vers $+\infty$ dans (4.4), il vient donc

$$c \int_X e \, d\mu \leq \alpha.$$

Cette égalité étant vérifiée pour tout $0 < c < 1$, il s'en suit que

$$\int_X e \, d\mu \leq \alpha,$$

et ceci pour toute fonction étagée e telle que $0 \leq e \leq f$. En revenant à la définition 4.29 de l'intégrale des fonctions positives, on voit que

$$\int_X f \, d\mu \leq \alpha$$

et donc, en rapprochant de (4.3), que

$$\int_X f \, d\mu = \alpha = \lim_{n\to+\infty} \int_X f_n \, d\mu.$$

□

De ce résultat découle facilement le résultat suivant :

Théorème 4.40. *(Théorème de Beppo-Levi). Soit $(X, \mathcal{T}, \mu)$ un espace mesuré et $(f_n)_{n\in\mathbb{N}}$ une suite* monotone *(c'est-à-dire soit croissante, soit décroissante) de fonctions intégrables. Soit f la limite simple de la suite $(f_n)_{n\in\mathbb{N}}$. La fonction f est mesurable ; elle est intégrable si et seulement si la suite réelle monotone $\int_X f_n \, d\mu$ admet une limite finie ; si c'est le cas, on a alors*

$$\int_X f \, d\mu = \lim_{n\to+\infty} \int_X f_n \, d\mu.$$

Preuve. Laissée en exercice. □

4.4.2 Lemme de Fatou

Lemme 4.41. *(Lemme de Fatou). Soit $(X, \mathcal{T}, \mu)$ un espace mesuré. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables positives. Alors*

$$\int_X \liminf_{n\to+\infty} f_n \, d\mu \leq \liminf_{n\to+\infty} \int_X f_n \, d\mu.$$

Preuve. La suite de fonctions $(g_n)_{n\in\mathbb{N}}$ définie par $g_n = \inf\limits_{k\geq n} f_k$ est une suite croissante de fonctions positives qui converge simplement vers la fonction $\liminf\limits_{n\to+\infty} f$. Il découle du théorème de convergence monotone que

$$\int_X \liminf_{n\to+\infty} f\, d\mu = \lim_{n\to+\infty} \int_X g_n\, d\mu.$$

Par ailleurs, on a pour tout $n \in \mathbb{N}$ et tout $k \geq n$, $g_n \leq f_k$, d'où il résulte $\displaystyle\int_X g_n\, d\mu \leq \int_X f_k\, d\mu$. On en déduit que

$$\int_X g_n\, d\mu \leq \inf_{k\geq n} \int_X f_k\, d\mu.$$

Finalement

$$\int_X \liminf_{n\to+\infty} f\, d\mu = \lim_{n\to+\infty} g_n\, d\mu \leq \liminf_{n\to+\infty} \int_X f_k\, d\mu.$$

□

4.4.3 Convergence dominée

Théorème 4.42. *(Théorème de convergence dominée). Soit $(X, \mathcal{T}, \mu)$ un espace mesuré. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables et g une fonction intégrable. On suppose que*

$$\begin{cases} f_n(x) \underset{n\to+\infty}{\longrightarrow} f(x) & \textit{presque partout} \\ \forall n \in \mathbb{N}, \quad |f_n(x)| \leq g(x) & \textit{presque partout,} \end{cases}$$

et que f est mesurable. Alors f est intégrable et

$$\int_X f_n\, d\mu \underset{n\to+\infty}{\longrightarrow} \int_X f\, d\mu.$$

Remarque 4.43. On peut éliminer l'hypothèse sur la mesurabilité de f à condition que la tribu $\mathcal{T}$ soit complète (ce qui signifie que tout ensemble inclus dans un ensemble négligeable est lui-même négligeable). C'est le cas de la tribu de Lebesgue sur $\mathbb{R}^d$, qui est la complétion de la tribu borélienne (cf. remarque 4.21).

Preuve du théorème 4.42 . Commençons par supposer que la convergence $f_n \longrightarrow f$ et la majoration $|f_n| \leq g$ ont lieu en tout point de X. En passant à la limite simple dans l'inégalité $|f_n| \leq g$, on obtient $|f| \leq g$. La fonction g étant intégrable par hypothèse, f l'est donc aussi. En outre

$$\left| \int_X f_n\, d\mu - \int_X f\, d\mu \right| \leq \int_X |f_n - f|\, d\mu.$$

Posons $h_n = 2g - |f_n - f|$. La fonction h_n est positive et la suite $(h_n)_{n\in\mathbb{N}}$ converge simplement vers $2g$. En appliquant le lemme de Fatou, il vient

$$\int_X 2g\, d\mu \leq \liminf_{n\to+\infty} \int_X (2g - |f_n - f|)\, d\mu,$$

d'où l'on déduit

$$\limsup_{n\to+\infty} \int_X |f_n - f|\, d\mu \leq 0.$$

Ceci montre que

$$\int_X |f_n - f|\, d\mu \longrightarrow 0,$$

ce qui prouve le résultat escompté. Si maintenant les convergences $f_n \longrightarrow f$ et $|f_n| \leq g$ n'ont lieu que presque partout, introduisons $A_0 = \{x \in X,\ (f_n(x))_{n\in\mathbb{N}}$ ne converge pas vers $f(x)\}$ et $A_n = \{x \in X,\ |f_n(x)| > g(x)\}$. L'ensemble $A = \bigcup_{n\in\mathbb{N}} A_n$ est négligeable (c'est une union dénombrable d'ensembles négligeables). Sur $X \setminus A$ on a par ce qui précède

$$\int_{X\setminus A} |f_n - f|\ d\mu \longrightarrow 0,$$

On conclut en utilisant la proposition 4.32 :

$$\int_X |f_n - f|\ d\mu = \int_{X\setminus A} |f_n - f|\ d\mu \longrightarrow 0.$$

$\square$

Remarque 4.44. Dans convergence dominée, il y a dominée ; attention donc à ne pas appliquer ce théorème sans avoir vérifié l'hypothèse de domination ($|f_n(x)| \leq g(x)$ presque partout). Pour éviter cette erreur, il est utile de garder en tête les exemples ci-dessous. Considérons les trois suites de fonctions positives $(f_n)_{n\in\mathbb{N}^*}$, $(g_n)_{n\in\mathbb{N}^*}$ et $(h_n)_{n\in\mathbb{N}^*}$ définies respectivement pour tout $n \in \mathbb{N}^*$ et tout $x \in \mathbb{R}$ par

$$f_n(x) = n\,\chi(nx), \qquad g_n(x) = n^{-1}\,\chi(x/n), \qquad h_n(x) = \chi(x-n),$$

où $\chi \in C^0(\mathbb{R})$ désigne une fonction positive, à support compact, d'intégrale égale à 1. Chacune de ces suites vérifie l'hypothèse de convergence

$$f_n(x) \underset{n\to+\infty}{\longrightarrow} f(x) = 0 \qquad \text{pour tout } x \in \mathbb{R}\setminus\{0\} \ \text{ (donc presque partout)},$$

$$g_n(x) \underset{n\to+\infty}{\longrightarrow} g(x) = 0 \qquad \text{pour tout } x \in \mathbb{R},$$
$$h_n(x) \underset{n\to+\infty}{\longrightarrow} h(x) = 0 \qquad \text{pour tout } x \in \mathbb{R}.$$

Puisque ces fonctions sont positives, le lemme de Fatou assure que

$$\int_{\mathbb{R}} f \leq \liminf_{n\to+\infty} \int_{\mathbb{R}} f_n, \qquad \int_{\mathbb{R}} g \leq \liminf_{n\to+\infty} \int_{\mathbb{R}} g_n, \qquad \int_{\mathbb{R}} h \leq \liminf_{n\to+\infty} \int_{\mathbb{R}} h_n.$$

Si l'hypothèse de domination était vérifiée, ces limites inf seraient en fait des limites et ces inégalités des égalités. Mais ce n'est pas le cas puisque pour tout $n \in \mathbb{N}^*$,

$$\int_{\mathbb{R}} f_n(x)\,dx = \int_{\mathbb{R}} g_n(x)\,dx = \int_{\mathbb{R}} h_n(x)\,dx = \int_{\mathbb{R}} \chi(x)\,dx = 1,$$

et que par ailleurs

$$\int_{\mathbb{R}} f(x)\,dx = \int_{\mathbb{R}} g(x)\,dx = \int_{\mathbb{R}} h(x)\,dx = 0.$$

Les "pertes de masse" ou "pertes de compacité" que l'on observe sur ces exemples (l'intégrale de la limite est strictement inférieure à la limite de l'intégrale), sont dues aux phénomènes suivants :

- *concentration :* lorsque n tend vers l'infini, la distribution de masse représentée par la fonction f_n se concentre en $\{0\}$, qui est un ensemble de mesure nulle ;
- *évanescence :* le support de la fonction g_n devient infini et dans le même temps $(g_n)_{n\in\mathbb{N}}$ converge uniformément vers 0 ;
- *fuite à l'infini :* le support de la fonction h_n reste borné mais part à l'infini.

Notons que les deux derniers phénomènes ne peuvent se produire que parce que l'on travaille sur un domaine non borné.

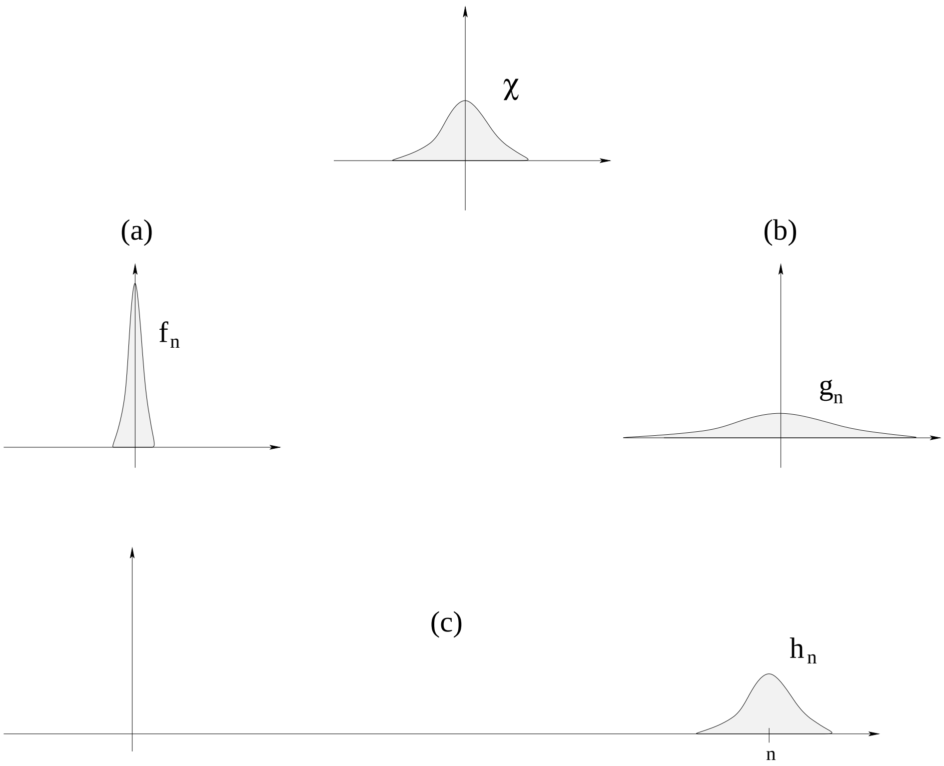


FIG. 4.3 – Perte de compacité par concentration (a), évanescence (b) et fuite à l'infini (c)

4.5 Fonctions définies par une intégrale

On considère un espace mesuré $(Y, \mathcal{T}, \mu)$, Ω un ouvert de $\mathbb{R}^d$, et

$$\begin{array}{rccc} f \,:\, & \Omega \times Y & \longrightarrow & \mathbb{R} \\ & (x,y) & \mapsto & f(x,y) \end{array}$$

une fonction intégrable par rapport à la variable y pour tout $x \in \Omega$. On peut alors définir sur Ω la fonction

$$F(x) = \int_Y f(x,y)\, d\mu(y).$$

Théorème 4.45. *(Continuité de F). Soit $x_0 \in \Omega$. Si pour presque tout $y \in Y$, f est continue en x au voisinage de x_0, i.e. si*

$$f(x,y) \underset{x \to x_0}{\longrightarrow} f(x_0,y) \qquad \text{pour presque tout } y \in Y,$$

et si il existe une fonction $g : Y \longrightarrow \mathbb{R}$ intégrable vérifiant

$$\forall x \in \Omega, \qquad |f(x,y)| \leq g(y) \qquad \text{pour presque tout } y \in Y,$$

alors F est continue en x_0.

Théorème 4.46. *(Différentiabilité de F). Si pour presque tout $y \in Y$, f est de classe C^1 en x sur Ω et si il existe une fonction $g : Y \longrightarrow \mathbb{R}$ intégrable vérifiant*

$$\forall x \in \Omega,\ \forall 1 \leq i \leq n, \qquad \left|\frac{\partial f}{\partial x_i}(x,y)\right| \leq g(y) \qquad \textit{pour presque tout } y \in Y,$$

alors F est de classe C^1 sur Ω et

$$\forall x \in \Omega,\ \forall 1 \leq i \leq n, \quad \frac{\partial F}{\partial x_i}(x) = \int_Y \frac{\partial f}{\partial x_i}(x,y)\, d\mu(y).$$

Preuve des théorèmes 4.45 et 4.46. Ces deux résultats découlent directement du théorème de convergence dominée 4.42. □

4.6 Changement de variables dans les intégrales

Théorème 4.47. *Soit ω et Ω deux ouverts de $\mathbb{R}^d$, $f : \Omega \longrightarrow \mathbb{R}$ une fonction intégrable, $\Phi : \omega \longrightarrow \Omega$ un C^1-difféomorphisme, et $J(y)$ le jacobien de Φ en $y \in \omega$. Alors la fonction $(f \circ \Phi)\,|J|$ est une fonction intégrable sur ω et*

$$\int_\Omega f(x)\, dx = \int_\omega f(\Phi(y))\, |J(y)|\, dy.$$

Par ailleurs, l'égalité ci-dessus est toujours vraie si f est une fonction mesurable positive (dans ce cas, les deux membres peuvent prendre la valeur $+\infty$).

Illustrons ce théorème par quatre exemples à connaître par cœur.

- Coordonnées polaires dans $\mathbb{R}^2$: $\Omega = \mathbb{R}^2 \setminus \{(x,0),\ x \in \mathbb{R}_+\}$, $\omega = \mathbb{R}_+^* \times]0, 2\pi[$ et

$$\begin{array}{cccc} \Phi & : & \omega & \longrightarrow & \Omega \\ & & (r,\theta) & \mapsto & (r\,\cos\theta, r\,\sin\theta)\,. \end{array}$$

  Un calcul simple montre que $J(r,\theta) = r$. Comme la demi-droite $\{(x,0),\ x \in \mathbb{R}_+\}$ est de mesure nulle dans $\mathbb{R}^2$, on obtient que pour toute fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ intégrable,

$$\int_{\mathbb{R}^2} f(x,y)\, dx\, dy = \int_0^{+\infty} \int_0^{2\pi} f(r\,\cos\theta, r\,\sin\theta)\, r\, dr\, d\theta.$$

- Coordonnées cylindriques dans $\mathbb{R}^3$: $\Omega = \mathbb{R}^3 \setminus \{(x,0,z),\ x \in \mathbb{R}_+,\ z \in \mathbb{R}\}$, $\omega = \mathbb{R}_+^* \times]0, 2\pi[\times \mathbb{R}$ et

$$\begin{array}{cccc} \Phi & : & \omega & \longrightarrow & \Omega \\ & & (r,\theta,z) & \mapsto & (r\,\cos\theta, r\,\sin\theta, z)\,. \end{array}$$

  On vérifie facilement que $J(r,\theta,r) = r$. Comme le demi-plan $\{(x,0,z),\ x \in \mathbb{R}_+,\ z \in \mathbb{R}\}$ est de mesure nulle dans $\mathbb{R}^3$, on obtient que pour toute fonction $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ intégrable,

$$\int_{\mathbb{R}^3} f(x,y,z)\, dx\, dy = \int_0^{+\infty} \int_0^{2\pi} \int_{-\infty}^{+\infty} f(r\,\cos\theta, r\,\sin\theta, z)\, r\, dr\, d\theta\, dz.$$

– Coordonnées sphériques dans $\mathbb{R}^3$: $\Omega = \mathbb{R}^3 \setminus \{(x,0,z),\ x \in \mathbb{R}_+,\ z \in \mathbb{R}\}$, $\omega = \mathbb{R}_+^* \times]0,\pi[\times]0,2\pi[$ et

$$\begin{array}{cccc} \Phi & : & \omega & \longrightarrow & \Omega \\ & & (r,\theta,\phi) & \mapsto & (r\,\sin\theta\,\cos\phi, r\,\sin\theta\,\sin\phi, r\,\cos\theta)\,. \end{array}$$

On vérifie facilement que $J(r,\theta,\phi) = r^2\,\sin\theta$. Comme le demi-plan $\{(x,0,z),\ x \in \mathbb{R}_+,\ z \in \mathbb{R}\}$ est de mesure nulle dans $\mathbb{R}^3$, on obtient que pour toute fonction $f\,:\,\mathbb{R}^3 \longrightarrow \mathbb{R}$ intégrable,

$$\int_{\mathbb{R}^3} f(x,y,z)\,dx\,dy\,dz = \int_0^{+\infty}\int_0^{\pi}\int_0^{2\pi} f(r\,\sin\theta\,\cos\phi, r\,\sin\theta\,\sin\phi, r\,\cos\theta)\,r^2\,\sin\theta\,dr\,d\theta\,d\phi.$$

– Changement d'échelle dans $\mathbb{R}^d$: $\Omega = \mathbb{R}^d$, $\omega = \mathbb{R}^d$, et

$$\begin{array}{ccccc} \Phi & : & \omega & \longrightarrow & \Omega \\ & & x & \mapsto & \lambda x, \end{array}$$

où λ est un réel strictement positif fixé. Il vient $J(x) = \lambda^d$, et on obtient ainsi que pour toute fonction $f\,:\,\mathbb{R}^d \longrightarrow \mathbb{R}$ intégrable,

$$\int_{\mathbb{R}^d} f(x)\,dx = \lambda^d \int_{\mathbb{R}^d} f(\lambda x)\,dx.$$

4.7 Mesures produits et théorème de Fubini

4.7.1 Mesures produits

Considérons deux espaces mesurés $(X_1, \mathcal{T}_1, \mu_1)$ et $(X_2, \mathcal{T}_2, \mu_2)$.

On a envie de définir une mesure sur l'ensemble $X = X_1 \times X_2$ en prolongeant la mesure définie sur l'ensemble des pavés

$$\{A = A_1 \times A_2, \quad A_1 \in \mathcal{A}_1,\ A_2 \in \mathcal{A}_2\}$$

par

$$\mu(A) = \mu(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2), \tag{4.5}$$

et étendue à l'ensemble des unions finies de pavés

$$\mathcal{A} = \left\{ \bigcup_{k=1}^{N} A_1^k \times A_2^k, \quad A_1^k \in \mathcal{A}_1,\ A_2^k \in \mathcal{A}_2 \right\} \tag{4.6}$$

en utilisant la relation

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B) \tag{4.7}$$

(une intersection de pavés est un pavé). Ceci est parfaitement licite :

Théorème 4.48.

1. *L'ensemble $\mathcal{A}$ défini par (4.6) est une algèbre sur $X = X_1 \times X_2$ et la fonction μ définie par (4.5)-(4.7) est une mesure sur $\mathcal{A}$.*
2. *Si μ_1 et μ_2 sont σ-finies, μ se prolonge en une mesure, appelée* mesure produit *et notée $\mu_1 \otimes \mu_2$, sur la tribu engendrée par $\mathcal{A}$. Cette tribu, appelée* tribu produit*, est notée $\mathcal{T}_1 \times \mathcal{T}_2$; on a*

$$\mathcal{T}_1 \times \mathcal{T}_2 = \left\{ A = \bigcup_{n \in \mathbb{N}} A_1^n \times A_2^n, \quad A_1^n \in \mathcal{A}_1,\ A_2^n \in \mathcal{A}_2 \right\}.$$

Preuve. Laissée en exercice. □

Remarque 4.49. Considérons l'espace $\mathbb{R}^d$, qu'on peut identifier à l'espace produit $\mathbb{R}^p \times \mathbb{R}^q$ si $p+q=d$. La tribu borélienne de $\mathbb{R}^d$ coïncide bien avec la tribu produit des tribus boréliennes de $\mathbb{R}^p$ et de $\mathbb{R}^q$ (exercice 4.10).

4.7.2 Théorème de Fubini

Le théorème de Fubini donne le cadre dans lequel on peut intervertir deux intégrales.

Théorème 4.50. *(Théorème de Fubini). Soit $(X_1, \mathcal{T}_1, \mu_1)$ et $(X_2, \mathcal{T}_2, \mu_2)$ deux espaces mesurés σ-finis et $f : X_1 \times X_2 \longrightarrow \mathbb{R}$ mesurable par rapport à la tribu produit $\mathcal{T}_1 \times \mathcal{T}_2$. Alors*

1. *si f est positive, les fonctions*

$$x_1 \mapsto \int_{X_2} f(x_1, x_2)\, d\mu_2 \quad et \quad x_2 \mapsto \int_{X_1} f(x_1, x_2)\, d\mu_1$$

sont mesurables et

$$\int_{X_1 \times X_2} f(x_1, x_2)\, d(\mu_1 \otimes \mu_2) = \int_{X_1} \left(\int_{X_2} f(x_1, x_2)\, d\mu_2 \right) d\mu_1 = \int_{X_2} \left(\int_{X_1} f(x_1, x_2)\, d\mu_1 \right) d\mu_2\ ;$$

2. *si f est intégrable sur $X_1 \times X_2$, les fonctions*

$$x_1 \mapsto \int_{X_2} f(x_1, x_2)\, d\mu_2 \quad et \quad x_2 \mapsto \int_{X_1} f(x_1, x_2)\, d\mu_1$$

sont définies presque partout et on a

$$\int_{X_1 \times X_2} f(x_1, x_2)\, d(\mu_1 \otimes \mu_2) = \int_{X_1} \left(\int_{X_2} f(x_1, x_2)\, d\mu_2 \right) d\mu_1 = \int_{X_2} \left(\int_{X_1} f(x_1, x_2)\, d\mu_1 \right) d\mu_2.$$

Preuve. Pour prouver la première assertion, considérons l'application

$$\begin{array}{rcl} \nu\ :\ \mathcal{T}_1 \times \mathcal{T}_2 & \longrightarrow & \overline{\mathbb{R}}_+ \\ A & \mapsto & \displaystyle\int_{X_1} \left(\int_{X_2} \chi_A(x_1, x_2)\, d\mu_2 \right) d\mu_1. \end{array}$$

En s'appuyant sur le théorème de convergence monotone 4.39, on montre que ν est une mesure sur $\mathcal{T}_1 \times \mathcal{T}_2$. Par ailleurs, si $A = A_1 \times A_2$ est un pavé, on a

$$\begin{aligned} \nu(A) &= \int_{X_1} \left(\int_{X_2} \chi_A(x_1, x_2)\, d\mu_2 \right) d\mu_1 \\ &= \int_{X_1} \left(\int_{X_2} \chi_{A_1}(x_1)\, \chi_{A_2}(x_2)\, d\mu_2 \right) d\mu_1 \\ &= \left(\int_{X_1} \chi_{A_1}(x_1)\, d\mu_1 \right) \left(\int_{X_2} \chi_{A_2}(x_2)\, d\mu_2 \right) \\ &= \mu_1(A_1)\, \mu_2(A_2). \end{aligned}$$

La mesure ν coïncide donc avec la mesure μ définie par (4.5) sur l'ensemble des pavés, et donc par (4.7) sur l'algèbre $\mathcal{A}$ des réunions finies de pavés. Par unicité du prolongement des mesures

σ-finies (cf. théorème 4.13), les mesures μ et ν sont égales. On a donc pour toute fonction caractéristique d'ensemble $\mathcal{T}_1 \times \mathcal{T}_2$-mesurable

$$\int_{X_1 \times X_2} f(x_1, x_2)\, d(\mu_1 \otimes \mu_2) = \int_{X_1} \left(\int_{X_2} f(x_1, x_2)\, d\mu_2 \right) d\mu_1$$

Par linéarité, cette égalité reste vraie pour toute fonction étagée positive ; on l'étend à toutes les fonctions positives mesurables en appliquant le théorème de convergence monotone 4.39. La deuxième assertion se prouve en décomposant f en la somme de sa partie positive et de sa partie négative. □

4.8 Exercices

▷ **4.1.** Prouver la proposition 4.5. En déduire que si $(X, \mathcal{T}, \mu)$ est un espace mesuré et si A et B sont deux ensembles mesurables,

$$A \subset B \quad \Rightarrow \quad \mu(A) \leq \mu(B).$$

▷ **4.2.** Montrer que $\mathbb{Q}$ est un borélien de $\mathbb{R}$. Quelle est sa mesure de Lebesgue ?

▷ **4.3.** Montrer que la tribu engendrée par les intervalles de $\mathbb{R}$ est la tribu borélienne de $\mathbb{R}$.

▷ **4.4.** Montrer que la tribu engendrée par les pavés de $\mathbb{R}^d$ est la tribu borélienne de $\mathbb{R}^d$.

▷ **4.5.** On chercher à construire un ensemble non mesurable pour la mesure de Lebesgue sur $\mathbb{R}$ qu'on note ici λ. On considère l'intervalle $I = [0, 1[$ (de mesure de Lebesgue égale à 1) et on note $\sim$ la relation d'équivalence sur I définie par $x \sim y \;\Leftrightarrow\; x - y \in \mathbb{Q}$. Soit maintenant A un sous-ensemble de I construit en prenant un point et un seul dans chaque classe d'équivalence de la relation d'équivalence $\sim$ (c'est là qu'intervient l'axiome du choix, cf. remarque 4.21).

1. Montrer que $I = \bigcup_{r \in \mathbb{Q}} A_r$ avec $A_r = \mod(\{r + A\}, 1)$ (la notation $\{r + A\}$ désigne l'ensemble $\{r + \alpha,\ \alpha \in A\}$ et $\mod(\alpha, 1)$ le réel "α modulo 1", qui est bien un élément de I). Supposons que les ensembles A_r soient mesurables.
2. Montrer que si $r \neq r'$, $A_r \cap A_{r'} = \emptyset$. En déduire que
$$\sum_{r \in \mathbb{Q}} \lambda(A_r) = \lambda(I) = 1.$$
3. En utilisant l'invariance par translation de la mesure de Lebesgue, montrer que $\lambda(A_r)$ est une constante indépendante de r.
4. Conclure.

▷ **4.6.** L'objet de cet exercice est de prouver la proposition 4.24. Soit donc f et g deux fonctions mesurables d'un espace mesurable $(X, \mathcal{T})$ dans $\overline{\mathbb{R}}$.

1. Montrer que
$$(\max(f, g))^{-1}(]\alpha, +\infty]) = f^{-1}(]\alpha, +\infty]) \cup g^{-1}(]\alpha, +\infty]).$$
En déduire que la fonction $\max(f, g)$ est mesurable. Montrer de même que $\min(f, g)$ est mesurable.

2. Montrer que pour tout $\alpha \in \mathbb{R}$
$$(f+g)^{-1}(]\alpha,+\infty]) = \bigcup_{q\in\mathbb{Q}} \left(f^{-1}(]q,+\infty]) \cap g^{-1}(]\alpha - q, +\infty])\right).$$
En déduire que la fonction $f+g$ est mesurable.
3. En utilisant le résultat du point 1, montrer que $f_+ = \max(f,0)$ et $f_- = -\min(f,0)$ sont mesurables ; en déduire que $|f|$ est une fonction mesurable.
4. Montrer que pour tout $\alpha \in \mathbb{R}$, αf est mesurable. En déduire que la fonction $\alpha f + \beta g$ pour α et β réels est mesurable si elle est bien définie (c'est-à-dire si pour tout $x \in X$, on n'a jamais $(\alpha f, \beta g) = (+\infty, -\infty)$ ou $(\alpha f, \beta g) = (-\infty, +\infty)$).
5. Montrer que la fonction fg est mesurable.
6. Montrer que si f ne s'annule jamais, la fonction $1/f$ est mesurable ; en déduire que sous cette condition, f/g l'est aussi.

▷ **4.7.** Soit $[a,b]$ un intervalle borné de $\mathbb{R}$ et $\chi_{\mathbb{Q}\cap[a,b]}$ la fonction caractéristique de l'ensemble des nombres rationnels inclus dans $[a,b]$. Montrer que $\chi_{\mathbb{Q}\cap[a,b]}$ est Lebesgue intégrable, mais n'est pas Riemann intégrable.

▷ **4.8.** Montrer que la fonction
$$f(x) = \frac{\sin x}{1+|x|}$$
est intégrable sur $\mathbb{R}$ au sens de Riemann (généralisé). Est-elle intégrable sur $\mathbb{R}$ au sens de Lebesgue ?

▷ **4.9.** Montrer que
$$\int_0^1 \left(\int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}\,dy\right)dx = \frac{\pi}{4}.$$
En déduire que la fonction $f(x,y) = \dfrac{x^2-y^2}{(x^2+y^2)^2}$ n'est pas intégrable dans $[0,1]\times[0,1]$.

▷ **4.10.** Considérons l'espace $\mathbb{R}^d$, qu'on peut identifier à l'espace produit $\mathbb{R}^p \times \mathbb{R}^q$ si $p+q=d$. Vérifier que la tribu borélienne de $\mathbb{R}^d$ coïncide bien avec la tribu produit des tribus boréliennes de $\mathbb{R}^p$ et de $\mathbb{R}^q$.

Chapitre 5

Espaces L^p

Ce chapitre présente la construction et certaines propriétés des espaces fonctionnels L^p qui jouent un rôle fondamental en analyse et dans les applications. Nous construisons ici les espaces $L^p(\Omega)$ où Ω est un ouvert (ou plus généralement un borélien) de $\mathbb{R}^d$ mais la même démarche permet de définir les espaces $L^p(X, \mathcal{T}, \mu)$ où $(X, \mathcal{T}, \mu)$ un espace mesuré quelconque. Par la suite, on note λ la mesure de Lebesgue.

5.1 Fonctions égales presque partout

Proposition 5.1. *Soit Ω un ouvert (ou plus généralement un borélien) de $\mathbb{R}^d$. La relation "$f = g$ presque partout", qui se traduit par "l'ensemble $\{x \in \Omega,\ f(x) \neq g(x)\}$ est négligeable (pour la mesure de Lebesgue)", est une relation d'équivalence sur l'ensemble des fonctions mesurables.*

Preuve. Notons $\sim$ la relation d'"égalité presque partout". Montrons d'abord que $\sim$ est une relation d'équivalence. Considérons pour cela trois fonctions mesurables f, g et h.

- Il est clair que $f \sim f$ (la relation est réflexive) ;
- il est également évident que si $f \sim g$ alors $g \sim f$ (la relation est symétrique) ;
- supposons maintenant que $f \sim g$ et $g \sim h$. Soit A et B deux ensembles négligeables tels que $\{x \in \Omega,\ f(x) \neq g(x)\} \subset A$ et $\{x \in \Omega,\ g(x) \neq h(x)\} \subset B$. On voit tout de suite que
$$\{x \in \Omega,\ f(x) \neq h(x)\} \subset \{x \in \Omega,\ f(x) \neq g(x)\} \cup \{x \in \Omega,\ g(x) \neq h(x)\} \subset A \cup B.$$
Pour montrer la transitivité de la relation $\sim$, il suffit de montrer que $A \cup B$ est négligeable. Cela découle de l'inégalité
$$\lambda(A \cup B) \leq \lambda(A) + \lambda(B) = 0.$$

□

La relation d'équivalence d'"égalité presque partout" est compatible avec l'intégration :

Proposition 5.2. *Soit Ω un ouvert (ou plus généralement un borélien) de $\mathbb{R}^d$. Soit f et g deux fonctions intégrables égales presque partout. Alors $\displaystyle\int_\Omega f = \int_\Omega g$.*

Preuve. Soit A négligeable tel que $\{x \in \Omega,\ f(x) \neq g(x)\} \subset A$. On a
$$\left| \int_\Omega f - \int_\Omega g \right| \leq \int_\Omega |f - g| = \int_A |f - g| = 0,$$
cette dernière égalité étant une conséquence de la proposition 4.32. □

5.2 Espace L^1

Notons $L^1(\Omega)$ le quotient (l'ensemble des classes d'équivalence) de l'ensemble $\mathcal{L}^1(\Omega)$ par la relation d'équivalence d'"égalité presque partout". Soulignons deux points importants :

1. Si $\alpha \in \mathbb{R}$ et si f_1 et f_2 sont deux fonctions de $\mathcal{L}^1(\Omega)$ égales presque partout, alors αf_1 et αf_2 sont aussi égales presque partout ; si $f \in L^1(\Omega)$ et $\alpha \in \mathbb{R}$, on peut donc définir le produit $\alpha f \in L^1(\Omega)$: ce sera la classe d'équivalence du produit de α par n'importe quel représentant de f.
2. Soit maintenant f_1, f_2, g_1, g_2 dans $\mathcal{L}^1(\Omega)$ telles que $f_1 = f_2$ et $g_1 = g_2$ presque partout. On sait (cf. remarque 4.1) qu'on n'a pas le droit *a priori* de définir $f_1 + g_1$ ou $f_2 + g_2$. Mais on peut en revanche définir la somme $f + g$ où f désigne la classe de f_1 (et de f_2) et où g désigne la classe de g_1 (et de g_2). En effet, comme f_1 (resp. g_1) est intégrable, on sait que les points où f_1 (resp. g_1) est infinie est de mesure nulle ; notons-le A_1 (resp. B_1). Soit $\tilde{f}_1$ (resp. $\tilde{g}_1$) la fonction égale à f_1 (resp. à g_1) en tout point de $\Omega \setminus A_1$ (resp. $\Omega \setminus B_1$) et égale à 0 sur A_1 (resp. sur B_1). Comme les fonctions $\tilde{f}_1$ et $\tilde{g}_1$ sont finies sur tout Ω, il n'y a pas d'obstacle à définir leur somme. On peut vérifier que si l'on fait de même avec f_2 et g_2, on obtiendra $\tilde{f}_1 + \tilde{g}_1 = \tilde{f}_2 + \tilde{g}_2$ presque partout. La classe de fonctions correspondante sera par définition la somme $f + g$ dans $L^1(\Omega)$.

On vérifie facilement que les opérations de multiplication par un scalaire et d'addition définies ci-dessus confèrent à $L^1(\Omega)$ une structure d'espace vectoriel.

Définition 5.3. *On note $L^1(\Omega)$ l'espace vectoriel obtenu en quotientant l'ensemble $\mathcal{L}^1(\Omega)$ par la relation d'équivalence d'"égalité presque partout". Pour $f \in L^1(\Omega)$, la notation $\int_\Omega f$ désigne l'intégrale de l'une quelconque des fonctions appartenant à la classe de f.*

C'est la proposition 5.2 qui garantit que pour $f \in L^1(\Omega)$, $\int_\Omega f$ est défini de manière unique. Dans toute la suite, on désignera par la même notation (f, g, ...) une fonction et la classe d'équivalence à laquelle cette fonction appartient, et on parlera (par abus de langage) de fonctions pour désigner les éléments de $L^1(\Omega)$.

La plus-value fondamentale qu'il y a à quotienter l'ensemble $\mathcal{L}^1$ par la relation d'équivalence d'"égalité presque partout" est que l'espace quotient L^1 que l'on obtient ainsi est un espace *de Banach* ; on dispose donc pour travailler dans L^1 des outils "géométriques" établis au chapitre 2.

Théorème 5.4. *Muni de la norme définie par*

$$\|f\|_{L^1} = \int_\Omega |f|,$$

l'espace vectoriel $L^1(\Omega)$ est un espace de Banach.

Preuve. Vérifions d'abord que $\|f\|_{L^1}$ est effectivement une norme sur L^1. Les points *(ii)* et *(iii)* de la définition 1.1 sont faciles à vérifier. Reste le point *(i)*. Pour le montrer, considérons un représentant de f, encore noté f, et définissons pour tout $k \in \mathbb{N}^*$ l'ensemble mesurable $A_k = \{x \in \Omega,\ |f(x)| \geq 1/k\}$; la fonction étagée $e_k = 1/k \times \chi_{A_k}$ vérifie $0 \leq e_k \leq |f|$. On a donc

$$0 \leq \frac{\lambda(A_k)}{k} = \int_\Omega e_k \leq \int_\Omega |f| = 0.$$

Donc $\lambda(A_k) = 0$. Il en résulte que l'ensemble

$$\{x \in \Omega, \quad f(x) \neq 0\} = \bigcup_{k \in \mathbb{N}^*} A_k$$

est un ensemble négligeable (c'est une union dénombrable d'ensembles négligeables). D'où $f = 0$ dans L^1.

Pour montrer la complétude de L^1, on va utiliser la caractérisation des espaces de Banach fournie par le théorème 2.16. Soit u_n le terme général d'une série normalement convergente ($\sum_{n\in\mathbb{N}} \|u_n\|_{L^1} < +\infty$) ; montrons que la série $\sum_{n\in\mathbb{N}} u_n$ converge dans L^1. Posons $U_N = \sum_{n=0}^{N} |u_n|$ et $U_\infty = \sum_{n=0}^{+\infty} |u_n|$. La suite $(U_N)_{N\in\mathbb{N}}$ est une suite croissante de fonctions positives. Il en résulte (cf. théorème 4.39) que

$$\int_\Omega U_\infty = \int_\Omega \lim_{N\to+\infty} U_N = \lim_{N\to+\infty} \int_\Omega U_N = \lim_{N\to+\infty} \sum_{n=0}^{N} \|u_n\|_{L^1} = \sum_{n=0}^{+\infty} \|u_n\|_{L^1} < +\infty.$$

La fonction U_∞ est donc intégrable ; on en déduit qu'elle est finie presque partout (cf. théorème 4.33). Presque partout, la série de terme général $u_n(x)$ est donc absolument convergente donc convergente ; notons alors $v(x)$ sa limite et prolongeons la fonction v par 0 sur l'ensemble négligeable sur lequel U_∞ est infinie. Posons enfin $w_N = \left| v - \sum_{n=0}^{N} u_n \right|$. En réunissant les résultats que l'on a établis, on voit que

$$\begin{cases} w_N(x) \underset{N\to+\infty}{\longrightarrow} 0 \quad \text{presque partout,} \\ \forall N \in \mathbb{N}, \quad |w_N(x)| \leq 2\, U_\infty \quad \text{sur } \Omega. \end{cases}$$

En appliquant le théorème de convergence dominée 4.42, on obtient $v \in L^1(\Omega)$ et

$$\left\| v - \sum_{n=0}^{N} u_n \right\|_{L^1} = \int_\Omega \left| v - \sum_{n=0}^{N} u_n \right| \underset{N\to+\infty}{\longrightarrow} 0,$$

ce qui termine la preuve. □

Pour identifier concrètement la limite d'une suite de fonctions $(f_n)_{n\in\mathbb{N}}$ dont on sait qu'elle converge en norme L^1, on pourra utiliser le résultat suivant :

Théorème 5.5. *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de $L^1(\Omega)$ qui converge dans $L^1(\Omega)$ vers une fonction f. La suite $(f_n)_{n\in\mathbb{N}}$ admet alors une sous-suite $(f_{n_k})_{k\in\mathbb{N}}$ qui converge simplement vers f presque partout dans Ω.*

Preuve. Construisons une sous-suite $(f_{n_k})_{k\in\mathbb{N}}$ de la suite $(f_n)_{n\in\mathbb{N}}$ telle que

$$\forall k \in \mathbb{N}, \quad \left\| f_{n_{k+1}} - f_{n_k} \right\|_{L^1} \leq \frac{1}{2^k},$$

et posons

$$g_k(x) = \sum_{j=1}^{k} \left| f_{n_{j+1}}(x) - f_{n_j}(x) \right|.$$

On voit que la suite $(g_k)_{k\in\mathbb{N}}$ est mesurable, positive et croissante et qu'il découle donc du théorème 4.39 que sa limite simple $g(x) = \sum_{j=1}^{+\infty} \left| f_{n_{j+1}}(x) - f_{n_j}(x) \right|$ vérifie

$$\int_\Omega g = \lim_{k\to+\infty} \int_\Omega g_k = \sum_{j=1}^{+\infty} \left\| f_{n_{k+1}} - f_{n_k} \right\|_{L^1} \leq \sum_{j=1}^{+\infty} \frac{1}{2^j} = 1.$$

On déduit donc de la proposition 4.33 que la fonction g est finie sur un ensemble $A \subset \Omega$ de complémentaire négligeable. On a alors

$$\forall x \in A, \quad \forall 0 \leq k \leq l, \quad |f_{n_l}(x) - f_{n_k}(x)| \leq g(x) - g_k(x),$$

ce qui implique

1. que pour tout $x \in A$ la suite $(f_{n_k}(x))_{k \in \mathbb{N}}$, qui est de Cauchy, converge vers $h(x)$ dans $\mathbb{R}$,
2. que h, prolongée par 0 dans $\Omega \setminus A$, est mesurable (c'est une limite simple de fonctions mesurables),
3. que pour tout $x \in A$, $|h(x) - f_{n_k}(x)| \leq g(x)$.

Il résulte donc du théorème de convergence dominée que la sous-suite $(f_{n_k})_{k \in \mathbb{N}}$ converge en norme L^1 vers h, et donc que $h = f$. Finalement, la sous-suite $(f_{n_k})_{k \in \mathbb{N}}$ converge donc presque partout vers f. □

5.3 Espace L^2

En réalisant cette même opération de "quotientage" non plus sur l'ensemble des fonctions intégrables mais sur l'ensemble des fonctions de *carré* intégrable, on obtient un espace vectoriel qu'on peut munir d'une structure hilbertienne.

Considérons l'ensemble

$$\mathcal{L}^2(\Omega) = \left\{ f \text{ mesurable}, \quad \int_\Omega |f|^2 < +\infty \right\},$$

des fonctions de carré intégrable.

Définition 5.6. *On note $L^2(\Omega)$ l'espace vectoriel obtenu en quotientant l'ensemble $\mathcal{L}^2(\Omega)$ par la relation d'équivalence d'"égalité presque partout".*

En utilisant l'inégalité

$$|fg| \leq \frac{1}{2} \left(|f|^2 + |g|^2 \right)$$

on voit que si f et g sont dans L^2, alors le produit fg est dans L^1 ; ceci permet de définir sur $L^2(\Omega)$ le produit scalaire

$$(f, g)_{L^2} = \int_\Omega fg$$

(la bilinéarité, la symétrie et la positivité sont claires ; enfin, $(f, f)_{L^2} = 0$ si et seulement si $|f|^2 = 0$ presque partout et donc si et seulement si $f = 0$ dans L^2). Qui plus est,

Théorème 5.7. *Muni du produit scalaire défini par*

$$(f, g)_{L^2} = \int_\Omega fg$$

l'espace $L^2(\Omega)$ est un espace de Hilbert.

Preuve. La démonstation de la complétude de l'espace L^2 s'apparente à celle de la complétude de l'espace L^1. □

Terminons cette section en remarquant que l'inégalité de Cauchy-Schwarz (3.2) prend dans l'espace L^2 la forme suivante : pour tout f et g dans $L^2(\Omega)$

$$\left| \int_\Omega fg \right| \leq \left(\int_\Omega |f|^2 \right)^{1/2} \left(\int_\Omega |g|^2 \right)^{1/2}. \tag{5.1}$$

5.4 Espace L^∞

Définition 5.8. *On dit qu'une fonction mesurable f est* essentiellement bornée *s'il existe un réel positif M tel que*

$$\lambda\left(\{x \in \Omega,\ |f(x)| \geq M\}\right) = 0.$$

On note $\mathcal{L}^\infty(\Omega)$ l'ensemble des fonctions essentiellement bornées.

Définition 5.9. *On note $L^\infty(\Omega)$ l'espace vectoriel obtenu en quotientant l'ensemble $\mathcal{L}^\infty(\Omega)$ par la relation d'équivalence d'"égalité presque partout". On désigne par $\|\cdot\|_{L^\infty} : L^\infty(\Omega) \to \mathbb{R}_+$ l'application définie par*

$$\|f\|_{L^\infty} = \inf\left\{M \geq 0, \quad \lambda\left(\{x \in \Omega,\ |f(x)| \geq M\}\right) = 0\right\}. \tag{5.2}$$

Nous allons maintenant vérifier que l'application $\|\cdot\|_{L^\infty}$ définie ci-dessus définit bien une norme sur l'espace vectoriel $L^\infty(\Omega)$. Mieux, cette norme confère à $L^\infty(\Omega)$ une structure d'espace de Banach. Pour cela, commençons par le lemme suivant.

Lemme 5.10. *Pour tout $f \in L^\infty(\Omega)$, on a*

$$|f(x)| \leq \|f\|_{L^\infty} \quad \textit{presque partout dans } \Omega.$$

Preuve. Voir exercice 5.1. □

Une conséquence immédiate du lemme 5.10 est le fait que l'application $\|\cdot\|_{L^\infty}$ définit bien une norme sur $L^\infty(\Omega)$. Passons maintenant à la complétude de $L^\infty(\Omega)$.

Théorème 5.11. *Muni de la norme $\|\cdot\|_{L^\infty}$ définie par (5.2), l'espace vectoriel $L^\infty(\Omega)$ est un espace de Banach.*

Preuve. Voir exercice 5.2. □

5.5 Autres espaces L^p

Soit un réel p tel que $1 \leq p < +\infty$. On définit l'ensemble $\mathcal{L}^p(\Omega)$ par

$$\mathcal{L}^p(\Omega) = \left\{f \text{ mesurables}, \quad \int_\Omega |f|^p < +\infty\right\}.$$

Définition 5.12. *On note $L^p(\Omega)$ l'espace vectoriel obtenu en quotientant $\mathcal{L}^p(\Omega)$ par la relation d'équivalence d'"égalité presque partout". On désigne par $\|\cdot\|_{L^p} : L^p(\Omega) \to \mathbb{R}_+$ l'application définie par*

$$\|f\|_{L^p} = \left(\int_\Omega |f|^p\right)^{\frac{1}{p}}. \tag{5.3}$$

Remarque 5.13. Le fait que $L^p(\Omega)$ est un espace vectoriel se prouve en observant que

$$|f(x) + g(x)|^p \leq 2^p \max\{|f(x)|^p, |g(x)|^p\} \leq 2^p(|f(x)|^p + |g(x)|^p)$$

si bien que $f \in L^p(\Omega)$ et $g \in L^p(\Omega)$ implique que $f + g \in L^p(\Omega)$.

Définition 5.14. *Soit un réel p avec $1 \leq p \leq +\infty$. On appelle réel conjugué de p le réel p' défini par la relation*

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

En particulier, on a $1 \leq p' \leq +\infty$, $p' = +\infty$ si $p = 1$, $p' = 1$ si $p = +\infty$ si $p' = 1$ et $p' = 2$ si $p = 2$. De plus, on a $(p')' = p$.

Théorème 5.15. *(Inégalité de Hölder). Si $f \in L^p(\Omega)$ et $g \in L^{p'}(\Omega)$, alors le produit fg est dans $L^1(\Omega)$ et*

$$\|fg\|_{L^1} \leq \|f\|_{L^p} \, \|g\|_{L^{p'}}. \tag{5.4}$$

Preuve. Pour $p = 1$ et pour $p = +\infty$, l'inégalité (5.4) prend la forme suivante : pour $f \in L^1(\Omega)$ et $g \in L^\infty(\Omega)$,

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

ce qui est une conséquence immédiate du lemme 5.10. Considérons maintenant le cas $1 < p < +\infty$, et soit p' le réel conjugué de p. Soit $f \in L^p(\Omega)$ et $g \in L^{p'}(\Omega)$. Si $f = 0$ ou $g = 0$, l'inégalité de Hölder est évidemment vérifiée. Dans le cas contraire, on utilise l'inégalité de Young (1.6) avec $u = \dfrac{|f(x)|}{\|f\|_{L^p}}$ et $u = \dfrac{|g(x)|}{\|g\|_{L^{p'}}}$ pour établir que

$$\frac{|f(x)|}{\|f\|_{L^p}} \, \frac{|g(x)|}{\|g\|_{L^{p'}}} \leq \frac{1}{p} \left(\frac{|f(x)|}{\|f\|_{L^p}} \right)^p + \frac{1}{p'} \left(\frac{|g(x)|}{\|g\|_{L^{p'}}} \right)^{p'} \qquad \text{presque partout.}$$

En intégrant sur Ω, on obtient

$$\frac{1}{\|f\|_{L^p} \, \|g\|_{L^{p'}}} \int_\Omega |f| \, |g| \leq \frac{1}{p} + \frac{1}{p'} = 1.$$

D'où l'inégalité (5.4). □

Notons que l'inégalité de Cauchy-Schwarz (5.1) est un cas particulier de l'inégalité de Hölder (correspondant à $p = p' = 2$). Une conséquence importante de l'inégalité de Hölder est l'inégalité de Minkowski dont découle le fait que l'application $\|\cdot\|_{L^p}$ définie par (5.3) est une norme sur $L^p(\Omega)$.

Proposition 5.16. *Soit f et g dans $L^p(\Omega)$. Alors, on a*

$$\|f + g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p}.$$

Preuve. Voir exercice 5.3. □

Théorème 5.17. *Equipé de la norme $\|\cdot\|_{L^p}$, $L^p(\Omega)$ est un espace de Banach.*

Preuve. Procéder comme dans la preuve de la complétude de $L^1(\Omega)$. □

L'inégalité de Hölder a également plusieurs conséquences importantes sur les relations d'inclusion entre les divers espaces L^p. En voici deux.

Proposition 5.18. *Soit $\Omega \subset \mathbb{R}^d$ tel que $\lambda(\Omega) < +\infty$ et $1 \leq p < q \leq +\infty$, alors*

$$L^q(\Omega) \subset L^p(\Omega).$$

Preuve. On applique l'inégalité de Hölder avec l'exposant $\alpha = \frac{q}{p}$ (qui est bien compris entre 1 et $+\infty$). Le réel conjugué α' vaut $\frac{q}{q-p}$. Il vient

$$\|f\|_{L^p}^p = \int_\Omega |f|^p \leq \left(\int_\Omega |f|^{p\alpha}\right)^{\frac{1}{\alpha}} \left(\int_\Omega 1^{\alpha'}\right)^{\frac{1}{\alpha'}} = \|f\|_{L^q}^p \lambda(\Omega)^{\frac{q-p}{q}}.$$

D'où

$$\|f\|_{L^p} \leq \|f\|_{L^q} \lambda(\Omega)^{\frac{q-p}{pq}}.$$

□

Proposition 5.19. *(Interpolation entre espaces L^p). Si $f \in L^p(\Omega) \cap L^q(\Omega)$ pour $1 \leq p \leq q \leq +\infty$, alors $f \in L^r(\Omega)$ pour tout $p \leq r \leq q$ et on a l'inégalité d'interpolation*

$$\|f\|_{L^r} \leq \|f\|_{L^p}^\alpha \|f\|_{L^q}^{1-\alpha}, \tag{5.5}$$

où $\alpha \in]0,1[$ est le réel tel que $\frac{1}{r} = \frac{\alpha}{p} + \frac{1-\alpha}{q}$.

Preuve. Voir exercice 5.4. □

La proposition suivante affine l'inégalité triangulaire pour la norme $\|\cdot\|_{L^p}$.

Proposition 5.20. *(Inégalité de Hanner). Soit $1 \leq p \leq +\infty$ et soit deux fonctions f et g dans $L^p(\Omega)$. Si $1 \leq p \leq 2$, on a*

$$\|f+g\|_{L^p}^p + \|f-g\|_{L^p}^p \geq (\|f\|_{L^p} + \|g\|_{L^p})^p + \big|\|f\|_{L^p} - \|g\|_{L^p}\big|^p, \tag{5.6}$$

alors que cette inégalité est renversée pour $2 \leq p < +\infty$.

Preuve. Voir exercice 5.11. □

Lorsque $p = 2$, l'inégalité (5.6) est en fait une égalité qui porte le nom *d'identité du parallélogramme* ; elle s'écrit sous la forme

$$\|f\|_{L^2}^2 + \|g\|_{L^2}^2 = \frac{1}{2}\|f+g\|_{L^2}^2 + \frac{1}{2}\|f-g\|_{L^2}^2, \tag{5.7}$$

ce qui se démontre simplement en développant le membre de droite. Par ailleurs, en remplaçant f par $f+g$ et g par $f-g$ dans (5.6), il vient

$$(\|f+g\|_{L^p} + \|f-g\|_{L^p})^p + \big|\|f+g\|_{L^p} - \|f-g\|_{L^p}\big|^p \leq 2^p(\|f\|_{L^p}^p + \|g\|_{L^p}^p) \tag{5.8}$$

pour $p < 2$, l'inégalité étant renversée pour $p > 2$.

Nous admettrons le résultat suivant (voir référence [Brézis 83], pages 57 et 61) :

Théorème 5.21. *Soit $1 < p < +\infty$ et p' le réel conjugué de p. L'espace vectoriel $L^{p'}(\Omega)$ s'identifie au dual topologique de l'espace $L^p(\Omega)$. Plus précisément, pour toute forme linéaire continue φ sur $L^p(\Omega)$, il existe un unique $u \in L^{p'}(\Omega)$ tel que*

$$\forall v \in L^p(\Omega), \qquad \langle \varphi, v \rangle_{(L^p)', L^p} = \int_\Omega u\, v.$$

De plus, $\|u\|_{L^{p'}} = \|\varphi\|_{(L^p)'}$.

Remarque 5.22. On peut également montrer que $L^\infty(\Omega)$ s'identifie au dual topologique de $L^1(\Omega)$. En revanche, $L^1(\Omega)$ <u>ne s'identifie pas</u> au dual topologique de $L^\infty(\Omega)$.

5.5.1 Espaces L^p_{loc}

Définition 5.23.

1. *On dit qu'une fonction mesurable f est* localement intégrable *sur Ω si $f\chi_K \in L^1(\Omega)$ pour tout compact K inclus dans Ω. On note $L^1_{loc}(\Omega)$ l'espace vectoriel des fonctions localement intégrables.*
2. *On note $L^p_{loc}(\Omega)$ l'espace vectoriel des fonctions mesurables f telles que $f\chi_K \in L^p(\Omega)$ pour tout compact K inclus dans Ω.*
3. *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de $L^p_{loc}(\Omega)$ et $f \in L^p_{loc}(\Omega)$. On dit que (f_n) converge vers f dans $L^p_{loc}(\Omega)$ si pour tout compact K inclus dans Ω, la suite $(f_n|_K)$ converge vers $f|_K$ dans $L^p(K)$.*

Proposition 5.24. *Si $1 \le p < q \le +\infty$, alors*

$$L^q_{loc}(\Omega) \subset L^p_{loc}(\Omega).$$

Preuve. Conséquence de la proposition 5.18. □

5.6 Convolution dans $L^1(\mathbb{R}^d)$

Définition - Théorème 5.25. *Soit f et g dans $L^1(\mathbb{R}^d)$. La fonction notée $f \star g$ définie preque partout sur $\mathbb{R}^d$ par*

$$(f \star g)(x) = \int_{\mathbb{R}^d} f(x-y) g(y)\, dy$$

est appelée la convolée *ou encore le* produit de convolution *de f et g. La fonction $f \star g$ est dans $L^1(\mathbb{R}^d)$ et on a*

$$\|f \star g\|_{L^1} \le \|f\|_{L^1}\, \|g\|_{L^1}. \tag{5.9}$$

Preuve. Laissée en exercice (des indications sont fournies dans l'énoncé de l'exercice 5.14). □

Proposition 5.26. *Soit f, g et h dans $L^1(\mathbb{R}^d)$; soit α dans $\mathbb{R}$. On a les propriétés suivantes*

$$f \star g = g \star f, \quad f \star (g+h) = (f \star g) + (f \star h), \quad f \star (\alpha g) = \alpha\, (f \star g), \quad (f \star g) \star h = f \star (g \star h).$$

Muni du produit de convolution, l'espace $L^1(\mathbb{R}^d)$ est donc une algèbre commutative.

Preuve. La commutativité résulte de l'invariance par translation de la mesure de Lebesgue ; les deux points suivants découlent de la linéarité de l'intégrale ; l'associativité se montre en utilisant le théorème de Fubini et l'invariance par translation de la mesure de Lebesgue. □

L'opération de convolution, qui peut s'étendre à une certaine classe de distributions, intervient naturellement dans les phénomènes qui présentent une invariance par translation (dans le temps ou dans l'espace par exemple) ; elle joue un rôle très important en analyse et dans les applications, notamment en traitement du signal.

Par ailleurs, un des intérêts pratiques de la convolution est qu'elle permet de "régulariser" une fonction ou une distribution peu régulière. Ainsi le théorème 6.15 (de densité des fonctions C^∞ à support compact dans l'espace L^p pour tout $1 \leq p < +\infty$) qui joue un rôle important dans la théorie des distributions (cf. chapitre 6 et suivants) se prouve par une technique de régularisation par convolution. On montre ainsi par exemple (mais c'est un peu technique) que si $f \in L^1(\mathbb{R}^d)$ et si $\chi \in C^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ est une fonction dont l'intégrale sur $\mathbb{R}^d$ vaut 1, la suite $(f_k)_{k\in\mathbb{N}^*}$ définie par $f_k = f \star \chi_k$ où

$$\forall k \in \mathbb{N}^*, \quad \forall x \in \mathbb{R}^d, \qquad \chi_k(x) = k^d \chi(kx).$$

est dans $C^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ et tend vers f en norme L^1 (ce qui montre la densité de $C^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ dans $L^1(\mathbb{R}^d)$ et fournit un procédé de construction d'une suite de $C^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ qui converge en norme L^1 vers une fonction donnée de $L^1(\mathbb{R}^d)$).

5.7 Exercices

▷ **5.1.** Prouver le lemme 5.10. *Indication : introduire une suite M_n qui tend vers $\|f\|_{L^\infty}$ par valeurs supérieures et considérer l'ensemble de mesure nulle A_n sur lequel $|f(x)| > M_n$.*

▷ **5.2.** Prouver le théorème 5.11. *Indication : considérer une suite de Cauchy $(f_n)_{n\in\mathbb{N}}$ dans $L^\infty(\Omega)$ et pour tout entier $k \geq 1$, considérer l'entier n_k et l'ensemble de mesure nulle A_k tels que*

$$\forall x \in \Omega \setminus A_k, \quad \forall m, n \geq n_k, \quad |f_m(x) - f_n(x)| \leq \frac{1}{k}.$$

▷ **5.3.** Prouver la proposition 5.16. *Indication : observer que*

$$|f(x) + g(x)|^p \leq |f(x)|\,|f(x) + g(x)|^{p-1} + |g(x)|\,|f(x) + g(x)|^{p-1}$$

et que $|f + g|^{p-1} \in L^{p'}(\Omega)$ où p' est le réel conjugué de p, puis utiliser l'inégalité de Hölder.

▷ **5.4.** Prouver la proposition 5.19. *Indication : appliquer l'inégalité de Hölder aux fonctions $|f|^{r\alpha}$ et $|f|^{r(1-\alpha)}$ avec l'exposant $\gamma = \frac{p}{r\alpha}$.*

▷ **5.5.** Pourquoi ne considère-t-on pas des espaces $L^p(\Omega)$ pour $p < 1$? *Indication : penser à l'inégalité triangulaire.*

▷ **5.6.** Soit k réels $(p_1, \dots, p_k)$ tous compris entre 1 et $+\infty$. On pose $p = (\sum_{i=1}^k \frac{1}{p_i})^{-1}$ et on suppose que $p \geq 1$. Soit k fonctions $(f_1, \dots, f_k)$ telles que pour tout $1 \leq i \leq k$, $f_i \in L^{p_i}(\Omega)$. Montrer que le produit $f = f_1 \dots f_k$ est dans $L^p(\Omega)$ et que

$$\|f\|_{L^p} \leq \|f_1\|_{L^{p_1}} \dots \|f_k\|_{L^{p_k}}.$$

Indication : utiliser l'inégalité de Hölder.

▷ **5.7.** Montrer que pour tout $1 \leq p < q \leq +\infty$,

$$L^\infty(\Omega) \cap L^p(\Omega) \subset L^\infty(\Omega) \cap L^q(\Omega).$$

Plus généralement, montrer que pour tout $1 \leq p < q < r \leq +\infty$,

$$L^r(\Omega) \cap L^p(\Omega) \subset L^r(\Omega) \cap L^q(\Omega).$$

▷ **5.8.** Soit $f \in L^\infty(\Omega)$. On suppose qu'il existe un réel $1 \leq q \leq +\infty$ tel que $f \in L^q(\Omega)$. Montrer que

$$\lim_{p \to +\infty, p \geq q} \|f\|_{L^p} = \|f\|_{L^\infty}.$$

Note : cette propriété justifie la notation $\|f\|_{L^\infty}$.

▷ **5.9.** On considère la fonction $f(x) = (x \log^2(1/x))^{-1}$ sur $\Omega =]0, \frac{1}{2}[$. Vérifier que $f \in L^1(\Omega)$ mais que f n'est pas dans $\bigcup_{p>1} L^p(\Omega)$.

▷ **5.10.** Soit f et g dans $L^p(\Omega)$ avec $1 < p < +\infty$. L'objet de cette exercice est d'étudier la differentiabilité de la fonction

$$N : \mathbb{R} \ni t \longmapsto N(t) = \|f + tg\|_{L^p}^p \in \mathbb{R}.$$

– Soit deux réels a et b. Montrer que

$$\lim_{t \to 0} \frac{1}{t}(|a + tb|^p - |a|^p) = p|a|^{p-2}ab.$$

(Noter que pour $p > 1$, $|a|^{p-2}a$ peut être prolongé par continuité en 0 pour $a = 0$.)

– Montrer que

$$\frac{d}{dt} N|_{t=0} = \int_\Omega p|f|^{p-2} fg.$$

Indication : utiliser le théorème de convergence dominée pour justifier la dérivation sous le signe somme. On montrera au préalable que pour des réels a et b,

$$|a|^p - |a - b|^p \leq \frac{1}{t}(|a + tb|^p - |a|^p) \leq |a + b|^p - |a|^p,$$

ce qui se prouve en utilisant la convexité de la fonction $s \mapsto |s|^p$ sur $\mathbb{R}_+$.

▷ **5.11.** L'objet de cet exercice est de prouver la proposition 5.20. On suit ici la preuve proposée par Ball, Carlen et Lieb. Sans perte de généralité, on peut supposer que $\|f\|_{L^p} = 1$ et $\|g\|_{L^p} \leq 1$.

– Pour $r \in]0, 1[$, on pose

$$\alpha(r) = (1 + r)^{p-1} + (1 - r)^{p-1},$$
$$\beta(r) = [(1 + r)^{p-1} - (1 - r)^{p-1}]r^{1-p}.$$

Soit $R \in]0, 1[$. Montrer que la fonction $\Phi_R(r) = \alpha(r) + \beta(r)R^p$ est maximale (resp., minimale) en $r = R$ pour $p < 2$ (resp., pour $p > 2$).

– En déduire que pour deux réels positifs a et b,

$$\alpha(r)|a|^p + \beta(r)|b|^p \leq |a + b|^p + |a - b|^p$$

pour $p < 2$, l'inégalité étant renversée pour $p > 2$. Montrer que ces ingéalités s'étendent au cas de réels a et b de signe quelconque.

– Conclure.

▷ **5.12.** (Inégalité de Clarkson et uniforme convexité de la boule unité de L^p).

– Soit $2 \leq p < +\infty$. Montrer que pour tout $f, g \in L^p(\Omega)$,

$$\|\tfrac{1}{2}(f+g)\|_{L^p}^p + \|\tfrac{1}{2}(f-g)\|_{L^p}^p \leq \frac{1}{2}(\|f\|_{L^p}^p + \|g\|_{L^p}^p).$$

Indication : utiliser la convexité de la fonction $(s^2+1)^{p/2} - s^p - 1$ *sur* $\mathbb{R}_+$.

– Soit $\epsilon > 0$ fixé. Montrer que si f et g sont dans la boule unité de L^p et si $\|f-g\|_{L^p} > \epsilon$, alors $\|\tfrac{1}{2}(f+g)\|_{L^p} < 1 - \delta(\epsilon)$ avec $\delta(\epsilon) = 1 - (1-(\epsilon/2)^p)^{1/p}$.

– Interpréter géométriquement la propriété ci-dessus. Peut-elle être étendue au cas $p = +\infty$?

▷ **5.13.** Soit $1 < p < +\infty$. L'objet de cet exercice est de prouver le résultat suivant : toute suite $(f_n)_{n\in\mathbb{N}}$ de $L^p(\Omega)$ telle que pour tout $u \in L^{p'}(\Omega)$, la suite $(\int_\Omega u f_n)_{n\in\mathbb{N}}$ est bornée, est elle-même bornée. On raisonne par l'absurde en supposant que la suite $(f_n)_{n\in\mathbb{N}}$ n'est pas bornée.

– Expliquer pourquoi modulo l'extraction d'une sous-suite et une renormalisation, on peut supposer que $\|f_n\|_{L^p} = 4^n$.

– On pose $u_n(x) = |f_n(x)|^{p-2} f_n(x) / \|f_n\|_{L^p}^{p-1}$. Vérifier que $\|u_n\|_{L^{p'}} = 1$ pour tout $n \in \mathbb{N}$.

– On pose $\epsilon_0 = 1$, et pour $n \geq 1$, on construit la suite $(\epsilon_n)_{n\in\mathbb{N}}$ par récurrence en choisissant $\epsilon_n \in \{-1, +1\}$ de sorte que $\epsilon_n \int_\Omega u_n f_n$ ait le même signe que $\sum_{k=0}^{n-1} 3^{-k} \epsilon_k \int_\Omega u_k f_n$. Montrer que la série $\sum_{n\in\mathbb{N}} 3^{-n} \epsilon_n u_n$ converge dans $L^{p'}(\Omega)$ et que pour tout $n \in \mathbb{N}$,

$$\left| \sum_{k=0}^{n} 3^{-n} \epsilon_n \int_\Omega u_n f_n \right| \geq \left(\frac{4}{3}\right)^n.$$

– Montrer que $|\int_\Omega u f_n| \geq \frac{1}{2}(\frac{4}{3})^n$. Conclure.

▷ **5.14.** Prouver le théorème 5.25. *Indication : montrer dans un premier temps que*

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |f(x-y)|\,|g(y)|\,dx\,dy = \|f\|_{L^1}\,\|g\|_{L^1}.$$

▷ **5.15.** Soit $f \in L^1(\mathbb{R}^d)$ et $g \in L^p(\mathbb{R}^d)$ avec $1 \leq p \leq +\infty$.

– Montrer que pour presque tout $x \in \mathbb{R}^d$, la fonction $y \mapsto f(x-y)g(y)$ est intégrable sur $\mathbb{R}^d$.

– En notant $f \star g(x)$ la valeur de l'intégrale de la fonctin ci-dessus, montrer que $f \star g \in L^p(\mathbb{R}^d)$ et que

$$\|f \star g\|_{L^p} \leq \|f\|_{L^1}\,\|g\|_{L^p}.$$

Indication : utiliser l'inégalité de Hölder.

Chapitre 6

Notion de distribution

Les espaces fonctionnels que nous avons construits jusqu'à présent, à savoir les espaces C^k et les espaces L^p, sont inadéquats pour analyser la plupart des problèmes mathématiques issus des sciences de l'ingénieur, en particulier ceux relevant de la théorie des EDP.

La théorie des distributions, introduite par Laurent Schwartz en 1946, généralise la notion de fonction et offre un cadre conceptuel unifié particulièrement élégant permettant de construire des espaces fonctionnels adaptés à ces problèmes. Qui plus est, les distributions se révèlent à l'usage plus souples à manipuler que les fonctions : toute distribution est dérivable et sa dérivée est encore une distribution, les critères de convergence sont faciles à vérifier, la dérivée de la limite est toujours égale à la limite des dérivées, ...

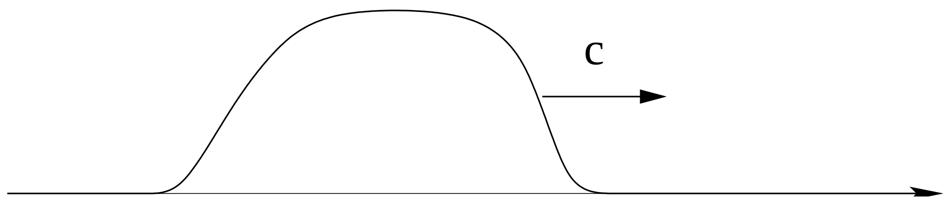
Dans tout ce chapitre, Ω désigne un ouvert (borné ou non) de $\mathbb{R}^d$ avec $d \geq 1$.

6.1 Insuffisances de la notion de fonction

Soit $f_{reg} : \mathbb{R} \to \mathbb{R}$ une fonction régulière, disons de classe C^2. La fonction

$$\begin{array}{rccl} u_{reg} : & \mathbb{R} \times \mathbb{R}^+ & \to & \mathbb{R} \\ & (x,t) & \mapsto & u_{reg}(x,t) = f_{reg}(x - ct) \end{array}$$

représente d'un point de vue physique une onde qui se déplace sans déformation à la célérité c dans le sens des x croissants



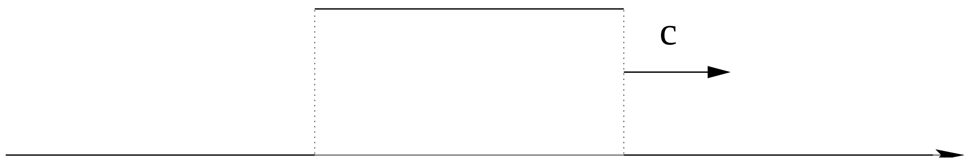
et vérifie l'équation (dite des ondes)

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0. \tag{6.1}$$

Soit maintenant $f_{sing} : \mathbb{R} \to \mathbb{R}$ une fonction plus singulière, par exemple le créneau $f_{sing}(x) = H(x) - H(x - L)$ où H désigne la fonction de Heaviside définie presque partout par

$$H(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 & \text{si } x > 0. \end{cases}$$

$u_{sing}(x,t) = f_{sing}(x-ct) = H(x-ct) - H(x-ct-L)$ représente toujours physiquement une onde de célérité c se déplaçant sans déformation dans le sens des x croissants. Pourtant on ne peut pas écrire que u_{sing} vérifie l'équation des ondes (6.1), tout au moins au sens usuel, parce que cette fonction manque de régularité ; en effet, si on s'en tient à la définition usuelle de la dérivée, H est dérivable en tout point de $\mathbb{R}^*$ (et $H'(x) = 0$ pour tout $x \neq 0$) mais n'est pas dérivable en 0.



Afin d'effectuer les premiers calculs de mécanique quantique, Dirac a introduit en 1926 la "fonction" qui porte son nom et qu'on note usuellement δ ou δ_0, "définie" par

$$\delta(x) = \begin{cases} 0 & \text{si } x \neq 0, \\ +\infty & \text{si } x = 0, \end{cases} \qquad \text{et} \qquad \int_{\mathbb{R}} \delta(x)\,dx = 1.$$

Formellement, $H' = \delta$ puisque $H(x) = \int_{-\infty}^{x} \delta(y)\,dy$ presque partout. Cette définition manque clairement de rigueur mathématique mais Dirac est allé encore plus loin : il a utilisé dans ses calculs la dérivée de la "fonction" δ, qu'on notera δ' sans essayer pour le moment de la définir plus précisément.

En appliquant ce formalisme à notre exemple, on voit que

$$\frac{\partial u_{sing}}{\partial t}(x,t) = -c\,\delta(x-ct) + c\delta(x-ct-L), \qquad \frac{\partial u_{sing}}{\partial x}(x,t) = \delta(x-ct) - \delta(x-ct-L)$$

$$\frac{\partial^2 u_{sing}}{\partial t^2}(x,t) = c^2\,\delta'(x-ct) - c^2\delta'(x-ct-L), \qquad \frac{\partial^2 u_{sing}}{\partial x^2}(x,t) = \delta'(x-ct) - \delta'(x-ct-L)$$

d'où il résulte que

$$\frac{1}{c^2}\frac{\partial^2 u_{sing}}{\partial t^2} - \frac{\partial^2 u_{sing}}{\partial x^2} = 0.$$

Si donc on relâche (provisoirement) l'exigence de rigueur mathématique, on peut "prouver" que l'onde u_{sing} vérifie bien l'équation des ondes (6.1) : c'est un bon point sur un plan théorique (à une réalité physique donnée correspond toujours la même équation) aussi bien que sur un plan pratique (il est bien utile de disposer d'équations pour faire de la physique ou de la mécanique). C'est la théorie des distributions qui permettra de donner un sens parfaitement rigoureux à la "fonction" de Dirac (qui n'est pas une fonction au sens usuel mais un objet plus général : une distribution) et à ses dérivées, et de justifier tous les calculs ci-dessus.

Il a fallu vingt ans aux mathématiciens pour justifier les calculs de Dirac, qui étaient tous justes ! Fort heureusement pour le développement des sciences, scientifiques et ingénieurs n'attendent pas de disposer d'outils mathématiques parfaitement polis pour effectuer leurs calculs (c'est d'ailleurs souvent pour justifier des calculs non rigoureux que de nouveaux outils mathématiques sont développés). Cela dit, disposer du bon cadre mathématique donne assurément une compréhension plus profonde des équations et fournit des gardes-fous au commun des utilisateurs (tout le monde ne s'appelle pas Dirac). En effet, tant qu'on reste dans un cadre mathématique bien défini, on sait exactement quelles sont les opérations licites qu'on peut effectuer avec les objets qu'on manipule. Dès qu'on sort d'un tel cadre, il n'y a plus que l'intuition qui permet de savoir ce qui est autorisé et ce qui ne l'est pas ; il faut donc redoubler de prudence car on travaille sans filet. Pour illustrer cela, effectuons un petit calcul sur la fonction de Heaviside et sa dérivée δ qui sonne juste mais qui est pourtant faux :

$$H^2 = H \quad \Rightarrow \quad 2H\delta = \delta \quad \Rightarrow \quad 2H^2\delta = H\delta \quad \Rightarrow \quad 2H\delta = H\delta \quad \Rightarrow \quad H\delta = 0.$$

Donc $\delta = 2H\delta = 0$. Mais $\delta \neq 0$ puisque $\displaystyle\int_{\mathbb{R}} \delta(x)\,dx = 1$.

6.2 Changeons de point de vue

On a vu ci-dessus que si l'on tentait de définir la "fonction" δ comme une fonction usuelle, c'est-à-dire en donnant ses valeurs ponctuelles, on aboutissait à une impasse. Cela nécessite de changer de point de vue.

Dorénavant, on ne va plus considérer les valeurs ponctuelles d'une fonction $f : \Omega \longrightarrow \mathbb{R}$ mais uniquement ses valeurs moyennes "contre" des fonctions test ϕ, autrement dit les quantités

$$\int_\Omega f\,\phi.$$

Comme ensemble des fonctions test, on prendra l'espace vectoriel $\mathcal{D}(\Omega)$ des fonctions C^∞ sur Ω à support compact (la définition précise sera donnée dans la définition 6.2). La suite justifiera ce choix un peu surprenant au premier abord.

On associe ainsi à la fonction f la forme linéaire sur $\mathcal{D}(\Omega)$ notée T_f et définie par

$$\begin{array}{rcl} T_f \ : \ \mathcal{D}(\Omega) & \to & \mathbb{R} \\ \phi & \mapsto & \langle T_f, \phi\rangle = \int_\Omega f\,\phi. \end{array}$$

Remarquons que T_f est définie pour toute fonction f de $L^1_{loc}(\Omega)$. Par ailleurs, si f et g sont deux fonctions de $L^1_{loc}(\Omega)$,

$$T_f = T_g \quad \Leftrightarrow \quad \forall \phi \in \mathcal{D}(\Omega), \quad \int_\Omega f\phi = \int_\Omega g\phi \quad \Leftrightarrow \quad f = g \quad \text{presque partout.}$$

Cette dernière équivalence n'est pas triviale : elle fait l'objet du théorème 7.1, qui sera démontré au chapitre suivant. Elle signifie que l'espace des fonctions test est suffisamment riche pour qu'on puisse distinguer deux fonctions f et g de L^1_{loc} en ne considérant que l'action de T_f et T_g sur les fonctions test.

Ce changement de point de vue se révèle extrêmement fécond :

1. Tout d'abord, il donne lieu à une généralisation naturelle de la notion de fonction. On définit une distribution sur Ω comme une forme linéaire continue (en un sens qui sera précisé plus loin) sur $\mathcal{D}(\Omega)$ et on note $\mathcal{D}'(\Omega)$ l'espace vectoriel des distributions sur Ω, autrement dit le dual topologique de $\mathcal{D}(\Omega)$. On vient de voir qu'il existe une injection naturelle $f \mapsto T_f$ de $L^1_{loc}(\Omega)$ dans $\mathcal{D}'(\Omega)$. Mais il existe des distributions qui ne peuvent pas être représentées par des fonctions (l'application $f \mapsto T_f$ est loin d'être surjective) :
   – il en est ainsi par exemple de la distribution δ_a (masse de Dirac),

   $$\begin{array}{rcl} \delta_a \ : \ \mathcal{D}(\mathbb{R}) & \to & \mathbb{R} \\ \phi & \mapsto & \langle \delta_a, \phi\rangle = \phi(a). \end{array}$$

   Il est clair en effet (le montrer en exercice) qu'on ne peut pas trouver de fonction dans $f \in L^1_{loc}(\mathbb{R})$ qui vérifie pour tout $\phi \in \mathcal{D}'(\Omega)$, $\int_\mathbb{R} f(x)\phi(x)\,dx = \phi(a)$;
   – il en va de même pour la distribution δ'_a définie par

   $$\begin{array}{rcl} \delta'_a \ : \ \mathcal{D}(\mathbb{R}) & \to & \mathbb{R} \\ \phi & \mapsto & \langle \delta'_a, \phi\rangle = -\phi'(a), \end{array}$$

   dont on verra dans un moment que c'est bien la "dérivée" (au sens des distributions) de δ_a.

2. En deuxième lieu, on peut utiliser la régularité des fonctions test pour dériver n'importe quelle distribution en faisant porter la dérivation sur la fonction test. C'est le concept de dérivée faible qui s'appuie sur la remarque suivante : pour $f \in C^1(\mathbb{R})$, on a

$$\forall \phi \in \mathcal{D}(\mathbb{R}), \qquad \int_{\mathbb{R}} \frac{df}{dx}\phi = -\int_{\mathbb{R}} f\frac{d\phi}{dx}.$$

Pour une fonction qui n'est pas dérivable, par exemple la fonction H de Heaviside, le membre de gauche n'est pas défini au sens usuel, mais le membre de droite l'est. On va donc définir la dérivée au sens des distributions dans $\mathcal{D}'(\mathbb{R})$ par

$$\forall \phi \in \mathcal{D}(\mathbb{R}), \qquad \langle \frac{dT}{dx}, \phi\rangle := -\langle T, \frac{d\phi}{dx}\rangle$$

On a ainsi

$$\forall \phi \in \mathcal{D}(\mathbb{R}), \qquad \langle \frac{dH}{dx}, \phi\rangle = -\langle H, \frac{d\phi}{dx}\rangle = -\int_0^{+\infty} \frac{d\phi}{dx}(x)\,dx = \phi(0) = \langle \delta_0, \phi\rangle$$

On en déduit, déjà un peu plus rigoureusement que précédemment, que $\dfrac{dH}{dx} = \delta_0$ dans $\mathcal{D}'(\mathbb{R})$.

3. Enfin, la topologie de $\mathcal{D}'(\Omega)$ qui sera définie plus loin *via* la notion de suite convergente, permettra de montrer des théorèmes à la fois très puissants et très simples à utiliser qui font des distributions un outil de premier choix pour faire de l'analyse.

6.3 Définitions et premier exemples

Venons en maintenant à la définition rigoureuse de ce qu'est une distribution. Nous avons dit plus haut qu'une distribution était une forme linéaire continue sur l'espace vectoriel $\mathcal{D}(\Omega)$ des fonctions C^∞ à support compact dans Ω. Rappelons donc pour commencer ce qu'est le support d'une fonction continue.

Définition 6.1. *Soit Ω ouvert de $\mathbb{R}^d$ et $\phi : \Omega \longrightarrow \mathbb{R}$ continue. On définit le* support *de ϕ par*

$$\mathrm{Supp}(\phi) = \overline{\{x \in \Omega,\ \phi(x) \neq 0\}}^{\Omega}.$$

Précisons que la notation $\overline{A}^{\Omega}$ signifie "fermeture de l'ensemble A dans Ω" ; c'est donc l'ensemble des points de Ω qui sont limites de suites de points de A. Afin d'éclairer cette définition, donnons quelques exemples :

1. $\Omega = \mathbb{R}$, $\phi(x) = \sin x$

$$\{x \in \mathbb{R},\ \phi(x) \neq 0\} = \mathbb{R} \setminus \pi\mathbb{Z}, \qquad \mathrm{Supp}(\phi) = \overline{\{x \in \mathbb{R},\ \phi(x) \neq 0\}}^{\mathbb{R}} = \mathbb{R}.$$

2. $\Omega =]-\pi, 2\pi[\,\cup\,]2\pi, 4\pi[$, $\phi(x) = \begin{cases} \sin x & \text{si } x \in]-\pi,\pi[\\ 0 & \text{si } x \in [\pi, 2\pi[\\ x - 3\pi & \text{si } x \in]2\pi, 4\pi[\end{cases}$

$$\{x \in \Omega,\ \phi(x) \neq 0\} =]-\pi, 0[\,\cup\,]0,\pi[\,\cup\,]2\pi, 3\pi[\,\cup\,]3\pi, 4\pi[,$$

$$\mathrm{Supp}(\phi) = \overline{\{x \in \Omega,\ \phi(x) \neq 0\}}^{\Omega} =]-\pi,\pi]\cup]2\pi, 4\pi[.$$

3. $\Omega =]0,3[$, $\phi(x) = \begin{cases} x & \text{si } x \in]0,1] \\ 2-x & \text{si } x \in]1,2] \\ 0 & \text{si } x \in]2,3[\end{cases}$

$$\{x \in \Omega,\ \phi(x) \neq 0\} =]0,2[, \qquad \mathrm{Supp}(\phi) = \overline{\{x \in \Omega,\ \phi(x) \neq 0\}}^{\Omega} =]0,2].$$

4. $\Omega =]-1/2,3[$, $\phi(x) = \begin{cases} 0 & \text{si } x \in]-1/2,0] \\ x & \text{si } x \in]0,1] \\ 2-x & \text{si } x \in]1,2] \\ 0 & \text{si } x \in]2,3[\end{cases}$

$$\{x \in \Omega,\ \phi(x) \neq 0\} =]0,2[, \qquad \mathrm{Supp}(\phi) = \overline{\{x \in \Omega,\ \phi(x) \neq 0\}}^{\Omega} = [0,2].$$

Il n'y a que dans le quatrième exemple que ϕ soit à support compact (rappelons à toutes fins utiles que les compacts de $\mathbb{R}^d$ sont les fermés bornés).

Définition 6.2. *On note $\mathcal{D}(\Omega)$ l'espace des fonctions indéfiniment dérivables sur Ω (i.e. C^∞) et à support compact. Ces fonctions sont appelées* fonctions d'essai *ou* fonctions test *dans Ω. Pour K compact dans Ω, on note $\mathcal{D}_K(\Omega)$ l'espace des fonctions test à support dans K.*

Remarque 6.3. On emploie parfois la terminologie “fonctions à support compact dans Ω” pour définir l'espace $\mathcal{D}(\Omega)$. Dans ce cas, le terme “dans Ω” ne fait bien sûr pas référence à la compacité du support mais au fait que celui-ci est pris dans Ω.

Définition 6.4. *On dit que T est une* distribution *dans l'ouvert Ω si T est une forme linéaire sur $\mathcal{D}(\Omega)$ qui vérifie la propriété de continuité suivante : pour tout compact K de Ω, il existe un entier p et une constante C tels que*

$$\forall \phi \in \mathcal{D}_K(\Omega), \qquad |\langle T, \phi\rangle| \leq C \sup_{x \in K,\ |\alpha| \leq p} |\partial^\alpha \phi(x)| \tag{6.2}$$

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

Remarque 6.5. Sur un espace vectoriel normé $(E, \|\cdot\|)$, la propriété de continuité des formes linéaires s'écrit tout simplement

$$\forall x \in E, \qquad |\langle T, x\rangle| \leq C\|x\|.$$

La forme plus complexe de la propriété de continuité (6.2) vient du fait que la topologie dont il faut munir l'espace vectoriel $\mathcal{D}(\Omega)$ pour que son dual $\mathcal{D}'(\Omega)$ ait de bonnes propriétés n'est pas une topologie issue d'une norme.

Définition 6.6. *Lorsque l'entier p peut être choisi indépendamment de K on dit que la distribution T est d'*ordre fini*, et la plus petite valeur de p possible est appelée l'*ordre de *T.*

Remarque 6.7. Heuristiquement, plus l'ordre d'une distribution est élevé, plus celle-ci est singulière.

Premiers exemples de distributions :

- Fonctions localement sommables : soit $f \in L^1_{loc}(\Omega)$. On note T_f (ou tout simplement f) la distribution définie par
$$\forall \phi \in \mathcal{D}(\Omega), \qquad \langle T_f, \phi\rangle = \int_\Omega f\,\phi. \tag{6.3}$$
- Mesures : soit μ une mesure borélienne localement bornée (i.e. une mesure μ définie sur l'ensemble $\mathcal{B}(\Omega)$ des boréliens de Ω telle que $\mu(K) < +\infty$ pour tout compact $K \subset \Omega$). On note T_μ (ou tout simplement μ) la distribution définie par
$$\forall \phi \in \mathcal{D}(\Omega), \qquad \langle T_\mu, \phi\rangle = \int_\Omega \phi\, d\mu. \tag{6.4}$$
- Pour tout $a \in \mathbb{R}$, on note δ'_a la distribution définie par
$$\forall \phi \in \mathcal{D}(\mathbb{R}), \qquad \langle \delta'_a, \phi\rangle = -\phi'(a). \tag{6.5}$$

L'objet de l'exercice 6.2 est de prouver que T_f, T_μ et δ'_a sont effectivement des distributions, que T_f et T_μ sont d'ordre 0 et que δ'_a est d'ordre 1.

6.4 Dérivation au sens des distributions

Considérons une fonction $f \in C^1(\mathbb{R})$. On a pour une telle fonction la formule d'intégration par parties
$$\forall \phi \in \mathcal{D}(\mathbb{R}), \qquad \int_{\mathbb{R}} \frac{df}{dx}\phi = -\int_{\mathbb{R}} f\frac{d\phi}{dx}.$$
De même, si f est une fonction de $C^1(\mathbb{R}^d)$, on a
$$\forall 1 \le i \le d,\ \ \forall \phi \in \mathcal{D}(\mathbb{R}^d), \qquad \int_{\mathbb{R}^d} \frac{\partial f}{\partial x_i}\phi = -\int_{\mathbb{R}^d} f\frac{\partial \phi}{\partial x_i}.$$
On peut en s'inspirant de cette formule définir la dérivée de n'importe quelle distribution :

Définition 6.8. *Soit Ω un ouvert de $\mathbb{R}^d$ et $T \in \mathcal{D}'(\Omega)$. On définit la dérivée de la distribution T par rapport à la variable x_i, que l'on note $\dfrac{\partial T}{\partial x_i}$, par la relation*
$$\forall \phi \in \mathcal{D}(\Omega), \qquad \langle \frac{\partial T}{\partial x_i}, \phi\rangle = -\langle T, \frac{\partial \phi}{\partial x_i}\rangle.$$

On vérifiera à titre d'exercice que la forme linéaire $\dfrac{\partial T}{\partial x_i}$ est bien une distribution, autrement dit qu'elle vérifie la propriété de continuité énoncée dans la définition 6.4.

6.5 Espace des fonctions test

L'espace des fonctions test joue un rôle central dans la construction et la manipulation des distributions. Pour montrer par exemple que la distribution δ'_a est d'ordre 1 (cf. exercice 6.2), nous avons besoin de postuler l'existence d'une fonction test $\phi_0 \in \mathcal{D}(\mathbb{R})$ paire et valant 1 au voisinage de 0. A l'extrême, on peut se demander si du fait que l'on impose des conditions très fortes sur les fonctions test, l'espace $\mathcal{D}(\Omega)$ ne se réduit pas à la fonction nulle. Nous verrons que ce n'est pas le cas et qu'en fait l'espace des fonctions test $\mathcal{D}(\Omega)$ est suffisamment riche pour être dense dans un espace aussi "gros" que $L^1(\Omega)$.

Cette section constitue un catalogue de lemmes qui permettent d'affirmer qu'il existe effectivement une fonction test vérifiant telle ou telle propriété. Il est donc fondamental pour la suite de connaître les résultats établis dans cette section ; les preuves, souvent assez techniques, peuvent en revanche être sautées en première lecture.

Lemme 6.9. *Soit Ω ouvert de $\mathbb{R}^d$ et $B_a(r)$ la boule de centre $a \in \Omega$ de rayon $r > 0$. Si $B_a(r) \subset \Omega$, il existe $\phi \in \mathcal{D}(\Omega)$ à support dans $B_a(r)$ vérifiant $\phi \geq 0$ et $\int_\Omega \phi = 1$.*

Preuve. Soit $\eta : \mathbb{R}^d \longrightarrow \mathbb{R}$ définie par

$$\eta(x) = \begin{cases} \exp\left(\dfrac{1}{|x|^2 - 1}\right) & \text{pour} \quad |x| < 1 \\ 0 & \text{pour} \quad |x| \geq 1. \end{cases}$$

On vérifie aisément que $\eta \in \mathcal{D}(\mathbb{R}^d)$. Soit maintenant $\phi_{a,r} : \Omega \longrightarrow \mathbb{R}$ définie par

$$\forall x \in \Omega, \qquad \phi_{a,r}(x) = \frac{\eta\left(\frac{x-a}{r/2}\right)}{(r/2)^d \int_{\mathbb{R}^d} \eta}.$$

La fonction $\phi_{a,r}$ vérifie les conditions requises. En effet

- $\phi_{a,r} \in \mathcal{D}(\Omega)$;
- $\mathrm{Supp}(\phi_{a,r}) = \overline{B_a(r/2)} \subset B_a(r)$;
- $\phi_{a,r} \geq 0$;
- enfin

$$\begin{aligned} \int_\Omega \phi_{a,r} &= \frac{1}{(r/2)^d \int_{\mathbb{R}^d} \eta} \int_\Omega \eta\left(\frac{x-a}{r/2}\right) dx \\ &= \frac{1}{(r/2)^d \int_{\mathbb{R}^d} \eta} \int_{\mathbb{R}^d} \eta\left(\frac{x-a}{r/2}\right) dx \\ &= \frac{1}{(r/2)^d \int_{\mathbb{R}^d} \eta} \int_{\mathbb{R}^d} \eta(y)(r/2)^d \, dy = 1. \end{aligned}$$

□

Lemme 6.10. *Soit K compact inclus dans Ω. Il existe une fonction test $\phi \in \mathcal{D}(\Omega)$ vérifiant les propriétés suivantes :*

- $0 \leq \phi \leq 1$ *sur* Ω *;*
- $\phi \equiv 1$ *sur* K*.*

Preuve. Soit r assez petit pour que

$$\left\{x \in \mathbb{R}^d, \;\; d(x, K) < 3r\right\} \subset \Omega.$$

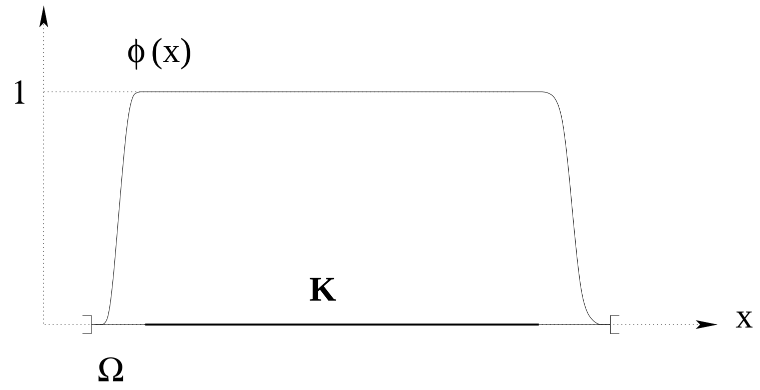
Soit χ la fonction caractéristique de l'ouvert

$$\left\{x \in \mathbb{R}^d, \;\; d(x, K) < r\right\}$$

et soit $\phi : \Omega \to \mathbb{R}$ la fonction définie par

$$\forall x \in \Omega, \qquad \phi(x) = \int_{\mathbb{R}^d} \phi_{0,r}(y) \, \chi(x - y) \, dy,$$

$\phi_{0,r}$ étant donnée dans la preuve du lemme 6.9 et prolongée par 0 sur $\mathbb{R}^d \setminus \Omega$. Alors ϕ satisfait les conditions requises. En effet,



- ϕ est de classe C^∞ (convolution d'une fonction L^∞ à support compact et d'une fonction C^∞) ;
- $\mathrm{Supp}(\phi) \subset \{x \in \mathbb{R}^d,\ \ d(x,K) \leq 2r\} \subset \Omega$;
- soit $x \in K$. On a

$$\begin{aligned}
\phi(x) &= \int_{\mathbb{R}^d} \chi(y)\, \phi_{0,r}(x-y)\, dy \\
&= \int_{B_x(r)} \chi(y)\, \phi_{0,r}(x-y)\, dy \\
&= \int_{B_x(r)} \phi_{0,r}(x-y)\, dy \\
&= \int_{\mathbb{R}^d} \phi_{0,r} = 1 \; ;
\end{aligned}$$

- enfin comme $\chi \geq 0$, $\phi_{0,r} \geq 0$ et $0 \leq \chi \leq 1$, on a pour tout $x \in \mathbb{R}^d$,

$$0 \leq \phi(x) = \int_{\mathbb{R}^d} \chi(x-y)\, \phi_{0,r}(y)\, dy \leq \int_{\mathbb{R}^d} \phi_{0,r}(y)\, dy = 1.$$

□

Lemme 6.11. *(Lemme de partition de l'unité) Soit $\Omega_1, \cdots, \Omega_n$ des ouverts de $\mathbb{R}^d$ et K un compact de $\mathbb{R}^d$ tel que*

$$K \subset \bigcup_{k=1}^{n} \Omega_k.$$

Il existe des fonctions test $\alpha_1, \cdots, \alpha_n$ telles que :
- $\mathrm{Supp}(\alpha_i) \subset \Omega_i$ *;*
- $0 \leq \alpha_i \leq 1$ *;*
- $\displaystyle\sum_{i=1}^{n} \alpha_i = 1$ *sur* K*.*

Preuve. Première étape. Montrons qu'il existe n compacts K_i, $1 \leq i \leq n$ tels que :
- pour tout $1 \leq i \leq n$, $K_i \subset \Omega_i$;
- $\displaystyle\bigcup_{i=1}^{n} K_i = K$.

Soit $x \in K$ et $B_x(r(x))$ une boule de centre x et de rayon $r(x) > 0$ contenue dans l'un des Ω_i.

$$\bigcup_{x \in K} B_x(r(x)/2)$$

est un recouvrement d'ouverts de K. K étant compact, on peut en extraire un sous-recouvrement fini

$$K \subset \bigcup_{k=1}^{p} B_{x_k}(r(x_k)/2).$$

On conclut en posant

$$K_i = \bigcup_{B_{x_k}(r(x_k)) \subset \Omega_i} (B_{x_k}(r(x_k)/2) \cap K).$$

Deuxième étape. D'après le lemme 6.9, on peut trouver $\phi_i \in \mathcal{D}(\mathbb{R}^d)$ vérifiant les propriétés suivantes :
- $\mathrm{Supp}(\phi_i) \subset \Omega_i$;
- $0 \le \phi_i \le 1$;
- $\phi_i \equiv 1$ sur K_i.

On a donc $\sum_{i=1}^{n} \phi_i > 0$ dans un voisinage V de K. Soit maintenant $\theta \in \mathcal{D}(\mathbb{R}^d)$ vérifiant
- $\mathrm{Supp}(\theta) \subset V$;
- $0 \le \theta \le 1$;
- $\theta \equiv 1$ sur K ;

et $\phi_0 = 1 - \theta$. On voit que les fonctions

$$\alpha_i = \frac{\phi_i}{\sum_{i=0}^{n} \phi_i}$$

satisfont les propriétés voulues. □

Lemme 6.12. *(Lemme de Borel) Pour toute suite $(a_\alpha)_{\alpha \in \mathbb{N}^d}$ il existe $\phi \in \mathcal{D}(\mathbb{R}^d)$ telle que*

$$\forall \alpha \in \mathbb{N}^d, \qquad \partial^\alpha \phi(0) = a_\alpha.$$

Preuve. Soit $(b_\alpha)_{\alpha \in \mathbb{N}^d}$ une suite de réels vérifiant

$$\forall \alpha \in \mathbb{N}^d, \qquad b_\alpha > 0, \quad |b_\alpha| \le \frac{1}{1 + |a_\alpha|}$$

$$\lim_{|\alpha| \to +\infty} b_\alpha = 0.$$

Soit par ailleurs $\psi \in \mathcal{D}(\mathbb{R}^d)$ telle que $\mathrm{Supp}(\psi) \subset B_0(1)$ et $\psi \equiv 1$ sur $B_0(1/2)$. Posons

$$\phi(x) = \sum_{\alpha \in \mathbb{N}^d} \frac{a_\alpha}{\alpha!} x^\alpha \psi\left(\frac{x}{b_\alpha}\right).$$

Alors ϕ satisfait les conditions requises. En effet,
- cette série converge car en tout point $x \neq 0$, elle ne comporte qu'un nombre fini de termes ;

– il y a de plus convergence normale donc uniforme sur $\mathbb{R}^d$. En effet,

$$\begin{aligned}
\sup_{x\in\mathbb{R}}\left|\frac{a_\alpha}{\alpha!}x^\alpha\psi\left(\frac{x}{b_\alpha}\right)\right| &= \left|\frac{a_\alpha\, b_\alpha^{|\alpha|}}{\alpha!}\right|\sup_{y\in\mathbb{R}}|y\psi(y)| \\
&\leq \frac{1}{\alpha!}\left|a_\alpha\, b_\alpha^{|\alpha|}\right|\sup|\psi| \leq \frac{\sup|\psi|}{\alpha!}.
\end{aligned}$$

La fonction ϕ est donc continue ;

– de même il y a convergence normale donc uniforme sur $\mathbb{R}$ de la série des dérivées ∂^α pour tout $\alpha \in \mathbb{N}^d$ (le vérifier). La fonction ϕ est donc de classe C^∞ ;
– le support de ϕ est contenue dans la boule unité ;
– enfin, on a pour tout $\beta \in \mathbb{N}^d$,

$$\begin{aligned}
\partial^\beta\phi(0) &= \sum_{\alpha\in\mathbb{N}^d}\frac{a_\alpha}{\alpha!}\partial^\beta\left(x^\alpha\psi\left(\frac{x}{b_\alpha}\right)\right)(0) \\
&= \sum_{\alpha\in\mathbb{N}^d}\frac{a_\alpha}{\alpha!}\sum_{|\gamma|\leq|\beta|}\left(\begin{array}{c}\beta\\ \gamma\end{array}\right)\partial^\gamma(x^\alpha)(0)\partial^{\beta-\gamma}\left(\psi\left(\frac{x}{b_\alpha}\right)\right)(0) \\
&= a_\beta.
\end{aligned}$$

□

Lemme 6.13. *(Lemme de Hadamard). Soit $\phi \in \mathcal{D}(\mathbb{R})$ telle que $\phi^{(k)}(0) = 0$ pour tout entier $0 \leq k \leq n$. Alors il existe $\psi \in \mathcal{D}(\mathbb{R})$ telle que pour tout $x \in \mathbb{R}$, $\phi(x) = x^{n+1}\psi(x)$.*

Preuve. Il suffit d'écrire la formule de Taylor avec reste intégral. □

Remarque 6.14. Un résultat analogue reste vrai en dimension supérieure.

Nous admettons le théorème suivant :

Théorème 6.15. *$\mathcal{D}(\Omega)$ est dense dans $L^p(\Omega)$ pour tout $1 \leq p < +\infty$.*

Remarque 6.16. Attention ! $\mathcal{D}(\Omega)$ n'est pas dense dans $L^\infty(\Omega)$. Exemple : sur $\Omega =]0,1[$, soit $f \in L^\infty(]0,1[)$ définie par $f(x) = 1$ presque partout. On a pour tout $\phi \in \mathcal{D}(]0,1[)$, $\|f-\phi\|_{L^\infty} \geq 1$.

Corollaire 6.17. *Pour tout entier $k \geq 0$ et pour tout $1 \leq p < +\infty$, $C^k(\Omega)\cap L^p(\Omega)$ est dense dans $L^p(\Omega)$.*

Preuve. Conséquence immédiate du théorème 6.15 puisque

$$\mathcal{D}(\Omega) \subset C^k(\Omega)\cap L^p(\Omega) \subset L^p(\Omega).$$

□

6.6 Exercices

▷ **6.1.** Soit f et g continues de $\mathbb{R}^d$ dans $\mathbb{R}$.

– Montrer que
$$\mathrm{Supp}(fg) \subset \mathrm{Supp}(f) \cap \mathrm{Supp}(g)$$
et que l'inclusion peut être stricte.

– Montrer que
$$\mathrm{Supp}(f+g) \subset \mathrm{Supp}(f) \cup \mathrm{Supp}(g)$$
et que l'inclusion peut être stricte.

▷ **6.2.**

– Montrer que la forme linéaire T_f définie par (6.3) est une distribution d'ordre 0 sur Ω.
– Montrer que la forme linéaire T_μ définie par (6.4) est une distribution d'ordre 0 sur Ω.
– Montrer que la forme linéaire δ'_a définie par (6.5) est une distribution d'ordre 1 sur $\mathbb{R}$ (*Indication : on pourra considérer son action sur la suite de fonctions d'essai* $\phi_j(x) = \phi_0(x)\ atan(jx)$ *où* $\phi_0 \in \mathcal{D}(\mathbb{R})$ *est paire et vaut* 1 *au voisinage de* 0 *; une telle fonction existe bien, cf. lemme 6.9*). Vérifier que δ'_a est la dérivée de δ_a au sens des distributions.

▷ **6.3.** Calculer la dérivée et la dérivée seconde au sens des distributions de la fonction “chapeau” définie sur $\mathbb{R}$ par
$$f(x) = \begin{cases} 0 & \text{si } x < -1 \\ x+1 & \text{si } -1 < x < 0 \\ 1-x & \text{si } 0 < x < 1 \\ 0 & \text{si } x > 1 \end{cases}$$

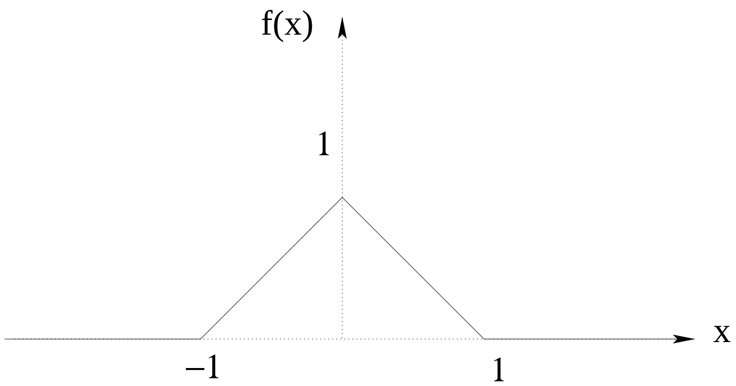


FIG. 6.1 – Fonction chapeau

Chapitre 7

Distributions : exemples et principales propriétés

L'objectif de ce chapitre est d'une part de donner de nombreux exemples de distributions pour bien ancrer dans l'esprit du lecteur l'idée que les distributions constituent une généralisation naturelle de la notion de fonction et d'autre part de présenter quelques propriétés importantes des distributions liées notamment aux notions de convergence et de dérivation.

L'espace des distributions contient des objets que nous connaissons déjà : les fonctions localement intégrables et plus généralement les mesures boréliennes localement finies. Nous allons voir que ces derniers objets engendrent en fait l'espace vectoriel des distributions d'ordre 0. Nous verrons ensuite des exemples de distributions plus singulières (d'ordre supérieur ou égal à 1), choisis parmi celles qui sont les plus fréquemment rencontrées en pratique (multipôles et multicouches, valeurs principales et parties finies). Nous définirons enfin l'espace des distributions à support compact et caractériserons l'ensemble des distributions dont le support est réduit à un point.

L'objet de la dernière partie de ce chapitre est d'introduire la notion fondamentale de convergence dans l'espace des distributions et d'approfondir le concept de dérivation au sens des distributions qui a déjà été introduit au chapitre 6. La simplicité avec laquelle ces notions se manipulent permettra au lecteur d'apprécier la puissance de cet outil d'analyse que sont les distributions. Nous verrons par exemple comment la formule des sauts dans l'espace, souvent utilisée en mécanique et dans les applications de la physique, s'obtient de manière naturelle dans le cadre de la théorie des distributions. On étudiera enfin la multiplication des distributions par des fonctions C^∞.

7.1 Fonctions localement sommables

L'espace des fonctions test est suffisamment riche pour que l'application naturelle $f \mapsto T_f$ de $L^1_{loc}(\Omega)$ dans $\mathcal{D}'(\Omega)$ soit injective. Ce fait très important, que nous avons mentionné dès le début du chapitre précédent, fait l'objet du théorème suivant :

Théorème 7.1. *Soit f et g dans $L^1_{loc}(\Omega)$. Alors*

$$f = g \quad \Leftrightarrow \quad T_f = T_g \ \textit{dans}\ \mathcal{D}'(\Omega).$$

Preuve. L'implication $\Rightarrow$ est évidente. Nous allons démontrer la réciproque en nous limitant au cas où $\Omega = \mathbb{R}^d$ et où f et g sont dans $L^1(\mathbb{R}^d)$ (ceci pour éviter des complications techniques).

Considérons donc f et g dans $L^1(\mathbb{R}^d)$ vérifiant $T_f = T_g$ et posons $h = f - g$; il s'agit de vérifier que $h = 0$ presque partout, sachant qu'on a

$$\forall \phi \in \mathcal{D}(\mathbb{R}^d), \qquad \int_\Omega h\phi = 0.$$

Soit $\epsilon > 0$ et $\psi \in \mathcal{D}(\mathbb{R}^d)$ vérifiant

$$\|h - \psi\|_{L^1} \leq \epsilon/3.$$

L'existence d'un tel ψ résulte de la densité de $\mathcal{D}(\mathbb{R}^d)$ dans $L^1(\mathbb{R}^d)$ (cf. théorème 6.15). Soit une fonction $\chi \in \mathcal{D}(\mathbb{R}^d)$ positive à support dans la boule unité et d'intégrale égale à 1. On construit la suite de fonctions $(\chi_n)_{n\in\mathbb{N}^*}$ définie par

$$\chi_n(x) = n^d \chi(n\,x).$$

Une telle suite s'appelle une approximation de l'identité (cf. section 7.5). Considérons maintenant la suite de fonctions $(h_n)_{n\in\mathbb{N}^*}$ définie par

$$h_n(x) = (h \star \chi_n)(x) = \int_{\mathbb{R}^d} h(y)\,\chi_n(x-y)\,dy.$$

Comme pour tout $n \in \mathbb{N}^*$, $\chi_n \in \mathcal{D}(\mathbb{R}^d)$, la suite $(h_n)_{n\in\mathbb{N}^*}$ est en fait identiquement nulle sur $\mathbb{R}^d$. On a donc

$$\|h\|_{L^1} = \|h - h \star \chi_n\|_{L^1} \leq \|h - \psi\|_{L^1} + \|\psi - \psi \star \chi_n\|_{L^1} + \|(h-\psi) \star \chi_n\|_{L^1}.$$

En utilisant l'inégalité (5.9), on obtient

$$\|(h-\psi) \star \chi_n\|_{L^1} \leq \|h - \psi\|_{L^1}\ \|\chi_n\|_{L^1} \leq \epsilon/3,$$

(on vérifie en effet que $\|\chi_n\|_{L^1} = 1$ pour tout $n \in \mathbb{N}^*$). Par ailleurs (ce point sera démontré plus loin),

$$\|\psi - \psi \star \chi_n\|_{L^1} \underset{n\to+\infty}{\longrightarrow} 0. \tag{7.1}$$

On peut donc choisir n assez grand pour que

$$\|\psi - \psi \star \chi_n\|_{L^1} \leq \epsilon/3,$$

ce qui conduit finalement à

$$\|h\|_{L^1} \leq \epsilon.$$

Cette majoration est valable pour tout $\epsilon > 0$. Donc $h = 0$ dans L^1. Terminons la preuve en démontrant (7.1). Remarquons d'abord que

$$|\psi \star \chi_n(x)| = \left| \int_{\mathbb{R}^d} \psi(x-y)\,\chi_n(y)\,dy \right| \leq \sup|\psi|,$$

et que

$$\mathrm{Supp}(\psi \star \chi) \subset \mathrm{Supp}(\psi) + \mathrm{Supp}(\chi_n) \subset \mathrm{Supp}(\psi) + \mathrm{Supp}(\chi).$$

On obtient d'autre part en utilisant le changement de variable $z = y/n$,

$$\psi \star \chi_n(x) = n^{-d} \int_{\mathbb{R}^d} \psi(x-y)\,\chi(y/n)\,dy = \int_{\mathbb{R}^d} \psi(x - z/n)\,\chi(z)\,dz$$

Il résulte alors du théorème de convergence dominée 4.42 que $\psi \star \chi_n(x)$ tend vers $\psi(x)$ pour tout $x \in \mathbb{R}^d$. En regroupant ces résultats, on aboutit à

$$\left\{ \begin{array}{l} |\psi - \psi \star \chi_n| \leq 2\sup|\psi|\chi_{\mathrm{Supp}(\psi)+\mathrm{Supp}(\chi)} \quad \text{pour tout } n \in \mathbb{N}^*, \\ (\psi - \psi \star \chi_n)(x) \underset{n\to+\infty}{\longrightarrow} 0 \quad \text{pour tout } x \in \mathbb{R}^d. \end{array} \right.$$

La convergence (7.1) découle donc à nouveau du théorème de convergence dominée. □

On peut donc identifier une fonction de $L^1_{loc}(\Omega)$ et la distribution qui lui est associée et noter f à la place de T_f. On peut ainsi reformuler le théorème précédent de la façon suivante :

Théorème 7.2. *Soit f et g dans $L^1_{loc}(\Omega)$. Alors*

$$f = g \quad \Leftrightarrow \quad f = g \text{ dans } \mathcal{D}'(\Omega).$$

Remarque 7.3. On écrira

$$L^1_{loc}(\Omega) \subset \mathcal{D}'(\Omega),$$

de la même façon qu'on écrit

$$\mathbb{N} \subset \mathbb{R} \subset \mathbb{C}.$$

La notation $A \subset B$ signifie dans ce cadre qu'il existe une injection naturelle de A dans B.

Remarque 7.4. Attention ! La notion de distribution généralise en un certain sens la notion de fonction mais cela ne veut pas dire que toutes les fonctions sont des distributions. Seules les fonctions de L^1_{loc} sont en effet des distributions. Ainsi, la fonction $f(x) = \dfrac{1}{|x|}$ définit une distribution dans $\mathbb{R}^d$ pour $d \geq 2$ mais pas dans $\mathbb{R}$.

7.2 Mesures de Radon

Théorème 7.5. *Soit $T : \phi \mapsto \langle T, \phi \rangle$ une forme linéaire positive sur $\mathcal{D}(\Omega)$ (positive signifie $\phi \geq 0 \Rightarrow \langle T, \phi \rangle \geq 0$). Alors T est une distribution dans Ω d'ordre 0.*

Preuve. Remarquons tout d'abord qu'une forme linéaire positive est croissante :

$$\phi \leq \psi \quad \Rightarrow \quad \langle T, \phi \rangle \leq \langle T, \psi \rangle.$$

Soit maintenant K un compact et $\psi \in \mathcal{D}(\Omega)$ telle que $0 \leq \psi \leq 1$ et $\psi = 1$ sur K (une telle fonction existe en vertu du lemme 6.9). On a pour tout $\phi \in \mathcal{D}_K(\Omega)$

$$-(\sup_{x \in K} |\phi(x)|)\psi \leq \phi \leq (\sup_{x \in K} |\phi(x)|)\psi.$$

Donc

$$-(\sup_{x \in K} |\phi(x)|)\langle T, \psi \rangle \leq \langle T, \phi \rangle \leq (\sup_{x \in K} |\phi(x)|)\langle T, \psi \rangle.$$

Posons $C = \langle T, \psi \rangle$. Il vient

$$\forall \phi \in \mathcal{D}_K(\Omega), \qquad |\langle T, \phi \rangle| \leq C \sup_{x \in K} |\phi(x)|.$$

□

Définition 7.6. *On appelle* mesure de Radon positive *une forme linéaire positive sur $\mathcal{D}(\Omega)$ et* mesure de Radon *(tout court) une combinaison linéaire $T_1 - T_2$ (ou $T_1 - T_2 + iT_3 - iT_4$ dans le cas complexe) de mesures de Radon positives.*

Remarque 7.7. La terminologie *mesure de Radon* vient du fait (non trivial) que l'ensemble des mesures de Radon positives définies par la définition 7.6 s'identifie avec l'ensemble des mesures (au sens de la définition 4.8) boréliennes localement finies sur Ω (les qualificatifs boréliennes et localement finies signifient respectivement définies sur la tribu borélienne de Ω et finies sur tout compact, i.e. $\mu(K) < +\infty$ pour tout compact inclus dans Ω). En effet, il est facile de vérifier que pour toute mesure borélienne μ localement finie, l'application

$$\begin{array}{rcl} T_\mu : \mathcal{D}(\Omega) & \longrightarrow & \mathbb{R} \\ \phi & \mapsto & \langle T_\mu, \phi \rangle = \displaystyle\int_\Omega \phi \, d\mu \end{array}$$

définit une mesure de Radon positive. On peut prouver que la réciproque est vraie : à toute mesure de Radon positive T on peut associer une unique mesure borélienne μ localement finie telle que $T = T_\mu$ (voir par exemple [Rudin 92]).

Outre les distributions définies par des fonctions de L^1_{loc}, les mesures de Radon les plus fréquemment rencontrées en pratique sont les mesures portées par des variétés.

Exemples :

- Mesures ponctuelles (mesures portées par des points).

  La distribution
  $$\sum_{k=1}^{N} m_k \, \delta_{x_k}$$
  est une mesure de Radon. C'est une mesure de Radon positive si et seulement si $m_k \geq 0$ pour tout k.

- Mesures linéiques (mesures portées par des courbes).

  Soit $g \in L^1_{loc}(\mathbb{R})$ et T la distribution dans $\mathbb{R}^3$ définie par
  $$\forall \phi \in \mathcal{D}(\mathbb{R}^3), \qquad \langle T, \phi \rangle = \int_{\mathbb{R}} g(z) \, \phi(0,0,z) \, dz.$$
  T est une mesure de Radon portée par l'axe Oz. C'est une mesure de Radon positive si et seulement si $g \geq 0$ presque partout.

  Plus généralement, on définit la mesure de Lebesgue d'une courbe paramétrée γ de $\mathbb{R}^3$
  $$\gamma = \left\{ M(u) = \begin{pmatrix} x(u) \\ y(u) \\ z(u) \end{pmatrix}, \quad u \in \,]a,b[\, \subset \mathbb{R} \right\}$$
  de classe C^1 comme la distribution notée parfois $\delta_\gamma \otimes ds$ et définie par
  $$\langle \delta_\gamma \otimes ds, \phi \rangle = \int_a^b \phi(x(u), y(u), z(u)) \, \sqrt{x'(u)^2 + y'(u)^2 + z'(u)^2} \, du.$$
  On écrit aussi
  $$\langle \delta_\gamma \otimes ds, \phi \rangle = \int_\gamma \phi \, ds.$$
  Cette distribution est une mesure de Radon positive invariante par changement de paramètre. La distribution linéique de densité g sur γ est définie par
  $$\langle \delta_\gamma \otimes g \, ds, \phi \rangle = \int_a^b g(x(u), y(u), z(u)) \, \phi(x(u), y(u), z(u)) \, \sqrt{x'(u)^2 + y'(u)^2 + z'(u)^2} \, du.$$
  On écrit aussi
  $$\langle \delta_\gamma \otimes g \, ds, \phi \rangle = \int_\gamma \phi g \, ds.$$

– Mesures surfaciques (mesures portées par des surfaces).

Soit $g \in L^1_{loc}(\mathbb{R}^2)$ et T la distribution dans $\mathbb{R}^3$ définie par

$$\forall \phi \in \mathcal{D}(\mathbb{R}^3), \qquad \langle T, \phi \rangle = \int_{\mathbb{R}^2} g(y,z)\,\phi(0,y,z)\,dy\,dz.$$

T définit une mesure de Radon portée par le plan d'équation $x = 0$. C'est une mesure de Radon positive si et seulement si $g \geq 0$ presque partout.

Plus généralement, on définit la mesure de Lebesgue d'une surface paramétrée Σ de $\mathbb{R}^3$

$$\Sigma = \left\{ M(u,v) = \begin{pmatrix} x(u,v) \\ y(u,v) \\ z(u,v) \end{pmatrix}, \quad (u,v) \in \omega \subset \mathbb{R}^2 \right\}$$

de classe C^1 comme la distribution notée parfois $\delta_\Sigma \otimes d\sigma$ et définie par

$$\langle \delta_\Sigma \otimes d\sigma, \phi \rangle = \int_\omega \phi(x(u,v), y(u,v), z(u,v)) \left\| \frac{\partial M}{\partial u} \times \frac{\partial M}{\partial v} \right\| du\,dv.$$

On écrit aussi

$$\langle \delta_\Sigma \otimes d\sigma, \phi \rangle = \int_\Sigma \phi\, d\sigma.$$

Cette distribution est une mesure de Radon positive invariante par changement de paramètres. La distribution dite *de simple couche* de densité g portée par Σ, est définie par

$$\langle \delta_\Sigma \otimes g\,d\sigma, \phi \rangle = \int_\omega g(x(u,v), y(u,v), z(u,v))\phi(x(u,v), y(u,v), z(u,v)) \left\| \frac{\partial M}{\partial u} \times \frac{\partial M}{\partial v} \right\| du\,dv.$$

On écrit aussi

$$\langle \delta_\Sigma \otimes g\,d\sigma, \phi \rangle = \int_\Sigma \phi g\, d\sigma.$$

7.3 Exemples de distributions plus singulières

7.3.1 Multipôles, multicouches

Une conséquence immédiate du théorème 7.5 est qu'une combinaison linéaire de mesures positives et négatives finies ne peut produire qu'une distribution d'ordre 0. Une façon de construire des distributions plus singulières consiste à faire intervenir des compensations entre mesures (penser à des charges électriques) infinies de signes contraires. A titre d'exemples, mentionnons :

– les *multipôles*, qui sont les combinaisons linéaires finies de dérivées successives de la masse de Dirac, dont l'exemple le plus simple est le dipôle. Le dipôle de moment dipolaire p situé au point a et orienté selon le vecteur e est ainsi défini par

$$\langle T, \phi \rangle = p\,e \cdot \nabla\phi(a).$$

On montre facilement (le faire en exercice) que

$$T = -p\,e \cdot \nabla\delta_a \;;$$

– les *multicouches* qui sont les dérivées successives des distributions de simples couches ; ainsi, une distribution *de double couche* de densité $g \in L^1_{loc}(\mathbb{R}^2)$ portée par le plan d'équation $x = 0$ est définie par

$$\langle T, \phi \rangle = \int_{\mathbb{R}^2} g(y,z) \frac{\partial \phi}{\partial x}(0,y,z)\,dy\,dz.$$

On peut voir que $T = \dfrac{\partial S}{\partial x}$ où S est la distribution de simple couche de densité $-g$ portée par le plan d'équation $x = 0$. La distribution T est aussi la limite dans $\mathcal{D}'(\mathbb{R}^3)$ (cette notion sera définie au chapitre suivant, section 7.5) de la famille de distributions

$$T_\epsilon = SC(g/\epsilon, \epsilon) + SC(-g/\epsilon, 0)$$

lorsque ϵ tend vers 0, où $SC(h, x_0)$ désigne la distribution de simple couche de densité $h \in L^1_{loc}(\mathbb{R}^2)$ portée par le plan d'équation $x = x_0$, définie par

$$\forall \phi \in \mathcal{D}(\mathbb{R}^3), \qquad \langle SC(h, x_0), \phi \rangle = \int_{\mathbb{R}^2} h(y,z)\, \phi(x_0, y, z)\, dy\, dz.$$

Les distributions T_ϵ étant composées de deux simples couches, on comprend l'origine de la dénonimation "double couche" utilisée pour désigner leur limite T.

7.3.2 Valeurs principales et parties finies

Les valeurs principales et les parties finies de fonctions sont des distributions qui apparaissent naturellement quand on cherche à interpréter en tant que distributions des fonctions singulières. Nous nous bornons ici à donner deux exemples.

Valeur principale de $\frac{1}{x}$

On définit la valeur principale de la fonction $\frac{1}{x}$ par

$$\forall \phi \in \mathcal{D}(\mathbb{R}), \qquad \langle \mathrm{vp}\left(\frac{1}{x}\right), \phi \rangle = \lim_{\epsilon \to 0+} \int_{|x|>\epsilon} \frac{\phi(x)}{x}\, dx. \tag{7.2}$$

On montre que $\mathrm{vp}\left(\frac{1}{x}\right)$ définit une distribution d'ordre 1 (cf. exercice 7.2).

Partie finie de $H(x)x^\alpha$

Soit $\phi \in \mathcal{D}(\mathbb{R})$. L'intégrale

$$\int_{\mathbb{R}} H(x) x^\alpha \phi(x)\, dx = \int_0^{+\infty} x^\alpha \phi(x)\, dx$$

est définie pour $\alpha > -1$, mais pas pour $\alpha \leq -1$ (si $\phi(0) \neq 0$).

En utilisant le lemme de Hadamard, on montre que pour $\epsilon > 0$,

$$\int_\epsilon^{+\infty} \phi(x) x^\alpha\, dx = P_\phi(\epsilon) + R_\phi(\epsilon)$$

où P_ϕ est une combinaison linéaire de puissances négatives de ϵ, et de $\mathrm{Log}(\epsilon)$ dans le cas où α est un entier strictement négatif, et où $R_\phi(\epsilon)$ tend vers une limite finie lorsque ϵ tend vers 0.

On définit alors la partie finie de la fonction $H(x)x^\alpha$, notée $\mathrm{Pf}(H(x)x^\alpha)$ par la formule

$$\forall \phi \in \mathcal{D}(\mathbb{R}), \qquad \langle \mathrm{Pf}(H(x)x^\alpha), \phi \rangle = \lim_{\epsilon \to 0+} R_\phi(\epsilon).$$

On peut montrer (cf. exercice 7.4) que $\mathrm{Pf}(H(x)x^\alpha)$ est une distribution d'ordre égal à la partie entière de α.

7.4 Distributions à support compact

Définition 7.8. *Soit $T \in \mathcal{D}'(\Omega)$.*

1. *Soit ω ouvert inclus dans Ω. On dit que T est nulle sur ω si pour toute fonction $\phi \in \mathcal{D}(\Omega)$ telle que* $\mathrm{Supp}(\phi) \subset \omega$*, on a $\langle T, \phi\rangle = 0$.*
2. *Le* support *de T est le complémentaire dans Ω de la réunion des ouverts de Ω sur lesquels T est nulle.*

Définition 7.9. *On note $\mathcal{E}'(\Omega)$ l'espace vectoriel des distributions sur Ω à support compact.*

Théorème 7.10. *Si une distribution de $\mathcal{D}'(\Omega)$ est à support compact, elle est d'ordre fini.*

Preuve. Soit K le support de T et $\alpha = d(K, \mathbb{R}^d \setminus \Omega)$ (dans le cas où $\Omega = \mathbb{R}^d$, on prendra $\alpha = +\infty$). Posons $\beta = \inf(1, \alpha)$ et considérons les ensembles

$$K' = \left\{x \in \mathbb{R}^d, \quad d(x, K) \leq \frac{\beta}{3}\right\},$$

et

$$\Omega' = \left\{x \in \mathbb{R}^d, \quad d(x, K) < \frac{2\beta}{3}\right\}.$$

Il est clair que K' est compact et que Ω' est un ouvert de fermeture compacte. De plus, on a

$$K \subset K' \subset \Omega' \subset \overline{\Omega'} \subset \Omega.$$

Soit p un entier et C une constante réelle tels que

$$\forall \phi \in \mathcal{D}_{\overline{\Omega'}}(\Omega), \qquad |\langle T, \phi\rangle| \leq C \sup_{x \in \overline{\Omega'},\, |\alpha| \leq p} |\partial^\alpha \phi(x)|.$$

Soit maintenant $\rho \in \mathcal{D}(\Omega)$ égale à 1 sur K' et à support dans Ω'. On a pour tout $\phi \in \mathcal{D}(\Omega)$,

$$\langle T, \phi\rangle = \langle T, \rho\phi\rangle + \langle T, (1-\rho)\phi\rangle$$

et $\langle T, (1-\rho)\phi\rangle = 0$ puisque les supports de T et de $(1-\rho)\phi$ sont disjoints. De plus, comme $\mathrm{Supp}(\rho\phi) \subset \overline{\Omega'}$, on a

$$\forall \phi \in \mathcal{D}(\Omega), \qquad |\langle T, \phi\rangle| \leq C \sup_{x \in \Omega,\, |\alpha| \leq p} |\partial^\alpha (\rho\phi)(x)|.$$

D'après la formule de Leibniz,

$$\forall \phi \in \mathcal{D}(\Omega), \qquad |\langle T, \phi\rangle| \leq C' \sup_{x \in \Omega,\, |\alpha| \leq p} |\partial^\alpha \phi(x)|$$

avec

$$C' = C \sup_{x \in \Omega,\, |\alpha| \leq p,\, \beta \leq \alpha} \frac{\alpha!}{\beta!\,(\alpha-\beta)!} |\partial^\beta \rho(x)|.$$

Donc T est d'ordre fini inférieur ou égal à p. □

Remarque 7.11. *Soit $T \in \mathcal{E}'(\Omega)$ une distribution à support compact, p son ordre, K un voisinage compact de Supp T et $\chi \in \mathcal{D}(\Omega)$ valant 1 sur K. Posons pour tout $\phi \in C^\infty(\Omega)$,*

$$\langle T, \phi \rangle_{\mathcal{E}', C^\infty} = \langle T, \chi\phi \rangle.$$

Cette définition est indépendante de χ : soit en effet χ_1 et χ_2 dans $\mathcal{D}(\Omega)$ valant 1 sur K ; on a

$$\langle T, \chi_1\phi \rangle - \langle T, \chi_2\phi \rangle = \langle T, (\chi_1 - \chi_2)\phi \rangle.$$

La fonction $\widetilde{\phi} = (\chi_1 - \chi_2)\phi$ étant nulle sur K voisinage de Supp u, on a

$$\langle T, (\chi_1 - \chi_2)\phi \rangle = 0.$$

On a ainsi associé à une distribution à support compact une forme linéaire sur $C^\infty(\Omega)$.

Nous montrerons au chapitre 10 d'autres résultats relatifs aux distributions à support compact (cf. le théorème 10.23 et la section 10.3). Pour l'heure, intéressons-nous au problème particulier de la caractérisation des distributions à support compact dont le support est réduit à un point.

Proposition 7.12. *Soit $T \in \mathcal{D}'(\Omega)$ d'ordre 0 vérifiant* $\mathrm{Supp}(T) = \{a\}$. *Alors T est de la forme*

$$T = c\,\delta_a,$$

où c est une constante.

Preuve. Soit $\epsilon > 0$ tel que $B_a(\epsilon) \subset \Omega$. Soit $K = \overline{B_a(\epsilon/2)}$ et C une constante telle que

$$\forall \phi \in \mathcal{D}_K(\Omega), \qquad |\langle T, \phi \rangle| \leq C \sup_{x \in K} |\phi(x)|.$$

Soit maintenant $0 < r \underline{<} \epsilon$ et $\rho \in \mathcal{D}(\Omega)$ une fonction à support compact dans $B_a(r)$ vérifiant $0 \leq \rho \leq 1$ et $\rho \equiv 1$ dans $\overline{B_a(r/2)}$. Pour tout $\phi \in \mathcal{D}(\Omega)$, on a par linéarité

$$\langle T, \phi \rangle = \langle T, \rho\phi \rangle + \langle T, (1 - \rho)\phi \rangle.$$

La fonction $(1 - \rho)\phi$ étant nulle au voisinage de $\{a\} = \mathrm{Supp}(T)$, il est clair que $\langle T, (1 - \rho)\phi \rangle = 0$. Puisque en outre, $\mathrm{Supp}(\rho\phi) \subset K$, il en résulte que

$$|\langle T, \phi \rangle| = |\langle T, \rho\phi \rangle| \leq C \sup_{\Omega} |\rho\phi| \leq C \sup_{B_a(r)} |\phi|.$$

La constante C étant indépendante de r, on obtient, en faisant tendre r vers 0,

$$|\langle T, \phi \rangle| \leq C|\phi(a)|. \tag{7.3}$$

Soit maintenant $\psi \in \mathcal{D}(\Omega)$ quelconque. On peut toujours décomposer ψ en la somme

$$\psi = \psi(a)\rho + (\psi - \psi(a)\rho)$$

dans laquelle la fonction $\phi = \psi - \psi(a)\rho$ est telle que $\phi(a) = 0$. En utilisant l'inégalité (7.3), il s'en suit que $\langle T, \psi - \psi(a)\rho \rangle = 0$. Finalement,

$$\langle T, \psi \rangle = \langle T, \psi(a)\rho \rangle = \langle T, \rho \rangle\ \psi(a).$$

D'où l'on conclut que $T = c\delta_a$, c désignant la constante $\langle T, \rho \rangle$. □

La proposition suivante (admise) généralise la proposition 7.12 en caractérisant l'ensemble des distributions (d'ordre quelconque) dont le support est réduit à un point.

Proposition 7.13. *Soit $T \in \mathcal{D}'(\Omega)$ vérifiant* $\mathrm{Supp}(T) = \{a\}$. *Alors, T est de la forme*

$$T = \sum_{|\alpha| \leq p} c_\alpha \partial^\alpha \delta_a,$$

où p est l'ordre de T et où les c_α, $|\alpha| \leq p$, désignent des constantes.

7.5 Convergence des distributions

Définition 7.14. *Soit* $(T_n)_{n\in\mathbb{N}}$ *une suite de distributions dans* Ω. *On dit qu'une suite* $(T_n)_{n\in\mathbb{N}}$ converge vers T dans $\mathcal{D}'(\Omega)$ *si et seulement si*

$$\forall\phi \in \mathcal{D}(\Omega), \qquad \langle T_n, \phi\rangle \underset{n\to+\infty}{\longrightarrow} \langle T, \phi\rangle.$$

Exemple : soit $\chi \in \mathcal{D}(\mathbb{R}^d)$ positive, à support dans la boule unité et d'intégrale égale à 1. Posons pour $n \geq 1$

$$\chi_n(x) = n^d \chi(n\,x).$$

On a

$$\chi_n \underset{n\to+\infty}{\longrightarrow} \delta_0 \text{ dans } \mathcal{D}'(\mathbb{R}^d).$$

Définition 7.15. *Une suite* $(\chi_n)_{n\in\mathbb{N}}$ *de* $\mathcal{D}(\mathbb{R}^d)$ *vérifiant*

$$\chi_n \underset{n\to+\infty}{\longrightarrow} \delta_0 \qquad \textit{dans } \mathcal{D}'(\mathbb{R}^d),$$

est appelée une approximation de l'identité.

Remarque 7.16. Cette dénomination vient du fait que la distribution δ_0 est l'identité pour l'opération de convolution que nous avons définie pour des fonctions de $L^1(\mathbb{R}^d)$ au chapitre 4 et qui peut être étendue à une certaine classe de distributions.

Proposition 7.17. *Soit* $1 \leq p \leq +\infty$. *La convergence dans* $L^p_{loc}(\Omega)$ *implique la convergence dans* $\mathcal{D}'(\Omega)$.

Preuve. Soit $(f_n)_{n\in\mathbb{N}}$ tendant vers f dans $L^p_{loc}(\Omega)$. Soit $\phi \in \mathcal{D}(\Omega)$. En se servant de l'inégalité de Hölder, on obtient pour $1 < p < +\infty$

$$\begin{aligned}
|\langle f_n, \phi\rangle - \langle f, \phi\rangle| &= \left|\int_\Omega (f_n - f)\phi\right| \\
&\leq \int_\Omega |f_n - f|\,|\phi| \\
&\leq \sup|\phi| \int_{\mathrm{Supp}(\phi)} |f_n - f| \\
&\leq \sup|\phi| \left(\int_{\mathrm{Supp}(\phi)} |f_n - f|^p\right)^{1/p} \left(\int_{\mathrm{Supp}(\phi)} 1^{p'}\right)^{1/p'} \\
&\leq \sup|\phi|\, \|f_n - f\|_{L^p(\mathrm{Supp}(\phi))}\, |\,\mathrm{Supp}(\phi)|^{1/p'} \longrightarrow 0.
\end{aligned}$$

Pour $p = 1$ ou $p = +\infty$, la preuve est encore plus facile ; elle est laissée en exercice au lecteur. □

Remarque 7.18. La convergence presque partout *n'implique pas* la convergence dans $\mathcal{D}'$ et réciproquement. Ainsi :

1. il peut y avoir convergence dans $\mathcal{D}'$ et pas convergence presque partout ;
2. il peut y avoir convergence presque partout et pas convergence dans $\mathcal{D}'$;

3. il peut y avoir convergence presque partout et convergence dans $\mathcal{D}'$ sans que les deux limites soient égales.

L'exercice 7.10 fournit des exemples de tels comportements.

L'espace $\mathcal{D}'(\Omega)$ muni de la topologie définie ci-dessus possède la propriété remarquable que les opérateurs de dérivation y sont continus :

Proposition 7.19. *Si la suite (T_n) tend vers T dans $\mathcal{D}'(\Omega)$, alors pour tout $\alpha \in \mathbb{N}^d$, $(\partial^\alpha T_n)$ converge vers $\partial^\alpha T$ dans $\mathcal{D}'(\Omega)$.*

Preuve. Soit $\phi \in \mathcal{D}(\Omega)$. On a

$$\begin{aligned} \langle \partial^\alpha T_n, \phi \rangle &= (-1)^{|\alpha|} \langle T_n, \partial^\alpha \phi \rangle \\ &\longrightarrow (-1)^{|\alpha|} \langle T, \partial^\alpha \phi \rangle \\ &= \langle \partial^\alpha T, \phi \rangle. \end{aligned}$$

Donc $\partial^\alpha T_n \longrightarrow \partial^\alpha T$ dans $\mathcal{D}'(\Omega)$. □

Comme corollaire immédiat, énonçons la proposition suivante :

Proposition 7.20. *Soit $(T_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{D}'(\Omega)$. Supposons que la série $\sum_{n\in\mathbb{N}} T_n$ converge dans $\mathcal{D}'(\Omega)$ vers une distribution T. Alors, pour tout $\alpha \in \mathbb{N}^d$, la série $\sum_{n\in\mathbb{N}} \partial^\alpha T_n$ converge dans $\mathcal{D}'(\Omega)$ et on a*

$$\partial^\alpha T = \sum_{n\in\mathbb{N}} \partial^\alpha T_n.$$

Dans $\mathcal{D}'(\Omega)$, on peut donc dériver sous le signe somme sans se poser de questions.

Nous admettrons le résultat suivant :

Théorème 7.21. *Soit $(T_n)_{n\in\mathbb{N}}$ une suite de distributions dans Ω. On suppose que pour tout $\phi \in \mathcal{D}(\Omega)$, la suite $(\langle T_n, \phi \rangle)_{n\in\mathbb{N}}$ converge vers une limite l_ϕ. Alors, l'application T définie par*

$$\langle T, \phi \rangle = l_\phi$$

est une distribution sur Ω.

Ce résultat permet de faire l'économie de la vérification de la continuité de l'application T (qui est manifestement une forme linéaire sur $\mathcal{D}'(\Omega)$). C'est l'un des résultats qui rend les distributions si faciles à manier.

7.6 Dérivation en dimension 1

Dans cette section, nous revenons sur la dérivation des distributions (qui a été introduite dans la définition 6.8) afin de faire le lien avec la notion usuelle de dérivation pour les fonctions C^1 puis de montrer comment la dérivation au sens des distributions permet de définir la dérivée d'une fonction qui n'est que C^1 par morceaux.

7.6.1 Cas des fonctions C^1

Soit $f \in C^1(]a,b[)$; f et $\dfrac{df}{dx}$ sont dans $L^1_{loc}(]a,b[)$, donc dans $\mathcal{D}'(]a,b[)$. On a

$$\begin{aligned}\langle \frac{d}{dx}T_f, \phi\rangle &= -\int_a^b f\frac{d\phi}{dx}\\ &= \int_a^b \frac{df}{dx}\phi\\ &= \langle T_{\frac{df}{dx}}, \phi\rangle.\end{aligned}$$

On en déduit que

$$\frac{d}{dx}T_f = T_{\frac{df}{dx}},$$

i.e. la dérivation au sens des distributions coïncide avec la dérivation au sens usuel pour les fonctions de classe C^1.

7.6.2 Cas des fonctions C^1 par morceaux

Définition 7.22. *On dit qu'une fonction f est* de classe C^k par morceaux *sur $]a,b[$ si pour tout intervalle compact $[\alpha,\beta]$ inclus dans $]a,b[$, il existe un nombre fini de points $\alpha = a_0 < a_1 < \cdots < a_N < a_{N+1} = \beta$ tels que :*
- *dans chacun des intervalles $]a_i, a_{i+1}[$, f est de classe C^k ;*
- *f et ses dérivées jusqu'à l'ordre k sont prolongeables par continuité à droite et à gauche en les points $a_1, \cdots, a_N$.*

On notera $f(a_i+0)$ la limite à droite de f en a_i et $f(a_i-0)$ la limite à gauche de f en a_i.

Remarquons qu'une fonction de classe C^1 par morceaux sur $]a,b[$ a un nombre au plus dénombrable de points de discontinuité (le montrer en exercice).

Exemples et contre-exemples :
- une fonction en escalier est C^1 par morceaux ;
- la fonction $x \mapsto |x|$ est C^1 par morceaux sur $\mathbb{R}$;
- la fonction $x \mapsto \tan x$ n'est pas C^1 par morceaux sur $\mathbb{R}$;
- la fonction $x \mapsto \sqrt{|x|}$ est continue mais n'est pas C^1 par morceaux sur $\mathbb{R}$.

Théorème 7.23. *Soit f de classe C^1 par morceaux sur $]a,b[$. On a, avec les notations de la définition ci-dessus, la* formule des sauts *suivante :*

$$f' = f'_{\mathrm{reg}} + \sum_{i\in\mathcal{I}} [f(c_i+0) - f(c_i-0)]\,\delta_{c_i},$$

où on a noté f' la dérivée de f au sens des distributions, $(c_i)_{i\in\mathcal{I}}$ l'ensemble des points de discontinuité de f sur $]a,b[$ (notons que $\mathcal{I}$ est un ensemble au plus dénombrable dont les seuls points d'accumulation éventuels sont a et b), f'_{reg} la fonction continue par morceaux qui est la dérivée (au sens usuel) de f en dehors des points c_i et δ_{c_i} la masse de Dirac en c_i.

Preuve. Ce résultat s'obtient par simple intégration par parties. □

7.6.3 Equations différentielles linéaires dans $\mathcal{D}'(]a,b[)$

Théorème 7.24. *Soit $]a,b[$ un intervalle ouvert.*

1. *Les distributions T sur $]a,b[$ vérifiant $T'=0$ dans $\mathcal{D}'(]a,b[)$ sont les fonctions constantes ;*
2. *Pour tout $S \in \mathcal{D}'(]a,b[)$, il existe $T \in \mathcal{D}'(]a,b[)$ telle que $T' = S$.*

Preuve. Remarquons d'abord que si $\phi \in \mathcal{D}(]a,b[)$, ϕ possède une primitive dans $\mathcal{D}(]a,b[)$ si et seulement si $\int_a^b \phi = 0$, la primitive étant alors $\int_a^x \phi$. Soit $\rho \in \mathcal{D}(]a,b[)$ telle que $\int_a^b \rho = 1$. Comme $\phi - (\int_a^b \phi)\rho$ est d'intégrale nulle, il existe $\psi \in \mathcal{D}(]a,b[)$ telle que

$$(1) \qquad \psi' = \phi - \left(\int_a^b \phi\right)\rho.$$

Soit maintenant $T \in \mathcal{D}'(]a,b[)$ vérifiant $T' = 0$. On a

$$\langle T,\phi\rangle = \left(\int_a^b \phi\right)\langle T,\rho\rangle + \langle T,\psi'\rangle = \left(\int_a^b \phi\right)\langle T,\rho\rangle - \langle T',\psi\rangle = \left(\int_a^b \phi\right)\langle T,\rho\rangle.$$

En appelant C la constante $\langle T,\rho\rangle$, on a donc $T = C$. Soit maintenant $S \in \mathcal{D}'(]a,b[)$. Posons pour $\phi \in \mathcal{D}(]a,b[)$,

$$\langle T,\phi\rangle = -\langle S,\psi\rangle,$$

où ψ est définie de manière unique dans $\mathcal{D}(]a,b[)$ par (1). On vérifiera à titre d'exercice que T est une distribution et que $T' = S$. □

Le théorème 7.24 permet notamment de montrer des résultats d'unicité sur des équations différentielles dans $\mathcal{D}'$. L'exemple de l'équation différentielle $xT' + T = 0$ est traité en exercice (exercice 7.14).

7.6.4 Rapport entre la dérivée au sens usuel et la dérivée au sens des distributions

Pour une fonction $f \in L^1_{loc}$, les rapports entre dérivée au sens des distributions et dérivée au sens usuel peuvent être complexes :

1. pour une fonction C^1, il y a identité entre les deux concepts ;
2. pour une fonction C^1 par morceaux, la dérivée usuelle est définie presque partout, mais ne rend pas correctement compte des variations de f puisqu'elle ignore les sauts : c'est la dérivée au sens des distributions qui est le bon concept ;
3. lorsque f est dérivable en tout point sans être de classe C^1, la situation est plus subtile : lorsque la dérivée est localement sommable, les deux concepts coïncident ; dans le cas contraire, la dérivée au sens des distributions est en quelque sorte une "partie finie" de la dérivée usuelle (cf. exercice 7.15).

7.7 Dérivation en dimension quelconque

7.7.1 Théorème de Schwarz

Théorème 7.25. *(Théorème de Schwarz) Soit Ω un ouvert de $\mathbb{R}^d$ et $T \in \mathcal{D}'(\Omega)$. Soient $\alpha \in \mathbb{N}^d$ et $\beta \in \mathbb{N}^d$ deux multi-indices. On a*

$$\partial^\alpha \partial^\beta T = \partial^\beta \partial^\alpha T = \partial^{\alpha+\beta} T.$$

Preuve. Immédiate. □

7.7.2 Cas des fonctions C^1

Comme pour le cas de la dimension 1, on vérifie aisément que la dérivation au sens des distributions coïncide avec la dérivation usuelle pour les fonctions de classe C^1. En particulier, si Ω est un ouvert de $\mathbb{R}^d$ et $f \in C^1(\Omega)$, on a

$$\forall 1 \leq i \leq d, \qquad \frac{\partial}{\partial x_i} T_f = T_{\frac{\partial f}{\partial x_i}}.$$

7.7.3 Cas des fonctions C^1 par morceaux

Rappelons que la formule d'intégration par parties

$$\int_a^b f'(x)\, g(x)\, dx = (f(b)g(b) - f(a)g(a)) - \int_a^b f(x)\, g'(x)\, dx,$$

bien connue en dimension 1, s'étend de la façon suivante à la dimension $d \geq 2$: soit Ω un ouvert borné suffisamment régulier de $\mathbb{R}^d$ et f et g dans $C^1(\overline{\Omega})$. On a

$$\int_\Omega \frac{\partial f}{\partial x_i}(x)\, g(x)\, dx = \int_{\partial\Omega} f(x)\, g(x)\, (n(x) \cdot e_i)\, d\sigma(x) - \int_\Omega f(x)\, \frac{\partial g}{\partial x_i}(x)\, dx \qquad (7.4)$$

où $n(x)$ désigne le vecteur normal sortant au point $x \in \partial\Omega$ et $d\sigma$ la mesure de Lebesgue sur $\partial\Omega$. Un corollaire immédiat de cette formule d'intégration par parties, que nous utiliserons au chapitre 9, est la formule de Green

$$\int_\Omega \Delta f(x)\, g(x)\, dx = -\int_\Omega \nabla f(x) \cdot \nabla g(x)\, dx + \int_{\partial\Omega} \frac{\partial f}{\partial n}(x)\, g(x)\, d\sigma(x) \qquad (7.5)$$

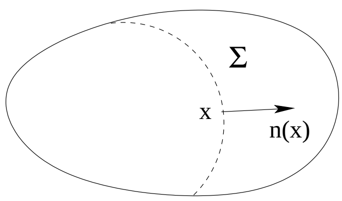
où $\frac{\partial f}{\partial n}(x) = \nabla f(x) \cdot n(x)$.

Proposition 7.26. *(Formule des sauts dans l'espace). Soit Ω un ouvert borné de $\mathbb{R}^d$ et $\Sigma \subset \Omega$ une hypersurface de codimension 1 régulière (de classe C^1), orientée fermée dans Ω et sans bord. Soit f de classe C^1 sur $\Omega \setminus \Sigma$ telle que f et ∇f soit prolongeables par continuité de part et d'autre de Σ. Pour $x \in \Sigma$, on note*

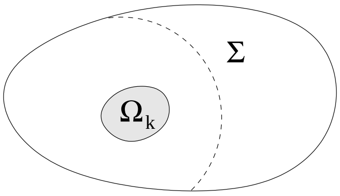
$$[f](x) = f_{ext}(x) - f_{int}(x) = \lim_{\epsilon \to 0^+} f(x + \epsilon n(x)) - \lim_{\epsilon \to 0^+} f(x - \epsilon n(x)).$$

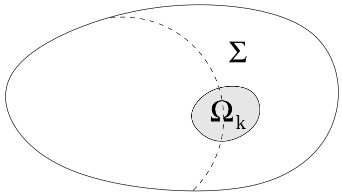
On a alors

$$\frac{\partial f}{\partial x_i} = \frac{\partial f_{reg}}{\partial x_i} + \delta_\Sigma \otimes [f](n \cdot e_i)\, d\sigma.$$



Preuve. Soit K un compact de Ω. On considère un recouvrement d'ouverts de K fini $K \subset \bigcup_{k=1}^n \Omega_k$ tel qu'on soit pour tout k dans l'une ou l'autre de ces deux situations (cf. figure ci-dessus) :
- ou bien $\Omega_k \cap \Sigma = \emptyset$;
- ou bien Σ coupe Ω_k en deux.





De par le lemme 6.11 de partition de l'unité, il existe $\alpha_1, \cdots, \alpha_n$ dans $\mathcal{D}(\Omega)$ telles que
- $\mathrm{Supp}(\alpha_k) \subset \Omega_k$;
- $0 \leq \alpha_k \leq 1$;
- $\sum_{i=1}^{n} \alpha_k = 1$ sur K.

Soit $\phi \in \mathcal{D}_K(\Omega)$. En utilisant la formule d'intégration par parties (7.4)

$$\begin{aligned}
\langle \frac{\partial f}{\partial x_i}, \phi \rangle &= \langle \frac{\partial f}{\partial x_i}, \sum_{k=1}^{n} \alpha_k \phi \rangle \\
&= -\sum_{k=1}^{n} \langle f, \frac{\partial(\alpha_k \phi)}{\partial x_i} \rangle \\
&= -\sum_{k=1}^{n} \int_{\Omega_k} f \frac{\partial(\alpha_k \phi)}{\partial x_i} \\
&= -\sum_{k \,/\, \Omega_k \cap \Sigma = \emptyset} \int_{\Omega_k} f \frac{\partial(\alpha_k \phi)}{\partial x_i} - \sum_{k \,/\, \Omega_k \cap \Sigma \neq \emptyset} \int_{\Omega_k} f \frac{\partial(\alpha_k \phi)}{\partial x_i} \\
&= -\sum_{k \,/\, \Omega_k \cap \Sigma = \emptyset} \int_{\Omega_k} f \frac{\partial(\alpha_k \phi)}{\partial x_i} - \sum_{k \,/\, \Omega_k \cap \Sigma \neq \emptyset} \left(\int_{\Omega_k^+} f \frac{\partial(\alpha_k \phi)}{\partial x_i} + \int_{\Omega_k^-} f \frac{\partial(\alpha_k \phi)}{\partial x_i} \right) \\
&= \sum_{k \,/\, \Omega_k \cap \Sigma = \emptyset} \int_{\Omega_k} \frac{\partial f_{reg}}{\partial x_i} \alpha_k \phi \\
&\quad - \sum_{k \,/\, \Omega_k \cap \Sigma \neq \emptyset} \left(- \int_{\Omega_k^+} \frac{\partial f_{reg}}{\partial x_i} \alpha_k \phi + \int_{\partial \Omega_k^+} f_{ext}(x) \alpha_k \phi \, (-n(x) \cdot e_i) \right. \\
&\quad \left. - \int_{\Omega_k^-} \frac{\partial f_{reg}}{\partial x_i} \alpha_k \phi + \int_{\partial \Omega_k^+} f_{int}(x) \alpha_k \phi \, (n(x) \cdot e_i) \right) \\
&= \int_{\Omega} \frac{\partial f_{reg}}{\partial x_i} \phi + \int_{\Sigma} [f](n \cdot e_i) \phi \, d\sigma.
\end{aligned}$$

□

Remarque 7.27. La formule des sauts dans l'espace permet de démontrer très simplement la condition de Rankhine-Hugoniot (cf. exercice 7.18), qui caractérise les solutions faibles de classe C^1 par morceaux d'équations hyperboliques du type

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0.$$

7.8 Multiplication par des fonctions C^∞

Définition 7.28. *Soit $T \in \mathcal{D}'(\Omega)$ et $g \in C^\infty(\Omega)$. On définit la distribution gT par la formule*

$$\forall \phi \in \mathcal{D}(\Omega), \qquad \langle gT, \phi \rangle = \langle T, g\phi \rangle.$$

On vérifiera en exercice que gT est bien une distribution.

Remarque 7.29. Attention ! On ne peut pas multiplier deux distributions. Ainsi la fonction $f(x) = \frac{1}{\sqrt{|x|}}$ est dans $L^1_{loc}(\mathbb{R})$ et définit donc une distribution mais il n'en va pas de même pour la fonction f^2 ($f^2(x) = \frac{1}{|x|}$ n'est pas dans $L^1_{loc}(\mathbb{R})$). De façon générale, le produit T_1T_2 pour $T_1 \in \mathcal{D}'(\Omega)$ et $T_2 \in \mathcal{D}'(\Omega)$ n'a aucun sens (cf. exercice 7.9).

7.9 Exercices

▷ **7.1.** Quel est le support de δ_a, de vp $\left(\frac{1}{x}\right)$?

▷ **7.2.** Montrer que la valeur principale de la fonction $\frac{1}{x}$, définie par la formule (7.2), est une distribution d'ordre 1. *Indication : montrer qu'elle est d'ordre inférieur ou égal à 1, puis qu'elle est exactement d'ordre 1 ; pour cela, on pourra calculer son action sur la suite $\phi_j(x) = \phi_0(x)\, atan(jx)$ où $\phi_0 \in \mathcal{D}(\mathbb{R})$ est paire et vaut 1 au voisinage de l'origine.*

▷ **7.3.** Montrer que la fonction $\mathrm{Log}(|x|)$ définit une distribution dans $\mathcal{D}'(\mathbb{R})$ et calculer sa dérivée au sens des distributions.

▷ **7.4.** Montrer que la partie finie de la fonction $H(x)x^\alpha$ (cf. section 7.3.2) définit une distribution d'ordre égal à la partie entière de α.

▷ **7.5.** Montrer que $x\,\mathrm{vp}\left(\frac{1}{x}\right) = 1$ et que $x\,\delta_0 = 0$.

▷ **7.6.** Soit $a \in \mathbb{R}^d$, $p > 0$ et $e \in \mathbb{R}^d$ tel que $|e| = 1$. Identifier la limite dans $\mathcal{D}'(\mathbb{R}^d)$ de la famille de distributions

$$T_\epsilon = \frac{p}{\epsilon}\delta_{a+\epsilon e} - \frac{p}{\epsilon}\delta_a$$

lorsque ϵ tend vers 0 (cf. section 7.3.1).

▷ **7.7.** Soit $g \in L^1_{loc}(\mathbb{R}^2)$. Identifier la limite dans $\mathcal{D}'(\mathbb{R}^3)$ de la famille de distributions

$$T_\epsilon = SC(g/\epsilon, \epsilon) - SC(-g/\epsilon, 0)$$

lorsque ϵ tend vers 0, où $SC(h, x_0)$ désigne la distribution de simple couche de densité $h \in L^1_{loc}(\mathbb{R}^2)$ portée par le plan d'équation $x = x_0$, définie par

$$\forall \phi \in \mathcal{D}(\mathbb{R}^3), \qquad \langle SC(h, x_0), \phi \rangle = \int_{\mathbb{R}^2} h(y,z)\,\phi(x_0, y, z)\,dz.$$

Avant de faire cet exercice, on pourra (re)lire la section 7.3.1.

▷ **7.8.** Résoudre dans $\mathcal{D}'(\mathbb{R})$ l'équation $xT = 0$ (i.e. chercher l'ensemble des $T \in \mathcal{D}'(\mathbb{R})$ tels que $xT = 0$ dans $\mathcal{D}'(\mathbb{R})$).

▷ **7.9.** Expliquer pourquoi les expressions

$$\left(x\,\mathrm{vp}\left(\frac{1}{x}\right)\right)\delta$$

et

$$(x\,\delta)\,\mathrm{vp}\left(\frac{1}{x}\right).$$

sont bien définies dans $\mathcal{D}'(\mathbb{R})$ et les calculer. Conclure quant à la possibilité de définir le produit de deux distributions arbitraires.

▷ **7.10.**

1. Soit $f_n(x) = e^{inx}$. Etudier la convergence presque partout et la convergence dans $\mathcal{D}'$. Réponse f_n ne converge pas presque partout mais converge vers 0 dans $\mathcal{D}'(\mathbb{R})$.
2. Soit $f_n = \displaystyle\sum_{k=1}^{n} \chi_{1/n^2}(x - \frac{1}{k})$ (où χ désigne une approximation de l'identité). Etudier la convergence presque partout et la convergence dans $\mathcal{D}'$. Réponse : $f_n \to 0$ presque partout mais f_n ne converge pas dans $\mathcal{D}'(\mathbb{R})$.
3. Soit $f_n(x) = \frac{1}{\sigma_n\sqrt{2\pi}} e^{-x^2/2\sigma_n^2}$ avec $\sigma_n \to 0$. Etudier la convergence presque partout et la convergence dans $\mathcal{D}'$. Réponse : $f_n \to 0$ presque partout et $f_n \to \delta$ dans $\mathcal{D}'(\mathbb{R})$.

▷ **7.11.** Montrer que

$$\sum_{n\in\mathbb{Z}} \delta_n \text{ converge dans } \mathcal{D}'(\mathbb{R}).$$

▷ **7.12.** La série

$$\sum_{n=1}^{+\infty} \delta_{1/n}$$

converge-t-elle dans $\mathcal{D}'(\mathbb{R})$? Même question avec $\mathcal{D}'(]0, +\infty[)$.

▷ **7.13.** Pour tout $\phi \in \mathcal{D}(\mathbb{R})$, on pose

$$\langle\frac{1}{x+i0}, \phi\rangle = \lim_{\substack{\epsilon\to 0\\ \epsilon>0}} \int_{\mathbb{R}} \frac{\phi(x)}{x+i\epsilon}\,dx \quad \text{ et } \quad \langle\frac{1}{x-i0}, \phi\rangle = \lim_{\substack{\epsilon\to 0\\ \epsilon>0}} \int_{\mathbb{R}} \frac{\phi(x)}{x-i\epsilon}\,dx$$

Montrer que $\frac{1}{x+i0}$ et $\frac{1}{x-i0}$ définissent des distributions d'ordre 1 et les exprimer en fonction de $\mathrm{vp}(\frac{1}{x})$. Vérifier en particulier que

$$\frac{1}{x+i0} - \frac{1}{x-i0} = 2i\pi\delta.$$

▷ **7.14.** Résoudre dans $\mathcal{D}'(\mathbb{R})$ l'équation différentielle $xT' + T = 0$. (Considérer $S = xT$).

▷ **7.15.**

1. Soit $f \in L^1_{loc}$ et $F(x) = \int_0^x f$. Montrer que la dérivée de F au sens des distributions est égale à f.
2. Soit $f(x) = \mathrm{Log}(|x|)$. Vérifier que f définit une distribution et calculer sa dérivée.

3. Soit $-1 < \alpha < 0$. Montrer que la dérivée au sens des distributions de la fonction $H(x)x^\alpha$ est égale à $\alpha \mathrm{Pf}(H(x)x^{\alpha-1})$.

▷ **7.16.** Ecrire la dérivée seconde d'une fonction C^2 par morceaux.

$$f'' = f''_{\text{reg}} + \sum_{i=1}^{N} [f'(a_i+0) - f'(a_i-0)]\, \delta_{a_i} + \sum_{i=1}^{N} [f(a_i+0) - f(a_i-0)]\, \delta'_{a_i}$$

▷ **7.17.** Soit ρ une densité de charge dans $\mathbb{R}^3$ régulière et finie (par exemple C^∞ à support compact). Le potentiel électrostatique V engendré par ρ est donné par la formule

$$V(x) = \int_{\mathbb{R}^3} \frac{\rho(y)}{4\pi\epsilon_0 |x-y|}\, dy$$

où ϵ_0 désigne la permitivité diélectrique du vide et vérifie au sens usuel l'équation de Poisson

$$-\Delta V = \rho/\epsilon_0.$$

Pour une charge ponctuelle q située en $\bar{x}$, on a

$$V(x) = \frac{q}{4\pi\epsilon_0 |x-\bar{x}|}.$$

Le but de cet exercice est de montrer que

$$-\Delta V = \frac{q}{\epsilon_0}\, \delta_{\bar{x}} \qquad \text{dans } \mathcal{D}'(\mathbb{R}^3).$$

1. Montrer que $-\Delta \frac{1}{4\pi|x|} = 0$ (au sens classique) dans $\mathbb{R}^3 \setminus \{0\}$;
2. Montrer que si Ω est un ouvert borné régulier de $\mathbb{R}^3$ et si f et g sont dans $C^2(\overline{\Omega})$, alors
$$\int_\Omega (f\, \Delta g - g\, \Delta f) = \int_{\partial\Omega} \left(f\, \frac{\partial g}{\partial n} - \frac{\partial f}{\partial n}\, g \right).$$
3. Ecrire cette formule pour $\Omega = B_0(R) \setminus \overline{B_0(r)}$ avec $0 < r < R < +\infty$, $f = \phi \in \mathcal{D}(\mathbb{R}^3)$ et $g = \frac{1}{4\pi|x|}$;
4. Passer à la limite dans cette dernière expression en faisant tendre r vers 0 et R vers $+\infty$. Conclure.

▷ **7.18.** L'objet de cet exercice est de démontrer la relation de Rankhine-Hugoniot en dimension 1 d'espace (cf. cours de calcul scientifique). On se place dans le demi-plan $\Omega = \{(x,t) \in \mathbb{R} \times \mathbb{R}^{+*}\}$ (cf. figure ci-dessous) et on se donne une fonction $u(x,t)$ de classe C^1 dans $\Omega \setminus \Gamma$, telle que u soit prolongeable par continuité de part et d'autre de Γ. On suppose que Γ est une courbe qui peut être paramétrée par

$$\Gamma = \big\{(\phi(t), t), \quad t \in \mathbb{R}^{+*}\big\}$$

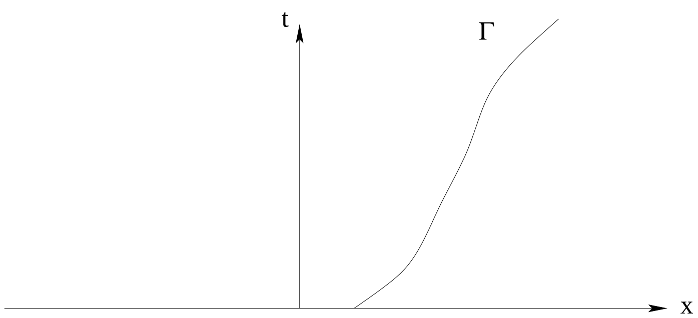
avec ϕ de classe C^1 sur $\mathbb{R}^{+*}$.

On suppose enfin que u vérifie au sens des distributions l'équation hyperbolique

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0 \tag{7.6}$$

sur Ω, où f est une fonction de classe C^1 de $\mathbb{R}$ dans $\mathbb{R}$.

1. Vérifier que l'équation (7.6) a bien un sens dans $\mathcal{D}'(\Omega)$.
2. Soit (x_*, t_*) un point de la courbe de choc Γ. On note u_g et u_d les prolongements par continuité de u à gauche et droite (respectivement) du point (x_*, t_*). Soit enfin $s = \dfrac{d\phi}{dt}(t_*)$. Etablir une relation entre u_g, u_d ; $f(u_g)$, $f(u_d)$ et s.


t
Γ
x

Chapitre 8

Espaces de Sobolev

Les espaces de Sobolev jouent un rôle central dans la théorie des EDP. Ce sont des sous-espaces de l'espace vectoriel des distributions qui possèdent une structure hilbertienne leur conférant des propriétés supplémentaires issues des théorèmes fondamentaux de l'analyse hilbertienne (inégalité de Cauchy-Schwarz, théorème de projection, théorème de Riesz). Ces propriétés permettront d'obtenir très simplement des résultats puissants sur les EDP. Signalons également que certains espaces de Sobolev s'introduisent de façon naturelle en considérant le modèle physique sous-jacent à une EDP : dans nombre de cas, le plus gros espace fonctionnel permettant de donner un sens mathématique à l'énergie d'un système est précisément un espace de Sobolev.

8.1 Espaces $H^k(\Omega)$

Définition 8.1. *Soit Ω un ouvert de $\mathbb{R}^d$ et soit k un entier positif. On note $H^k(\Omega)$ l'ensemble des fonctions de $L^2(\Omega)$ dont toutes les dérivées (au sens des distributions) jusqu'à l'ordre k sont dans $L^2(\Omega)$*

$$H^k(\Omega) = \left\{ f \in L^2(\Omega), \quad \forall\alpha,\ |\alpha| \leq k, \quad \partial_\alpha f \in L^2(\Omega) \right\}.$$

Théorème 8.2. *$H^k(\Omega)$ est un espace vectoriel. Muni du produit scalaire*

$$(f, g)_{H^k} = \sum_{|\alpha| \leq k} \int_\Omega \partial^\alpha f(x)\, \partial^\alpha g(x)\, dx,$$

$H^k(\Omega)$ est un espace de Hilbert. Sa norme est notée $\|\cdot\|_{H^k}$.

Preuve. Il est clair que $H^k(\Omega)$ est un espace vectoriel et que $(\cdot,\cdot)_{H_k}$ est un produit scalaire sur $H^k(\Omega)$. Il reste à vérifier que $H^k(\Omega)$ est complet pour la norme $\|\cdot\|_{H^k}$. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy de $H^k(\Omega)$. Comme

$$\|f_p - f_q\|_{L^2} \leq \|f_p - f_q\|_{H^k},$$

on voit que (f_n) est de Cauchy dans L^2, donc converge dans L^2 vers $f \in L^2$. De même pour tout $|\alpha| \leq k$,

$$\|\partial^\alpha f_p - \partial^\alpha f_q\|_{L^2} \leq \|f_p - f_q\|_{H^k}.$$

Donc $(\partial^\alpha f_n)$ converge dans L^2 vers une fonction g_α. Par ailleurs, la convergence dans L^2 implique la convergence dans $\mathcal{D}'$. Donc

$$f_n \longrightarrow f \qquad \text{dans } \mathcal{D}'$$

et

$$\partial^\alpha f_n \longrightarrow g_\alpha \qquad \text{dans } \mathcal{D}'.$$

De la première assertion, on tire

$$\partial^\alpha f_n \longrightarrow \partial^\alpha f \qquad \text{dans } \mathcal{D}'.$$

Par unicité de la limite dans $\mathcal{D}'(\Omega)$, on trouve $\partial^\alpha f = g_\alpha \in L^2$. Donc $f \in H^k$ et comme $\partial^\alpha f_n \longrightarrow \partial^\alpha f$ dans L^2, il s'en suit que $f_n \longrightarrow f$ dans H^k. □

8.2 Espace $H_0^1(\Omega)$

Proposition 8.3.

1. *Si $\Omega = \mathbb{R}^d$, $\mathcal{D}(\mathbb{R}^d)$ est dense dans $H^1(\mathbb{R}^d)$.*
2. *Si $\Omega \subset \mathbb{R}^d$, $\Omega \neq \mathbb{R}^d$, $\mathcal{D}(\Omega)$ n'est pas dense dans $H^1(\Omega)$.*

Preuve.

1. La première assertion, assez technique, se montre par troncature et régularisation.
2. On se contente de prouver ce résultat en dimension 1 dans le cas où $\Omega =]0,1[$. Soit $\phi \in \mathcal{D}(]0,1[)$. On a
$$\forall x \in]0,1[, \qquad \phi(x) = \int_0^x \phi'(t)\,dt.$$
Donc
$$\forall x \in]0,1[, \quad |\phi(x)| \leq \int_0^x |\phi'(t)|\,dt \leq \left(\int_0^x |\phi'(t)|^2\,dt\right)^{1/2} \left(\int_0^x 1^2\,dt\right)^{1/2} \leq \|\phi'\|_{L^2}.$$
Il en résulte que
$$\|\phi\|_{L^2} = \left(\int_0^1 |\phi(t)|^2\,dt\right)^{1/2} \leq \|\phi'\|_{L^2}.$$
On a donc finalement
$$\forall \phi \in \mathcal{D}(]0,1[), \qquad \|\phi\|_{L^2} \leq \|\phi'\|_{L^2}.$$
Si $\mathcal{D}(]0,1[)$ était dense dans $H^1(]0,1[)$, cette inégalité resterait vraie pour tout $\phi \in H^1(]0,1[)$. Or en prenant par exemple $f \equiv 1$ sur $]0,1[$, on voit que cette inégalité est violée.

□

Définition 8.4. *On note $H_0^1(\Omega)$ la fermeture de $\mathcal{D}(\Omega)$ dans $H^1(\Omega)$.*

Remarque 8.5. En vertu de la proposition ci-dessus, $H_0^1(\mathbb{R}^d) = H^1(\mathbb{R}^d)$, mais $H_0^1(\Omega)$ n'est qu'un sous-espace vectoriel de $H^1(\Omega)$ (strictement inclus dans $H^1(\Omega)$) si $\Omega \subset \mathbb{R}^d$ avec $\Omega \neq \mathbb{R}^d$.

Proposition 8.6. *L'espace $H_0^1(\Omega)$ est un espace vectoriel. C'est un espace de Hilbert pour le produit scalaire $(\cdot,\cdot)_{H^1}$.*

Preuve. Il est clair que $H_0^1(\Omega)$ est un espace vectoriel (c'est la fermeture d'un espace vectoriel pour une certaine norme). Le produit scalaire de $H^1(\Omega)$ restreint à $H_0^1(\Omega)$ définit évidemment un produit scalaire sur $H_0^1(\Omega)$. Comme $H_0^1(\Omega)$ est fermé dans un espace complet, il est complet. □

Enonçons maintenant un résultat important d'analyse fonctionnelle dont nous aurons besoin par la suite (cf. section 9.3) :

Théorème 8.7. *(Inégalité de Poincaré). Soit Ω un ouvert borné de $\mathbb{R}^d$. Il existe une constante C_Ω telle que*

$$\forall u \in H_0^1(\Omega), \qquad \|u\|_{L^2} \leq C_\Omega \, \|\nabla u\|_{L^2}.$$

Dans le langage de l'analyse fonctionnelle, on dit que la norme L^2 d'une fonction de $H_0^1(\Omega)$ (Ω borné) est *contrôlée* par la norme L^2 de son gradient.

Preuve du théorème 8.7. L'ouvert Ω étant borné, on peut trouver $L > 0$ tel que $\Omega \subset [-L, L]^n$. Soit $\phi \in \mathcal{D}(\Omega)$ et ψ le prolongement de ϕ par 0 à tout $\mathbb{R}^d$. Il est clair que $\psi \in C^\infty(\mathbb{R}^d)$ et que

$$\forall x \in [-L, L]^n, \qquad \psi(x) = \int_{-L}^{x_1} \frac{\partial \psi}{\partial x_1}(t, x_2, \cdots, x_n)\, dt.$$

Il en résulte que

$$\begin{aligned} |\psi(x)|^2 &= \left(\int_{-L}^{x_1} \frac{\partial \psi}{\partial x_1}(t, x_2, \cdots, x_n)\, dt \right)^2 \\ &\leq \left(\int_{-L}^{x_1} \left| \frac{\partial \psi}{\partial x_1}(t, x_2, \cdots, x_n) \right|^2 dt \right) \left(\int_{-L}^{x_1} 1^2\, dt \right) \\ &\leq 2L \int_{-L}^{L} \left| \frac{\partial \psi}{\partial x_1}(t, x_2, \cdots, x_n) \right|^2 dt. \end{aligned}$$

En intégrant l'inégalité obtenue sur $[-L, L]^n$, on obtient

$$\int_{[-L,L]^n} |\psi|^2 \leq 4L^2 \int_{[-L,L]^n} \left| \frac{\partial \psi}{\partial x_1} \right|^2 \leq 4L^2 \int_{[-L,L]^n} |\nabla \psi|^2$$

Comme $\mathrm{Supp}(\psi) = \mathrm{Supp}(\phi) \subset \Omega$, $\mathrm{Supp}(\nabla\psi) = \mathrm{Supp}(\nabla\phi) \subset \Omega$, et comme $\phi = \psi$ et $\nabla\phi = \nabla\psi$ sur Ω, il en résulte que

$$\forall \phi \in \mathcal{D}(\Omega), \qquad \|\phi\|_{L^2} \leq 2L \, \|\nabla \phi\|_{L^2}. \tag{8.1}$$

Il est clair par ailleurs que les applications

$$\begin{array}{rcl} H^1(\Omega) & \longrightarrow & \mathbb{R} \\ u & \mapsto & \|u\|_{L^2} \end{array} \quad \text{et} \quad \begin{array}{rcl} H^1(\Omega) & \longrightarrow & \mathbb{R} \\ u & \mapsto & \|\nabla u\|_{L^2} \end{array}$$

sont continues. Comme $\mathcal{D}(\Omega)$ est dense dans $H_0^1(\Omega)$ pour la norme H^1, il en résulte que l'inégalité (8.1) reste valable pour tous les éléments de $H_0^1(\Omega)$. $\square$

Terminons cette section par quelques remarques sur la définition de $H_0^1(\Omega)$ *via* la notion de trace : on dit parfois que $H_0^1(\Omega)$ est l'espace des fonctions H^1 "nulles au bord". Ceci appelle quelques commentaires. Introduisons pour commencer la notion de trace. Pour une fonction $u \in C^0(\overline{\Omega})$, la trace de u sur $\partial\Omega$ est définie par

$$\begin{array}{rrcl} \gamma(u) \,:\, & \partial\Omega & \longrightarrow & \mathbb{R} \\ & x & \mapsto & u(x). \end{array}$$

En d'autres termes, $\gamma(u) = u|_{\partial\Omega}$. Notons que l'application trace

$$\begin{array}{rrcl} \gamma \,:\, & C^0(\overline{\Omega}) & \longrightarrow & C^0(\partial\Omega) \\ & u & \mapsto & \gamma(u) \end{array}$$

est linéaire et continue. La question qui se pose maintenant est la suivante : peut-on étendre la notion de trace à des fonctions moins régulières ? Voici quelques éléments de réponse :

1. On ne peut pas définir la trace d'une fonction de $L^2(\Omega)$.

   Considérons en effet, à titre d'exemple, la fonction $u : x \mapsto \sin(1/x)$ sur $\Omega =]0,1[$. L'ensemble $\partial\Omega$ est alors composé des deux points 0 et 1. La fonction u est continue en 1 et on peut donc définir sa trace en 1 (c'est le réel $\sin 1$) ; en revanche, l'ensemble des valeurs d'adhérence de u en 0 est l'intervalle $[-1,1]$; il n'y a donc pas de façon naturelle de définir la trace de u (la valeur de u) en 0 ;
2. En dimension 1, une fonction qui est dans $H^1(]a,b[)$ est dans $C^0([a,b])$. La démonstration de cette assertion fait l'objet de l'exercice 8.6. On peut donc définir la trace en a et en b d'une fonction de $H^1(]a,b[)$;
3. En dimension ≥ 2, une fonction qui est dans $H^1(\Omega)$ n'est pas nécessairement continue (cf. exercice 8.2). On peut cependant définir la trace sur $\partial\Omega$ d'une fonction de $H^1(\Omega)$. Plus précisément, il existe une application linéaire et continue

$$\begin{array}{rccc} \gamma \,:\, & H^1(\Omega) & \longrightarrow & L^2(\partial\Omega) \\ & u & \mapsto & \gamma(u) \end{array}$$

   vérifiant $\forall u \in H^1(\Omega) \cap C^0(\overline{\Omega})$, $\gamma(u) = u|_{\partial\Omega}$. En fait, l'application γ prend ses valeurs dans un sous-espace de $L^2(\partial\Omega)$ noté $H^{1/2}(\partial\Omega)$. Cette notation, qui se base sur la théorie des espaces de Sobolev à exposant fractionnaire, sera éclairée dans la section 10.4.2.

Ce troisième point permet d'énoncer le résultat suivant :

Proposition 8.8. *On a la caractérisation suivante :*

$$H_0^1(\Omega) = \left\{ u \in H^1(\Omega), \quad \gamma(u) = 0 \right\}.$$

8.3 Espace $H^{-1}(\Omega)$

Définition 8.9. *Soit Ω un ouvert de $\mathbb{R}^d$. On note $H^{-1}(\Omega)$ l'espace vectoriel des distributions $T \in \mathcal{D}'(\Omega)$ telles qu'il existe une constante C pour laquelle*

$$\forall \phi \in \mathcal{D}(\Omega), \qquad |\langle T, \phi \rangle| \leq C \|\phi\|_{H^1}.$$

Remarque 8.10. Il est clair que $L^2(\Omega) \subset H^{-1}(\Omega)$. En effet, si $f \in L^2(\Omega)$, on a

$$\forall \phi \in \mathcal{D}(\Omega), \qquad |\langle f, \phi \rangle| = \left| \int_\Omega f\phi \right| \leq \|f\|_{L^2} \|\phi\|_{L^2} \leq \|f\|_{L^2} \|\phi\|_{H^1}.$$

Théorème 8.11. *On peut identifier $H^{-1}(\Omega)$ au dual topologique de $H_0^1(\Omega)$.*

Preuve. Soit $T \in H^{-1}(\Omega)$. L'application linéaire

$$\phi \mapsto \langle T, \phi \rangle$$

est continue sur $\mathcal{D}(\Omega)$ *muni de la norme* H^1. Par conséquent, cette application se prolonge en une forme linéaire continue sur $H_0^1(\Omega)$ (cf. théorème 1.24) notée

$$\phi \mapsto \langle T, \phi \rangle_{H^{-1}, H_0^1},$$

qui vérifie en particulier

$$\forall \phi \in \mathcal{D}(\Omega), \qquad \langle T, \phi \rangle_{H^{-1}, H_0^1} = \langle T, \phi \rangle.$$

De plus, ce prolongement est unique car $\mathcal{D}(\Omega)$ est dense dans $H_0^1(\Omega)$ pour la norme de H^1. On peut donc associer à tout $T \in H^{-1}(\Omega)$ un élément du dual topologique de $H_0^1(\Omega)$ (i.e. de l'espace vectoriel des formes linéaires continues sur $H_0^1(\Omega)$). On définit ainsi

$$\begin{array}{rcl} \alpha \,:\, H^{-1}(\Omega) & \longrightarrow & (H_0^1(\Omega))' \\ T & \mapsto & \langle T, \cdot \rangle_{H^{-1},H_0^1}. \end{array}$$

Réciproquement, soit $L \in (H_0^1(\Omega))'$. Il existe une constante C telle que

$$\forall \phi \in H_0^1(\Omega), \qquad |L(\phi)| \leq C \|\phi\|_{H^1}.$$

En notant $L_{|\mathcal{D}(\Omega)}$ la restriction de L à $\mathcal{D}(\Omega) \subset H_0^1(\Omega)$, on définit une forme linéaire sur $\mathcal{D}(\Omega)$ qui vérifie

$$\forall \phi \in \mathcal{D}(\Omega), \qquad |L_{|\mathcal{D}(\Omega)}(\phi)| \leq C \|\phi\|_{H^1}.$$

Il reste à vérifier que $L_{|\mathcal{D}(\Omega)}$ est une distribution (c'est-à-dire qu'elle est continue sur $\mathcal{D}(\Omega)$ pour la topologie de $\mathcal{D}(\Omega)$). Soit donc K compact inclus dans Ω et $\phi \in \mathcal{D}_K(\Omega)$. Il vient

$$|L_{|\mathcal{D}(\Omega)}(\phi)| \leq C \|\phi\|_{H^1} \leq C \left(\|\phi\|_{L^2}^2 + \|\nabla \phi\|_{L^2}^2 \right)^{1/2}.$$

Or $\|\phi\|_{L^2} \leq \sqrt{|K|} \sup_K |\phi|$ et $\|\nabla \phi\|_{L^2} \leq \sqrt{|K|} \sup_K |\nabla \phi|$. Donc

$$|L_{|\mathcal{D}(\Omega)}(\phi)| \leq C' \sup_{|\alpha| \leq 1,\, x \in K} |\partial^\alpha \phi|.$$

Donc $L_{|\mathcal{D}(\Omega)}$ définit une distribution (d'ordre ≤ 1). On définit ainsi

$$\begin{array}{rcl} \beta \,:\, (H_0^1(\Omega))' & \longrightarrow & H^{-1}(\Omega) \\ L & \mapsto & L_{|\mathcal{D}(\Omega)}. \end{array}$$

On vérifie sans difficulté que $\alpha \circ \beta = I_{(H_0^1(\Omega))'}$ et que $\beta \circ \alpha = I_{H^{-1}(\Omega)}$. □

Proposition 8.12. *(Caractérisation des éléments de H^{-1}). Soit Ω un ouvert de $\mathbb{R}^d$. Une distribution T appartient à $H^{-1}(\Omega)$ si et seulement si il existe, pour tout $|\alpha| \leq 1$ une fonction $g_\alpha \in L^2(\Omega)$ telle que*

$$T = \sum_{|\alpha| \leq 1} \partial^\alpha g_\alpha.$$

Preuve. Il est clair que si T est de la forme

$$T = \sum_{|\alpha| \leq 1} \partial^\alpha g_\alpha$$

avec $g_\alpha \in L^2(\Omega)$, on a

$$\begin{array}{rcl} \forall \phi \in \mathcal{D}(\Omega), \qquad |\langle T, \phi \rangle| & = & |\langle \sum_{|\alpha| \leq 1} \partial^\alpha g_\alpha, \phi \rangle| \\ & \leq & |\sum_{|\alpha| \leq 1} (-1)^{|\alpha|} \langle g_\alpha, \partial^\alpha \phi \rangle| \\ & \leq & \sum_{|\alpha| \leq 1} \|g_\alpha\|_{L^2} \|\partial^\alpha \phi\|_{L^2} \\ & \leq & \left(\sum_{|\alpha| \leq 1} \|g_\alpha\|_{L^2} \right) \|\phi\|_{H^1}. \end{array}$$

Donc $T \in H^{-1}(\Omega)$. La réciproque est plus délicate (voir [Bony 01]). □

Conséquences : on a les inclusions

$$\mathcal{D}(\Omega) \subset H_0^1(\Omega) \subset L^2(\Omega) \equiv (L^2(\Omega))' \subset H^{-1}(\Omega) \subset \mathcal{D}'(\Omega)$$

et on a par ailleurs pour $T \in L^2(\Omega)$ et $\phi \in \mathcal{D}(\Omega)$,

$$\langle T, \phi \rangle = \langle T, \phi \rangle_{H^{-1}, H_0^1} = \langle T, \phi \rangle_{(L^2)', L^2} = (T, \phi)_{L^2} = \int_\Omega T\,\phi.$$

L^2 est «l'espace pivot» de ces dualités.

8.4 Exercices

▷ **8.1.**

1. Pour quels k les fonctions suivantes appartiennent-elles à l'espace $H^k(]-1,1[)$: $x \mapsto \sin x$, $x \mapsto |x|$, $x \mapsto \mathrm{sgn}(x)$?
2. Pour quels k, quels α et quels d la fonction $\frac{1}{|x|^\alpha}$ appartient-elle à l'espace $H^k(B_0(1))$ dans $\mathbb{R}^d$?
3. Pour quels k, quels α et quels d la fonction $\frac{1}{|x|^\alpha}$ appartient-elle à l'espace $H^k(\mathbb{R}^d \setminus \overline{B_0(1)})$?

▷ **8.2.** Soit Ω le disque centré en 0 et de rayon $1/e$ dans $\mathbb{R}^2$. Soit la fonction

$$u(x,y) = \mathrm{Log}\left(-\mathrm{Log}\left(\sqrt{x^2+y^2}\right)\right).$$

Montrer que $u \in H_0^1(\Omega)$ mais que u n'est pas continue en 0.

▷ **8.3.** On considère l'espace vectoriel

$$V = \{\, v \in L^2(\Omega);\ \Delta v \in L^2(\Omega)\,\}$$

équipé du produit scalaire $(v,w)_V = (v,w)_{L^2} + (\Delta v, \Delta w)_{L^2}$. Montrer que V est un espace de Hilbert.

▷ **8.4.** Soit p un réel tel que $1 \leq p \leq +\infty$. Soit un entier k positif. On pose

$$W^{k,p}(\Omega) = \{u \in L^p(\Omega), \quad \forall \alpha,\ |\alpha| \leq k, \quad \partial_\alpha u \in L^p(\Omega)\},$$

les dérivées étant prises au sens des distributions. On équipe cet espace de la norme

$$\|u\|_{W^{k,p}} = \sum_{|\alpha| \leq k} \|\partial^\alpha u\|_{L^p}.$$

- Justifier le fait que l'on puisse considérer les dérivées de u au sens des distributions.
- Montrer que $W^{k,p}(\Omega)$ équipé de la norme ci-dessus est un espace de Banach.

▷ **8.5.** Soit $u \in H^1(\mathbb{R})$. Montrer que

$$\|u\|_{L^\infty} \leq \|u\|_{L^2}^{\frac{1}{2}} \|u'\|_{L^2}^{\frac{1}{2}}.$$

Indication : utiliser la densité de $\mathcal{D}(\mathbb{R})$ dans $H^1(\mathbb{R})$.

▷ **8.6.** Soit a et b deux réels tels que $a < b$. Montrer que toute fonction de $H^1(]a,b[)$ admet un représentant continu.

▷ **8.7.** Soit a et b deux réels tels que $a < b$. Soit $f \in H^1(]a,b[)$ Montrer que $|f| \in H^1(]a,b[)$. *Indication : on montrera en particulier que*

$$\frac{d|f|}{dx}(x) = \begin{cases} (f(x)/|f(x)|)\frac{df}{dx}(x) & \textit{si } f(x) \neq 0, \\ 0 & \textit{sinon.} \end{cases}$$

En déduire que si f et g sont deux fonctions de $H^1(]a,b[)$, les fonctions $\max(f,g)$ et $\min(f,g)$ sont dans $H^1(]a,b[)$.

▷ **8.8.** Soit I un intervalle ouvert de $\mathbb{R}$. On désigne par $\chi_I(s)$ la fonction caractéristique de I et on pose $\pi_I(s) = \int_{-\infty}^{s} \chi_I(t)dt$. Soit Ω un ouvert de $\mathbb{R}^d$. Montrer que pour $f, g \in H^1(\Omega)$,

$$\int_\Omega g(x)\chi_I(f(x))\nabla f(x)\,dx = \int_\Omega \nabla g(x)\pi_I(f(x))\,dx.$$

Indication : construire une suite de fonctions $(\chi_n)_{n\in\mathbb{N}}$ continues, à valeurs dans $[0,1]$ et convergeant simplement vers χ_I sur $\mathbb{R}$.

▷ **8.9.** Soit $u \in H^1(\Omega)$ et $\psi \in C^\infty(\Omega)$. On suppose que ψ et ses dérviées premières sont bornées sur Ω. Montrer que le produit ψu défini pour presque tout $x \in \Omega$ par $(\psi u)(x) = \psi(x)u(x)$ est dans $H^1(\Omega)$.

▷ **8.10.** Soit $u \in H^1(\mathbb{R}^d)$. On désigne par $(e_1, \ldots, e_d)$ la base cartésienne de $\mathbb{R}^d$. Montrer que pour tout $1 \leq i \leq d$,

$$\int_{\mathbb{R}^d} \left|\frac{\partial u}{\partial x_i}\right|^2 = \lim_{t\to 0} \frac{1}{t} \int_{\mathbb{R}^d} |u(x + te_i) - u(x)|^2\,dx.$$

▷ **8.11.** Soit $u \in H^1(\mathbb{R})$ et soit $\Phi : \mathbb{R} \to \mathbb{R}$ de classe C^1 et telle que $\Phi(0) = 0$. Montrer que la fonction $\Phi \circ u$, définie pour presque tout $x \in \mathbb{R}$ par $(\Phi \circ u)(x) = \Phi(u(x))$, est dans $H^1(\mathbb{R})$.

▷ **8.12.** (Inégalité de Hardy sur $]0,1[$). Soit $\Omega =]0,1[$ et $u \in H^1_0(\Omega)$. Montrer que la fonction $\frac{u(x)}{x(1-x)}$ est dans $L^2(\Omega)$ et que

$$\left\|\frac{u(x)}{x(1-x)}\right\|_{L^2} \leq c\|u'\|_{L^2}$$

où c est une constante indépendante de u.

Chapitre 9

Problèmes aux limites elliptiques linéaires

Ce chapitre est axé sur l'analyse mathématique de *problèmes aux limites elliptiques linéaires*. Le terme *problème aux limites* désigne un problème mathématique composé
- d'une (ou plusieurs) équation(s) aux dérivées partielles (EDP),
- de conditions aux bords (ou de conditions asymptotiques à l'infini).

Pour les problèmes dépendant du temps (dont nous ne parlerons pas ici) s'ajoutent des conditions initiales. Le terme *elliptique* désigne une certaine classe d'équations aux dérivées partielles dont l'archétype est l'*équation de Poisson* $-\Delta u = f$. Enfin le qualificatif *linéaire* signifie que la ou les EDP ainsi que les conditions aux bords sont linéaires ou affines.

Dans ce chapitre, nous démontrerons notamment que le problème aux limites[1]

$$\begin{cases} -\Delta u = f & \text{dans } \mathcal{D}'(\Omega), \\ u = 0 & \text{sur } \partial\Omega, \end{cases} \tag{9.1}$$

admet, pour f donnée dans $H^{-1}(\Omega)$, une solution u et une seule dans $H^1(\Omega)$, et que cette solution dépend continûment du second membre f. Ce résultat sera obtenu en deux étapes :

1. on associera au problème (9.1) une formulation faible (ou variationnelle) qui lui est équivalente

$$\begin{cases} \text{Chercher } u \in V \text{ tel que} \\ \forall v \in V, \quad a(u,v) = b(v), \end{cases} \tag{9.2}$$

   avec $V = H_0^1(\Omega)$, $a(u,v) = \displaystyle\int_\Omega \nabla u \cdot \nabla v$ et $b(v) = \displaystyle\int_\Omega fv$. Les problèmes (9.1) et (9.2) sont équivalents en ce sens que toute solution dans H^1 du problème (9.1) est solution du problème (9.2), et vice versa ;

2. on montrera à l'aide du *théorème de Lax-Milgram* que le problème variationnel (9.2) admet une solution et une seule.

D'autres cas plus compliqués faisant intervenir des opérateurs non symétriques seront également étudiés dans ce chapitre.

[1]dont la résolution numérique par la méthode des éléments finis constitue l'un des thèmes majeurs du cours de Calcul Scientifique

9.1 Version "symétrique" du théorème de Lax-Milgram

Définition 9.1. *Soit H un espace de Hilbert et a une forme bilinéaire sur H. On dit que a est* coercive *s'il existe un réel $\alpha > 0$ tel que*

$$\forall u \in H, \qquad a(u,u) \geq \alpha \|u\|^2.$$

Théorème 9.2. *(Version "symétrique" du théorème de Lax-Milgram). Soit H un espace de Hilbert, a une forme bilinéaire sur H, symétrique, continue et coercive et b une forme linéaire continue sur H. Alors le problème*

$$\left\{ \begin{array}{l} \text{Chercher } u \in H \text{ tel que} \\ \forall v \in H, \qquad a(u,v) = b(v), \end{array} \right. \tag{9.3}$$

admet une solution et une seule. En outre l'unique solution u de (9.3) est également l'unique solution du problème de minimisation

$$\left\{ \begin{array}{l} \text{Chercher } u \in H \text{ tel que} \\ J(u) = \inf\limits_{v \in H} J(v), \end{array} \right. \tag{9.4}$$

où la fonctionnelle $J(v)$ (dite fonctionnelle d'énergie) est définie par

$$J(v) = \frac{1}{2} a(v,v) - b(v).$$

Preuve. Posons

$$(u,v)_a = a(u,v).$$

Il est clair que $(\cdot,\cdot)_a$ est un produit scalaire sur H. Par ailleurs, les normes $\|\cdot\|_H$ et $\|\cdot\|_a$ sont équivalentes ; en effet, de la coercivité et de la continuité de la forme bilinéaire a, nous déduisons que

$$\forall v \in H, \qquad \alpha \|v\|^2 \leq \|v\|_a^2 = a(v,v) \leq M \|v\|^2.$$

L'espace H, qui est complet pour la norme $\|\cdot\|$, l'est donc aussi pour la norme $\|\cdot\|_a$. Donc H muni du produit scalaire $(\cdot,\cdot)_a$ est un espace de Hilbert. La forme linéaire b est continue pour la norme $\|\cdot\|$ donc aussi pour la norme $\|\cdot\|_a$ qui lui est équivalente. D'après le théorème de Riesz 3.11, il existe donc un unique élément $u \in H$ vérifiant

$$\forall v \in H, \qquad a(u,v) = (u,v)_a = b(v).$$

Ceci montre l'existence et l'unicité de la solution de (9.3).

Soit maintenant u solution de (9.3) et $v \in H$. Posons $h = v - u$. On a

$$\begin{array}{rcl} J(v) & = & J(u+h) \\ & = & \frac{1}{2} a(u+h, u+h) - b(u+h) \\ & = & J(u) + a(u,h) - b(h) + \frac{1}{2} a(h,h) \\ & = & J(u) + \frac{1}{2} a(h,h) \\ & \geq & J(u). \end{array}$$

Donc u est solution de (9.4).

Réciproquement, si u est solution de (9.4), on a alors pour tout $v \in H$ et pour tout $\lambda \in \mathbb{R}$

$$J(u) \leq J(u+\lambda v) = J(u) + \lambda(a(u,v) - b(v)) + \frac{1}{2}\lambda^2 a(v,v).$$

Donc pour tout $\lambda \in \mathbb{R}$,

$$\lambda(a(u,v) - b(v)) + \frac{1}{2}\lambda^2 a(v,v) \geq 0.$$

Ceci n'est possible que si $a(u,v) = b(v)$ pour tout $v \in H$. Donc u est solution de (9.3). $\square$

9.2 Résolution d'un problème aux limites elliptique symétrique

Commençons, pour fixer les idées, par nous intéresser au problème modèle qui consiste à

$$\begin{cases} \text{Chercher } u \ : \ \Omega \longrightarrow \mathbb{R} \text{ vérifiant} \\ -\Delta u + \lambda u = f \quad \text{ dans } \Omega, \\ u = 0 \quad \text{ sur } \partial\Omega, \end{cases} \tag{9.5}$$

où Ω est un ouvert de $\mathbb{R}^d$, λ un réel strictement positif et f est une fonction donnée. Le premier travail à faire est de construire un cadre fonctionnel adapté à l'étude du problème (9.5), c'est-à-dire de préciser un espace fonctionnel F dans lequel on s'autorise à choisir la donnée f et un espace fonctionnel V dans lequel on va chercher la solution u.

Bien souvent, la physique fournit des modèles sans les espaces fonctionnels qui vont avec ! C'est au mathématicien de trouver les bons espaces, c'est-à-dire ceux pour lesquels on saura dire des choses sur les solutions du problème (existence, unicité, régularité, ...), pour des données f aussi générales que possible. Un bon cadre fonctionnel doit en principe assurer l'existence et l'unicité d'une solution dans V pour tout $f \in F$.

Nous allons dans cette section construire un cadre fonctionnel naturel, celui des *espaces d'énergie* (la raison pour laquelle on parle d'espace d'énergie apparaîtra clairement plus loin). Commençons par un raisonnement formel. Multiplions les deux membres de l'équation $-\Delta u + \lambda u = f$ par u (sans se préoccuper pour l'instant de rigueur mathématique) et intégrons sur Ω, ce qui donne

$$\int_\Omega (-\Delta u + \lambda u)\, u = \int_\Omega fu. \tag{9.6}$$

En utilisant la formule de Green (7.5) et en tenant compte de la nullité de u sur le bord $\partial\Omega$, on obtient

$$\int_\Omega |\nabla u|^2 + \lambda u^2 = \int_\Omega fu.$$

Pour une donnée f très régulière, par exemple $f \in \mathcal{D}(\Omega)$, les deux termes de l'égalité ont un sens si et seulement si $u \in H^1(\Omega)$. Comme en outre $u = 0$ sur $\partial\Omega$, un bon choix consiste à prendre $V = H_0^1(\Omega)$. On voit alors qu'on peut donner un sens formel au membre de droite de l'équation ci-dessus dès que f est dans le dual de $H_0^1(\Omega)$, c'est-à-dire dans $H^{-1}(\Omega)$. On choisira donc $F = H^{-1}(\Omega)$. On aboutit donc à la formulation mathématique suivante du problème (9.5)

$$\begin{cases} \text{Chercher } u \in H_0^1(\Omega) \text{ vérifiant} \\ -\Delta u + \lambda u = f \quad \text{ dans } \mathcal{D}'(\Omega), \end{cases} \tag{9.7}$$

où f est donnée dans $H^{-1}(\Omega)$. Le problème de Poisson (qui correspond au cas où $\lambda = 0$) sera examiné ultérieurement.

Formulation faible du problème 9.7.

Posons

1. $H = H_0^1(\Omega)$
2. $a(u,v) = \displaystyle\int_\Omega \nabla u \cdot \nabla v + \lambda \int_\Omega u\, v$
3. $b(v) = \langle f, v\rangle_{H^{-1},H_0^1}$

et considérons le problème

$$\left\{\begin{array}{l} \text{Chercher } u \in H \text{ tel que} \\ \forall v \in H, \qquad a(u,v) = b(v). \end{array}\right. \tag{9.8}$$

Proposition 9.3. *Les problèmes (9.7) et (9.8) sont équivalents.*

Preuve. Montrons d'abord que (9.8) $\Rightarrow$ (9.7). Soit u solution de (9.8). On a déjà $u \in H_0^1(\Omega)$, donc d'une part $u \in H^1(\Omega)$ et d'autre part $u = 0$ sur $\partial\Omega$. Soit maintenant $\phi \in \mathcal{D}(\Omega)$. On a $\phi \in H_0^1(\Omega)$ et

$$\begin{aligned} \langle f, \phi\rangle &= \langle f, \phi\rangle_{H^{-1},H_0^1} \\ &= b(\phi) \\ &= a(u,\phi) \\ &= \int_\Omega \nabla u \cdot \nabla \phi + \lambda \int_\Omega u\,\phi \\ &= \langle \nabla u, \nabla \phi\rangle + \lambda \langle u, \phi\rangle \\ &= \langle -\Delta u + \lambda u, \phi\rangle. \end{aligned}$$

Donc $-\Delta u + \lambda u = f$ dans $\mathcal{D}'(\Omega)$. Donc u est solution de (9.7).

Réciproquement, soit u solution de (9.7). On a d'abord $u \in H^1(\Omega)$ et $u = 0$ sur $\partial\Omega$. Donc $u \in H_0^1(\Omega)$. Soit $\phi \in \mathcal{D}(\Omega)$. On a

$$\begin{aligned} b(\phi) &= \langle f, \phi\rangle_{H^{-1},H_0^1} \\ &= \langle f, \phi\rangle \\ &= \langle -\Delta u + \lambda u, \phi\rangle \\ &= \langle \nabla u, \nabla \phi\rangle + \lambda \langle u, \phi\rangle \\ &= a(u,\phi). \end{aligned}$$

Donc

$$\forall \phi \in \mathcal{D}(\Omega), \qquad a(u,\phi) = b(\phi).$$

Vérifions maintenant que a est continue sur $H \times H$ et que b est continue sur H :

$$\forall v \in H = H_0^1(\Omega), \qquad |b(v)| = |\langle f, v\rangle_{H^{-1},H_0^1}| \leq \|f\|_{H^{-1}}\, \|v\|_{H^1},$$

$$\begin{aligned} \forall (u,v) \in H, \qquad |a(u,v)| &= \left| \int_\Omega \nabla u \cdot \nabla v + \lambda \int_\Omega u\, v \right| \\ &\leq \left| \int_\Omega \nabla u \cdot \nabla v \right| + \lambda \left| \int_\Omega u\, v \right| \\ &\leq \left(\int_\Omega |\nabla u|^2 \right)^{1/2} \left(\int_\Omega |\nabla v|^2 \right)^{1/2} + \lambda \left(\int_\Omega u^2 \right)^{1/2} \left(\int_\Omega v^2 \right)^{1/2} \\ &\leq (1+\lambda) \|u\|_{H^1}\, \|v\|_{H^1}. \end{aligned}$$

A fortiori, $a(u,\cdot)$ et b sont continues sur $H = H_0^1(\Omega)$. Comme en outre $\mathcal{D}(\Omega)$ est dense dans $H_0^1(\Omega)$, on a

$$\forall v \in H, \qquad a(u,v) = b(v).$$

Donc u est solution de (9.8). □

Vérification des hypothèses du théorème de Lax-Milgram version symétrique.
- $H = H_0^1(\Omega)$ est un espace de Hilbert ;
- b est continue sur H ;
- a est bilinéaire symétrique et continue sur $H \times H$;
- a est en outre coercive sur H. En effet, pour tout $u \in H$

$$\begin{aligned} a(u,u) &= \int_\Omega |\nabla u|^2 + \lambda \int_\Omega |u|^2 \\ &\geq \min(1,\lambda) \left(\int_\Omega |\nabla u|^2 + \int_\Omega |u|^2 \right) \\ &= \min(1,\lambda) \|u\|_{H^1}^2. \end{aligned}$$

Conclusion. Le problème (9.8) a une solution et une seule de par le théorème de Lax-Milgram version symétrique. Les problèmes (9.7) et (9.8) sont équivalents. Le problème (9.7) a donc une solution et une seule. Remarquons que la fonctionnelle d'énergie associée au problème (9.7) est donnée par

$$J(v) = \frac{1}{2} \int_\Omega |\nabla v|^2 + \frac{\lambda}{2} \int_\Omega |v|^2 - \langle f, v \rangle_{H^{-1}, H_0^1}.$$

9.3 Résolution de l'équation de Poisson sur un ouvert borné

Soit Ω un ouvert borné de $\mathbb{R}^d$ et $f \in H^{-1}(\Omega)$. On cherche $u \in H^1(\Omega)$ solution de

$$\begin{cases} -\Delta u = f & \text{dans } \mathcal{D}'(\Omega), \\ u = 0 & \text{sur } \partial\Omega. \end{cases} \tag{9.9}$$

Formulation faible du problème (9.9).

Posons
1. $H = H_0^1(\Omega)$
2. $a(u,v) = \displaystyle\int_\Omega \nabla u \cdot \nabla v$
3. $b(v) = \langle f, v \rangle_{H^{-1}, H_0^1}$

et considérons le problème

$$\begin{cases} \text{Chercher } u \in H \text{ tel que} \\ \forall v \in H, \qquad a(u,v) = b(v). \end{cases} \tag{9.10}$$

Proposition 9.4. *Les problèmes (9.9) et (9.10) sont équivalents.*

Preuve. Transcription directe de la preuve de la proposition 9.3. □

Vérification des hypothèses du théorème de Lax-Milgram version symétrique.
- $H = H_0^1(\Omega)$ est un espace de Hilbert ;
- b est continue sur H ;
- a est bilinéaire symétrique et continue sur $H \times H$;
- il reste à prouver que a est coercive sur H. En utilisant l'inégalité de Poincaré 8.7, on voit que pour tout $u \in H_0^1(\Omega)$,

$$\|u\|_{H^1}^2 = \|u\|_{L^2}^2 + \|\nabla u\|_{L^2}^2 \leq (1 + C_\Omega^2) \|\nabla u\|_{L^2}^2 = (1 + C_\Omega^2) a(u,u).$$

Donc

$$a(u,u) \geq \frac{1}{1 + C_\Omega^2} \|u\|_{H^1}^2.$$

Conclusion. Le problème (9.10) a une solution et une seule de par le théorème de Lax-Milgram version symétrique. Les problèmes (9.9) et (9.10) sont équivalents. Le problème (9.9) a donc une solution et une seule. Remarquons que la fonctionnelle d'énergie associée au problème (9.9) est donnée par

$$J(v) = \frac{1}{2}\int_\Omega |\nabla v|^2 - \langle f, v\rangle_{H^{-1},H^1_0}.$$

9.4 Théorème de Lax-Milgram

L'hypothèse de symétrie de a est contraignante car de nombreux problèmes aux limites, en particulier ceux dans lesquels figurent des termes d'*advection* (cf. cours de calcul scientifique) sont non symétriques. Heureusement, on peut s'en affranchir.

Théorème 9.5. *(Théorème de Lax-Milgram). Soit H un espace de Hilbert, a une forme bilinéaire sur H continue et coercive et b une forme linéaire continue sur H. Alors le problème*

$$\left\{\begin{array}{l} \text{Chercher } u \in H \text{ tel que} \\ \forall v \in H, \qquad a(u,v) = b(v), \end{array}\right. \tag{9.11}$$

admet une solution et une seule.

Remarque 9.6. Le théorème de Lax-Milgram "contient" le théorème de Lax-Milgram version symétrique.

Preuve. On considère l'application linéaire continue

$$\begin{array}{rcl} \Phi : H & \longrightarrow & H' \\ u & \mapsto & \Phi(u) = a(u,\cdot). \end{array}$$

Il est clair que u est solution de (9.11) si et seulement si $\Phi(u) = b$. Montrer l'existence et l'unicité de la solution de (9.11) équivaut donc à montrer que Φ est bijective.

1. Injectivité (unicité de la solution de (P)).

   Soit $u \in H$ tel que $\Phi(u) = 0$. On a pour tout $v \in H$, $a(u,v) = 0$, donc en particulier $a(u,u) = 0$. Or en utilisant la coercivité,

   $$0 = a(u,u) \geq \alpha\|u\|^2.$$

   Donc $u = 0$.
2. Surjectivité (existence de la solution de (P)).

   Soit $V = \mathrm{Im}(\Phi) \subset H'$. On va montrer que V est à la fois dense et fermé (ce qui implique $V = H'$).

   (a) Montrons que $V = \mathrm{Im}(\Phi)$ est fermé. Soit $w \neq 0$. On a

   $$\|\Phi(w)\|_{H'} = \sup_{v\in H,\ v\neq 0} \frac{|a(w,v)|}{\|v\|} \geq \frac{a(w,w)}{\|w\|} \geq \alpha\|w\|.$$

   Donc

   $$\|\Phi(w)\|_{H'} \geq \alpha\|w\|$$

et cette inégalité reste vraie pour $w = 0$. Soit maintenant $(b_n)_{n\in\mathbb{N}}$ une suite d'éléments de V qui converge dans H' vers $b \in H'$. Il faut montrer que $b \in V$. Soit w_n tel que $\Phi(w_n) = b_n$. Comme $(b_n)_{n\in\mathbb{N}}$ est de Cauchy, $(w_n)_{n\in\mathbb{N}}$ l'est aussi. En effet,

$$\|b_p - b_q\|_{H'} = \|\Phi(w_p) - \Phi(w_q)\|_{H'} = \|\Phi(w_p - w_q)\|_{H'} \geq \alpha\|w_p - w_q\|.$$

Donc $(w_n)_{n\in\mathbb{N}}$ converge dans H vers un certain $w \in H$. Comme en outre Φ est continue,

$$b_n = \Phi(w_n) \longrightarrow \Phi(w) \qquad \text{dans } H'.$$

Donc $b = \Phi(w) \in V$ (unicité de la limite).

(b) Montrons que $V = \mathrm{Im}(\Phi)$ est dense dans H'. Notons

$$V^0 = \{u \in H, \quad \forall f \in V, \qquad f(w) = 0\}.$$

Soit $w \in V^\circ$. On a alors

$$\forall v \in H, \quad a(v,w) = \Phi(v)\,(w) = 0.$$

En particulier pour $v = w$, il vient $a(w,w) = 0$. Donc $w = 0$ par l'hypothèse de coercivité. Donc $V^\circ = \{0\}$, ce qui implique que V est dense dans H.

□

9.5 Résolution d'un problème aux limites non symétrique

Soit Ω un ouvert borné de $\mathbb{R}^d$, $c : \Omega \to \mathbb{R}^d$ un champ de vecteurs de classe $C^1(\overline{\Omega})$ à divergence nulle ($\operatorname{div} c = 0$ dans Ω) et $f \in H^{-1}(\Omega)$. On cherche $u \in H^1(\Omega)$ solution de

$$\left\{\begin{array}{ll} -\Delta u + c\cdot\nabla u = f & \text{dans } \mathcal{D}'(\Omega), \\ u = 0 & \text{sur } \partial\Omega. \end{array}\right. \tag{9.12}$$

Remarquons que le produit $c\cdot\nabla u$ est bien défini dans $\mathcal{D}'(\Omega)$; ceci provient du fait que ∇u est en fait une fonction de $L^2(\Omega)$, donc de $L^1(\Omega)$ puisque Ω est borné, et que c est une fonction bornée (c'est une fonction continue sur un compact). Le produit $c\cdot\nabla u$ est donc dans $L^1(\Omega)$, donc dans $\mathcal{D}'(\Omega)$.

Formulation faible du problème (P).

Posons

1. $H = H_0^1(\Omega)$;
2. $a(u,v) = \displaystyle\int_\Omega \nabla u\cdot\nabla v + \int_\Omega (c\cdot\nabla u)\,v$;
3. $b(v) = \langle f, v\rangle_{H^{-1},H_0^1}$;

et considérons le problème

$$\left\{\begin{array}{l} \text{Chercher } u \in H \text{ tel que} \\ \forall v \in H, \qquad a(u,v) = b(v). \end{array}\right. \tag{9.13}$$

Proposition 9.7. *Les problèmes (9.12) et (9.13) sont équivalents.*

Preuve. On a déjà montré la continuité de b. Pour tout $(u,v) \in H$,

$$\begin{aligned} |a(u,v)| &= \left| \int_\Omega \nabla u \cdot \nabla v + \int_\Omega (c \cdot \nabla u)\, v \right| \\ &\leq \left| \int_\Omega \nabla u \cdot \nabla v \right| + \left| \int_\Omega (c \cdot \nabla u)\, v \right| \\ &\leq \|\nabla u\|_{L^2} \|\nabla v\|_{L^2} + \sup_{\overline{\Omega}} |c| \|\nabla u\|_{L^2} \|v\|_{L^2} \\ &\leq (1 + \sup_{\overline{\Omega}} |c|) \|u\|_{H^1} \|v\|_{H^1}. \end{aligned}$$

Donc a est continue sur $H \times H$.

Le reste de la preuve est une simple transcription de la démonstration de la proposition 9.3. □

Montrons qu'on peut appliquer le théorème de Lax-Milgram.

- $H = H_0^1(\Omega)$ est un espace de Hilbert ;
- b est continue sur H ;
- a est bilinéaire et continue sur $H \times H$;
- montrons que a est coercive sur H. On remarque que pour tout $\phi \in \mathcal{D}(\Omega)$,

$$\int_\Omega c \cdot \nabla \left(\frac{\phi^2}{2} \right) + \int_\Omega \frac{\phi^2}{2} \mathrm{div}\, c = \int_\Omega \mathrm{div}\, \left(\frac{\phi^2}{2} c \right) = \int_{\partial\Omega} \frac{\phi^2}{2} (c \cdot n) = 0.$$

En conséquence,

$$\begin{aligned} a(\phi,\phi) &= \int_\Omega |\nabla \phi|^2 + \int_\Omega (c \cdot \nabla \phi)\phi \\ &= \int_\Omega |\nabla \phi|^2 + \int_\Omega c \cdot \nabla \frac{\phi^2}{2} \\ &= \int_\Omega |\nabla \phi|^2 - \int_\Omega \frac{\phi^2}{2} \mathrm{div}\, c \\ &= \int_\Omega |\nabla \phi|^2 \\ &\geq \frac{1}{1 + C_\Omega^2} \|\phi\|_{H^1}. \end{aligned}$$

Conclusion. Le problème (9.13) a une solution et une seule de par le théorème de Lax-Milgram. Les problèmes (9.12) et (9.13) sont équivalents. Le problème (9.12) a donc une solution et une seule.

9.6 Exercices

▷ **9.1.** Soit Ω un ouvert borné de $\mathbb{R}^d$. On se donne

- un champ de matrices $A : \Omega \longrightarrow \mathcal{M}(d,d)$ de $(L^\infty(\Omega))^{d^2}$,
- deux champs de vecteurs sur b et c de $(L^\infty(\Omega))^d$,
- un champ scalaire $d \in L^\infty(\Omega)$.

Ecrire une formulation faible du problème aux limites : chercher $u \in H_0^1(\Omega)$ vérifiant

$$-\mathrm{div}\, (A(x)\nabla u(x)) + \mathrm{div}\, (bu) + c \cdot \nabla u + du = f \qquad \text{dans } \mathcal{D}'(\Omega).$$

On suppose que le champ de matrices A est uniformément coercif sur Ω, c'est-à-dire qu'il existe une constante $\alpha > 0$ telle que

$$\forall x \in \Omega, \quad \forall y \in \mathbb{R}^d, \qquad (A(x)y, y) \geq \alpha \|y\|^2.$$

A quelles conditions sur A, b, c et d le problème aux limites ci-dessus a-t-il une solution et une seule dans $H_0^1(\Omega)$?

▷ **9.2.** Soit $u_0 \in L^2(\mathbb{R}^d)$ et τ un réel strictement positif. On considère la suite $(u_n)_{n\in\mathbb{N}^*}$ de fonctions de $H^1(\mathbb{R}^d)$ définie à partir de u_0 par la relation de récurrence

$$\forall n \in \mathbb{N}, \qquad \frac{u_{n+1} - u_n}{\tau} = \Delta u_{n+1}. \tag{9.14}$$

1. Montrer que la suite $(u_n)_{n\in\mathbb{N}^*}$ est bien définie de manière unique par la relation de récurrence (9.14).

2. On se place dans le cas où $d = 1$ et on considère la fonction u définie sur $\mathbb{R}$ par

$$u(x) = \left| \begin{array}{ll} \cos(\omega x) & \text{si } x \leq -2\pi/\omega \\ 1 & \text{si } -2\pi/\omega < x < 2\pi/\omega \\ \cos(\omega x) & \text{si } x \geq 2\pi/\omega \end{array} \right.$$

   où $\omega = \tau^{-1/2}$. Calculer u'' (dérivée seconde de u au sens des distributions).

3. En utilisant la question précédente, construire une fonction $u_0 \in L^2(\mathbb{R})$ telle que le problème

$$\left\{ \begin{array}{l} \text{Chercher } u_1 \in H^1(\mathbb{R}) \text{ tel que} \\ \dfrac{u_1 - u_0}{\tau} = u_0'' \end{array} \right.$$

   n'ait pas de solution.

4. Selon vous, dans quel contexte est utilisée la relation de récurrence (9.14) ? Que suggère le résultat de la question 3 ?

▷ **9.3.** Pour tout α réel strictement positif, on note a_α la forme bilinéaire sur $H_0^1(]-1,1[)$ définie par

$$\forall (u,v) \in H_0^1(]-1,1[) \times H_0^1(]-1,1[), \qquad a_\alpha(u,v) = \alpha \int_{-1}^{0} u'v' + \int_0^1 u'v'.$$

On note b la forme linéaire sur $H_0^1(]-1,1[)$ définie par

$$\forall v \in H_0^1(]-1,1[), \qquad b(v) = \int_{-1}^{1} v.$$

1. Montrer que pour tout $\alpha > 0$, le problème

$$\left\{ \begin{array}{l} \text{Chercher } u \in H_0^1(]-1,1[) \text{ tel que} \\ \forall v \in H_0^1(]-1,1[), \quad a_\alpha(u,v) = b(v) \end{array} \right.$$

   admet une solution et une seule, que l'on notera u_α.

2. En introduisant la fonction

$$\eta_\alpha(x) = \left| \begin{array}{ll} \alpha & \text{si } x \in]-1,0[\\ 1 & \text{si } x \in]0,1[\end{array} \right.$$

   ecrire une EDP dans $\mathcal{D}'(]-1,1[)$ dont u_α est solution.

3. Expliciter u_α. On vérifiera que pour $\alpha = 1$, la solution est $(1+x)(1-x)/2$.

   Indication : on cherchera u_α sous la forme

$$u_\alpha(x) = \left| \begin{array}{ll} a_1(x+1)^2 + b_1(x+1) & \text{si } x \in]-1,0[\\ a_2(x-1)^2 + b_2(x-1) & \text{si } x \in]0,1[, \end{array} \right.$$

   où a_1, b_1, a_2 et b_2 sont des coefficients réels.

▷ **9.4.** Ecrire une formulation faible de la forme

$$\left\{ \begin{array}{l} \text{Chercher } u \in V \text{ tel que} \\ \forall v \in V, \quad a(u,v) = b(v) \end{array} \right.$$

pour chacun des problèmes aux limites suivants. On vérifiera dans chaque cas qu'il y a bien équivalence entre le problème aux limites et la formulation faible et que les hypothèses du théorème de Lax-Milgram sont satisfaites.

– Soit $a \in C^0([0,1], \mathbb{R})$ telle que $\inf\limits_{[0,1]} a = \alpha > 0$ et $f \in L^2(0,1)$.

$$(I) \quad \left\{ \begin{array}{l} \text{Chercher } u \in H_0^1(]0,1[) \text{ vérifiant} \\ -\dfrac{d}{dx}\left(a(x) \dfrac{du}{dx}(x) \right) = f(x) \qquad \text{sur }]0,1[\end{array} \right.$$

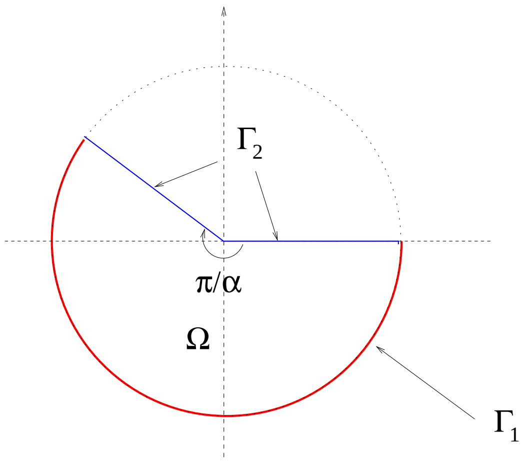
– Soit $f \in L^2(0,1)$.

$$(II) \quad \left\{ \begin{array}{l} \text{Chercher } u \in H_0^1(0,1) \text{ vérifiant} \\ -u'' + u' + u = f(x) \qquad \text{sur }]0,1[\end{array} \right.$$

▷ **9.5.** On considère un réel $\alpha \in]-1, -\dfrac{1}{2}[$ et le domaine Ω défini en coordonnées polaires par

$$\Omega = \left\{ (r,\theta),\ 0 < r < 1,\ \frac{\pi}{\alpha} < \theta < 0 \right\}.$$

Soit $\Gamma_1 = \left\{ (r,\theta),\ r = 1,\ \dfrac{\pi}{\alpha} < \theta < 0 \right\}$ et $\Gamma_2 = \partial\Omega \setminus \Gamma_1$ (voir figure ci-dessous).



On considère le problème aux limites

$$\left\{ \begin{array}{l} \text{Chercher } u \text{ vérifiant} \\ -\Delta u = 0 \quad \text{dans } \Omega, \\ u = \sin(\alpha\theta) \quad \text{sur } \Gamma_1, \\ u = 0 \quad \text{sur } \Gamma_2. \end{array} \right. \tag{9.15}$$

1. Soit $\varphi_1 = r^\alpha \sin(\alpha\theta)$ et $\varphi_2 = r^{-\alpha}\sin(\alpha\theta)$. Montrer que φ_1 et φ_2 sont solutions de (9.15).
   Indication : on rappelle qu'en coordonnées polaires
$$\Delta\varphi = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\varphi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\varphi}{\partial\theta^2}.$$
2. Montrer que φ_1 et φ_2 sont dans $L^2(\Omega)$.
3. Considérons maintenant le problème
$$\left\{\begin{array}{l} \text{Chercher } u \in H^1(\Omega) \text{ vérifiant} \\ -\Delta u = 0 \quad \text{dans } \mathcal{D}'(\Omega), \\ u = \sin(\alpha\theta) \quad \text{sur } \Gamma_1, \\ u = 0 \quad \text{sur } \Gamma_2. \end{array}\right. \tag{9.16}$$
   Montrer que φ_2 est solution du problème (9.16). Qu'en est-il de la fonction φ_1 ?

$\triangleright$ **9.6.** Soit Ω un domaine borné de $\mathbb{R}^d$. Soit w_1, w_2 et f trois fonctions données de $L^2(\Omega)$. Soit λ et μ deux réels positifs ou nuls. Soit a la forme bilinéaire définie formellement par
$$a(u,v) = \int_\Omega \nabla u \cdot \nabla v + \lambda \int_\Omega uv - \mu \left(\int_\Omega w_1 u\right)\left(\int_\Omega w_2 v\right),$$
et b la forme linéaire définie formellement par
$$b(v) = \int_\Omega fv.$$

1. Montrer que a et b sont définies et continues sur $H_0^1(\Omega)$.
2. On note C_Ω la constante de Poincaré de Ω, c'est-à-dire la plus petite constante réelle vérifiant
$$\forall u \in H_0^1(\Omega), \quad \|u\|_{L^2} \le C_\Omega \|\nabla u\|_{L^2}.$$
   Montrer que si $\mu\,\|w_1\|_{L^2}\,\|w_2\|_{L^2} < \lambda + \dfrac{1}{C_\Omega^2}$, le problème
$$\left\{\begin{array}{l} \text{Chercher } u \in H_0^1(\Omega) \text{ tel que} \\ \forall v \in H_0^1(\Omega), \quad a(u,v) = b(v) \end{array}\right. \tag{9.17}$$
   admet une solution et une seule.
3. Montrer que si $\mu\,\|w_1\|_{L^2}\,\|w_2\|_{L^2} < \lambda + \dfrac{1}{C_\Omega^2}$, l'unique solution u du problème (9.17) est solution d'une EDP de la forme
$$-\Delta u + L(u) = f \quad \text{dans } \mathcal{D}'(\Omega)$$
   où $L \in \mathcal{L}\left(L^2(\Omega)\right)$ (i.e. où L est une application linéaire continue de $L^2(\Omega)$ dans lui-même). et donner un majorant de la norme de L.

$\triangleright$ **9.7.** Principe du maximum

Soit Ω un ouvert borné régulier de $\mathbb{R}^d$. On considère l'opérateur linéaire L défini pour une fonction $u \in C^2(\Omega)$ par
$$(Lu)(x) = \sum_{i,j=1}^d a_{ij}(x)\,\frac{\partial^2 u}{\partial x_i \partial x_j}(x) + \sum_{i=1}^d b_i(x)\,\frac{\partial u}{\partial x_i}(x) + c(x)\,u(x)$$
où les champs scalaires a_{ij}, b_i et c sont dans $C^0(\overline{\Omega})$ et vérifient les deux propriétés suivantes :

– il existe un réel $\alpha > 0$ tel que pour tout $x \in \Omega$ et tout $\xi \in \mathbb{R}^d$,

$$\sum_{i,j=1}^{d} a_{ij}(x)\, \xi_i\, \xi_j \geq \alpha\, |\xi|^2 \ ;$$

– le champ c est négatif ou nul sur Ω.

Soit u une fonction de $C^2(\Omega) \cap C^0(\overline{\Omega})$ telle que

$$\begin{cases} Lu < 0 \text{ dans } \Omega \\ u \geq 0 \text{ sur } \partial\Omega. \end{cases}$$

Montrer que $u \geq 0$ dans Ω (par approximation, on peut montrer que ce résultat reste vrai lorsqu'on a seulement $Lu \leq 0$). *Indication : supposer que* $\min\limits_{x \in \overline{\Omega}} u(x) < 0$ *et considérer un* $x_0 \in \overline{\Omega}$ *en lequel le minimum est atteint.*

▷ **9.8.** Soit μ un réel et β un champ de vecteurs de classe C^∞ sur $\mathbb{R}^d$ (autrement dit une application de $\mathbb{R}^d$ à valeurs dans $\mathbb{R}^d$ de classe C^∞) tel que $\|\beta\|_{L^\infty} < +\infty$ et $\|\text{div}\ \beta\|_{L^\infty} < +\infty$.

1. Soit $u \in L^2(\mathbb{R}^d)$. Montrer que les expressions

$$Tu = \mu u + \beta \cdot \nabla u$$

et

$$T^* u = \mu u - \text{div}\ (u\beta)$$

ont un sens dans $\mathcal{D}'(\mathbb{R}^d)$.

2. Soit

$$D = \left\{ u \in L^2(\mathbb{R}^d) \quad \text{tel que} \quad Tu \in L^2(\mathbb{R}^d) \right\}.$$

Montrer que D, muni de la norme $\|\cdot\|_\Gamma$, définie par

$$\forall u \in D, \quad \|u\|_\Gamma = \left(\|u\|_{L^2}^2 + \|Tu\|_{L^2}^2 \right)^{1/2},$$

est un espace de Hilbert.

3. Soit

$$D^* = \left\{ u \in L^2(\mathbb{R}^d) \quad \text{tel que} \quad T^* u \in L^2(\mathbb{R}^d) \right\}.$$

Vérifier que $\mathcal{D}(\mathbb{R}^d) \subset D$, que $\mathcal{D}(\mathbb{R}^d) \subset D^*$ et que

$$\forall (u, v) \in \mathcal{D}(\mathbb{R}^d) \times \mathcal{D}(\mathbb{R}^d), \quad (Tu, v)_{L^2} = (u, T^* v)_{L^2},$$

où $(\cdot, \cdot)_{L^2}$ désigne le produit scalaire usuel sur $L^2(\mathbb{R}^d)$.

4. On admet que $\mathcal{D}(\mathbb{R}^d)$ est dense dans D pour la norme $\|\cdot\|_\Gamma$ et que $\mathcal{D}(\mathbb{R}^d)$ est dense dans D^* pour la norme $\|\cdot\|_{\Gamma^*}$ définie par

$$\forall u \in D^*, \quad \|u\|_{\Gamma^*} = \left(\|u\|_{L^2}^2 + \|T^* u\|_{L^2}^2 \right)^{1/2}.$$

Déduire de la question précédente que

$$\forall (u, v) \in D \times D^*, \quad (Tu, v)_{L^2} = (u, T^* v)_{L^2}.$$

5. Montrer que $D = D^*$.

Commentaire : les opérateurs T est T^* sont des opérateurs linéaires non-bornés sur $L^2(\mathbb{R}^d)$ (un opérateur linéaire non-borné sur $L^2(\mathbb{R}^d)$ est un application linéaire définie sur un sous-espace vectoriel de $L^2(\mathbb{R}^d)$ - le domaine de l'opérateur - à valeurs dans $L^2(\mathbb{R}^d)$). L'opérateur T^* (défini sur le domaine D^*) est appelé l'adjoint de l'opérateur T (défini sur le domaine D).

▷ **9.9.** Régularité elliptique

On considère le problème

$$\left\{ \begin{array}{l} \text{Chercher } u \in H^1(\mathbb{R}^d) \text{ tel que} \\ -\Delta u + u = f \end{array} \right. \tag{9.18}$$

où f est une fonction donnée de $L^2(\mathbb{R}^d)$.

1. Montrer l'existence et l'unicité de la solution de (9.18).

2. Soit $\phi \in C_c^\infty(\mathbb{R}^d)$. Montrer que pour tout $1 \leq i, j \leq d$,

$$\left\| \frac{\partial \phi}{\partial x_i} \right\|_{L^2} \leq \left(\|\phi\|_{L^2}^2 + \|\Delta \phi\|_{L^2}^2 \right)^{1/2}$$

   et

$$\left\| \frac{\partial^2 \phi}{\partial x_i \partial x_j} \right\|_{L^2} \leq \left(\|\phi\|_{L^2}^2 + \|\Delta \phi\|_{L^2}^2 \right)^{1/2}.$$

   On pourra utiliser la transformée de Fourier de ϕ définie par $\hat{\phi}(\xi) = \int_{\mathbb{R}^d} \phi(x)\, e^{-ix\cdot\xi}\, dx$.

3. En admettant que $C_c^\infty(\mathbb{R}^d)$ est dense dans

$$\mathcal{H} = \left\{ v \in L^2(\mathbb{R}^d), \quad \Delta v \in L^2(\mathbb{R}^d) \right\}$$

   muni de la norme définie par

$$\|v\| = \left(\|v\|_{L^2}^2 + \|\Delta v\|_{L^2}^2 \right)^{1/2}$$

   montrer que $\mathcal{H} = H^2(\mathbb{R}^d)$ et que la norme $\|\cdot\|$ est une norme sur $H^2(\mathbb{R}^d)$ équivalente à la norme usuelle définie par

$$\|v\|_{H^2} = \left(\|v\|_{L^2}^2 + \|\nabla v\|_{L^2}^2 + \sum_{i,j=1}^d \left\| \frac{\partial^2 v}{\partial x_i \partial x_j} \right\|_{L^2} \right)^{1/2}.$$

4. Montrer que la solution u de (9.18) est dans $H^2(\mathbb{R}^d)$ et que l'application linéaire qui à f associe l'unique solution u de (9.18) est continue de $L^2(\mathbb{R}^d)$ dans $H^2(\mathbb{R}^d)$.

5. On suppose maintenant que $f \in H^m(\mathbb{R}^d)$ pour un certain $m \geq 1$. Montrer que la solution u de (9.18) est dans $H^{m+2}(\mathbb{R}^d)$. *Indication : on pourra écrire une équation vérifiée par* $\frac{\partial u}{\partial x_i}$.

6. Soit Ω un ouvert borné de $\mathbb{R}^d$, une fonction $u \in H^m(\Omega)$ avec $m \in \mathbb{N}^*$ telle que $\Delta u \in H^m(\Omega)$. Montrer que pour tout ouvert ω tel que $\overline{\omega} \subset \Omega$, on a $u \in H^{m+2}(\Omega)$.

▷ **9.10.** On considère le problème consistant à

$$\left\{\begin{array}{l}\text{Chercher } u \in V \text{ tel que} \\ \forall v \in V, \quad a_\lambda(u,v) = b(v)\end{array}\right. \tag{9.19}$$

avec $V = H^2(\mathbb{R}^2)$

$$a_\lambda(u,v) = \int_{\mathbb{R}^2} \Delta u \, \Delta v + \lambda \int_{\mathbb{R}^2} \nabla u \cdot \nabla v + \int_{\mathbb{R}^2} u\, v \qquad \text{et} \qquad b(v) = \int_{\mathbb{R}^2} f\, v.$$

1. Montrer que pour u et v dans $H^2(\mathbb{R}^2)$,

$$\int_{\mathbb{R}^2} \Delta u \, \Delta v = \sum_{i,j=1}^{2} \int_{\mathbb{R}^2} \frac{\partial^2 u}{\partial x_i \partial x_j} \, \frac{\partial^2 v}{\partial x_i \partial x_j}.$$

2. Montrer que si $\lambda > 0$, le problème (9.19) est bien posé.

3. Montrer que pour tout $\phi \in C_c^\infty(\mathbb{R}^2)$,

$$\int_{\mathbb{R}^2} |\nabla \phi|^2 = - \int_{\mathbb{R}^2} \phi \, \Delta \phi,$$

   et de là que pour tout $u \in H^2(\mathbb{R}^2)$,

$$\int_{\mathbb{R}^2} |\nabla u|^2 \leq \frac{1}{2} \left(\int_{\mathbb{R}^2} u^2 + \int_{\mathbb{R}^2} |\Delta u|^2 \right).$$

   En déduire que si $\lambda > -2$, le problème (9.19) est bien posé.

4. Montrer que les solutions de $\Delta(\Delta u) + 2\,\Delta u + u = f$ dans $H^2(\mathbb{R}^2)$ sont reliées aux solutions dans $L^2(\mathbb{R}^2)$ de $\Delta w + w = f$. Exhiber une fonction $f \in L^2(\mathbb{R}^2)$ pour laquelle cette dernière équation n'a pas de solution dans $L^2(\mathbb{R}^2)$. *Indication : on pourra utiliser la transformation de Fourier.*

Chapitre 10

Transformation de Fourier

La transformation de Fourier est un outil incontournable de l'analyse des phénomènes linéaires, constamment utilisé dans de nombreuses applications comme le traitement du signal, l'électronique ou l'acoustique mais aussi la mécanique des fluides (théorie de la turbulence de Kolmogorov) ou la physique statistique.

Nous commençons par définir la transformée de Fourier d'une fonction de $L^1(\mathbb{R}^d)$ (section 10.1), puis nous étendons cette définition aux éléments de l'espace $\mathcal{S}'(\mathbb{R}^d)$ des distributions tempérées (sections 10.2 et 10.3). Cette extension est réalisée par dualité, à partir d'un espace de fonctions "de petite taille", l'espace de Schwartz $\mathcal{S}(\mathbb{R}^d)$, qui possède la propriété d'être stable par transformation de Fourier.

Nous montrons ensuite (section 10.4) que la transformation de Fourier ainsi définie est une isométrie de $L^2(\mathbb{R}^d)$ et nous utilisons cette propriété pour construire de nouveaux espaces fonctionnels, les espaces de Sobolev fractionnaires $H^s(\mathbb{R}^d)$, $s \in \mathbb{R}$. Ces espaces nous permettrons de mettre en évidence le lien entre la régularité locale d'une fonction et le comportement à l'infini de sa transformée de Fourier, et de fonder rigoureusement la notion de trace définie formellement à la section 8.2.

La section 10.5 traite des distributions périodiques. Nous revisitons la théorie des séries de Fourier dans L^2 à l'aide du concept de base hilbertienne introduit à la section 3.4, puis nous relions le développement en série de Fourier d'une distribution périodique à sa transformée de Fourier. Enfin, nous utilisons ces résultats théoriques pour établir un résultat très utile dans les applications, le théorème d'échantillonage de Shannon.

Dans ce dernier chapitre, nous manipulerons des fonctions à valeurs complexes. Le produit scalaire sur $L^2(\mathbb{R}^d)$ que nous utiliserons sera celui défini par

$$(f, g)_{L^2} = \int_{\mathbb{R}^d} \overline{f}\, g.$$

10.1 Transformation de Fourier dans L^1

10.1.1 Définition

Si f est une fonction de $L^1(\mathbb{R}^d)$, l'intégrale

$$\int_{\mathbb{R}^d} f(x)\, e^{-i\xi\cdot x}\, dx$$

est définie pour tout $\xi \in \mathbb{R}^d$. Cette observation permet d'énoncer la

Définition 10.1. *Soit* $f \in L^1(\mathbb{R}^d)$. *On appelle* transformée de Fourier de f *la fonction notée* $\mathcal{F}f$ *ou* $\hat{f}$ *définie en tout point* $\xi \in \mathbb{R}^d$ *par*

$$(\mathcal{F}f)(\xi) = \hat{f}(\xi) = \int_{\mathbb{R}^d} f(x)\, e^{-i\xi\cdot x}\, dx. \tag{10.1}$$

Proposition 10.2. *Soit* $f \in L^1(\mathbb{R}^d)$. *La transformée de Fourier de* f *vérifie les propriétés suivantes*

1. $\hat{f} \in L^\infty(\mathbb{R}^d)$ *et* $\|\hat{f}\|_{L^\infty} \leq \|f\|_{L^1}$ *;*
2. $\hat{f} \in C^0(\mathbb{R}^d)$ *et* $\hat{f}$ *tend vers* 0 *à l'infini.*

Avant de donner la preuve de cette proposition, tirons-en un corollaire (presque) immédiat :

Corollaire 10.3. *L'application* $\mathcal{F}$ *est une application linéaire continue de* $L^1(\mathbb{R}^d)$ *sur* $L^\infty(\mathbb{R}^d)$ *et*

$$\|\mathcal{F}\|_{\mathcal{L}(L^1(\mathbb{R}^d), L^\infty(\mathbb{R}^d))} = 1.$$

Preuve de la proposition 10.2. Pour tout $\xi \in \mathbb{R}^d$,

$$|\hat{f}(\xi)| = \left| \int_{\mathbb{R}^d} f(x)\, e^{-i\xi\cdot x}\, dx \right| \leq \int_{\mathbb{R}^d} |f(x)|\, dx \leq \|f\|_{L^1}.$$

Ceci prouve la première assertion.

La continuité de la fonction $\xi \mapsto \hat{f}(\xi)$ résulte du théorème 4.45. Enfin, la limite de $\hat{f}$ à l'infini peut s'obtenir par exemple de la façon suivante :

- Si $\phi \in \mathcal{D}(\mathbb{R}^d)$, on montre par intégration par parties que

  $$\hat{\phi}(\xi) = \int_{\mathbb{R}^d} \phi(x) e^{-ix\cdot\xi}\, dx = -\frac{1}{|\xi|^2} \int_{\mathbb{R}^d} \Delta\phi(x) e^{-ix\cdot\xi}\, dx.$$

  Il s'en suit que

  $$|\hat{\phi}(\xi)| \leq \frac{\|\Delta\phi\|_{L^\infty}}{|\xi|^2} \underset{|\xi|\to+\infty}{\longrightarrow} 0.$$

- Comme $\mathcal{D}(\mathbb{R}^d)$ est dense dans $L^1(\mathbb{R}^d)$ (théorème 6.15), on peut trouver, pour tout $f \in L^1(\mathbb{R}^d)$, une suite $(\phi_n)_{n\in\mathbb{N}}$ de fonctions de $\mathcal{D}(\mathbb{R}^d)$ qui converge vers f dans L^1. Comme $\hat{f} - \hat{\phi}_n = \widehat{f - \phi_n}$, il résulte de la première assertion que

  $$\|\hat{f} - \hat{\phi}_n\|_{L^\infty} = \|\widehat{f - \phi_n}\|_{L^\infty} \leq \|f - \phi_n\|_{L^1} \underset{n\to+\infty}{\longrightarrow} 0.$$

  Cela signifie donc que la suite de fonctions $(\hat{\phi}_n)_{n\in\mathbb{N}}$ converge uniformément sur $\mathbb{R}^d$ vers la fonction $\hat{f}$. Comme chaque $\hat{\phi}_n$ tend vers 0 à l'infini, il en est de même pour $\hat{f}$.

$\square$

Preuve du corollaire 10.3. Il est clair que l'application $\mathcal{F}$ est linéaire. On déduit en outre de l'assertion 1 de la proposition 10.2 que $\mathcal{F}$ envoie $L^1(\mathbb{R}^d)$ dans $L^\infty(\mathbb{R}^d)$ et qu'on a pour tout $f \in L^1(\mathbb{R}^d)$,

$$\|\mathcal{F}f\|_{L^\infty} \leq \|f\|_{L^1}.$$

Ceci prouve que $\mathcal{F}$ est continue de $L^1(\mathbb{R}^d)$ dans $L^\infty(\mathbb{R}^d)$ et de norme inférieure ou égale à 1. Pour voir que la norme de $\mathcal{F}$ est exactement égale à 1, il suffit de remarquer que pour tout $f \in L^1(\mathbb{R}^d)$ *positive sur* $\mathbb{R}^d$,

$$\|f\|_{L^1} = \int_{\mathbb{R}^d} f(x)\, dx = (\mathcal{F}f)(0) \leq \|\mathcal{F}f\|_{L^\infty}.$$

$\square$

10.1.2 Transformée de Fourier des gaussiennes

Pour des raisons profondes liées à l'universalité du théorème de la limite centrale, les gaussiennes jouent un rôle fondamental en mathématiques. Ces fonctions ont un comportement remarquable vis-à-vis de la transformation de Fourier qui sera exploité par la suite.

Proposition 10.4. *Soit $\alpha > 0$. On a la relation*

$$\left(\mathcal{F}\left(e^{-\alpha|x|^2}\right)\right)(\xi) = \left(\frac{\pi}{\alpha}\right)^{d/2} e^{-|\xi|^2/4\alpha}.$$

On déduit de la proposition précédente que la transformée de Fourier de la gaussienne $g(x) = C\,\exp(-\frac{|x|^2}{2\sigma^2})$ de variance σ^2 est la gaussienne $\hat{g}(\xi) = \left(\frac{\pi}{a}\right)^{d/2} C\,\exp(-\frac{|\xi|^2}{2/\sigma^2})$ de variance $1/\sigma^2$. La transformée de Fourier d'une gaussienne très piquée sera donc une gaussienne très étalée et réciproquement.

Preuve de la proposition 10.4. Commençons par prouver ce résultat en dimension 1. Par intégration par parties, on montre que la transformée de Fourier

$$\hat{f}(\xi) = \int_{\mathbb{R}} e^{-\alpha\, x^2} e^{-ix\xi}\, dx$$

est de classe C^1 (et même C^∞) et est solution sur $\mathbb{R}$ de l'équation différentielle du premier ordre

$$\hat{f}'(\xi) + \frac{\xi}{2\alpha}\,\hat{f}(\xi) = 0.$$

En intégrant cette équation, on en déduit que

$$\hat{f}(\xi) = C\, e^{-\xi^2/4\alpha}$$

où C est une constante à déterminer, dont on obtient la valeur en écrivant :

$$\hat{f}(0) = C = \int_{\mathbb{R}} e^{-\alpha\, x^2}\, dx = \sqrt{\frac{\pi}{\alpha}}.$$

Pour étendre ce résultat en dimension $n \geq 2$, il suffit de remarquer que pour une gaussienne (comme d'ailleurs pour toute fonction de n variables qui s'écrit comme le produit de fonctions d'une seule variable), on peut séparer les variables dans la définition de la transformée de Fourier :

$$\begin{aligned}
\left(\mathcal{F}\left(e^{-\alpha|x|^2}\right)\right)(\xi) &= \int_{\mathbb{R}^d} e^{-\alpha|x|^2} e^{-i\xi\cdot x}\, dx \\
&= \int_{\mathbb{R}^d} \left(\prod_{j=1}^d e^{-\alpha x_j^2}\right)\left(\prod_{j=1}^d e^{-i\xi_j\, x_j}\right) \prod_{j=1}^d dx_j \\
&= \prod_{j=1}^d \left(\int_{\mathbb{R}} e^{-\alpha x_j^2} e^{-i\xi_j\, x_j}\, dx_j\right) \\
&= \prod_{j=1}^d \left(\sqrt{\frac{\pi}{\alpha}}\, e^{-|\xi_j|^2/4\alpha}\right) \\
&= \left(\frac{\pi}{\alpha}\right)^{d/2} e^{-|\xi|^2/4\alpha}.
\end{aligned}$$

□

10.1.3 Propriétés élémentaires

Un des intérêts majeurs de la transformation de Fourier est que celle-ci permet de transformer certaines opérations linéaires fondamentales (dérivation, convolution, translation) en des opérations plus simples. Les trois théorèmes suivants synthétisent les principaux résultats à connaître :

Théorème 10.5. *Soit* $f \in L^1(\mathbb{R}^d)$.

– *Si* $\dfrac{\partial f}{\partial x_j} \in L^1(\mathbb{R}^d)$, *alors*

$$\left(\mathcal{F}\frac{\partial f}{\partial x_j}\right)(\xi) = i\xi_j(\mathcal{F}f)(\xi).$$

– *Si* $x_j f \in L^1(\mathbb{R}^d)$, *alors*

$$\left(\mathcal{F}(x_j f)\right)(\xi) = i\frac{\partial(\mathcal{F}f)}{\partial \xi_j}(\xi).$$

Théorème 10.6. *Si* $f \in L^1(\mathbb{R}^d)$ *et* $g \in L^1(\mathbb{R}^d)$, *alors* $f \star g \in L^1(\mathbb{R}^d)$ *et*

$$\mathcal{F}(f \star g) = \mathcal{F}(f)\mathcal{F}(g).$$

Théorème 10.7. *Soit* $f \in L^1(\mathbb{R}^d)$ *et* $a \in \mathbb{R}^d$. *Soit* $\tau_a f$ *la translatée de* f *selon* a *définie pour tout* $x \in \mathbb{R}^d$ *par* $\tau_a f(x) = f(x-a)$. *On a*

$$\mathcal{F}(\tau_a f) = e^{-i\xi\cdot a}(\mathcal{F}f).$$

Les preuves des théorèmes 10.5 et 10.7 sont élémentaires ; la première s'appuie sur le théorème 4.46 et la seconde s'obtient par le changement de variable $x \mapsto x - a$. Détaillons la

Preuve du théorème 10.6. Posons $h = f \star g$. On a

$$h(x) = \int_{\mathbb{R}^d} f(y)\, g(x-y)\, dy.$$

L'expression ci-dessus est définie presque partout et la fonction h est de classe L^1 (cela résulte du théorème 4.50 de Fubini). On a donc, toujours en utilisant le théorème 4.50 et le changement de variable $z = x - y$,

$$\begin{aligned}
\hat{h}(\xi) &= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} f(y)\, g(x-y)\, dy\right) e^{-ix\cdot\xi}\, dx \\
&= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} f(y)\, g(x-y)\, dy\right) e^{-i(x-y)\cdot\xi} e^{-iy\cdot\xi}\, dx \\
&= \left(\int_{\mathbb{R}^d} f(y) e^{-iy\cdot\xi}\, dy\right)\left(\int_{\mathbb{R}^d} g(z) e^{-iz\cdot\xi}\, dz\right) \\
&= \hat{f}(\xi)\, \hat{g}(\xi).
\end{aligned}$$

□

10.1.4 Formule d'inversion de Fourier

Définition 10.8. *Soit $f \in L^1(\mathbb{R}^d)$. On note $(\mathcal{F}^{-1}f)$ la fonction définie en tout $\xi \in \mathbb{R}^d$ par*

$$(\mathcal{F}^{-1}f)(\xi) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} f(x)\, e^{i\xi\cdot x}\, dx. \tag{10.2}$$

Nous allons maintenant énonçer le théorème d'inversion de Fourier dans L^1. Nous admettrons provisoirement ce résultat, qui apparaîtra par la suite comme un simple corollaire du théorème 10.26 d'inversion de Fourier dans l'espace des distributions tempérées.

Théorème 10.9. *Soit f dans $L^1(\mathbb{R}^d)$ telle que $\hat{f}$ soit aussi dans $L^1(\mathbb{R}^d)$. Alors*

$$f = \mathcal{F}^{-1}\mathcal{F}f \tag{10.3}$$

et

$$f = \mathcal{F}\mathcal{F}^{-1}f. \tag{10.4}$$

L'espace des fonctions L^1 dont la transformée de Fourier est aussi dans L^1 est donc stable par transformation de Fourier et la transformation $\mathcal{F}^{-1}$ est l'inverse de la transformation de Fourier $\mathcal{F}$ sur cet espace. Notons que la formule (10.3) s'écrit aussi

$$f(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{f}(\xi)\, e^{i\xi\cdot x}\, d\xi.$$

10.2 L'espace $\mathcal{S}$ de Schwartz

Notre objectif, que nous atteindrons à la section 10.3, est d'étendre la transformation de Fourier à des objets plus généraux que les fonctions L^1. Le procédé va être similaire à celui que nous avons utilisé à la section 6.4 pour définir la dérivation dans l'espace des distributions. Nous allons en effet construire dans cette section un "petit" espace de fonctions tests stable par transformation de Fourier (l'espace de Schwartz) qui nous permettra de définir par dualité la transformation de Fourier sur un "grand" espace (l'espace des distributions tempérées).

10.2.1 Définition de l'espace $\mathcal{S}$

Définition 10.10. *On dit qu'une fonction $\phi : \mathbb{R}^d \longrightarrow \mathbb{R}$ de classe C^∞ est* à décroissance rapide *si pour tout $p \geq 0$,*

$$\mathcal{N}_p(\phi) = \sup_{|\alpha|\leq p} \sup_{|\beta|\leq p} \|x^\alpha \partial^\beta \phi(x)\|_{L^\infty} < +\infty.$$

On note $\mathcal{S}(\mathbb{R}^d)$ l'espace vectoriel des fonctions C^∞ à décroissance rapide.

Définition 10.11. *On dit qu'une suite $(\phi_n)_{n\in\mathbb{N}}$ de fonctions de $\mathcal{S}$* converge dans $\mathcal{S}$ *vers $\phi \in \mathcal{S}$ si et seulement si*

$$\forall p \in \mathbb{N}, \quad \mathcal{N}_p(\phi_n - \phi) \underset{n\to+\infty}{\longrightarrow} 0.$$

On a évidemment $\mathcal{D}(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d)$. Mais on a en fait un résultat plus fort :

Proposition 10.12. *$\mathcal{D}(\mathbb{R}^d)$ est dense dans $\mathcal{S}(\mathbb{R}^d)$.*

Comme exemple de fonctions de $\mathcal{S}(\mathbb{R}^d)$ qui ne sont pas dans $\mathcal{D}(\mathbb{R}^d)$, citons en particulier les gaussiennes.

Preuve. Soit $\phi \in \mathcal{S}(\mathbb{R}^d)$ et $\chi \in \mathcal{D}(\mathbb{R}^d)$ telle que $\chi = 1$ sur la boule unité. On considère la suite $(\phi_n)_{n\in\mathbb{N}^*}$ définie par

$$\phi_n(x) = \phi(x)\chi(x/n).$$

Il est clair que pour tout $n \in \mathbb{N}^*$, $\phi_n \in \mathcal{D}(\mathbb{R}^d)$. On a donc pour tout $\beta \geq 0$,

$$\partial^\beta(\phi - \phi_n)(x) = \partial^\beta\phi(x)(1 - \chi(x/n)) - \sum_{|\gamma|\geq 1,\ \gamma\leq\beta} \begin{pmatrix} \beta \\ \gamma \end{pmatrix} \frac{1}{n^{|\gamma|}} \partial^{\beta-\gamma}\phi(x)\, \partial^\gamma\chi(x/n).$$

D'où,

$$\left\|x^\alpha\partial^\beta(\phi - \phi_n)(x)\right\|_{L^\infty} \leq \sup_{|x|\geq n} |x^\alpha\partial^\beta\phi(x)| + \frac{C_{\alpha,\beta}}{n} \sum_{|\gamma|\geq 1,\ \gamma\leq\beta} \left\|x^\alpha\partial^{\beta-\gamma}\phi(x)\right\|_{L^\infty}.$$

Pour α et β fixés, chacun des deux termes tend vers 0 lorsque n tend vers $+\infty$. On a donc pour tout $\alpha \geq 0$ et $\beta \geq 0$,

$$\left\|x^\alpha\partial^\beta(\phi - \phi_n)(x)\right\|_{L^\infty} \underset{n\to+\infty}{\longrightarrow} 0.$$

La suite $(\phi_n)_{n\geq 1}$ converge donc vers ϕ dans $\mathcal{S}(\mathbb{R}^d)$. □

Proposition 10.13. *L'espace $\mathcal{S}(\mathbb{R}^d)$ est stable par dérivation et par multiplication par des polynômes. De plus, il existe une constante $C \in \mathbb{R}_+$ telle que pour tout $p \in \mathbb{N}$*

$$\forall\phi \in \mathcal{S}(\mathbb{R}^d), \quad \sup_{|\alpha|\leq p}\sup_{|\beta|\leq p} \|x^\alpha\partial^\beta\phi(x)\|_{L^1} \leq C\mathcal{N}_{p+d+1}(\phi). \tag{10.5}$$

Preuve. Les premières affirmations sont des conséquences directes de la définition 10.10. Pour montrer l'inégalité (10.5), commençons par remarquer que

$$I = \int_{\mathbb{R}^d} \frac{1}{1 + \sum_{1\leq j\leq d} |x_j|^{d+1}}\, dx < +\infty.$$

On en déduit que pour $p \geq 0$, $|\alpha| \leq p$ et $|\beta| \leq p$,

$$\begin{aligned}
\|x^\alpha\partial^\beta\phi(x)\|_{L^1} &= \left\|(1 + \sum_{1\leq j\leq d} |x_j|^{d+1})x^\alpha\partial^\beta\phi(x)\, \frac{1}{1 + \sum_{1\leq j\leq d} |x_j|^{d+1}}\right\|_{L^1} \\
&\leq \left\|(1 + \sum_{1\leq j\leq d} |x_j|^{d+1})x^\alpha\partial^\beta\phi(x)\right\|_{L^\infty} \left\|\frac{1}{1 + \sum_{1\leq j\leq d} |x_j|^{d+1}}\right\|_{L^1} \\
&\leq C\mathcal{N}_{p+d+1}(\phi).
\end{aligned}$$

□

10.2.2 Transformation de Fourier dans $\mathcal{S}$

Toute fonction de $\mathcal{S}(\mathbb{R}^d)$ étant intégrable, on peut définir la transformation de Fourier sur $\mathcal{S}(\mathbb{R}^d)$ en restreignant la définition 10.1 à l'espace $\mathcal{S}(\mathbb{R}^d)$. Le point remarquable est que la transformée de Fourier d'une fonction de $\mathcal{S}(\mathbb{R}^d)$ est elle-aussi dans $\mathcal{S}(\mathbb{R}^d)$.

Théorème 10.14. *L'espace $\mathcal{S}(\mathbb{R}^d)$ est stable par transformée de Fourier et pour tout $p \in \mathbb{N}$ il existe une constante C'_p telle que*

$$\forall \phi \in \mathcal{S}(\mathbb{R}^d), \quad \mathcal{N}_p(\hat{\phi}) \le C'_p \mathcal{N}_{p+d+1}(\phi).$$

De plus la transformée de Fourier définit un isomorphisme séquentiellement bicontinu de $\mathcal{S}(\mathbb{R}^d)$ dans lui-même d'inverse l'application $\mathcal{F}^{-1}$ définie par (10.2).

Preuve. Soit $p \ge 0$. Pour tout $|\alpha| \le p$ et $|\beta| \le p$,

$$\begin{aligned} \left\| \xi^\alpha \partial^\beta \hat{\phi}(\xi) \right\|_{L^\infty} &= \left\| \mathcal{F}\left(\partial^\alpha (x^\beta \phi(x))\right) \right\|_{L^\infty} \\ &\le \left\| \partial^\alpha (x^\beta \phi(x)) \right\|_{L^1} \\ &\le \widetilde{C}_p \sup_{|\alpha'| \le p} \sup_{|\beta'| \le p} \left\| x^{\alpha'} \partial^{\beta'} \phi(x) \right\|_{L^1} \\ &\le C'_p \mathcal{N}_{p+d+1}(\phi). \end{aligned}$$

On en déduit que la transformation de Fourier est un endomorphisme séquentiellement continu sur $\mathcal{S}(\mathbb{R}^d)$. Il reste à montrer que pour tout $f \in \mathcal{S}(\mathbb{R}^d)$,

$$\mathcal{F}^{-1}(\mathcal{F} f) = f \quad \text{et} \quad \mathcal{F}(\mathcal{F}^{-1} f) = f.$$

On voit déjà qu'il suffit de montrer l'une des deux égalités, par exemple la première (l'autre s'en déduit en remplaçant f par son complexe conjugué). Cela se fait en deux étapes :

1. Soit g et h dans $\mathcal{S}(\mathbb{R}^d)$. Comme la fonction

$$(x, y) \mapsto g(x)\,\overline{h(y)}\,e^{-ix\cdot y}$$

est dans $L^1(\mathbb{R}^d \times \mathbb{R}^d)$, on a

$$\begin{aligned} \int_{\mathbb{R}^d} (\mathcal{F} g)(y) \overline{h(y)}\, dy &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} g(x)\, e^{-ix\cdot y}\, dx \right) \overline{h(y)}\, dy \\ &= \int_{\mathbb{R}^d} g(x) \left(\int_{\mathbb{R}^d} \overline{h(y)}\, e^{-ix\cdot y}\, dy \right) dx \\ &= \int_{\mathbb{R}^d} g(x)\, \overline{\left(\int_{\mathbb{R}^d} h(y)\, e^{ix\cdot y}\, dy \right)}\, dx, \end{aligned}$$

ce qui s'écrit aussi

$$\int_{\mathbb{R}^d} (\mathcal{F} g)(y)\, \overline{h(y)}\, dy = (2\pi)^d \int_{\mathbb{R}^d} g(x)\, \overline{(\mathcal{F}^{-1} h)(x)}\, dx.$$

A partir de cette relation, on obtient facilement

$$\int_{\mathbb{R}^d} (\mathcal{F}^{-1}(\mathcal{F} g))(x)\, \overline{h(x)}\, dx = \int_{\mathbb{R}^d} g(x)\, \overline{\left(\mathcal{F}^{-1}\left(\mathcal{F} h\right)\right)(x)}\, dx. \tag{10.6}$$

2. Soit $x_0 \in \mathbb{R}^d$. Appliquons la relation (10.6) avec $g = f$ et $h = \chi_\epsilon$ où $\chi_\epsilon(x) = \dfrac{e^{-|x-x_0|^2/2\epsilon}}{\epsilon^{d/2}}$ (on pourra vérifier en exercice que l'espace $\mathcal{S}(\mathbb{R}^d)$ contient bien les fonctions gaussiennes). En utilisant la proposition 10.4 et le théorème 10.7, on établit facilement que
$$\forall x \in \mathbb{R}^d, \quad \left(\mathcal{F}^{-1}\left(\mathcal{F}\chi_\epsilon\right)\right)(x) = \chi_\epsilon(x).$$
Il en résulte que
$$\int_{\mathbb{R}^d} f(x)\,\frac{e^{-|x-x_0|^2/2\epsilon}}{\epsilon^{d/2}}\,dx = \int_{\mathbb{R}^d} (\mathcal{F}^{-1}(\mathcal{F}f))(x)\,\frac{e^{-|x-x_0|^2/2\epsilon}}{\epsilon^{d/2}}\,dx.$$
En faisant tendre ϵ vers 0, on obtient
$$(\mathcal{F}^{-1}\mathcal{F}f)(x_0) = f(x_0),$$
ce qui permet de conclure la preuve.

□

10.3 L'espace $\mathcal{S}'$ des distributions tempérées

10.3.1 Définition de l'espace $\mathcal{S}'$

Définition 10.15. *On note $\mathcal{S}'(\mathbb{R}^d)$ l'espace vectoriel des formes linéaires sur $\mathcal{S}(\mathbb{R}^d)$ qui vérifient la propriété de continuité suivante : il existe un entier p et une constante C tels que*
$$\forall \phi \in \mathcal{S}(\mathbb{R}^d), \qquad |\langle T, \phi\rangle_{\mathcal{S}',\mathcal{S}}| \le C\mathcal{N}_p(\phi). \tag{10.7}$$

Théorème 10.16. *La restriction à $\mathcal{D}(\mathbb{R}^d)$ d'un élément de $\mathcal{S}'(\mathbb{R}^d)$ définit une distribution.*

Preuve. Soit $T \in \mathcal{S}'(\mathbb{R}^d)$, $p \in \mathbb{N}$ et $C \ge 0$ tels que
$$\forall \phi \in \mathcal{S}(\mathbb{R}^d), \qquad |\langle T, \phi\rangle_{\mathcal{S}',\mathcal{S}}| \le C\mathcal{N}_p(\phi).$$
$\mathcal{D}(\mathbb{R}^d)$ étant un sous-ensemble de $\mathcal{S}(\mathbb{R}^d)$, il est clair que la restriction de T à $\mathcal{D}(\mathbb{R}^d)$ est une forme linéaire sur $\mathcal{D}(\mathbb{R}^d)$. Soit K un compact de $\mathbb{R}^d$ et $R \ge 1$ tel que $K \subset \overline{B_R(0)}$. On a pour tout $\phi \in \mathcal{D}(\mathbb{R}^d)$ à support dans K,
$$|\langle T, \phi\rangle_{\mathcal{S}',\mathcal{S}}| \le C\mathcal{N}_p(\phi) = C \sup_{|\alpha|\le p}\sup_{|\beta|\le p} \|x^\alpha \partial^\beta \phi(x)\|_{L^\infty} \le C\,R^p \sup_{|\beta|\le p} \|\partial^\beta \phi(x)\|_{L^\infty}.$$
La restriction de T à $\mathcal{D}(\mathbb{R}^d)$ définit donc une distribution sur $\mathbb{R}^d$ d'ordre inférieur ou égal à p. □

Définition 10.17. *Les éléments de $\mathcal{S}'(\mathbb{R}^d)$ sont appelés les* distributions tempérées*, ou parfois les* distributions à croissance lente.

Remarque 10.18. Le théorème 10.16 montre qu'il existe une application naturelle de l'espace $\mathcal{S}'(\mathbb{R}^d)$ des distributions tempérées dans l'espace $\mathcal{D}'(\mathbb{R}^d)$ des distributions. Le fait que cette application soit une injection résulte de ce qu'une distribution tempérée $T \in \mathcal{S}'(\mathbb{R}^d)$ est complètement caractérisée par sa restriction à $\mathcal{D}(\mathbb{R}^d)$ du fait de la densité de $\mathcal{D}(\mathbb{R}^d)$ dans $\mathcal{S}(\mathbb{R}^d)$ (proposition 10.12). Réciproquement, si T est une distribution telle qu'il existe $p \in \mathbb{N}$ et $C \in \mathbb{R}_+$ tel que
$$\forall \phi \in \mathcal{D}(\Omega), \quad |\langle T, \phi\rangle| \le C\,\mathcal{N}_p(\phi),$$

alors il existe une unique distribution tempérée qui prolonge T. Par abus de notation, on notera par la même lettre une distribution tempérée et sa restriction à $\mathcal{D}(\mathbb{R}^d)$. Cela permet notamment d'écrire

$$\forall \phi \in \mathcal{D}(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d), \quad \langle T, \phi \rangle_{\mathcal{S}',\mathcal{S}} = \langle T, \phi \rangle. \tag{10.8}$$

Si en particulier, f est une fonction de $L^p(\mathbb{R}^d)$ avec $1 \leq p \leq +\infty$, et si $\phi \in \mathcal{S}(\mathbb{R}^d)$, on écrira

$$\langle f, \phi \rangle_{\mathcal{S}',\mathcal{S}} = \int_{\mathbb{R}^d} f\phi \tag{10.9}$$

(nous établirons en effet à la section suivante que toute fonction de $L^p(\mathbb{R}^d)$ définit au travers de (10.9) une distribution tempérée sur $\mathbb{R}^d$).

10.3.2 Convergence et dérivation dans $\mathcal{S}'$

Définition 10.19. *On dit que la suite $(u_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{S}'(\mathbb{R}^d)$* converge dans $\mathcal{S}'(\mathbb{R}^d)$ *vers u si on a*

$$\forall \phi \in \mathcal{S}(\mathbb{R}^d), \quad \langle u_n, \phi \rangle_{\mathcal{S}',\mathcal{S}} \underset{n\to+\infty}{\longrightarrow} \langle u, \phi \rangle_{\mathcal{S}',\mathcal{S}}.$$

Théorème 10.20. *Soit $(T_n)_{n\in\mathbb{N}}$ une suite de distributions tempérées qui converge dans $\mathcal{S}'(\mathbb{R}^d)$ vers la distribution tempérée T. Alors $(T_n)_{n\in\mathbb{N}}$ converge aussi dans $\mathcal{D}'(\mathbb{R}^d)$ vers la distribution T.*

Preuve. Résulte immédiatement de la relation (10.8) et des définitions 7.14 et 10.19.

Définition - Théorème 10.21. *Soit $T \in \mathcal{S}'(\mathbb{R}^d)$. La* dérivée de T par rapport à la variable x_j *est l'élément de $\mathcal{S}'(\mathbb{R}^d)$ noté $\dfrac{\partial T}{\partial x_j}$ défini par*

$$\forall \phi \in \mathcal{S}(\mathbb{R}^d), \quad \langle \frac{\partial T}{\partial x_j}, \phi \rangle_{\mathcal{S}',\mathcal{S}} = -\langle T, \frac{\partial \phi}{\partial x_j} \rangle_{\mathcal{S}',\mathcal{S}}. \tag{10.10}$$

La distribution définie par $\dfrac{\partial T}{\partial x_j}$ est la dérivée (au sens des distributions) de la distribution définie par T.

Preuve. Il est clair que $\dfrac{\partial T}{\partial x_j}$ définit une forme linéaire sur $\mathcal{S}(\mathbb{R}^d)$. Comme

$$\begin{aligned} |\langle \frac{\partial T}{\partial x_j}, \phi \rangle_{\mathcal{S}',\mathcal{S}}| &= |\langle T, \frac{\partial \phi}{\partial x_j} \rangle_{\mathcal{S}',\mathcal{S}}| \\ &\leq C\mathcal{N}_p(\frac{\partial \phi}{\partial x_j}) \\ &\leq C\mathcal{N}_{p+1}(\phi), \end{aligned}$$

on voit que $\dfrac{\partial T}{\partial x_j} \in \mathcal{S}'(\mathbb{R}^d)$. Le fait que la restriction à $\mathcal{D}(\mathbb{R}^d)$ de la distribution tempérée $\dfrac{\partial T}{\partial x_j}$ soit la dérivée au sens des distributions de la restriction de T à $\mathcal{D}(\mathbb{R}^d)$, résulte de la relation (10.8) et des définitions 6.8 et 10.21. □

Théorème 10.22. *Si $u_n \underset{n\to+\infty}{\longrightarrow} u$ dans $\mathcal{S}'(\mathbb{R}^d)$, alors $\partial^\alpha u_n \underset{n\to+\infty}{\longrightarrow} \partial^\alpha u$ dans $\mathcal{S}'(\mathbb{R}^d)$.*

Preuve. Soit $\phi \in \mathcal{S}(\mathbb{R}^d)$. On a

$$\langle \partial^\alpha u_n, \phi \rangle_{\mathcal{S}',\mathcal{S}} = (-1)^{|\alpha|} \langle u_n, \partial^\alpha \phi \rangle_{\mathcal{S}',\mathcal{S}} \underset{n\to+\infty}{\longrightarrow} (-1)^{|\alpha|} \langle u, \partial^\alpha \phi \rangle_{\mathcal{S}',\mathcal{S}} = \langle \partial^\alpha u, \phi \rangle_{\mathcal{S}',\mathcal{S}}.$$

□

10.3.3 Distributions tempérées particulières

Théorème 10.23. *L'espace $\mathcal{S}'(\mathbb{R}^d)$ contient*

1. *les polynômes sur $\mathbb{R}^d$;*
2. *les fonctions L^p pour tout $1 \leq p \leq +\infty$ et l'injection*
$$L^p(\mathbb{R}^d) \hookrightarrow \mathcal{S}'(\mathbb{R}^d)$$
est séquentiellement continue ;
3. *l'espace $\mathcal{E}'(\mathbb{R}^d)$ des distributions à support compact.*

Preuve.

1. Soit $p(x) = \sum_{\alpha \in \mathbb{N}^d,\ |\alpha| \leq N} a_\alpha x^\alpha$ un polynôme sur $\mathbb{R}^d$ de degré total inférieur ou égal à N. On a, au vu de l'inégalité (10.5),
$$\forall \phi \in \mathcal{S}(\mathbb{R}^d), \qquad \left| \int_{\mathbb{R}^d} p\phi \right| \leq \sum_{\alpha \in \mathbb{N}^d,\ |\alpha| \leq N} |a_\alpha| \; \|x^\alpha \phi\|_{L^1} \leq C_p \mathcal{N}_{N+d+1}(\phi),$$
où C_p est une constante qui ne dépend que de p. Ceci prouve la première assertion.
2. Soit $f \in L^p(\mathbb{R}^d)$. Comme $\mathcal{S}(\mathbb{R}^d) \subset L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$, f définit une forme linéaire sur $\mathcal{S}(\mathbb{R}^d)$ (ceci découle de l'inégalité de Hölder (5.4). Vérifions la propriété de continuité.
   (a) Si $f \in L^1(\mathbb{R}^d)$ on a
$$|\langle f, \phi \rangle_{\mathcal{S}',\mathcal{S}}| = \left| \int_{\mathbb{R}^d} f\phi \right| \leq \|f\|_{L^1} \, \|\phi\|_{L^\infty} = \|f\|_{L^1} \mathcal{N}_0(\phi).$$
   (b) Si $f \in L^\infty(\mathbb{R}^d)$ on a
$$|\langle f, \phi \rangle_{\mathcal{S}',\mathcal{S}}| = \left| \int_{\mathbb{R}^d} f\phi \right| \leq \|f\|_{L^\infty} \, \|\phi\|_{L^1} = C_0 \|f\|_{L^\infty} \mathcal{N}_{d+1}(\phi).$$
   (c) Si $f \in L^p(\mathbb{R}^d)$ avec $1 < p < +\infty$, on peut écrire
$$f = f\chi_{|f|<1} + f\chi_{|f|\geq 1} = f_\infty + f_1 \qquad \text{avec} \qquad f_\infty \in L^\infty \text{ et } f_1 \in L^1$$
   On a donc
$$|\langle f, \phi \rangle_{\mathcal{S}',\mathcal{S}}| = |\langle f_\infty, \phi \rangle_{\mathcal{S}',\mathcal{S}} + \langle f_1, \phi \rangle_{\mathcal{S}',\mathcal{S}}| \leq \|f_1\|_{L^1} \mathcal{N}_0(\phi) + C_0 \|f\|_{L^\infty} \mathcal{N}_{d+1}(\phi) \leq C \mathcal{N}_{d+1}(\phi).$$

   Soit maintenant une suite $(f_n)_{n \in \mathbb{N}}$ qui converge vers f dans L^p. On a pour tout $\phi \in \mathcal{S}(\mathbb{R}^d)$
$$|\langle f_n, \phi \rangle_{\mathcal{S}',\mathcal{S}} - \langle f_n, \phi \rangle_{\mathcal{S}',\mathcal{S}}| \leq \left| \int_{\mathbb{R}^d} (f_n - f)\phi \right| \leq \|f_n - f\|_{L^p} \, \|\phi\|_{L^{p'}} \underset{n \to +\infty}{\longrightarrow} 0.$$
3. Soit enfin $u \in \mathcal{E}'(\mathbb{R}^d)$. Soit p son ordre, K un voisinage compact de Supp u et $\chi \in \mathcal{D}(\mathbb{R}^d)$ valant 1 sur K. Posons pour tout $\phi \in C^\infty(\mathbb{R}^d)$,
$$\langle u, \phi \rangle_{\mathcal{E}',C^\infty} = \langle u, \chi\phi \rangle.$$
Cette définition est indépendante de χ : soit en effet χ_1 et χ_2 dans $\mathcal{D}(\mathbb{R}^d)$ valant 1 sur K ; on a
$$\langle u, \chi_1\phi \rangle - \langle u, \chi_2\phi \rangle = \langle u, (\chi_1 - \chi_2)\phi \rangle.$$

La fonction $\widetilde{\phi} = (\chi_1 - \chi_2)\phi$ étant nulle sur K voisinage de Supp u, on a

$$\langle u, (\chi_1 - \chi_2)\phi \rangle = 0.$$

On a ainsi associé à une distribution à support compact une forme linéaire sur C^∞. Comme $\mathcal{S}(\mathbb{R}^d) \subset C^\infty(\mathbb{R}^d)$, on peut définir (voir remarque 7.11)

$$\langle u, \phi \rangle_{\mathcal{S}',\mathcal{S}} = \langle u, \phi \rangle_{\mathcal{E}',C^\infty}.$$

On a

$$|\langle u, \phi \rangle_{\mathcal{S}',\mathcal{S}}| = |\langle u, \chi\phi \rangle| \leq C \sup_{|\alpha|\leq p} \|\partial^\alpha(\chi\phi)\|_{L^\infty} \leq C' \sup_{|\alpha|\leq p} \|\partial^\alpha \phi\|_{L^\infty},$$

où C et C' désignent deux réels positifs ne dépendant pas de ϕ.

$\square$

Remarque 10.24. On peut donc écrire

$$\mathcal{D}(\mathbb{R}^d) \hookrightarrow \mathcal{S}(\mathbb{R}^d) \hookrightarrow L^2(\mathbb{R}^d) \hookrightarrow \mathcal{S}'(\mathbb{R}^d) \hookrightarrow \mathcal{D}'(\mathbb{R}^d)$$

où la notation $A \hookrightarrow B$ signifie que A s'injecte dans B et que l'injection de A dans B est séquentiellement continue. Là encore, l'espace L^2 joue un rôle d'espace pivot.

10.3.4 Transformation de Fourier dans $\mathcal{S}'$

Définition - Théorème 10.25. *Soit $u \in \mathcal{S}'(\mathbb{R}^d)$. La* transformée de Fourier *de u est la distribution tempérée notée $\hat{u}$ ou $\mathcal{F}u$ définie par*

$$\forall \phi \in \mathcal{S}(\mathbb{R}^d), \qquad \langle \hat{u}, \phi \rangle_{\mathcal{S}',\mathcal{S}} = \langle u, \hat{\phi} \rangle_{\mathcal{S}',\mathcal{S}}.$$

La transformation de Fourier ainsi définie est une extension de la définition classique de la transformation de Fourier sur $L^1(\mathbb{R}^d)$.

Preuve. Vérifions d'abord que $\hat{u}$ est bien une distribution tempérée. Il est clair que $\hat{u}$ est une forme linéaire sur $\mathcal{S}(\mathbb{R}^d)$. Il reste à s'assurer que $\hat{u}$ vérifie la propriété de continuité (10.7). Soit $\phi \in \mathcal{S}(\mathbb{R}^d)$. On a

$$\begin{aligned} |\langle \hat{u}, \phi \rangle| &= |\langle u, \hat{\phi} \rangle| \\ &\leq C\mathcal{N}_p(\hat{\phi}) \\ &\leq C'\mathcal{N}_{p+d+1}(\phi). \end{aligned}$$

On a donc bien $\hat{u} \in \mathcal{S}'(\mathbb{R}^d)$. Soit maintenant $u \in L^1(\mathbb{R}^d)$. On a pour tout $\phi \in \mathcal{S}(\mathbb{R}^d)$,

$$\begin{aligned} \langle \hat{u}, \phi \rangle_{\mathcal{S}',\mathcal{S}} &= \langle u, \hat{\phi} \rangle_{\mathcal{S}',\mathcal{S}} \\ &= \int_{\mathbb{R}^d} u\hat{\phi} \\ &= \int_{\mathbb{R}^d} u(x) \left(\int_{\mathbb{R}^d} \phi(y) e^{-ix\cdot y}\, dy \right) dx. \end{aligned}$$

Comme $u(x)\phi(y)e^{-ix\cdot y} \in L^1(\mathbb{R}^d \times \mathbb{R}^d)$, on obtient en appliquant le théorème 4.50 de Fubini

$$\begin{aligned} \langle \hat{u}, \phi \rangle_{\mathcal{S}',\mathcal{S}} &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} u(x) e^{-ix\cdot y}\, dx \right) \phi(y)\, dy \\ &= \langle \int_{\mathbb{R}^d} u(x) e^{-ix\cdot y}\, dx, \phi \rangle_{\mathcal{S}',\mathcal{S}}. \end{aligned}$$

Il en résulte que

$$\hat{u}(y) = \int_{\mathbb{R}^d} u(x) e^{-ix\cdot y}\, dx.$$

□

Les propriétés de la transformée de Fourier énoncées dans les sections 10.1.3 et 10.1.4, valables dans L^1, donc dans $\mathcal{S}$, peuvent ainsi se transporter sur $\mathcal{S}'$. Nous avons notamment les résultats suivants :

Théorème 10.26. *La transformée de Fourier est un isomorphisme séquentiellement bicontinu de $\mathcal{S}'(\mathbb{R}^d)$ sur lui-même, d'inverse $\mathcal{F}^{-1}$ défini par*

$$\forall \phi \in \mathcal{S}(\mathbb{R}^d), \quad \langle \mathcal{F}^{-1}u, \phi\rangle_{\mathcal{S}',\mathcal{S}} = \langle u, \mathcal{F}^{-1}\phi\rangle_{\mathcal{S}',\mathcal{S}}.$$

Preuve. La continuité séquentielle de $\mathcal{F}$ se vérifie sans difficultés : soit $(u_n)_{n\in\mathbb{N}}$ une suite de $\mathcal{S}'(\mathbb{R}^d)$ et $\phi \in \mathcal{S}(\mathbb{R}^d)$, on a

$$\langle \mathcal{F}u_n, \phi\rangle_{\mathcal{S}',\mathcal{S}} = \langle u_n, \hat{\phi}\rangle_{\mathcal{S}',\mathcal{S}} \underset{n\to+\infty}{\longrightarrow} \langle u, \hat{\phi}\rangle_{\mathcal{S}',\mathcal{S}} = \langle \mathcal{F}u, \phi\rangle_{\mathcal{S}',\mathcal{S}}.$$

On vérifie de même que $\mathcal{F}^{-1}$ est un endormorphisme séquentiellement continu sur $\mathcal{S}'(\mathbb{R}^d)$. Soit maintenant $u \in \mathcal{S}'(\mathbb{R}^d)$. On a pour tout $\phi \in \mathcal{S}(\mathbb{R}^d)$

$$\langle \mathcal{F}^{-1}\mathcal{F}u, \phi\rangle_{\mathcal{S}',\mathcal{S}} = \langle \mathcal{F}u, \mathcal{F}^{-1}\phi\rangle_{\mathcal{S}',\mathcal{S}} = \langle u, \mathcal{F}\mathcal{F}^{-1}\phi\rangle_{\mathcal{S}',\mathcal{S}} = \langle u, \phi\rangle_{\mathcal{S}',\mathcal{S}}.$$

Donc $\mathcal{F}^{-1}\mathcal{F}u = u$ et de même $\mathcal{F}\mathcal{F}^{-1}u = u$. □

Théorème 10.27. *Soit $u \in \mathcal{S}'(\mathbb{R}^d)$. Alors $\dfrac{\partial u}{\partial x_j} \in \mathcal{S}'(\mathbb{R}^d)$ et $x_j u \in \mathcal{S}'(\mathbb{R}^d)$ et*

$$\mathcal{F}\left(\frac{\partial u}{\partial x_j}\right) = i\xi_j \mathcal{F}(u),$$

$$\mathcal{F}\left(x_j u\right) = i\frac{\partial\left(\mathcal{F}(u)\right)}{\partial \xi_j}$$

Preuve. Soit $\phi \in \mathcal{S}(\mathbb{R}^d)$. On a

$$\begin{aligned}
\langle \mathcal{F}\left(\frac{\partial u}{\partial x_j}\right), \phi\rangle_{\mathcal{S}',\mathcal{S}} &= \langle \frac{\partial u}{\partial x_j}, \mathcal{F}\phi\rangle_{\mathcal{S}',\mathcal{S}} \\
&= -\langle u, \frac{\partial}{\partial x_j}(\mathcal{F}\phi)\rangle_{\mathcal{S}',\mathcal{S}} \\
&= -\langle u, (\mathcal{F}(-i\xi_j\phi))\rangle_{\mathcal{S}',\mathcal{S}} \\
&= -\langle \mathcal{F}u, -i\xi_j\phi\rangle_{\mathcal{S}',\mathcal{S}} \\
&= \langle i\xi_j\mathcal{F}u, \phi\rangle_{\mathcal{S}',\mathcal{S}}.
\end{aligned}$$

De même,

$$\begin{aligned}
\langle \mathcal{F}(x_j u), \phi\rangle_{\mathcal{S}',\mathcal{S}} &= \langle x_j u, \mathcal{F}\phi\rangle_{\mathcal{S}',\mathcal{S}} \\
&= \langle u, x_j \mathcal{F}\phi\rangle_{\mathcal{S}',\mathcal{S}} \\
&= \langle u, (\mathcal{F}(-i\frac{\partial \phi}{\partial \xi_j}))\rangle_{\mathcal{S}',\mathcal{S}} \\
&= \langle \mathcal{F}u, -i\frac{\partial \phi}{\partial \xi_j}\rangle_{\mathcal{S}',\mathcal{S}} \\
&= \langle i\frac{\partial(\mathcal{F}(u))}{\partial \xi_j}, \phi\rangle_{\mathcal{S}',\mathcal{S}}.
\end{aligned}$$

□

10.4 Transformation de Fourier dans L^2

Nous venons de définir la transformée de Fourier sur l'espace $\mathcal{S}'$ qui contient (largement !) L^2. Nous allons maintenant étudier des propriétés spécifiques à la transformée de Fourier dans L^2, qui ont de nombreuses applications en analyse.

10.4.1 Propriété d'isométrie

Théorème 10.28. *Les transformations* $\frac{1}{(2\pi)^{d/2}}\mathcal{F}$ *et* $(2\pi)^{d/2}\mathcal{F}^{-1}$ *sont des isométries de* $L^2(\mathbb{R}^d)$ *inverses l'une de l'autre. On a donc*

$$\forall f \in L^2(\mathbb{R}^d), \quad \mathcal{F}f \in L^2(\mathbb{R}^d) \;\; et \;\; \frac{1}{(2\pi)^{d/2}}\|\mathcal{F}f\|_{L^2} = \|f\|_{L^2}, \tag{10.11}$$

$$\forall f \in L^2(\mathbb{R}^d), \quad \mathcal{F}^{-1}f \in L^2(\mathbb{R}^d) \;\; et \;\; (2\pi)^{d/2}\|\mathcal{F}^{-1}f\|_{L^2} = \|f\|_{L^2}.$$

Preuve. Soit $\phi \in \mathcal{S}(\mathbb{R}^d)$ et $\psi \in \mathcal{S}(\mathbb{R}^d)$. La fonction $e^{-ix\cdot\xi}\phi(x)\overline{\psi(\xi)}$ étant dans $L^1(\mathbb{R}^d \times \mathbb{R}^d)$, on obtient en appliquant le théorème 4.50 de Fubini

$$\begin{aligned}
\frac{1}{(2\pi)^d}(\mathcal{F}\phi, \psi)_{L^2} &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \overline{(\mathcal{F}\phi)(\xi)}\,\psi(\xi)\,d\xi \\
&= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \overline{\left(\int_{\mathbb{R}^d}\phi(x)\,e^{-ix\cdot\xi}\,dx\right)}\,\psi(\xi)\,d\xi \\
&= \int_{\mathbb{R}^d}\overline{\phi(x)}\left(\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\psi(\xi)\,e^{ix\cdot\xi}\,d\xi\right)dx \\
&= \int_{\mathbb{R}^d}\overline{\phi(x)}\,(\mathcal{F}^{-1}\psi)(x)\,dx \\
&= (\phi, \mathcal{F}^{-1}\psi)_{L^2}.
\end{aligned}$$

Appliquons cette formule avec $\psi = \mathcal{F}\phi$; il vient

$$\frac{1}{(2\pi)^d}\|\mathcal{F}\phi\|_{L^2}^2 = \frac{1}{(2\pi)^d}(\mathcal{F}\phi, \mathcal{F}\phi)_{L^2} = (\phi, \mathcal{F}^{-1}\mathcal{F}\phi)_{L^2} = (\phi, \phi)_{L^2} = \|\phi\|_{L^2}^2.$$

Il en résulte que $\frac{1}{(2\pi)^{d/2}}\mathcal{F}$ est une isométrie sur $\mathcal{S}$.

Soit maintenant $f \in L^2(\mathbb{R}^d)$. L'espace $\mathcal{S}(\mathbb{R}^d)$, qui contient $\mathcal{D}(\mathbb{R}^d)$, est dense dans $L^2(\mathbb{R}^d)$. Considérons donc une suite $(\phi_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{S}(\mathbb{R}^d)$ qui converge vers f dans L^2. La suite $(\phi_n)_{n\in\mathbb{N}}$ converge, donc est de Cauchy dans $L^2(\mathbb{R}^d)$. La suite des transformées de Fourier $(\hat{\phi}_n)_{n\in\mathbb{N}}$ est alors elle-aussi une suite de Cauchy puisque

$$\forall 0 \leq m \leq n, \quad \|\hat{\phi}_m - \hat{\phi}_n\|_{L^2} = \|\mathcal{F}(\phi_m - \phi_n)\|_{L^2} = (2\pi)^{d/2}\|\phi_m - \phi_n\|_{L^2}.$$

De ce fait, la suite $(\hat{\phi}_n)_{n\in\mathbb{N}}$ converge dans $L^2(\mathbb{R}^d)$ vers une certaine fonction $g \in L^2(\mathbb{R}^d)$. Or on sait par le théorème 10.23 que la convergence dans L^2 entraîne la convergence dans $\mathcal{S}'$. Donc $(\hat{\phi}_n)_{n\in\mathbb{N}}$ converge dans $S'(\mathbb{R}^d)$ vers g. Mais d'autre part, $(\phi_n)_{n\in\mathbb{N}}$ converge vers f dans L^2, donc dans $\mathcal{S}'$; il résulte du théorème 10.26 que $(\hat{\phi}_n)_{n\in\mathbb{N}}$ converge vers $\hat{f}$ dans $\mathcal{S}'$. Il s'en suit que $\hat{f} = g$ et donc que $\hat{f} \in L^2(\mathbb{R}^d)$ et que $(\hat{\phi}_n)_{n\in\mathbb{N}}$ converge vers $\hat{f}$ dans $L^2(\mathbb{R}^d)$. En passant à la limite dans la relation

$$\forall n \in \mathbb{N}, \quad \frac{1}{(2\pi)^d}\|\mathcal{F}\phi_n\|_{L^2}^2 = \|\mathcal{F}\phi_n\|_{L^2}^2,$$

on obtient l'égalité

$$\frac{1}{(2\pi)^d}\|\mathcal{F}f\|_{L^2}^2 = \|f\|_{L^2}^2$$

L'application $\dfrac{1}{(2\pi)^{d/2}}\mathcal{F}$ est donc une isométrie sur L^2. En remplaçant ϕ par $\mathcal{F}^{-1}\phi$, on montre qu'il en est de même de $(2\pi)^{d/2}\mathcal{F}^{-1}$. □

Remarque 10.29. L'égalité (10.11) s'écrit aussi pour $f \in L^2(\mathbb{R}^d)$

$$\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} |\hat{f}(\xi)|^2\, d\xi = \int_{\mathbb{R}^d} |f(x)|^2\, dx. \tag{10.12}$$

Il arrive souvent en physique ou en mécanique que l'énergie d'un champ sur $\mathbb{R}^d$ s'exprime précisément sous la forme d'une intégrale sur l'espace du carré du champ (penser à l'énergie cinétique d'un écoulement, à l'énergie d'un champ électromagnétique, à l'énergie de déformation élastique linéaire d'un matériau homogène isotrope). L'égalité (10.12) signifie que l'énergie du champ peut alors être calculée indifféremment dans l'espace réel ou dans l'espace réciproque.

La proposition suivante fournit une méthode de calcul de la transformée de Fourier des fonctions de classe L^2 qui ne sont pas dans L^1.

Proposition 10.30. *Soit $f \in L^2(\mathbb{R}^d)$. Si*

$$g_R(\xi) = \int_{|x|<R} f(x)\, e^{-ix\cdot\xi}\, d\xi \underset{R\to+\infty}{\longrightarrow} g(\xi) \text{ presque partout,}$$

alors $g \in L^2(\mathbb{R}^d)$ et $\mathcal{F}f = g$.

Preuve. Soit $f_R = f\,\chi_{\{|x|<R\}}$. Il est clair que famille de fonctions $(f_R)_{R>0}$ tend vers f dans $L^2(\mathbb{R}^d)$. Il en résulte que $\hat{f}_R = g_R$ converge vers $\hat{f}$ dans $L^2(\mathbb{R}^d)$ lorsque R tend vers l'infini. Il existe donc une suite extraite $(f_{R_n})_{n\in\mathbb{N}}$ avec $\lim\limits_{k\to+\infty} R_n = +\infty$ qui converge presque partout vers $\hat{f}$. On en déduit que $\hat{f} = g$. □

10.4.2 Espaces de Sobolev fractionnaires

La tranformation de Fourier permet de définir sur $\mathbb{R}^d$ des espaces de Sobolev d'exposant réel qui coïncident avec ceux construits au chapitre précédent lorsque l'exposant est un entier naturel. Grâce à cette nouvelle définition, nous allons pouvoir relier les propriétés de régularité locale d'une fonction L^2 (son caractère continu, C^1, C^2, ...) au comportement de sa transformée de Fourier à l'infini (i.e. dans le domaine des hautes fréquences).

Définition - Théorème 10.31. *Pour tout $s \in \mathbb{R}$, on pose*

$$H^s(\mathbb{R}^d) = \left\{ u \in \mathcal{S}'(\mathbb{R}^d), \quad \hat{u} \in L^1_{loc}(\mathbb{R}^d), \quad \int_{\mathbb{R}^d} (1+|\xi|^2)^s \, |\hat{u}(\xi)|^2 \, d\xi < +\infty \right\}. \tag{10.13}$$

Muni du produit scalaire noté $(\cdot,\cdot)_{H^s}$ et défini par

$$\forall (u,v) \in H^s(\mathbb{R}^d) \times H^s(\mathbb{R}^d), \quad (u,v)_{H^s} = \int_{\mathbb{R}^d} (1+|\xi|^2)^s \, \overline{\hat{u}(\xi)} \, \hat{v}(\xi) \, d\xi, \tag{10.14}$$

l'espace $H^s(\mathbb{R}^d)$ est un espace de Hilbert.

Preuve. Il est clair que $(\cdot,\cdot)_{H^s}$ est un produit scalaire. Pour montrer la complétude, il suffit de remarquer que pour toute suite $(u_n)_{n\in\mathbb{N}}$ de $H^s(\mathbb{R}^d)$, la suite $(v_n)_{n\in\mathbb{N}}$ définie par $v_n(\xi) = (1 + |\xi|^2)^{s/2} \hat{u}_n(\xi)$ est une suite de $L^2(\mathbb{R}^d)$ vérifiant

$$\forall 0 \le k \le l, \quad \|u_k - u_l\|_{H^s} = \|v_k - v_l\|_{L^2}.$$

Si la suite $(u_n)_{n\in\mathbb{N}}$ est de Cauchy dans $H^s(\mathbb{R}^d)$, la suite $(v_n)_{n\in\mathbb{N}}$ est de Cauchy dans L^2 ; cette dernière converge donc dans l'espace $L^2(\mathbb{R}^d)$ vers $v \in L^2(\mathbb{R}^d)$. Posons $u = \mathcal{F}^{-1}\left(\dfrac{v}{(1+|\xi|^2)^{s/2}}\right)$. Il est clair que $u \in H^s(\mathbb{R}^d)$ et que $(u_n)_{n\in\mathbb{N}}$ tend vers u dans $H^s(\mathbb{R}^d)$. □

Comme annoncé, cette définition des espaces de Sobolev $H^s(\mathbb{R}^d)$ coïncide, lorsque s est un entier naturel, avec la définition 8.1 donnée au chapitre précédent :

Proposition 10.32. *Soit $s \in \mathbb{N}$. On a*

$$H^s(\mathbb{R}^d) = \left\{ u \in L^2(\mathbb{R}^d), \quad \partial^\alpha u \in L^2(\mathbb{R}^d), \quad \forall \alpha \in \mathbb{N}^d, \quad |\alpha| \le s \right\}, \tag{10.15}$$

et la norme définie sur H^s par le produit (10.14) est équivalente à celle définie par le produit scalaire

$$(u,v) = \sum_{\alpha \le \mathbb{N}^d, \, |\alpha| \le s} \int_{\mathbb{R}^d} \overline{\partial^\alpha u(x)} \, \partial^\alpha v(x) \, dx.$$

Preuve. Nous montrons le résultat pour $H^1(\mathbb{R}^d)$ et laissons le cas général en exercice au lecteur. Soit u appartenant à l'espace H^1 défini par (10.15). On a $u \in L^2(\mathbb{R}^d)$ et $\dfrac{\partial u}{\partial x_j} \in L^2(\mathbb{R}^d)$ pour tout $1 \le j \le n$. Il en résulte que

- $u \in \mathcal{S}'(\mathbb{R}^d)$ (puisque $L^2(\mathbb{R}^d) \subset \mathcal{S}'(\mathbb{R}^d)$),
- $\hat{u} \in L^2(\mathbb{R}^d) \subset L^1_{loc}(\mathbb{R}^d)$ (puisque $u \in L^2(\mathbb{R}^d)$),
- $\xi_j \hat{u} \in L^2(\mathbb{R}^d)$ pour tout $1 \le j \le n$ (puisque $\xi_j \hat{u} = -i\mathcal{F}\left(\dfrac{\partial f}{\partial x_j}\right)$).

La fonction u appartient donc à l'espace H^1 défini par (10.13). La réciproque se démontre en utilisant les mêmes arguments. $\square$

La transformation de Fourier permet de montrer très facilement le résultat suivant qui relie la "décroissance" ou plus exactement l'"intégrabilité" de la transformée de Fourier à l'infini à la régularité locale de la fonction d'origine :

Théorème 10.33. *Soit $d \in \mathbb{N}^*$, $k \in \mathbb{N}$ et $s \in \mathbb{R}$ vérifiant $s > k + d/2$. Si $u \in H^s(\mathbb{R}^d)$, alors u est de classe C^k et toutes ses dérivées jusqu'à l'ordre k tendent vers zéro à l'infini.*

Preuve. Nous donnons la preuve de ce résultat pour $k = 0$ et laissons le lecteur compléter la preuve (c'est immédiat). Soit donc $s > d/2$ et $u \in H^s(\mathbb{R}^d)$. On a

$$\hat{u}(\xi) = \left((1 + |\xi|^2)^{s/2}\hat{u}(\xi)\right) \frac{1}{(1 + |\xi|^2)^{s/2}}.$$

Comme $u \in H^s(\mathbb{R}^d)$, on a par hypothèse

$$(1 + |\xi|^2)^{s/2}\hat{u}(\xi) \in L^2(\mathbb{R}^d)$$

et d'autre part

$$\frac{1}{(1 + |\xi|^2)^{s/2}} \in L^2(\mathbb{R}^d)$$

puisque $s > d/2$. On a donc $\hat{u} \in L^1(\mathbb{R}^d)$, d'où il découle que $u = \mathcal{F}^{-1}\hat{u}$ est continue et tend vers zéro à l'infini (il est clair que la proposition 10.2 fonctionne aussi pour $\mathcal{F}^{-1}$). $\square$

Pour le moment, nous n'avons traité que de fonctions définies sur $\mathbb{R}^d$. Comme la régularité d'une fonction est une propriété locale, on s'attend à ce que le théorème précédent s'adapte au cas d'un ouvert quelconque et, en effet, c'est bien le cas.

Définition 10.34. *Soit Ω un ouvert de $\mathbb{R}^d$ et s un réel. On note*

$$H^s_{\text{loc}}(\Omega) = \left\{u \in \mathcal{D}'(\Omega), \quad \forall \phi \in \mathcal{D}(\Omega), \quad \phi u \in H^s(\mathbb{R}^d)\right\}.$$

Cette définition appelle plusieurs commentaires. Tout d'abord, précisons que la distribution ϕu apparaissant dans l'expression $\phi u \in H^s(\mathbb{R}^d)$ est définie par

$$\forall \psi \in \mathcal{D}(\mathbb{R}^d), \qquad \langle \phi u, \psi \rangle_{\mathcal{D}'(\mathbb{R}^d), \mathcal{D}(\mathbb{R}^d)} = \langle u, \phi\, \psi|_\Omega \rangle_{\mathcal{D}'(\Omega), \mathcal{D}(\Omega)}.$$

En second lieu, on voit sans difficulté que pour $s = 0$, la définition 10.34 est consistante avec la définition de l'espace $L^2_{\text{loc}}(\Omega)$ donnée dans la section 5.5.1. Enfin, on pourra vérifier que pour s entier naturel, la définition 10.34 est équivalente à la définition

$$H^s_{\text{loc}}(\Omega) = \left\{u \in L^2_{\text{loc}}(\Omega), \quad \forall \alpha \in \mathbb{N}^d, \quad |\alpha| \le s, \quad \partial^\alpha u \in L^2_{\text{loc}}(\Omega)\right\}.$$

Théorème 10.35. *Soit Ω un ouvert de $\mathbb{R}^d$ et $u \in H^s_{\text{loc}}(\Omega)$ pour un certain $s \in \mathbb{R}$. Alors si $s > k + d/2$ pour un certain $k \in \mathbb{N}$, u est de classe C^k dans Ω.*

Preuve. Laissée en exercice. $\square$

Notons bien que le théorème 10.35 ne dit rien sur la régularité de u au bord de Ω.

Remarque 10.36. *Pour tout réel $0 < s < 1$, il est possible de définir l'espace $H^s(\mathbb{R}^d)$ par*

$$H^s(\mathbb{R}^d) = \left\{ u \in L^2(\mathbb{R}^d), \quad \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{|u(x) - u(y)|^2}{|x-y|^{2s+d}} \, dx \, dy < +\infty \right\}, \tag{10.16}$$

cette définition étant équivalente à la définition (10.13). On peut alors définir l'espace $H^s(\mathbb{R}^d)$ pour tout réel positif s par (10.15) si $s \in \mathbb{N}$ et par

$$H^s(\mathbb{R}^d) = \left\{ u \in L^2(\mathbb{R}^d), \quad \forall \alpha \in \mathbb{N}^d, \ |\alpha| \leq [s], \ \partial^\alpha u \in L^2(\mathbb{R}^d), \quad \forall \alpha \in \mathbb{N}^d, \ |\alpha| = [s], \ \partial^\alpha u \in H^{s-[s]}(\mathbb{R}^d) \right\} \tag{10.17}$$

si $s \in [1, +\infty[\setminus \mathbb{N}$, $[s]$ désignant la partie entière de s. Là encore, la définition (10.17) coïncide avec la définition (10.13). L'intérêt des définitions (10.16) et (10.17) est qu'elles s'étendent

- *aux ouverts $\Omega \subset \mathbb{R}^d$, ainsi qu'aux sous-variétés différentielles de $\mathbb{R}^d$ (en particulier aux bords des domaines réguliers) ;*
- *aux espaces L^p avec $p \neq 2$.*

10.4.3 Retour sur la notion de trace

A la section 8.2, nous avions énoncé sans démonstration le résultat suivant : si Ω est un ouvert régulier, on peut construire une application "trace" linéaire et continue

$$\begin{array}{rccc} \gamma \, : & H^1(\Omega) & \longrightarrow & L^2(\partial\Omega) \\ & u & \longmapsto & \gamma(u) \end{array}$$

vérifiant $\forall u \in H^1(\Omega) \cap C^0(\overline{\Omega})$, $\gamma(u) = u|_{\partial\Omega}$. Ceci nous avait permis d'interpréter les fonctions de $H^1_0(\Omega)$ comme les fonctions de $H^1(\Omega)$ "nulles au bord". Nous allons esquisser la preuve de ce résultat (et en fait d'un résultat beaucoup plus fort) dans le cas particulier où Ω est le demi-espace $]0, +\infty[\times \mathbb{R}^{d-1}$. Pour cela, commençons par établir le théorème suivant :

Théorème 10.37. *Soit $s > 1/2$. L'opérateur linéaire qui à une fonction $\phi \in \mathcal{S}(\mathbb{R}^d)$ associe la fonction $\phi(0, \cdot) \in \mathcal{S}(\mathbb{R}^{d-1})$ se prolonge de manière unique en un opérateur linéaire continu de $H^s(\mathbb{R}^d)$ dans $H^{s-1/2}(\mathbb{R}^{d-1})$, qu'on notera $\widetilde{\gamma}$. De plus, $\widetilde{\gamma}$ est surjectif.*

Preuve. Soit $\phi \in \mathcal{S}(\mathbb{R}^d)^d$. De par la formule d'inversion de Fourier,

$$\forall (x_1, x_2, \cdots, x_d) \in \mathbb{R}^d, \quad \phi(x_1, x_2, \cdots, x_d) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{\phi}(\xi_1, \cdots, \xi_d) \exp\left(i \sum_{j=1}^d \xi_j x_j \right) dx_1 \cdots dx_d.$$

Donc pour tout $(x_2, \cdots, x_d) \in \mathbb{R}^{d-1}$,

$$\begin{aligned} [\widetilde{\gamma}(\phi)] (x_2, \cdots, x_d) &= \phi(0, x_2, \cdots, x_d) \\ &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{\phi}(\xi_1, \cdots, \xi_d) \exp\left(i \sum_{j=2}^d \xi_j x_j \right) d\xi_1 \cdots d\xi_d \\ &= \frac{1}{(2\pi)^{d-1}} \int_{\mathbb{R}^{d-1}} \left(\frac{1}{2\pi} \int_{\mathbb{R}} \hat{\phi}(\xi_1, \cdots, \xi_d) \, d\xi_1 \right) \exp\left(i \sum_{j=1}^d \xi_j x_j \right) d\xi_2 \cdots d\xi_d. \end{aligned}$$

D'où

$$\forall \xi' \in \mathbb{R}^{d-1}, \quad \widehat{[\widetilde{\gamma}(\phi)]}(\xi') = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{\phi}(\xi_1, \xi')\, d\xi_1,$$

ce qui conduit après calcul à l'inégalité

$$\int_{\mathbb{R}^{d-1}} (1+|\xi'|^2)^{s-1/2} \left| \widehat{[\widetilde{\gamma}(\phi)]}(\xi') \right|^2 d\xi' \leq \frac{C_s}{4\pi^2} \int_{\mathbb{R}^d} (1+|\xi|^2)^s \left| \widehat{\phi}(\xi) \right|^2 d\xi$$

avec $C_s = \displaystyle\int_{\mathbb{R}} \frac{dt}{(1+t^2)^s}\, dt$. Il s'en suit que

$$\forall \phi \in \mathcal{S}(\mathbb{R}^d), \qquad \|\widetilde{\gamma}(\phi)\|_{H^{s-1/2}} \leq \frac{\sqrt{C_s}}{2\pi} \|\phi\|_{H^s}.$$

Comme $\mathcal{S}(\mathbb{R}^d)$ est dense dans $H^s(\mathbb{R}^d)$ (le montrer en exercice), il en résulte que $\widetilde{\gamma}$ est définie de manière unique sur $H^s(\mathbb{R}^d)$ et que

$$\forall u \in H^s(\mathbb{R}^d), \qquad \|\widetilde{\gamma}(u)\|_{H^{s-1/2}} \leq \frac{\sqrt{C_s}}{2\pi} \|u\|_{H^s}.$$

Soit enfin $v \in H^{s-1/2}(\mathbb{R}^{d-1})$. On pourra vérifier en exercice que la fonction

$$u = \frac{2\pi}{C_z} \mathcal{F}^{-1} \left(\frac{(1+|\xi'|^2)^{z-1/2}}{(1+|\xi|^2)^z} \hat{v}(\xi') \right).$$

(où par convention de notation $\xi = (\xi_1, \xi')$, avec $\xi_1 \in \mathbb{R}$, $\xi' \in \mathbb{R}^{d-1}$) est dans $H^s(\mathbb{R}^d)$ et vérifie $\widetilde{\gamma}(u) = v$ dès que $z > s/2 + 1/4$. Cela prouve la surjectivité de $\widetilde{\gamma}$. □

Ce théorème a deux conséquences importantes :

- si on dispose d'une fonction de $H^s(\mathbb{R}^d)$ avec $s > 1/2$, on peut considérer sa trace sur l'hyperplan $x_1 = 0$, et par suite, sur n'importe quel hyperplan affine de $\mathbb{R}^d$;
- si on dispose d'une fonction v de $H^{s-1/2}(\mathbb{R}^{d-1})$ avec $s > 1/2$, on peut construire une fonction u de $H^s(\mathbb{R}^d)$ dont la trace sur l'hyperplan $x_1 = 0$ (et par suite sur un hyperplan affine donné de $\mathbb{R}^d$) est égale à v. On appelle une telle fonction u un *relèvement* de v dans $H^s(\mathbb{R}^d)$.

A partir du théorème 10.37, on établir le résultat suivant :

Théorème 10.38. *Soit $\Omega =]0, +\infty[\times \mathbb{R}^{d-1}$. L'opérateur linéaire qui à une fonction de $C^\infty(\overline{\Omega})$ à support compact fait correspondre sa restriction à l'hyperplan $\partial\Omega = \{0\} \times \mathbb{R}^{d-1}$ se prolonge de manière unique en un opérateur linéaire γ, continu et surjectif de $H^1(\Omega)$ dans $H^{1/2}(\partial\Omega)$. De plus $H_0^1(\Omega) = Ker(\gamma)$.*

Remarque 10.39. En utilisant des outils de géométrie différentielle, il est possible de définir les espaces $H^s(\Omega)$ et $H^s(\partial\Omega)$ pour un ouvert régulier et de montrer que le théorème 10.38 s'étend en fait à tout ouvert Ω de classe C^1 de $\mathbb{R}^d$ de bord $\partial\Omega$ borné. L'argument heuristique est que le bord d'un ouvert régulier ressemble localement à un hyperplan. Pour faire les choses proprement, il faut travailler dans des cartes locales dans lesquelles le bord $\partial\Omega$ est effectivement représenté par un hyperplan, et montrer que les résultats obtenus ne dépendent pas du jeu de cartes locales choisi.

10.5 Distributions périodiques et séries de Fourier

10.5.1 Séries de Fourier

Nous proposons dans cette section une lecture "géométrique" de la théorie des séries de Fourier : l'ensemble

$$L^2_{\text{per}}(]0,T[) = \left\{ f \in L^2_{\text{loc}}(\mathbb{R},\mathbb{C}), \quad f(\cdot + T) = f(\cdot) \text{ p.p.} \right\},$$

des fonctions de $\mathbb{R}$ dans $\mathbb{C}$, L^2_{loc} et T-périodiques, forme un espace vectoriel possédant une structure d'espace de Hilbert complexe pour le produit scalaire

$$(f,g)_{L^2} = \int_0^T \overline{f(x)}\, g(x)\, dx,$$

et la famille de fonctions

$$\left(e_n(x) = \frac{1}{\sqrt{T}} e^{2i\pi nx/T} \right)_{n\in\mathbb{Z}}$$

constitue une base hilbertienne de cet espace. La décomposition d'une fonction de $L^2_{\text{per}}(]0,T[)$ dans cette base n'est autre que la décomposition de Fourier de cette fonction.

Définition 10.40. *Pour $f \in L^2_{\text{per}}(]0,T[)$, on appelle* coefficients de Fourier *de f les nombres*

$$c_n(f) = (e_n, f)_{L^2} = \frac{1}{\sqrt{T}} \int_0^T f(x)\, e^{-2i\pi nx/T}\, dx. \tag{10.18}$$

Théorème 10.41. *(Théorème de décomposition de Fourier). La famille $B = (e_n)_{n\in\mathbb{Z}}$ est une base hilbertienne de $H = L^2_{\text{per}}(]0,T[)$.*

De ce théorème et des résultats établis à la section 3.4 découle immédiatement le corollaire suivant :

Corollaire 10.42.

1. *Pour $f \in L^2_{\text{per}}(]0,T[)$, on a au sens de la convergence L^2_{loc}*

$$f = \sum_{n=-\infty}^{+\infty} c_n(f)\, e_n.$$

2. *Pour $f \in L^2_{\text{per}}(]0,T[)$, on a* l'égalité de Parseval

$$\sum_{n=-\infty}^{+\infty} |c_n(f)|^2 = \int_0^T |f(x)|^2\, dx.$$

Remarque 10.43. Soit $f \in L^2_{\text{per}}(]0,T[)$ et $g \in L^2_{\text{per}}(]0,T[)$. En appliquant l'égalité de Parseval à $f+g$ puis à $if+g$, on obtient

$$\int_0^T \overline{f(x)}\, g(x)\, dx = \sum_{n\in\mathbb{Z}} \overline{c_n(f)}\, c_n(g).$$

Preuve du théorème 10.41. On voit facilement que $(e_n)_{n\in\mathbb{Z}}$ est une famille orthonormée. En effet

$$(e_p, e_q)_{L^2} = \frac{1}{T}\int_0^T e^{-2i\pi px/T} e^{2i\pi qx/T}\, dx = \frac{1}{T}\int_0^T e^{2i\pi(q-p)x/T}\, dx = \delta_{pq}.$$

Soit $V = \overline{\mathrm{Vect}((e_n)_{n\in\mathbb{N}})}$ (la barre horizontale désigne l'adhérence dans $L^2_{\mathrm{per}}(]0,T[)$ et non la conjugaison complexe). V est un sous-espace fermé de $L^2_{\mathrm{per}}(]0,T[)$. Soit maintenant $f \in C^2(\mathbb{R}) \cap L^2_{\mathrm{per}}(]0,T[)$. La série

$$\sum_{n\in\mathbb{Z}} c_n(f)\, e_n$$

converge dans $L^2_{loc}(\mathbb{R})$ vers la projection orthogonale de f sur V. Pour évaluer sa somme, on part de la relation

$$\sum_{n=-N}^{N} e^{2i\pi nz/T} = \frac{\sin((2N+1)\pi z/T)}{\sin(\pi z/T)}$$

qui, intégrée sur $]0,T[$, conduit à l'égalité

$$1 = \frac{1}{T}\int_0^T \sum_{n=-N}^{N} e^{2i\pi nz/T}\, dz = \frac{1}{2\pi}\int_0^T \frac{\sin((2N+1)\pi z/T)}{\sin(\pi z/T)}\, dz.$$

On en déduit

$$\begin{aligned}
f(x) - \sum_{n=-N}^{N} c_n(f)\, e_n(x) &= \frac{1}{T}\int_0^T \left(\sum_{n=-N}^{N} e^{2i\pi n(x-y)/T}\, dy\right) f(x) - \frac{1}{T}\sum_{n=-N}^{N}\left(\int_0^T f(y) e^{-2i\pi ny/T}\, dy\right) e^{2i\pi nx/T} \\
&= \frac{1}{T}\int_0^T \left(\sum_{n=-N}^{N} e^{2i\pi n(x-y)/T}\right) (f(x) - f(y))\, dy \\
&= \frac{1}{T}\int_0^T \frac{f(x)-f(y)}{\sin(\pi(x-y)/T)} \sin((2N+1)\pi(x-y)/T)\, dy.
\end{aligned}$$

Comme la fonction

$$y \mapsto \frac{f(x)-f(y)}{\sin(\pi(x-y)/T)}$$

est de classe C^1, l'intégrale ci-dessus tend vers 0 lorsque N tend vers $+\infty$ (on le voit en intégrant par parties), ce qui implique que pour tout $x \in]0,T[$,

$$f(x) = \sum_{n\in\mathbb{Z}} c_n(f)\, e_n(x).$$

Donc $f \in V$. Il en résulte que $C^2(\mathbb{R}) \cap L^2_{\mathrm{per}}(]0,T[) \subset V$. Comme $C^2(\mathbb{R}) \cap L^2_{\mathrm{per}}(]0,T[)$ est dense dans $L^2_{\mathrm{per}}(]0,T[)$ (d'après une adaptation du corollaire 6.17), il vient $V = L^2\mathrm{per}(]0,T[)$. $\square$

10.5.2 Représentation des distributions périodiques

Rappelons que pour $T \in \mathbb{R}$, la translation τ_T opère sur une distribution $U \in \mathcal{D}'(\mathbb{R})$ de la façon suivante :

$$\forall \phi \in \mathcal{D}(\mathbb{R}), \qquad \langle \tau_T U, \phi\rangle = \langle U, \phi(\cdot + T)\rangle = \langle U, \tau_{-T}\phi\rangle.$$

Définition 10.44. *Soit $U \in \mathcal{D}'(\mathbb{R})$. On dit que U est T-périodique si $\tau_T U = U$. On note $\mathcal{D}'_T(\mathbb{R})$ l'espace vectoriel des distributions T-périodique.*

Définition 10.45. *On dit qu'une suite $(\gamma_n)_{n\in\mathbb{Z}}$ est à croissance lente s'il existe $p \in \mathbb{N}$ et $C \geq 0$ tels que*

$$\forall n \in \mathbb{Z}, \quad |\gamma_n| \leq C\,(1+|n|)^p.$$

Théorème 10.46. *Soit $T > 0$ et $(\gamma_n)_{n\in\mathbb{Z}}$ une suite réelle à croissance lente. La série*

$$\sum_{n=-\infty}^{+\infty} \gamma_n\, e^{2i\pi n\, t/T}$$

définit une distribution T-périodique.

Preuve. Soit $p \in \mathbb{N}$ et $C \geq 0$ tel que

$$\forall n \in \mathbb{Z}, \quad |\gamma_n| \leq C\,(1+|n|)^p.$$

La série de fonction

$$\sum_{n\in\mathbb{Z}^*} \frac{\gamma_n}{(2i\pi n/T)^{p+2}}\, e^{2i\pi n\, t/T}$$

est normalement donc uniformément convergente sur $\mathbb{R}$, et converge donc uniformément vers une fonction continue f. Cette série converge donc aussi vers f dans $\mathcal{D}'(\mathbb{R})$. Dérivons $p+2$ fois dans $\mathcal{D}'$ la relation

$$f = \sum_{n\in\mathbb{Z}^*} \frac{\gamma_n}{(2i\pi n/T)^{p+2}}\, e^{2i\pi n\, t/T}.$$

On obtient ainsi que la série

$$\sum_{n\in\mathbb{Z}^*} \gamma_n\, e^{2i\pi n\, t/T}$$

converge dans $\mathcal{D}'$ (accessoirement vers $f^{(p+2)}$). Il en va donc de même pour la série

$$\sum_{n\in\mathbb{Z}} \gamma_n\, e^{2i\pi n\, t/T}.$$

Comme les sommes partielles

$$\sum_{n=-N}^{N} \gamma_n\, e^{2i\pi n\, t/T}$$

sont toutes des distributions T-périodiques, il en est de même pour la série. $\square$

Théorème 10.47. *Soit U une distribution T-périodique. Il existe une unique suite réelle à croissance lente $(\gamma_n)_{n\in\mathbb{Z}}$ telle que*

$$U = \sum_{n=-\infty}^{+\infty} \gamma_n\, e^{2i\pi n\, t/T}.$$

Les $(\gamma_n)_{n\in\mathbb{Z}}$ sont appelés coefficients de Fourier de la distribution U.

Pour faciliter la lecture, nous avons découpé la preuve de ce théorème en une collection de lemmes.

Lemme 10.48. *(Théorème de Fejer-Dirichlet). Soit $f \in C^\infty$ une fonction T-périodique. On note*

$$c_n(f) = \frac{1}{T}\int_0^T f(t)\, e^{-2i\pi n\, t/T}\, dt.$$

1. *La série de Fourier*
$$\sum_{n=-\infty}^{+\infty} c_n(f)\, e^{2i\pi n\, t/T}$$
converge uniformément vers f.

2. *Pour tout $p \geq 0$, la série de Fourier*
$$\sum_{n=-\infty}^{+\infty} c_n(f)\, (2i\pi n/T)^p\ e^{2i\pi n\, t/T}$$
converge uniformément vers $f^{(p)}$.

3. *La suite $(c_n(f))_{n\in\mathbb{Z}}$ est à décroissance rapide : pour tout N, il existe une constante C_N telle que*
$$|c_n(f)| \leq \frac{C_N}{(1+|n|)^N}.$$

Preuve. Laissée en exercice. □

Lemme 10.49. *Il existe $\chi \in \mathcal{D}(\mathbb{R})$ telle que*
$$\sum_{k\in\mathbb{Z}} \tau_{kT}\chi = 1.$$

Preuve. Soit $\phi \in \mathcal{D}(\mathbb{R})$ positive, égale à 1 sur l'intervalle $[0,T]$. La fonction
$$\sum_{k\in\mathbb{Z}} \tau_{kT}\phi$$
est de classe C^∞, T-périodique et strictement positive sur $\mathbb{R}$. On vérifie facilement que la fonction
$$\chi = \frac{\phi}{\sum_{k\in\mathbb{Z}} \tau_{kT}\phi}$$
répond aux critères. □

Lemme 10.50. *Soit $\phi \in \mathcal{D}(\mathbb{R})$ et $\widetilde{\phi} = \displaystyle\sum_{k\in\mathbb{Z}} \tau_{kT}\phi$. La fonction $\widetilde{\phi}$ est C^∞, T-périodique et*
$$c_n(\widetilde{\phi}) = \frac{1}{T}\int_{-\infty}^{+\infty} \phi(t)\, e^{-2i\pi n\, t/T}\, dt.$$

Preuve. Il est clair que $\widetilde{\phi}$ est C^∞ et T-périodique. De plus

$$\begin{aligned}
c_n(\widetilde{\phi}) &= \frac{1}{T}\int_0^T \left(\sum_{k\in\mathbb{Z}} \tau_{kT}\phi\right)(t)\, e^{-2i\pi n\, t/T}\, dt \\
&= \frac{1}{T}\sum_{k\in\mathbb{Z}}\int_0^T \phi(t-kT)\, e^{-2i\pi n\, t/T}\, dt \\
&= \frac{1}{T}\sum_{k\in\mathbb{Z}}\int_{-kT}^{(k+1)T} \phi(t)\, e^{-2i\pi n\,(t+kT)/T}\, dt \\
&= \frac{1}{T}\sum_{k\in\mathbb{Z}}\int_{-kT}^{(k+1)T} \phi(t)\, e^{-2i\pi n\, t/T}\, dt \\
&= \frac{1}{T}\int_{-\infty}^{+\infty} \phi(t)\, e^{-2i\pi n\, t/T}\, dt.
\end{aligned}$$

□

Lemme 10.51. *Soit $(\gamma_n)_{n\in\mathbb{Z}}$ une suite réelle à croissance lente et $\phi\in\mathcal{D}(\mathbb{R}^d)$. On a*

$$\langle\sum_{n\in\mathbb{Z}}\gamma_n e^{2i\pi n\, t/T},\phi\rangle = T\sum_{n\in\mathbb{Z}}\gamma_n c_{-n}(\widetilde{\phi}).$$

Preuve. On a pour tout $N\in\mathbb{N}$

$$\langle\sum_{n=-N}^{N}\gamma_n e^{2i\pi n\, t/T},\phi\rangle = \sum_{n=-N}^{N}\gamma_n\langle e^{2i\pi n\, t/T},\phi\rangle = T\sum_{n=-N}^{N}\gamma_n c_{-n}(\widetilde{\phi}).$$

□

Lemme 10.52. *L'application*

$$\begin{array}{ccc} SCL & \longrightarrow & \mathcal{D}'_T(\mathbb{R}) \\ (\gamma_n)_{n\in\mathbb{Z}} & \mapsto & \displaystyle\sum_{n\in\mathbb{Z}}\gamma_n e^{2i\pi n\, t/T} \end{array}$$

est injective.

Preuve. Supposons qu'il existe deux suites à croissance lente $(\gamma_n)_{n\in\mathbb{Z}}$ et $(\gamma'_n)_{n\in\mathbb{Z}}$ telles que

$$U = \sum_{k\in\mathbb{Z}}\gamma_n e^{2i\pi n\, t/T} = \sum_{k\in\mathbb{Z}}\gamma'_n e^{2i\pi n\, t/T}.$$

Pour tout $\phi\in\mathcal{D}(\mathbb{R})$,

$$\sum_{k\in\mathbb{Z}}(\gamma_n-\gamma'_n)c_{-n}(\widetilde{\phi}) = 0.$$

Appliquons cette égalité avec $\phi(t) = \chi(t)\, e^{2i\pi p_0\, t/T}$ où χ est la fonction construite au lemme 2 et $p_0\in\mathbb{Z}$. Comme

$$\widetilde{\phi}(t) = \sum_{k\in\mathbb{Z}}\tau_{kT}\left(\chi\, e^{2i\pi p_0\, t/T}\right) = \sum_{k\in\mathbb{Z}}\tau_{kT}\chi\cdot\tau_{kT}\left(e^{2i\pi p_0\, t/T}\right) = \sum_{k\in\mathbb{Z}}\tau_{kT}\chi\cdot e^{2i\pi p_0\, t/T} = e^{2i\pi p_0\, t/T},$$

il vient

$$\begin{aligned} c_n(\widetilde{\phi}) &= \frac{1}{T}\int_0^T \widetilde{\phi}(t)e^{-2i\pi n\,t/T}\,dt \\ &= \frac{1}{T}\int_0^T e^{2i\pi(p_0-n)\,t/T}\,dt \\ &= \delta_{p_0 n} \end{aligned}$$

On en déduit $\gamma_{p_0} = \gamma'_{p_0}$. Cela est vrai pour tout p_0. Donc les suites $(\gamma_n)_{n\in\mathbb{Z}}$ et $(\gamma'_n)_{n\in\mathbb{Z}}$ sont identiques. □

Lemme 10.53. *Soit $U \in \mathcal{D}'_T(\mathbb{R})$. Il existe $u \in \mathcal{E}'(\mathbb{R})$ telle que*

$$U = \sum_{n\in\mathbb{Z}} \tau_{nT} u.$$

Soit

$$\gamma_n = \frac{1}{T}\langle u, e^{-2i\pi n\,t/T}\rangle_{\mathcal{E}',C^\infty}.$$

Alors $(\gamma_n)_{n\in\mathbb{Z}}$ est à croissance lente et

$$\langle U, \phi\rangle = T\sum_{n\in\mathbb{Z}} \gamma_n c_{-n}(\widetilde{\phi}).$$

Preuve. On note χ la fonction construite au lemme 10.49. On a

$$U = 1\cdot U = \left(\sum_{k\in\mathbb{Z}} \tau_{kT}\chi\right)\cdot U = \sum_{k\in\mathbb{Z}} (\tau_{kT}\chi)\cdot U.$$

Posons $u = \chi\cdot U$. On a bien $u \in \mathcal{E}'(\mathbb{R})$. De plus

$$\tau_{kT}u = \tau_{kT}\,(\chi\cdot U) = (\tau_{kT}\chi)\cdot(\tau_{kT}U) = (\tau_{kT}\chi)\cdot U$$

puisque U est périodique. Il en résulte que

$$U = \sum_{k\in\mathbb{Z}} \tau_{kT}u.$$

Soit p l'ordre de u et K un voisinage compact de Supp u. Pour tout $n \in \mathbb{Z}$,

$$\begin{aligned} |\gamma_n| &= \left|\frac{1}{T}\langle u, e^{-2i\pi n\,t/T}\rangle_{\mathcal{E}',C^\infty}\right| \\ &\leq C\sup_{0\leq k\leq p}\sup_{t\in K}\left|\frac{d^k}{dt^k}e^{-2i\pi n\,t/T}\right| \\ &\leq C\,(1+|n|)^p. \end{aligned}$$

Enfin pour tout $\phi \in \mathcal{D}(\mathbb{R})$,

$$\begin{aligned} \langle U, \phi\rangle &= \langle\sum_{n\in\mathbb{Z}} \tau_{nT}u, \phi\rangle \\ &= \sum_{n\in\mathbb{Z}}\langle\tau_{nT}u, \phi\rangle \\ &= \sum_{n\in\mathbb{Z}}\langle u, \tau_{-nT}\phi\rangle \\ &= \langle u, \sum_{n\in\mathbb{Z}} \tau_{-nT}\phi\rangle \\ &= \langle u, \widetilde{\phi}\rangle. \end{aligned}$$

D'autre part, d'après le lemme 10.48,

$$\widetilde{\phi}(t) = \sum_{n\in\mathbb{Z}} c_{-n}(\widetilde{\phi})e^{-2i\pi n\, t/T}.$$

D'où,

$$\langle U, \phi\rangle = \langle u, \sum_{n\in\mathbb{Z}} c_{-n}(\widetilde{\phi})e^{-2i\pi n\, t/T}\rangle = \sum_{n\in\mathbb{Z}} \langle u, e^{-2i\pi n\, t/T}\rangle c_{-n}(\widetilde{\phi}) = T\sum_{n\in\mathbb{Z}} \gamma_n c_{-n}(\widetilde{\phi}).$$

□

Preuve du théorème 10.47. Soit $\phi \in \mathcal{D}(\mathbb{R})$. On a

$$\langle U, \phi\rangle = \sum_{n\in\mathbb{Z}} \gamma_n c_{-n}(\widetilde{\phi}) = \langle \sum_{n\in\mathbb{Z}} \gamma_n e^{2i\pi n\, t/T}, \phi\rangle.$$

Donc

$$U = \sum_{n\in\mathbb{Z}} \gamma_n e^{2i\pi n\, t/T}.$$

La suite $(\gamma_n)_{n\in\mathbb{Z}}$ est unique d'après le lemme 10.52 et donc indépendante du choix de la distribution u. □

10.5.3 Transformée de Fourier des distributions périodiques

Proposition 10.54. *Soit $a \in \mathbb{R}$. On a*

$$\mathcal{F}(e^{iat}) = 2\pi\delta_a, \qquad \mathcal{F}(\delta_a) = e^{-iat},$$

$$\mathcal{F}^{-1}(\delta_a) = \frac{1}{2\pi}e^{iat}, \qquad \mathcal{F}^{-1}(e^{-iat}) = \delta_a.$$

Preuve. Laissée en exercice. □

Théorème 10.55. *Soit U une distribution T-périodique. U est une distribution tempérée et sa tranformée de Fourier est donnée par*

$$\mathcal{F}U = 2\pi \sum_{k\in\mathbb{Z}} \gamma_k \delta_{k\omega}$$

où les γ_k sont les coefficients de Fourier de U et où $\omega = 2\pi/T$.

Preuve. On sait que la suite $(\gamma_k)_{k\in\mathbb{Z}}$ des coefficients de Fourier de U ($U = \sum_{k\in\mathbb{Z}} \gamma_k e^{2i\pi kt/T}$) est à croissance lente. On en déduit que

$$u = 2\pi \sum_{k\in\mathbb{Z}} \gamma_k \delta_{k\omega}$$

est dans $\mathcal{S}'$. Soit en effet $p \in \mathbb{N}$ et $C \geq 0$ tels que

$$\forall k \in \mathbb{Z}, \quad |\gamma_k| \leq C\,(1 + |k|)^p.$$

Soit $\phi \in \mathcal{S}$. On a pour tout $k \in \mathbb{Z}^*$

$$\begin{array}{rcl} |\langle \gamma_k \delta_{k\omega}, \phi \rangle| & = & |\gamma_k| \, |\phi(k\omega)| \\ & \leq & C\,(1+|k|)^p \, |\phi(k\omega)| \\ & \leq & C'(1+|k\omega|)^p \, |\phi(k\omega)| \\ & \leq & \dfrac{C''}{k^2} \sup\limits_{0 \leq \alpha \leq p+2} \|x^\alpha \phi(x)\|_{L^\infty} \\ & \leq & \dfrac{C''}{k^2} \mathcal{N}_{p+2}(\phi). \end{array}$$

On a donc

$$|\langle \sum_{k\in\mathbb{Z}} \gamma_k \delta_{k\omega}, \phi \rangle| \leq |\gamma_0| \, \mathcal{N}_0(\phi) + C''' \mathcal{N}_{p+2}(\phi) \leq C \mathcal{N}_{p+2}(\phi).$$

Soit maintenant

$$u_n = 2\pi \sum_{k=-n}^{n} \gamma_k \delta_{k\omega}.$$

On a

$$\mathcal{F}^{-1}(u_n) = 2\pi \sum_{k=-n}^{n} \gamma_k \mathcal{F}^{-1}(\delta_{k\omega}) = \sum_{k=-n}^{n} \gamma_k e^{ik\omega t} = \sum_{k=-n}^{n} \gamma_k e^{2i\pi kt/T}.$$

Comme u_n tend vers u dans $\mathcal{S}'$, il en résulte que $\bar{\mathcal{F}}(u_n)$ tend vers $\bar{\mathcal{F}}(u)$ dans $\mathcal{S}'$, donc dans $\mathcal{D}'$. Mais on a par ailleurs

$$\mathcal{F}^{-1}(u_n) \longrightarrow U \quad \text{dans } \mathcal{D}'(\mathbb{R}^d).$$

On en déduit que $U \in \mathcal{S}'(\mathbb{R}^d)$ et que

$$\mathcal{F}^{-1}(u) = U.$$

En appliquant $\mathcal{F}$ aux deux membres, on aboutit finalement à

$$\mathcal{F}(U) = u.$$

□

Remarque 10.56. Le théorème ci-dessus met en évidence le lien entre la théorie des séries de Fourier et celle de la transformée de Fourier : les coefficients de la série de Fourier d'une distribution périodique sont les poids des dents du peigne de Dirac correspondant à la transformée de Fourier de cette distribution.

10.5.4 Application : théorème d'échantillonage de Shannon

On considère un signal à une dimension (le temps), c'est-à-dire une fonction

$$s \,:\, \mathbb{R} \longrightarrow \mathbb{R}$$

qu'on suppose ici de classe H^1. On aura donc en particulier
- $s \in C^0(\mathbb{R})$
- $s \longrightarrow 0$ en $-\infty$ et $+\infty$.

On a

$$s(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{s}(\omega) \, e^{i\omega t} \, d\omega.$$

Mais s étant réel, $\hat{s}$ vérifie

$$\hat{s}(-\omega) = \int_{-\infty}^{+\infty} s(t)\, e^{i\omega t}\, dt = \overline{\int_{-\infty}^{+\infty} s(t)\, e^{-i\omega t}\, dt} = \overline{\hat{s}(\omega)}.$$

En écrivant $\hat{s}(\omega)$ sous la forme

$$\hat{s}(\omega) = \pi A(\omega) e^{i\phi(\omega)}$$

il en résulte

$$\begin{aligned} s(t) &= \frac{1}{2\pi}\int_0^{+\infty} \left(\hat{s}(\omega)\, e^{i\omega t} + \hat{s}(-\omega)\, e^{-i\omega t}\right)\, d\omega \\ &= \frac{1}{2\pi}\int_0^{+\infty} \left(\hat{s}(\omega)\, e^{i\omega t} + \overline{\hat{s}(\omega)}\, e^{-i\omega t}\right)\, d\omega \\ &= \frac{1}{2}\int_0^{+\infty} A(\omega)\, \left(e^{i(\omega t+\phi(\omega))} + e^{-i(\omega t+\phi(\omega))}\right)\, d\omega \\ &= \int_0^{+\infty} A(\omega)\, \cos(\omega t + \phi(\omega))\, d\omega. \end{aligned}$$

On a ainsi décomposé le signal s en une superposition de signaux harmoniques d'amplitude $A(\omega)$, de pulsation ω présentant un déphasage relatif $\phi(\omega)$.

La bonne représentation mathématique de l'échantillonage du signal s à la pulsation $\omega_\tau = \frac{2\pi}{\tau}$ (i.e. à la fréquence $\nu_\tau = \frac{1}{\tau}$) est le peigne de Dirac

$$s_\tau = \sum_{k\in\mathbb{Z}} \tau\, s(k\tau)\, \delta_{k\tau}.$$

On a ainsi la

Proposition 10.57. *$s_\tau \longrightarrow s$ dans $\mathcal{S}'$ quand $\tau \longrightarrow 0$.*

Preuve. Soit $\phi \in \mathcal{S}(\mathbb{R})$.

$$\langle s_\tau, \phi\rangle_{\mathcal{S}',\mathcal{S}} = \sum_{k\in\mathbb{Z}} \tau\, s(k\tau)\phi(k\tau) \longrightarrow \int_{-\infty}^{+\infty} s(t)\phi(t)\, dt = \langle s, \phi\rangle_{\mathcal{S}',\mathcal{S}}.$$

□

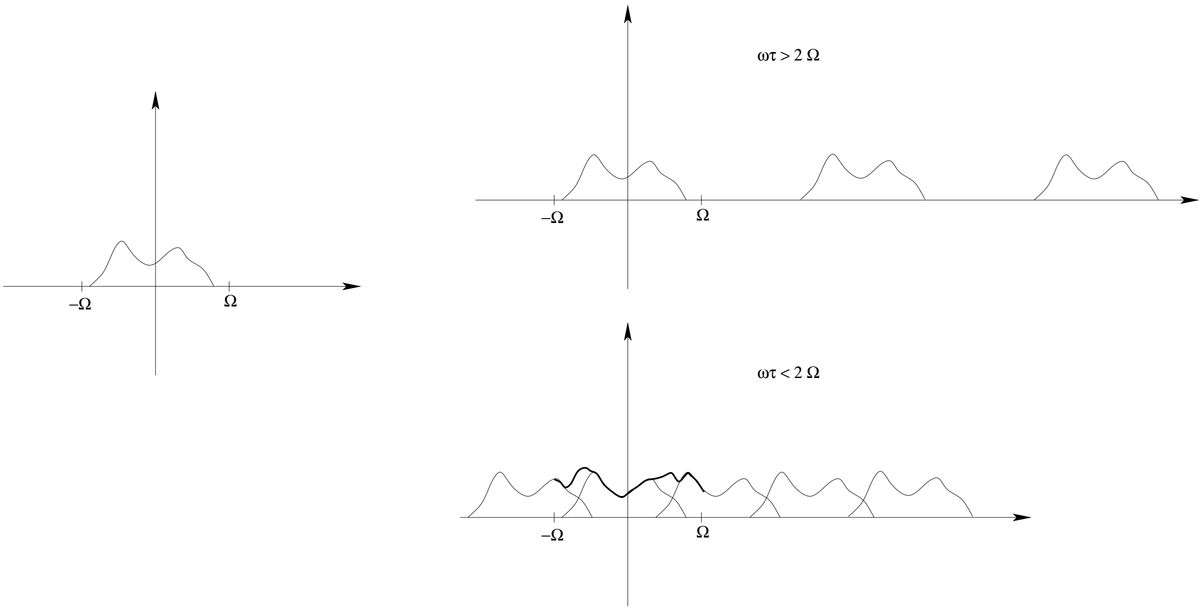
Définition 10.58. *On dit qu'un signal $s(t) \in \mathcal{S}'(\mathbb{R})$ est à bande limitée si sa transformée de Fourier est à support compact.*

Théorème 10.59. *Soit $s(t)$ à bande limitée. La transformée de Fourier du signal échantilloné*

$$s_\tau = \sum_{k\in\mathbb{Z}} \tau\, s(k\tau)\, \delta_{k\tau}$$

est reliée à la transformée de Fourier du signal d'origine par la relation

$$\hat{s}_\tau = \sum_{k\in\mathbb{Z}} \tau_{k\omega_\tau}(\hat{s}) \qquad \text{où} \quad \omega_\tau = \frac{2\pi}{\tau}.$$



Représentation graphique.

On voit que si $\omega_\tau < 2\omega$, il y a “repliement” de spectre (phénomène d'*aliasing*).

Exemple. On considère le signal à valeurs complexes $s(t) = e^{i\omega_0 t}$. On a

$$\hat{s} = 2\pi\delta_{\omega_0},$$

d'où l'on déduit

$$\hat{s}_\tau = 2\pi \sum_{k\in\mathbb{Z}} \delta_{\omega_0 + k\omega_\tau}.$$

Théorème 10.60. *(de Shannon). Soit $s(t)$ un signal de classe H^1 à bande limitée $[-\Omega,\Omega]$. Soit s_τ l'échantillonage de s à la pulsation ω_τ. Si $\omega_\tau \geq 2\Omega$, $\hat{s}_\tau = \hat{s}$ dans la bande $[-\Omega,\Omega]$. De plus, on a la formule de reconstruction du signal*

$$s(t) = \sum_{k\in\mathbb{Z}} s(k\tau)\, h_\tau(t-k\tau)$$

où la fonction h_τ est définie par

$$\forall t \in \mathbb{R}, \quad h_\tau(t)(t) = \frac{\sin(\pi t/\tau)}{\pi t/\tau}.$$

Preuve. La première assertion est une conséquence directe du théorème 10.59. Elle entraîne en particulier que

$$\hat{s} = \hat{s}_\tau\, \chi_{]-\pi/\tau,\pi/\tau[}$$

où $\chi_{]-\pi/\tau,\pi/\tau[}$ désigne la fonction caractéristique de l'intervalle $]-\pi/\tau,\pi/\tau[$. On vérifie par un calcul direct que $\mathcal{F}^{-1}\left(\chi_{]-\pi/\tau,\pi/\tau[}\right) = h_\tau$, et on en déduit que

$$\hat{s} = \hat{s}_\tau\, \hat{h}_\tau,$$

ce qui s'écrit aussi (on vérifiera que le membre de droite a bien un sens)

$$s = s_\tau \star h_\tau.$$

Ceci prouve la deuxième assertion. □

Application à la numérisation du son. Les fréquences entendues par les humains s'étallent entre 16 Hz et 20 000 Hz. Echantilloner à 40 000 Hz est donc suffisant en théorie. En pratique, on échantillonne à 48 000 Hz (sur CD) pour plusieurs raisons, notamment parce que
- il faut filtrer le signal avant d'échantilloner
- un signal numérique suréchantilloné est plus facile à convertir en signal analogique qu'un signal échantilloné à la fréquence minimale.

La preuve du théorème 10.59 repose sur le résultat suivant, qui est intéressant en lui-même.

Lemme 10.61. *(formule de Poisson). Soit $f \in \mathcal{S}(\mathbb{R})$. On a l'égalité*

$$\sum_{k\in\mathbb{Z}} f(k) = \sum_{k\in\mathbb{Z}} \hat{f}(2k\pi)$$

Preuve. Comme $f \in \mathcal{S}$, la série de fonctions

$$\sum_{k\in\mathbb{Z}} \tau_k f$$

est normalement convergente sur tout compact et converge vers une fonction U continue et 1-périodique. Calculons les coefficients de Fourier de U :

$$\begin{aligned}
\gamma_k &= \int_0^1 U(t)\, e^{-2ik\pi t}\, dt \\
&= \int_0^1 \sum_{l\in\mathbb{Z}} \tau_l f(t)\, e^{-2ik\pi t}\, dt \\
&= \sum_{l\in\mathbb{Z}} \int_0^1 f(t-l)\, e^{-2ik\pi t}\, dt \\
&= \sum_{l\in\mathbb{Z}} \int_{-l}^{-l+1} f(t)\, e^{-2ik\pi t}\, dt \\
&= \int_{-\infty}^{+\infty} f(t)\, e^{-2ik\pi t}\, dt \\
&= \hat{f}(2k\pi).
\end{aligned}$$

On a donc

$$\sum_{k\in\mathbb{Z}} f(k) = \left(\sum_{k\in\mathbb{Z}} \tau_k f\right)(0) = U(0) = \sum_{k\in\mathbb{Z}} \gamma_k = \sum_{k\in\mathbb{Z}} \hat{f}(2k\pi).$$

□

Preuve du théorème 10.59. Supposons d'abord que $s \in \mathcal{S}(\mathbb{R})$ (i.e. que $\hat{s} \in \mathcal{S}(\mathbb{R})$). On a

$$\begin{aligned}
\hat{s}_\tau(y) &= \mathcal{F}\left(\sum_{k\in\mathbb{Z}} \tau s(k\tau)\delta_{k\tau}\right)(y) \\
&= \sum_{k\in\mathbb{Z}} \tau s(k\tau)e^{-ik\tau y} \\
&= \sum_{k\in\mathbb{Z}} f_y(k)
\end{aligned}$$

où l'on a posé

$$f_y(x) = \tau s(x\tau)e^{-ix\tau y}.$$

En utilisant la formule de Poisson, on obtient

$$\hat{s}_\tau(y) = \sum_{k\in\mathbb{Z}} \hat{f}_y(2k\pi).$$

Mais

$$\begin{aligned}
\hat{f}_y(2k\pi) &= \int_{-\infty}^{+\infty} f_y(x)e^{-2ik\pi x}\,dx \\
&= \int_{-\infty}^{+\infty} \tau s(x\tau)e^{-ix\tau y}e^{-2ik\pi x}\,dx \\
&= \int_{-\infty}^{+\infty} s(t)e^{-ity}e^{-2ik\pi t/\tau}\,dt \\
&= \int_{-\infty}^{+\infty} s(t)e^{-it(y+2k\pi/\tau)}\,dt \\
&= \hat{s}(y+2k\pi/\tau) = \hat{s}(y+k\omega_\tau) = \tau_{k\omega_\tau}\hat{s}(y).
\end{aligned}$$

D'où

$$\hat{s}_\tau(y) = \sum_{k\in\mathbb{Z}} \tau_{k\omega_\tau}\hat{s}(y).$$

Par densité, la formule ci-dessus reste vraie pour tout signal à bande limitée (la série est finie). $\square$

10.5.5 Transformée de Fourier discrète

Considérons un signal échantillonné de longueur N représenté par une suite $(x_k)_{k\in|[0,N-1]|}$. Sa tranformée de Fourier discrète est par définition la suite $(X_k)_{k\in|[0,N-1]|}$ définie par

$$X_k = \sum_{l=0}^{N-1} x_l e^{-2i\pi kl/N}.$$

La transformée de Fourier discrète est reliée à la transformée de Fourier dans $\mathcal{S}'(\mathbb{R})$ de la façon suivante : on considère le signal échantilloné

$$s_\tau = \sum_{k=0}^{N-1} \tau\, x_k\, \delta_{k\tau},$$

et on le T-périodise avec $T = N\tau$. On obtient ainsi un peigne de Dirac périodique U, dont la transformée de Fourier est aussi un peigne de Dirac d'espacement $2\pi/T$, et périodique de période $2\pi/\tau$. Il en résulte que

$$\hat{U} = 2\pi \sum_{k\in\mathbb{Z}} \gamma_k \delta_{2\pi k/T}.$$

où les coefficients γ_k sont tels que

$$\forall k \in \mathbb{Z}, \quad \gamma_{k+N} = \gamma_k,$$

et on a pour tout $0 \leq k \leq N-1$

$$\begin{aligned}
\gamma_k &= \frac{1}{T}\langle s_\tau,\, e^{-2i\pi kt/T}\rangle_{\mathcal{E}',C^\infty} \\
&= \frac{1}{T}\sum_{l=0}^{N-1} \tau\, x_l\, e^{-2i\pi kl\tau/T} \\
&= \frac{1}{N}\sum_{l=0}^{N-1} x_l\, e^{-2i\pi kl/N} \\
&= \frac{X_k}{N}.
\end{aligned}$$

On obtient ainsi la formule d'inversion

$$x_k = \frac{1}{N}\sum_{l=0}^{N-1} X_l e^{2i\pi kl/N}.$$

En traitement du signal numérique, la transformée de Fourier discrète est l'opération de base qui permet de passer de la représentation en temps à la représentation en fréquences.

A priori, $4N(N-1)$ multiplications complexes et $4N(N-1)$ additions complexes sont nécessaires pour calculer la transformée de Fourier discrète d'une suite $(x_k)_{k\in|[0,N-1]|}$. Si N est une puissance de 2 (ex. $N = 2^{10} = 1024$), le calcul de la transformée de Fourier discrète peut en fait être effectué en $2N \log_2 N$ multiplications réelles et $3N \log_2 N$ additions réelles par un algorithme FFT (*Fast Fourier Transform*).

10.6 Exercices

$\triangleright$ **10.1.** Montrer que si f et g sont dans $L^2(\mathbb{R}^d)$,

$$(2\pi)^d \int_{\mathbb{R}^d} f(x)\,\overline{g(x)}\,dx = \int_{\mathbb{R}^d} \hat{f}(x)\,\overline{\hat{g}(x)}\,dx.$$

$\triangleright$ **10.2.** Montrer que la suite des sommes partielles

$$\sum_{n=0}^{N} \frac{x^n}{n!}$$

est une suite d'éléments de $\mathcal{S}'(\mathbb{R})$ qui converge dans $\mathcal{D}'(\mathbb{R})$ mais pas dans $\mathcal{S}'(\mathbb{R})$.

$\triangleright$ **10.3.** Soit $(a_n)_{n\in\mathbb{N}}$ une suite réelle. Montrer que la suite $a_n\delta_n$ converge dans $\mathcal{D}'(\mathbb{R})$. A quelle condition converge-t-elle dans $\mathcal{S}'(\mathbb{R})$?

$\triangleright$ **10.4.** Soit $T > 0$ et $f : \mathbb{R} \to \mathbb{C}$ une fonction continue telle que $xf(x)$ soit dans $L^1(\mathbb{R})$ et telle que le support de $\hat{f}$ soit inclus dans $[-\pi/T, \pi/T]$.

1. Montrer que $\hat{f}$ est de classe C^1, et en déduire que $\hat{f}$ et f sont dans $L^2(\mathbb{R})$. En utilisant le théorème de Shannon, montrer que
$$f(t) = \sum_{k\in\mathbb{Z}} f(kT)h_T(t-kT)$$
avec $h_T(t) = \frac{\sin(\pi t/T)}{\pi t/T}$.
2. On suppose que le signal f ne peut être mesuré qu'à ϵ près de sorte que la fonction reconstruite est de la forme
$$f_1(t) = \sum_{k\in\mathbb{Z}} (f(kT) + \epsilon_k)\, h_T(t-kT)$$
avec $|\epsilon_k| \leq \epsilon$. Montrer que sous cette seule hypothèse, on ne peut pas borner l'erreur de reconstruction $\|f_1 - f\|_{L^\infty}$.
3. Soit
$$f_2(t) = \sum_{|k|\leq T} (f(kT) + \epsilon_k)\ h_T(t-kT).$$
On suppose que $f(t)$ est telle qu'il existe K et C_K satisfaisant
$$\left|\sum_{|t-kT|>K} f(kT)h_T(t-kT)\right| \leq C_K\epsilon.$$
En déduire une majoration de l'erreur de reconstruction $\|f_2 - f\|_{L^\infty}$. *Indication : commencer par montrer que*
$$\sum_{T\leq|t-kT|<K,|t-kT|} \frac{1}{|t-kT|} \leq C\log(K/T).$$
4. Pour diminuer l'erreur de reconstruction, on peut suréchantilloner le signal. Soit $\hat{g}$ une fonction de classe C^∞ égale à 1 sur $[-\pi/T, \pi/T]$ et à support dans $[-\pi/T_s, \pi/T_s]$ avec $T_s \leq T$. Montrer que $g = \mathcal{F}^{-1}(\hat{g})$ est bien définie et est dans $\mathcal{S}$.
5. En adaptant la démonstration du théorème de Shannon, montrer que
$$f(t) = \sum_{k\in\mathbb{Z}} f(kT_s)g(t-kT_s).$$
6. Expliquer le processus par lequel le filtre g permet de diminuer l'erreur de reconstruction.
7. On pose
$$f_3(t) = \sum_{k\in\mathbb{Z}} (f(kT_s) + \epsilon_k)g(t-kT_s).$$
Montrer que
$$\|f - f_3\|_\infty \leq \frac{\epsilon}{T_s}(\|g\|_1 + T_s\|g'\|_1).$$
Pour cela, on utilisera en la justifiant l'égalité
$$\sum_k g(t-kT_s) = \sum_k \int_{t-(k+1/2)T_s}^{t-(k-1/2)T_s} (g(t-kT_s) - g(u))du + \sum_k \int_{t-(k+1/2)T_s}^{t-(k-1/2)T_s} g(u)du$$
et la dérivabilité de g.

▷ **10.5.** On pose $f(x) = \ln|x|$.

1. Montrer que la fonction f définit une distribution sur $\mathbb{R}$.

2. Montrer qu'au sens des distributions, on a pour toute fonction test $\varphi \in \mathcal{D}(\mathbb{R})$,

$$\langle f', \varphi \rangle \;=\; \lim_{\delta \to 0^+} \int_{\mathbb{R}\setminus[-\delta,\delta]} \left(-\ln|x|\varphi'(x)\right)\, dx$$

On introduit pour $\delta \geq 0$ la notation

$$1_{[-\delta,\delta]}(x) = \begin{cases} 1 & \text{si} \quad |x| \leq \delta \\ 0 & \text{si} \quad |x| > \delta \end{cases}$$

3. En déduire que pour tout $\varepsilon > 0$, on a

$$\langle f', \varphi \rangle = \int_{\mathbb{R}} \frac{\varphi(x) - \varphi(0)\ 1_{[-\varepsilon,\varepsilon]}(x)}{x}\, dx$$

On note $\text{vp}\left(\frac{1}{x}\right) = f'$.

4. Soit $\psi \in \mathcal{D}(\mathbb{R})$ fixée vérifiant $\psi(-x) = \psi(x)$ et $\psi(x) = 1$ si $x \in [-1, 1]$. On admet que f' et $\psi f'$ sont dans $\mathcal{S}'(\mathbb{R})$. Montrer que pour $\phi \in \mathcal{S}(\mathbb{R})$, et $\varepsilon \in]0, 1]$, on a

$$\langle \mathcal{F}(\psi f'), \phi \rangle \;=\; \int_{\mathbb{R}} \left(\frac{\hat{\phi}(x) - \hat{\phi}(0)\ 1_{[-\varepsilon,\varepsilon]}(x)}{x} \right) \psi(x)\, dx$$

où $\mathcal{F}(g)$ et $\hat{g}$ désignent indifféremment la transformée de Fourier de la distribution tempérée g.

5. En déduire que

$$\langle \mathcal{F}(\psi f'), \phi \rangle = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \phi(\xi)\psi(x) M_1(x,\xi)\, d\xi \right)\, dx$$

avec pour une fonction $l(x)$:

$$M_l(x,\xi) = \left(\frac{e^{-i\xi\cdot x} - 1_{[-\varepsilon,\varepsilon]}(x)}{x} \right) l(x)$$

6. En déduire que

$$\langle \mathcal{F}(\psi f'), \phi \rangle = \int_{\mathbb{R}} \phi(\xi)\, m_\psi(\xi)\, d\xi$$

où

$$m_\psi(\xi) = \int_{\mathbb{R}} M_\psi(x,\xi)\, dx$$

7. On admet qu'en prenant $\psi = 1$ on a encore

$$\langle \mathcal{F}(f'), \phi \rangle = \int_{\mathbb{R}} \phi(\xi)\, m_1(\xi)\, d\xi$$

avec

$$m_1(\xi) = \lim_{R \to +\infty} \int_{[-R,R]} M_1(x,\xi)\, dx$$

Montrer qu'on a

$$\mathcal{F}\left(\text{vp}\left(\frac{1}{x}\right)\right)(\xi) = -ic\ \text{sgn}(\xi)$$

où

$$\operatorname{sgn}(\xi) = \begin{cases} 1 & \text{si} \quad \xi > 0 \\ \\ -1 & \text{si} \quad \xi < 0 \end{cases}$$

et

$$c = \int_{\mathbb{R}} \frac{\sin y}{y}\, dy.$$

▷ **10.6.** L'objet de ce problème est l'étude de l'équation de Poisson

$$-\Delta u = f \tag{10.19}$$

posée dans $\mathcal{D}'(\mathbb{R}^3)$. Dans ce qui suit, on note $|y|$ la norme euclidienne du vecteur $y \in \mathbb{R}^3$.

Les parties A et B sont indépendantes (à l'exception de la question 5d).

Partie A : résolution par transformation de Fourier

Question 1. Soit $v \in \mathcal{S}'(\mathbb{R}^3)$. Montrer que $\Delta v \in \mathcal{S}'(\mathbb{R}^3)$ et que

$$\mathcal{F}(\Delta v)(\xi) = -|\xi|^2\, \hat{v}(\xi).$$

Question 2. On suppose dans cette question que $f \in \mathcal{S}'(\mathbb{R}^3)$ et que $\hat{f} \in L^\infty(\mathbb{R}^3)$.

2a On note χ_η la fonction définie sur $\mathbb{R}^3$ par

$$\chi_\eta(\xi) = \left| \begin{array}{ll} 1 & \text{si } |\xi| \leq \eta \\ 0 & \text{si } |\xi| > \eta. \end{array} \right.$$

Montrer que la fonction $g_\eta(\xi) = |\xi|^{-2}\chi_\eta(\xi)\hat{f}(\xi)$ est dans $L^1(\mathbb{R}^3)$ et que la fonction $h_\eta(\xi) = |\xi|^{-2}(1-\chi_\eta(\xi))\hat{f}(\xi)$ est dans $L^2(\mathbb{R}^3)$.

2b En déduire que la fonction $v(\xi) = |\xi|^{-2}\hat{f}(\xi)$ définit une distribution tempérée et que $u = \mathcal{F}^{-1}(v)$ est solution de l'équation (10.19) dans $\mathcal{S}'(\mathbb{R}^3)$.

2c Montrer que

$$\lim_{\eta \to 0} \|g_\eta\|_{L^1} = 0.$$

2d Dans cette question, ainsi que dans la question 2e, u désigne la solution de l'équation (10.19) construite à la question 2b. En utilisant les résultats des questions 2a, 2b et 2c, montrer que pour tout $\epsilon > 0$, il existe une fonction $u_2 \in L^2(\mathbb{R}^3)$ et une fonction $u_\infty \in L^\infty(\mathbb{R}^3)$ telles que

$$u = u_2 + u_\infty \qquad \text{avec} \qquad \|u_\infty\|_{L^\infty} \leq \epsilon.$$

2e Montrer que tout $v \in \mathcal{S}'(\mathbb{R}^3)$ vérifiant $-\Delta v = f$ est tel que $v = u + w$ avec $\Delta w = 0$ et $\operatorname{Supp}(\hat{w}) \subset \{0\}$. Déduire de la Proposition 7.13 du cours que w est un polynôme vérifiant $\Delta w = 0$ (un tel polynôme est appelé un polynôme harmonique). Donner un exemple de polynôme harmonique de degré 2.

2f Les distributions f suivantes vérifient-elles les hypothèses en vigueur dans la question 2 ($f \in \mathcal{S}'(\mathbb{R}^3)$ et $\hat{f} \in L^\infty(\mathbb{R}^3)$) ?

- $f = \delta_0$
- $f(x) = e^{-|x|^2}$
- $f(x) = 1$
- $f(x) = \frac{1}{1+|x|^4}$

Partie B : résolution dans l'espace de Sobolev à poids

$$W^1(\mathbb{R}^3) = \left\{ u \in L^2_{\text{loc}}(\mathbb{R}^3), \quad \frac{u(x)}{|x|} \in L^2(\mathbb{R}^3), \quad \nabla u \in (L^2(\mathbb{R}^3))^3 \right\}.$$

On rappelle que $L^2_{\text{loc}}(\mathbb{R}^3)$ désigne l'espace vectoriel des fonctions $u : \mathbb{R}^3 \longrightarrow \mathbb{R}$ mesurables pour la tribu borélienne et telles que $\int_K |u|^2 < \infty$ pour tout compact $K \subset \mathbb{R}^3$. Comme $L^2_{\text{loc}}(\mathbb{R}^3) \subset L^1_{\text{loc}}(\mathbb{R}^3)$, ∇u est bien défini, au sens des distributions.

Pour $(u, v) \in W^1(\mathbb{R}^3) \times W^1(\mathbb{R}^3)$, on pose

$$(u, v)_{W^1} = \int_{\mathbb{R}^3} \frac{u(x)\, v(x)}{|x|^2}\, dx + \int_{\mathbb{R}^3} \nabla u \cdot \nabla v. \tag{10.20}$$

Question 3. Montrer que pour $(u, v) \in W^1(\mathbb{R}^3) \times W^1(\mathbb{R}^3)$, le terme de droite de l'expression (10.20) est bien défini, puis que $(\cdot, \cdot)_{W^1}$ définit un produit scalaire sur $W^1(\mathbb{R}^3)$.

Question 4. Montrer que, muni du produit scalaire $(\cdot, \cdot)_{W^1}$, $W^1(\mathbb{R}^3)$ est un espace de Hilbert. *Indication : on pourra s'inspirer que la preuve du Théorème 8.2 du cours.*

Question 5. On suppose dans cette question que f est une fonction de $\mathbb{R}^3$ à valeurs dans $\mathbb{R}$ (mesurable pour la tribu borélienne) telle que

$$\int_{\mathbb{R}^3} |x|^2 |f(x)|^2\, dx < +\infty.$$

On admet que $\mathcal{D}(\mathbb{R}^3)$ est dense dans $W^1(\mathbb{R}^3)$ (lorsqu'on munit $W^1(\mathbb{R}^3)$ de la norme $\|\cdot\|_{W^1}$ associée au produit scalaire $(\cdot, \cdot)_{W^1}$). On considère les problèmes consistant à

$$\left\{ \begin{array}{l} \text{chercher } u \in W^1(\mathbb{R}^3) \text{ tel que} \\ -\Delta u = f \quad \text{dans } \mathcal{D}'(\mathbb{R}^3) \end{array} \right. \tag{10.21}$$

et à

$$\left\{ \begin{array}{l} \text{chercher } u \in W^1(\mathbb{R}^3) \text{ tel que} \\ \forall v \in W^1(\mathbb{R}^3), \quad a(u, v) = b(v) \end{array} \right. \tag{10.22}$$

où

$$a(u, v) = \int_{\mathbb{R}^3} \nabla u \cdot \nabla v, \qquad b(v) = \int_{\mathbb{R}^3} f v.$$

5a Montrer que a définit une forme bilinéaire continue sur $W^1(\mathbb{R}^3)$ (pour la norme $\|\cdot\|_{W^1}$) et que b définit une forme linéaire continue sur $W^1(\mathbb{R}^3)$ (pour la norme $\|\cdot\|_{W^1}$).

5b Montrer que les problèmes (10.21) et (10.22) sont équivalents.

5c On peut montrer (voir la question 6) que pour tout $u \in W^1(\mathbb{R}^3)$,

$$\int_{\mathbb{R}^3} \frac{|u(x)|^2}{|x|^2}\, dx \leq 4 \int_{\mathbb{R}^3} |\nabla u|^2. \tag{10.23}$$

En déduire que la forme bilinéaire a est coercive sur $W^1(\mathbb{R}^3)$, et de là que le problème (10.21) admet une unique solution u.

5d Montrer que la solution u de (10.21) est dans $\mathcal{S}'(\mathbb{R}^3)$. On suppose que $\hat{f} \in L^\infty(\mathbb{R}^3)$. Montrer que la solution de (10.21) coïncide avec la solution construite à la question 2b.

Question 6. Preuve de l'inégalité de Hardy (10.23).

6a Soit $f \in C^\infty([0, +\infty[)$ à support compact. Montrer que

$$\int_0^{+\infty} f(r)^2\, dr = -2 \int_0^{+\infty} r\, f'(r)\, f(r)\, dr$$

et en déduire que

$$\int_0^{+\infty} f(r)^2\, dr \leq 4 \int_0^{+\infty} r^2\, f'(r)^2\, dr.$$

6b En déduire que pour tout $\psi \in \mathcal{D}(\mathbb{R}^3)$,

$$\int_{\mathbb{R}^3} \frac{\psi(x)^2}{|x|^2}\, dx \leq 4 \int_{\mathbb{R}^3} |\nabla \psi|^2.$$

Indication : remarquer qu'en coordonnées sphériques

$$\int_{\mathbb{R}^3} \frac{\psi(x)^2}{|x|^2}\, dx = \int_0^{\pi} \int_0^{2\pi} \left(\int_0^{+\infty} \psi(r, \theta, \phi)^2\, dr \right) \sin\theta\, d\theta\, d\phi,$$

et utiliser la question 6a. On rappelle qu'en coordonnées sphériques

$$\nabla \psi = \frac{\partial \psi}{\partial r} e_r + \frac{1}{r} \frac{\partial \psi}{\partial \theta} e_\theta + \frac{1}{r\, \sin\theta} \frac{\partial \psi}{\partial \phi} e_\phi,$$

et que la base (e_r, e_θ, e_ϕ) *est orthonormée.*

6c Montrer par densité que pour tout $u \in W^1(\mathbb{R}^3)$,

$$\left\| \frac{u}{|x|} \right\|_{L^2} \leq 2\, \|\nabla u\|_{L^2}.$$

Bibliographie

[AF 87] J.-M. Arnaudiès et H. Fraysse, *Cours de mathématiques. I - Algèbre*, Bordas 1987.

[AF 88] J.-M. Arnaudiès et H. Fraysse, *Cours de mathématiques. II - Analyse*, Bordas 1988.

[Bony 01] J.-M. Bony, *Cours d'analyse. Théorie des distributions et analyse de Fourier*, Editions de l'Ecole polytechnique 2001.

[Revuz 94] D. Revuz, *Mesure et intégration*, Hermann 1994.

[Rudin 92] W. Rudin, *Analyse réelle et complexe*, Masson 1992.

[Brézis 83] H. Brézis, *Analyse fonctionnelle, théorie et applications*, Masson 1983.

Index