

Outils élémentaires d'analyse pour les Equations aux Dérivées Partielles

Anne-Sophie Bonnet-Ben Dhia, Laurent Bourgeois, Christophe Hazard

▶ **To cite this version:**

Anne-Sophie Bonnet-Ben Dhia, Laurent Bourgeois, Christophe Hazard. Outils élémentaires d'analyse pour les Equations aux Dérivées Partielles. École d'ingénieur. Cours MIA 102 - Année 2022-2023, France. 2021, pp.132. hal-04061707

HAL Id: hal-04061707
https://hal.science/hal-04061707

Submitted on 7 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Cours MA 102

Année 2022–2023

Outils élémentaires d'analyse pour les Equations aux Dérivées Partielles

Anne-Sophie BONNET-BEN DHIA, Laurent BOURGEOIS
et Christophe HAZARD

Introduction

Ce cours présente quelques uns des principaux outils de l'analyse pour l'étude mathématique des équations aux dérivées partielles issues des sciences physiques ou humaines.

Le chapitre 1 rappelle quelques notions essentielles concernant la *topologie des espaces vectoriels normés*.

Le chapitre 2 est une introduction succincte à l'*intégrale de Lebesgue*, l'idée étant ici de comprendre quels sont les avantages de cette intégrale par rapport à celle de Riemann, et non pas de détailler la théorie sous-jacente.

Le chapitre 3 expose les bases de la théorie des *distributions*, due à Laurent Schwartz, qui généralise la notion de fonction.

Le chapitre 4 présente les propriétés fondamentales de la *transformation de Fourier* pour les fonctions intégrables et les fonctions de carré intégrable.

Le chapitre 5 est consacré aux *espaces de Hilbert* qui, dans le cas de la dimension infinie, fournissent un cadre de travail analogue aux espaces euclidiens.

Le chapitre 6 donne des exemples d'espaces de Hilbert qui jouent un rôle important dans l'étude des équations aux dérivées partielles : les *espaces de Sobolev*.

Enfin, le chapitre 7 pose les bases de l'*analyse variationnelle* des problèmes elliptiques, qui ouvre notamment la porte à la méthode des *éléments finis*. Mais c'est là une autre histoire...

Table des matières

1 Topologie des espaces vectoriels normés **9**
- 1.1 Espaces vectoriels normés . . . 9
- 1.2 Espaces complets . . . 13
- 1.3 Quelques résultats utiles . . . 17
  - 1.3.1 Prolongement d'une application linéaire . . . 17
  - 1.3.2 Séries convergentes et normalement convergentes . . . 18

2 L'intégrale de Lebesgue **21**
- 2.1 De Riemann à Lebesgue . . . 22
  - 2.1.1 L'intégrale de Riemann . . . 22
  - 2.1.2 Définition de l'intégrale de Lebesgue . . . 23
  - 2.1.3 Espaces $\mathcal{L}^1$ et L^1 . . . 28
- 2.2 Critères d'intégrabilité . . . 30
  - 2.2.1 L'intégrale n'est pas une affaire de signe ! . . . 30
  - 2.2.2 Les fonctions intégrables au sens de Riemann . . . 31
  - 2.2.3 Le théorème de convergence monotone . . . 32
  - 2.2.4 Le théorème de convergence dominée . . . 32
- 2.3 Intégrales multiples . . . 34
  - 2.3.1 Le théorème de Fubini . . . 34
  - 2.3.2 Produit de convolution . . . 36
- 2.4 Quelques outils pratiques . . . 37
  - 2.4.1 Changement de variable . . . 37
  - 2.4.2 Formule de Stokes : un premier cas simple . . . 40
  - 2.4.3 Formule de Stokes : cas général . . . 43
  - 2.4.4 Formules de Green . . . 45

3 Les distributions **47**
- 3.1 Définition et premiers exemples . . . 48
  - 3.1.1 De la physique aux mathématiques . . . 48
  - 3.1.2 L'espace des distributions . . . 49
  - 3.1.3 Quelques exemples de distributions . . . 51
- 3.2 Dérivation des distributions . . . 53
  - 3.2.1 Définition et exemples . . . 53

3.2.2 La formule des sauts en dimension 1 . . . 55
3.2.3 Primitive d'une fonction de $L^1(\mathbb{R})$ . . . 57
3.3 Multiplication et division des distributions . . . 58
3.3.1 Multiplication des distributions . . . 58
3.3.2 Division des distributions . . . 60
3.4 Convergence des distributions . . . 62
3.A Annexe : produit de convolution . . . 63
3.A.1 Support d'une distribution . . . 64
3.A.2 Convolution de fonctions localement intégrables . . . 67
3.A.3 Convolution de distributions . . . 69

4 Transformation de Fourier **73**
4.1 Transformation de Fourier des fonctions L^1 . . . 74
4.1.1 Le théorème de Riemann-Lebesgue . . . 74
4.1.2 Propriétés de la transformation de Fourier dans L^1 . . . 76
4.1.3 Exemples . . . 78
4.1.4 Convolution et transformation de Fourier dans L^1 . . . 80
4.2 Le cas des fonctions de carré intégrable . . . 80
4.2.1 Définition et propriétés de L^2 . . . 81
4.2.2 Transformation de Fourier dans L^2 . . . 83

5 Espaces de Hilbert **89**
5.1 Propriétés élémentaires . . . 89
5.2 Le théorème de projection . . . 92
5.3 Bases hilbertiennes . . . 97

6 Espaces de Sobolev **101**
6.1 Les espaces $H^m(\Omega)$ . . . 101
6.2 Le cas particulier $H^1(\Omega)$ . . . 103
6.2.1 Notion de trace . . . 104
6.2.2 Formules de Green . . . 106
6.2.3 Caractérisation de $H_0^1(\Omega)$ . . . 107
6.2.4 Le cas $d = 1$ . . . 108

7 Etude variationnelle des problèmes elliptiques **111**
7.1 Quelques exemples de problèmes aux limites . . . 111
7.1.1 Le problème de Dirichlet . . . 111
7.1.2 Le problème de Neumann . . . 112
7.1.3 Autres exemples . . . 113
7.2 Formulations variationnelles . . . 114
7.2.1 Un cas simplifié dans l'espace libre . . . 114
7.2.2 Le problème de Dirichlet . . . 116
7.2.3 Le problème de Neumann . . . 118
7.2.4 Formulations variationnelles et énergie . . . 119

7.3 Le théorème de Lax-Milgram 120
7.4 Application aux problèmes aux limites elliptiques 123
7.4.1 Problème simplifié dans l'espace libre 123
7.4.2 Le problème de Dirichlet 123
7.4.3 Le problème de Neumann 126

Bibliographie **129**

Index **131**

Chapitre 1

Topologie des espaces vectoriels normés

Dans ce chapitre, nous rappelons quelques notions élémentaires relatives aux espaces vectoriels munis d'une *norme*, qui permet de définir une *topologie* sur ces espaces. Nous nous contentons ici d'évoquer deux notions fondamentales : la *convergence* (d'une suite) et la *continuité* (d'une application). Nous introduisons ensuite la notion d'espace *complet* qui joue un rôle essentiel en analyse.

1.1 Espaces vectoriels normés

Dans tout ce chapitre, E désigne un espace vectoriel que l'on supposera défini sur $\mathbb{R}$ pour simplifier. On pourrait considérer des espaces définis sur $\mathbb{C}$: tous les résultats qui sont donnés ici restent valables, la notation $|\cdot|$ désignant alors le module d'un nombre complexe au lieu de la valeur absolue d'un nombre réel. En pratique, les espaces qui vont nous intéresser dans ce cours sont des espaces de fonctions, de *dimension infinie*, dont la définition repose principalement sur les notions de continuité et de dérivabilité (exemples suivants), mais aussi sur celle d'intégrabilité (chapitres suivants).

Exemple 1.1 Soit D un domaine inclus dans $\mathbb{R}^d$ (avec $d \geq 1$). On désigne par $\mathcal{C}^0(D)$ l'espace vectoriel formé des fonctions continues de D dans $\mathbb{R}$. Rappelons qu'une fonction $u : \ D \to \mathbb{R}$ est continue si elle est continue en tout point $x \in D$, c'est-à-dire si

$$\forall x \in D,\ \forall \varepsilon > 0,\ \exists \eta > 0,\ \forall y \in D,\ |y - x| \leq \eta \Longrightarrow |u(y) - u(x)| \leq \varepsilon.$$

Remarquons que lorsque $D \neq \mathbb{R}^d$, cette définition contient des informations sur le comportement de u au voisinage du bord de D (soit $\overline{D} \backslash \overset{\circ}{D}$, c'est-à-dire l'ensemble des points qui appartiennent à l'adhérence de D mais pas à son intérieur). En effet, si D est fermé (et donc contient son bord), les fonctions de $\mathcal{C}^0(D)$ sont continues jusqu'au bord, alors que si D est ouvert, elles peuvent avoir un comportement singulier au voisinage du bord (par exemple tendre vers l'infini lorsqu'on s'en rapproche). •

Exemple 1.2 Plus généralement, si Ω est un ouvert de $\mathbb{R}^d$ et $m \geq 1$, un entier, on désigne par $\mathcal{C}^m(\Omega)$ l'espace des fonctions m fois continûment dérivables dans Ω, et $\mathcal{C}^\infty(\Omega)$, l'espace des fonctions indéfiniment dérivables. Que peut-on dire si l'on remplace Ω par un fermé, par exemple son adhérence $\overline{\Omega}$? A la différence du cas $m = 0$, définir des fonctions "dérivables jusqu'au bord" est délicat car la définition même de la dérivée en un point nécessite que la fonction soit définie dans un voisinage de ce point, ce qui n'est pas le cas si celui-ci est situé sur le bord. On peut contourner la difficulté en définissant $\mathcal{C}^m(\overline{\Omega})$ comme l'espace des restrictions à $\overline{\Omega}$ des fonctions de $\mathcal{C}^m(\mathbb{R}^d)$ (et de même pour $\mathcal{C}^\infty(\overline{\Omega})$). •

Définition 1.3 *Soit E un espace vectoriel.*

(i) *Une application $u \mapsto \|u\|$ de E dans $[0, +\infty[$ est une norme sur E si elle vérifie les trois propriétés suivantes (pour tous $u, v \in E$ et $\lambda \in \mathbb{R}$) :*

$$\|\lambda\, u\| = |\lambda|\, \|u\| \quad \textit{(homogénéité)}, \tag{1.1}$$

$$\|u + v\| \leq \|u\| + \|v\| \quad \textit{(inégalité triangulaire)}, \tag{1.2}$$

$$(\|u\| = 0) \Longrightarrow (u = 0) \quad \textit{(séparation)}. \tag{1.3}$$

(ii) *Si les propriétés* (1.1) *et* (1.2) *sont vérifiées, mais pas* (1.3)*, l'application est appelée une semi-norme.*

Un espace vectoriel normé est alors simplement un espace muni d'une norme particulière. Bien sûr, un même espace vectoriel peut être muni de plusieurs normes. Deux normes $\|\cdot\|_1$ et $\|\cdot\|_2$ sont dites *équivalentes* s'il existe deux constantes $\alpha > 0$ et $\beta > 0$ telles que pour tout $u \in E$, on a

$$\alpha\, \|u\|_1 \leq \|u\|_2 \leq \beta\, \|u\|_1.$$

Si E est de *dimension finie*, toutes les normes sont équivalentes. Mais ce n'est plus vrai en dimension infinie, comme le montre l'exemple suivant.

Exemple 1.4 Il est facile de vérifier que les deux applications suivantes constituent deux normes sur $\mathcal{C}^0([a, b])$, où $a, b \in \mathbb{R}$ avec $a < b$:

$$\|u\|_\infty = \sup_{x \in [a,b]} |u(x)| \quad \text{(norme de la convergence uniforme)}, \tag{1.4}$$

$$\|u\|_1 = \int_a^b |u(x)|\, \mathrm{d}x \quad \text{(norme de la convergence en moyenne)}. \tag{1.5}$$

Mais ces deux normes ne sont pas équivalentes. On a clairement $\|u\|_1 \leq (b - a)\, \|u\|_\infty$, mais il n'est en revanche pas possible de trouver une constante $\alpha > 0$ telle que $\|u\|_\infty \leq \alpha\, \|u\|_1$. Pour s'en convaincre, il suffit par exemple de considérer la suite de fonctions "chapeaux" $(u_n)_{n \geq 1} \subset \mathcal{C}^0([a, b])$ définie par

$$u_n(x) = \begin{cases} n - n^2 |x - c| & \text{si } |x - c| \leq 1/n, \\ 0 & \text{sinon}, \end{cases} \tag{1.6}$$

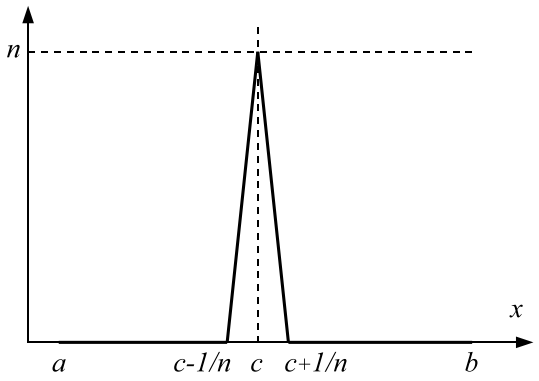


FIGURE 1.1 – Fonction "chapeau"

où $c \in]a,b[$ est fixé (voir Figure 1.1). On remarque alors que $\|u_n\|_1 = 1$ pour tout $n \geq 1$, tandis que $\|u_n\|_\infty = n$. •

Se donner une norme sur un espace vectoriel E, c'est définir une *topologie* sur cet espace, c'est-à-dire un cadre qui donne un sens précis à l'idée un peu vague de "proximité" entre différents éléments de E, principalement au travers de la notion de *convergence* (d'une suite d'éléments de E), et celle de *continuité* (d'une application agissant sur les éléments de E).

Définition 1.5 *Soit $(E, \|\cdot\|)$ un espace vectoriel normé. Une suite $(u_n)_{n\in\mathbb{N}}$ d'éléments de E converge vers $u \in E$ (ce qu'on notera $u_n \overset{E}{\to} u$, ou plus simplement $u_n \to u$) si $\lim_{n\to\infty} \|u_n - u\| = 0$, c'est-à-dire*

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ n \geq N \Rightarrow \|u_n - u\| \leq \varepsilon.$$

Définition 1.6 *Soit $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés. Une application $f : E \to F$ est dite continue en $u \in E$ si $\lim_{v\to u} \|f(v) - f(u)\|_F = 0$, c'est-à-dire*

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall v \in E,\ \|v - u\|_E \leq \eta \Rightarrow \|f(v) - f(u)\|_F \leq \varepsilon. \tag{1.7}$$

Et elle est dite continue sur E si elle est continue en tout $u \in E$. Enfin, elle est dite uniformément continue sur E si

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall u \in E,\ \forall v \in E,\ \|v - u\|_E \leq \eta \Rightarrow \|f(v) - f(u)\|_F \leq \varepsilon. \tag{1.8}$$

Notons qu'entre le critère de continuité sur E (soit (1.7) pour tout $u \in E$) et celui de continuité uniforme (1.8), la seule différence est la place de "$\forall u \in E$" : dans le second cas, η ne dépend pas de u, alors qu'il peut en dépendre dans le premier.

Les topologies associées à deux normes équivalentes sont les mêmes, ce qui signifie qu'on peut utiliser indifféremment l'une ou l'autre de ces normes dans les deux définitions ci-dessus sans en changer le sens. Par contre, il n'en va plus de même si les deux normes ne sont pas équivalentes. Pour l'illustrer, on peut considérer les deux normes définies sur $\mathcal{C}^0([a,b])$ dans l'exemple 1.4 et choisir la suite $(v_n)_{n\in\mathbb{N}}$ définie par $v_n(x) = u_n(x)/n$ où u_n est la fonction "chapeau" donnée par (1.6). On vérifie immédiatement que v_n converge en moyenne vers 0, mais pas uniformément.

Dans le cas des applications *linéaires*, c'est-à-dire telles que

$$f(\lambda\, u + \mu\, v) = \lambda\, f(u) + \mu\, f(v) \quad \forall u, v \in E,\ \forall \lambda, \mu \in \mathbb{R},$$

le critère de continuité en un point équivaut à la continuité sur tout l'espace, et prend une forme simplifiée donnée dans la proposition suivante.

Proposition 1.7 *Soit $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés. Une application linéaire $f : E \to F$ est continue sur E si et seulement si il existe une constante $C > 0$ telle que*

$$\|f(u)\|_F \leq C\ \|u\|_E \quad \forall u \in E. \tag{1.9}$$

Démonstration. Supposons tout d'abord que f est continue sur E. Elle est donc continue en 0, ce qui montre en particulier (en choisissant $\varepsilon = 1$ dans (1.7)) qu'il existe $\eta > 0$ tel que

$$\forall v \in E,\ \|v\|_E \leq \eta \Rightarrow \|f(v)\|_F \leq 1.$$

Pour tout $u \in E$ non nul, on peut poser $v = \eta\, u/\|u\|_E$ qui est tel que $\|v\|_E = \eta$. L'inégalité $\|f(v)\|_F \leq 1$ s'écrit encore $\|f(u)\|_F \leq \eta^{-1}\ \|u\|_E$ en utilisant la linéarité de f. Donc (1.9) est vérifiée pour $C = \eta^{-1}$.

Réciproquement, si (1.9) est vérifiée, alors

$$\|f(v) - f(u)\|_F = \|f(v-u)\|_F \leq C\ \|v-u\|_E \quad \forall u, v \in E,$$

ce qui signifie que f est uniformément continue (on a bien (1.8) en prenant $\eta = \varepsilon/C$). □

La propriété (1.9) caractérise la notion d'application linéaire *bornée*. La proposition 1.7 nous dit que cette notion est équivalente à celle de continuité (et même de continuité uniforme). L'ensemble des applications linéaires bornées entre deux espaces vectoriels normés $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$, noté $\mathcal{L}(E,F)$ ou plus simplement $\mathcal{L}(E)$ lorsque $E = F$, apparaît lui-même comme un espace normé si on le munit de la norme

$$\|f\| = \sup_{u\in E;\, u\neq 0} \frac{\|f(u)\|_F}{\|u\|_E} = \sup_{u\in E;\, \|u\|_E \leq 1} \|f(u)\|_F = \sup_{u\in E;\, \|u\|_E = 1} \|f(u)\|_F.$$

Remarquons que $\|f\|$ est la plus petite constante C qui peut être choisie dans (1.9).

1.2 Espaces complets

La notion de *complétude* d'un espace vectoriel normé $(E, \|\cdot\|)$ joue un rôle fondamental en analyse, qui peut s'expliquer de façon très simple. Comment montre-t-on en pratique qu'une suite $(u_n)_{n\in\mathbb{N}}$ converge ? De deux choses l'une. Ou bien nous connaissons (ou avons l'intuition de) la limite u, et nous devons alors estimer $\|u_n - u\|$. Ou bien nous ne connaissons pas cette limite. Dans ce cas, la notion de *suite de Cauchy* va nous être d'un grand secours. De façon imagée, on peut dire qu'une suite de Cauchy est une suite qui "arrête d'avancer", au sens où la distance qui sépare deux éléments u_n et u_m devient de plus en plus petite lorsque n et m deviennent grands. Et une suite convergente est une suite "qui va quelque part". Clairement, une suite qui va quelque part est bien forcée d'arrêter d'avancer. Mais qu'en est-il de la réciproque ? Une suite qui arrête d'avancer va-t-elle nécessairement quelque part ? Toute la difficulté vient du fait qu'elle peut s'approcher de plus en plus de quelque chose qui n'est pas dans l'espace, une sorte de "trou" dans l'espace. Dans certains cas, il peut y avoir beaucoup de trous, il peut y en avoir "partout", au sens où tout élément de l'espace peut être entouré de trous qui sont infiniment proches, il peut même y avoir beaucoup plus de trous que d'éléments dans l'espace. Un espace complet est un espace sans trou, où toute suite qui arrête d'avancer est garantie d'aller quelque part, bref un espace où l'on peut se déplacer en toute sécurité !

L'exemple le plus simple qui illustre cette notion est l'espace $\mathbb{Q}$ des nombre rationnels, muni de la valeur absolue comme norme. Cet espace n'est pas complet, et c'est précisément en cherchant un espace complet qui le contient qu'on construit l'espace $\mathbb{R}$ des nombres réels. Il faut cependant prendre garde de ne pas se laisser piéger par cet exemple qui peut laisser penser que l'existence de ces "trous" est liée intrinsèquement à l'espace E lui-même, indépendamment de la norme dont on le munit. C'est vrai en dimension finie, bien sûr, mais pas en dimension infinie, où un espace peut être complet pour une certaine norme, et infesté de trous pour une autre !

Définition 1.8 *Une suite $(u_n)_{n\in\mathbb{N}}$ d'un espace vectoriel normé $(E, \|\cdot\|)$ est une suite de Cauchy si*

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ (n \geq N \text{ et } m \geq N) \Rightarrow \|u_n - u_m\| \leq \varepsilon, \tag{1.10}$$

ce qu'on peut noter plus simplement

$$\lim_{n,m\to+\infty} \|u_n - u_m\| = 0.$$

Une suite convergente est bien sûr une suite de Cauchy, ce qui découle immédiatement de l'inégalité triangulaire :

$$\|u_n - u_m\| \leq \|u_n - u\| + \|u_m - u\|.$$

Mais la réciproque n'est pas vraie en général. Elle le devient dans le cadre de la définition suivante.

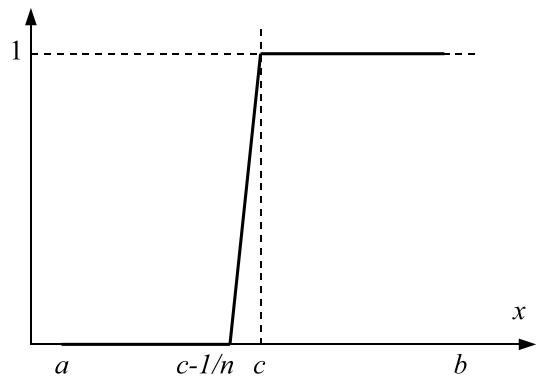


FIGURE 1.2 – La "marche d'escalier"

Définition 1.9 *Un espace vectoriel normé $(E, \|\cdot\|)$ est dit complet si toute suite de Cauchy converge. Un espace vectoriel normé complet est appelé espace de Banach.*

Exemple 1.10 Tout espace vectoriel (sur $\mathbb{R}$ ou $\mathbb{C}$) de dimension finie est complet. •

Proposition 1.11 *$\mathcal{C}^0([a,b])$ muni de la norme $\|\cdot\|_\infty$ définie en* (1.4) *est complet.*

DÉMONSTRATION. Soit (u_n) une suite de Cauchy de $(\mathcal{C}^0([a,b]), \|\cdot\|_\infty)$. On sait donc que

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall x \in [a,b],\ (n \geq N \text{ et } m \geq N) \Rightarrow |u_n(x) - u_m(x)| \leq \|u_n - u_m\|_\infty \leq \varepsilon.$$

Ainsi, pour tout $x \in [a,b]$ fixé, la suite $(u_n(x))$ est de Cauchy dans $\mathbb{R}$, et est par conséquent convergente. Notons $u(x)$ sa limite. En faisant tendre m vers l'infini dans la relation ci-dessus, on a alors

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall x \in [a,b],\ n \geq N \Rightarrow |u_n(x) - u(x)| \leq \varepsilon,$$

ce qui montre que u_n converge uniformément vers u. Et u est continue puisque

$$\begin{aligned} |u(y) - u(x)| &\leq |u(y) - u_n(y)| + |u_n(y) - u_n(x)| + |u_n(x) - u(x)| \\ &\leq 2\,\|u - u_n\|_\infty + |u_n(y) - u_n(x)|, \end{aligned}$$

où cette dernière quantité peut être rendue aussi petite que l'on veut en choisissant d'abord un n assez grand, puis $|y - x|$ assez petit. □

Proposition 1.12 *$\mathcal{C}^0([a,b])$ muni de la norme $\|\cdot\|_1$ définie en* (1.5) *n'est pas complet.*

DÉMONSTRATION. On peut par exemple considérer la suite (u_n) définie par

$$u_n(x) = \begin{cases} 0 & \text{si } x < c - 1/n, \\ n(x-c) + 1 & \text{si } c - 1/n \leq x \leq c, \\ 1 & \text{si } x > c, \end{cases}$$

où $c \in]a,b[$ est fixé (voir Figure 1.2). Cette suite est de Cauchy pour la norme $\|\cdot\|_1$ car

$$\int_a^b |u_n(x) - u_m(x)|\,\mathrm{d}x = \left|\frac{1}{2n} - \frac{1}{2m}\right|.$$

De plus, elle converge pour la norme $\|\cdot\|_1$ vers la fonction u définie par

$$u(x) = \begin{cases} 0 & \text{si } x \leq c, \\ 1 & \text{si } x > c. \end{cases}$$

En effet

$$\int_a^b |u_n(x) - u(x)|\,\mathrm{d}x = \frac{1}{2n}.$$

Or u n'appartient pas à $\mathcal{C}^0([a,b])$ puisqu'elle est discontinue en $x = c$. Peut-on en déduire que (u_n) ne converge pas dans $\mathcal{C}^0([a,b])$? Raisonnons par l'absurde en supposant qu'il existe une autre fonction $v \in \mathcal{C}^0([a,b])$ telle que $\lim_{n\to\infty} \|u_n - v\|_1 = 0$. En utilisant le fait que

$$\|u - v\|_1 \leq \|u_n - u\|_1 + \|u_n - v\|_1,$$

on en déduit, en passant à la limite $n \to \infty$, que $\|u - v\|_1 = 0$. Ceci entraîne que $u = v$ dans chacun des intervalles $[a,c[$ et $]c,b]$ (sur lesquels $u - v$ est continue), ce qui contredit le fait que $v \in \mathcal{C}^0([a,b])$. Ainsi (u_n) ne converge pas dans $\mathcal{C}^0([a,b])$. □

Au vu de cette dernière proposition, il semble naturel de se demander si on peut trouver un espace plus grand que $\mathcal{C}^0([a,b])$ qui soit complet pour la norme $\|\cdot\|_1$, autrement dit si on peut boucher les trous de $(\mathcal{C}^0([a,b]), \|\cdot\|_1)$. Cette question est une des motivations principales du chapitre 2 où nous introduisons l'intégrale de Lebesgue. De façon plus générale, peut-on boucher les trous d'un espace vectoriel normé incomplet ? Intuitivement, il s'agit de lui ajouter "toutes les limites possibles des suites de Cauchy", de sorte que l'espace incomplet soit *dense* dans l'espace complété, au sens de la définition suivante.

Définition 1.13 *Un sous-ensemble A d'un espace vectoriel normé $(E, \|\cdot\|)$ est dit dense si pour tout $u \in E$, il existe une suite (u_n) d'éléments de A qui converge vers u.*

On peut répondre affirmativement à la question précédente : il est toujours possible de compléter un espace vectoriel normé incomplet $(E, \|\cdot\|)$, plus précisément de construire un espace vectoriel normé complet $\mathcal{E}$ tel que E s'identifie à un sous-espace dense de $\mathcal{E}$. Nous rappelons ici brièvement le processus abstrait de construction de $\mathcal{E}$, qui est identique à celui qu'on utilise pour construire $\mathbb{R}$ à partir de $\mathbb{Q}$ à l'aide des suites de Cauchy. Notons qu'il s'agit là d'une curiosité qui dépasse légèrement le cadre de ce cours. Un lecteur un peu pressé pourra passer directement au paragraphe suivant.

Le point de départ consiste à définir une relation d'équivalence [1] sur l'espace des suites de Cauchy de E : on dit que deux suites de Cauchy $U = (u_n)$ et $V = (v_n)$ sont équivalentes

1. Rappelons qu'une *relation d'équivalence* $\sim$ sur un ensemble $\mathcal{U}$ est une relation binaire (autrement dit, qui caractérise des couples d'éléments de $\mathcal{U}$) qui est *réflexive* (soit $U \sim U$ pour tout $U \in \mathcal{U}$),

si $\lim_{n\to\infty} \|u_n - v_n\| = 0$. Appelons $\mathcal{E}$ l'espace des classes d'équivalence pour cette relation : la classe $\dot{U} \in \mathcal{E}$ associée à une suite de Cauchy $U = (u_n)$ est l'ensemble de toutes les suites de Cauchy $V = (v_n)$ telles que $\lim_{n\to\infty} \|u_n - v_n\| = 0$. Pour une classe $\dot{U}$, la quantité

$$\|\dot{U}\|_{\mathcal{E}} = \lim_{n\to\infty} \|u_n\| \tag{1.11}$$

existe (car $(\|u_n\|)$ est une suite de Cauchy de $\mathbb{R}$) et ne dépend pas du représentant $U = (u_n)$ choisi dans la classe (car $\lim_{n\to\infty} \|u_n\| = \lim_{n\to\infty} \|v_n\|$ dès que $\lim_{n\to\infty} \|u_n - v_n\| = 0$). On vérifie sans difficulté que $\|\cdot\|_{\mathcal{E}}$ définit une norme sur $\mathcal{E}$.

Remarquons par ailleurs qu'on peut identifier E à un sous-espace de $\mathcal{E}$ en considérant, pour tout $u \in E$, la suite constante (donc de Cauchy), notée $\tilde{u}$, dont tous les termes sont égaux à u. L'ensemble $\tilde{E}$ des classes d'équivalences correspondantes, soit $\tilde{E} = \{\dot{\tilde{u}} \in \mathcal{E};\ u \in E\}$, est un sous-espace de $\mathcal{E}$, et on a par construction

$$\|\dot{\tilde{u}}\|_{\mathcal{E}} = \|u\| \quad \forall u \in E. \tag{1.12}$$

Proposition 1.14 *Avec les notations ci-dessus, l'espace $\mathcal{E}$ muni de la norme $\|\cdot\|_{\mathcal{E}}$ définie en* (1.11) *est complet, et $\tilde{E}$ est dense dans $\mathcal{E}$.*

Démonstration. Vérifions tout d'abord que $\tilde{E}$ est dense dans $\mathcal{E}$. Soit $U = (u_n)$ une suite de Cauchy de E. Pour tout $m \in \mathbb{N}$, notons comme précédemment $\widetilde{u_m}$ la suite constante égale à u_m. Montrons que

$$\lim_{m\to\infty} \|\dot{\widetilde{u_m}} - \dot{U}\|_{\mathcal{E}} = 0, \tag{1.13}$$

autrement dit que

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ m > N \Rightarrow \lim_{n\to\infty} \|u_m - u_n\| \leq \varepsilon.$$

Il s'agit en fait d'une conséquence immédiate de la définition (1.10) d'une suite de Cauchy où il suffit de passer à la limite $n \to +\infty$.

Montrons maintenant que $\mathcal{E}$ muni de la norme $\|\cdot\|_{\mathcal{E}}$ est complet. Soit $(\dot{U}_n)$ une suite de Cauchy de $\mathcal{E}$ (c'est-à-dire une suite de Cauchy de suites de Cauchy !). D'après le résultat de densité qui précède, pour tout $n \in \mathbb{N}$, on peut trouver $u_n \in E$ tel que $\|\dot{\widetilde{u_n}} - \dot{U}_n\|_{\mathcal{E}} \leq 1/n$. Ainsi

$$\lim_{n\to\infty} \|\dot{\widetilde{u_n}} - \dot{U}_n\|_{\mathcal{E}} = 0. \tag{1.14}$$

La suite $(\dot{U}_n)$ étant de Cauchy, il s'ensuit que $(\dot{\widetilde{u_n}})$ l'est aussi puisque

$$\|\dot{\widetilde{u_m}} - \dot{\widetilde{u_n}}\|_{\mathcal{E}} \leq \|\dot{\widetilde{u_m}} - \dot{U}_m\|_{\mathcal{E}} + \|\dot{U}_m - \dot{U}_n\|_{\mathcal{E}} + \|\dot{\widetilde{u_n}} - \dot{U}_n\|_{\mathcal{E}}.$$

D'après (1.12), ceci signifie que (u_n) est une suite de Cauchy de E. Notons U cette suite. D'après (1.13) et (1.14), on conclut que $\lim_{n\to\infty} \|\dot{U}_n - \dot{U}\|_{\mathcal{E}} = 0$. □

symétrique (soit $U \sim V \Rightarrow V \sim U$ pour tout $(U, V) \in \mathcal{U}^2$) et *transitive* (soit ($U \sim V$ et $V \sim W$) $\Rightarrow$ $U \sim W$ pour tout $(U, V, W) \in \mathcal{U}^3$). La *classe d'équivalence* $\dot{U}$ associée à un élément $U \in \mathcal{U}$ est l'ensemble de tous les $V \in \mathcal{U}$ tels que $V \sim U$, appelés représentants de la classe. Enfin, l'ensemble de ces classes d'équivalences est appelé l'*ensemble quotient* de $\mathcal{U}$ par la relation d'équivalence.

Notons qu'avec cette construction de $\mathcal{E}$, nous n'avons pas tout à fait répondu à la question initiale qui était de trouver ce qu'on pourrait appeler un "sur-espace" de E. Nous avons en fait construit un espace qui n'est pas de même nature que l'espace de départ, où le second s'identifie à un sous-espace du premier. En effet, $\mathcal{E}$ contient non pas des éléments de E mais des classes d'équivalence de suites de Cauchy de E, ce qui complique sérieusement les choses ! Cette difficulté n'est pas seulement due au caractère abstrait de la construction : elle est liée à une difficulté de fond dans la question elle-même (quel sens peut-on donner à l'idée de "sur-espace" ?), mais elle est surtout significative de difficultés pratiques qui peuvent apparaître lorsqu'on applique ce processus à un cas concret. Nous rencontrerons ce type de difficultés dans le chapitre 2, lors de la construction d'un espace complet pour la norme $\|\cdot\|_1$.

1.3 Quelques résultats utiles

1.3.1 Prolongement d'une application linéaire

La proposition suivante est très utile en pratique. Elle montre que la complétude nous permet d'étendre à tout un espace une application définie initialement sur un sous-espace dense de cet espace.

Proposition 1.15 *Soient $(E, \|\cdot\|_E)$ un espace vectoriel normé, A un sous-espace dense de E, $(F, \|\cdot\|_F)$ un espace vectoriel normé complet, et $f : A \to F$ une application linéaire. Supposons qu'il existe une constante $C > 0$ telle que*

$$\|f(u)\|_F \leq C \ \|u\|_E \quad \forall u \in A, \tag{1.15}$$

alors il existe une unique application linéaire continue $\tilde{f} : E \to F$ qui prolonge f, au sens où $\tilde{f}(u) = f(u)$ pour tout $u \in A$.

Démonstration. Soit $u \in E$. Comme A est dense dans E, on peut trouver une suite (u_n) de A qui tend vers u. Cette suite est donc de Cauchy dans E, et d'après (1.15), la suite $(f(u_n))$ est de Cauchy dans F (puisque $\|f(u_n) - f(u_m)\|_F \leq C \ \|u_n - u_m\|_E$). Comme F est complet, elle est convergente. Notons $\tilde{f}(u)$ sa limite, ce qui est licite car cette limite ne dépend pas de la suite (u_n) mais seulement de u. En effet, si (v_n) est une autre suite de A qui tend vers u, on aura

$$\|f(u_n) - f(v_n)\|_F \leq C \ \|u_n - v_n\|_E \leq C \ \left(\|u_n - u\|_E + \|v_n - u\|_E\right),$$

où le terme de droite tend vers 0. L'application $\tilde{f} : E \to F$ ainsi construite est clairement linéaire, et il suffit de passer à la limite sur l'inégalité $\|f(u_n)\|_F \leq C \ \|u_n\|_E$ pour obtenir sa continuité. Il s'agit bien d'un prolongement de f : si $u \in A$, on peut choisir la suite (u_n) constante égale à u, ce qui montre que $\tilde{f}(u) = f(u)$. Enfin ce prolongement est bien unique, dès lors qu'on lui demande d'être continu : s'il en existait deux, leur différence $\tilde{d}$ serait telle que $\tilde{d}(u) = 0$ pour tout $u \in A$, et donc aussi pour tout $u \in E$ d'après la continuité de $\tilde{d}$ et la densité de A dans E. □

1.3.2 Séries convergentes et normalement convergentes

Soit $(E, \|\cdot\|)$ un espace vectoriel normé. Pour toute suite (u_n) de E, on appelle série de terme général u_n la suite (S_n) de E définie par

$$S_n = \sum_{k=0}^{n} u_k.$$

La quantité S_n est appelée la somme partielle d'indice n de la série. On dit que cette série converge si la suite (S_n) converge dans E. La limite S de cette suite est appelée la somme de la série et on note $S = \sum_{k=0}^{\infty} u_k$.

Définition 1.16 *On dit que la série (S_n) est normalement convergente si la série réelle positive $\left(\sum_{k=0}^{n} \|u_k\|\right)$ est convergente.*

Proposition 1.17 *Si $(E, \|\cdot\|)$ est un espace de Banach, toute série normalement convergente est convergente et*

$$\left\| \sum_{k=0}^{\infty} u_k \right\| \leq \sum_{k=0}^{\infty} \|u_k\| \tag{1.16}$$

Démonstration. Soit $(S_n) = \left(\sum_{k=0}^{n} u_k\right)$ une série normalement convergente. Comme $(E, \|\cdot\|)$ est complet, il suffit de vérifier que (S_n) est de Cauchy, ce qui revient à dire que

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \geq N,\ \forall p \geq 1,\ \|S_{n+p} - S_n\| = \left\| \sum_{k=n+1}^{n+p} u_k \right\| \leq \varepsilon.$$

Ceci découle d'une part du fait que comme (S_n) est normalement convergente, la série réelle $\left(\sum_{k=0}^{n} \|u_k\|\right)$ est de Cauchy, soit

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \geq N,\ \forall p \geq 1,\ \sum_{k=n+1}^{n+p} \|u_k\| \leq \varepsilon,$$

et d'autre part de l'inégalité triangulaire

$$\left\| \sum_{k=n+1}^{n+p} u_k \right\| \leq \sum_{k=n+1}^{n+p} \|u_k\|.$$

On sait donc que la série (S_n) converge. Notons S sa somme.

L'inégalité triangulaire nous montre par ailleurs que

$$\|S_n\| \leq \sum_{k=0}^{n} \|u_k\|.$$

La propriété (1.16) annoncée revient simplement à passer à la limite $n \to \infty$ dans cette inégalité. Pour cela, il suffit de remarquer que (toujours d'après l'inégalité triangulaire)

$$\Big| \|S_n\| - \|S\| \Big| \leq \|S_n - S\|,$$

d'où l'on déduit que $\lim_{n\to\infty} \|S_n\| = \|S\|$. □

On le voit dans cette démonstration : le caractère complet de $(E, \|\cdot\|)$ est tout à fait essentiel pour assurer la convergence d'une série normalement convergente. Comme le montre la proposition suivante, réciproque de la proposition 1.17, il s'agit d'une condition nécessaire et suffisante.

Proposition 1.18 *Soit $(E, \|\cdot\|)$ un espace normé. Si toute série normalement convergente est convergente, alors E est complet.*

Démonstration. Soit (u_n) une suite de Cauchy de E. Montrons qu'elle converge. Pour cela, on va construire une série normalement convergente dont la somme sera la limite de (u_n).

La suite (u_n) étant de Cauchy, on sait que pour tout $\ell \in \mathbb{N}$, il existe un entier n_ℓ tel que

$$\forall n \geq n_\ell,\ \forall m \geq n_\ell,\ \|u_n - u_m\| \leq \frac{1}{2^\ell}.$$

La suite d'entiers $(n_\ell)_{\ell\in\mathbb{N}}$ ainsi construite peut évidemment être choisie strictement croissante. On considère alors la série (S_ℓ) de terme général v_ℓ où $v_0 = u_{n_0}$ et $v_\ell = u_{n_\ell} - u_{n_{\ell-1}}$ pour tout $\ell \geq 1$. On a par construction

$$S_\ell = \sum_{k=0}^{\ell} v_k = u_{n_\ell} \quad \text{et} \quad \sum_{k=0}^{\ell} \|v_k\| \leq \|u_{n_0}\| + \sum_{k=1}^{\ell} \frac{1}{2^{k-1}}.$$

Cette dernière inégalité nous montre que la série (S_ℓ) est normalement convergente. Elle est donc convergente par hypothèse. D'après l'égalité ci-dessus, sa limite S n'est autre que la limite de la sous-suite $(u_{n_\ell})_{\ell\in\mathbb{N}}$ lorsque $\ell \to \infty$. En remarquant par ailleurs que pour tout $n \geq n_\ell$, on a

$$\|u_n - S\| \leq \|u_n - u_{n_\ell}\| + \|u_{n_\ell} - S\| \leq \frac{1}{2^\ell} + \sum_{k=\ell+1}^{\infty} \frac{1}{2^{k-1}} = \frac{3}{2^\ell},$$

on en déduit que S est la limite de la suite (u_n) (non pas seulement de la sous-suite (u_{n_ℓ})). □

Chapitre 2

L'intégrale de Lebesgue

Ce chapitre tient une place assez particulière dans ce cours, dans la mesure où il ne s'agit pas à proprement parler d'un "cours" sur l'intégrale de Lebesgue, mais plutôt d'un "guide d'utilisation" de cette intégrale, introduite par Lebesgue au début du 20e siècle afin de généraliser et combler certaines lacunes laissées par la théorie de l'intégration proposée par Riemann un demi siècle plus tôt. De même qu'il n'est pas nécessaire de connaître le fonctionnement d'un moteur pour conduire une voiture, de même on peut manipuler l'intégrale de Lebesgue sans en connaître toutes les subtilités. Conduire une voiture nécessite toute de même un permis. On pourrait dire que ce chapitre est une aide à l'obtention du permis "L", qui donne le droit d'utiliser l'intégrale de Lebesgue !

Ce point de vue constitue en quelque sorte une voie médiane entre un point de vue minimaliste qui se contenterait de travailler avec l'intégrale de Riemann enseignée en classes préparatoires, et un point de vue maximaliste qui détaillerait la construction et les propriétés de l'intégrale de Lebesgue.

Les partisans du premier pourraient avancer que l'intégrale de Riemann est tout à fait suffisante pour nos besoins. De fait, c'est vrai dans la grande majorité des cas où nous aurons à calculer une intégrale. Les quelques rares exemples de fonctions intégrables au sens de Lebesgue mais pas au sens de Riemann ont un intérêt très limité en pratique. Ainsi, en introduisant l'intégrale de Lebesgue, notre motivation première est d'ordre théorique. Comme nous l'avons évoqué au chapitre 1, elle est liée principalement à la recherche d'un espace de fonctions qui soit *complet* pour la norme de la convergence en moyenne. En outre, on dispose avec l'intégrale de Lebesgue d'outils particulièrement efficaces pour les besoins de l'analyse, notamment pour permuter une limite et une intégrale, dériver une intégrale dépendant d'un paramètre, ou permuter deux intégrales. Ce chapitre est essentiellement destiné à se familiariser avec ces outils.

Les puristes rétorqueront que l'on risque d'utiliser ces outils de façon incorrecte si l'on n'a pas compris les fondements de l'intégrale de Lebesgue. C'est probablement vrai dans les domaines où ces fondements jouent un rôle capital. Par exemple, les probabilistes ne trouveront pas dans ce chapitre les éléments de *théorie de la mesure* dont ils ont besoin. Ce chapitre vise un public généraliste, et il ne nous semble pas raisonnable de demander à des étudiants de première année d'école d'ingénieurs de

maîtriser les aspects théoriques de l'intégrale de Lebesgue. La construction détaillée de cette intégrale est difficile, et nécessiterait à elle seule plusieurs séances de cours, au détriment des autres thèmes qui ont été retenus. Celles et ceux qui voudraient aller plus loin pourront consulter l'excellent livre de W. Rudin [3].

2.1 De Riemann à Lebesgue

2.1.1 L'intégrale de Riemann

Commençons par rappeler brièvement le principe de la construction de l'intégrale de Riemann. Nous nous limitons pour simplifier au cas d'une seule dimension d'espace, en considérant une fonction u définie sur un intervalle borné $[a, b] \subset \mathbb{R}$ à valeurs dans $\mathbb{R}$. On suppose de plus que u est *continue*. Dans une représentation graphique de u, l'intégrale de u correspond à la mesure de la surface comprise entre l'axe des abscisses et la courbe représentant u (cette surface étant comptée positivement "au dessus" de l'axe des abscisses et négativement "en dessous"). L'intégrale de Riemann de u revient à ramener la mesure de cette surface à la mesure de surfaces élémentaires, en l'occurrence des rectangles, grâce à un découpage "vertical" de cette surface. L'idée consiste simplement à approcher u par une suite de fonctions en escalier, autrement dit constantes par morceaux, en introduisant des subdivisions de $[a, b]$, c'est-à-dire des découpages de $[a, b]$ en "petits" intervalles (au sens où leur longueur est destinée à tendre vers 0). On peut par exemple choisir des subdivisions régulières : pour un entier $n \geq 1$, on définit les points

$$x_i^{(n)} = a + i\,\frac{b-a}{n} \quad \text{pour } i = 0 \text{ à } n.$$

Ainsi les intervalles $[x_{i-1}^{(n)}, x_i^{(n)}]$, pour $i = 1$ à n, sont tous de longueur $(b-a)/n$. Sur chacun d'eux, on va approcher $u(x)$ par une valeur constante, par exemple $u(x_{i-1}^{(n)})$. La fonction u étant continue, il est clair que l'erreur que l'on commet devient de plus en plus petite lorsque la longueur des intervalles diminue, soit $n \to +\infty$. L'intégrale de Riemann de u est alors définie par

$$\int_a^b u(x)\,\mathrm{d}x = \lim_{n\to+\infty} \frac{b-a}{n} \sum_{i=1}^{n} u(x_{i-1}^{(n)}). \tag{2.1}$$

On peut vérifier que cette limite existe pour toute fonction continue sur $[a, b]$ (ou continue par morceaux) et que cette définition ne dépend ni de la famille de subdivisions choisie, ni de la valeur particulière de x choisie dans chaque intervalle $[x_{i-1}^{(n)}, x_i^{(n)}]$. Cette définition s'étend à des fonctions qui ne sont pas nécessairement continues. Il est par contre difficile de caractériser l'espace de toutes les fonctions qui sont intégrables au sens de Riemann. Cet espace est complet pour la topologie de la convergence uniforme, ce qui découle essentiellement de la proposition suivante qui nous dit que l'ensemble

des fonctions intégrables au sens de Riemann est fermé pour la norme de la convergence uniforme.

Proposition 2.1 *Soit $(u_n)_{n\in\mathbb{N}}$ une suite de fonctions de $[a,b]$ dans $\mathbb{R}$ intégrables au sens de Riemann qui converge uniformément vers une fonction u, soit*

$$\lim_{n\to+\infty} \|u_n - u\|_\infty = \lim_{n\to+\infty} \sup_{x\in[a,b]} |u_n(x) - u(x)| = 0.$$

Alors u est intégrable au sens de Riemann et

$$\int_a^b u(x)\,\mathrm{d}x = \lim_{n\to+\infty} \int_a^b u_n(x)\,\mathrm{d}x.$$

En revanche, si l'on munit l'espace des fonctions intégrables au sens de Riemann de la norme de la convergence en moyenne (cf. (1.5)) définie par l'intégrale de Riemann elle-même, cet espace n'est pas complet. C'est là un inconvénient majeur de l'intégrale de Riemann, qui la rend insuffisante pour les besoins de l'analyse. L'intégrale de Lebesgue permet notamment de remédier à cet inconvénient. Elle conduit à étendre la notion d'intégrabilité à une classe de fonctions plus large. En particulier, l'intégrale de Riemann se limite à des fonctions bornées sur des domaines bornés, certaines situations plus générales pouvant être traitées par la notion d'intégrale généralisée (appelée aussi "impropre") qui revient à passer à la limite sur le domaine d'intégration. L'intégrale de Lebesgue permet quant à elle d'intégrer directement des fonctions pouvant être non bornées sur des domaines bornés ou non.

2.1.2 Définition de l'intégrale de Lebesgue

Une expression nouvelle : "presque partout"

La définition de l'intégrale de Lebesgue que nous allons donner plus bas est reliée au processus général de construction du complété d'un espace incomplet, évoqué brièvement dans le paragraphe 1.2. Nous avons vu que ce complété contient des classes d'équivalence de suites de Cauchy. Que représentent ici ces objets abstraits ? Notre objectif étant d'intégrer des *fonctions*, on peut s'attendre à ce que chaque classe d'équivalence puisse être identifiée à une fonction. C'est *presque* vrai. En fait, une telle classe d'équivalence va être identifiée non pas à *une* fonction, mais à toute une famille de fonctions dont les différences ne peuvent être perçues par une intégrale : ces fonctions sont égales "presque partout". Nous allons commencer par expliquer le sens de cette expression, qui marque le passage de Riemann à Lebesgue. Son utilisation récurrente (on la trouve *presque partout* dans un cours sur l'intégrale de Lebesgue !) est souvent mal comprise. Elle est liée au fait que l'intégrale de Lebesgue va nous permettre d'intégrer des fonctions beaucoup moins régulières que l'intégrale de Riemann.

Pour le comprendre, considérons la fonction indicatrice d'un ensemble $A \subset \mathbb{R}$, soit

$$1_A(x) = \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{si } x \in \mathbb{R} \setminus A. \end{cases}$$

Si A est réduit à un point, il est clair que 1_A est intégrable sur $\mathbb{R}$ (au sens de Riemann, car c'est une fonction continue par morceaux) et que $\int_{\mathbb{R}} 1_A(x)\,\mathrm{d}x = 0$. C'est évidemment encore vrai si A est composé d'un nombre fini de points. Si A est un ensemble dénombrable de points, par exemple l'ensemble $\mathbb{Q}$ des nombres rationnels [1], la fonction 1_A n'est plus intégrable au sens de Riemann. Mais l'interprétation géométrique de l'intégrale nous pousse à croire que si on sait donner un sens à l'intégrale de 1_A, celle-ci sera nulle (c'est une somme dénombrable de 0 !). Jusqu'à quel type d'ensemble A peut-on aller ? Bien sûr, si A contient un intervalle ouvert, aussi petit soit-il, l'intégrale de 1_A ne sera plus nulle. Intuitivement, pour que cette intégrale s'annule, il faut que la "longueur" de A soit nulle, et on dira dans ce cas que A est *négligeable*. Mais comment mesurer cette longueur lorsque A est plus compliqué qu'un intervalle, ou une union d'intervalles ? La définition suivante n'utilise que la mesure d'intervalles.

Définition 2.2 (i) *Un ensemble $A \subset \mathbb{R}$ est dit négligeable (ou de mesure nulle) si, pour tout $\varepsilon > 0$, on peut trouver une famille finie ou dénombrable d'intervalles $]a_j, b_j[\subset \mathbb{R}$, où $j \in \mathbb{J} \subset \mathbb{N}$, telle que*

$$A \subset \bigcup_{j\in\mathbb{J}}]a_j, b_j[\quad \textit{et} \quad \sum_{j\in\mathbb{J}} (b_j - a_j) \leq \varepsilon.$$

(ii) *On dit qu'une propriété dépendant d'un paramètre $x \in X \subset \mathbb{R}$ est vraie presque partout (ou vraie pour presque tout $x \in X$; en abrégé : p.p.) si elle est vraie pour tout $x \in X \setminus A$ où $A \subset \mathbb{R}$ est négligeable.*

Exemple 2.3 Cette définition permet évidemment de retrouver qu'un singleton $\{c\} \subset \mathbb{R}$ est négligeable : il suffit de prendre un seul intervalle $]c - \varepsilon/2, c + \varepsilon/2[$. De même, un ensemble fini de points $\{c_j;\ j \in \mathbb{J}\}$ est négligeable : on peut choisir la famille finie d'intervalles $]c_j - \varepsilon/(2n), c_j + \varepsilon/(2n)[$, pour $j \in \mathbb{J}$. Dans le cas où cet ensemble n'est plus fini, mais dénombrable (soit $\mathbb{J} = \mathbb{N}$), on pourra prendre les intervalles $]c_j - \varepsilon_j, c_j + \varepsilon_j[$ avec $\varepsilon_j = \varepsilon/2^{j+1}$ pour tout $j \in \mathbb{N}$. Ainsi, un ensemble dénombrable de points est négligeable. Notons enfin qu'il existe des ensembles négligeables qui ne sont pas dénombrables (mais leur construction sort du cadre de ce cours). •

Exemple 2.4 Nous savons maintenant qu'un ensemble dénombrable est négligeable. Donc $\mathbb{Q}$ est négligeable, et nous pouvons alors affirmer que *les nombres réels sont presque partout irrationnels.* •

1. La fonction $1_{\mathbb{Q}}$ est un exemple classique de fonction qui est intégrable au sens de Lebesgue, mais pas au sens de Riemann.

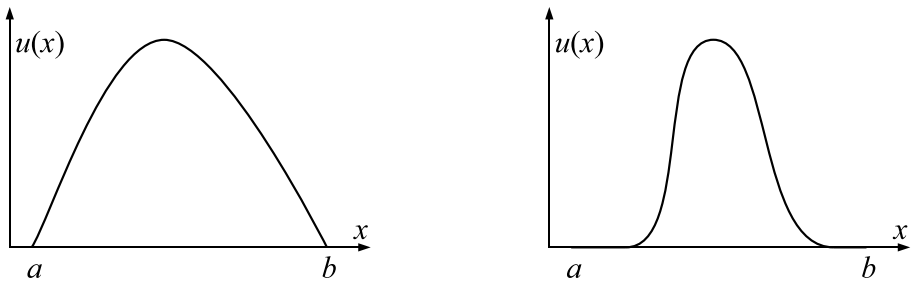


FIGURE 2.1 – Exemples de fonctions $u \notin \mathcal{D}^0(]a,b[)$ (à gauche) et $u \in \mathcal{D}^0(]a,b[)$ (à droite)

La définition ci-dessus se généralise immédiatement au cas de plusieurs dimensions d'espace, soit $A \subset \mathbb{R}^d$. Il suffit de remplacer les intervalles ouverts $]a_j, b_j[$ par des pavés ouverts de la forme $]a_1^{(j)}, b_1^{(j)}[\times \ldots \times]a_n^{(j)}, b_n^{(j)}[$.

Le rôle crucial que jouent ces notions dans le cadre de l'intégrale de Lebesgue réside dans le fait qu'elles nous fournissent un critère pour la nullité d'une intégrale : dire que $\int |u(x)|\,\mathrm{d}x = 0$ est équivalent à dire que $u(x)$ s'annule presque partout. En d'autres termes, si l'on modifie les valeurs prises par une fonction sur un sous-ensemble négligeable du domaine d'intégration, on ne modifie pas la valeur de l'intégrale.

Une définition abstraite

Nous sommes maintenant en mesure de donner une définition de l'intégrale de Lebesgue sur un domaine $\Omega \subset \mathbb{R}^d$. Pour éviter les difficultés relatives au bord du domaine d'intégration, et pour traiter de la même façon les domaines bornés et non bornés, nous allons introduire l'espace $\mathcal{D}^0(\Omega)$ formé des fonctions définies de Ω dans $\mathbb{R}$ qui sont continues et dont le *support* est un *compact* strictement inclus dans Ω, c'est-à-dire inclus dans l'intérieur de Ω. Rappelons que le *support* d'une fonction $u : \Omega \to \mathbb{R}$ désigne la zone où elle ne s'annule pas, plus exactement l'adhérence de l'ensemble des points où elle ne s'annule pas, soit

$$\mathrm{Supp}(u) = \overline{u^{-1}(\mathbb{R}^*)} = \overline{\{x \in \Omega;\ u(x) \neq 0\}}, \tag{2.2}$$

la notation $\overline{A}$ désignant ici l'adhérence d'un sous-ensemble A de $\mathbb{R}^d$. En d'autres termes, $\mathrm{Supp}(u)$ *est le plus petit fermé à l'extérieur duquel* u *s'annule*. Dans le cas où $\Omega =]a,b[$, la figure 2.1 donne un exemple de fonction qui n'appartient pas à $\mathcal{D}^0(]a,b[)$ (à gauche) puisque son support est l'intervalle fermé $[a,b]$, alors que celle de droite a son support contenu dans $]a,b[$.

Définition 2.5 *Soit Ω un domaine de $\mathbb{R}^d$, borné ou non. Une fonction $u : \Omega \to \mathbb{R}$ est dite intégrable au sens de Lebesgue s'il existe une suite (u_n) de fonctions de $\mathcal{D}^0(\Omega)$ telle que*

(i) $u_n(x) \to u(x)$ *pour presque tout* $x \in \Omega$;

(ii) *cette suite est de Cauchy pour la norme de la convergence en moyenne, soit*

$$\lim_{n,m\to+\infty} \int_\Omega |u_n(x) - u_m(x)|\,\mathrm{d}x = 0.$$

Considérons alors une fonction u qui satisfait cette définition, approchée par une suite de Cauchy (u_n). La suite de réels (r_n) définie par $r_n = \int_\Omega u_n(x)\,\mathrm{d}x$ est de Cauchy puisque

$$|r_n - r_m| = \left| \int_\Omega (u_n(x) - u_m(x))\,\mathrm{d}x \right| \leq \int_\Omega |u_n(x) - u_m(x)|\,\mathrm{d}x,$$

et elle converge donc vers un nombre réel r. On peut montrer (ce qui n'est pas évident !) que ce nombre ne dépend pas de la suite de Cauchy choisie pour approcher u : c'est la valeur de l'intégrale de u, et nous conserverons la notation habituelle pour la désigner, d'où la définition suivante.

Définition 2.6 *Soient $u : \Omega \to \mathbb{R}$ une fonction intégrable au sens de Lebesgue et (u_n) une suite de fonctions de $\mathcal{D}^0(\Omega)$ qui approche u au sens de la définition* 2.5*. Alors l'intégrale de Lebesgue de u est définie par*

$$\int_\Omega u(x)\,\mathrm{d}x = \lim_{n\to+\infty} \int_\Omega u_n(x)\,\mathrm{d}x. \tag{2.3}$$

Sur un plan pratique, les deux définitions ci-dessus sont difficilement utilisables, dans la mesure où elles ne nous fournissent pas directement un moyen simple pour vérifier si telle ou telle fonction est intégrable au sens de Lebesgue. Nous verrons dans le paragraphe 2.2 quelques critères qui permettent de le faire.

Remarque 2.7 *La condition* (i) *de la définition* 2.5 *caractérise les fonctions dites mesurables. Nous avons choisi de passer sous silence cette notion assez subtile. La raison de ce choix est simple : dans la pratique, nous ne risquons pas de nous retrouver face à une fonction non mesurable, dont la construction ne peut se faire qu'en faisant appel à l'axiome du choix... Nous allons donc faire comme si toutes les fonctions étaient mesurables, ce qui n'est ni totalement faux, ni totalement vrai ! Tout dépend du système d'axiomes dans lequel on choisit de se placer. Pour les puristes, il suffit généralement de remplacer le mot "fonction" par "fonction mesurable" pour que les énoncés des résultats donnés dans ce chapitre deviennent parfaitement corrects.*

S'agit-il bien d'une intégrale ?

On peut se demander si notre choix de notation dans (2.3) est licite, si cette limite représente vraiment une intégrale. Il est en effet possible, par un passage à la limite similaire, de construire des objets qui ne sont pas intégrables. Pour le voir, considérons la suite de fonctions "chapeaux" $(u_n) \subset \mathcal{C}^0([a,b])$ de l'exemple 1.4. Il est clair que

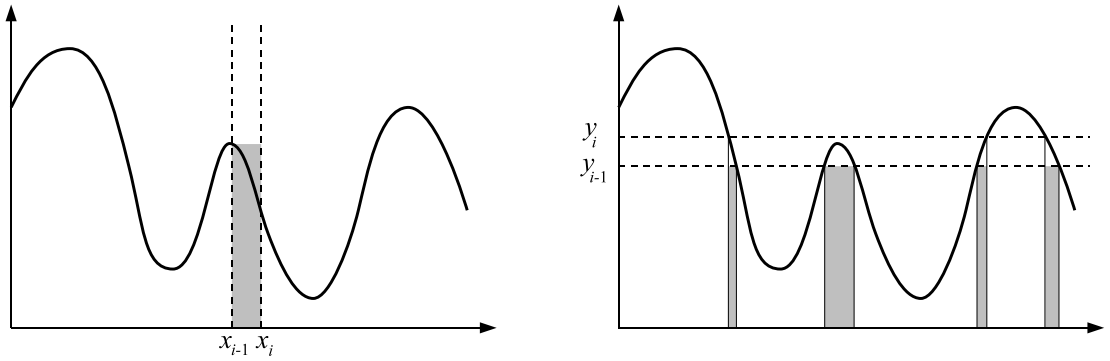


FIGURE 2.2 – L'intégrale de Riemann (à gauche) et celle de Lebesgue (à droite)

$u_n \to 0$ presque partout, puisque si $x \neq c$, on a $u_n(x) = 0$ dès que $n > 1/|x - c|$. De plus $\int_a^b u_n(x)\,\mathrm{d}x = 1$ pour tout $n \geq 1$. Ainsi, (u_n) tend presque partout vers la fonction nulle alors que son intégrale reste constante égale à 1. Où est l'erreur ? Pourquoi cet exemple ne rentre-t-il pas dans le cadre de la définition 2.5 ? La réponse réside bien sûr dans le fait que la suite (u_n) n'est pas de Cauchy pour la norme de la convergence en moyenne. En effet, on vérifie facilement que

$$\int_a^b |u_{2n}(x) - u_n(x)|\,\mathrm{d}x > \int_{c-1/(3n)}^{c+1/(3n)} |u_{2n}(x) - u_n(x)|\,\mathrm{d}x = \frac{1}{3}.$$

Ainsi l'objet que l'on décrit par ce passage à la limite ne peut être une fonction intégrable. Ce n'est même pas une *fonction*. On peut comprendre la signification de cet objet étrange en interprétant $u_n(x)$ comme une densité de masse. Pour tout $n \geq 1$, la masse totale, représentée par l'intégrale, est égale à 1. A la limite, la densité vaut 0 partout sauf en $x = c$ où la masse s'est concentrée. Mais cette fois la mesure de cette masse n'est plus représentée par une intégrale. Ainsi, cet objet décrit la densité de masse d'une masse ponctuelle, objet bien connu des physiciens : c'est une *mesure de Dirac*. Nous verrons au chapitre 3 qu'on peut en donner une interprétation mathématique, à l'aide de la théorie des *distributions*.

Pour revenir à (2.3), l'utilisation de la notation habituelle d'une intégrale est justifiée par le fait que la suite qui approche u est de Cauchy précisément pour la norme définie par l'intégrale elle-même. Bien sûr, l'interprétation de l'intégrale de Riemann en tant que limite d'intégrales de fonctions en escalier (voir (2.1)) n'est en général plus adaptée ici. En fait, l'interprétation donnée initialement par Lebesgue consiste en un procédé d'approximation similaire, où l'on considère une subdivision non plus de l'ensemble de départ, c'est-à-dire du domaine d'intégration, mais de l'ensemble d'arrivée $u(\Omega)$. Considérons le cas de la dimension 1 avec $\Omega =]a, b[$ et supposons pour simplifier que $u(]a, b[)$ est contenu dans un intervalle borné. Pour $n \geq 1$, notons $\{y_i^{(n)}; i = 0 \text{ à } n\}$ une famille de subdivisions (par exemple régulières) de cet intervalle, dont le pas tend vers

0 lorsque $n \to +\infty$. L'idée consiste alors à remplacer la formule (2.1) par

$$\int_a^b u(x)\,\mathrm{d}x = \lim_{n\to+\infty} \sum_{i=1}^{n} y_{i-1}^{(n)}\ \mathrm{mes}\left(u^{-1}([y_{i-1}^{(n)}, y_i^{(n)}[)\right), \tag{2.4}$$

où mes(A) désigne la mesure d'un ensemble $A \subset \mathbb{R}$, autrement dit sa "longueur". La figure 2.2 illustre le passage de (2.1) à (2.4). La figure de gauche correspond à une subdivision *verticale* associée à l'intégrale de Riemann : la partie grisée représente la contribution d'un terme de la somme dans (2.1). La figure de droite montre qu'à l'intégrale de Lebesgue est associée une subdivision *horizontale*, la partie grisée représentant cette fois la contribution d'un terme de la somme dans (2.4). On voit ici que la structure de l'ensemble $u^{-1}([y_{i-1}^{(n)}, y_i^{(n)}[)$ peut devenir particulièrement complexe. Il faut donc au préalable savoir donner un sens à la mesure d'un tel ensemble (alors que l'intégrale de Riemann se contente de la mesure d'intervalles). Ceci constitue l'objet de la *théorie de la mesure* que nous n'aborderons pas ici, mais qui est nécessaire pour comprendre la représentation (2.4) de l'intégrale de Lebesgue. Cette théorie fournit un moyen constructif pour introduire l'intégrale de Lebesgue indépendamment de celle de Riemann (le lecteur intéressé pourra par exemple consulter le célèbre livre de W. Rudin [3]). Il existe d'autres méthodes, qui donnent un cadre mathématique rigoureux à la démarche formelle que nous avons adoptée ici (voir par exemple [5]) : elles permettent de généraliser l'intégrale de Riemann en s'appuyant sur la construction du complété d'un espace incomplet (voir paragraphe 1.2).

2.1.3 Espaces $\mathcal{L}^1$ et L^1

Pour un domaine $\Omega \subset \mathbb{R}^d$, nous noterons $\mathcal{L}^1(\Omega)$ l'ensemble des fonctions intégrables sur Ω au sens de Lebesgue. Au vu des définitions 2.5 et 2.6, l'intégrale de Lebesgue hérite des propriétés usuelles de l'intégrale de Riemann. La proposition suivante regroupe trois de ces propriétés.

Proposition 2.8 *Soient Ω un domaine de $\mathbb{R}^d$, et u, v deux fonctions de $\mathcal{L}^1(\Omega)$.*

(i) *Pour tous $\lambda, \mu \in \mathbb{R}$, on a $\lambda\, u + \mu\, v \in \mathcal{L}^1(\Omega)$ et*

$$\int_\Omega (\lambda\, u(x) + \mu\, v(x))\,\mathrm{d}x = \lambda \int_\Omega u(x)\,\mathrm{d}x + \mu \int_\Omega v(x)\,\mathrm{d}x.$$

(ii) *On a $|u| \in \mathcal{L}^1(\Omega)$ et* $\displaystyle\left|\int_\Omega u(x)\,\mathrm{d}x\right| \le \int_\Omega |u(x)|\,\mathrm{d}x.$

(iii) *Si $u \le v$ presque partout dans Ω, alors* $\displaystyle\int_\Omega u(x)\,\mathrm{d}x \le \int_\Omega v(x)\,\mathrm{d}x.$

Démonstration. Pour $u \in \mathcal{L}^1(\Omega)$, notons (u_n) une suite de fonctions de $\mathcal{D}^0(\Omega)$ qui approche u au sens de la définition 2.5. Si $\lambda \in \mathbb{R}$, il est clair que $\lambda\, u_n(x) \to \lambda\, u(x)$ pour

presque tout $x \in \Omega$, et que la suite $(\lambda\, u_n)$ est de Cauchy pour la norme de la convergence en moyenne, puisque $|\lambda\, u_n(x) - \lambda\, u_m(x)| = |\lambda|\, |u_n(x) - u_m(x)|$. Ainsi $u \in \mathcal{L}^1(\Omega)$ et la linéarité de l'intégrale de Riemann nous montre que

$$\int_\Omega \lambda\, u(x)\, \mathrm{d}x = \lim_{n\to+\infty} \int_\Omega \lambda\, u_n(x)\, \mathrm{d}x = \lambda \lim_{n\to+\infty} \int_\Omega u_n(x)\, \mathrm{d}x = \lambda \int_\Omega u(x)\, \mathrm{d}x.$$

On procède de la même façon pour $u + v$, et (i) en découle. Pour (ii), on suit le même raisonnement avec $|u|$, en remarquant que $||u_n(x)| - |u_m(x)|| \leq |u_n(x) - u_m(x)|$. Enfin d'après (i), le point (iii) revient à dire que si $u \geq 0$ presque partout dans Ω, alors $\int_\Omega u(x)\, \mathrm{d}x \geq 0$, ce qui découle de (ii) puisque $u = |u|$ presque partout. □

Cette proposition entraîne que $\mathcal{L}^1(\Omega)$ est un espace vectoriel, et que l'application $u \mapsto \int_\Omega |u(x)|\, \mathrm{d}x$ définit une *semi-norme* sur $\mathcal{L}^1(\Omega)$ (voir la définition 1.3 d'une semi-norme). Par contre, cette application ne définit pas une *norme* sur $\mathcal{L}^1(\Omega)$, car la nullité de l'intégrale n'entraîne pas la nullité de la fonction. C'est précisément la raison d'être de l'expression "presque partout" introduite au début du paragraphe 2.1.2. Il découle en effet de la définition de l'intégrale de Lebesgue que si u est nulle presque partout, elle appartient à $\mathcal{L}^1(\Omega)$ et son intégrale est nulle. La réciproque est vraie (mais pas évidente à démontrer) si u ne change pas de signe, ce qu'on peut résumer sous la forme suivante.

Proposition 2.9 *Pour toute fonction* $u : \Omega \to \mathbb{R}$, *on a*

$$\left(u \in \mathcal{L}^1(\Omega) \text{ et } \int_\Omega |u(x)|\, \mathrm{d}x = 0\right) \iff (u = 0 \text{ presque partout dans } \Omega).$$

Cette proposition nous montre que l'intégrale ne fait pas de distinction entre deux fonctions qui diffèrent sur un ensemble négligeable. Ceci suggère de passer à l'espace quotient (voir note 1 page 16) de $\mathcal{L}^1(\Omega)$ par la relation d'égalité presque partout, espace qui sera noté $L^1(\Omega)$. La relation

$$u \sim v \iff (u = v \text{ presque partout}) \tag{2.5}$$

définit en effet une relation d'équivalence sur $\mathcal{L}^1(\Omega)$. Une classe d'équivalence $\dot{u}$ associée à une fonction u donnée est l'ensemble de toutes les fonctions v telles que $v = u$ presque partout, ce qui équivaut à dire que $\int_\Omega |u(x) - v(x)|\, \mathrm{d}x = 0$. Ainsi, la valeur de l'intégrale $\int_\Omega v(x)\, \mathrm{d}x$ ne dépend pas du représentant choisi dans une même classe, et cette valeur commune peut donc légitimement être notée $\int_\Omega \dot{u}(x)\, \mathrm{d}x$. Avec cette convention, la proposition 2.9 se récrit

$$\int_\Omega |\dot{u}(x)|\, \mathrm{d}x = 0 \iff \dot{u} = \dot{0},$$

ce qui montre que l'application $\dot{u} \mapsto \int_\Omega |\dot{u}(x)|\, \mathrm{d}x$ définit une norme sur l'espace quotient $L^1(\Omega)$.

Dans la suite, nous allons omettre le point dans les notations ci-dessus, ce qui revient à ne plus faire de distinction entre une fonction et sa classe d'équivalence. Ainsi, nous parlerons d'une "fonction" $u \in L^1(\Omega)$, étant entendu qu'il s'agit là d'un raccourci pour désigner un représentant d'une classe d'équivalence de $L^1(\Omega)$. Grâce à l'expression "presque partout", cette petite simplification de langage et de notation n'est en général source d'aucune ambiguïté. Elle nous permet surtout d'abandonner à partir de maintenant l'espace $\mathcal{L}^1(\Omega)$, qui était un passage obligé pour parvenir à $L^1(\Omega)$.

On peut résumer cette construction de $L^1(\Omega)$ sous la forme suivante.

Proposition 2.10 (i) *Pour un domaine $\Omega \subset \mathbb{R}^d$, l'espace quotient $L^1(\Omega)$ de $\mathcal{L}^1(\Omega)$ par la relation d'équivalence* (2.5) *est un espace vectoriel et l'application*

$$u \mapsto \|u\|_{L^1} = \int_\Omega |u(x)|\,\mathrm{d}x$$

définit une norme sur $L^1(\Omega)$, appelée norme de la convergence en moyenne.

(ii) *L'espace $\mathcal{D}^0(\Omega)$ des fonctions continues dont le support est un compact inclus dans l'intérieur de Ω est dense dans $L^1(\Omega)$.*

Nous admettrons le théorème suivant (dont la démonstration n'est pas évidente !) qui nous montre que nous avons atteint notre objectif initial.

Théorème 2.11 (Fischer–Riesz) *Pour un domaine $\Omega \subset \mathbb{R}^d$, l'espace $L^1(\Omega)$ est complet pour la norme de la convergence en moyenne.*

2.2 Critères d'intégrabilité

2.2.1 L'intégrale n'est pas une affaire de signe !

La proposition 2.8-(ii) nous montre que si u appartient à $L^1(\Omega)$, alors $|u|$ aussi. La réciproque est vraie [2], ce qui explique le titre de ce paragraphe. Plus généralement, pour vérifier qu'une fonction u appartient à $L^1(\Omega)$, il suffit de vérifier que $|u|$ est majoré par une fonction intégrable. En résumé, on a la

Proposition 2.12 (i) *Une fonction $u : \Omega \to \mathbb{R}$ appartient à $L^1(\Omega)$ si et seulement si*

$$\int_\Omega |u(x)|\,\mathrm{d}x < \infty.$$

(ii) *Si $|u| \le v$ presque partout dans Ω avec $v \in L^1(\Omega)$, alors $u \in L^1(\Omega)$.*

2. L'énoncé de la proposition 2.12 est correct à un détail près, lié à la notion de *mesurabilité* (voir la remarque 2.7). En toute rigueur, l'implication $|u| \in L^1(\Omega) \Longrightarrow u \in L^1(\Omega)$ ainsi que le point (ii) de la proposition 2.12 sont vrais à condition que u soit mesurable.

2.2.2 Les fonctions intégrables au sens de Riemann

Bien évidemment, si u est une fonction continue à support compact, elle est intégrable au sens de Lebesgue : dans la définition 2.5, il suffit de choisir la suite de Cauchy constante $u_n = u$. On peut aller plus loin. Les deux propositions suivantes nous montrent que non seulement l'intégrale de Lebesgue généralise celle de Riemann, mais aussi la notion d'intégrale de Riemann *absolument convergente* (qui fait partie des intégrales souvent appelées *généralisées* ou *impropres*, ces termes signifiant qu'on passe à la limite sur le domaine d'intégration). Ainsi, il n'est pas nécessaire de remettre en cause toutes les habitudes prises avec l'intégrale de Riemann : elles restent valables dans le cadre de l'intégrale de Lebesgue. Ces deux propositions sont énoncées ici en dimension 1, mais restent valables en dimensions supérieures.

Proposition 2.13 *Soient $]a,b[\subset \mathbb{R}$ un intervalle borné et $u :]a,b[\to \mathbb{R}$ une fonction bornée. Si u est intégrable au sens de Riemann sur $]a,b[$, alors elle est intégrable au sens de Lebesgue, et les deux intégrales coïncident.*

Proposition 2.14 *Soient $]a,b[\subset \mathbb{R}$ un intervalle borné ou non et $u :]a,b[\to \mathbb{R}$ une fonction localement intégrable au sens de Riemann, c'est-à-dire intégrable sur tout intervalle compact $[c,d] \subset]a,b[$. Alors, la limite*

$$\lim_{c\to a,\ d\to b} \int_c^d |u(x)|\,\mathrm{d}x \tag{2.6}$$

existe et est bornée si et seulement si u est intégrable au sens de Lebesgue, et dans ce cas, l'intégrale de u coïncide avec $\lim_{c\to a,\ d\to b} \int_c^d u(x)\,\mathrm{d}x$.

Exemple 2.15 Pour tout $x_0 > 0$, la fonction $u(x) = x^\alpha$ est intégrable au sens de Lebesgue sur l'intervalle $]0,x_0[$ si et seulement si $\alpha > -1$, et sur $]x_0,+\infty[$ si et seulement si $\alpha < -1$, et on a

$$\int_0^{x_0} x^\alpha\,\mathrm{d}x = \frac{x_0^{\alpha+1}}{\alpha+1} \text{ si } \alpha > -1 \quad \text{et} \quad \int_{x_0}^{+\infty} x^\alpha\,\mathrm{d}x = \frac{x_0^{\alpha+1}}{-\alpha-1} \text{ si } \alpha < -1.$$

•

Qu'en est-il des intégrales *semi-convergentes*, c'est-à-dire le cas de fonctions analogues à celles de la proposition 2.14 pour lesquelles $\lim_{c\to a,\ d\to b} \int_c^d u(x)\,\mathrm{d}x$ existe et est bornée, alors que (2.6) (soit la même limite, mais pour $|u|$) est infinie ? Un exemple classique est donné par la fonction $\sin x/x$ sur $]a,b[= \mathbb{R}$. L'existence de la limite est due aux compensations entre les contributions positives et négatives, compensations qui n'ont plus lieu pour l'intégrale de la valeur absolue. Ce type de fonction ne rentre pas dans la théorie de Lebesgue, précisément parce que l'intégrabilité au sens de Lebesgue n'est pas une affaire de signe (proposition 2.12). Il s'agit d'une limite d'intégrale, rien de plus !

2.2.3 Le théorème de convergence monotone

Dans ce paragraphe et le suivant, nous énonçons deux résultats importants qui permettent d'intervertir un passage à la limite et une intégrale. Le théorème ci-dessous constitue un des piliers de la théorie de l'intégrale de Lebesgue : il fournit un critère d'intégrabilité ou de non-intégrabilité d'une fonction définie comme limite d'une suite croissante de fonctions positives. Il est cependant moins utilisé en pratique que le *théorème de convergence dominée* qui en découle, et que nous présentons dans le paragraphe suivant.

Théorème 2.16 (de convergence monotone : Beppo–Levi) *Soient Ω un domaine de $\mathbb{R}^d$ et (u_n) une suite de fonctions de Ω à valeurs dans $[0, +\infty[$. Si cette suite est croissante, c'est-à-dire si $u_{n+1}(x) \geq u_n(x)$ pour tout $n \in \mathbb{N}$ et presque tout $x \in \Omega$, alors on a*

$$\int_\Omega \lim_{n\to+\infty} u_n(x)\,\mathrm{d}x = \lim_{n\to+\infty} \int_\Omega u_n(x)\,\mathrm{d}x,$$

que ces quantités soient finies ou infinies.

Notons que dans cet énoncé du théorème de convergence monotone, nous ne supposons pas que les fonctions u_n sont intégrables. La quantité $\int_\Omega u_n(x)\,\mathrm{d}x$ est donc autorisée à prendre la valeur $+\infty$, et dans ce cas, le théorème fournit un critère de non-intégrabilité.

2.2.4 Le théorème de convergence dominée

D'un point de vue pratique, on pourrait presque dire que le théorème suivant fournit la raison d'être de la théorie de l'intégrale de Lebesgue. A partir d'hypothèses très peu contraignantes (à comparer à celles de la proposition 2.1), il permet d'intervertir un passage à la limite et une intégrale.

Théorème 2.17 (de convergence dominée : Lebesgue) *Soient Ω un domaine de $\mathbb{R}^d$ et (u_n) une suite de fonctions de Ω dans $\mathbb{R}$. On suppose que*

(i) *(u_n) converge presque partout vers une fonction $u : \Omega \to \mathbb{R}$, soit*

$$\lim_{n\to\infty} u_n(x) = u(x) \quad \textit{pour presque tout } x \in \Omega;$$

(ii) *(u_n) est "dominée", au sens où il existe une fonction $v \in L^1(\Omega)$ telle que*

$$|u_n(x)| \leq v(x) \quad \textit{pour presque tout } x \in \Omega.$$

Alors u appartient à $L^1(\Omega)$ et

$$\int_\Omega u(x)\,\mathrm{d}x = \lim_{n\to+\infty} \int_\Omega u_n(x)\,\mathrm{d}x.$$

Remarquons que d'après la proposition 2.12-(ii), l'hypothèse (ii) entraîne que les fonctions u_n appartiennent à $L^1(\Omega)$ pour tout n.

Nous donnons ci-dessous deux corollaires très utiles de ce théorème, qui concernent le cas d'intégrales dépendant d'un paramètre. Par souci de simplicité, nous les énonçons ici pour un paramètre réel λ (ils se généralisent sans difficulté à des situations plus compliquées). Considérons donc une fonction $u : \Omega \times \Lambda \mapsto \mathbb{R}$ où Ω est un domaine de $\mathbb{R}^d$ et $\Lambda =]\lambda_1, \lambda_2[$ un intervalle de $\mathbb{R}$. Il s'agit d'étudier les propriétés de la fonction

$$\lambda \mapsto U(\lambda) = \int_\Omega u(x, \lambda)\, \mathrm{d}x.$$

Corollaire 2.18 (continuité sous le signe somme) *On suppose que*

(i) *pour presque tout $x \in \Omega$, la fonction $\lambda \mapsto u(x, \lambda)$ est continue sur Λ;*

(ii) *il existe une fonction $v \in L^1(\Omega)$ telle que $|u(x, \lambda)| \leq v(x)$ pour presque tout $x \in \Omega$ et pour tout $\lambda \in \Lambda$.*

Alors la fonction $\lambda \mapsto U(\lambda)$ est continue sur Λ.

Démonstration. Il suffit de montrer que pour toute suite (λ_n) qui converge vers un point $\lambda \in \Lambda$ quelconque mais fixé, on a $U(\lambda_n) \to U(\lambda)$, ce qui résulte directement du théorème de convergence dominée. □

Corollaire 2.19 (dérivabilité sous le signe somme) *On suppose que*

(i) *pour tout $\lambda \in \Lambda$, la fonction $x \mapsto u(x, \lambda)$ appartient à $L^1(\Omega)$;*

(ii) *pour presque tout $x \in \Omega$ et pour tout $\lambda \in \Lambda$, la dérivée partielle $\partial u/\partial\lambda(x, \lambda)$ existe ;*

(iii) *il existe une fonction $v \in L^1(\Omega)$ telle que $|\partial u/\partial\lambda(x, \lambda)| \leq v(x)$ pour presque tout $x \in \Omega$ et pour tout $\lambda \in \Lambda$.*

Alors la fonction $\lambda \mapsto U(\lambda)$ est dérivable sur Λ et

$$U'(\lambda) = \int_\Omega \frac{\partial u}{\partial \lambda}(x, \lambda)\, \mathrm{d}x.$$

Démonstration. L'hypothèse (i) assure que $U(\lambda)$ est bien défini pour tout $\lambda \in \Lambda$. Pour montrer que cette fonction est dérivable, remarquons que pour $\lambda \in \Lambda$ fixé et pour tout $\mu \in \mathbb{R}$ tel que $\lambda + \mu \in \Lambda$, on a

$$\frac{U(\lambda + \mu) - U(\lambda)}{\mu} = \int_\Omega w_\mu(x)\, \mathrm{d}x \quad \text{où} \quad w_\mu(x) = \frac{u(x, \lambda + \mu) - u(x, \lambda)}{\mu}.$$

D'après l'hypothèse (ii), on sait que lorsque μ tend vers 0, $w_\mu(x)$ tend vers $\partial u/\partial\lambda(x, \lambda)$ pour presque tout $x \in \Omega$ et

$$|w_\mu(x)| \leq \sup_{t \in [0,1]} \left| \frac{\partial u}{\partial \lambda}(x, \lambda + t\mu) \right|,$$

en vertu du théorème des accroissements finis. L'hypothèse (iii) nous montre que le second membre de cette inégalité est majoré par $v(x)$ (pour presque tout $x \in \Omega$). Et la conclusion découle alors du théorème de convergence dominée. □

Remarque 2.20 *L'énoncé du corollaire* 2.19 *soulève une petite ambiguïté qu'il est important de lever pour pouvoir utiliser correctement ce résultat de dérivation sous le signe somme. D'après la définition* 2.2*, l'expression "presque partout" dans les propriétés* (ii) *et* (iii) *signifie que ces deux propriétés doivent être satisfaites pour tout $x \in \Omega \setminus A$, où A un ensemble négligeable. Ces deux propriétés dépendant aussi du paramètre λ, il est fondamental que l'ensemble A n'en dépende pas. Il est en effet possible de construire un contre-exemple d'application de ce corollaire dans le cas où pour tout λ, les propriétés* (ii) *et* (iii) *sont satisfaites en dehors d'un seul point $x \in \Omega$ dépendant de λ.*

2.3 Intégrales multiples

2.3.1 Le théorème de Fubini

On s'intéresse dans ce paragraphe aux intégrales sur un domaine *produit* du type $X \times Y$ où X et Y sont respectivement des domaines de $\mathbb{R}^d$ et $\mathbb{R}^{d'}$. Considérons une fonction $u : X \times Y \to \mathbb{R}$, qui peut donc être vue comme une fonction de deux variables $u(x, y)$ où $x \in X$ et $y \in Y$. La question qu'on se pose ici est la suivante : comment peut-on exploiter la structure *produit* du domaine d'intégration pour affirmer l'existence et éventuellement effectuer le calcul de l'intégrale

$$\mathcal{I}_{X\times Y} = \int_{X\times Y} u(x, y)\,\mathrm{d}x\,\mathrm{d}y \ ?$$

L'idée naturelle consiste à utiliser des intégrales itérées, autrement dit à chercher à intégrer d'abord par rapport à l'une des deux variables, et ensuite, par rapport à l'autre. Selon l'ordre dans lequel on procède, on est donc amené à considérer

$$\mathcal{I}_{X(Y)} = \int_X \left(\int_Y u(x, y)\,\mathrm{d}y \right) \mathrm{d}x \quad \text{ou bien}$$
$$\mathcal{I}_{Y(X)} = \int_Y \left(\int_X u(x, y)\,\mathrm{d}x \right) \mathrm{d}y.$$

Quels sont les liens entre ces trois intégrales ?

Avant de répondre à cette question, il nous faut comprendre le sens de chacune d'elles. La première, $\mathcal{I}_{X\times Y}$, ne soulève aucune difficulté. Elle rentre dans le cadre de l'intégrabilité au sens de Lebesgue pour le domaine $\Omega = X \times Y$ (voir le paragraphe 2.1). Cette intégrale est définie si et seulement si u appartient à $L^1(X \times Y)$, autrement dit si

$$\int_{X\times Y} |u(x, y)|\,\mathrm{d}x\,\mathrm{d}y < \infty.$$

Les deux autres sont plus délicates à interpréter. Prenons le cas de $\mathcal{I}_{X(Y)}$ (le cas de $\mathcal{I}_{Y(X)}$ s'en déduit en permutant les rôles de X et Y). Dire que cette intégrale est finie signifie deux choses.

1) L'intégrale "intérieure" a un sens. Il s'agit d'une intégrale sur Y dépendant du paramètre $x \in X$, et il suffit que celle-ci soit définie pour presque tout x, ce qui s'écrit

$$\text{pour presque tout } x \in X, \text{ la fonction } y \mapsto u(x,y) \text{ appartient à } L^1(Y),$$

et on peut donc parler de $\int_Y u(x,y)\,\mathrm{d}y$ pour presque tout $x \in X$.

2) L'intégrale "extérieure" a un sens, ce qui signifie que

$$\text{la fonction } x \mapsto \int_Y u(x,y)\,\mathrm{d}y \text{ appartient à } L^1(X).$$

Le théorème suivant établit quels sont les liens entre les trois intégrales $\mathcal{I}_{X\times Y}$, $\mathcal{I}_{X(Y)}$ et $\mathcal{I}_{Y(X)}$.

Théorème 2.21 (de Fubini) *Soit u une fonction définie de $X \times Y$ dans $\mathbb{R}$.*

(i) *Si $u \geq 0$ presque partout, alors on a toujours $\mathcal{I}_{X\times Y} = \mathcal{I}_{X(Y)} = \mathcal{I}_{Y(X)}$, au sens où si l'une de ces intégrales est infinie, les autres le sont aussi, et si l'une d'elles est finie, les autres le sont aussi et prennent la même valeur.*

(ii) *Si $u \in L^1(X \times Y)$, les trois intégrales $\mathcal{I}_{X\times Y}$, $\mathcal{I}_{X(Y)}$ et $\mathcal{I}_{Y(X)}$ sont finies et égales.*

Ce théorème permet en pratique de permuter les deux intégrales sur X et Y. Comment l'utilise-t-on concrètement ? Si u est à valeurs positives, le premier volet (souvent appelé théorème de Tonelli) répond à la question : on peut calculer ou majorer l'une des intégrales, et le résultat restera valable pour les autres. Si u n'est pas de signe constant, il faut au préalable vérifier que $u \in L^1(X \times Y)$ pour pouvoir appliquer le second volet. Pour cela, il suffit d'appliquer le premier volet à $|u|$, puisque dire que $u \in L^1(X \times Y)$ revient à dire que $\int_{X\times Y} |u(x,y)|\,\mathrm{d}x\,\mathrm{d}y$ est fini. Ainsi dans le cas général, on procède en deux temps : on applique tout d'abord (i) à $|u|$, puis (ii) à u.

La permutation des deux intégrales sur X et Y est en général interdite dans la situation qui n'est pas traitée par le théorème de Fubini, c'est-à-dire pour une fonction de signe variable qui n'appartient pas à $L^1(X \times Y)$. L'intégrale $\mathcal{I}_{X\times Y}$ n'est alors pas définie, mais il peut arriver que les deux intégrales $\mathcal{I}_{X(Y)}$ et $\mathcal{I}_{Y(X)}$ soient finies, et prennent des valeurs différentes, comme le montre l'exemple suivant.

Exemple 2.22 Soit $u :]0,1[\times]0,1[\to \mathbb{R}$ définie par $u(x,y) = 1/y^2$ si $0 < x < y < 1$, et $u(x,y) = -1/x^2$ si $0 < y < x < 1$. On a

$$\begin{aligned}\int_0^1 \left(\int_0^1 u(x,y)\,\mathrm{d}y\right)\mathrm{d}x &= \int_0^1 \left(\int_0^x \frac{-\mathrm{d}y}{x^2} + \int_x^1 \frac{\mathrm{d}y}{y^2}\right)\mathrm{d}x \\ &= \int_0^1 \left(\frac{-1}{x} + \left(\frac{1}{x} - 1\right)\right)\mathrm{d}x \;=\; -1,\end{aligned}$$

tandis que

$$\begin{aligned}\int_0^1 \left(\int_0^1 u(x,y)\,\mathrm{d}x\right)\mathrm{d}y &= \int_0^1 \left(\int_0^y \frac{\mathrm{d}x}{y^2} + \int_y^1 \frac{-\mathrm{d}x}{x^2}\right)\mathrm{d}y \\ &= \int_0^1 \left(\frac{1}{y} + \left(1-\frac{1}{y}\right)\right)\mathrm{d}y \;=\; 1.\end{aligned}$$

Ceci montre que la fonction u n'appartient pas à $L^1(]0,1[^2)$. •

2.3.2 Produit de convolution

Nous présentons dans ce paragraphe une conséquence du théorème de Fubini qui concerne un outil clef de l'analyse que l'on retrouve dans toutes les branches de la physique. La *convolution* fournit en effet une représentation simple de la solution des équations aux dérivées partielles linéaires à coefficients constants. Ceci n'apparaîtra clairement que lorsque nous aurons compris dans le chapitre 3 comment la généraliser à des objets qui ne sont plus des fonctions (voir paragraphe 3.A). Nous commençons ici par étudier comment elle s'applique à des fonctions intégrables.

Théorème 2.23 *Soient u et v deux fonctions de $L^1(\mathbb{R}^d)$. Pour presque tout $x \in \mathbb{R}^d$, on peut définir*

$$(u \star v)(x) = \int_{\mathbb{R}^d} u(x-y)\,v(y)\,\mathrm{d}y = \int_{\mathbb{R}^d} u(y)\,v(x-y)\,\mathrm{d}y. \tag{2.7}$$

La fonction $u \star v$ ainsi définie est appelée le produit de convolution de u et v. Elle appartient à $L^1(\mathbb{R}^d)$ et

$$\|u \star v\|_{L^1} \le \|u\|_{L^1}\ \|v\|_{L^1}.$$

Démonstration. L'existence des intégrales dans (2.7) n'est pas évidente : à x fixé, il s'agit d'intégrer le produit de deux fonctions intégrables. Mais un tel produit n'est pas intégrable en général. Dans ce cas particulier, nous allons le déduire du théorème de Fubini. On a

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} |u(x-y)\,v(y)|\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}^d} |v(y)| \left(\int_{\mathbb{R}^d} |u(x-y)|\,\mathrm{d}x\right)\mathrm{d}y.$$

En effectuant le changement de variable $z = x - y$ (à y fixé), il vient

$$\int_{\mathbb{R}^d} |u(x-y)|\,\mathrm{d}x = \int_{\mathbb{R}^d} |u(z)|\,\mathrm{d}z = \|u\|_{L^1}.$$

Ainsi l'égalité précédente s'écrit

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} |u(x-y)\,v(y)|\,\mathrm{d}x\,\mathrm{d}y = \|u\|_{L^1} \int_{\mathbb{R}^d} |v(y)|\mathrm{d}y = \|u\|_{L^1}\ \|v\|_{L^1}.$$

Il s'ensuit que la fonction $(x,y) \mapsto u(x-y)\,v(y)$ appartient à $L^1(\mathbb{R}^d \times \mathbb{R}^d)$. Le théorème de Fubini nous permet donc de dire que pour presque tout x, la fonction $y \mapsto u(x-y)\,v(y)$ appartient à $L^1(\mathbb{R}^d)$, ce qui donne un sens à la première intégrale dans (2.7), la seconde s'en déduisant par le changement de variable $z = x - y$ (à x fixé). Il reste à constater que

$$\|u \star v\|_{L^1} = \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} u(x-y)\,v(y)\,\mathrm{d}y \right| \mathrm{d}x \leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |u(x-y)\,v(y)|\,\mathrm{d}x\,\mathrm{d}y,$$

où l'on vient de voir que la dernière intégrale n'est autre que $\|u\|_{L^1}\ \|v\|_{L^1}$. □

2.4 Quelques outils pratiques

2.4.1 Changement de variable

Comment peut-on réexprimer une intégrale de la forme $\int_\Omega u(x)\,\mathrm{d}x$ lorsque le domaine d'intégration Ω, ouvert de $\mathbb{R}^d$, apparaît comme l'image d'un autre ouvert $\tilde{\Omega} \subset \mathbb{R}^d$ par une transformation connue ? Nous noterons Φ cette transformation :

$$\begin{array}{rccl} \Phi : & \tilde{\Omega} & \to & \Omega \\ & \tilde{x} & \mapsto & x = \Phi(\tilde{x}) = (\Phi_1(\tilde{x}), \ldots, \Phi_d(\tilde{x})) \end{array}$$

et nous supposerons que Φ est un $\mathcal{C}^1$-*difféomorphisme*, c'est-à-dire une bijection qui est de classe $\mathcal{C}^1$ par rapport aux coordonnées spatiales et dont l'inverse Φ^{-1} est aussi de classe $\mathcal{C}^1$. La *matrice jacobienne* de cette transformation est la matrice $d \times d$ formée des dérivées partielles $\partial\Phi_i/\partial\tilde{x}_j\,(\tilde{x})$, pour $i,j = 1$ à d. Son déterminant est appelé le *jacobien* de Φ et sera noté $J_\Phi(\tilde{x})$, soit

$$J_\Phi(\tilde{x}) = \begin{vmatrix} \partial\Phi_1/\partial\tilde{x}_1\,(\tilde{x}) & \ldots & \partial\Phi_1/\partial\tilde{x}_d\,(\tilde{x}) \\ \vdots & \ddots & \vdots \\ \partial\Phi_d/\partial\tilde{x}_1\,(\tilde{x}) & \ldots & \partial\Phi_d/\partial\tilde{x}_d\,(\tilde{x}) \end{vmatrix}.$$

La proposition suivante est très utile en pratique : elle permet d'éviter de vérifier le caractère $\mathcal{C}^1$ de Φ^{-1}.

Proposition 2.24 *Soient Ω et $\tilde{\Omega}$ deux ouverts de $\mathbb{R}^d$ et $\Phi : \tilde{\Omega} \to \Omega$ une bijection de classe $\mathcal{C}^1$. Si $J_\Phi(\tilde{x}) \neq 0$ pour tout $\tilde{x} \in \tilde{\Omega}$, alors Φ est un $\mathcal{C}^1$-difféomorphisme.*

Sous l'hypothèse que notre transformation ϕ est bien un $\mathcal{C}^1$-difféomorphisme, le théorème suivant (que nous admettrons) fournit l'outil de passage de $\int_\Omega u(x)\,\mathrm{d}x$ à une intégrale sur $\tilde{\Omega}$. Il s'énonce sous une forme analogue au théorème de Fubini.

Théorème 2.25 (de changement de variable) *Soit u une fonction définie de Ω dans $\mathbb{R}$.*

(i) *Si $u \geq 0$ presque partout, alors on a toujours*

$$\int_\Omega u(x)\,\mathrm{d}x = \int_{\tilde\Omega} u(\Phi(\tilde x))\,|J_\Phi(\tilde x)|\,\mathrm{d}\tilde x, \tag{2.8}$$

au sens où si l'une de ces deux intégrales est infinie, l'autre l'est aussi, et si l'une d'elles est finie, l'autre l'est aussi et prend la même valeur.

(ii) *$u \in L^1(\Omega)$ si et seulement si la fonction $\tilde x \mapsto u(\Phi(\tilde x))\,|J_\Phi(\tilde x)|$ appartient à $L^1(\tilde\Omega)$. Dans ce cas, l'égalité* (2.8) *a lieu (les deux intégrales étant finies).*

Exemple 2.26 Un cas d'application important de ce théorème concerne le passage des coordonnées cartésiennes aux coordonnées polaires. La transformation Φ est dans ce cas donnée par

$$\Phi : \tilde x = (r,\theta) \mapsto x = (x_1,x_2) = (r\cos\theta, r\sin\theta).$$

Le déterminant de la matrice jacobienne est alors

$$J_\Phi(\tilde x) = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r.$$

Pour montrer que Φ est un $\mathcal{C}^1$-difféomorphisme, il nous suffit de vérifier que c'est une bijection de classe $\mathcal{C}^1$ entre deux ouverts (ici de $\mathbb{R}^2$) dont le jacobien ne s'annule pas. La nécessité de considérer des domaines *ouverts* nous amène à supposer que (r,θ) appartient à $]0,+\infty[\times]-\pi,+\pi[$, ce qui revient à exclure la demi-droite $\Gamma = \{(x_1,0);\ x_1 \leq 0\}$ dans le plan d'arrivée. Il est clair que Φ définit ainsi une bijection de classe $\mathcal{C}^1$ de $]0,+\infty[\times]-\pi,+\pi[$ dans $\mathbb{R}^2 \setminus \Gamma$ et que son jacobien ne s'annule jamais.

La coupure formée par la demi-droite Γ n'a aucune répercussion sur la valeur d'une intégrale sur $\mathbb{R}^2$ puisque c'est un *sous-ensemble négligeable* de $\mathbb{R}^2$. Pour montrer cette dernière affirmation, il suffit par exemple de constater que pour tout $\varepsilon > 0$, Γ est contenue dans $\bigcup_{n\in\mathbb{N}} R_n(\varepsilon)$ où $R_n(\varepsilon)$ désigne le rectangle

$$R_n(\varepsilon) = \,]-(n+1), -n+1/2[\,\times\,]-\varepsilon/2^n, +\varepsilon/2^n[,$$

dont la surface est $3\varepsilon/2^n$. Ainsi la surface totale de $\bigcup_{n\in\mathbb{N}} R_n(\varepsilon)$ est inférieure à 6ε.

Ainsi, avec le théorème 2.25, nous retrouvons la formule bien connue

$$\int_{\mathbb{R}^2} u(x)\,\mathrm{d}x = \int_0^{+\infty}\int_{-\pi}^{+\pi} u(r\cos\theta, r\sin\theta)\,r\,\mathrm{d}r\,\mathrm{d}\theta,$$

valable pour toute fonction u positive ou intégrable sur $\mathbb{R}^2$.

Terminons par une remarque concernant l'introduction de la coupure Γ qui peut paraître artificielle. Il peut sembler plus naturel de se contenter de ne retirer que l'origine, associée à $r = 0$, puisque c'est le seul point où le jacobien s'annule. La transformation Φ est en effet bijective de $]0,+\infty[\times]-\pi,+\pi]$ dans $\mathbb{R}^2 \setminus \{0\}$. Mais ce n'est pas un $\mathcal{C}^1$-difféomorphisme car son inverse est discontinu le long de la demi-droite $\Gamma = \{(x_1,0);\ x_1 \leq 0\}$ puisque l'argument θ d'un point est discontinu lorsqu'on la traverse. On comprend ici que la proposition 2.24 n'est pas vraie si l'on ne suppose pas que Ω et $\tilde\Omega$ sont deux *ouverts*. •

Exemple 2.27 L'exemple précédent se généralise à une dimension quelconque d'espace. Contentons-nous de rappeler le cas des coordonnées sphériques en dimension 3. On peut choisir Φ sous la forme

$$\Phi : \tilde{x} = (r, \theta, \phi) \mapsto x = (x_1, x_2, x_3) = (r\cos\theta\sin\phi, r\sin\theta\sin\phi, r\cos\phi)$$

et le déterminant de la matrice jacobienne est

$$J_\Phi(\tilde{x}) = \begin{vmatrix} \cos\theta\sin\phi & -r\sin\theta\sin\phi & r\cos\theta\cos\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\phi & 0 & -r\sin\phi \end{vmatrix} = -r^2\sin\phi.$$

En utilisant les mêmes arguments que dans l'exemple précédent, on vérifie sans difficulté que la transformation Φ est un $\mathcal{C}^1$-difféomorphisme de $]0, +\infty[\times]-\pi, +\pi[\times]0, \pi[$ dans $\mathbb{R}^3 \setminus \{(x_1, 0, x_3);\ x_1 \le 0\}$. La formule de changement de variable s'écrit donc

$$\int_{\mathbb{R}^3} u(x)\,\mathrm{d}x = \int_0^{+\infty}\int_{-\pi}^{+\pi}\int_0^{\pi} u(r\cos\theta\sin\phi, r\sin\theta\sin\phi, r\cos\phi)\, r^2\sin\phi\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi.$$

Elle est valable pour toute fonction u positive ou intégrable. •

Le théorème 2.25 nous fournit un outil très efficace pour vérifier l'intégrabilité d'une fonction. En particulier, lorsqu'une fonction devient singulière au voisinage d'un point, il peut être judicieux d'utiliser le changement de variable de l'un des exemples ci-dessus pour vérifier son intégrabilité.

Exemple 2.28 Pour $\alpha \in \mathbb{R}$, considérons la fonction

$$u(x) = |x|^\alpha \quad \text{pour } x \in \mathbb{R}^d.$$

Pour quelles valeurs de α cette fonction est-elle intégrable sur la boule unité $B_d = \{x \in \mathbb{R}^d; |x| < 1\}$? Les exemples ci-dessus nous montrent que pour $d = 2$ ou $d = 3$, on a

$$\int_{B_d} u(x)\,\mathrm{d}x = m_d \int_0^1 r^\alpha\, r^{d-1}\,\mathrm{d}r,$$

où m_d désigne la mesure de la frontière de la boule unité B_d (soit $m_2 = 2\pi$ et $m_3 = 4\pi$). Ainsi la fonction $|x|^\alpha$ est intégrable dans un voisinage de l'origine si et seulement si $\alpha + d - 1 > -1$, soit $\alpha > -d$. Ce résultat est valable en dimension quelconque. On voit que plus la dimension est grande, plus la singularité peut être forte.

De la même façon, la fonction $|x|^\alpha$ est intégrable dans $\mathbb{R}^d \setminus B_d$ si et seulement si $\alpha < -d$. Ainsi, plus la dimension augmente, plus la décroissance de la fonction à l'infini doit être forte. •

2.4.2 Formule de Stokes : un premier cas simple

La *formule de Stokes*, appelée aussi formule de Green-Ostrogradski ou théorème de la divergence, joue un rôle essentiel en analyse ainsi qu'en géométrie différentielle. Il s'agit de la généralisation aux dimensions supérieures à 1 de la formule fondamentale de l'intégration

$$\int_a^b u'(x)\,\mathrm{d}x = u(b) - u(a). \tag{2.9}$$

Pour l'énoncer, il faut tout d'abord introduire la notion de *normale extérieure* à la frontière d'un ouvert $\Omega \subset \mathbb{R}^d$ ainsi que la notion d'*intégrale de surface* sur cette frontière. D'un point de vue intuitif, ces deux notions semblent très simples. Mais leur définition rigoureuse est en réalité assez délicate et s'appuie sur un cas simple que nous allons traiter dans ce paragraphe. Nous évoquerons le cas général dans le paragraphe suivant, sans rentrer dans les détails techniques.

Commençons par rappeler que si Ω est un ouvert de $\mathbb{R}^d$, sa frontière, que nous noterons $\partial\Omega$, est définie par

$$\partial\Omega = \overline{\Omega} \setminus \Omega \quad \text{où } \overline{\Omega} \text{ désigne l'adhérence de } \Omega.$$

Nous allons supposer dans ce paragraphe que $\partial\Omega$ peut être représentée par le graphe d'une fonction $F : \mathbb{R}^{d-1} \to \mathbb{R}$ (voir la figure 2.3). Dans un repère orthonormé donné, on note $x = (\underline{x}, x_d)$ les coordonnées d'un point courant de $\mathbb{R}^d$, où $\underline{x} = (x_1, \ldots, x_{d-1})$ représente les $d-1$ premières coordonnées. Considérons le domaine

$$\Omega = \{(\underline{x}, x_d) \in \mathbb{R}^d;\ x_d > F(\underline{x})\} \tag{2.10}$$

dont la frontière est simplement

$$\partial\Omega = \{(\underline{x}, x_d) \in \mathbb{R}^d;\ x_d = F(\underline{x})\}.$$

Nous supposerons que $F \in \mathcal{C}^1(\mathbb{R}^{d-1})$ (voir l'exemple 1.2). Ainsi, pour deux points voisins $\underline{x}$ et $\underline{y}$ de $\mathbb{R}^{d-1}$, le développement limité à l'ordre 1 de F s'écrit

$$F(\underline{y}) - F(\underline{x}) = (\underline{y} - \underline{x}) \cdot \underline{\nabla} F(\underline{x}) + o(|\underline{y} - \underline{x}|),$$

où on a noté $\underline{\nabla} F = (\partial F/\partial x_1, \ldots, \partial F/\partial x_{d-1})$ le gradient de F par rapport aux coordonnées $\underline{x}$. Ce développement montre que le plan [3] tangent à la frontière $\partial\Omega$ au point $(\underline{x}, F(\underline{x}))$ est l'ensemble des points $(\underline{y}, y_d) \in \mathbb{R}^d$ qui vérifient l'équation

$$y_d - F(\underline{x}) = (\underline{y} - \underline{x}) \cdot \underline{\nabla} F(\underline{x}).$$

La droite normale à ce plan est engendrée par exemple par le vecteur $(\underline{\nabla} F(\underline{x}), -1)$, qui est orienté vers l'extérieur de Ω. Ainsi, la *normale unitaire extérieure* en un point $x = (\underline{x}, F(\underline{x})) \in \partial\Omega$ est le vecteur $n = n(x) \in \mathbb{R}^d$ défini par

$$n(x) = \frac{1}{\sqrt{1 + |\underline{\nabla} F(\underline{x})|^2}}\ (\underline{\nabla} F(\underline{x}), -1)\,. \tag{2.11}$$

3. On adopte ici la terminologie de la dimension 3. En dimension quelconque, il faudrait plutôt parler d'un *l'hyperplan*.

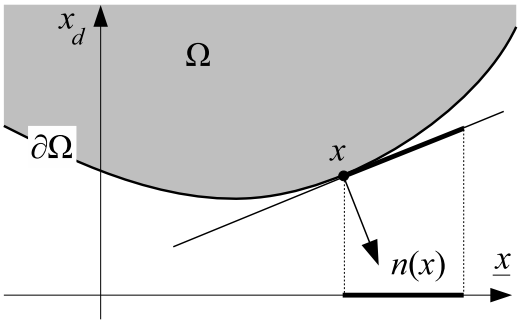


FIGURE 2.3 – Cas d'un domaine Ω dont la frontière est définie par une fonction

Comme nous avons supposé que F est de classe $\mathcal{C}^1$, il est clair que $n(x)$ est une fonction continue sur $\partial\Omega$.

On définit par ailleurs l'*intégrale de surface* sur $\partial\Omega$ par

$$\int_{\partial\Omega} u(x)\,\mathrm{d}\sigma(x) = \int_{\mathbb{R}^{d-1}} u(\underline{x}, F(\underline{x}))\,\sqrt{1+|\underline{\nabla} F(\underline{x})|^2}\,\mathrm{d}\underline{x}, \tag{2.12}$$

par exemple pour une fonction u continue et à support compact. Cette définition se comprend aisément dès lors que l'on admet que pour mesurer des éléments de surface infiniment petits au voisinage d'un point $x \in \partial\Omega$ donné, la frontière peut être confondue localement avec son plan tangent (ce qui est justifié par l'hypothèse $F \in \mathcal{C}^1(\mathbb{R}^{d-1})$). La quantité $\sqrt{1+|\underline{\nabla} F(\underline{x})|^2}$ représente le rapport entre la mesure d'un élément de surface sur ce plan tangent par la mesure de sa projection orthogonale sur le plan d'équation $x_d = 0$. La figure 2.3 illustre le cas de la dimension $d = 2$, où ce rapport est simplement la norme du vecteur $(1, F'(\underline{x}))$. Dans le cas de la dimension 3, ce rapport est donné par la surface du parallélogramme qui s'appuie sur les deux vecteurs $(1, 0, \partial F/\partial x_1(\underline{x}))$ et $(0, 1, \partial F/\partial x_2(\underline{x}))$, c'est-à-dire la norme du produit vectoriel de ces deux vecteurs. Ceci se généralise à une dimension quelconque.

Proposition 2.29 (Formule de Stokes – cas simple) *Soient Ω le domaine de $\mathbb{R}^d$ défini par* (2.10) *et $V : \overline{\Omega} \mapsto \mathbb{R}^d$ une application à valeurs vectorielles (ou "champ de vecteurs") : $V(x) = (V_1(x), \ldots, V_d(x)) \in \mathbb{R}^d$ pour tout $x \in \Omega$. On suppose que chaque composante V_i, pour $i = 1$ à d, appartient à $\mathcal{C}^1(\overline{\Omega})$ (voir l'exemple 1.2) et a un support compact. On a alors*

$$\int_\Omega \nabla \cdot V\,\mathrm{d}x = \int_{\partial\Omega} V \cdot n\,\mathrm{d}\sigma,$$

où on a noté[4] $\nabla \cdot V = \partial V_1/\partial x_1 + \ldots + \partial V_d/\partial x_d$ *la divergence de V.*

Remarque 2.30 *Cette proposition montre que si Ω contient les supports des composantes V_i de V (autrement dit, si V s'annule dans un voisinage de $\partial\Omega$), alors*

$$\int_\Omega \nabla \cdot V\,\mathrm{d}x = 0.$$

4. Au lieu de $\nabla \cdot V$, la divergence d'un champ de vecteurs est souvent notée $\mathrm{div}\, V$.

Et ceci reste évidemment vrai si on remplace Ω par n'importe quel ouvert $\Omega' \subset \mathbb{R}^d$ contenant les supports des V_i.

DÉMONSTRATION. Commençons par remarquer que les supports des V_i étant supposés compacts, toutes les intégrales que nous allons manipuler portent toujours sur des domaines d'intégration bornés, bien que les domaines indiqués pour les écrire soient non bornés. Nous allons calculer séparément chacune des intégrales $\int_\Omega \partial V_i/\partial x_i \, \mathrm{d}x$ pour $i = 1, \ldots, d$.

Dans le cas où $i = d$, commençons par remarquer que notre choix particulier de domaine Ω nous permet de séparer les intégrations selon $\underline{x}$ et x_d en écrivant

$$\int_\Omega u(x) \, \mathrm{d}x = \int_{\mathbb{R}^{d-1}} \left(\int_{F(\underline{x})}^{+\infty} u(\underline{x}, x_d) \, \mathrm{d}x_d \right) \mathrm{d}\underline{x},$$

pour toute fonction $u \in L^1(\Omega)$. Cette formule découle du théorème 2.21 de Fubini. Il suffit en effet d'introduire la fonction $\tilde{u} : \mathbb{R}^d \to \mathbb{R}$ définie par $\tilde{u}(x) = u(x)$ si $x \in \Omega$ et $\tilde{u}(x) = 0$ sinon. Ceci permet de nous ramener à une intégrale sur un domaine produit : $\int_\Omega u(x) \, \mathrm{d}x = \int_{\mathbb{R}^{d-1}\times\mathbb{R}} \tilde{u}(\underline{x}, x_d) \, \mathrm{d}\underline{x} \, \mathrm{d}x_d$. En posant $u = \partial V_d/\partial x_d$, on peut utiliser la formule fondamentale (2.9) dans l'intégrale intérieure, ce qui montre que

$$\int_\Omega \frac{\partial V_d}{\partial x_d}(x) \, \mathrm{d}x = - \int_{\mathbb{R}^{d-1}} V_d(\underline{x}, F(\underline{x})) \, \mathrm{d}\underline{x}. \tag{2.13}$$

Considérons maintenant le cas où $i = 1, \ldots, d-1$. On effectue tout d'abord le changement de variable associé à la transformation

$$\begin{array}{rccl} \Phi : & \mathbb{R}^{d-1}\times]0, +\infty[& \mapsto & \Omega \\ & (\underline{x}, t) & \mapsto & x = (\underline{x}, F(\underline{x}) + t) \end{array}$$

dont le jacobien est égal à 1. Le théorème 2.25 nous montre que

$$\int_\Omega \frac{\partial V_i}{\partial x_i}(x) \, \mathrm{d}x = \int_{\mathbb{R}^{d-1}\times]0,+\infty[} \frac{\partial V_i}{\partial x_i}(\Phi(\underline{x}, t)) \, \mathrm{d}\underline{x}\mathrm{d}t.$$

Définissons alors la fonction $\tilde{V}_i$ par

$$\tilde{V}_i(\underline{x}, t) = V_i(\Phi(\underline{x}, t)) = V_i(\underline{x}, F(\underline{x}) + t),$$

qui est clairement de classe $\mathcal{C}^1$ et à support compact. D'après la formule de dérivation d'une fonction composée, on a

$$\frac{\partial \tilde{V}_i}{\partial x_i}(\underline{x}, t) = \frac{\partial V_i}{\partial x_i}(\Phi(\underline{x}, t)) + \frac{\partial V_i}{\partial x_d}(\Phi(\underline{x}, t)) \, \frac{\partial F}{\partial x_i}(\underline{x}) \quad \text{et} \quad \frac{\partial \tilde{V}_i}{\partial t}(\underline{x}, t) = \frac{\partial V_i}{\partial x_d}(\Phi(\underline{x}, t)).$$

On en déduit que

$$\int_\Omega \frac{\partial V_i}{\partial x_i}(x) \, \mathrm{d}x = \int_{\mathbb{R}^{d-1}\times]0,+\infty[} \frac{\partial \tilde{V}_i}{\partial x_i}(\underline{x}, t) \, \mathrm{d}\underline{x}\mathrm{d}t - \int_{\mathbb{R}^{d-1}\times]0,+\infty[} \frac{\partial \tilde{V}_i}{\partial t}(\underline{x}, t) \, \frac{\partial F}{\partial x_i}(\underline{x}) \, \mathrm{d}\underline{x}\mathrm{d}t.$$

La première intégrale du second membre est nulle. En effet, en intégrant d'abord selon x_i (grâce au théorème de Fubini), l'intégrale intérieure s'annule puisque V_i est à support compact. Pour la seconde, on intègre d'abord selon t, ce qui conduit à

$$\int_\Omega \frac{\partial V_i}{\partial x_i}(x)\,\mathrm{d}x = \int_{\mathbb{R}^{d-1}} \tilde{V}_i(\underline{x},0)\,\frac{\partial F}{\partial x_i}(\underline{x})\,\mathrm{d}\underline{x} = \int_{\mathbb{R}^{d-1}} V_i(\underline{x},F(\underline{x}))\,\frac{\partial F}{\partial x_i}(\underline{x})\,\mathrm{d}\underline{x}. \tag{2.14}$$

Il reste juste à recoller les morceaux en sommant pour $i = 1$ à d les intégrales données par les formules (2.13) et (2.14), ce qui donne le résultat annoncé, compte tenu des définitions (2.11) et (2.12). □

2.4.3 Formule de Stokes : cas général

Lorsqu'on considère un ouvert $\Omega \subset \mathbb{R}^d$ quelconque, il n'est en général pas possible de représenter *globalement* sa frontière par le graphe d'une fonction, mais il peut être possible de le faire *localement*, c'est-à-dire au voisinage d'un point donné de $\partial\Omega$, en choisissant un système de coordonnées adapté. C'est ce qui motive la définition suivante, qui caractérise les ouverts pour lesquels on peut se ramener au cas du paragraphe précédent au voisinage de tout point de la frontière. Nous nous limitons ici au cas d'ouverts de classe $\mathcal{C}^1$ (qui correspond à l'hypothèse $F \in \mathcal{C}^1(\mathbb{R}^{d-1})$ du paragraphe précédent), mais on pourrait de la même façon définir les ouverts de classe $\mathcal{C}^m$ pour tout $m > 1$.

Définition 2.31 *On dit qu'un ouvert $\Omega \subset \mathbb{R}^d$ est de classe $\mathcal{C}^1$ si pour tout point $x^* \in \partial\Omega$, on peut trouver*
• *un système de coordonnées orthonormales $(x_1,\dots,x_{d-1},x_d) = (\underline{x},x_d)$ (dans lequel on notera $x^* = (\underline{x}^*,x_d^*)$),*
• *un ouvert $\mathcal{O} \subset \mathbb{R}^{d-1}$ contenant $\underline{x}^*$ et un intervalle ouvert $J \subset \mathbb{R}$ contenant x_d^*, qui définissent un domaine $\omega \subset \mathbb{R}^d$ représenté par $\mathcal{O}\times J$ dans les coordonnées $(\underline{x},x_d)$,*
• *une application $F : \mathcal{O} \mapsto J$ de classe $\mathcal{C}^1$,*
tels que

$$\Omega \cap \omega = \{(\underline{x},x_d) \in \mathcal{O}\times J;\ x_d > F(\underline{x})\}.$$

Les données associées à chaque point $x^* \in \partial\Omega$ (soit : les coordonnées $(\underline{x},x_d)$, les ouverts $\mathcal{O}$ et J ainsi que l'application F) constituent ce qu'on appelle une *carte locale* de la frontière. La figure 2.4 illustre cette notion. La définition ci-dessus n'est pas très économique puisqu'elle demande autant de cartes que de points sur la frontière ! En pratique, si la frontière $\partial\Omega$ est bornée, un nombre fini de cartes (appelé un *atlas*) est suffisant pour la décrire. En effet, comme $\partial\Omega$ est un ensemble compact contenu dans l'union des ouverts ω lorsque x^* parcourt $\partial\Omega$, on peut en extraire une union finie qui recouvre $\partial\Omega$.

La définition 2.31 nous permet de définir la *normale unitaire extérieure* en tout point $x^* \in \partial\Omega$: il suffit d'appliquer la définition (2.11) dans le système de coordonnées

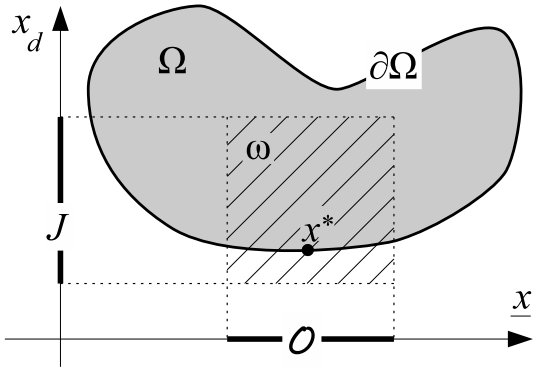


FIGURE 2.4 – Carte locale en un point $x^* \in \partial\Omega$

associé à x^*. Ceci soulève cependant une question importante. En effet, si l'on reste dans le système de coordonnées associé à x^*, on peut aussi appliquer la définition (2.11) à tout point $x \in \partial\Omega \cap \omega$. La normale ainsi obtenue est-elle la même que si on se plaçait dans le système de coordonnées associé à ce point x particulier, où à tout autre point dont la carte contient x ? La réponse est évidemment positive : cette définition est intrinsèque, au sens où la normale ne dépend que de l'ouvert Ω et non du système de coordonnées choisi pour l'exprimer. La démonstration de ce résultat n'est pas très compliquée mais elle est un peu technique et ne présente pas beaucoup d'intérêt pour notre propos. Nous l'admettrons donc.

De façon similaire, la notion d'*intégrale de surface* $\int_{\partial\Omega} u(x)\,\mathrm{d}\sigma(x)$ va découler de la définition (2.12). En quelque sorte, il suffit d'intégrer carte par carte et de sommer les résultats. Pour ce faire, nous admettrons une propriété très utile qui permet de décomposer une fonction en une somme de fonctions dont les supports sont contenus dans des ouverts fixés.

Théorème 2.32 (Partition de l'unité) *Soient $\{\omega_p;\ p = 1, P\}$ une famille d'ouverts de $\mathbb{R}^d$ et $K \subset \bigcup_{p=1,P} \omega_p$ un compact. Il existe des fonctions $\psi_p \in \mathcal{C}^\infty(\mathbb{R}^d)$ à valeurs dans $[0,1]$, pour $p = 1$ à P, telles que le support de chaque ψ_p est contenu dans ω_p et*

$$\sum_{p=1}^{P} \psi_p = 1 \text{ dans un voisinage de } K.$$

Considérons une fonction $u : \partial\Omega \mapsto \mathbb{R}$ continue et à support compact. Comme nous l'avons vu plus haut, on peut extraire une famille finie de cartes qui recouvre $\mathrm{Supp}(u)$. Notons $\{\omega_p;\ i = 1, P\}$ la famille d'ouverts ainsi construite et $\{\psi_p;\ p = 1, P\}$ une partition de l'unité associée. On peut ainsi poser

$$\int_{\partial\Omega} u(x)\,\mathrm{d}\sigma(x) = \sum_{p=1}^{P} \int_{\partial\Omega\cap\omega_p} u(x)\,\psi_p(x)\,\mathrm{d}\sigma(x),$$

où chacune des intégrales peut être calculée par la formule (2.12). Ici encore se pose la question de l'indépendance du résultat par rapport à la famille de cartes utilisées. Ici encore, nous admettrons le caractère intrinsèque de cette définition.

Nous sommes maintenant en mesure d'énoncer la forme générale de la proposition 2.29.

Théorème 2.33 (Formule de Stokes) *Soient Ω un ouvert borné de $\mathbb{R}^d$ de classe $\mathcal{C}^1$ et $V : \overline{\Omega} \mapsto \mathbb{R}^d$ un champ de vecteurs dont chaque composante V_i appartient à $\mathcal{C}^1(\overline{\Omega})$ (voir l'exemple 1.2). On a alors*

$$\int_\Omega \nabla \cdot V \, \mathrm{d}x = \int_{\partial\Omega} V \cdot n \, \mathrm{d}\sigma,$$

où $n = n(x)$ désigne la normale unitaire extérieure à $\partial\Omega$ et $\nabla \cdot V = \partial V_1/\partial x_1 + \ldots + \partial V_d/\partial x_d$ la divergence de V.

Démonstration. Le domaine $\overline{\Omega}$ étant fermé et borné, donc compact, il peut être recouvert par une union finie d'ouverts ω_p, pour $p = 1$ à P, associés aux cartes locales de la définition 2.31, complétés par un ouvert borné $\omega_0 \subset \Omega$. En considérant une partition de l'unité associée $\{\psi_p;\ p = 0, P\}$, il suffit de vérifier la formule de Stokes pour chacun des champs $\psi_p V$. Lorsque $p > 0$, on se place dans le système de coordonnées de la carte locale correspondante et on applique simplement la proposition 2.29. Et pour $p = 0$, c'est une conséquence de la remarque 2.30. □

Remarque 2.34 *La formule de Stokes est énoncée ici pour un domaine borné, mais elle reste évidemment valable pour un domaine non borné si l'on suppose que chacune des composantes V_i a un support compact.*

2.4.4 Formules de Green

Nous donnons ici quelques conséquences très utiles de la formule de Stokes, que nous regroupons sous la terminologie de *formules de Green*. Il s'agit simplement de généralisations de la formule d'*intégration par parties*.

Corollaire 2.35 *Soient Ω un ouvert borné de $\mathbb{R}^d$ de classe $\mathcal{C}^1$ et u, v deux fonctions de $\mathcal{C}^1(\overline{\Omega})$. On a*

$$\int_\Omega u \, \frac{\partial v}{\partial x_i} \, \mathrm{d}x = -\int_\Omega \frac{\partial u}{\partial x_i} \, v \, \mathrm{d}x + \int_{\partial\Omega} u \, v \, n \cdot e_i \, \mathrm{d}\sigma \quad \textit{pour } i = 1, \ldots, d,$$

où n est la normale unitaire extérieure à $\partial\Omega$ et $(e_1, \ldots, e_d)$ la base canonique de $\mathbb{R}^d$.

Démonstration. Il suffit d'appliquer la formule de Stokes (théorème 2.33) au champ de vecteurs $V = u\, v\, e_i$. □

Corollaire 2.36 *Soient Ω un ouvert borné de $\mathbb{R}^d$ de classe $\mathcal{C}^1$, $u \in \mathcal{C}^2(\overline{\Omega})$ et $v \in \mathcal{C}^1(\overline{\Omega})$. On a*

$$\int_\Omega (\Delta u)\, v \, \mathrm{d}x = -\int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x + \int_{\partial\Omega} \frac{\partial u}{\partial n}\, v \, \mathrm{d}\sigma,$$

où $\nabla u = (\partial u/\partial x_1, \ldots, \partial u/\partial x_d)$ désigne le gradient de u, $\Delta u = \partial^2 u/\partial x_1^2 + \ldots + \partial^2 u/\partial x_d^2$ est le laplacien de u, et $\partial u/\partial n = \nabla u \cdot n$ est la dérivée normale de u.

DÉMONSTRATION. On applique la formule de Stokes au champ $V = (\nabla u)\, v$. □

Corollaire 2.37 *Soient Ω un ouvert borné de $\mathbb{R}^d$ de classe $\mathcal{C}^1$ et u, v deux fonctions de $\mathcal{C}^2(\overline{\Omega})$. On a*

$$\int_\Omega \{(\Delta u)\, v - u\, (\Delta v)\} \, \mathrm{d}x = \int_{\partial\Omega} \left(\frac{\partial u}{\partial n}\, v - u\, \frac{\partial v}{\partial n} \right) \mathrm{d}\sigma,$$

avec les mêmes notations que dans le corollaire 2.36.

DÉMONSTRATION. On applique le corollaire 2.36 aux couples (u, v) et (v, u), puis on effectue la différence. □

Chapitre 3

Les distributions

La théorie des *distributions* a été introduite par Laurent Schwartz au milieu du 20e siècle. Cette théorie a fait progresser considérablement l'étude des équations aux dérivées partielles, et est devenue aujourd'hui un outil incontournable de l'analyse. La notion de distribution généralise celle de *fonction*. Cette généralisation a fourni aux physiciens un cadre rigoureux pour manipuler des objets dont ils se servaient depuis longtemps sans savoir en donner une définition précise. L'exemple le plus évident est celui de la *mesure de Dirac* dont nous avons parlé au paragraphe 2.1.2 en passant à la limite sur la suite de fonctions "chapeaux" de l'exemple 1.4. L'objet résultant de ce passage à la limite représente la densité de masse associée à une masse ponctuelle. Un physicien pourrait décrire cet objet, noté δ (pour une masse placée à l'origine), en disant qu'il s'agit d'une fonction généralisée qui vaut 0 partout sauf à l'origine où elle vaut l'infini, mais pas n'importe quel infini, car celui-ci doit être choisi de sorte que "l'intégrale" de cet objet vaille 1 ! En résumé, on pourrait écrire

$$\text{“ } \delta(x) = 0 \text{ pour tout } x \neq 0 \quad \text{et} \quad \int \delta(x)\,\mathrm{d}x = 1 \text{ ”},$$

ce qui en toute rigueur n'a pas de sens puisque δ n'est pas une fonction et il n'est donc pas possible de parler de son intégrale. Cette description incorrecte de δ permet cependant d'en donner une vision intuitive. Et si nous nous interdisons dans la suite d'écrire ce genre de chose, rien n'empêche de le penser ! C'est souvent très utile pour guider un raisonnement où l'on manipule une masse de Dirac.

Cet exemple nous donne en fait une première façon d'interpréter les distributions. Comme la mesure de Dirac, toute distribution peut être vue comme la limite d'une suite de fonctions. Ce passage à la limite a lieu en un sens très faible et permet ainsi d'atteindre des objets très variés. On pourrait dire que toute la richesse des distributions réside dans le fait que l'espace qui contient ces objets est tellement vaste que tout (ou presque) y devient permis ! Plus sérieusement, de nombreuses opérations qui requièrent habituellement des hypothèses restrictives deviennent licites pour toutes les distributions sans aucune hypothèse. L'opération qui nous intéresse le plus ici est la *dérivation* : les distributions possèdent l'étonnante propriété d'être indéfiniment

dérivables ! On peut ainsi donner un sens à de nombreuses équations aux dérivées partielles issues de la physique sans faire d'hypothèse sur leurs solutions. La théorie des distributions fournit alors des outils de calcul très puissants pour les résoudre. Ce chapitre est destiné à donner un premier aperçu sur ces outils. Pour plus de détails, on pourra par exemple consulter [4] ou [1].

3.1 Définition et premiers exemples

3.1.1 De la physique aux mathématiques

L'idée fondamentale de la définition des distributions peut être expliquée par une analogie avec un processus expérimental. Considérons un système physique occupant une région $\Omega \subset \mathbb{R}^d$ dont l'état peut être décrit par une *fonction* $u : \Omega \to \mathbb{R}$. On peut par exemple penser à un champ de température, de pression, de densité de masse... Ce mode de représentation d'un état par une fonction est très commode pour le physicien qui cherche à modéliser un phénomène par des équations. Mais pour l'expérimentateur, une telle fonction est généralement inaccessible à l'expérience. Du fait de son étendue dans l'espace, un instrument de mesure n'a pas accès aux valeurs ponctuelles d'une grandeur physique, mais plutôt à une moyenne pondérée de ces valeurs dans une région de l'espace, soit une quantité du type

$$\int_\Omega u(x)\,\varphi(x)\,\mathrm{d}x,$$

où φ représente la fonction de pondération, localisée dans une petite région de l'espace. Déplacer l'instrument de mesure ou modifier les réglages de celui-ci revient alors à changer de fonction de pondération φ, que nous appellerons dans la suite *fonction test* (ou fonction d'essai), dans la mesure où chacune d'elle permet de tester le système, c'est-à-dire d'obtenir une information sur celui-ci à partir d'une expérience donnée. L'ensemble de ces fonctions caractérise ainsi le champ des expériences possibles.

En réalité, l'expérimentateur n'a pas besoin de savoir que l'état du système qu'il observe peut être décrit par une fonction u, qui n'est en fait qu'une pure création mathématique. Oublions donc un instant cette fonction, et appelons simplement T notre système dans l'état dans lequel il est observé. Pour ce système particulier, l'expérimentateur peut choisir une condition expérimentale, soit une fonction test φ, et il obtient le résultat de la mesure, que nous supposerons réel. Nous noterons cette mesure $\langle T, \varphi\rangle$: c'est la réponse du système T à la sollicitation φ. Effectuer une batterie de tests revient ainsi avoir accès à l'application

$$\varphi \mapsto \langle T, \varphi\rangle \in \mathbb{R}.$$

En quelque sorte, la théorie des distributions repose sur l'idée que cette application est le "bon" objet pour représenter l'état du système, plutôt que la fonction u. Plus

précisément, cette représentation va permettre de décrire non seulement des états du système qui peuvent être représentés par une fonction, autrement dit tels que

$$\langle T, \varphi \rangle = \int_\Omega u(x)\, \varphi(x)\, \mathrm{d}x \quad \text{pour tout } \varphi,$$

mais aussi toute une famille d'états pour lesquels aucune fonction ne permet d'obtenir une telle relation. C'est en ce sens que les distributions généralisent la notion de fonction.

Nous allons supposer que l'ensemble des fonctions tests est un espace vectoriel, noté $\mathcal{D}$, et que le processus expérimental que nous venons de décrire est *linéaire* et *continu*. La linéarité traduit un principe de superposition du processus, soit

$$\langle T, \lambda\varphi + \mu\psi \rangle = \lambda\, \langle T, \varphi \rangle + \mu\, \langle T, \psi \rangle \quad \forall \varphi, \psi \in \mathcal{D},\ \forall \lambda, \mu \in \mathbb{R},$$

et la continuité, une propriété de stabilité : deux conditions expérimentales voisines produisent des résultats voisins. Ces hypothèses reviennent à dire que T est une *forme linéaire continue sur* $\mathcal{D}$. L'ensemble des formes linéaires continues sur $\mathcal{D}$ est appelé le *dual* (topologique) de $\mathcal{D}$, et il est noté $\mathcal{D}'$. Dans l'analogie ci-dessus, $\mathcal{D}$ est significatif de l'appareil de mesure (il décrit l'ensemble des conditions expérimentales possibles), alors que $\mathcal{D}'$ représente l'ensemble des états qu'il est possible d'observer avec cet appareil. Plus l'appareil est précis, plus $\mathcal{D}$ est petit (et est muni d'une topologie plus fine), et plus son dual $\mathcal{D}'$ est grand [1]. En quelque sorte, l'appareil de mesure proposé par Laurent Schwartz est infiniment précis. Il permet d'observer des états qui ne sont plus des fonctions, parmi lesquels figure la mesure de Dirac évoquée plus haut, mais aussi bien d'autres objets exotiques !

3.1.2 L'espace des distributions

Commençons par définir notre espace de fonctions tests, noté [2] $\mathcal{D}(\Omega)$.

Définition 3.1 *Soit Ω un ouvert de $\mathbb{R}^d$, borné ou non.*

(i) *On note $\mathcal{D}(\Omega)$ l'espace des fonctions de Ω dans $\mathbb{R}$ indéfiniment dérivables et dont le support est un compact inclus dans Ω (voir* (2.2) *pour la définition du support d'une fonction).*

(ii) *Pour un compact donné $K \subset \Omega$, on note $\mathcal{D}_K(\Omega)$ le sous-espace de $\mathcal{D}(\Omega)$ constitué des fonctions dont le support est inclus dans K (autrement dit qui s'annulent à l'extérieur de K).*

1. Il est clair que si $\mathcal{D}_1 \subset \mathcal{D}_2$ la restriction à $\mathcal{D}_1$ d'une forme linéaire $\varphi_2 \mapsto \langle T_2, \varphi_2 \rangle$ sur $\mathcal{D}_2$ est une forme linéaire sur $\mathcal{D}_1$. Et on peut comprendre (mais nous ne nous attarderons pas sur ce fait) que si la topologie de $\mathcal{D}_1$ est plus fine que celle de $\mathcal{D}_2$, la continuité de T_2 sur $\mathcal{D}_2$ entraîne celle de $T_2|_{\mathcal{D}_1}$ sur $\mathcal{D}_1$. En résumé : $\mathcal{D}_1 \subset \mathcal{D}_2 \Longrightarrow \mathcal{D}_2' \subset \mathcal{D}_1'$.

2. $\mathcal{D}(\Omega)$ est parfois noté $\mathcal{C}_0^\infty(\Omega)$ dans la littérature.

Rappelons que comme dans les chapitres précédents, on pourrait aussi bien considérer des fonctions à valeurs dans $\mathbb{C}$: tous les résultats qui sont donnés ici restent valables, la notation $|\cdot|$ désignant alors le module d'un nombre complexe au lieu de la valeur absolue d'un nombre réel.

Il n'est pas évident de construire des fonctions $\varphi \in \mathcal{D}(\Omega)$. En effet, à la frontière du support de φ, toutes les dérivées de φ doivent s'annuler, sinon φ ne peut être indéfiniment dérivable. Nous donnons ci-dessous un exemple de fonction de $\mathcal{D}(\mathbb{R})$.

Exemple 3.2 Soit $H : \mathbb{R} \to \mathbb{R}$ la fonction dite de Heaviside, définie par

$$H(x) = \begin{cases} 1 & \text{si } x > 0, \\ 0 & \text{sinon.} \end{cases} \tag{3.1}$$

On vérifie facilement que la fonction $f(x) = H(x)\,\mathrm{e}^{-1/x}$ est indéfiniment dérivable. Il suffit en effet de remarquer que toutes ses dérivées sont de la forme $H(x)\,P(1/x)\,\mathrm{e}^{-1/x}$ où P est un polynôme. Ainsi la fonction $\varphi(x) = f(1-x^2)$ est indéfiniment dérivable et son support est $[-1,+1]$. Elle appartient donc à $\mathcal{D}(\mathbb{R})$. •

L'espace des distributions va être défini comme étant le dual (topologique) de l'espace des fonctions tests, c'est-à-dire l'espace des formes linéaires continues sur $\mathcal{D}(\Omega)$. Rappelons qu'une forme linéaire sur $\mathcal{D}(\Omega)$ est une application linéaire de $\mathcal{D}(\Omega)$ dans $\mathbb{R}$. Comme nous l'avons déjà fait dans le paragraphe précédent, si T désigne une telle application, nous noterons $\langle T, \varphi\rangle$ au lieu de $T(\varphi)$ l'image de φ par T. Les crochets $\langle\,\cdot,\cdot\,\rangle$ représentent un *produit de dualité :* cette notation est à rapprocher des "bra" $\langle\cdot|$ et des "ket" $|\cdot\rangle$ introduits par Dirac en mécanique quantique.

Comment exprime-t-on la continuité d'une forme linéaire T sur $\mathcal{D}(\Omega)$? La question est loin d'être triviale. Notons tout d'abord qu'il n'est pas possible de munir $\mathcal{D}(\Omega)$ d'une norme qui en ferait un espace complet. En fait, la définition de la topologie dont est muni habituellement cet espace est très compliquée, et sort largement du cadre de ce cours. Nous allons donc nous contenter de parachuter la définition ci-dessous, où nous utilisons la notation condensée suivante pour désigner une dérivée partielle d'une fonction $\varphi \in \mathcal{D}(\Omega)$. Pour un multi-indice $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$, on note

$$\partial^\alpha \varphi = \frac{\partial^{|\alpha|}\varphi}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}} \quad \text{où } |\alpha| = \sum_{i=1}^{d} \alpha_i.$$

Définition 3.3 (distributions) *Une forme linéaire T sur $\mathcal{D}(\Omega)$ est dite continue si, pour tout compact $K \subset \Omega$, il existe une constante $C_K > 0$ et un entier m_K tels que*

$$|\langle T, \varphi\rangle| \leq C_K \sup_{|\alpha|\leq m_K} \sup_{x\in K} |\partial^\alpha \varphi(x)| \quad \forall \varphi \in \mathcal{D}_K(\Omega). \tag{3.2}$$

On appelle distribution dans l'ouvert Ω toute forme linéaire continue sur $\mathcal{D}(\Omega)$, et on note $\mathcal{D}'(\Omega)$ l'espace des distributions dans Ω.

Définition 3.4 (ordre d'une distribution) *Dans la définition précédente, si m_K peut être choisi indépendamment de K, la distribution est dite d'ordre fini, et son ordre est la plus petite valeur de m_K qu'il est possible de choisir.*

L'inégalité (3.2) est à rapprocher du critère de continuité (1.9) d'une application linéaire entre deux espaces vectoriels normés. Mais ici, $\mathcal{D}(\Omega)$ n'étant pas normé, la continuité est assurée non pas par *une* mais par toute une collection d'inégalités du type (1.9). Le second membre de (3.2) définit, pour chaque compact $K \subset \Omega$ et chaque entier m_K une semi-norme sur $\mathcal{D}(\Omega)$ (cf. définition 1.3). Et c'est à partir de cette famille de semi-normes qu'est définie la topologie de $\mathcal{D}(\Omega)$.

3.1.3 Quelques exemples de distributions

Fonctions localement intégrables

On désigne par $L^1_{\text{loc}}(\Omega)$ l'ensemble des fonctions localement intégrables sur Ω, c'est-à-dire dont la restriction à n'importe quel *compact* $K \subset \Omega$ appartient à $L^1(K)$. Pour une fonction $u \in L^1_{\text{loc}}(\Omega)$, considérons la forme linéaire T_u définie sur $\mathcal{D}(\Omega)$ par

$$\langle T_u, \varphi \rangle = \int_\Omega u(x)\, \varphi(x)\, \mathrm{d}x \quad \forall \varphi \in \mathcal{D}(\Omega). \tag{3.3}$$

Si le support de φ est contenu dans un compact $K \subset \Omega$, on a clairement

$$|\langle T_u, \varphi \rangle| \leq \int_K |u(x)|\, \mathrm{d}x \ \sup_{x \in K} |\varphi(x)|,$$

ce qui montre que T_u vérifie la propriété de continuité (3.2) avec $m_K = 0$ et $C_K = \int_K |u(x)|\, \mathrm{d}x$. Ainsi T_u est une distribution d'ordre 0.

Il est clair que deux fonctions égales presque partout définissent une même distribution (puisque l'intégrale dans (3.3) est inchangée). La proposition suivante (que nous admettrons) nous dit que l'application $u \mapsto T_u$ est injective.

Proposition 3.5 *Pour tous $u, v \in L^1_{\text{loc}}(\Omega)$, on a*

$$T_u = T_v \Longrightarrow u(x) = v(x) \text{ presque partout dans } \Omega.$$

Ce résultat nous permet d'abandonner la notation T_u en notant directement u la distribution associée à la fonction u, ce qui revient à considérer $L^1_{\text{loc}}(\Omega)$ comme un sous-espace de $\mathcal{D}'(\Omega)$:

$$L^1_{\text{loc}}(\Omega) \subset \mathcal{D}'(\Omega).$$

Ainsi, au lieu de (3.3), on écrira

$$\langle u, \varphi \rangle = \int_\Omega u(x)\, \varphi(x)\, \mathrm{d}x \quad \forall \varphi \in \mathcal{D}(\Omega). \tag{3.4}$$

C'est en ce sens que les distributions généralisent la notion de fonction. L'espace $\mathcal{D}'(\Omega)$ contient les fonctions localement intégrables, mais aussi beaucoup d'autres choses !

Mesure de Dirac et multipôles

Soit a un point de $\Omega \subset \mathbb{R}^d$. La forme linéaire δ_a définie par

$$\langle \delta_a, \varphi \rangle = \varphi(a) \quad \forall \varphi \in \mathcal{D}(\Omega), \tag{3.5}$$

est appelée *mesure* (ou masse) *de Dirac* au point a. Elle vérifie clairement la relation de continuité (3.2). On peut en effet choisir $C_K = 1$ et $m_K = 0$, ce qui montre que δ_a est une distribution d'ordre 0. Cette distribution est souvent notée simplement δ lorsque $a = 0$.

De la même façon, si α est un multi-indice de $\mathbb{N}^d$, la forme linéaire T définie par

$$\langle T, \varphi \rangle = \partial^\alpha \varphi(a) \quad \forall \varphi \in \mathcal{D}(\Omega), \tag{3.6}$$

est une distribution d'ordre inférieur ou égal à $|\alpha|$. Et on peut montrer qu'elle est exactement d'ordre $|\alpha|$. Nous verrons plus loin le lien entre cette distribution, appelée multipôle, et la mesure de Dirac (voir l'exemple 3.8).

Valeur principale

Il est possible de définir des distributions à partir de fonctions qui ne sont pas localement intégrables. Un exemple très utile est donné en dimension 1 ($\Omega = \mathbb{R}$) par la fonction $x \mapsto 1/x$. Cette fonction n'est évidemment pas intégrable au voisinage de 0, et la définition (3.4) ne permet donc pas de lui associer une distribution. La notion de *valeur principale* (de Cauchy) permet de le faire : il suffit en quelque sorte de retirer la source du problème. Posons en effet

$$\langle \mathrm{vp}(1/x), \varphi \rangle = \lim_{\varepsilon \to 0} \int_{\mathbb{R} \setminus [-\varepsilon, +\varepsilon]} \frac{\varphi(x)}{x} \, \mathrm{d}x \quad \forall \varphi \in \mathcal{D}(\mathbb{R}). \tag{3.7}$$

Pour vérifier que la limite existe bien, rappelons la formule de Taylor avec reste intégral à l'ordre 0 :

$$\varphi(x) = \varphi(0) + x \int_0^1 \varphi'(tx) \, \mathrm{d}t.$$

On a par conséquent

$$\int_{\mathbb{R} \setminus [-\varepsilon, +\varepsilon]} \frac{\varphi(x)}{x} \, \mathrm{d}x = \varphi(0) \int_{\varepsilon < |x| < R} \frac{\mathrm{d}x}{x} + \int_{\varepsilon < |x| < R} \int_0^1 \varphi'(tx) \, \mathrm{d}t \, \mathrm{d}x.$$

où $R > 0$ est choisi tel que le support de φ est inclus dans $[-R, +R]$. La première intégrale du second membre est nulle (par symétrie) et la seconde admet évidemment une limite lorsque $\varepsilon \to 0$ qui n'est autre que $\int_{-R}^{+R} \int_0^1 \varphi'(tx) \, \mathrm{d}t \, \mathrm{d}x$. De plus

$$\left| \int_{-R}^{+R} \int_0^1 \varphi'(tx) \, \mathrm{d}t \, \mathrm{d}x \right| \leq 2R \sup_{x \in [-R, +R]} |\varphi'(x)|,$$

ce qui montre que la relation de continuité (3.2) est vérifiée avec $C_K = 2R$ et $m_K = 1$. Ainsi (3.7) définit bien une distribution $\mathrm{vp}(1/x) \in \mathcal{D}'(\mathbb{R})$ d'ordre inférieur ou égal à 1. Et on peut montrer qu'elle est exactement d'ordre 1.

3.2 Dérivation des distributions

Les distributions possèdent une propriété essentielle qui rend leur manipulation très confortable : *toutes les distributions sont indéfiniment dérivables*. Ceci peut paraître choquant si on se rappelle que les fonctions de $L^1_{\mathrm{loc}}(\Omega)$ sont toutes des distributions, mais ne sont pas toutes dérivables. La réponse réside dans le fait que quand on parle de la dérivation d'une fonction, la dérivée doit être elle-même une fonction, alors que pour les distributions, la dérivée sera une distribution. Ainsi, la dérivée au sens des distributions d'une fonction non dérivable au sens des fonctions sera une distribution qui ne peut être associée à une fonction. Il reste à comprendre comment "dériver au sens des distributions", ce qui va apparaître comme une extension de la dérivation des fonctions.

3.2.1 Définition et exemples

Considérons donc pour commencer une fonction u continûment dérivable dans un ouvert $\Omega \subset \mathbb{R}^d$, et une fonction test $\varphi \in \mathcal{D}(\Omega)$. D'après la formule d'intégration par parties en dimension 1, ou la formule de Green en dimensions supérieures (voir le corollaire 2.35), on a

$$\int_\Omega \frac{\partial u}{\partial x_i}(x)\,\varphi(x)\,\mathrm{d}x = -\int_\Omega u(x)\,\frac{\partial \varphi}{\partial x_i}(x)\,\mathrm{d}x \quad (1 \le i \le d),$$

où aucun terme de bord n'apparaît puisque φ est à support compact dans Ω. A chacune des deux fonctions u et $\partial u/\partial x_i$, on peut associer une distribution par la formule (3.4). En notant encore u et $\partial u/\partial x_i$ ces deux distributions, la formule précédente peut aussi s'écrire

$$\langle \partial u/\partial x_i, \varphi\rangle = -\langle u, \partial\varphi/\partial x_i\rangle. \tag{3.8}$$

Dériver au sens des distributions revient simplement à remplacer dans cette formule la fonction u par une distribution quelconque. Considérons en effet une distribution $T \in \mathcal{D}'(\Omega)$. La relation de continuité (3.2) étant vérifiée, on en déduit que pour tout $\varphi \in \mathcal{D}_K(\Omega)$,

$$|\langle T, \partial\varphi/\partial x_i\rangle| \le C_K \sup_{|\alpha|\le m_K} \sup_{x\in K} |\partial^\alpha(\partial\varphi/\partial x_i)(x)| \le C_K \sup_{|\alpha|\le m_K+1} \sup_{x\in K} |\partial^\alpha \varphi(x)|,$$

ce qui montre que la forme linéaire $\varphi \mapsto -\langle T, \partial\varphi/\partial x_i\rangle$ définit une distribution dans Ω. Et d'après (3.8), il est légitime de noter $\partial T/\partial x_i$ cette distribution : nous ne faisons que généraliser une propriété vérifiée pour des fonctions. Ceci justifie la définition suivante.

Définition 3.6 *Pour tout $T \in \mathcal{D}'(\Omega)$, on note $\partial T/\partial x_i$ la distribution définie par*

$$\langle \partial T/\partial x_i, \varphi\rangle = -\langle T, \partial\varphi/\partial x_i\rangle \quad \forall \varphi \in \mathcal{D}(\Omega). \tag{3.9}$$

Plus généralement, pour tout multi-indice $\alpha \in \mathbb{N}^d$, on note $\partial^\alpha T$ la distribution définie par

$$\langle \partial^\alpha T, \varphi\rangle = (-1)^{|\alpha|}\,\langle T, \partial^\alpha \varphi\rangle \quad \forall \varphi \in \mathcal{D}(\Omega). \tag{3.10}$$

La dérivation au sens des distributions n'est donc qu'une extension de la formule d'intégration par parties. Donnons-en quelques exemples.

Exemple 3.7 Appliquons la définition (3.9) à la fonction de Heaviside (cf (3.1)) :

$$\langle H', \varphi \rangle = -\langle H, \varphi' \rangle = -\int_{\mathbb{R}} H(x)\, \varphi'(x)\, \mathrm{d}x = -\int_0^{+\infty} \varphi'(x)\, \mathrm{d}x = \varphi(0),$$

pour tout $\varphi \in \mathcal{D}(\mathbb{R})$. Ainsi, d'après la définition (3.5) de la mesure de Dirac, on a $H' = \delta$. •

Exemple 3.8 Les multipôles définis en (3.6) ne sont autres que des dérivées de la mesure de Dirac (au signe près). En effet, pour $a \in \mathbb{R}^d$ et pour tout multi-indice $\alpha \in \mathbb{N}^d$, on a

$$\langle \partial^\alpha \delta_a, \varphi \rangle = (-1)^{|\alpha|}\, \langle \delta_a, \partial^\alpha \varphi \rangle = (-1)^{|\alpha|}\, \partial^\alpha \varphi(a) \quad \forall \varphi \in \mathcal{D}(\Omega).$$

•

Exemple 3.9 Quelle est la dérivée au sens des distributions de la fonction $\log|x|$? Cette fonction est localement intégrable sur $\mathbb{R}$ et définit donc bien une distribution de $\mathcal{D}'(\mathbb{R})$ dont la dérivée, que nous notons pour l'instant T, vérifie

$$\langle T, \varphi \rangle = -\langle \log|x|, \varphi' \rangle = -\int_{\mathbb{R}} \log|x|\, \varphi'(x)\, \mathrm{d}x = -\lim_{\varepsilon \to 0} \int_{|x|>\varepsilon} \log|x|\, \varphi'(x)\, \mathrm{d}x,$$

où la dernière égalité résulte du théorème 2.17 de convergence dominée. Ainsi, en intégrant par parties, il vient

$$\langle T, \varphi \rangle = \lim_{\varepsilon \to 0} \left\{ (\varphi(+\varepsilon) - \varphi(-\varepsilon)) \log \varepsilon + \int_{|x|>\varepsilon} \frac{\varphi(x)}{x}\, \mathrm{d}x \right\}.$$

Il reste à remarquer que φ étant dérivable, $\varphi(+\varepsilon) - \varphi(-\varepsilon) = O(\varepsilon)$, donc le premier terme entre accolades tend vers 0, et le second n'est autre que celui qui intervient dans la définition (3.7) de $\mathrm{vp}(1/x)$. En résumé, on a

$$\frac{\mathrm{d}}{\mathrm{d}x} \log|x| = \mathrm{vp}\left(\frac{1}{x}\right).$$

•

Terminons ce paragraphe par une propriété importante qui, en dimension 1, nous fournit la clef de la résolution d'équations différentielles posées au sens des distributions.

Proposition 3.10 *Les seules distributions $T \in \mathcal{D}'(\mathbb{R})$ qui vérifient $T' = 0$ sont les fonctions constantes.*

Démonstration. Remarquons tout d'abord que si $\varphi \in \mathcal{D}(\mathbb{R})$ est une fonction à moyenne nulle, c'est-à-dire telle que

$$\int_{\mathbb{R}} \varphi(x)\,\mathrm{d}x = 0,$$

alors il existe une fonction $\psi \in \mathcal{D}(\mathbb{R})$ telle que $\psi' = \varphi$. Il suffit en effet de poser

$$\psi(x) = \int_{-\infty}^{x} \varphi(y)\,\mathrm{d}y.$$

Cette fonction est clairement indéfiniment dérivable et son support est compact. En effet, si $[a, b]$ contient le support de φ, on a

$$\psi(x) = 0 \text{ si } x < a \quad \text{et} \quad \psi(x) = \int_{a}^{b} \varphi(x)\,\mathrm{d}x = 0 \text{ si } x > b.$$

Ainsi, si $T' = 0$, on aura

$$\langle T, \varphi\rangle = \langle T, \psi'\rangle = -\langle T', \psi\rangle = 0.$$

Considérons maintenant une fonction $\xi \in \mathcal{D}(\mathbb{R})$ telle que $\int_{\mathbb{R}} \xi(x)\,\mathrm{d}x = 1$. Pour tout $\varphi \in \mathcal{D}(\mathbb{R})$, la fonction

$$\varphi_\xi(x) = \varphi(x) - \left(\int_{\mathbb{R}} \varphi(y)\,\mathrm{d}y\right) \xi(x)$$

est évidemment à moyenne nulle. Ainsi, d'après ce qui précède, $\langle T, \varphi_\xi\rangle = 0$, autrement dit,

$$\langle T, \varphi\rangle = \langle T, \xi\rangle \int_{\mathbb{R}} \varphi(y)\,\mathrm{d}y = \langle T, \xi\rangle\ \langle 1, \varphi\rangle = \langle\langle T, \xi\rangle, \varphi\rangle,$$

ce qui signifie que T est la fonction constante égale à $\langle T, \xi\rangle$. □

3.2.2 La formule des sauts en dimension 1

Dans ce paragraphe, nous allons introduire une formule très utile qui généralise le cas de l'exemple 3.7 : elle permet de dériver au sens des distributions des fonctions qui ne sont pas continûment dérivables (au sens des fonctions) globalement sur leur domaine de définition mais qui le sont "par morceaux". Nous ne traitons ici que le cas de la dimension 1. Considérons pour commencer une fonction $u : \mathbb{R} \to \mathbb{R}$ continûment dérivable sauf en $x = 0$, plus précisément telle que ses restrictions $u|_{]-\infty,0[}$ et $u|_{]0,+\infty[}$ appartiennent respectivement à $\mathcal{C}^1(]-\infty, 0])$ et $\mathcal{C}^1([0, +\infty[)$ (cf. exemple 1.2). Ceci signifie en particulier que u admet des limites à gauche et à droite en 0, que nous

noterons respectivement $u(0^-)$ et $u(0^+)$. Une telle fonction appartient évidemment à $L^1_{\mathrm{loc}}(\mathbb{R})$, et constitue donc une distribution de $\mathcal{D}'(\mathbb{R})$ dont la dérivée est définie par

$$\langle u', \varphi \rangle = -\int_{\mathbb{R}} u(x)\,\varphi'(x)\,\mathrm{d}x = -\int_{-\infty}^{0} u(x)\,\varphi'(x)\,\mathrm{d}x - \int_{0}^{+\infty} u(x)\,\varphi'(x)\,\mathrm{d}x,$$

pour tout $\varphi \in \mathcal{D}(\mathbb{R})$. Il est licite d'intégrer par parties sur chacun des intervalles $]-\infty, 0[$ et $]0, +\infty[$, puisque u y est dérivable. Il vient

$$\langle u', \varphi \rangle = \int_{-\infty}^{0} u'(x)\,\varphi(x)\,\mathrm{d}x + \int_{0}^{+\infty} u'(x)\,\varphi(x)\,\mathrm{d}x + (u(0^+) - u(0^-))\,\varphi(0).$$

Notons alors $\{u'\}$ la fonction continue sauf en $x = 0$ égale à la dérivée usuelle de u sur les deux intervalles $]-\infty, 0[$ et $]0, +\infty[$, de sorte que la somme des deux intégrales du second membre peut s'écrire

$$\int_{\mathbb{R}} \{u'\}(x)\,\varphi(x)\,\mathrm{d}x = \langle \{u'\}, \varphi \rangle.$$

D'autre part, par définition de la mesure de Dirac,

$$(u(0^+) - u(0^-))\,\varphi(0) = (u(0^+) - u(0^-))\,\langle \delta, \varphi \rangle.$$

En résumé, nous venons de montrer que

$$u' = \{u'\} + (u(0^+) - u(0^-))\,\delta,$$

ce qui signifie que la dérivée au sens des distributions de u est égale à la somme de sa dérivée usuelle en dehors de 0 et d'une masse de Dirac en 0 pondérée par le saut de u en 0.

Ce résultat se généralise sans difficulté au cas d'une fonction possédant plusieurs points de discontinuité, d'où la proposition suivante.

Proposition 3.11 (Formule des sauts) *Soit $u : \mathbb{R} \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ par morceaux, ce qui signifie qu'il existe un nombre fini de points $a_1 < a_2 < \ldots < a_n$ tels que les restrictions de u aux intervalles $]-\infty, a_1[$, $]a_1, a_2[$, $\ldots$, $]a_n, +\infty[$ appartiennent respectivement à $\mathcal{C}^1(]-\infty, a_1])$, $\mathcal{C}^1([a_1, a_2])$, $\ldots$, $\mathcal{C}^1([a_n, +\infty[)$. En particulier u admet des limites à gauche et à droite en chaque point a_i, notées respectivement $u(a_i^-)$ et $u(a_i^+)$. Alors la dérivée au sens des distributions de u est donnée par*

$$u' = \{u'\} + \sum_{i=1}^{n} (u(a_i^+) - u(a_i^-))\,\delta_{a_i}, \tag{3.11}$$

où $\{u'\}$ désigne la fonction continue par morceaux égale à la dérivée usuelle de u sur chacun des intervalles $]-\infty, a_1[$, $]a_1, a_2[$, $\ldots$, $]a_n, +\infty[$.

Remarquons que si u est de classe $\mathcal{C}^2$ par morceaux, on peut obtenir une formule analogue pour u''. Il suffit de dériver (3.11) au sens des distributions et exprimer $\{u'\}'$ à l'aide de cette même formule. On obtient

$$u'' = \{u''\} + \sum_{i=1}^{n} (u'(a_i^+) - u'(a_i^-))\, \delta_{a_i} + \sum_{i=1}^{n} (u(a_i^+) - u(a_i^-))\, \delta'_{a_i}.$$

Et on peut bien sûr continuer pour exprimer la dérivée m-ème de u, lorsque celle-ci est de classe $\mathcal{C}^m$ par morceaux.

3.2.3 Primitive d'une fonction de $L^1(\mathbb{R})$

Il est bien connu du lecteur que si $f \in \mathcal{C}^0(\mathbb{R})$, alors pour tout $x_0 \in \mathbb{R}$, la fonction

$$g(x) = \int_{x_0}^{x} f(t)\,\mathrm{d}t$$

appartient à $\mathcal{C}^1(\mathbb{R})$ et vérifie $g' = f$. C'est une primitive de f. Le lemme suivant généralise cette propriété à des fonctions $L^1(\mathbb{R})$.

Lemme 3.12 *Soit f une fonction de $L^1(\mathbb{R})$. La fonction g définie par*

$$g(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$$

est continue sur $\mathbb{R}$ et sa dérivée au sens des distributions est f.

Par ailleurs, si $f \in L^1(\mathbb{R})$ satisfait de plus $f' \in L^1(\mathbb{R})$, alors pour tout $x \in \mathbb{R}$,

$$f(x) = \int_{-\infty}^{x} f'(t)\,\mathrm{d}t.$$

Démonstration. On commence par montrer que g est continue. Soient $x \in \mathbb{R}$ et (x_n) une suite réelle tendant vers x. On a

$$g(x_n) - g(x) = \int_{x}^{x_n} f(t)\,\mathrm{d}t = \int_{\mathbb{R}} 1_{[x,x_n]}(t) f(t)\,\mathrm{d}t,$$

où $1_{[a,b]}$ désigne la fonction caractéristique de l'intervalle $[a,b]$. En posant $f_n(t) = 1_{[x,x_n]}(t)f(t)$, on voit que la suite $(f_n(t))$ converge presque partout vers 0 et que la fonction $|f_n|$ est majorée par $|f|$, qui est intégrable. On en déduit par le théorème de convergence dominée que la suite $(g(x_n))$ converge vers $g(x)$, ce qui achève la preuve de la continuité.

Calculons maintenant g' au sens des distributions. Pour $\varphi \in \mathcal{D}(\mathbb{R})$, on a

$$\langle g', \varphi \rangle = -\langle g, \varphi' \rangle = -\int_{\mathbb{R}} g\,\varphi'\,\mathrm{d}x = -\int_{-\infty}^{+\infty} \left(\int_{-\infty}^{x} f(t)\,\mathrm{d}t \right) \varphi'(x)\,\mathrm{d}x$$

$$= -\int_{\mathbb{R}} \left(\int_{\mathbb{R}} 1_{\{t<x\}} f(t)\,\mathrm{d}t \right) \varphi'(x)\,\mathrm{d}x,$$

où $1_{\{t<x\}}$ est la fonction caractéristique du demi-plan $\{t < x\}$. La fonction $(x,t) \mapsto 1_{\{t<x\}} f(t)\varphi'(x)$ étant intégrable sur $\mathbb{R}^2$, le théorème de Fubini conduit à

$$\langle g', \varphi \rangle = -\int_{-\infty}^{+\infty} \left(\int_t^{+\infty} \varphi'(x)\,\mathrm{d}x \right) f(t)\,\mathrm{d}t = \int_{\mathbb{R}} f(t)\varphi(t)\,\mathrm{d}t = \langle f, \varphi \rangle,$$

d'où $g' = f$ au sens des distributions.

Si l'on suppose enfin que $f' \in L^1(\mathbb{R})$, en utilisant le résultat précédent on obtient que la dérivée au sens des distributions de la fonction $x \mapsto f(x) - \int_{-\infty}^{x} f'(t)\,\mathrm{d}t$ est nulle. D'après la proposition 3.10, il existe donc une constante C telle que pour presque tout $x \in \mathbb{R}$,

$$f(x) = \int_{-\infty}^{x} f'(t)\,\mathrm{d}t + C.$$

La fonction f est donc continue et admet des limites finies en $-\infty$ et en $+\infty$. Comme de plus f est intégrable, ces limites sont nécessairement nulles, d'où $C = 0$. □

3.3 Multiplication et division des distributions

3.3.1 Multiplication des distributions

Puisque tout semble permis dans les espaces de distributions, on pourrait s'attendre à ce qu'il soit possible de multiplier entre elles n'importe quelles distributions. Ce n'est pas le cas, loin de là ! Même si ces distributions sont deux fonctions localement intégrables, leur produit n'est pas nécessairement une distribution puisqu'il n'a aucune raison d'être lui aussi localement intégrable (par exemple, si $\Omega = \mathbb{R}$, la fonction $u(x) = 1/\sqrt{|x|}$ est localement intégrable mais son carré ne l'est pas). En fait, le seul produit qui soit toujours autorisé est celui d'une distribution par une fonction indéfiniment dérivable.

L'idée de départ est très simple : il s'agit de généraliser pour une distribution quelconque ce qu'on sait déjà faire pour une fonction. Si $u \in L^1_{\mathrm{loc}}(\Omega)$ et $f \in \mathcal{C}^\infty(\Omega)$, le produit usuel fu appartient clairement à $L^1_{\mathrm{loc}}(\Omega)$. C'est la distribution définie par

$$\langle fu, \varphi \rangle = \int_\Omega \Big(f(x)\,u(x) \Big) \varphi(x)\,\mathrm{d}x \quad \forall \varphi \in \mathcal{D}(\Omega),$$

ce qu'on peut écrire aussi

$$\langle fu, \varphi \rangle = \int_\Omega u(x) \Big(f(x)\,\varphi(x) \Big)\,\mathrm{d}x \quad \forall \varphi \in \mathcal{D}(\Omega),$$

ou encore

$$\langle fu, \varphi \rangle = \langle u, f\varphi \rangle \quad \forall \varphi \in \mathcal{D}(\Omega),$$

où le second membre a bien un sens puisque $f\varphi \in \mathcal{D}(\Omega)$ (c'est le produit de deux fonctions indéfiniment dérivables, et son support est compact car contenu dans celui de φ). Et le tour est joué ! Il reste juste à remplacer u par une distribution arbitraire.

Proposition 3.13 *Pour toute distribution $T \in \mathcal{D}'(\Omega)$ et toute fonction $f \in \mathcal{C}^\infty(\Omega)$, il existe une distribution de $\mathcal{D}'(\Omega)$, notée fT, définie par*

$$\langle fT, \varphi\rangle = \langle T, f\varphi\rangle \quad \forall \varphi \in \mathcal{D}(\Omega). \tag{3.12}$$

La formule habituelle de dérivation d'un produit s'applique encore, soit

$$\frac{\partial}{\partial x_i}(fT) = \frac{\partial f}{\partial x_i}\, T + f\, \frac{\partial T}{\partial x_i}. \tag{3.13}$$

Démonstration. Il nous faut vérifier que le second membre de (3.12) entre bien dans le cadre de la définition 3.3. Comme $T \in \mathcal{D}'(\Omega)$, on sait que pour tout compact $K \subset \Omega$, il existe une constante $C_K > 0$ et un entier m_K tels que

$$|\langle T, \varphi\rangle| \leq C_K \sup_{|\alpha|\leq m_K} \sup_{x\in K} |\partial^\alpha \varphi(x)| \quad \forall \varphi \in \mathcal{D}_K(\Omega),$$

et par conséquent

$$|\langle T, f\varphi\rangle| \leq C_K \sup_{|\alpha|\leq m_K} \sup_{x\in K} |\partial^\alpha (f\varphi)(x)| \leq C'_K \sup_{|\alpha|\leq m_K} \sup_{x\in K} |\partial^\alpha \varphi(x)|,$$

où la seconde inégalité résulte de la formule de Leibniz. Ainsi (3.12) définit bien une distribution.

Montrons maintenant (3.13). En utilisant la définition (3.9) de la dérivée au sens des distributions puis la définition (3.12) de fT, on a

$$\left\langle \frac{\partial}{\partial x_i}(fT), \varphi \right\rangle = -\left\langle fT, \frac{\partial \varphi}{\partial x_i} \right\rangle = -\left\langle T, f\frac{\partial \varphi}{\partial x_i} \right\rangle.$$

En remarquant alors que $f\,\partial\varphi/\partial x_i = \partial(f\varphi)/\partial x_i - \varphi\,\partial f/\partial x_i$, il s'ensuit que

$$\left\langle \frac{\partial}{\partial x_i}(fT), \varphi \right\rangle = -\left\langle T, \frac{\partial}{\partial x_i}(f\varphi) \right\rangle + \left\langle T, \varphi\, \frac{\partial f}{\partial x_i} \right\rangle.$$

En utilisant à nouveau (3.9) et (3.12), on obtient

$$\left\langle \frac{\partial}{\partial x_i}(fT), \varphi \right\rangle = \left\langle f\, \frac{\partial T}{\partial x_i}, \varphi \right\rangle + \left\langle \frac{\partial f}{\partial x_i}\, T, \varphi \right\rangle,$$

d'où le résultat. □

Exemple 3.14 Si T est une masse de Dirac en un point $a \in \Omega$, alors $f\,\delta_a = f(a)\,\delta_a$. En effet

$$\langle f\,\delta_a, \varphi\rangle = \langle \delta_a, f\varphi\rangle = f(a)\,\varphi(a) = f(a)\,\langle \delta_a, \varphi\rangle = \langle f(a)\,\delta_a, \varphi\rangle.$$

•

Exemple 3.15 Dans le cas de la dimension 1, on a $x\ \mathrm{vp}(1/x) = 1$. En effet, d'après la définition (3.7) de $\mathrm{vp}(1/x)$, on a

$$\langle x\,\mathrm{vp}(1/x), \varphi\rangle = \langle \mathrm{vp}(1/x), x\,\varphi\rangle = \lim_{\varepsilon\to 0} \int_{\mathbb{R}\setminus[-\varepsilon,+\varepsilon]} \frac{x\,\varphi(x)}{x}\,\mathrm{d}x,$$

soit, d'après le théorème de convergence dominée,

$$\langle x\,\mathrm{vp}(1/x), \varphi\rangle = \int_{\mathbb{R}} \varphi(x)\,\mathrm{d}x.$$

•

3.3.2 Division des distributions

Nous savons maintenant comment multiplier une distribution par une fonction indéfiniment dérivable. Nous allons nous intéresser ici à l'opération inverse, plus précisément à la résolution d'équation du type

$$\text{Trouver } T \in \mathcal{D}'(\Omega) \text{ tel que } f\,T = S,$$

où $f \in \mathcal{C}^\infty(\Omega)$ et $S \in \mathcal{D}'(\Omega)$ sont donnés. Bien évidemment, cette équation étant linéaire, si on en connaît une solution particulière, toute autre solution s'en déduira en lui rajoutant une solution de l'équation homogène $f\,T = 0$. C'est cette dernière équation que nous allons considérer dans ce paragraphe, en nous limitant au cas de la dimension 1, soit $\Omega = \mathbb{R}$.

La situation la plus simple est celle où la fonction f ne s'annule pas. On retrouve alors le même résultat que lorsque les solutions recherchées sont des fonctions.

Proposition 3.16 *Soit $f \in \mathcal{C}^\infty(\mathbb{R})$ telle que $f(x) \neq 0$ pour tout $x \in \mathbb{R}$. Alors l'unique solution $T \in \mathcal{D}'(\mathbb{R})$ de l'équation $f\,T = 0$ est la solution triviale $T = 0$.*

Démonstration. Il suffit de remarquer que pour tout $\varphi \in \mathcal{D}(\mathbb{R})$, la fonction $\psi = \varphi/f$ appartient aussi à $\mathcal{D}(\mathbb{R})$ puisque $1/f \in \mathcal{C}^\infty(\mathbb{R})$. Ainsi

$$\langle T, \varphi\rangle = \langle T, f\,\psi\rangle = \langle f\,T, \psi\rangle = 0$$

pour tout $\varphi \in \mathcal{D}(\mathbb{R})$, d'où le résultat. □

Les choses se compliquent dès que f s'annule, comme nous le montre la

Proposition 3.17 *Soit $a \in \mathbb{R}$. Toutes les solutions $T \in \mathcal{D}'(\mathbb{R})$ de l'équation*

$$(x-a)\,T = 0$$

sont de la forme $T = k\,\delta_a$ avec $k \in \mathbb{R}$.

Démonstration. Remarquons tout d'abord que d'après l'exemple 3.14, pour tout $k \in \mathbb{R}$, la distribution $T = k\,\delta_a$ vérifie bien $(x-a)\,T = 0$. Il reste à vérifier qu'il n'y a pas d'autres solutions.

Soit $\rho \in \mathcal{D}(\mathbb{R})$ telle que $\rho(x) = 1$ dans un voisinage de $x = a$. Commençons par montrer que pour tout $\varphi \in \mathcal{D}(\mathbb{R})$, il existe $\psi \in \mathcal{D}(\mathbb{R})$ tel que

$$\varphi(x) = \varphi(a)\,\rho(x) + (x-a)\,\psi(x) \quad \text{et} \quad \psi(a) = \varphi'(a). \tag{3.14}$$

On peut pour cela utiliser la formule de Taylor avec reste intégral, soit

$$\xi(x) = \xi(a) + (x-a)\int_0^1 \xi'(tx + (1-t)a)\,\mathrm{d}t,$$

que l'on applique à $\xi = \varphi - \varphi(a)\,\rho$. En constatant que la fonction

$$\psi(x) = \int_0^1 \xi'(tx + (1-t)a)\,\mathrm{d}t$$

appartient à $\mathcal{D}(\mathbb{R})$, on en déduit (3.14).

Cette propriété nous montre que

$$\langle T, \varphi\rangle = \langle T, \varphi(a)\,\rho\rangle + \langle T, (x-a)\,\psi\rangle,$$

où le dernier terme est nul par hypothèse (puisque $\langle T, (x-a)\,\psi\rangle = \langle (x-a)\,T, \psi\rangle$). Ainsi

$$\langle T, \varphi\rangle = \langle T, \rho\rangle\ \varphi(a) = \langle T, \rho\rangle\ \langle \delta_a, \varphi\rangle.$$

La quantité $\langle T, \rho\rangle$ ne dépend pas de φ et joue donc ici le rôle d'une constante, d'où la conclusion. □

Remarque 3.18 *La proposition ci-dessus reste vraie si on remplace $(x-a)$ par une fonction $f \in \mathcal{C}^\infty(\mathbb{R})$ qui possède comme seule racine une racine simple en $x = a$. Dans ce cas, la fonction $f/(x-a)$ ainsi que son inverse appartiennent à $\mathcal{C}^\infty(\mathbb{R})$ et ne s'annulent pas. La proposition* 3.16 *nous dit alors que résoudre $f\,T = 0$ est équivalent à résoudre $(x-a)\,T = 0$.*

Remarque 3.19 *Dans le cas où $f \in \mathcal{C}^\infty(\mathbb{R})$ admet plusieurs racines simples, les solutions de $f\,T = 0$ s'obtiennent simplement en superposant les solutions obtenues pour chaque racine. Par exemple, si $f(x) = (x-a)(x-b)$ avec $a \neq b$, la proposition* 3.17 *nous dit que d'une part $(x-b)\,T = k\,\delta_a$ avec $k \in \mathbb{R}$ et d'autre part $(x-a)\,T = k'\,\delta_b$ avec $k' \in \mathbb{R}$. En effectuant la différence entre ces deux équations, on en conclut que $T = k_a\,\delta_a + k_b\,\delta_b$ avec $(k_a, k_b) \in \mathbb{R}^2$. En utilisant l'argument de la remarque* 3.18*, on voit que ce résultat reste vrai pour une fonction possédant deux racines simples distinctes.*

Le cas où f possède des racines multiples n'est guère plus compliqué. Nous n'énonçons ici que le cas d'une racine double, mais le cas général s'en déduit sans difficulté.

Proposition 3.20 *Soit $a \in \mathbb{R}$. Toutes les solutions $T \in \mathcal{D}'(\mathbb{R})$ de l'équation*

$$(x-a)^2\,T = 0$$

sont de la forme $T = k\,\delta_a + k'\,\delta'_a$ avec $(k, k') \in \mathbb{R}^2$.

Démonstration. Il est facile de vérifier que les distributions de la forme $T = k\,\delta_a + k'\,\delta'_a$ sont bien solutions de $(x-a)^2\,T = 0$. Pour vérifier qu'il n'y en a pas d'autres, remarquons tout d'abord que d'après la proposition 3.17, on sait qu'il existe $k \in \mathbb{R}$ tel que

$$(x-a)\,T = k\,\delta_a.$$

Par ailleurs, d'après (3.14), on a

$$\langle T, \varphi\rangle = \varphi(a)\,\langle T, \rho\rangle + \langle T, (x-a)\,\psi\rangle,$$

où

$$\langle T, (x-a)\,\psi\rangle = \langle (x-a)\,T, \psi\rangle = \langle k\,\delta_a, \psi\rangle = k\,\psi(a) = k\,\varphi'(a),$$

d'où le résultat. □

3.4 Convergence des distributions

Définition 3.21 *On dit qu'une suite $(T_n) \subset \mathcal{D}'(\Omega)$ converge vers $T \in \mathcal{D}'(\Omega)$ si*

$$\lim_{n\to+\infty} \langle T_n, \varphi\rangle = \langle T, \varphi\rangle \quad \forall \varphi \in \mathcal{D}(\Omega). \tag{3.15}$$

Remarque 3.22 *Cette définition très simple de la convergence dans $\mathcal{D}'(\Omega)$ est en même temps très puissante. En particulier, si (T_n) converge vers une distribution T, alors pour tout multi-indice $\alpha \in \mathbb{N}^d$, la suite $(\partial^\alpha T_n)$ converge vers $\partial^\alpha T$, ce qui signifie que la dérivation est une opération continue dans $\mathcal{D}'(\Omega)$. En effet, d'après* (3.10) *et* (3.15)*,*

$$\langle \partial^\alpha T_n, \varphi\rangle = (-1)^{|\alpha|}\,\langle T_n, \partial^\alpha \varphi\rangle \longrightarrow (-1)^{|\alpha|}\,\langle T, \partial^\alpha \varphi\rangle = \langle \partial^\alpha T, \varphi\rangle.$$

Dans la définition 3.21, la limite T est supposée connue a priori. On peut en fait faire l'économie de cette hypothèse, comme nous le dit le théorème suivant (que nous admettrons).

Théorème 3.23 *Soit (T_n) une suite de distributions de $\mathcal{D}'(\Omega)$. Si pour tout $\varphi \in \mathcal{D}(\Omega)$, la suite réelle $(\langle T_n, \varphi\rangle)$ admet une limite $\ell(\varphi)$, alors la famille $\{\ell(\varphi);\ \varphi \in \mathcal{D}(\Omega)\}$ définit une distribution de $\mathcal{D}'(\Omega)$ au sens où il existe $T \in \mathcal{D}'(\Omega)$ telle que*

$$\langle T, \varphi\rangle = \ell(\varphi) \quad \forall \varphi \in \mathcal{D}(\Omega).$$

Ce théorème est très utile pour vérifier qu'une forme linéaire est continue sur $\mathcal{D}(\Omega)$, sans passer par la définition 3.3. Il nous montre en effet que la continuité de T est assurée par l'existence des limites $\ell(\varphi)$ indépendamment les unes des autres.

Pour illustrer la définition ci-dessus, revenons sur l'exemple de la *mesure de Dirac* évoqué en introduction, qui apparaissait formellement comme la limite de la suite de fonctions "chapeaux" de l'exemple 1.4. La proposition suivante justifie ce passage à la limite.

Proposition 3.24 *Soit $u \in L^1(\mathbb{R})$ telle que $\int_{\mathbb{R}} u(x)\,\mathrm{d}x = 1$. Alors la suite $(u_n) \subset L^1(\mathbb{R})$ définie par $u_n(x) = n\,u(nx)$ converge vers δ dans $\mathcal{D}'(\mathbb{R})$.*

Démonstration. Pour tout $\varphi \in \mathcal{D}(\mathbb{R})$, on a

$$\langle u_n, \varphi\rangle = \int_{\mathbb{R}} u_n(x)\,\varphi(x)\,\mathrm{d}x = \int_{\mathbb{R}} n\,u(nx)\,\varphi(x)\,\mathrm{d}x = \int_{\mathbb{R}} u(y)\,\varphi(y/n)\,\mathrm{d}y.$$

En remarquant que

$$\lim_{n\to+\infty} u(y)\,\varphi(y/n) = u(y)\,\varphi(0) \quad \text{presque partout dans } \mathbb{R}, \text{ et}$$
$$|u(y)\,\varphi(y/n)| \le v(y) \quad \text{avec } v = |u| \sup_{x\in\mathbb{R}} |\varphi(x)| \in L^1(\mathbb{R}),$$

le théorème de convergence dominée 2.17 nous permet de conclure :

$$\lim_{n\to+\infty} \langle u_n, \varphi\rangle = \int_{\mathbb{R}} u(x)\,\mathrm{d}x\ \varphi(0) = \langle \delta, \varphi\rangle.$$

□

3.A Annexe : produit de convolution

Cette dernière section ne fait désormais plus partie du programme du cours d'analyse. Cependant, vu l'importance pratique de la notion de convolution des distributions, nous avons choisi de la conserver sous la forme d'une annexe.

Jusqu'ici, le seul produit de convolution que nous sachions effectuer est celui de deux fonctions u et v de $L^1(\mathbb{R}^d)$ (voir le paragraphe 2.3.2). Pour presque tout $x \in \mathbb{R}$, celui-ci est défini par

$$(u \star v)(x) = \int_{\mathbb{R}^d} u(x-y)\, v(y)\, \mathrm{d}y. \tag{3.16}$$

Notre objectif ici est de comprendre comment cette définition peut être étendue à des distributions. De même qu'on peut dériver indéfiniment toute distribution, on pourrait s'attendre à ce qu'il soit possible de convoler tout couple de distributions. Ce n'est malheureusement pas le cas : la possibilité de convoler deux distributions nécessite une certaine compatibilité entre elles. Le critère de compatibilité que nous allons introduire ici repose sur la notion de *support* d'une distribution. Nous allons commencer par définir cette notion et nous verrons ensuite comment l'application de ce critère à la convolution de fonctions localement intégrables conduit à la convolution des distributions.

3.A.1 Support d'une distribution

Commençons par rappeler la définition du *support* d'une fonction $u : \Omega \to \mathbb{R}$ continue (voir (2.2)) : il s'agit de l'adhérence de l'ensemble des points où u ne s'annule pas, soit

$$\mathrm{Supp}(u) = \overline{u^{-1}(\mathbb{R}^*)} = \overline{\{x \in \Omega;\ u(x) \neq 0\}}.$$

Cette définition n'est pas adaptée pour définir le support d'une distribution, pour la simple raison qu'il n'est en général pas possible de parler de la valeur prise par une distribution en un point. Pour les distributions, il suffit de renverser le point de vue : plutôt que de chercher à caractériser directement la zone où une distribution ne s'annule pas, on va regarder là où elle s'annule, ce qui est possible grâce à la notion de *restriction* d'une distribution.

Définition 3.25 *Soient ω et Ω deux ouverts de $\mathbb{R}^d$ tels que $\omega \subset \Omega$. Pour toute distribution $T \in \mathcal{D}'(\Omega)$, la restriction $T|_\omega$ de T à ω est la distribution de $\mathcal{D}'(\omega)$ définie par*

$$\langle T|_\omega, \varphi \rangle = \langle T, \varphi^\Omega \rangle \quad \forall \varphi \in \mathcal{D}(\omega), \tag{3.17}$$

où $\varphi^\Omega \in \mathcal{D}(\Omega)$ désigne le prolongement par 0 de φ en dehors de ω, soit $\varphi^\Omega|_\omega = \varphi$ et $\varphi^\Omega|_{\Omega\setminus\omega} = 0$.

Cette définition a bien un sens : il est clair que dans (3.17), le second membre $\langle T, \varphi^\Omega \rangle$ définit une forme linéaire continue sur $\mathcal{D}(\omega)$ (au sens de la définition 3.3).

Ainsi, on peut dire qu'une distribution $T \in \mathcal{D}'(\Omega)$ s'annule sur un ouvert $\omega \subset \Omega$ si $T|_\omega = 0$, ce qui revient à dire que

$$\langle T, \varphi \rangle = 0 \quad \forall \varphi \in \mathcal{D}(\Omega) \text{ tel que } \mathrm{Supp}(\varphi) \subset \omega. \tag{3.18}$$

Mais pour une distribution T donnée, comment peut-on caractériser toute la zone où elle s'annule ? Y-a-t'il un plus grand ouvert $\omega \subset \Omega$ où cette propriété est vérifiée ? La réponse est donnée par le

Lemme 3.26 *Pour $T \in \mathcal{D}'(\Omega)$, l'union ω_T de tous les ouverts $\omega \subset \Omega$ tels que $T|_\omega = 0$ est le plus grand ouvert où T s'annule.*

DÉMONSTRATION. Comme ω_T est un ouvert (en tant qu'union d'ouverts), il suffit de vérifier que $T|_{\omega_T} = 0$. Soit $\varphi \in \mathcal{D}(\Omega)$ tel que $\mathrm{Supp}(\varphi) \subset \omega_T$. Comme $\mathrm{Supp}(\varphi)$ est compact [3], on peut extraire de la famille $\{\omega \subset \Omega;\ T|_\omega = 0\}$ une famille finie $\{\omega_i;\ i = 1, n\}$ telle que $\mathrm{Supp}(\varphi) \subset \bigcup_{i=1,n} \omega_i$. Considérons alors une partition de l'unité $\{\psi_i \in \mathcal{D}(\omega_i);\ i = 1, n\}$ fournie par le théorème 2.32. Il reste à remarquer que

$$\langle T, \varphi \rangle = \sum_{i=1}^{n} \langle T, \psi_i \varphi \rangle = 0,$$

puisque $\mathrm{Supp}(\psi_i \varphi) \subset \omega_i$, d'où la conclusion. □

Nous sommes enfin en mesure de donner la

Définition 3.27 *Le support d'une distribution $T \in \mathcal{D}'(\Omega)$ est le complémentaire dans Ω du plus grand ouvert où T s'annule :* $\mathrm{Supp}(T) = \Omega \setminus \omega_T$.

Exemple 3.28 Le support de la masse de Dirac $\delta_a \in \mathcal{D}'(\mathbb{R}^d)$ est $\{a\}$. En effet, pour tout $\varphi \in \mathcal{D}(\mathbb{R}^d)$ tel que $a \notin \mathrm{Supp}(\varphi)$, on a

$$\langle \delta_a, \varphi \rangle = \varphi(a) = 0,$$

ce qui montre que la restriction de δ_a à $\mathbb{R}^d \setminus \{a\}$ est nulle. Comme $\delta_a \neq 0$, le plus grand ouvert où δ_a s'annule est bien $\mathbb{R}^d \setminus \{a\}$, d'où la conclusion. •

Exemple 3.29 Soit $u :\ \Omega \to \mathbb{R}$ une fonction localement intégrable. D'après (3.18), dire que $u|_\omega = 0$ pour un ouvert $\omega \subset \Omega$ signifie que

$$\int_\Omega u(x)\, \varphi(x)\, \mathrm{d}x = 0 \quad \forall \varphi \in \mathcal{D}(\Omega) \text{ tel que } \mathrm{Supp}(\varphi) \subset \omega,$$

autrement dit,

$$\int_\omega u(x)\, \varphi(x)\, \mathrm{d}x = 0 \quad \forall \varphi \in \mathcal{D}(\omega).$$

D'après la proposition 3.5, ceci revient à dire que u s'annule *presque partout* dans ω. Ainsi, le support de u est le complémentaire du plus grand ouvert où u s'annule *presque partout*. Et si u est continue, on retrouve la définition usuelle. •

Terminons ce paragraphe par une propriété qui va jouer un rôle essentiel dans la définition de la convolution des distributions (pour cette raison, on se limite au cas où $\Omega = \mathbb{R}^d$). Il s'agit de la possibilité d'étendre le produit de dualité $\langle T, \varphi \rangle$ à certaines

3. On rappelle qu'un ensemble $K \subset \mathbb{R}^d$ est *compact* si, de tout recouvrement de K par une famille d'ouverts de $\mathbb{R}^d$, on peut extraire un sous-recouvrement fini. C'est cette définition qui est utilisée ici. Le théorème de Borel–Lebesgue montre qu'elle est équivalente à dire que K est fermé et borné.

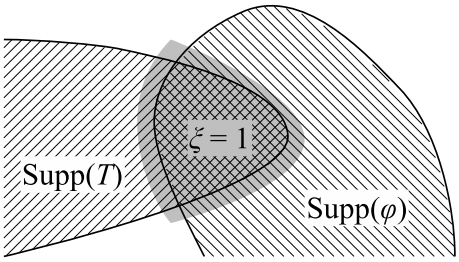


FIGURE 3.1 – Exemple d'une situation où le produit de dualité étendu $\langle T, \varphi \rangle_{\mathrm{e}}$ peut être défini. La zone grisée représente un voisinage de $\mathrm{Supp}(T) \cap \mathrm{Supp}(\varphi)$ où la fonction ξ est constante et égale à 1

fonctions $\varphi \in \mathcal{C}^\infty(\mathbb{R}^d)$ qui n'appartiennent pas à $\mathcal{D}(\mathbb{R}^d)$ (autrement dit, dont le support n'est pas compact). Cette propriété généralise simplement le fait que dans le cas où T est une fonction, soit $T = u \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, l'intégrale $\int_{\mathbb{R}^d} u\,\varphi\,\mathrm{d}x$ ne dépend que de la restriction de φ à l'ensemble $\mathrm{Supp}(u) \cap \mathrm{Supp}(\varphi)$. Par conséquent, si on modifie φ à l'extérieur de cet ensemble, l'intégrale est inchangée. La proposition suivante traduit le fait que, dans le cas d'une distribution quelconque $T \in \mathcal{D}'(\mathbb{R}^d)$, *le produit de dualité* $\langle T, \varphi \rangle$ *ne dépend que de la restriction de* φ *à un voisinage de* $\mathrm{Supp}(T) \cap \mathrm{Supp}(\varphi)$. Contrairement au cas des fonctions, on ne peut en général pas supprimer les mots "*un voisinage de*" dans cette affirmation. On s'en convainc en considérant par exemple le cas d'une dérivée d'une masse de Dirac.

Proposition 3.30 *Soit* T *une distribution de* $\mathcal{D}'(\mathbb{R}^d)$. *Pour toute fonction* $\varphi \in \mathcal{C}^\infty(\mathbb{R}^d)$ *telle que* $\mathrm{Supp}(T) \cap \mathrm{Supp}(\varphi)$ *est compact, on peut définir le produit de dualité étendu de* T *et* φ *par*

$$\langle T, \varphi \rangle_{\mathrm{e}} = \langle T, \xi\varphi \rangle, \tag{3.19}$$

où ξ *est une fonction de* $\mathcal{D}(\mathbb{R}^d)$ *telle que* $\xi = 1$ *dans un voisinage de* $\mathrm{Supp}(T) \cap \mathrm{Supp}(\varphi)$ *(voir figure* 3.1*). Cette définition prolonge le produit de dualité, au sens où*

$$\langle T, \varphi \rangle_{\mathrm{e}} = \langle T, \varphi \rangle \quad \forall \varphi \in \mathcal{D}(\mathbb{R}^d), \tag{3.20}$$

et elle est indépendante de la fonction ξ *choisie.*

DÉMONSTRATION. Remarquons tout d'abord qu'on peut bien choisir une fonction ξ dans $\mathcal{D}(\mathbb{R}^d)$ telle que $\xi = 1$ dans un voisinage de $\mathrm{Supp}(T) \cap \mathrm{Supp}(\varphi)$ car ce dernier est supposé compact. Comme $\varphi \in \mathcal{C}^\infty(\mathbb{R}^d)$, il s'ensuit que $\xi\varphi \in \mathcal{D}(\mathbb{R}^d)$, ce qui montre que le second membre de (3.19) est bien défini. Pour vérifier (3.20), notons que $1 - \xi = 0$ dans un voisinage de $\mathrm{Supp}(T) \cap \mathrm{Supp}(\varphi)$, ce qui entraîne que $\mathrm{Supp}((1-\xi)\varphi) \subset \mathbb{R}^d \setminus \mathrm{Supp}(T)$ (voir figure 3.1), donc $\langle T, (1-\xi)\varphi \rangle = 0$, d'où le résultat. Enfin si ξ_1 et ξ_2 sont deux fonctions de $\mathcal{D}(\mathbb{R}^d)$ telles que $\xi_1 = \xi_2 = 1$ dans un voisinage de $\mathrm{Supp}(T) \cap \mathrm{Supp}(\varphi)$, on en déduit de la même façon que $\langle T, (\xi_1 - \xi_2)\varphi \rangle = 0$ puisque $\xi_1 - \xi_2 = 0$ dans un

voisinage de $\mathrm{Supp}(T) \cap \mathrm{Supp}(\varphi)$. Ainsi $\langle T, \varphi \rangle_{\mathrm{e}}$ est bien indépendant du choix de la fonction ξ. □

3.A.2 Convolution de fonctions localement intégrables

Afin de comprendre comment définir la convolution de deux distributions, nous allons commencer par traiter le cas de fonctions localement intégrables. Peut-on étendre la définition (3.16) à des fonctions de $L^1_{\mathrm{loc}}(\mathbb{R}^d)$? La réponse est évidemment négative dans le cas général. Pour que cette formule ait un sens en un point $x \in \mathbb{R}^d$ donné, il faut que la fonction $y \mapsto u(x-y)\, v(y)$ soit intégrable. Lorsque u et v appartiennent à $L^1(\mathbb{R}^d)$, nous avons déduit cette propriété du théorème de Fubini. Mais elle n'est bien sûr pas vérifiée pour tous les couples de fonctions de $L^1_{\mathrm{loc}}(\mathbb{R}^d)$. Par exemple si u et v sont constantes, cette propriété n'est vérifiée pour aucun x.

On peut identifier différentes catégories de couples $(u, v) \in L^1_{\mathrm{loc}}(\mathbb{R}^d)^2$ pour lesquels cette propriété est satisfaite. Une telle catégorie repose sur une certaine compatibilité des comportements asymptotiques de $u(x-y)$ et $v(y)$ lorsque $|y|$ tend vers l'infini. La catégorie que nous allons considérer ici est fondée sur une compatibilité des *supports* de u et v. L'idée sous-jacente est très simple : on va s'intéresser aux cas où, pour chaque x, l'intégrale dans (3.16) porte en réalité sur un domaine *borné*. Ceci se produit évidemment si u ou v est à support compact, mais aussi dans d'autres situations que nous allons identifier grâce à la définition suivante.

Définition 3.31 (Ensembles convolutifs) (i) *Deux ensembles fermés A et B de $\mathbb{R}^d$ sont dits convolutifs si pour tout $R > 0$, il existe $\rho(R) > 0$ tel que*

$$(x \in A,\ y \in B \text{ et } |x+y| \le R) \Longrightarrow (|x| \le \rho(R) \text{ et } |y| \le \rho(R)).$$

(ii) *Plus généralement une famille de n ensembles fermés $A_i \subset \mathbb{R}^d$, $i = 1$ à n, est dite convolutive si pour tout $R > 0$, il existe $\rho(R) > 0$ tel que*

$$\left(x_i \in A_i,\ i = 1, n \text{ et } \left| \sum_{i=1}^{n} x_i \right| \le R \right) \Longrightarrow (|x_i| \le \rho(R) \quad \forall i = 1, n).$$

En d'autres termes, deux ensembles fermés A et B sont convolutifs si

$$\forall R > 0, \quad (A \times B) \ \cap\ \big\{(x,y) \in \mathbb{R}^d \times \mathbb{R}^d;\ |x+y| \le R\big\} \text{ est borné.}$$

Exemple 3.32 La figure 3.2 donne une interprétation géométrique de cette propriété dans le cas où $d = 1$. La bande hachurée représente l'ensemble $H_R = \{(x,y) \in \mathbb{R}^2;\ |x+y| \le R\}$ pour R fixé, les traits épais sur les axes indiquent les ensembles A et B, et la zone grisée correspond au produit $A \times B$. A gauche, on a choisi $A = [a, +\infty[$ et $B =]-\infty, b]$ et on voit que dans ce cas, $(A \times B) \cap H_R$ n'est jamais borné : A et B ne sont pas convolutifs. A droite, on a toujours $A = [a, +\infty[$ mais cette fois $B = [b', +\infty[$, et il est alors clair que $(A \times B) \cap H_R$ est borné pour tout R : ces ensembles sont convolutifs. •

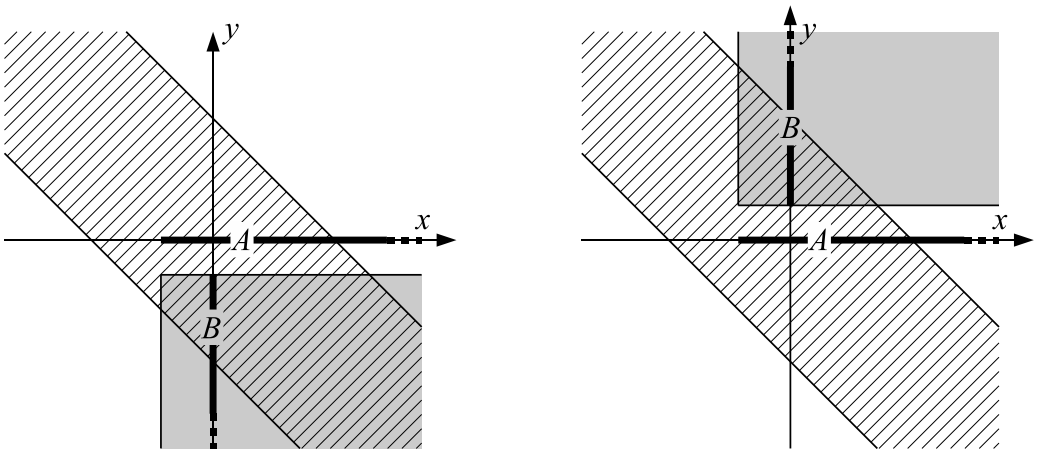


FIGURE 3.2 – Exemples de sous-ensembles A et B de $\mathbb{R}$ non convolutifs (à gauche) et convolutifs (à droite)

On vérifie facilement les propriétés suivantes.

Proposition 3.33 *Les familles suivantes (d'ensembles fermés) sont convolutives.*

(i) *une famille composée d'ensembles compacts, sauf un au plus ;*

(ii) *une sous-famille d'une famille convolutive ;*

(iii) *une famille convolutive complétée d'un ensemble compact ;*

(iv) *en dimension 1, une famille d'intervalles de la forme* $[a_i, +\infty[$.

Nous sommes maintenant en mesure de répondre à notre question initiale.

Théorème 3.34 *Soient u et v deux fonctions de $L^1_{\mathrm{loc}}(\mathbb{R}^d)$ dont les supports sont convolutifs. Pour presque tout $x \in \mathbb{R}^d$, la fonction $y \mapsto u(x-y)\,v(y)$ est intégrable sur $\mathbb{R}^d$, et la fonction $u \star v$ définie par* (3.16) *appartient à $L^1_{\mathrm{loc}}(\mathbb{R}^d)$.*

DÉMONSTRATION. La démarche est très voisine de la démonstration du théorème 2.23. Il s'agit de vérifier que pour tout $R > 0$, l'expression (3.16) définit bien une fonction intégrable sur $B_R = \{x \in \mathbb{R}^d;\ |x| < R\}$. On va montrer pour cela que la fonction $(x, y) \mapsto u(x-y)\,v(y)$ est intégrable sur $B_R \times \mathbb{R}^d$. Les supports de u et v étant supposés convolutifs, on sait qu'il existe $\rho = \rho(R)$ tel que

$$(x - y \in \mathrm{Supp}(u),\ y \in \mathrm{Supp}(v) \text{ et } |x| \leq R) \Longrightarrow (|x-y| \leq \rho \text{ et } |y| \leq \rho).$$

Ainsi, on a

$$\int_{B_R \times \mathbb{R}^d} |u(x-y)\,v(y)|\,\mathrm{d}x\,\mathrm{d}y = \int_{B_\rho} |v(y)| \left(\int_{B_R} |u(x-y)|\,\mathrm{d}x \right) \mathrm{d}y,$$

avec

$$\int_{B_R} |u(x-y)|\,\mathrm{d}x \leq \int_{B_\rho} |u(z)|\,\mathrm{d}z = \|u\|_{L^1(B_\rho)}.$$

Il s'ensuit que

$$\int_{B_R\times\mathbb{R}^d} |u(x-y)\,v(y)|\,\mathrm{d}x\,\mathrm{d}y \leq \|u\|_{L^1(B_\rho)}\ \|v\|_{L^1(B_\rho)},$$

d'où le résultat annoncé. D'après le théorème de Fubini 2.21, on sait alors que pour presque tout x, la fonction $y \mapsto u(x-y)\,v(y)$ appartient à $L^1(\mathbb{R}^d)$, ce qui donne un sens à l'expression (3.16). Il reste à constater que

$$\|u\star v\|_{L^1(B_R)} = \int_{B_R}\left|\int_{\mathbb{R}^d} u(x-y)\,v(y)\,\mathrm{d}y\right|\mathrm{d}x \leq \int_{B_R\times\mathbb{R}^d} |u(x-y)\,v(y)|\,\mathrm{d}x\,\mathrm{d}y,$$

où l'on vient de voir que la dernière intégrale est majorée par $\|u\|_{L^1(B_\rho)}\ \|v\|_{L^1(B_\rho)}$. □

Cette proposition nous montre donc que si u et v sont des fonctions de $L^1_{\mathrm{loc}}(\mathbb{R}^d)$ dont les supports sont convolutifs, leur produit de convolution peut être vu comme une distribution de $\mathcal{D}'(\mathbb{R}^d)$ définie par

$$\langle u\star v,\varphi\rangle = \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} u(x-y)\,v(y)\,\mathrm{d}y\right)\varphi(x)\,\mathrm{d}x,$$

pour tout $\varphi\in\mathcal{D}(\mathbb{R}^d)$. En utilisant le théorème de Fubini et en effectuant le changement de variable $z=x-y$, il vient

$$\langle u\star v,\varphi\rangle = \int_{\mathbb{R}^d\times\mathbb{R}^d} u(z)\,v(y)\,\varphi(z+y)\,\mathrm{d}z\,\mathrm{d}y,$$

ou encore

$$\langle u\star v,\varphi\rangle = \int_{\mathbb{R}^d} u(x)\left(\int_{\mathbb{R}^d} v(y)\,\varphi(x+y)\,\mathrm{d}y\right)\mathrm{d}x \tag{3.21}$$

$$= \int_{\mathbb{R}^d} v(y)\left(\int_{\mathbb{R}^d} u(x)\,\varphi(x+y)\,\mathrm{d}x\right)\mathrm{d}y. \tag{3.22}$$

Cette interprétation du produit de convolution va nous permettre de passer d'un couple de fonctions à un couple de distributions.

3.A.3 Convolution de distributions

L'idée consiste simplement à récrire le second membre de (3.21) ou (3.22) en remplaçant les intégrales par des produits de dualité. En choisissant la formule (3.21), on est alors conduit au

Théorème 3.35 *Soient S et T deux distributions de $\mathcal{D}'(\mathbb{R}^d)$ dont les supports sont convolutifs. Alors il existe une distribution $S \star T \in \mathcal{D}'(\mathbb{R}^d)$, appelée produit de convolution de S et T, définie par*[4]

$$\langle S \star T, \varphi \rangle = \langle S(x), \langle T(y), \varphi(x+y) \rangle \rangle \quad \forall \varphi \in \mathcal{D}(\mathbb{R}^d). \tag{3.23}$$

Tout la question est de comprendre le sens de cette "définition" du produit de convolution. Examinons cette formule.

Tout d'abord, nous avons noté ici $S(x)$ et $T(y)$ comme si S et T étaient des fonctions, alors que nous savons qu'il n'est en général pas possible de parler de la valeur prise par une distribution en un point. Cette notation a simplement pour but d'indiquer clairement que les deux produits de dualité qui apparaissent portent sur deux variables différentes. D'une part, le produit intérieur $\langle T(y), \varphi(x+y) \rangle$ porte sur la variable y, la variable x y jouant le rôle d'un paramètre. En notant φ_x la fonction de $\mathcal{D}(\mathbb{R}^d)$ définie par $\varphi_x(y) = \varphi(x+y)$, ce produit de dualité aurait pu prendre la forme plus conventionnelle $\langle T, \varphi_x \rangle$, ce qui définit donc une fonction de la variable x, que nous allons noter $\varphi^T(x)$. D'autre part, le produit extérieur faisant intervenir S porte sur la variable x, et nous pouvons maintenant l'écrire plus simplement $\langle S, \varphi^T \rangle$. Toute la difficulté réside dans la justification de ce dernier produit de dualité, qui n'a a priori de sens que si φ^T appartient à $\mathcal{D}(\mathbb{R}^d)$, ce qui n'est en général pas le cas ! Le lemme suivant montre qu'il s'agit d'une fonction de classe $\mathcal{C}^\infty$, mais nous verrons ensuite que son support n'est en général pas compact.

Lemme 3.36 (Dérivation sous le crochet) *Soient $T \in \mathcal{D}'(\mathbb{R}^d)$, $\varphi \in \mathcal{D}(\mathbb{R}^d)$, et φ_x la fonction de $\mathcal{D}(\mathbb{R}^d)$ définie pour tout $x \in \mathbb{R}^d$ par $\varphi_x(y) = \varphi(x+y)$. Alors la fonction $\varphi^T : x \mapsto \langle T, \varphi_x \rangle$ est indéfiniment dérivable dans $\mathbb{R}^d$. De plus, pour tout multi-indice $\alpha \in \mathbb{N}^d$, on a*

$$\partial^\alpha \varphi^T(x) = \partial^\alpha \langle T, \varphi_x \rangle = \langle T, \partial^\alpha \varphi_x \rangle. \tag{3.24}$$

Démonstration. Montrons tout d'abord que φ^T est continue dans tout ouvert borné $\mathcal{O} \subset \mathbb{R}^d$. Notons $K = \{y \in \mathbb{R}^d;\ \exists x \in \overline{\mathcal{O}},\ x+y \in \mathrm{Supp}(\varphi)\}$. Cet ensemble est évidemment compact et contient par construction le support de φ_x pour tout $x \in \mathcal{O}$. Comme T est une distribution (cf. (3.2)), on sait qu'il existe une constante $C_K > 0$ et un entier m_K tels que

$$|\langle T, \varphi_x - \varphi_{x'} \rangle| \leq C_K \sup_{|\beta| \leq m_K} \sup_{y \in K} |\partial_y^\beta (\varphi_x(y) - \varphi_{x'}(y))| \quad \forall x, x' \in \mathcal{O}.$$

4. Dans la définition (3.23), le second membre fait apparaître des produits de dualités emboîtés qui jouent pour les distributions un rôle analogue aux intégrales multiples pour les fonctions. C'est la notion de *produit tensoriel* qui se cache derrière cet emboîtement. Pour la simplicité de la présentation, nous avons choisi de ne pas développer cette notion.

En remarquant que $\varphi_x(y) - \varphi_{x'}(y) = \varphi(x+y) - \varphi(x'+y)$, on déduit du théorème des accroissements finis que

$$|\langle T, \varphi_x - \varphi_{x'}\rangle| \leq |x - x'|\ C_K \sup_{|\beta|\leq m_K+1} \sup_{y\in\mathbb{R}^d} |\partial^\beta \varphi(y)|,$$

où le second membre tend vers 0 lorsque $|x - x'| \to 0$.

Pour montrer que φ^T admet des dérivées partielles $\partial\varphi^T/\partial x_i$ pour tout $i = 1, d$, on suit la même démarche en remplaçant $\varphi_x - \varphi_{x'}$ par $(\varphi_{x+\tau e_i} - \varphi_x)/\tau - \partial\varphi_x/\partial x_i$ où $\tau \in \mathbb{R}$ et e_i désigne le i-ème vecteur de la base canonique de $\mathbb{R}^d$. Ainsi, pour $x \in \mathcal{O}$ et $|\tau|$ choisi assez petit de sorte que $x + \tau e_i \in \mathcal{O}$, on obtient

$$\left|\left\langle T, \frac{\varphi_{x+\tau e_i} - \varphi_x}{\tau} - \frac{\partial\varphi_x}{\partial x_i}\right\rangle\right| \leq \tau\ C_K \sup_{|\beta|\leq m_K+2} \sup_{y\in\mathbb{R}^d} |\partial^\beta \varphi(y)|.$$

En faisant tendre τ vers 0, on en déduit (3.24) pour $|\alpha| = 1$. Et on montre la continuité de $\partial\varphi^T/\partial x_i$ comme on l'a fait pour φ^T. Enfin, le même résultat pour les dérivées partielles d'ordres supérieurs s'obtient par récurrence. □

Que peut-on dire du support de φ^T ? Lorsque le support de la distribution T est compact, celui de φ^T l'est aussi. En effet, lorsque $|x|$ augmente, le support de φ_x (qui n'est autre que le support de φ translaté de $-x$) s'éloigne de l'origine, de sorte que pour $|x|$ assez grand, $\mathrm{Supp}(\varphi_x) \subset \mathbb{R}^d \setminus \mathrm{Supp}(T)$, donc $\varphi^T(x) = 0$. En revanche, si $\mathrm{Supp}(T)$ n'est pas compact, $\mathrm{Supp}(\varphi^T)$ ne l'est pas non plus. Comment peut-on alors donner un sens à $\langle S, \varphi^T\rangle$? La réponse réside dans la proposition 3.30 qui nous permet d'interpréter cette quantité comme un produit de dualité étendu sous la condition que $\mathrm{Supp}(S) \cap \mathrm{Supp}(\varphi^T)$ soit compact. Et cette condition est bien vérifiée. En effet, si $x \in \mathrm{Supp}(\varphi^T)$, alors $\mathrm{Supp}(T) \cap \mathrm{Supp}(\varphi_x)$ n'est pas vide, autrement dit, il existe $y \in \mathrm{Supp}(T)$ tel que $x + y \in \mathrm{Supp}(\varphi)$. Ainsi, dire que $x \in \mathrm{Supp}(S) \cap \mathrm{Supp}(\varphi^T)$ revient à dire qu'il existe $y \in \mathbb{R}^d$ tel que

$$x \in \mathrm{Supp}(S), \quad y \in \mathrm{Supp}(T) \quad \text{et} \quad x + y \in \mathrm{Supp}(\varphi).$$

Il reste à constater que comme les supports de S et T sont supposés convolutifs, ceci entraîne que x et y restent tout deux dans un ensemble borné, d'où le résultat.

Nous savons maintenant quel est le sens de la "définition" (3.23) qu'il faudrait plutôt écrire sous la forme

$$\langle S \star T, \varphi\rangle = \left\langle S, \varphi^T\right\rangle_{\mathrm{e}} = \langle S(x), \langle T(y), \varphi(x+y)\rangle\rangle_{\mathrm{e}} \quad \forall\varphi \in \mathcal{D}(\mathbb{R}^d),$$

où l'indice 'e' désigne le produit de dualité étendu de la proposition 3.30. Il est alors facile de vérifier que cette expression définit bien une distribution de $\mathcal{D}'(\mathbb{R}^d)$, au sens de la définition 3.3. Le théorème 3.35 est donc démontré.

Exemple 3.37 Donnons un exemple très simple, mais fondamental, qui nous dit que δ est élément neutre pour la convolution, soit $T \star \delta = T$ pour toute distribution $T \in \mathcal{D}'(\mathbb{R}^d)$. Il suffit en effet de constater que

$$\langle T \star \delta, \varphi\rangle = \langle T(x), \langle \delta(y), \varphi(x+y)\rangle\rangle_{\mathrm{e}} = \langle T(x), \varphi(x)\rangle_{\mathrm{e}} = \langle T, \varphi\rangle.$$

•

Les principales propriétés du produit de convolution des distributions sont regroupées dans la proposition suivante, que nous admettrons. Notons que compte tenu de notre choix de présentation (qui évite en particulier la notion de *produit tensoriel*, voir note 4 page 70), la commutativité du produit de convolution est difficile à démontrer.

Proposition 3.38 *Soient S, T et U des distributions de $\mathcal{D}'(\mathbb{R}^d)$ dont les supports sont convolutifs.*

(i) *Le produit de convolution est commutatif :* $S \star T = T \star S$.

(ii) *Le produit de convolution est associatif :* $(S \star T) \star U = S \star (T \star U)$.

(iii) $\mathrm{Supp}(S \star T) \subset \mathrm{Supp}(S) + \mathrm{Supp}(T) = \{x+y \in \mathbb{R}^d;\ x \in \mathrm{Supp}(S)$ *and* $y \in \mathrm{Supp}(T)\}$.

(iv) *Pour tout multi-indice* $\alpha \in \mathbb{N}^d$, *on a* $\partial^\alpha(S \star T) = (\partial^\alpha S) \star T = S \star (\partial^\alpha T)$.

Chapitre 4

Transformation de Fourier

La transformation de Fourier est un outil fondamental de l'analyse. Elle est utilisée dans des domaines très divers : l'étude des équations aux dérivées partielles, le traitement du signal... Elle permet (dans de nombreux cas que l'on précisera dans la suite) de représenter une fonction comme une somme de fonctions exponentielles complexes.

Une propriété importante de la transformation de Fourier est le fait qu'elle transforme les dérivations par rapport à une variable en multiplication par la variable de Fourier correspondante. Cela permet de ramener l'étude des équations différentielles (à coefficients constants) à des équations algébriques et les équations aux dérivées partielles à des équations différentielles ou même algébriques. Plus généralement, la transformation de Fourier transforme la convolution en produit. Ainsi la résolution d'une équation de convolution se ramène à un problème de division. Cette propriété simplifie grandement les calculs et permet d'accéder plus facilement à la résolution de nombreux problèmes.

Il est important de remarquer que la transformation de Fourier ne s'applique que lorsque l'ouvert Ω est égal à l'espace $\mathbb{R}^d$ tout entier. Cependant, on peut encore utiliser la transformation de Fourier lorsque $\Omega \neq \mathbb{R}^d$ à condition de multiplier au préalable les fonctions que l'on cherche à transformer par une fonction de $\mathcal{D}(\Omega)$. Ceci permet d'obtenir des informations locales sur la solution d'un problème posé dans Ω.

Nous commençons par décrire la transformation de Fourier dans le cas des fonctions intégrables. Puis nous la définissons dans le cas des fonctions de carré intégrable et établissons dans ce cadre des résultats essentiels comme la formule de réciprocité, la formule de Plancherel et le fait que la transformation de Fourier est une isométrie de l'espace L^2 des fonctions de carré intégrable. Il est à noter que dans ce chapitre les fonctions que nous considérons sont à valeurs dans $\mathbb{C}$.

La transformation de Fourier s'étend aux distributions, plus précisément à un sous-espace de $\mathcal{D}(\mathbb{R}^d)$ appelé l'espace des distributions tempérées. Cette extension, essentielle dans l'analyse des équations aux dérivées partielles, ne sera pas abordée ici. Le lecteur intéressé se reportera à [4] ou [1].

4.1 Transformation de Fourier des fonctions L^1

On commence par définir la transformée de Fourier pour les fonctions intégrables.

4.1.1 Le théorème de Riemann-Lebesgue

Dans ce qui suit, f désigne une fonction à valeurs complexes de la variable $x \in \mathbb{R}^d$.

Définition 4.1 *Si $f \in L^1(\mathbb{R})$, on appelle transformée de Fourier de f et on note $\widehat{f}$ ou $\mathcal{F}(f)$ la fonction à valeurs complexes de la variable réelle ξ définie par :*

$$\widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) \mathrm{e}^{-\mathrm{i}x\xi} \, \mathrm{d}x. \tag{4.1}$$

Plus généralement, pour $d \geq 1$ et $f \in L^1(\mathbb{R}^d)$, on appelle transformée de Fourier de f et on note $\widehat{f}$ (ou $\mathcal{F}(f)$) la fonction à valeurs complexes de la variable $\xi \in \mathbb{R}^d$ définie par :

$$\widehat{f}(\xi) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(x) \mathrm{e}^{-\mathrm{i}x\cdot\xi} \, \mathrm{d}x \tag{4.2}$$

où $x \cdot \xi = \sum_{j=1}^d x_j \xi_j$.

Il est clair que, l'exponentielle $\mathrm{e}^{-\mathrm{i}x\cdot\xi}$ étant de module 1, $\widehat{f}$ est bien définie quand $f \in L^1(\mathbb{R}^d)$.

Remarque 4.2 *On peut définir de la même manière la conjuguée de la transformée de Fourier $\overline{\mathcal{F}}(f)$ comme l'analogue de (4.2) en remplaçant $-\mathrm{i}$ par* i.

Remarque 4.3 *Dans le cas particulier d'une fonction à variables séparées :*

$$f(x) = f_1(x_1) f_2(x_2) \dots f_d(x_d),$$

avec $f_j \in L^1(\mathbb{R})$ pour $j = 1, \dots, d$, on vérifie facilement en utilisant le théorème de Fubini (théorème 2.21) que :

$$\widehat{f}(\xi) = \widehat{f_1}(\xi_1) \widehat{f_2}(\xi_2) \dots \widehat{f_d}(\xi_d).$$

Nous aurons dans la suite besoin du théorème de densité suivant, que nous admettrons :

Théorème 4.4 *$\mathcal{D}(\mathbb{R}^d)$ est dense dans $L^1(\mathbb{R}^d)$.*

On établit maintenant le résultat fondamental suivant, qui précise la nature de la fonction $\widehat{f}$:

Théorème 4.5 (de Riemann-Lebesgue) *Si $f \in L^1(\mathbb{R}^d)$, $\widehat{f}$ est une fonction continue de $\xi \in \mathbb{R}^d$ qui tend vers 0 à l'infini.*

DÉMONSTRATION. La continuité de $\widehat{f}$ est une conséquence du théorème de convergence dominée de Lebesgue (théorème 2.17). Soit en effet (ξ_n) une suite d'éléments de $\mathbb{R}^d$ qui converge vers ξ. On a

$$\widehat{f}(\xi_n) = \int_{\mathbb{R}^d} g_n(x)\,\mathrm{d}x,$$

avec

$$g_n(x) = \frac{1}{(2\pi)^{d/2}} f(x)\mathrm{e}^{-\mathrm{i}x\cdot\xi_n}.$$

Pour presque tout x, la suite $(g_n(x))$ converge vers $g(x) = f(x)\mathrm{e}^{-\mathrm{i}x\cdot\xi}/(2\pi)^{d/2}$ et la fonction $|g_n|$ est majorée par $|f|/(2\pi)^{d/2}$, qui est intégrable. On en déduit que la suite $(\widehat{f}(\xi_n))$ converge vers $\widehat{f}(\xi)$, d'où le résultat.

Prouvons maintenant que $\widehat{f}$ tend vers 0 à l'infini. Cela se démontre en deux temps. On établit tout d'abord le résultat pour les fonctions de $\mathcal{D}(\mathbb{R}^d)$ (voir la définition 3.1), puis on l'étend aux fonctions de $L^1(\mathbb{R}^d)$ en utilisant le théorème de densité 4.4.

Si $f \in \mathcal{D}(\mathbb{R}^d)$, une intégration par parties donne :

$$\mathrm{i}\xi_j \widehat{f}(\xi) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \frac{\partial f}{\partial x_j}(x)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\,\mathrm{d}x \text{ pour tout } j = 1,\dots,d, \tag{4.3}$$

d'où

$$\left(\max_{j=1,\dots,d} |\xi_j|\right) |\widehat{f}(\xi)| \leq \frac{1}{(2\pi)^{d/2}} \max_{j=1,\dots,d} \left\| \frac{\partial f}{\partial x_j} \right\|_{L^1}.$$

On a donc finalement

$$|\widehat{f}(\xi)| \leq \frac{1}{(2\pi)^{d/2}} \frac{1}{\max_{j=1,\dots,d} |\xi_j|} \max_{j=1,\dots,d} \left\| \frac{\partial f}{\partial x_j} \right\|_{L^1},$$

ce qui montre que $\widehat{f}$ tend vers 0 lorsque $|\xi|$ tend vers l'infini.

Soit maintenant $f \in L^1(\mathbb{R}^d)$ et $\varepsilon > 0$. Par le théorème 4.4, il existe une fonction $g \in \mathcal{D}(\mathbb{R}^d)$ telle que

$$\|f - g\|_{L^1} \leq (2\pi)^{d/2} \frac{\varepsilon}{2}.$$

Par ailleurs :

$$|\widehat{f}(\xi)| = \left| \widehat{g}(\xi) + \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} (f(x) - g(x))\mathrm{e}^{-\mathrm{i}x\cdot\xi}\,\mathrm{d}x \right| \leq |\widehat{g}(\xi)| + \frac{1}{(2\pi)^{d/2}} \|f - g\|_{L^1}.$$

D'après la première partie de la preuve, il existe A tel que

$$|\widehat{g}(\xi)| \leq \frac{\varepsilon}{2} \quad \text{si } |\xi| \geq A.$$

On a donc $|\widehat{f}(\xi)| \leq \varepsilon$ pour $|\xi| \geq A$, ce qui achève la démonstration du théorème. □

Une conséquence immédiate de l'expression (4.2) de $\widehat{f}$ est l'inégalité

$$\sup_{\xi\in\mathbb{R}^d} |\widehat{f}(\xi)| \leq \frac{1}{(2\pi)^{d/2}} ||f||_{L^1}, \tag{4.4}$$

qui implique la continuité de l'application $\mathcal{F}$ de l'espace L^1 dans l'espace des fonctions continues tendant vers 0 à l'infini muni de la norme de la convergence uniforme.

4.1.2 Propriétés de la transformation de Fourier dans L^1

Nous allons maintenant énoncer des propriétés plus générales concernant la régularité et le comportement à l'infini de la transformée de Fourier. On énonce un premier théorème qui précise celui de Riemann-Lebesgue dans le cas de fonctions plus régulières que simplement L^1. Nous utilisons la notation condensée des dérivées partielles introduite au paragraphe 3.1.2.

Théorème 4.6 *Si f et toutes ses dérivées partielles jusqu'à l'ordre m sont intégrables, alors :*

$$\mathcal{F}\left(\partial^\alpha f\right) = \widehat{\partial^\alpha f} = (\mathrm{i}\xi)^\alpha \mathcal{F} f, \tag{4.5}$$

pour tout multi-indice $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d) \in \mathbb{N}^d$ tel que

$$|\alpha| = \sum_{j=1,\ldots,d} \alpha_j \leq m.$$

Il en résulte en particulier que $\mathcal{F}f(\xi)$ tend vers 0 à l'infini plus vite que $1/|\xi|^m$.

Démonstration. On commence par montrer que si f est une fonction intégrable et $\partial f/\partial x_j$ (définie au sens des distributions) l'est également, alors

$$\mathcal{F}\left(\frac{\partial f}{\partial x_j}\right) = \widehat{\frac{\partial f}{\partial x_j}} = \mathrm{i}\xi_j \mathcal{F} f \text{ pour } j = 1, 2, \ldots, d. \tag{4.6}$$

Il s'agit en fait d'étendre la formule (4.3) valable pour les fonctions de $\mathcal{D}(\mathbb{R}^d)$ aux fonctions de $L^1(\mathbb{R}^d)$ dont les dérivées partielles sont également dans $L^1(\mathbb{R}^d)$. Supposons pour fixer les idées que $j = 1$. En posant $F(x) = f(x)\mathrm{e}^{-\mathrm{i}x\cdot\xi}$, comme les fonctions f et $\partial f/\partial x_1$ sont intégrables dans $\mathbb{R}^d$, c'est également le cas pour F et $\partial F/\partial x_1$. D'après le théorème de Fubini, pour presque tout $(x_2, x_3, \ldots, x_d)$, les fonctions d'une variable $x_1 \mapsto F(x_1, x_2, \ldots, x_d)$ et $x_1 \mapsto \partial F/\partial x_1(x_1, x_2, \ldots, x_d)$ sont intégrables dans $\mathbb{R}$. On en déduit d'après le lemme 3.12 que pour presque tout $(x_2, x_3, \ldots, x_d)$, la fonction $x_1 \mapsto F(x_1, x_2, \ldots, x_d)$ est continue et a des limites nulles à l'infini, avec

$$F(X_1, x_2, \ldots, x_d) = \int_{-\infty}^{X_1} \frac{\partial}{\partial x_1} F(x_1, x_2, \ldots, x_d)\, \mathrm{d}x_1.$$

En faisant $X_1 \to +\infty$ dans l'égalité précédente, on obtient

$$\int_{\mathbb{R}} \frac{\partial}{\partial x_1} \left(f(x) \mathrm{e}^{-\mathrm{i}x\cdot\xi} \right) \mathrm{d}x_1 = 0,$$

et enfin

$$\int_{\mathbb{R}} \left(-i\xi_1 f(x) + \frac{\partial f}{\partial x_1}(x) \right) \mathrm{e}^{-\mathrm{i}x\cdot\xi} \, \mathrm{d}x_1 = 0.$$

En intégrant cette identité par rapport à $(x_2, x_3, \ldots, x_d)$, on obtient finalement la formule (4.6). En appliquant ce résultat aux dérivées successives de f jusquà l'ordre m, le résultat est démontré. □

On énonce maintenant un second théorème, complémentaire du précédent, qui précise celui de Riemann Lebesgue dans le cas de fonctions décroissant plus vite à l'infini que les fonctions simplement L^1.

Théorème 4.7 *Si $x^\alpha f(x)$ est intégrable pour tout $\alpha \in \mathbb{N}^d$ tel que $|\alpha| \leq m$, alors $\mathcal{F}f$ est de classe $\mathcal{C}^m$ et :*

$$\partial^\alpha (\mathcal{F}f)(\xi) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} (-\mathrm{i}x)^\alpha f(x) \mathrm{e}^{-\mathrm{i}x\cdot\xi} \, \mathrm{d}x = \mathcal{F}\left((-\mathrm{i}x)^\alpha f \right)(\xi). \tag{4.7}$$

Démonstration. On commence par remarquer que si f est une fonction intégrable et $x_j f$ l'est également, en utilisant le théorème de dérivation sous le signe somme il vient que $\widehat{f}$ admet une dérivée partielle par rapport à ξ_j et que

$$\frac{\partial \widehat{f}}{\partial \xi_j}(\xi) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} (-\mathrm{i}x_j) f(x) \mathrm{e}^{-\mathrm{i}x\cdot\xi} \, \mathrm{d}x = \mathcal{F}(-\mathrm{i}x_j f)(\xi). \tag{4.8}$$

Cette dérivée partielle est de plus continue par le théorème de convergence dominée. On termine la preuve du théorème par application successive du théorème de dérivation sous le signe somme jusqu'à l'ordre m. □

Les deux théorèmes 4.6 et 4.7 nous indiquent que la transformation de Fourier "échange" régularité et décroissance à l'infini. De manière plus précise, plus une fonction est régulière, plus sa transformée de Fourier est décroissante à l'infini. Inversement, plus la fonction décroît vite à l'infini, plus sa transformée de Fourier est régulière.

Nous complétons ces deux résultats par l'étude de l'effet d'une translation et d'une dilatation sur la transformation de Fourier. On a d'abord le

Lemme 4.8 *Soient $f \in L^1(\mathbb{R}^d)$ et $a \in \mathbb{R}^d$. Alors :*

$$\mathcal{F}(\tau_a f) = \mathrm{e}^{-\mathrm{i}a\cdot\xi} \mathcal{F}f \tag{4.9}$$

$$\mathcal{F}\left(\mathrm{e}^{\mathrm{i}a\cdot x} f \right) = \tau_a \mathcal{F}f, \tag{4.10}$$

où τ_a est l'opérateur défini par

$$\tau_a f(x) = f(x - a).$$

DÉMONSTRATION. Ces deux identités s'obtiennent sans aucune difficulté. On a par exemple

$$\mathcal{F}(\tau_a f)(\xi) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(x-a) \mathrm{e}^{-\mathrm{i}x\cdot\xi} \, \mathrm{d}x = \frac{1}{(2\pi)^{d/2}} \mathrm{e}^{-\mathrm{i}a\xi} \int_{\mathbb{R}^d} f(y) \mathrm{e}^{-\mathrm{i}y\cdot\xi} \, \mathrm{d}y.$$

□

Enfin, on vérifie facilement le

Lemme 4.9 *Soit $f \in L^1(\mathbb{R}^d)$ et $\widehat{f}$ sa transformée de Fourier. Alors :*

$$\widehat{f_\lambda}(\xi) = \frac{1}{|\lambda|^d} \widehat{f}\left(\frac{\xi}{\lambda}\right) \quad \forall \lambda \in \mathbb{R}^*,$$

où f_λ désigne la fonction définie par :

$$f_\lambda(x) = f(\lambda x).$$

4.1.3 Exemples

Montrons quelques exemples classiques de transformées de Fourier. Notons que le premier sera utilisé plus tard dans la construction de la transformée de Fourier des fonctions de carré intégrable.

Exemple 4.10 Considérons la fonction gaussienne définie par $f_1(x) = \mathrm{e}^{-x^2/2}$. Il s'agit bien entendu d'une fonction intégrable et l'on a :

$$\widehat{f_1}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{-x^2/2} \mathrm{e}^{-\mathrm{i}x\xi} \, \mathrm{d}x.$$

De plus, on peut appliquer les théorèmes 4.6 et 4.7 jusqu'à un ordre m arbitraire car f_1 est de classe $\mathcal{C}^\infty$, toutes ses dérivées sont intégrables et $x^j f_1(x)$ est intégrable pour tout j. Il en résulte que $\widehat{f_1}$ est de classe $\mathcal{C}^\infty$ et qu'elle tend vers 0 plus vite que toute puissance de $1/\xi$.

Pour calculer explicitement $\widehat{f_1}$, on peut remarquer, en utilisant le théorème 4.7 et en effectuant une intégration par parties, que :

$$(\widehat{f_1})'(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (-\mathrm{i}x) \mathrm{e}^{-x^2/2} \mathrm{e}^{-\mathrm{i}x\xi} \, \mathrm{d}x = -\xi \widehat{f_1}(\xi).$$

On a donc :

$$\widehat{f_1}(\xi) = C\mathrm{e}^{-\xi^2/2}$$

avec

$$C = \widehat{f_1}(0) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{-x^2/2} \, \mathrm{d}x = 1.$$

On a finalement montré que

$$\mathcal{F}\left(\mathrm{e}^{-x^2/2}\right) = \mathrm{e}^{-\xi^2/2}. \tag{4.11}$$

Autrement dit, la fonction f_1 est égale à sa transformée de Fourier. En dimension supérieure, on obtient directement, en utilisant la remarque 4.3, que :

$$\mathcal{F}\left(\mathrm{e}^{-|x|^2/2}\right) = \mathrm{e}^{-|\xi|^2/2}. \tag{4.12}$$

Enfin, en utilisant successivement les lemmes 4.8 et 4.9, on obtient que pour $x_0 \in \mathbb{R}^d$ et $\sigma > 0$:

$$\mathcal{F}\left(\mathrm{e}^{-|x-x_0|^2/2\sigma^2}\right) = \sigma^d\, \mathrm{e}^{-\mathrm{i}x_0\cdot\xi}\, \mathrm{e}^{-\sigma^2|\xi|^2/2}. \tag{4.13}$$

•

Exemple 4.11 Considérons maintenant l'exemple de la fonction $f_2(x) = \mathrm{e}^{-|x|}$. Cette fonction est intégrable, ainsi que sa dérivée au sens des distributions :

$$f_2'(x) = -\mathrm{sgn}(x)\, \mathrm{e}^{-|x|},$$

où $\mathrm{sgn}(x)$ désigne le signe de x. En revanche sa dérivée seconde n'est pas une fonction. On conclut par le théorème 4.6 que $\widehat{f_2}$ décroît plus vite à l'infini que $1/|\xi|$. D'autre part, $x^j f_2$ est intégrable pour tout j, ce qui implique par le théorème 4.7 que $\widehat{f_2}$ est une fonction de classe $\mathcal{C}^\infty$. On a

$$\widehat{f_2}(\xi) = \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{0} \mathrm{e}^{-\mathrm{i}x\xi}\, \mathrm{e}^{x}\, \mathrm{d}x + \int_{0}^{+\infty} \mathrm{e}^{-\mathrm{i}x\xi}\, \mathrm{e}^{-x}\, \mathrm{d}x \right),$$

soit

$$\widehat{f_2}(\xi) = \frac{1}{\sqrt{2\pi}} \left(\frac{1}{1-\mathrm{i}\xi} + \frac{1}{1+\mathrm{i}\xi} \right) = \sqrt{\frac{2}{\pi}} \frac{1}{1+\xi^2}.$$

•

Exemple 4.12 Soit maintenant f_3 la fonction caractéristique de l'intervalle $]-1,1[$. Cette fonction est intégrable, mais sa dérivée n'est pas une fonction. Le théorème 4.6 ne permet donc pas d'améliorer la conclusion du théorème de Riemann-Lebesgue, à savoir que $\widehat{f_3}$ tend vers 0 à l'infini. D'autre part, $x^j f_3$ est intégrable pour tout j, ce qui implique à nouveau que $\widehat{f_3}$ est une fonction de classe $\mathcal{C}^\infty$. On obtient directement :

$$\widehat{f_3}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{1}^{1} \mathrm{e}^{-\mathrm{i}x\xi}\, \mathrm{d}x = \sqrt{\frac{2}{\pi}} \frac{\sin\xi}{\xi}.$$

On remarque en particulier sur cet exemple que la transformée de Fourier d'une fonction intégrable n'est en général pas intégrable. Plus généralement, si f est la fonction caractéristique du pavé $]-1,1[^d$, alors il résulte de la remarque 4.3 que :

$$\widehat{f}(\xi) = \left(\frac{2}{\pi}\right)^{d/2} \frac{\sin\xi_1}{\xi_1} \frac{\sin\xi_2}{\xi_2} \cdots \frac{\sin\xi_d}{\xi_d}.$$

•

4.1.4 Convolution et transformation de Fourier dans L^1

Dans ce paragraphe, nous allons établir le résultat fondamental suivant annoncé dans l'introduction : la transformée de Fourier d'un produit de convolution est égal au produit des transformées de Fourier. Plus précisément, on a le

Théorème 4.13 *Soient f et g deux fonctions de $L^1(\mathbb{R}^d)$. Alors*

$$\mathcal{F}(f * g) = (2\pi)^{d/2}\mathcal{F}(f)\mathcal{F}(g). \tag{4.14}$$

Démonstration. On a déjà vu que, si f et g sont dans $L^1(\mathbb{R}^d)$, leur produit de convolution appartient également à $L^1(\mathbb{R}^d)$ (voir le théorème 2.23). On peut donc calculer sa transformée de Fourier qui s'écrit :

$$\mathcal{F}(f * g)(\xi) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(y)g(x-y)\,\mathrm{d}y \right) \mathrm{e}^{-\mathrm{i}x\cdot\xi}\mathrm{d}x.$$

Le théorème de Fubini s'applique alors sans difficulté, car la fonction $f(y)g(x-y)\mathrm{e}^{-\mathrm{i}x\cdot\xi}$ est intégrable sur $\mathbb{R}^{2d}$ (nous l'avons vérifié dans la démonstration du théorème 2.23) et l'on a donc :

$$\mathcal{F}(f * g)(\xi) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(y) \left(\int_{\mathbb{R}^d} g(x-y)\mathrm{e}^{-\mathrm{i}x\cdot\xi} dx \right) \mathrm{d}y.$$

Le changement de variable $z = x - y$ donne finalement :

$$\begin{aligned}\mathcal{F}(f * g)(\xi) &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(y) \left(\int_{\mathbb{R}^d} g(z)\mathrm{e}^{-\mathrm{i}(z+y)\cdot\xi}\,\mathrm{d}z \right) \mathrm{d}y \\ &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(y)\mathrm{e}^{-\mathrm{i}y\cdot\xi}\,\mathrm{d}y \int_{\mathbb{R}^d} g(z)\mathrm{e}^{-\mathrm{i}z\cdot\xi}\,\mathrm{d}z,\end{aligned}$$

ce qui termine la preuve. □

4.2 Le cas des fonctions de carré intégrable

Dans ce paragraphe nous allons étendre la notion de transformation de Fourier aux fonctions dont le carré est intégrable. De telles fonctions ne sont en général pas intégrables, et par conséquent la définition 4.1 vue au paragraphe précédent ne leur est pas applicable. Nous commençons par décrire l'ensemble des fonctions de carré intégrable dans un domaine ouvert quelconque, avant d'introduire leur tranformée de Fourier dans le cas où cet ouvert est l'espace $\mathbb{R}^d$ entier.

4.2.1 Définition et propriétés de L^2

On considère donc un ouvert Ω de $\mathbb{R}^d$.

Définition 4.14 *Une fonction f définie sur Ω à valeurs complexes est dite de carré intégrable si $f^2 \in L^1(\Omega)$. On pose alors :*

$$||f||_{L^2} = \left(\int_\Omega |f(x)|^2 \, \mathrm{d}x \right)^{1/2}.$$

On note $L^2(\Omega)$ l'ensemble des fonctions de carré intégrable sur Ω.

Proposition 4.15 *L'ensemble $L^2(\Omega)$ satisfait les propriétés suivantes :*

(i) *$L^2(\Omega)$ est un espace vectoriel et $||f||_{L^2}$ est une norme sur $L^2(\Omega)$. En particulier, si $f \in L^2(\Omega)$ et $g \in L^2(\Omega)$, alors la somme $f+g$ appartient à $L^2(\Omega)$ et on a (inégalité de Minkowski)*

$$\left(\int_\Omega |f+g|^2 \, \mathrm{d}x \right)^{1/2} \leq \left(\int_\Omega |f|^2 \, \mathrm{d}x \right)^{1/2} + \left(\int_\Omega |g|^2 \, \mathrm{d}x \right)^{1/2}.$$

(ii) *Si $f \in L^2(\Omega)$ et $g \in L^2(\Omega)$, alors le produit fg appartient à $L^1(\Omega)$ et on a (inégalité de Cauchy-Schwarz)*

$$\int_\Omega |fg| \, \mathrm{d}x \leq \left(\int_\Omega |f|^2 \, \mathrm{d}x \right)^{1/2} \left(\int_\Omega |g|^2 \, \mathrm{d}x \right)^{1/2}.$$

(iii) *L'application*

$$(f,g) \to \int_\Omega f\bar{g} \, \mathrm{d}x$$

est un produit scalaire sur $L^2(\Omega)$ (voir la définition 5.8 au chapitre suivant) que l'on notera $(f,g)_{L^2}$.

Démonstration. Commençons par montrer que $L^2(\Omega)$ est un espace vectoriel. Soient f et g de carrés intégrables sur Ω. Pour tout $x \in \Omega$, on a

$$2|f(x)||g(x)| \leq |f(x)|^2 + |g(x)|^2,$$

d'où

$$|f(x)+g(x)|^2 = |f(x)|^2 + |g(x)|^2 + 2\,\mathfrak{Re}(f\bar{g}) \leq 2\left(|f(x)|^2 + |g(x)|^2\right).$$

On déduit de la première inégalité que fg est intégrable et de la seconde que $f+g$ est de carré intégrable, ce qui implique que $L^2(\Omega)$ est un espace vectoriel.

Pour démontrer l'inégalité de Cauchy-Schwarz, on peut remarquer que

$$\int_\Omega (|f| + t|g|)^2 \, \mathrm{d}x \geq 0$$

pour tout réel t. Or le membre de gauche est un polynôme de degré 2 en t, par conséquent son discriminant est nécessairement négatif ou nul, ce qui s'écrit

$$\left(\int_\Omega |fg|\,\mathrm{d}x\right)^2 - \left(\int_\Omega |f|^2\,\mathrm{d}x\right)\left(\int_\Omega |g|^2\,\mathrm{d}x\right) \leq 0,$$

d'où le résultat.

Pour démontrer l'inégalité de Minkowski, on écrit

$$\int_\Omega |f+g|^2\,\mathrm{d}x = \int_\Omega (f+g)\bar{f}\,\mathrm{d}x + \int_\Omega (f+g)\bar{g}\,\mathrm{d}x,$$

et on applique l'inégalité de Cauchy-Schwarz à chacun des deux termes. Le point (iii) de la proposition est immédiat. □

La topologie déduite de la norme

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

est appelée *topologie de la convergence en moyenne quadratique*. Nous allons maintenant établir le résultat essentiel suivant :

Théorème 4.16 *Soit (f_n) une suite de Cauchy dans $L^2(\Omega)$. Alors il existe $f \in L^2(\Omega)$ tel que*

$$||f - f_n||_{L^2} \to 0 \text{ quand } n \to +\infty.$$

L'espace $L^2(\Omega)$ est donc complet pour la topologie de la convergence en moyenne quadratique.

Démonstration. Soit (f_n) une suite de Cauchy dans $L^2(\Omega)$. Alors on peut construire une sous-suite extraite que nous noterons encore (f_n) telle que pour $n \geq 1$

$$||f_n - f_{n-1}||_{L^2} \leq \frac{1}{2^n}.$$

On pose alors

$$g_N(x) = \sum_{n=1}^{N} |f_n(x) - f_{n-1}(x)|,$$

d'où

$$||g_N||_{L^2} \leq \sum_{n=1}^{N} ||f_n - f_{n-1}||_{L^2} \leq \sum_{n=1}^{N} \frac{1}{2^n} = 1 - \frac{1}{2^N} \leq 1.$$

En posant $g(x) = \lim_{N\to+\infty} g_N(x)$, le théorème de convergence monotone (théorème 2.16) appliqué à la suite (g_N^2), qui est positive et croissante, donne

$$\int_\Omega g^2(x)\,\mathrm{d}x = \lim_{N\to+\infty} \int_\Omega g_N^2(x)\,\mathrm{d}x.$$

On en déduit que la fonction g est dans $L^2(\Omega)$. En particulier, l'ensemble A des points x tels que $g(x) = +\infty$ est de mesure nulle, ce qui implique qu'en tout point x du complémentaire de A, la suite numérique $(g_N(x))$ converge vers $g(x)$. Autrement dit, la suite (g_N) converge presque partout vers g. En remarquant que pour $p \geq 1$

$$|f_{N+p}(x) - f_N(x)| \leq \sum_{n=N+1}^{N+p} |f_n(x) - f_{n-1}(x)| = g_{N+p}(x) - g_N(x), \tag{4.15}$$

on en déduit en utilisant le critère de Cauchy dans $\mathbb{R}$ que la suite de fonctions (f_n) converge presque partout vers une fonction f. De plus, en passant à la limite sur p dans (4.15), on vérifie que pour presque tout $x \in \Omega$:

$$|f(x) - f_N(x)| \leq g(x).$$

Par conséquent, $(f - f_N)$ appartient à $L^2(\Omega)$ et donc f également. On obtient alors que f_N converge vers f dans $L^2(\Omega)$ en appliquant le théorème de convergence dominée de Lebesgue à la fonction $(f - f_N)^2$. On a ainsi montré qu'il existe une suite extraite de (f_n) qui converge vers un élément f dans $L^2(\Omega)$.

Il reste à montrer que toute la suite converge vers f dans $L^2(\Omega)$. On raisonne par l'absurde en supposant le contraire. On peut alors extraire une sous-suite de (f_n) dont les éléments restent éloignés de f d'une distance $\varepsilon > 0$ fixée. Cela contredit le fait que la suite (f_n) est de Cauchy, ce qui termine la démonstration. □

Ce théorème montre que $L^2(\Omega)$ est un *espace de Hilbert* (dont nous verrons la définition au chapitre 5).

Nous aurons souvent recours dans la suite au résultat de densité suivant, qui est le pendant du théorème 4.4 (lorsque $\Omega = \mathbb{R}^d$) pour les fonctions L^2 :

Théorème 4.17 *$\mathcal{D}(\Omega)$ est dense dans $L^2(\Omega)$.*

4.2.2 Transformation de Fourier dans L^2

On se limite désormais à $\Omega = \mathbb{R}^d$. Remarquons tout d'abord qu'une fonction f de carré intégrable dans $\mathbb{R}^d$ appartient à $L^1_{\text{loc}}(\mathbb{R}^d)$. En effet, si $f \in L^2(\mathbb{R}^d)$, l'inégalité de Cauchy-Schwarz montre que pour tout compact $K \subset \mathbb{R}^d$, on a :

$$\int_K |f| \, \mathrm{d}x \leq \|f\|_{L^2(\mathbb{R}^d)} \, (\mathrm{mes}(K))^{1/2},$$

où $\mathrm{mes}(K)$ est la mesure de Lebesgue du compact K. En revanche, on peut trouver des fonctions de $L^2(\mathbb{R}^d)$ qui n'appartiennent pas à $L^1(\mathbb{R}^d)$, et donc on ne peut pas calculer la transformée de Fourier d'une fonction de carré intégrable en utilisant la formule (4.2). Nous allons dans ce qui suit définir la transformée de Fourier d'une manière indirecte en utilisant le théorème de densité 4.17.

Nous commençons par remarquer que pour f et g deux fonctions de $L^1(\mathbb{R}^d)$, les fonctions $f\widehat{g}$ et $\widehat{f}g$ sont dans $L^1(\mathbb{R}^d)$ et

$$\int_{\mathbb{R}^d} f\,\widehat{g}\,\mathrm{d}x = \int_{\mathbb{R}^d} \widehat{f}\,g\,\mathrm{d}x. \tag{4.16}$$

On a effet

$$\int_{\mathbb{R}^d} f\,\widehat{g}\,\mathrm{d}x = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(x) \left(\int_{\mathbb{R}^d} g(y)\mathrm{e}^{-\mathrm{i}x\cdot y}\,\mathrm{d}y \right) \mathrm{d}x,$$

et la fonction $f(x)g(y)\mathrm{e}^{-\mathrm{i}x\cdot y}$ étant intégrable dans $\mathbb{R}^{2d}$, on obtient en utilisant le théorème de Fubini que

$$\int_{\mathbb{R}^d} f\,\widehat{g}\,\mathrm{d}x = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(x)\mathrm{e}^{-\mathrm{i}y\cdot x}\,\mathrm{d}x \right) g(y)\,\mathrm{d}y.$$

Nous aurons dans la suite besoin du lemme technique suivant :

Lemme 4.18 *Soient f et G_1 deux fonctions intégrables avec*

$$\int_{\mathbb{R}^d} G_1(x)\,\mathrm{d}x = 1.$$

On pose $G_\varepsilon(x) = \varepsilon^{-d}G_1(x/\varepsilon)$ pour $\varepsilon > 0$ et $f_\varepsilon = G_\varepsilon * f$. Alors f_ε converge vers f dans $L^1(\mathbb{R}^d)$ lorsque $\varepsilon \to 0$.

Démonstration. On commence par traiter le cas où $f = \varphi \in \mathcal{D}(\mathbb{R}^d)$, de support inclus dans la boule $B(0,R)$. En faisant le changement de variable $y = \varepsilon s$ dans la formule

$$\varphi_\varepsilon(x) = \int_{\mathbb{R}^d} \varphi(x-y)G_\varepsilon(y)\,\mathrm{d}y,$$

il vient que

$$\varphi_\varepsilon(x) = \int_{\mathbb{R}^d} \varphi(x-\varepsilon s)G_1(s)\,\mathrm{d}s.$$

On en déduit

$$\varphi_\varepsilon(x) - \varphi(x) = \int_{\mathbb{R}^d} (\varphi(x-\varepsilon s) - \varphi(x))G_1(s)\,\mathrm{d}s,$$

d'où

$$||\varphi_\varepsilon - \varphi||_{L^1(\mathbb{R}^d)} \leq \int_{\mathbb{R}^{2d}} |\varphi(x-\varepsilon s) - \varphi(x)|\,|G_1(s)|\,\mathrm{d}x\,\mathrm{d}s,$$

$$||\varphi_\varepsilon - \varphi||_{L^1(\mathbb{R}^d)} \leq \int_{\mathbb{R}^d} \int_{|x|\leq R+1} |\varphi(x-\varepsilon s) - \varphi(x)|\,|G_1(s)|\,\mathrm{d}x\,\mathrm{d}s$$

$$+ \int_{\mathbb{R}^d} \int_{|x|\geq R+1} |\varphi(x-\varepsilon s)|\,|G_1(s)|\,\mathrm{d}x\,\mathrm{d}s = I_1 + I_2.$$

Dans la seconde intégrale, on a $\varphi(x-\varepsilon s)=0$ lorsque $|\varepsilon s|\leq 1$, donc

$$I_2 \leq \int_{|s|\geq 1/\varepsilon}\int_{\mathbb{R}^d} |\varphi(x-\varepsilon s)|\,|G_1(s)|\,\mathrm{d}x\,\mathrm{d}s,$$

$$I_2 \leq \int_{|s|\geq 1/\varepsilon} |G_1(s)| \left(\int_{B(\varepsilon s,R)} |\varphi(x-\varepsilon s)|\,\mathrm{d}x\right)\,\mathrm{d}s,$$

et donc finalement

$$I_2 \leq ||\varphi||_{L^1(\mathbb{R}^d)} \int_{|s|\geq 1/\varepsilon} |G_1(s)|\,\mathrm{d}s,$$

cette dernière intégrale tendant vers 0 lorsque $\varepsilon\to 0$. Pour ce qui est de l'intégrale I_1, la fonction de (x,s) intégrée tend vers 0 presque partout en utilisant la continuité de φ. Elle est par ailleurs majorée par $2||\varphi||_{L^\infty}|G_1(s)|$ qui est intégrable sur $B(0,R+1)\times\mathbb{R}^d$. Par le théorème de convergence dominée de Lebesgue, on conclut donc que I_1 tend également vers 0 lorsque $\varepsilon\to 0$.

On considère maintenant une fonction de $L^1(\mathbb{R}^d)$. Pour $\alpha>0$, par densité de $\mathcal{D}(\mathbb{R}^d)$ dans $L^1(\mathbb{R}^d)$ on peut trouver $\varphi\in\mathcal{D}(\mathbb{R}^d)$ telle que $||f-\varphi||_{L^1}\leq\alpha/3$ et $||f-\varphi||_{L^1}||G_1||_{L^1}\leq\alpha/3$. En utilisant l'inégalité triangulaire, on a

$$||f*G_\varepsilon - f||_{L^1(\mathbb{R}^d)} \leq ||(f-\varphi)*G_\varepsilon||_{L^1(\mathbb{R}^d)} + ||\varphi*G_\varepsilon-\varphi||_{L^1(\mathbb{R}^d)} + ||\varphi-f||_{L^1(\mathbb{R}^d)}.$$

D'après le théorème 2.23, le premier terme se majore par $||f-\varphi||_{L^1}||G_1||_{L^1}$, soit par $\alpha/3$, il en est de même pour le troisième. Le second est majoré par $\alpha/3$ dès que ε est suffisamment petit d'après la première partie du théorème. On a donc finalement

$$||f*G_\varepsilon - f||_{L^1(\mathbb{R}^d)} \leq \alpha$$

dès que ε est suffisamment petit, d'où la conclusion. □

On utilise le lemme précédent pour montrer la formule de réciprocité pour les fonctions de L^1 dont la transformée de Fourier est dans L^1.

Proposition 4.19 (Formule de réciprocité) *Soit $f\in L^1(\mathbb{R}^d)$ telle que $\widehat{f}\in L^1(\mathbb{R}^d)$. On a*

$$f = \overline{\mathcal{F}}(\widehat{f}).$$

Démonstration. Si l'on explicite le membre de droite de l'égalité précédente, on obtient

$$\overline{\mathcal{F}}(\widehat{f})(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}y\cdot\xi} f(y)\,\mathrm{d}y\right)\mathrm{e}^{\mathrm{i}x\cdot\xi}\,\mathrm{d}\xi.$$

Nous serions tentés d'inverser l'ordre d'intégration dans cette intégrale double, mais la fonction $(\xi,y)\mapsto \mathrm{e}^{\mathrm{i}(x-y)\cdot\xi}f(y)$ n'est pas intégrable dans $\mathbb{R}^{2d}$. On introduit donc, pour $\varepsilon>0$ et $x\in\mathbb{R}^d$, la fonction $(\xi,y)\mapsto \mathrm{e}^{\mathrm{i}(x-y)\cdot\xi}\mathrm{e}^{-\varepsilon^2|\xi|^2/4}f(y)$, qui est elle intégrable dans $\mathbb{R}^{2d}$, et on peut ainsi considérer

$$I_\varepsilon(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^{2d}} \mathrm{e}^{\mathrm{i}(x-y)\cdot\xi}\mathrm{e}^{-\varepsilon^2|\xi|^2/4} f(y)\,\mathrm{d}y\,\mathrm{d}\xi. \tag{4.17}$$

Par application du théorème de Fubini, en intégrant d'abord par rapport à y on obtient

$$I_\varepsilon(x) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}x\cdot\xi} \mathrm{e}^{-\varepsilon^2|\xi|^2/4} \widehat{f}(\xi)\, \mathrm{d}\xi.$$

En utilisant le théorème de convergence dominée de Lebesgue, étant donnée l'hypothèse que $\widehat{f} \in L^1(\mathbb{R}^d)$, on obtient que $\lim_{\varepsilon\to 0} I_\varepsilon(x) = I(x)$, avec

$$I(x) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}x\cdot\xi} \widehat{f}(\xi)\, \mathrm{d}\xi.$$

On voit donc que $I_\varepsilon(x)$ converge vers $\overline{\mathcal{F}}(\widehat{f})(x)$ lorsque $\varepsilon \to 0$ en chaque point x. Pour $\varphi \in \mathcal{D}(\mathbb{R}^d)$, on a

$$\langle I_\varepsilon - \overline{\mathcal{F}}(\widehat{f}), \varphi\rangle = \int_{\mathbb{R}^d} (I_\varepsilon(x) - \overline{\mathcal{F}}(\widehat{f})(x))\varphi(x)\, \mathrm{d}x.$$

Comme I_ε est majorée au facteur $1/(2\pi)^{d/2}$ près par la constante $||\widehat{f}||_{L^1}$, une nouvelle application du théorème de convergence dominée de Lebesgue montre que I_ε converge vers $\overline{\mathcal{F}}(\widehat{f})$ au sens des distributions.

Si maintenant on intègre d'abord par rapport à ξ dans (4.17), on obtient

$$I_\varepsilon(x) = \int_{\mathbb{R}^d} G_\varepsilon(x-y) f(y)\, \mathrm{d}y,$$

avec

$$G_\varepsilon(z) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}z\cdot\xi} \mathrm{e}^{-\varepsilon^2|\xi|^2/4}\, \mathrm{d}\xi,$$

c'est-à-dire $I_\varepsilon = G_\varepsilon * f$. On voit que G_ε est au facteur $1/(2\pi)^{d/2}$ près la transformée de Fourier de la fonction $\mathrm{e}^{-\varepsilon^2|\xi|^2/4}$. En utilisant la transformée de Fourier de la Gaussienne donnée par (4.13), on obtient sans difficulté que

$$G_\varepsilon(z) = \frac{1}{\varepsilon^d} G_1\left(\frac{z}{\varepsilon}\right),$$

avec

$$G_1(z) = \frac{1}{\pi^{d/2}} \mathrm{e}^{-|z|^2}.$$

La fonction G_1 étant d'intégrale égale à 1, en utilisant le lemme 4.18 on obtient que $G_\varepsilon * f$ converge vers f dans $L^1(\mathbb{R}^d)$. On voit donc que I_ε converge vers f dans $L^1(\mathbb{R}^d)$, et donc en particulier au sens des distributions. Les fonctions f et $\overline{\mathcal{F}}(\widehat{f})$ sont égales en tant que distributions, et donc en tant qu'éléments de $L^1(\mathbb{R}^d)$. □

On en déduit le théorème suivant, qui définit la transformée de Fourier pour les fonctions de carré intégrable.

Théorème 4.20 *Pour $f \in L^2(\mathbb{R}^d)$, si (φ_n) est une suite de fonctions de $\mathcal{D}(\mathbb{R}^d)$ qui converge vers f dans L^2, la suite $(\widehat{\varphi}_n)$ converge dans L^2 vers une limite appelée la transformée de Fourier $\mathcal{F}$ de f, cette limite ne dépendant pas de la suite choisie. On définit de manière analogue la conjuguée de la transformée de Fourier $\overline{\mathcal{F}}$ de f. La transformation de Fourier $\mathcal{F}$ est une isométrie bijective de $L^2(\mathbb{R}^d)$ sur lui-même, d'inverse $\overline{\mathcal{F}}$.*

Démonstration. On voit tout d'abord que $\mathcal{F}$ conserve la norme L^2 des éléments de $\mathcal{D}(\mathbb{R}^d)$. Soient en effet deux éléments f et g de $\mathcal{D}(\mathbb{R}^d)$, et posons $h = \overline{\widehat{g}}$. Il est clair que $h \in L^1(\mathbb{R}^d)$ en vertu des théorèmes 4.6 et 4.7. En utilisant la relation (4.16) pour les fonctions f et h, on a

$$\int_{\mathbb{R}^d} f\,\widehat{h}\,\mathrm{d}x = \int_{\mathbb{R}^d} \widehat{f}\,h\,\mathrm{d}x.$$

On remarque ensuite que

$$\mathcal{F}h = \mathcal{F}\overline{\widehat{g}} = \overline{\overline{\mathcal{F}}\widehat{g}} = \overline{g},$$

cette dernière égalité provenant de la proposition 4.19. On a donc

$$\int_{\mathbb{R}^d} f\,\overline{g}\,\mathrm{d}x = \int_{\mathbb{R}^d} \widehat{f}\,\overline{\widehat{g}}\,\mathrm{d}x.$$

On voit que la transformée de Fourier conserve le produit scalaire, et donc la norme L^2, pour les éléments de $\mathcal{D}(\mathbb{R}^d)$.

On utilise maintenant le fait que $\mathcal{D}(\mathbb{R}^d)$ est dense dans $L^2(\mathbb{R}^d)$. Pour $f \in L^2(\mathbb{R}^d)$, soit donc (φ_n) une suite d'éléments de $\mathcal{D}(\mathbb{R}^d)$ qui converge vers f dans $L^2(\mathbb{R}^d)$. La suite (φ_n) est donc de Cauchy dans $L^2(\mathbb{R}^d)$, c'est-à-dire que pour tout $\varepsilon > 0$, on a $||\varphi_m - \varphi_n||_{L^2} \leq \varepsilon$ pour m, n assez grands. Comme $||\widehat{\varphi}_m - \widehat{\varphi}_n||_{L^2} = ||\varphi_m - \varphi_n||_{L^2}$ d'après ce qui précède, on en déduit que la suite $(\widehat{\varphi}_n)$ est également de Cauchy dans $L^2(\mathbb{R}^d)$, et converge dans $L^2(\mathbb{R}^d)$ puisque cet espace est complet d'après le théorème 4.16. On appelle $\mathcal{F}f$ la limite de cette suite, dont on vérifie facilement qu'elle ne dépend pas de la suite choisie. On peut définir $\overline{\mathcal{F}}f$ de la même manière. En passant à la limite dans les relations

$$||\widehat{\varphi}_n||_{L^2(\mathbb{R}^d)} = ||\varphi_n||_{L^2(\mathbb{R}^d)}, \quad \varphi_n = \overline{\mathcal{F}}(\widehat{\varphi}_n),$$

on obtient les mêmes relations avec f, ce qui permet de conclure que $\mathcal{F} : L^2 \to L^2$ est une isométrie d'inverse $\overline{\mathcal{F}}$.

Il reste à vérifier que la transformée de Fourier définie ainsi sur L^2 étend effectivement celle définie sur L^1 par la formule (4.2), c'est-à-dire que les deux définitions coïncident pour les fonctions de $L^1 \cap L^2$. On utilise à cette fin un raffinement des théorèmes 4.4 et 4.17 qui précise que pour tout $f \in L^1 \cap L^2$, il existe une suite de fonctions $\varphi_n \in \mathcal{D}(\mathbb{R}^d)$ telle que (φ_n) converge vers f simultanément dans L^1 et dans L^2. D'après (4.4), il en résulte que $\widehat{\varphi}_n$ converge uniformément vers la transformée de

Fourier de f calculée par la formule (4.2) (c'est un élément de $C^0(\mathbb{R}^d)$), tandis que $(\widehat{\varphi}_n)$ converge en norme L^2 vers la transformée de Fourier de f définie dans la première partie de la démonstration (c'est un élément de $L^2(\mathbb{R}^d)$). Ces deux limites coïncident donc presque partout. □

Compte tenu de l'importance de ce résultat, nous explicitons la conservation par la transformation de Fourier du produit scalaire dans L^2 (*identité de Parseval*) et de la norme associée (*identité de Plancherel*).

Théorème 4.21 (Identité de Parseval) *Pour tous f et g dans $L^2(\mathbb{R}^d)$, on a*

$$(f, g)_{L^2} = (\widehat{f}, \widehat{g})_{L^2}$$

Théorème 4.22 (Identité de Plancherel) *Pour tout f dans $L^2(\mathbb{R}^d)$, on a*

$$||f||_{L^2} = ||\widehat{f}||_{L^2}.$$

En pratique, lorsque l'on cherche à calculer la transformée de Fourier d'une fonction de $L^2(\mathbb{R}^d)$ qui n'est pas dans $L^1(\mathbb{R}^d)$, il n'est pas nécessaire de construire une suite d'approximations dans $\mathcal{D}(\mathbb{R}^d)$. En effet, on a le

Théorème 4.23 *Soit $f \in L^2(\mathbb{R}^d)$. Alors $\widehat{f}$ est la limite dans $L^2(\mathbb{R}^d)$ lorsque $n \to +\infty$ de la suite de fonctions*

$$\frac{1}{(2\pi)^{d/2}} \int_{|x|\leq n} \mathrm{e}^{-\mathrm{i}x\cdot\xi} f(x)\, \mathrm{d}x.$$

Démonstration. Il suffit de remarquer que la suite (f_n) des fonctions de $L^1 \cap L^2$ égales à f pour $|x| \leq n$ et 0 sinon tend vers f dans $L^2(\mathbb{R}^d)$, en vertu du théorème de convergence dominée de Lebesgue. Donc $\widehat{f_n}$ tend vers $\widehat{f}$ dans $L^2(\mathbb{R}^d)$ d'après l'identité de Plancherel. □

Chapitre 5

Espaces de Hilbert

Ce chapitre est consacré aux espaces de Hilbert dont nous donnons la définition et les principales propriétés. Il s'agit d'une classe particulière d'espaces vectoriels normés complets dont la norme dérive d'un produit scalaire, ce qui leur confère une structure très riche.

L'ensemble des propriétés spécifiques aux espaces de Hilbert repose sur un résultat fondamental : il s'agit du théorème de la projection qui prouve l'existence d'un opérateur de projection sur un sous-ensemble convexe et fermé. Comme on le verra, sa démonstration met en jeu une utilisation conjointe du critère de Cauchy et de propriétés élémentaires telles que le théorème de Pythagore.

Une conséquence importante en est le théorème d'identification de Riesz qui prouve qu'un espace de Hilbert est isomorphe à son dual, ce qui conduit dans de nombreux cas à identifier l'espace et son dual.

Plus généralement, nous démontrerons au chapitre 7 une conséquence très utile du théorème de la projection : il s'agit du théorème de Lax-Milgram qui constitue, dans le cadre des espaces de Sobolev, un outil puissant pour l'étude des problèmes aux limites dits elliptiques.

La notion de base hilbertienne est introduite à la fin du chapitre, et illustrée par l'exemple, très important en pratique, des séries de Fourier. Il est à noter que dans ce chapitre les espaces vectoriels que nous considérons sont construits sur le corps $\mathbb{C}$ des complexes.

5.1 Propriétés élémentaires

Soit H un espace vectoriel sur $\mathbb{C}$. Nous allons tout d'abord poser quelques définitions.

Définition 5.1 *Une application ℓ de H dans $\mathbb{C}$ est dite anti-linéaire si*

$$\ell(u+v) = \ell(u) + \ell(v) \text{ et } \ell(\lambda u) = \overline{\lambda}\ell(u), \ \forall u \in V, \ \forall v \in V, \ \forall \lambda \in \mathbb{C}.$$

Dans la suite, une application anti-linéaire de H dans $\mathbb{C}$ sera encore appelée forme

anti-linéaire sur H, tandis qu'une application linéaire de H dans $\mathbb{C}$ sera encore appelée forme linéaire sur H.

Définition 5.2 *On dit qu'une application de $H \times H$ dans $\mathbb{C}$ qui au couple (u, v) associe $b(u, v)$ est une forme sesquilinéaire sur $H \times H$ si pour tout $u \in H$, l'application $v \to b(v, u)$ est linéaire et l'application $v \to b(u, v)$ est anti-linéaire.*

On vérifie sans difficulté le

Lemme 5.3 (Identité du parallélogramme) *Soit b une forme sesquilinéaire sur $H \times H$. Alors pour tout couple (u, v) de $H \times H$,*

$$b(u + v, u + v) + b(u - v, u - v) = 2(b(u, u) + b(v, v)).$$

Définition 5.4 *Une forme sesquilinéaire b sur $H \times H$ est dite hermitienne si pour tout couple (u, v) de $H \times H$,*

$$b(u, v) = \overline{b(v, u)}.$$

Bien entendu, si b est une forme sesquilinéaire hermitienne sur H, alors $b(u, u) \in \mathbb{R}$ pour tout $u \in H$.

Définition 5.5 *Une forme sesquilinéaire hermitienne b sur $H \times H$ est dite positive si pour tout $u \in H$, $b(u, u) \geq 0$.*

Nous allons maintenant démontrer deux propriétés très utiles des formes sesquilinéaires hermitiennes positives.

Lemme 5.6 (Inégalité de Cauchy-Schwarz) *Soit b une forme sesquilinéaire hermitienne positive b sur $H \times H$. Alors, pour tout couple (u, v) de $H \times H$,*

$$|b(u, v)| \leq b(u, u)^{1/2}\, b(v, v)^{1/2}.$$

Démonstration. Il suffit de remarquer que pour tout t réel, $b(u + tv, u + tv) \geq 0$. Ceci s'écrit aussi :

$$b(u, u) + 2t\, \mathfrak{Re}(b(u, v)) + t^2 b(v, v) \geq 0.$$

Il s'agit d'un polynôme de degré 2 en t qui est positif ou nul pour tout réel t. On peut donc affirmer que son discriminant est négatif ou nul, ce qui s'écrit :

$$(\mathfrak{Re}\ b(u, v))^2 \leq b(u, u)\, b(v, v).$$

Si $b(u, v)$ est réel la démonstration est terminée. Sinon, on choisit $\lambda \in \mathbb{C}$ de module 1 tel que $b(u, \lambda v)$ soit réel (c'est toujours possible) et on pose $t = \lambda s$. On raisonne alors de la même façon sur la variable réelle s. □

De l'inégalité de Cauchy-Schwarz, on déduit immédiatement le

Lemme 5.7 (Inégalité de Minkowski) *Soit b est une forme sesquilinéaire hermitienne positive b sur $H \times H$. Alors, pour tout couple (u, v) de $H \times H$,*

$$b(u+v, u+v)^{1/2} \leq b(u,u)^{1/2} + b(v,v)^{1/2}.$$

DÉMONSTRATION. En effet, on a

$$b(u+v, u+v) = b(u+v, u) + b(u+v, v).$$

Il suffit alors d'appliquer l'inégalité de Cauchy-Schwarz aux deux termes à droite de l'égalité. □

Enfin, on a la

Définition 5.8 *Une forme sesquilinéaire hermitienne positive b sur $H \times H$ est dite définie si la propriété suivante est satisfaite : si $u \in H$ est tel que $b(u,u) = 0$, alors $u = 0$. On dit alors que b est un* produit scalaire *sur H.*

Nous pouvons maintenant montrer qu'un produit scalaire sur H permet de définir une norme sur H (voir la définition 1.3).

Lemme 5.9 *Si $(u, v)_H$ est un produit scalaire sur H, alors l'application*

$$u \to (u,u)_H^{1/2}$$

est une norme sur H.

DÉMONSTRATION. Il suffit de remarquer que l'inégalité de Minkowski n'est rien d'autre que l'inégalité triangulaire ; les deux autres propriétés sont immédiates. □

Ceci nous permet de poser la

Définition 5.10 *Un espace de Hilbert est un espace vectoriel muni d'un produit scalaire et qui est complet pour la norme associée.*

Nous avons déjà vu un exemple d'espace de Hilbert : si Ω est un ouvert de $\mathbb{R}^n$, l'espace $L^2(\Omega)$ est un espace de Hilbert (voir le paragraphe 4.2.1). Dans le chapitre 6, nous introduirons d'autres espaces de Hilbert très utiles pour l'étude des équations aux dérivées partielles.

Un autre exemple classique est l'espace ℓ^2 constitué des suites de nombres complexes $a = (a_n)_{n\in\mathbb{N}}$ telles que

$$\sum_{n\in\mathbb{N}} |a_n|^2 < +\infty.$$

Proposition 5.11 *L'espace ℓ^2, muni du produit scalaire*

$$(a,b) = \sum_{n\in\mathbb{N}} a_n\overline{b_n},$$

est un espace de Hilbert.

DÉMONSTRATION. Remarquons tout d'abord que si a et b sont dans ℓ^2, alors $|a_n + b_n|^2 \leq 2(|a_n|^2 + |b_n|^2)$, donc $(a+b) \in \ell^2$. Autrement dit, ℓ^2 est un espace vectoriel. Par ailleurs, pour tout n, $2|a_nb_n| \leq |a_n|^2 + |b_n|^2$, donc la série $\sum a_n\overline{b_n}$ est absolument convergente. On peut ainsi munir ℓ^2 du produit scalaire défini dans l'énoncé de la proposition.

Montrons maintenant que ℓ^2 muni de ce produit scalaire est un espace vectoriel complet. Soit $a^j = (a_n^j)$ une suite de Cauchy dans ℓ^2. Alors, pour tout $\varepsilon > 0$, on a pour j et k assez grands $\sum |a_n^j - a_n^k|^2 < \varepsilon^2$. Il en résulte que, pour chaque n fixé, la suite $(a_n^j)_j$ est de Cauchy dans $\mathbb{C}$ et converge donc vers un complexe que nous notons a_n.

Montrons enfin que la suite a de terme général a_n est dans ℓ^2 et que la suite a^j converge vers a dans ℓ^2. On a pour j et k assez grands et pour tout N

$$\sum_{n=0}^{N} |a_n^j - a_n^k|^2 < \varepsilon^2.$$

En passant à la limite sur k dans l'inégalité précédente, on trouve que pour j assez grand, on a pour tout N, $\sum_{n=0}^{N} |a_n^j - a_n|^2 < \varepsilon^2$, et donc pour j assez grand

$$\sum_{n=0}^{+\infty} |a_n^j - a_n|^2 < \varepsilon^2.$$

Ce qui précède prouve que la suite $a - a^j$ appartient à ℓ^2, par conséquent $a \in \ell^2$, et enfin que la suite a^j converge vers a dans ℓ^2. □

5.2 Le théorème de projection

A partir de maintenant, H désigne un espace de Hilbert. Son produit scalaire et la norme associée sont notés (u,v) et $\|u\|$.

On rappelle tout d'abord la

Définition 5.12 *On dit qu'un sous-ensemble K de H est convexe si pour tout couple (u,v) de $K \times K$ et pour tout réel t compris entre 0 et 1, $(1-t)u + tv$ appartient à K.*

Nous pouvons maintenant énoncer le résultat fondamental suivant :

Théorème 5.13 (de projection) *Soit $K \subset H$ un convexe fermé non vide.*

(i) *Pour tout u dans H, il existe un unique élément de K que l'on notera Pu tel que :*

$$||u - Pu|| \leq ||u - v||, \ \forall v \in K.$$

(ii) *De plus, Pu est caractérisé par la propriété suivante :*

$$Pu \in K \text{ et } \mathfrak{Re}(u - Pu, v - Pu) \leq 0 \quad \forall v \in K. \tag{5.1}$$

Pu est appelé la projection de u sur K et vérifie $P^2u = Pu$.

Démonstration. On commence par montrer l'existence de Pu : posons

$$d = \inf_{v \in K} ||u - v||$$

et considérons une suite minimisante (v_n), c'est-à-dire

$$v_n \in K \text{ et } ||u - v_n|| \to d \text{ quand } n \to +\infty.$$

En appliquant l'identité du parallélogramme à $a = u - v_n$ et $b = u - v_m$, on trouve :

$$||2u - v_n - v_m||^2 + ||v_m - v_n||^2 = 2\left(||u - v_n||^2 + ||u - v_m||^2\right).$$

Comme K est convexe, $(v_n + v_m)/2 \in K$ et par définition de d, on a

$$||2u - v_n - v_m||^2 = 4||u - \frac{v_n + v_m}{2}||^2 \geq 4d^2.$$

Par conséquent

$$||v_m - v_n||^2 \leq 2\left(||u - v_n||^2 + ||u - v_m||^2\right) - 4d^2.$$

Ceci prouve que la suite (v_n) est de Cauchy dans H. Comme H est complet, elle converge donc vers un élément de H que nous notons Pu tel que $||u - Pu|| = d$. Enfin, comme K est fermé, $Pu \in K$.

Montrons maintenant l'unicité de Pu : supposons qu'il existe deux éléments de K, notés Pu et $\tilde{P}u$, tels que :

$$||u - Pu|| = ||u - \tilde{P}u|| = d.$$

Alors, en appliquant l'identité du parallélogramme à $a = u - Pu$ et $b = u - \tilde{P}u$, on trouve :

$$||2u - Pu - \tilde{P}u||^2 + ||Pu - \tilde{P}u||^2 = 4d^2.$$

En utilisant à nouveau la convexité de K et la définition de d, on voit que :

$$||2u - Pu - \tilde{P}u||^2 = 4||u - \frac{Pu + \tilde{P}u}{2}||^2 \geq 4d^2,$$

d'où, nécessairement $||Pu - \tilde{P}u|| = 0$.

Montrons enfin la caractérisation de Pu. Vérifions d'abord que Pu satisfait (5.1). Soit $v \in K$ et $t \in]0,1]$. Alors $w = (1-t)Pu + tv \in K$ et par conséquent, $||u - w|| \geq d$. Or

$$\begin{aligned} ||u-w||^2 &= ||(u-Pu) + t(Pu-v)||^2 \\ &= ||u-Pu||^2 + 2t\,\mathfrak{Re}(u-Pu, Pu-v) + t^2||Pu-v||^2 \\ &= d^2 + 2t\,\mathfrak{Re}(u-Pu, Pu-v) + t^2||Pu-v||^2. \end{aligned}$$

On a donc

$$2t\,\mathfrak{Re}(u-Pu, Pu-v) + t^2||Pu-v||^2 \geq 0.$$

En divisant par t puis en faisant tendre t vers 0, on obtient finalement (5.1). Réciproquement, soit $w \in K$ tel que

$$\mathfrak{Re}(u-w, w-v) \geq 0, \ \forall v \in K.$$

Alors, pour tout $v \in K$

$$||u-v||^2 = ||(u-w)+(w-v)||^2 = ||u-w||^2 + ||w-v||^2 + 2\,\mathfrak{Re}(u-w, w-v)$$

et par conséquent $||u-v||^2 \geq ||u-w||^2$. Ceci montre que $w = Pu$. □

Bien entendu, le théorème précédent s'applique au cas où K est un sous-espace vectoriel fermé de H. L'opérateur P a dans ce cas plusieurs propriétés remarquables :

Lemme 5.14 *Soit E un sous-espace vectoriel fermé de H et P l'opérateur qui à $u \in H$ associe sa projection Pu sur le sous-espace E. Alors on a les propriétés suivantes :*

(i) *Pour tout $u \in H$, Pu est caractérisé par*

$$Pu \in E \text{ et } (u-Pu, v) = 0, \ \forall v \in E. \tag{5.2}$$

(ii) *P est un opérateur linéaire de H dans H.*

(iii) *Pour tout $u \in H$, $||Pu|| \leq ||u||$. P est donc un opérateur continu de H dans H. On dit que P est l'opérateur de projection orthogonale sur E.*

Démonstration. Commençons par montrer la caractérisation. Soient $u \in H$, $v \in E$ et $\lambda \in \mathbb{C}$. D'après le théorème précédent,

$$\mathfrak{Re}(u-Pu, w-Pu) \leq 0, \ \forall w \in E.$$

En particulier, si $w = Pu + \lambda v$ (qui appartient bien à E puisque E est un espace vectoriel), on trouve $\mathfrak{Re}(u-Pu, \lambda v) \leq 0$. En choisissant $\lambda = (u-Pu, v)$ on trouve finalement

$$|(u-Pu, v)|^2 \leq 0,$$

d'où $(u - Pu, v) = 0$. Réciproquement, si $w \in E$ est tel que $(u - w, v) = 0,\ \forall v \in E$, alors on a aussi, puisque $w - v \in E$,

$$(u - w, w - v) = 0.$$

Ainsi,

$$||u - v||^2 = ||u - w||^2 + ||w - v||^2,\ \forall v \in E,$$

d'où

$$||u - w|| = \min_{v \in E} ||u - v||.$$

Autrement dit, $w = Pu$.

Montrons maintenant que P est un opérateur linéaire de H dans H. Pour tout $u \in H$ et $\lambda \in \mathbb{C}$, on a $\lambda Pu \in E$ et

$$\lambda(u - Pu, v) = (\lambda u - \lambda Pu, v) = 0,\ \forall v \in E.$$

D'après la caractérisation (5.2), cela prouve que $P(\lambda u) = \lambda Pu$. On prouve de même que $P(u + v) = Pu + Pv, \forall u, v \in H$.

Enfin, pour tout u dans H, on a $(u - Pu, Pu) = 0$, donc $||u||^2 = ||u - Pu||^2 + ||Pu||^2$, d'où $||Pu|| \leq ||u||$. □

Avant d'énoncer quelques corollaires importants de ce lemme, il nous faut introduire la définition suivante :

Définition 5.15 *Soit F un sous-ensemble de H. On appelle espace orthogonal à F et on note $F^\perp$ le sous-espace vectoriel de H défini par :*

$$F^\perp = \{u \in H;\ (u, v) = 0,\ \forall v \in F\}.$$

Remarquons que $F^\perp$ est fermé. En effet, si une suite (u_n) d'éléments de $F^\perp$ converge vers $u \in H$, on a $(u_n - u, v) \to 0$ pour tout $v \in H$, d'après l'inégalité de Cauchy-Schwarz. Par conséquent $(u, v) = 0$ pour tout $v \in F$.

On a tout d'abord le

Corollaire 5.16 *Si E est un sous-espace vectoriel fermé de H, alors :*

$$H = E \oplus E^\perp.$$

Démonstration. Notons P la projection orthogonale sur E. Alors, si $u \in H$, $u = (u - Pu) + Pu$ avec $Pu \in E$ et $u - Pu \in E^\perp$ d'après le lemme 5.14 (i). Pour établir l'unicité d'une telle décomposition, on remarque que si

$$u_1 + u_2 = 0 \text{ avec } u_1 \in E \text{ et } u_2 \in E^\perp,$$

on a immédiatement $||u_1||^2 + ||u_2||^2 = 0$, d'où $u_1 = u_2 = 0$. □

Ce résultat nous donne une caractérisation très utile des sous-espaces denses d'un espace de Hilbert.

Corollaire 5.17 *Soit F un sous-espace vectoriel de H. Alors F est dense dans H si et seulement si $F^\perp = \{0\}$.*

DÉMONSTRATION. Notons E l'adhérence de F. Vérifions que

$$E^\perp = F^\perp.$$

En effet, il est clair que $F \subset E$ implique que $E^\perp \subset F^\perp$. Soit maintenant un élément u de $F^\perp$, on a donc $(u, v) = 0$ pour tout $v \in F$. Pour tout $v \in E$, il existe une suite (v_n) d'éléments de F qui converge vers v dans H. En passant à la limite dans l'égalité $(u, v_n) = 0$, il vient $(u, v) = 0$, soit $u \in E^\perp$. Par ailleurs, E est un sous-espace vectoriel fermé de H. Par conséquent, d'après le corollaire précédent,

$$H = E \oplus E^\perp = E \oplus F^\perp.$$

Ceci prouve que $E = H$ si et seulement si $F^\perp = \{0\}$. □

Nous allons conclure ce paragraphe en démontrant un résultat important pour la suite. En utilisant le théorème de la projection, nous allons montrer qu'il existe un isomorphisme entre un espace de Hilbert H et son dual, c'est-à-dire l'espace vectoriel des formes linéaires continues sur H. Ce résultat, dû à Riesz, conduira souvent à identifier un espace de Hilbert à son dual.

Théorème 5.18 (de Riesz) *Soit ℓ une forme linéaire continue sur H. Alors il existe un et un seul élément $u \in H$ tel que*

$$\ell(v) = (v, u) \ \forall v \in H.$$

De plus

$$|||\ell||| := \sup_{v \in H, v \neq 0} \frac{|\ell(v)|}{||v||} = ||u||.$$

DÉMONSTRATION. Notons tout d'abord que si $\ell = 0$, le résultat est immédiat. On suppose donc que $\ell \neq 0$ et on note E le noyau de la forme ℓ :

$$E = \{v \in H;\ \ell(v) = 0\}.$$

Comme ℓ est linéaire et continue, E est un sous-espace vectoriel fermé de H. Comme $E^\perp$ n'est pas réduit à $\{0\}$, il contient un élément non nul w. Cet élément n'appartient donc pas à E, soit $\ell(w) \neq 0$. Pour tout $v \in H$, posons

$$v = \frac{\ell(v)}{\ell(w)} w + \left(v - \frac{\ell(v)}{\ell(w)} w \right) = v_1 + v_2.$$

On a $v_1 \in E^\perp$ tandis que $v_2 \in E$. On en déduit

$$(v, w) = \frac{\ell(v)}{\ell(w)} ||w||^2.$$

On pose enfin

$$u = \frac{\overline{\ell(w)}}{||w||^2} w,$$

et l'on voit que pour tout $v \in H$, $(v, u) = \ell(v)$, ce qui achève la preuve de l'existence du u.

L'unicité est elle immédiate, de même que l'identité $|||\ell||| = ||u||$. □

5.3 Bases hilbertiennes

On se limite dans la suite aux espaces de Hilbert *séparables*, c'est-à-dire admettant un sous-ensemble dénombrable dense. La quasi-totalité des espaces de Hilbert associés aux problèmes de la physique sont séparables : ce cas "particulier" est en fait très général ! Tout espace de Hilbert *séparable* possède une base hilbertienne (dénombrable), qui généralise la base orthonormale en dimension finie. Ainsi par exemple, on peut montrer facilement, à l'aide du théorème de Stone-Weierstraβ, que les fonctions $e^{in\theta}/\sqrt{2\pi}$, pour $n \in \mathbb{Z}$, constituent une base hilbertienne de $L^2(]0, 2\pi[$. Plus généralement, la théorie spectrale des opérateurs autoadjoints compacts (qui n'est pas abordée dans ce cours) permet de construire diverses bases hilbertiennes, destinées à diagonaliser des opérateurs aux dérivées partielles. Nous considérons donc exclusivement dans ce paragraphe le cas d'un espace de Hilbert H séparable et de dimension infinie.

Rappelons tout d'abord les définitions essentielles suivantes.

Définition 5.19 *On dit qu'un espace de Hilbert H est séparable s'il existe un sous-ensemble dénombrable et dense dans H.*

Définition 5.20 *Soit H un espace de Hilbert séparable. On appelle base hilbertienne de H toute famille $(e_n)_{n\in\mathbb{N}}$ d'éléments de H de norme 1 et orthogonaux deux à deux, c'est-à-dire*

$$(e_m, e_n) = \delta_{mn},$$

telle que le sous-espace formé par les combinaisons linéaires finies d'éléments de la suite $(e_n)_{n\in\mathbb{N}}$ soit dense dans H.

On a le

Théorème 5.21 *Tout espace de Hilbert séparable de dimension infinie admet une base hilbertienne.*

DÉMONSTRATION. Soit H un espace de Hilbert séparable de dimension infinie et $(f_n)_{n\in\mathbb{N}}$ un sous-ensemble dense dans H. A plus forte raison, l'espace vectoriel F engendré par les f_n, c'est-à-dire formé par les combinaisons linéaires finies des f_n, est dense dans H. On peut alors extraire de la suite (f_n) une sous-suite dont tous les vecteurs f_n sont linéairement indépendants (en effet, si l'un d'eux est une combinaison linéaire finie des autres, on peut le supprimer de la suite (f_n) sans modifier l'espace F).

On applique alors le *procédé d'othonormalisation de Schmidt* à la sous-suite (f_n) : on pose $e_0 = f_0/||f_0||$ et par récurrence, pour tout $n \geq 0$,

$$e_{n+1} = \frac{f_{n+1} - \sum_{m=0}^{n}(f_{n+1}, e_m)e_m}{||f_{n+1} - \sum_{m=0}^{n}(f_{n+1}, e_m)e_m||}.$$

Par construction, la famille (e_n) est orthonormée et elle engendre l'espace F qui est dense dans H. C'est donc bien une base hilbertienne de H. □

Remarque 5.22 *La réciproque est également vraie. Si un espace de Hilbert H admet une base hilbertienne, alors il est séparable. En effet, l'ensemble dénombrable formé de toutes les combinaisons linéaires finies d'éléments de la base à coefficients dont les parties réelles et imaginaires sont rationnelles est dense dans H.*

On aura besoin dans la suite du lemme suivant.

Lemme 5.23 *Soit H un espace de Hilbert et (u_n) une suite d'éléments de H tels que les u_n soient deux à deux orthogonaux. Alors la série $\sum u_n$ converge si et seulement si la série $\sum ||u_n||^2$ converge, et*

$$\left\|\sum u_n\right\|^2 = \sum ||u_n||^2.$$

DÉMONSTRATION. Supposons que la série $\sum u_n$ soit convergente dans H et posons $S_N = \sum_{n=0}^{N} u_n$. Les u_n étant orthogonaux deux à deux, on a

$$||S_N||^2 = \sum_{n=0}^{N} ||u_n||^2.$$

La suite (S_N) converge par hypothèse, ce qui est donc aussi le cas pour la série numérique $\sum ||u_n||^2$.

Réciproquement, on suppose que la série $\sum ||u_n||^2$ converge. On a, les u_n étant deux à deux orthogonaux,

$$||S_{N+M} - S_N||^2 = \sum_{n=N+1}^{N+M} ||u_n||^2 = ||S_{N+M}||^2 - ||S_N||^2.$$

La suite numérique $(||S_N||)$ est convergente, donc de Cauchy dans $\mathbb{R}$. La suite (S_N) est donc également de Cauchy dans H, et finalement converge dans H. □

On en arrive maintenant au théorème le plus important de ce paragraphe.

Théorème 5.24 *Soit H un espace de Hilbert séparable et $(e_n)_{n\in\mathbb{N}}$ une base hilbertienne de H.*

(i) *Tout élément $u \in H$ peut s'écrire de manière unique sous la forme*

$$u = \sum (u, e_n) e_n,$$

avec

$$||u||^2 = \sum |(u, e_n)|^2.$$

(ii) *Réciproquement, étant donnée une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de $\mathbb{C}$ tels que $\sum |a_n|^2 < +\infty$, alors la série $\sum a_n e_n$ converge dans H vers un élément u tel que pour tout n, $a_n = (u, e_n)$.*

Démonstration. Pour montrer la partie (i) du théorème, considérons $u \in H$ et posons $S_N = \sum_{n=0}^{N} (u, e_n) e_n$. On a

$$(S_N, u) = \sum_{n=0}^{N} |(u, e_n)|^2 = ||S_N||^2.$$

On en déduit que $||S_N||^2 \leq ||S_N||\,||u||$, soit $||S_N|| \leq ||u||$, et la série numérique $\sum |(u, e_n)|^2$ est donc convergente. D'après le lemme 5.23, la série $\sum (u, e_n) e_n$ est donc convergente dans H. Si $v \in H$ désigne la limite, on a pour $N \geq n$, $(S_N, e_n) = (u, e_n)$. En passant à la limite $N \to +\infty$, il vient que $(v, e_n) = (u, e_n)$ pour tout n, soit $v = u$ d'après le corollaire 5.17.

Montrons maintenant la partie (ii). Si la série $\sum |a_n|^2 = \sum ||a_n e_n||^2$ converge, alors par le lemme 5.23, la série $\sum a_n e_n$ converge dans H vers un élément u. Posant $S_N = \sum_{n=0}^{N} a_n e_n$, on a pour $N \geq n$, $(S_N, e_n) = a_n$. En passant à la limite $N \to +\infty$, il vient que $a_n = (u, e_n)$ pour tout n, ce qui termine la preuve. □

Du théorème précédent on déduit le corollaire suivant :

Corollaire 5.25 *Tout espace de Hilbert séparable de dimension infinie est isomorphe à ℓ^2.*

En effet, soit H un espace de Hilbert séparable de dimension infinie et $(e_n)_{n\in\mathbb{N}}$ une base hilbertienne de H. Alors l'application T qui à $u \in H$ associe la suite (u, e_n) définit un isomorphisme de H dans ℓ^2.

Exemple 5.26 Nous allons terminer ce paragraphe par l'exemple des séries de Fourier dans l'espace $H = L^2(]0, 2\pi[)$. Pour tout $n \in \mathbb{Z}$, on note e_n la fonction suivante :

$$e_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}.$$

Montrons que la famille $(e_n)_{n\in\mathbb{Z}}$ forme une base hilbertienne de H. Il est facile de vérifier que les fonctions (e_n) forment une famille d'éléments de norme 1 et orthogonaux deux à deux. Pour montrer que cette famille est dense, il nous suffit d'après le corollaire 5.17 de vérifier que si $f \in H$ est tel que tous ses coefficients de Fourier sont nuls :

$$\int_0^{2\pi} f(x)e^{-\mathrm{i}nx}\,\mathrm{d}x = 0, \text{ pour tout } n \in \mathbb{Z},$$

alors la fonction f est identiquement nulle.

Soit donc f une telle fonction. On a alors

$$\int_0^{2\pi} f(x)p(x)\,\mathrm{d}x = 0$$

pour tout *polynôme trigonométrique* p, c'est-à-dire pour toute fonction de la forme suivante :

$$p(x) = \sum_{n=0}^{N}(a_n\cos(nx) + b_n\sin(nx)),\ a_n, b_n \in \mathbb{C}.$$

Or, d'après le Théorème de Stone-Weierstraβ, toute fonction g continue sur $]0, 2\pi[$ et telle que $g(0) = g(2\pi)$ est limite uniforme de polynômes trigonométriques. On a donc pour toute fonction g continue et périodique :

$$\int_0^{2\pi} f(x)g(x)\,\mathrm{d}x = 0.$$

On conclut à la nullité de f en utilisant par exemple la densité des fonctions continues à support compact dans $L^2(]0, 2\pi[)$, donnée par le théorème 4.17.

En conclusion, nous pouvons affirmer que pour toute fonction f de $L^2(]0, 2\pi[)$, la série de Fourier de f :

$$\sum_{n\in\mathbb{Z}}(f, e_n)e_n$$

converge dans L^2 vers f. On a a fortiori

$$\int_0^{2\pi} |f(x)|^2\,\mathrm{d}x = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}}\left|\int_0^{2\pi} f(x)e^{-\mathrm{i}nx}\,\mathrm{d}x\right|^2.$$

•

Chapitre 6

Espaces de Sobolev

Nous allons dans ce chapitre étudier des espaces vectoriels de distributions particuliers, appelés espaces de Sobolev. Une propriété importante est qu'il s'agit d'espaces de Hilbert. Pour de nombreuses équations aux dérivées partielles issues de la physique, l'espace de Sobolev adapté à la résolution de l'équation peut s'interpréter comme l'espace des solutions d'énergie finie. Nous nous contenterons de définir et étudier les espaces de Sobolev d'ordre entier $H^m(\Omega)$ dans le cas d'un ouvert Ω quelconque, en étant plus précis sur l'espace $H^1(\Omega)$, qui joue un rôle fondamental dans l'analyse des équations aux dérivées partielles d'ordre 2. Indiquons enfin que l'on peut généraliser la notion d'espace de Sobolev $H^s(\Omega)$ pour $s \in \mathbb{R}$, mais cela nécessite des outils non abordés dans ce cours.

6.1 Les espaces $H^m(\Omega)$

Les espaces de Sobolev sont définis à partir de l'espace $L^2(\Omega)$ introduit au paragraphe 4.2.1.

Définition 6.1 *Si Ω désigne un ouvert de $\mathbb{R}^d$ et m un entier naturel, on appelle espace de Sobolev d'ordre m et on note $H^m(\Omega)$ l'espace vectoriel suivant :*

$$H^m(\Omega) = \left\{ u \in L^2(\Omega); \partial^\alpha u \in L^2(\Omega),\ \forall \alpha \in \mathbb{N}^d, |\alpha| \leq m \right\}, \tag{6.1}$$

où

$$|\alpha| = \sum_{j=1,\dots,d} \alpha_j \text{ et } \partial^\alpha u = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}}.$$

Précisons que dans cette définition la dérivée $\partial^\alpha u$ est à prendre au sens de $\mathcal{D}'(\Omega)$. Il découle de la définition que

$$H^{m'}(\Omega) \subset H^m(\Omega) \subset H^0(\Omega) = L^2(\Omega), \quad \text{si} \quad 0 \leq m \leq m'.$$

Des propriétés de $L^2(\Omega)$, on déduit immédiatement la

Proposition 6.2 *$H^m(\Omega)$ est un espace vectoriel et l'application*

$$(u,v) \to \sum_{\alpha\in\mathbb{N}^d,|\alpha|\leq m} \int_\Omega \partial^\alpha u\, \partial^\alpha \overline{v}\, \mathrm{d}x$$

est un produit scalaire sur $H^m(\Omega)$ que l'on notera $(u,v)_{H^m}$. La norme associée est notée $||u||_{H^m}$.

Montrons maintenant qu'il s'agit d'espaces complets.

Proposition 6.3 *L'espace $H^m(\Omega)$, muni de la norme $||\cdot||_{H^m}$ est complet.*

Démonstration. Soit u_n une suite de Cauchy dans $H^m(\Omega)$. Alors $\partial^\alpha u_n$ est de Cauchy dans $L^2(\Omega)$ pour tout $\alpha \in \mathbb{N}^d$ tel que $|\alpha| \leq m$. Comme $L^2(\Omega)$ est complet, la suite u_n converge dans $L^2(\Omega)$ vers une fonction u. De même, il existe pour chaque $\alpha \neq 0$ une fonction $g_\alpha \in L^2(\Omega)$ telle que

$$\partial^\alpha u_n \to g_\alpha \in L^2(\Omega).$$

Pour montrer que $u_n \to u$ dans $H^m(\Omega)$, il suffit de vérifier que $\partial^\alpha u = g_\alpha$ dans $\mathcal{D}'(\Omega)$. Or, pour tout $\varphi \in \mathcal{D}(\Omega)$, on a :

$$\langle \partial^\alpha u, \varphi\rangle = \langle u, (-1)^{|\alpha|}\partial^\alpha\varphi\rangle = \lim\langle u_n, (-1)^{|\alpha|}\partial^\alpha\varphi\rangle = \lim\langle \partial^\alpha u_n, \varphi\rangle = \langle g_\alpha, \varphi\rangle,$$

ce qui termine la démonstration. □

Pour $m \in \mathbb{N}$, on définit l'espace $H_0^m(\Omega)$ comme suit.

Définition 6.4 *On appelle $H_0^m(\Omega)$ l'adhérence de $\mathcal{D}(\Omega)$ dans $H^m(\Omega)$.*

L'espace $H_0^m(\Omega)$ est par conséquent un sous-espace fermé de l'espace de Hilbert $H^m(\Omega)$, et donc un espace de Hilbert muni de la norme induite de $H^m(\Omega)$. Dans le cas où $\Omega = \mathbb{R}^d$, on admet le résultat suivant :

Théorème 6.5 *Pour tout entier naturel m, l'espace $H_0^m(\mathbb{R}^d)$ coïncide avec l'espace $H^m(\mathbb{R}^d)$.*

Notons que le théorème précédent revient exactement à dire que $\mathcal{D}(\mathbb{R}^d)$ est dense dans $H^m(\mathbb{R}^d)$. Il est important de noter que ce résultat n'est plus vrai pour $\Omega \neq \mathbb{R}^d$, sauf pour $m = 0$, ce que l'on comprendra en étudiant le cas $m = 1$ (paragraphe 6.2).

On termine ce paragraphe par une caractérisation du dual de $H_0^m(\Omega)$.

Définition 6.6 *Pour tout entier naturel m, on note $H^{-m}(\Omega)$ l'espace des distributions u telles qu'il existe une constante C avec*

$$\forall \varphi \in \mathcal{D}(\Omega), \quad |\langle u, \varphi\rangle| \leq C\, ||\varphi||_{H^m}.$$

Le théorème suivant permet d'identifier $H^{-m}(\Omega)$ au dual de $H_0^m(\Omega)$.

Théorème 6.7 *On a les deux résultats suivants.*

(i) *Pour tout $u \in H^{-m}(\Omega)$, l'application $\varphi \mapsto \langle u, \varphi \rangle$ définie de $\mathcal{D}(\Omega)$ dans $\mathbb{C}$ se prolonge en une forme linéaire continue sur $H_0^m(\Omega)$. On notera $\langle u, v \rangle_{H^{-m}, H^m}$ ce prolongement pour $v \in H_0^m(\Omega)$.*

(ii) *Le prolongement précédent identifie $H^{-m}(\Omega)$ au dual de $H_0^m(\Omega)$: pour toute forme linéaire ℓ continue sur $H_0^m(\Omega)$, il existe un et un seul $u \in H^{-m}(\Omega)$ tel que*

$$\forall v \in H^m(\Omega), \quad \ell(v) = \langle u, v \rangle_{H^{-m}, H^m}.$$

Démonstration. La première partie du théorème est une application directe de la proposition 1.15 avec $E = H_0^m(\Omega)$, $F = \mathbb{C}$, $A = \mathcal{D}(\Omega)$ et $f : \varphi \mapsto \langle u, \varphi \rangle$. Montrons maintenant la seconde partie. Soit ℓ une forme linéaire continue sur $H_0^m(\Omega)$. On a en particulier pour tout compact $K \subset \Omega$ et $\varphi \in \mathcal{D}_K(\Omega)$,

$$|\ell(\varphi)| \leq C \, ||\varphi||_{H^m} = C \sqrt{\sum_{|\alpha| \leq m} \int_\Omega |\partial^\alpha \varphi|^2 \, \mathrm{d}x} \leq C \sqrt{N} \sup_{|\alpha| \leq m} ||\partial^\alpha \varphi||_{L^2(\Omega)},$$

où N est le nombre de multi-indices concernés. On utilise maintenant le fait que

$$||\partial^\alpha \varphi||_{L^2(\Omega)} \leq \sqrt{\mathrm{mes}(K)} \sup_{x \in K} |\partial^\alpha \varphi(x)|,$$

les deux inégalités précédentes aboutissant au fait qu'il existe une constante C_K telle que

$$|\ell(\varphi)| \leq C_K \sup_{|\alpha| \leq m} \sup_{x \in K} |\partial^\alpha \varphi(x)|,$$

ce qui montre que ℓ restreinte à $\mathcal{D}(\Omega)$ est une distribution u. On a $|\langle u, \varphi \rangle| = |\ell(\varphi)| \leq C \, ||\varphi||_{H^m}$ pour tout $\varphi \in \mathcal{D}(\Omega)$, ce qui montre que $u \in H^{-m}(\Omega)$. Les deux formes linéaires ℓ et $v \mapsto \langle u, v \rangle_{H^{-m}, H^m}$ coïncident sur l'espace dense $\mathcal{D}(\Omega)$ et donc sur $H_0^m(\Omega)$. □

6.2 Le cas particulier $H^1(\Omega)$

La définition générale (6.1) devient dans le cas $m = 1$

$$H^1(\Omega) = \left\{ u \in L^2(\Omega); \frac{\partial u}{\partial x_j} \in L^2(\Omega), \, \forall j = 1, \ldots, d \right\},$$

le produit scalaire de $H^1(\Omega)$ étant

$$(u, v) \mapsto \int_\Omega \left(u \, \overline{v} + \sum_{j=1,\ldots,d} \frac{\partial u}{\partial x_j} \frac{\partial \overline{v}}{\partial x_j} \right) \mathrm{d}x = \int_\Omega \left(u \, \overline{v} + \nabla u \cdot \overline{\nabla v} \right) \mathrm{d}x,$$

où ∇u désigne le vecteur de $\mathbb{R}^d$ formé par les d composantes $\partial u / \partial x_j$.

6.2.1 Notion de trace

Dans un domaine borné Ω de classe $\mathcal{C}^1$ et pour une fonction u continue jusqu'au bord du domaine, soit $u \in \mathcal{C}^0(\overline{\Omega})$, la restriction de la fonction u sur le bord du domaine $\partial\Omega$ est naturellement parfaitement définie. Si l'on considère maintenant une fonction de $L^2(\Omega)$, une telle restriction n'a pas de sens car $\partial\Omega$ est de mesure nulle dans $\mathbb{R}^d$. En effet, d'après la définition 2.31, le bord $\partial\Omega$ est formé de la réunion de graphes de fonctions $\mathcal{C}^1$ de $\mathbb{R}^{d-1}$ dans $\mathbb{R}$, et le lecteur pourra se convaincre en exercice que de tels graphes sont des ensembles de mesure nulle dans $\mathbb{R}^d$. Dans ce paragraphe on veut montrer qu'en revanche on peut donner un sens à cette restriction pour une fonction de l'espace de Sobolev $H^1(\Omega)$, on parlera alors pour une telle fonction de trace sur $\partial\Omega$. On considère à cette fin un domaine borné Ω de $\mathbb{R}^d$ de classe $\mathcal{C}^1$ au sens de la définition 2.31. Nous noterons $L^2(\partial\Omega)$ l'espace des fonctions u définies presque partout sur $\partial\Omega$ et telles que

$$\int_{\partial\Omega} |u(x)|^2 \, \mathrm{d}\sigma(x) < +\infty,$$

où cette intégrale de surface est définie à l'aide de la formule locale (2.12) et de la partition de l'unité donnée par le théorème 2.32. On peut montrer facilement que $L^2(\partial\Omega)$ est un espace vectoriel, que l'application

$$(u, v) \mapsto (u, v)_{L^2(\partial\Omega)} = \int_{\partial\Omega} u\overline{v} \, \mathrm{d}\sigma$$

définit un produit scalaire sur $L^2(\partial\Omega)$ qui est un espace complet pour la norme associée. Nous allons montrer dans un premier temps la proposition suivante :

Proposition 6.8 *Soit Ω un ouvert borné de $\mathbb{R}^d$ de classe $\mathcal{C}^1$. Il existe une constante C telle que pour toute fonction u de $\mathcal{C}^1(\overline{\Omega})$,*

$$||u||_{L^2(\partial\Omega)} \leq C \, ||u||_{H^1(\Omega)}.$$

Démonstration. Reprenons les notations de la définition 2.31. Pour un point $x^* \in \partial\Omega$, on utilise donc un système de coordonnées locales $(\underline{x}, x_d)$ dans le voisinage $\omega = \mathcal{O} \times J$ du point x^*, le domaine Ω étant localement (c'est-à-dire dans ω) défini par $x_d > F(\underline{x})$. On note $J =]a, b[$ de sorte que $F(\overline{\mathcal{O}}) \subset]a, b[$ (voir figure 2.4). On suppose dans un premier temps que la fonction u est à support dans ω. Pour $(\underline{x}, x_d) \in \Omega \cap \omega$, on a

$$u(\underline{x}, F(\underline{x})) = u(\underline{x}, x_d) - \int_{F(\underline{x})}^{x_d} \frac{\partial u}{\partial x_d}(\underline{x}, t) \, \mathrm{d}t.$$

En utilisant l'inégalité de Cauchy-Schwarz, il vient que

$$|u(\underline{x}, F(\underline{x}))|^2 \leq 2|u(\underline{x}, x_d)|^2 + 2(b - F(\underline{x})) \int_{F(\underline{x})}^{b} \left| \frac{\partial u}{\partial x_d}(\underline{x}, t) \right|^2 \, \mathrm{d}t.$$

On intègre alors l'inégalité précédente en x_d dans l'intervalle $]F(\underline{x}), b[$. On obtient

$$(b - F(\underline{x}))\,|u(\underline{x}, F(\underline{x}))|^2 \leq 2\int_{F(\underline{x})}^{b} |u(\underline{x}, t)|^2\,\mathrm{d}t + 2(b - F(\underline{x}))^2 \int_{F(\underline{x})}^{b} \left|\frac{\partial u}{\partial x_d}(\underline{x}, t)\right|^2 \mathrm{d}t.$$

En utilisant le fait que $F(\overline{\mathcal{O}}) \subset]a, b[$, il existe une constante C telle que

$$|u(\underline{x}, F(\underline{x}))|^2 \leq C\int_{F(\underline{x})}^{b} |u(\underline{x}, t)|^2\,\mathrm{d}t + C\int_{F(\underline{x})}^{b} \left|\frac{\partial u}{\partial x_d}(\underline{x}, t)\right|^2 \mathrm{d}t.$$

En remarquant que la fonction $\underline{x} \mapsto |\underline{\nabla} F(\underline{x})|$ est bornée sur $\overline{\mathcal{O}}$ puisque F est de classe $\mathcal{C}^1$, il vient

$$|u(\underline{x}, F(\underline{x}))|^2 \sqrt{1 + |\underline{\nabla} F(\underline{x})|^2} \leq C\int_{F(\underline{x})}^{b} |u(\underline{x}, t)|^2\,\mathrm{d}t + C\int_{F(\underline{x})}^{b} \left|\frac{\partial u}{\partial x_d}(\underline{x}, t)\right|^2 \mathrm{d}t,$$

pour une nouvelle constante C. En introduisant enfin la mesure de surface

$$\mathrm{d}\sigma(x) = \sqrt{1 + |\underline{\nabla} F(\underline{x})|^2}\,\mathrm{d}\underline{x}$$

et en intégrant en $\underline{x}$ sur $\mathcal{O}$,

$$\int_{\partial\Omega\cap\omega} |u(x)|^2\,\mathrm{d}\sigma(x) \leq C\int_{\Omega\cap\omega} \left(|u(x)|^2 + |\nabla u(x)|^2\right)\mathrm{d}x.$$

Si l'on considère maintenant une fonction quelconque de $\mathcal{C}^1(\overline{\Omega})$, on utilise un recouvrement du compact $\partial\Omega$ (l'ouvert Ω étant borné) par une famille finie d'ouverts ω_p, $p = 1, \ldots, P$, et par le théorème 2.32 une partition de l'unité associée ψ_p, $p = 1, \ldots, P$. On remarque que les fonctions ψ_p satisfont la propriété

$$1 = \left(\sum_{p=1}^{P} \psi_p\right)^2 \leq P\sum_{p=1}^{P} \psi_p^2.$$

On a alors

$$\int_{\partial\Omega} |u(x)|^2\,\mathrm{d}\sigma(x) \leq P\sum_{p=1}^{P} \int_{\partial\Omega\cap\omega_p} |u(x)\psi_p(x)|^2\,\mathrm{d}\sigma(x)$$

et l'on utilise l'inégalité de trace précédente pour chaque fonction $u\,\psi_p$, dont on note que le support est dans ω_p. La proposition 6.8 est montrée. □

On admet le résultat de densité suivant.

Théorème 6.9 *Soit Ω un ouvert borné de $\mathbb{R}^d$ de classe $\mathcal{C}^1$. L'espace $\mathcal{C}^\infty(\overline{\Omega})$ est dense dans $H^1(\Omega)$.*

On en déduit immédiatement le théorème de trace suivant dans $H^1(\Omega)$.

Théorème 6.10 *Soit Ω un ouvert borné de $\mathbb{R}^d$ de classe $\mathcal{C}^1$. L'application $\gamma : \mathcal{C}^1(\overline{\Omega}) \to L^2(\partial\Omega)$ de restriction sur le bord se prolonge en une application linéaire continue, notée encore γ et appelée trace, de $H^1(\Omega)$ sur $L^2(\partial\Omega)$. Il existe donc une constante C telle que pour tout $u \in H^1(\Omega)$,*

$$||\gamma u||_{L^2(\partial\Omega)} \leq C\, ||u||_{H^1(\Omega)}.$$

DÉMONSTRATION. Le preuve est une application directe de la proposition 1.15 avec $E = H^1(\Omega)$, $F = L^2(\partial\Omega)$, $A = \mathcal{C}^1(\overline{\Omega})$ et $f = \gamma$, la continuité de γ résultant de la proposition 6.8 et la densité de $\mathcal{C}^1(\overline{\Omega})$ résultant du théorème 6.9. □

Remarque 6.11 *Le théorème de trace 6.10 que nous venons de démontrer n'est pas optimal au sens où l'application $\gamma : H^1(\Omega) \to L^2(\partial\Omega)$ n'est pas surjective : les éléments de $L^2(\partial\Omega)$ ne sont pas en général les traces sur $\partial\Omega$ d'élements de $H^1(\Omega)$. L'image de $H^1(\Omega)$ par γ est formé par l'espace de Sobolev fractionnaire $H^{1/2}(\partial\Omega)$, qui est strictement inclus dans $L^2(\partial\Omega)$ et qu'il est possible de définir au prix de notions supplémentaires non introduites dans le présent cours.*

6.2.2 Formules de Green

On considère à nouveau un domaine borné Ω de $\mathbb{R}^d$ de classe $\mathcal{C}^1$, le vecteur normal sortant étant noté n, et l'on veut étendre les formules de Green établies dans les corollaires 2.35, 2.36 et 2.37 pour des fonctions u et v dans $\mathcal{C}^m(\overline{\Omega})$ ($m = 1$ ou 2) au cas des fonctions u et v dans $H^m(\Omega)$ ($m = 1$ ou 2). En utilisant le théorème de densité 6.9 et la formule de Green du corollaire 2.35, on obtient immédiatement le

Théorème 6.12 *Pour tous u et v dans $H^1(\Omega)$ et tout $i = 1, \ldots, d$, on a*

$$\int_\Omega u \frac{\partial v}{\partial x_i}\, \mathrm{d}x = -\int_\Omega \frac{\partial u}{\partial x_i} v\, \mathrm{d}x + \int_{\partial\Omega} \gamma u\, \gamma v\, n_i\, \mathrm{d}\sigma.$$

Dorénavant, la trace de u sur $\partial\Omega$ sera notée simplement u au lieu de γu. A partir de la formule de Green du théorème 6.12, on déduit sans difficulté les deux suivantes : pour tous u dans $H^2(\Omega)$ et v dans $H^1(\Omega)$,

$$\int_\Omega (\Delta u) v\, \mathrm{d}x = -\int_\Omega \nabla u \cdot \nabla v\, \mathrm{d}x + \int_{\partial\Omega} \frac{\partial u}{\partial n} v\, \mathrm{d}\sigma \tag{6.2}$$

et : pour tous u et v dans $H^2(\Omega)$,

$$\int_\Omega \{(\Delta u)v - u(\Delta v)\}\, \mathrm{d}x = \int_{\partial\Omega} \left(\frac{\partial u}{\partial n} v - u \frac{\partial v}{\partial n}\right) \mathrm{d}\sigma, \tag{6.3}$$

où l'on rappelle que $\partial u/\partial n$ désigne $\nabla u \cdot n$, et les composantes du champ de vecteurs ∇u sur $\partial\Omega$ sont à comprendre au sens des traces.

6.2.3 Caractérisation de $H_0^1(\Omega)$

On rappelle que $H_0^1(\Omega)$ est l'adhérence de $\mathcal{D}(\Omega)$ dans $H^1(\Omega)$, qu'il constitue un espace de Hilbert muni de la norme induite de $H^1(\Omega)$, et qu'il diffère de $H^1(\Omega)$ si $\Omega \neq \mathbb{R}^d$. On souhaite maintenant caractériser l'espace $H_0^1(\Omega)$ à l'aide de la trace pour un domaine borné de classe $\mathcal{C}^1$. A cette fin, on admet le résultat suivant.

Théorème 6.13 *Soit Ω un ouvert borné de $\mathbb{R}^d$ de classe $\mathcal{C}^1$. L'espace $H_0^1(\Omega)$ coïncide avec celui des fonctions de $H^1(\Omega)$ dont le prolongement par 0 dans tout $\mathbb{R}^d$ est dans $H^1(\mathbb{R}^d)$.*

On notera que le théorème 6.13 revient à dire que l'espace $\mathcal{D}(\Omega)$ est dense dans le sous-espace des fonctions de $H^1(\Omega)$ dont le prolongement par 0 dans tout $\mathbb{R}^d$ est dans $H^1(\mathbb{R}^d)$. Donnons maintenant une nouvelle caractérisation de $H_0^1(\Omega)$ à l'aide de la trace.

Théorème 6.14 *Soit Ω un ouvert borné de $\mathbb{R}^d$ de classe $\mathcal{C}^1$. L'espace $H_0^1(\Omega)$ est l'ensemble des fonctions de $H^1(\Omega)$ dont la trace sur $\partial\Omega$ est nulle.*

Démonstration. Considérons tout d'abord $u \in H_0^1(\Omega)$, qui est donc la limite dans $H^1(\Omega)$ d'une suite (φ_n) d'éléments de $\mathcal{D}(\Omega)$. Comme la trace des φ_n sur $\partial\Omega$ est nulle, il vient d'après le théorème 6.10 que la trace de u sur $\partial\Omega$ est nulle également.

Réciproquement, considérons une fonction $u \in H^1(\Omega)$ telle que sa trace sur $\partial\Omega$ soit nulle. On appelle $\tilde{u}$ son prolongement par 0 dans $\mathbb{R}^d$. Il est clair que $\tilde{u} \in L^2(\mathbb{R}^d)$. Montrons que pour $i = 1, \ldots, d$, $\partial\tilde{u}/\partial x_i \in L^2(\mathbb{R}^d)$. Soit donc $\varphi \in \mathcal{D}(\mathbb{R}^d)$. Par définition de la dérivée au sens des distributions,

$$\langle \frac{\partial \tilde{u}}{\partial x_i}, \varphi \rangle = -\langle \tilde{u}, \frac{\partial \varphi}{\partial x_i} \rangle = -\int_\Omega u \frac{\partial \varphi}{\partial x_i} \, \mathrm{d}x.$$

En utilisant le théorème 6.12, il vient

$$\langle \frac{\partial \tilde{u}}{\partial x_i}, \varphi \rangle = \int_\Omega \frac{\partial u}{\partial x_i} \varphi \, \mathrm{d}x - \int_{\partial\Omega} u\varphi \, n_i \, \mathrm{d}\sigma.$$

Compte tenu que la trace de u sur $\partial\Omega$ est nulle, on obtient que pour tout $\varphi \in \mathcal{D}(\mathbb{R}^d)$,

$$\langle \frac{\partial \tilde{u}}{\partial x_i}, \varphi \rangle = \int_{\mathbb{R}^d} \widetilde{\frac{\partial u}{\partial x_i}} \varphi \, \mathrm{d}x,$$

où $\widetilde{\partial u/\partial x_i}$ désigne la prolongement par 0 de $\partial u/\partial x_i \in L^2(\Omega)$ dans $\mathbb{R}^d$. On conclut que $\partial\tilde{u}/\partial x_i$ s'identifie à un élément de $L^2(\mathbb{R}^d)$, et donc que $\tilde{u} \in H^1(\mathbb{R}^d)$. Il s'ensuit d'après le théorème 6.13 que $u \in H_0^1(\Omega)$, ce qui achève la preuve. □

6.2.4 Le cas $d = 1$

Dans le cas monodimensionnel, l'espace $H^1(I)$, où I est un intervalle de $\mathbb{R}$, bénéficie de propriétés supplémentaires par rapport à la dimension quelconque, ce qui justifie une étude spécifique. En particulier, les fonctions de $H^1(I)$ sont continues sur $\overline{I}$. Plus précisément, on a le

Théorème 6.15 *On a $H^1(I) \subset \mathcal{C}^0(\overline{I})$ et pour tout $u \in H^1(I)$,*

$$u(y) - u(x) = \int_x^y u'(t)\,\mathrm{d}t, \quad \forall x, y \in \overline{I}. \tag{6.4}$$

Démonstration. Pour x_0 fixé dans I, on pose pour $x \in I$

$$v(x) = \int_{x_0}^x u'(t)\,\mathrm{d}t,$$

v étant bien définie puisque $u' \in L^2(I) \subset L^1_{\mathrm{loc}}(I)$. En reprenant les arguments de la preuve du lemme 3.12, on voit que v est continue sur $\overline{I}$ est que la dérivée au sens des distributions de v coïncide avec u'. On a donc $(v-u)' = 0$ au sens des distributions dans I, soit $v = u + C$, où C est une constante. La fonction u coïncide donc presque partout avec une fonction continue sur $\overline{I}$. L'identité (6.4) découle immédiatement du fait que $v = u + C$. □

Remarque 6.16 *A partir de l'identité (6.4) on peut obtenir un résultat un peu plus précis que $H^1(I) \subset \mathcal{C}^0(\overline{I})$. En utilisant l'inégalité de Cauchy-Schwarz, on voit en effet que pour tous $x, y \in \overline{I}$,*

$$|u(x) - u(y)| \leq \sqrt{|x-y|}\sqrt{\int_x^y |u'(t)|^2\,\mathrm{d}t} \leq ||u||_{H^1(I)}\sqrt{|x-y|}.$$

Donc $H^1(I) \subset \mathcal{C}^{0,1/2}(\overline{I})$, où pour $\alpha \in]0,1[$ l'espace de Hölder $\mathcal{C}^{0,\alpha}(\overline{I})$ est formé par l'ensemble des fonctions u de $\mathcal{C}^0(\overline{I})$ telles que pour une constante C,

$$|u(x) - u(y)| \leq C\,|x-y|^\alpha, \quad \forall x, y \in \overline{I}.$$

On déduit immédiatement de (6.4) une formule d'intégration par parties pour les fonctions de $H^1(I)$.

Corollaire 6.17 *Pour tous u et v dans $H^1(I)$,*

$$\int_x^y u'v\,\mathrm{d}t = -\int_x^y uv'\,\mathrm{d}t + u(y)v(y) - u(x)v(x), \quad \forall x, y \in \overline{I}.$$

On voit donc que la notion de trace pour les fonctions de $H^1(I)$ est parfaitement définie en tout point $x_0 \in \overline{I}$ puisque ces fonctions sont continues sur $\overline{I}$. On précise ce constat dans le théorème 6.19, qui nécessite le résultat suivant sur le prolongement continue à $\mathbb{R}$ des fonctions de $H^1(I)$, que nous admettrons.

Proposition 6.18 *Pour tout intervalle $I \subset \mathbb{R}$, il existe un opérateur de prolongement $P : H^1(I) \to H^1(\mathbb{R})$ linéaire et continu tel que*

(i) *$Pu|_I = u$ pour tout $u \in H^1(I)$,*

(ii) *il existe une constante C telle que*

$$||Pu||_{H^1(\mathbb{R})} \leq C\,||u||_{H^1(I)}, \quad \forall u \in H^1(I).$$

Théorème 6.19 *On a $H^1(I) \subset \mathcal{C}_b^0(\overline{I})$, où $\mathcal{C}_b^0(\overline{I})$ est l'espace des fonctions continues et bornées sur $\overline{I}$, et il existe une constante C telle que pour tout $u \in H^1(I)$,*

$$\sup_{x\in\overline{I}} |u(x)| \leq C\,||u||_{H^1(I)}. \tag{6.5}$$

En particulier, l'application qui à une fonction de $H^1(I)$ associe sa valeur en un point $x_0 \in \overline{I}$ est une forme linéaire continue. Plus précisément, pour tout $u \in H^1(I)$,

$$|u(x_0)| \leq C\,||u||_{H^1(I)}.$$

Démonstration. On commence par le cas $I = \mathbb{R}$. Soit $u \in H^1(\mathbb{R})$. En vertu du théorème de densité 6.5, il existe une suite de fonctions $(\varphi_n) \in \mathcal{D}(\mathbb{R})$ telle que $\varphi_n \to u$ dans $H^1(\mathbb{R})$. On a pour tout $x \in \mathbb{R}$ et tout $\varphi \in \mathcal{D}(\mathbb{R})$,

$$\varphi^2(x) = 2\int_{-\infty}^{x} \varphi(t)\varphi'(t)\,\mathrm{d}t.$$

Par l'inégalité de Cauchy-Schwarz, il vient

$$|\varphi(x)|^2 \leq 2||\varphi||_{L^2(\mathbb{R})}||\varphi'||_{L^2(\mathbb{R})},$$

soit

$$|\varphi(x)| \leq \sqrt{||\varphi||^2_{L^2(\mathbb{R})} + ||\varphi'||^2_{L^2(\mathbb{R})}} = ||\varphi||_{H^1(\mathbb{R})}.$$

Il s'ensuit que pour tout $\varphi \in \mathcal{D}(\mathbb{R})$,

$$\sup_{x\in\mathbb{R}} |\varphi(x)| \leq ||\varphi||_{H^1(\mathbb{R})}.$$

La suite (φ_n) étant de Cauchy dans $H^1(\mathbb{R})$, elle est également de Cauchy dans $\mathcal{C}_b^0(\mathbb{R})$, qui est complet. Elle converge donc dans $\mathcal{C}_b^0(\mathbb{R})$ vers un certain $v \in \mathcal{C}_b^0(\mathbb{R})$. Comme (φ_n) converge simultanément vers u et v au sens des distributions, on a $u = v$. On obtient l'inégalité (6.5) en passant à la limite dans l'inégalité

$$\sup_{x\in\mathbb{R}} |\varphi_n(x)| \leq ||\varphi_n||_{H^1(\mathbb{R})}.$$

La deuxième partie du théorème résulte immédiatement de l'inégalité (6.5) et de la continuité de u.

On traite maintenant le cas d'un intervalle I quelconque. Soit P un opérateur de prolongement tel que défini par la proposition 6.18. En appliquant le résultat précédent montré pour $I = \mathbb{R}$, il vient que pour tout $u \in H^1(I)$,

$$\sup_{x\in\mathbb{R}} |Pu(x)| \leq C\, ||Pu||_{H^1(\mathbb{R})}.$$

Il existe donc une nouvelle constante C' telle que

$$\sup_{x\in\overline{I}} |u(x)| \leq \sup_{x\in\mathbb{R}} |Pu(x)| \leq C\, ||Pu||_{H^1(\mathbb{R})} \leq C'\, ||u||_{H^1(I)},$$

ce qui achève la preuve. □

On déduit du théorème 6.19 le fait que les fonctions de $H^1(\mathbb{R})$ tendent vers 0 à l'infini.

Corollaire 6.20 *Pour tout $u \in H^1(\mathbb{R})$,*

$$\lim_{x\to\pm\infty} u(x) = 0.$$

Démonstration. Soit $u \in H^1(\mathbb{R})$. Par le théorème de densité 6.5, il existe une suite de fonctions $(\varphi_n) \in \mathcal{D}(\mathbb{R})$ telle que $\varphi_n \to u$ dans $H^1(\mathbb{R})$. On déduit de l'inégalité (6.5) appliquée à $u - \varphi_n$ que $\sup_{x\in\mathbb{R}} |u(x) - \varphi_n(x)| \to 0$. Pour tout $\varepsilon > 0$, choisissons un n assez grand pour que $\sup_{x\in\mathbb{R}} |u(x) - \varphi_n(x)| < \varepsilon$. Pour $|x|$ assez grand on a $\varphi_n(x) = 0$, et donc $|u(x)| < \varepsilon$, ce qui termine la preuve. □

Remarque 6.21 *On a montré que les fonctions de $H^1(]-1,1[)$ sont continues sur $[-1,1]$. En revanche, les fonctions de $H^1(B_d(0,1))$, où $B_d(0,1)$ est la boule unité de $\mathbb{R}^d$, ne sont pas continues en général pour $d \geq 2$. On pourra pour s'en convaincre considérer pour $d = 2$ l'exemple suivant :*

$$u(x,y) = \theta(\sqrt{x^2+y^2}),$$

où $\theta(r) = (\log 1/2r)^k$, avec $k \in]0,1/2[$. Le lecteur vérifiera que la fonction u est dans $H^1(B_2(0,1))$ mais n'est pas continue en 0.

On peut ajouter que le corollaire 6.20 n'est pas valable pour les fonctions de $L^2(\mathbb{R})$, et qu'il n'est pas valable non plus pour les fonctions de $H^1(\mathbb{R}^d)$ avec $d \geq 2$.

Chapitre 7

Etude variationnelle des problèmes elliptiques

La notion de formulation variationnelle joue un rôle fondamental dans la résolution des équations aux dérivées partielles (EDP), en particulier des équations aux dérivées partielles elliptiques. Les EDP elliptiques sont celles dans lesquelles la variable temporelle est absente, soit parce que le phénomène physique décrit est stationnaire (indépendant du temps), soit parce le temps a été éliminé par application préalable de la transformée de Fourier (passage au domaine fréquentiel). Le méthode variationnelle fournit alors un cadre théorique simple pour montrer le caractère bien posé d'une EDP, c'est-à-dire l'existence et l'unicité de la solution, qui plus est la stabilité de cette solution vis-à-vis des données. La méthode variationnelle correspond à ce que les mécaniciens appellent le principe des travaux (ou puissances) virtuelles, principe qui se trouvera rigoureusement justifié dans le présent chapitre. Enfin, la méthode variationnelle est la base de la méthode des éléments finis : cette méthode numérique permet en pratique d'approcher la solution d'une EDP pour des géométries complexes.

7.1 Quelques exemples de problèmes aux limites

7.1.1 Le problème de Dirichlet

Considérons le problème d'électrostatique suivant. Dans une cavité Ω sont réparties des charges électriques selon une densité $\rho(x)$ connue et l'on cherche à calculer le champ électrique E satisfaisant

$$\operatorname{div}(\varepsilon E) = \rho \text{ et } \operatorname{rot} E = 0 \text{ dans } \Omega,$$

où ε est une constante qui désigne la permittivité diélectrique du milieu. De plus, si la paroi de la cavité est métallique, assimilable à un conducteur parfait, la composante tangentielle du champ électrique doit s'annuler sur cette paroi.

Pour résoudre ce problème, on remarque que E dérive d'un potentiel électrostatique, soit u :

$$E = \nabla u$$

qui doit vérifier l'équation suivante :

$$-\Delta u = f,$$

où $f = -\rho/\varepsilon$ (le signe "$-$" sera commode dans la suite). De plus le potentiel u doit prendre une valeur constante sur chaque composante connexe de la frontière $\partial\Omega$: en effet, sa dérivée tangentielle n'est autre que la trace tangentielle de E. Considérons pour simplifier le cas d'une cavité simplement connexe. L'ajout d'une constante au potentiel électrostatique ne modifiant pas le champ électrique, on peut imposer à u de s'annuler identiquement sur le bord de Ω. On est alors conduit à résoudre le problème suivant : trouver une fonction u telle que

$$\left\{ \begin{array}{lll} -\Delta u = f & \text{dans} & \Omega, \\ u = 0 & \text{sur} & \partial\Omega. \end{array} \right. \tag{7.1}$$

Ce problème est appelé *problème de Dirichlet.* L'équation aux dérivées partielles satisfaite par u est appelée *équation de Poisson.* La condition imposée à u sur le bord du domaine est appelée une *condition aux limites de type Dirichlet homogène.* Si l'on considère une cavité non simplement connexe, on peut imposer à u de prendre des valeurs constantes distinctes sur chacune des composantes connexes de $\partial\Omega$. Cela revient à imposer une différence de potentiel. La condition aux limites obtenue dans ce cas est dite de type Dirichlet *non homogène.*

7.1.2 Le problème de Neumann

Considérons maintenant le problème des vibrations acoustiques d'un fluide compressible dans une cavité Ω. En régime harmonique, la pression est de la forme

$$\mathcal{P}(x,t) = \mathfrak{Re}(u(x)\mathrm{e}^{-\mathrm{i}\omega t})$$

où u est une fonction de $x \in \Omega$ à valeurs dans $\mathbb{C}$ et elle satisfait l'équation suivante :

$$\frac{1}{c^2}\frac{\partial^2 \mathcal{P}}{\partial t^2} - \Delta \mathcal{P} + \alpha \frac{\partial \mathcal{P}}{\partial t} = \mathfrak{Re}(f(x)\mathrm{e}^{-\mathrm{i}\omega t})$$

où c désigne la célérité du son, α le coefficient de dissipation (supposé strictement positif) et f la densité volumique de sources acoustiques (à valeurs complexes). Si le bord de la cavité est rigide et vibre à la pulsation ω, on doit imposer une condition de glissement des particules sur la paroi qui s'exprime sous la forme suivante :

$$\frac{\partial \mathcal{P}}{\partial n} = \mathfrak{Re}(g(x)e^{-\mathrm{i}\omega t}) \text{ sur } \partial\Omega.$$

Il en résulte que la fonction u est solution du problème suivant : trouver une fonction u telle que

$$\begin{cases} -\Delta u + \nu\, u = f & \text{dans} \quad \Omega, \\ \dfrac{\partial u}{\partial n} = g & \text{sur} \quad \partial\Omega, \end{cases} \tag{7.2}$$

avec

$$\nu = -\frac{\omega^2}{c^2} - \mathrm{i}\omega\alpha.$$

Ce problème est dit *de Neumann.* L'équation aux dérivées partielles satisfaite par u est appelée *équation de Helmholtz.* La condition imposée à u sur le bord du domaine est appelée une *condition aux limites de type Neumann (non homogène* si $g \neq 0$ et *homogène* sinon).

7.1.3 Autres exemples

Nous avons présenté ci-dessus deux exemples fondamentaux de problèmes aux limites elliptiques. En réalité, il existe une grande variété de problèmes issus de la physique, conduisant à des équations aux dérivées partielles elliptiques différentes ou à des conditions aux limites différentes.

Ainsi par exemple, dans un milieu non homogène, les coefficients de l'équation aux dérivées partielles ne sont pas constants, mais dépendent des coordonnées d'espace. L'équation de Poisson prend alors la forme suivante :

$$\mathrm{div}\,(\mu \nabla u) = \rho f,$$

où μ et ρ désignent deux fonctions de la variable x. Ce sont des données du problème.

Bien entendu, les conditions aux limites ne sont pas toujours du même type sur tout le bord de la cavité. Ainsi, on peut avoir une condition de Dirichlet sur une partie du bord et une condition de Neumann sur l'autre partie. On dit alors que l'on a affaire à un problème mixte Dirichlet-Neumann.

De plus, les conditions aux limites ne sont pas toutes du type Dirichlet ou Neumann. Par exemple, on peut écrire sur le bord du domaine une condition qui relie la valeur de l'inconnue à sa dérivée normale. C'est le cas de la condition de *Fourier* (également appellée condition de *Robin*) qui s'écrit sous la forme

$$\alpha \frac{\partial u}{\partial n} + \beta u = g \text{ sur } \partial\Omega,$$

où α et β sont des coefficients donnés. C'est une condition intermédiaire entre la condition de Dirichlet et celle de Neumann : pour $\alpha = 0$ on retrouve la condition de Dirichlet, et pour $\beta = 0$ celle de Neumann.

Notons enfin que dans les exemples que nous avons considérés, l'inconnue du problème était toujours une fonction scalaire, à valeurs réelles (pour le potentiel électrostatique) ou complexes (pour la pression acoustique en régime harmonique). Il existe des problèmes pour lesquels l'inconnue est un champ de vecteurs. C'est le cas par

exemple des équations d'équilibre d'un corps élastique, dont la solution est le champ de déplacement. C'est aussi le cas des équations de Maxwell dont la solution est le champ électromagnétique.

7.2 Formulations variationnelles

Notre objectif est maintenant de montrer que les problèmes que nous avons introduits dans le paragraphe précédent sont *bien posés* à condition de choisir un cadre fonctionnel adéquat. On dit qu'un problème est bien posé lorsqu'il admet une solution, que cette solution est unique et qu'elle dépend continûment des données du problème. Autrement dit, pour montrer qu'un problème est bien posé, il faut établir un résultat d'*existence*, un résultat d'*unicité* et un résultat de *stabilité*.

Jusqu'à présent, nous sommes restés peu précis dans l'énoncé des problèmes de Dirichlet (7.1) et de Neumann (7.2). En effet, nous n'avons pas dit quelle régularité devait avoir la solution u. On pourrait être tenté de chercher une solution à ces problèmes dans un espace du type $\mathcal{C}^2(\overline{\Omega})$, ce qui reviendrait à chercher des solutions *classiques* de (7.1) et de (7.2) au sens où les dérivées secondes seraient interprétées au sens classique. Une bonne raison de ne pas le faire est qu'une théorie assurant la caratère bien posé de ces problèmes dans l'espace $\mathcal{C}^2(\overline{\Omega})$ n'existe pas ! La théorie que nous allons présenter ici assure la caractère bien posé de ces problèmes et conduit naturellement, comme on va le voir, à travailler dans le cadre des espaces de Sobolev définis au chapitre précédent.

Dans ce paragraphe, nous allons montrer que l'on peut écrire une formulation *faible* équivalente aux problèmes (7.1) et (7.2). Le terme *faible* signifie que le degré de dérivation de la solution a été abaissé. Plus précisément, la première formulation dite *forte* fait intervenir l'opérateur Laplacien qui est d'ordre 2, alors que la formulation faible ne comprendra que des dérivées d'ordre 1.

La démarche à suivre pour obtenir cette formulation faible n'est pas la même, suivant que l'on considère le problème de Dirichlet ou celui de Neumann. Nous les traiterons donc séparément. De plus pour les exemples physiques particuliers que nous avons considérés, la solution est à valeurs réelles pour le problème de Dirichlet et à valeurs complexes pour le problème de Neumann. Ces deux exemples nous permettront donc d'illustrer la méthode variationnelle dans un espace de Hilbert réel ou complexe. Mais au préalable nous examinons un cas simplifié dans $\mathbb{R}^d$ tout entier, ce qui permet de s'affranchir des conditions aux limites.

7.2.1 Un cas simplifié dans l'espace libre

On considère le problème suivant, pour une distribution f : trouver les distributions u à valeurs réelles satisfaisant (au sens des distributions) l'équation

$$-\Delta u + u = f \text{ dans } \mathbb{R}^d. \tag{7.3}$$

Les deux membres de l'équation ci-dessus étant des éléments de $\mathcal{D}'(\mathbb{R}^d)$, on peut leur appliquer une fonction test $\varphi \in \mathcal{D}(\mathbb{R}^d)$, ce qui donne

$$\langle \sum_{j=1,\dots,d} -\frac{\partial^2 u}{\partial {x_j}^2}, \varphi\rangle + \langle u, \varphi\rangle = \langle f, \varphi\rangle,$$

soit

$$\sum_{j=1,\dots,d} \langle \frac{\partial u}{\partial x_j}, \frac{\partial \varphi}{\partial x_j}\rangle + \langle u, \varphi\rangle = \langle f, \varphi\rangle.$$

Limitons-nous maintenant au cas où la donnée f est dans $L^2(\mathbb{R}^d)$ et supposons que u est solution de l'équation (7.3) dans $H^1(\mathbb{R}^d)$. On peut alors transformer les crochets de distribution en intégrales, soit

$$\int_{\mathbb{R}^d} \nabla u \cdot \nabla \varphi \, \mathrm{d}x + \int_{\mathbb{R}^d} u \, \varphi \, \mathrm{d}x = \int_{\mathbb{R}^d} f \, \varphi \, \mathrm{d}x.$$

Cette équation est de la forme

$$a(u, \varphi) = \ell(\varphi), \quad \forall \varphi \in \mathcal{D}(\mathbb{R}^d), \tag{7.4}$$

où a est le produit scalaire de $H^1(\mathbb{R}^d)$ et ℓ est l'application qui à un élement v de $H^1(\mathbb{R}^d)$ associe son produit scalaire dans $L^2(\mathbb{R}^d)$ avec f. Il est clair que a est une forme bilinéaire continue de $H^1(\mathbb{R}^d)$. Vérifions que ℓ est une forme linéaire continue sur $H^1(\mathbb{R}^d)$. En utilisant l'inégalité de Cauchy-Schwarz dans $L^2(\mathbb{R}^d)$, on a

$$|\ell(v)| = \left| \int_{\mathbb{R}^d} f \, v \, \mathrm{d}x \right| \leq ||f||_{L^2(\mathbb{R}^d)} ||v||_{L^2(\mathbb{R}^d)} \leq ||f||_{L^2(\mathbb{R}^d)} ||v||_{H^1(\mathbb{R}^d)}. \tag{7.5}$$

En utilisant la densité de $\mathcal{D}(\mathbb{R}^d)$ dans $H^1(\mathbb{R}^d)$ (voir théorème 6.5) on peut maintenant montrer la validité de l'équation (7.4) pour $v \in H^1(\mathbb{R}^d)$ et pas seulement $\varphi \in \mathcal{D}(\mathbb{R}^d)$. Soit $v \in H^1(\mathbb{R}^d)$ et considérons en effet une suite (φ_n) d'éléments de $\mathcal{D}(\mathbb{R}^d)$ convergeant vers v dans $H^1(\mathbb{R}^d)$. Par la continuité de a et de ℓ, on a pour deux constantes C_u et C_f

$$|a(u, \varphi_n - v)| \leq C_u ||\varphi_n - v||_{H^1(\mathbb{R}^d)}$$

et

$$|\ell(\varphi_n - v)| \leq C_f ||\varphi_n - v||_{H^1(\mathbb{R}^d)},$$

de sorte que $a(u, \varphi_n) \to a(u, v)$ et $\ell(\varphi_n) \to \ell(v)$ quand $n \to +\infty$. On voit donc qu'une solution $u \in H^1(\mathbb{R}^d)$ de l'équation (7.3) satisfait

$$\int_{\mathbb{R}^d} \nabla u \cdot \nabla v \, \mathrm{d}x + \int_{\mathbb{R}^d} u \, v \, \mathrm{d}x = \int_{\mathbb{R}^d} f \, v \, \mathrm{d}x, \quad \forall v \in H^1(\mathbb{R}^d). \tag{7.6}$$

L'équation (7.6) s'appelle une *formulation variationnelle*, dans laquelle la fonction v joue le rôle de *fonction test*. Comme on l'avait annoncé, cette formulation ne fait plus

intervenir que des dérivées d'ordre 1. On remarque par ailleurs que la fonction test vit dans le même espace que la solution : ici, l'espace $H^1(\mathbb{R}^d)$. Ceci garantit le fait qu'il y ait *autant* d'équations (à une fonction test correspond une équation) que d'inconnues. Par analogie avec la dimension finie, on peut penser qu'il s'agit d'une condition nécessaire pour que le problème soit bien posé.

Réciproquement, supposons que $u \in H^1(\mathbb{R}^d)$ soit solution de l'équation variationnelle (7.6). En choisissant $v = \varphi \in \mathcal{D}(\mathbb{R}^d) \subset H^1(\mathbb{R}^d)$ et en déroulant les étapes précédentes en sens inverse, on voit que u satisfait l'équation 7.3 au sens des distributions. On résume les résultats précédents en énonçant le

Théorème 7.1 *Pour $f \in L^2(\mathbb{R}^d)$, une fonction $u \in H^1(\mathbb{R}^d)$ est solution de l'équation (7.3) au sens des distributions si et seulement si elle satisfait la formulation variationnelle (7.6).*

Ce théorème établit l'équivalence entre l'équation (7.3) et la formulation variationnelle (7.6) pour une fonction u dans $H^1(\mathbb{R}^d)$. Il nous restera donc, pour montrer le caractère bien posé de (7.3), à vérifier le caractère bien posé de (7.6). Cela sera une conséquence du théorème de Lax-Milgram que nous montrerons plus loin.

7.2.2 Le problème de Dirichlet

Considérons maintenant un ouvert borné $\Omega \subset \mathbb{R}^d$ de classe $\mathcal{C}^1$ au sens de la définition 2.31. Supposons que pour $f \in L^2(\Omega)$, la fonction $u \in H^1(\Omega)$ soit solution du problème de Dirichlet (7.1), l'équation

$$-\Delta u = f \text{ dans } \Omega \tag{7.7}$$

étant satisfaite au sens des distributions et la condition aux limites $u = 0$ sur $\partial\Omega$ étant satisfaite au sens de la trace. En effet, la trace de la fonction u sur le bord $\partial\Omega$ a bien un sens comme fonction de $L^2(\partial\Omega)$ en vertu du théorème 6.10. Comme de plus cette trace est nulle, on a $u \in H^1_0(\Omega)$ en vertu du théorème 6.14. En appliquant les deux membres de l'équation (7.7), qui sont des éléments de $\mathcal{D}'(\Omega)$, à une fonction test $\varphi \in \mathcal{D}(\Omega)$, on obtient

$$\langle -\Delta u, \varphi \rangle = \langle f, \varphi \rangle,$$

et en procédant comme dans l'exemple précédent il vient

$$\int_\Omega \nabla u \cdot \nabla \varphi \, \mathrm{d}x = \int_\Omega f \, \varphi \, \mathrm{d}x.$$

On a donc à nouveau

$$a(u, \varphi) = \ell(\varphi), \quad \forall \varphi \in \mathcal{D}(\Omega),$$

la forme bilinéaire a et la forme linéaire ℓ étant définies sur $H^1_0(\Omega)$ par

$$a(u, v) = \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x, \quad \ell(v) = \int_\Omega f \, v \, \mathrm{d}x.$$

La forme bilinéaire $a(u,v)$ est continue : en effet, en appliquant l'inégalité de Cauchy-Schwarz, on obtient

$$|a(u,v)| \leq \left(\int_\Omega |\nabla u|^2 \,\mathrm{d}x\right)^{1/2} \left(\int_\Omega |\nabla v|^2 \,\mathrm{d}x\right)^{1/2} \leq \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}, \tag{7.8}$$

pour tout couple $(u,v) \in H_0^1(\Omega) \times H_0^1(\Omega)$. La forme linéaire $\ell(v)$ est une forme linéaire continue : en effet, en appliquant à nouveau l'inégalité de Cauchy-Schwarz, on vérifie que

$$|\ell(v)| \leq \left(\int_\Omega |f|^2 \,\mathrm{d}x\right)^{1/2} \left(\int_\Omega |v|^2 \,\mathrm{d}x\right)^{1/2} \leq C_\ell \|v\|_{H^1(\Omega)}, \tag{7.9}$$

pour tout $v \in H_0^1(\Omega)$, où l'on a posé

$$C_\ell = \|f\|_{L^2(\Omega)}.$$

Puisque l'espace $H_0^1(\Omega)$ est défini comme l'adhérence de $\mathcal{D}(\Omega)$ dans $H^1(\Omega)$, en procédant par densité comme dans l'exemple précédent il vient que $u \in H_0^1(\Omega)$ satisfait

$$\int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x = \int_\Omega f\, v \,\mathrm{d}x \quad \forall v \in H_0^1(\Omega). \tag{7.10}$$

Réciproquement, supposons que $u \in H_0^1(\Omega)$ soit solution du problème (7.10). La trace de u sur $\partial\Omega$ est nulle par le théorème 6.14 et en choisissant $v = \varphi \in \mathcal{D}(\Omega)$ dans (7.10) on obtient l'équation (7.7), satisfaite au sens des distributions. Résumons par le

Théorème 7.2 *Pour $f \in L^2(\Omega)$, une fonction $u \in H^1(\Omega)$ est solution du problème de Dirichlet (7.1) si et seulement si $u \in H_0^1(\Omega)$ et satisfait la formulation variationnelle (7.10).*

On remarque que dans la formulation variationnelle (7.10) la condition aux limites de Dirichlet est imposée *fortement* à la solution (par opposition à la condition de Neumann, comme on le verra plus loin). Elle intervient dans la définition de l'espace fonctionnel. On dit que c'est une condition *essentielle.*

Remarque 7.3 *On pourra noter que le problème (7.1) n'a plus de sens en général pour une fonction de $H^1(\Omega)$ si le domaine Ω est un ouvert quelconque, c'est-à-dire si on n'impose pas une certaine régularité à Ω, tout simplement parce que la trace sur le bord $\partial\Omega$ n'est alors pas définie. En revanche, la formulation variationnelle (7.10) a toujours un sens, car l'espace $H_0^1(\Omega)$ est défini pour un ouvert Ω quelconque.*

On peut d'ailleurs déjà remarquer que l'unicité du problème variationnel (7.10) est assurée. En effet, si u et $\tilde{u}$ sont deux solutions du problème (7.10), alors $u_0 = u - \tilde{u}$ est solution du problème homogène (i.e. correspondant à un second membre f nul) et vérifie :

$$\begin{cases} u_0 \in H_0^1(\Omega) \\ \displaystyle\int_\Omega \nabla u_0 \cdot \nabla v \,\mathrm{d}x = 0 \text{ pour tout } v \in H_0^1(\Omega). \end{cases}$$

En prenant $v = u_0$, on trouve que le gradient de u_0 est nul presque partout. Ceci prouve que u_0 est une fonction constante, mais comme elle s'annule sur le bord, elle est nécessairement identiquement nulle ; cela prouve l'unicité. L'existence et la stabilité résulteront du théorème de Lax-Milgram.

7.2.3 Le problème de Neumann

Le problème est à nouveau posé dans un ouvert $\Omega \subset \mathbb{R}^d$ de classe $\mathcal{C}^1$ au sens de la définition 2.31. Supposons maintenant, pour $f \in L^2(\Omega)$ et $g \in L^2(\partial\Omega)$, l'existence d'une solution $u \in H^2(\Omega)$ (à valeurs complexes) au problème de Neumann (7.2). La régularité de u (qui est $H^2(\Omega)$ au lieu de $H^1(\Omega)$ seulement) donne un sens à la condition aux limites $\partial u/\partial n = g$ sur le bord $\partial\Omega$, les deux membres de l'égalité ayant un sens dans $L^2(\partial\Omega)$. Si v est une fonction test appartenant à $H^1(\Omega)$, la formule de Green (6.2) conduit à l'identité suivante :

$$\int_\Omega (\nabla u \cdot \nabla \overline{v} + \nu\, u\, \overline{v})\ \mathrm{d}x - \int_{\partial\Omega} \frac{\partial u}{\partial n} \overline{v}\, \mathrm{d}\sigma = \int_\Omega f\, \overline{v}\, \mathrm{d}x.$$

On peut alors transformer le terme de bord en utilisant la condition aux limites. On trouve ainsi que $u \in H^2(\Omega)$ est solution du problème variationnel suivant :

$$\int_\Omega (\nabla u \cdot \nabla \overline{v} + \nu\, u\, \overline{v})\ \mathrm{d}x = \int_\Omega f\, \overline{v}\, \mathrm{d}x + \int_{\partial\Omega} g\, \overline{v}\, \mathrm{d}\sigma, \quad \forall v \in H^1(\Omega). \tag{7.11}$$

Réciproquement, supposons que $u \in H^2(\Omega)$ est solution de la solution variationnelle (7.11). On doit distinguer deux étapes dans la démarche :

(i) On choisit tout d'abord la fonction test $v = \varphi$ dans $\mathcal{D}(\Omega)$. On trouve alors que

$$\int_\Omega (\nabla u \cdot \nabla \overline{\varphi} + \nu\, u\, \overline{\varphi})\ \mathrm{d}x = \int_\Omega f\, \overline{\varphi}\, \mathrm{d}x,$$

car le terme de bord est nul dans ce cas. Ceci permet de montrer que l'équation de Helmholtz

$$-\Delta u + \nu\, u = f \tag{7.12}$$

est bien satisfaite au sens des distributions.

(ii) Il nous faut maintenant vérifier que u satisfait la condition aux limites sur le bord $\partial\Omega$. Comme nous avons supposé que $u \in H^2(\Omega)$, à l'aide de la formule de Green (6.2) on trouve alors que pour tout $v \in H^1(\Omega)$,

$$\int_\Omega (-\Delta u + \nu u)\, \overline{v}\, \mathrm{d}x + \int_{\partial\Omega} \frac{\partial u}{\partial n} \overline{v}\, \mathrm{d}\sigma = \int_\Omega f \overline{v}\, \mathrm{d}x + \int_{\partial\Omega} g\, \overline{v}\, \mathrm{d}\sigma.$$

D'après (7.12), cela donne

$$\int_{\partial\Omega} \left(\frac{\partial u}{\partial n} - g \right) \overline{v}\, \mathrm{d}\sigma = 0,$$

pour tout $v \in H^1(\Omega)$. Mais l'image de l'application trace sur $H^1(\Omega)$ est dense dans $L^2(\partial\Omega)$, ce qui implique

$$\frac{\partial u}{\partial n} = g \text{ sur } \partial\Omega.$$

On a montré le

Théorème 7.4 *Pour $f \in L^2(\Omega)$ et $g \in L^2(\partial\Omega)$, une fonction $u \in H^2(\Omega)$ est solution du problème de Neumann (7.2) si et seulement si elle satisfait la formulation variationnelle (7.11).*

Contrairement à la condition de Dirichlet, la condition de Neumann n'intervient pas dans l'espace fonctionnel (on ne peut d'ailleurs pas parler de la dérivée normale d'une fonction quelconque de $H^1(\Omega)$) mais elle est une conséquence de la formulation variationnelle (7.11) : on dit que c'est une condition *naturelle*.

Remarque 7.5 *Il peut paraître surprenant au lecteur d'avoir supposé $u \in H^2(\Omega)$ dans le théorème précédent alors que la formulation variationnelle (7.11) ne requiert que la régularité $H^1(\Omega)$. Cette régularité a été nécessaire pour donner un sens à la trace $\partial u/\partial n$ sur le bord $\partial\Omega$ (comme fonction de $L^2(\partial\Omega)$) et rendre licite la formule de Green (6.2). En fait, on peut montrer que le théorème est encore vérifié pour une fonction $u \in H^1(\Omega)$: il faut pour cela montrer que la trace $\partial u/\partial n$ sur le bord $\partial\Omega$ a un sens (plus faible) pour les fonctions u de $H^1(\Omega)$ à Laplacien dans $L^2(\Omega)$ et étendre la formule de Green (6.2) pour de telles fonctions.*

7.2.4 Formulations variationnelles et énergie

Pour comprendre le terme *variationnel,* considérons par exemple le problème de Dirichlet (7.1) et introduisons la fonctionnelle suivante :

$$J(v) = \frac{1}{2}\int_\Omega |\nabla v|^2 \,\mathrm{d}x - \int_\Omega f\, v \,\mathrm{d}x.$$

Il s'agit en fait de l'énergie électrostatique contenue dans Ω. Nous allons établir le

Lemme 7.6 *La fonction $u \in H_0^1(\Omega)$ est solution du problème (7.1) si et seulement si u minimise la fonctionnelle J sur l'espace $H_0^1(\Omega)$.*

Démonstration. Soit $u \in H_0^1(\Omega)$ une solution de (7.1), alors pour tout $v \in H_0^1(\Omega)$ on a

$$\begin{aligned} J(u+v) &= \frac{1}{2}\int_\Omega |\nabla(u+v)|^2\, dx - \int_\Omega f(u+v)\,\mathrm{d}x \\ &= J(u) + \left(\int_\Omega \nabla u\cdot\nabla v\,\mathrm{d}x - \int_\Omega f\, v\,\mathrm{d}x\right) + \frac{1}{2}\int_\Omega |\nabla v|^2\,\mathrm{d}x. \end{aligned}$$

Comme u est solution de (7.1), le terme entre parenthèses est nul, d'où

$$J(u+v) \geq J(u),$$

pour tout $v \in H_0^1(\Omega)$, ce qui montre que u minimise la fonctionnelle J.

Réciproquement, si u minimise la fonctionnelle J, alors $J(u+tv) - J(u) \geq 0$ pour tout $v \in H_0^1(\Omega)$ et tout $t \in \mathbb{R}$. L'identité précédente montre alors que

$$t\left(\int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x - \int_\Omega f\, v \,\mathrm{d}x\right) + \frac{t^2}{2}\int_\Omega |\nabla v|^2 \,\mathrm{d}x \geq 0,$$

pour tout $t \in \mathbb{R}$. En divisant par t pour $t > 0$ et $t < 0$ puis en faisant tendre t vers 0, on obtient que le terme entre parenthèses est nul, et on vérifie que u est bien solution de (7.1). □

En fait, imposer à u d'être solution du problème variationnel revient à imposer à la fonctionnelle J d'être *stationnaire* au voisinage de u. En effet, pour tout v et pour t proche de 0, on a au premier ordre

$$J(u+tv) - J(u) \simeq t\left(\int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x - \int_\Omega f\, v \,\mathrm{d}x\right).$$

Autrement dit, la formulation *variationnelle* est obtenue en calculant les petites *variations* de l'énergie J.

7.3 Le théorème de Lax-Milgram

Nous avons vu dans le paragraphe précédent trois exemples de problèmes variationnels, les deux derniers étant issus de la physique. Plus généralement, nous appellerons *formulation variationnelle* tout problème du type suivant :

$$\left\{\begin{array}{l} \text{Trouver } u \in V \text{ tel que} \\ a(u,v) = \ell(v) \text{ pour tout } v \in V \end{array}\right. \tag{7.13}$$

où

(i) V est un espace de Hilbert réel (ou complexe) dont le produit scalaire est noté $(u,v)_V$ et la norme associée $\|u\|_V$.

(ii) $a(u,v)$ est une forme bilinéaire (ou sesquilinéaire) continue sur $V \times V$, ce qui signifie qu'il existe une constante C_a telle que

$$|a(u,v)| \leq C_a \|u\|_V \|v\|_V,$$

pour tout couple $(u,v) \in V \times V$.

(iii) $\ell(v)$ est une forme linéaire (ou antilinéaire) continue sur V, ce qui signifie qu'il existe une constante C_ℓ telle que

$$|\ell(v)| \leq C_\ell \|v\|_V,$$

pour tout $v \in V$.

Le théorème de Lax-Milgram, que nous allons maintenant énoncer et démontrer, permet d'établir le caractère bien posé d'un problème de ce type lorsque la forme bilinéaire (ou sesquilinéaire) $a(u, v)$ vérifie une propriété supplémentaire : il s'agit de la coercivité que nous définissons ci-dessous.

Définition 7.7 *La forme bilinéaire (ou sesquilinéaire) $a(u, v)$ est dite coercive sur V s'il existe une constante strictement positive α telle que*

$$|a(v, v)| \geq \alpha \|v\|_V^2 \text{ pour tout } v \in V.$$

Remarque 7.8 *Si V est un espace de Hilbert réel ou si $a(u, v)$ est une forme hermitienne, il suffit pour établir la coercivité de montrer l'existence d'une constante strictement positive α telle que :*

$$a(v, v) \geq \alpha \|v\|_V^2 \text{ pour tout } v \in V.$$

La coercivité est souvent définie de cette manière dans la littérature.

Théorème 7.9 (de Lax-Milgram) *Si la forme bilinéaire (ou sesquilinéaire) $a(u, v)$ est coercive sur V, alors le problème (7.13) admet une solution unique $u \in V$ et il existe une constante C indépendante de la forme ℓ telle que :*

$$\|u\|_V \leq C \, |||\ell||| \tag{7.14}$$

Démonstration. Notons tout d'abord que l'estimation de stabilité (7.14) est immédiate. En effet, si u est la solution du problème, on a en vertu de la coercivité de la forme $a(u, v)$ et de la continuité de la forme $\ell(v)$:

$$\alpha \|u\|_V^2 \leq |a(u, u)| = |\ell(u)| \leq \left(\sup_{v \in V, v \neq 0} \frac{|\ell(v)|}{\|v\|_V} \right) \|u\|_V = |||\ell||| \, \|u\|_V.$$

On obtient donc (7.14) avec $C = 1/\alpha$.

La démonstration du caractère bien posé est plus délicate. Elle s'appuie essentiellement sur le théorème de Riesz vu au chapitre précédent. En effet, on remarque que d'après ce théorème :

(i) Comme ℓ est une forme linéaire (ou antilinéaire) continue sur V, il existe un unique $f \in V$ tel que

$$\ell(v) = (f, v)_V \text{ pour tout } v \in V.$$

(ii) Comme pour tout $u \in V$, l'application $v \to a(u,v)$ est une forme linéaire (ou antilinéaire) continue sur V, il existe un unique $w_u \in V$ tel que

$$a(u,v) = (w_u, v)_V \text{ pour tout } v \in V.$$

L'application A qui à u associe w_u est une application linéaire continue de V dans V. En effet

$$\|Au\|_V^2 = (w_u, w_u)_V = a(u, w_u) \leq C_a \|u\|_V \|w_u\|_V,$$

d'où :

$$\|Au\|_V \leq C_a \|u\|_V \text{ pour tout } u \in V.$$

Le problème (7.13) s'écrit maintenant

$$\begin{cases} \text{Trouver } u \in V \text{ tel que} \\ (Au, v)_V = (f, v)_V \text{ pour tout } v \in V, \end{cases}$$

soit encore

$$\text{Trouver } u \in V \text{ tel que } Au = f.$$

Autrement dit, pour démontrer le théorème, il suffit de montrer que l'opérateur A est bijectif. Comme V est a priori un espace de dimension infinie, nous allons montrer successivement l'injectivité et la surjectivité de A.

(i) Montrons que A est injectif. En effet, la coercivité de la forme $a(u,v)$ permet de montrer que, pour tout $u \in V$:

$$\alpha \|u\|_V^2 \leq |a(u,u)| = |(Au, u)_V| \leq \|Au\|_V \|u\|_V,$$

d'où

$$\|Au\|_V \geq \alpha \|u\|_V \text{ pour tout } u \in V. \tag{7.15}$$

(ii) Montrons que A est surjectif. A cette fin, considérons l'image de A, notée $R(A)$. Il s'agit d'un sous-espace vectoriel fermé de V : en effet si Au_n est une suite de Cauchy dans V alors d'après (7.15), u_n est également de Cauchy dans V. Par conséquent, u_n converge vers un élément u de V et, par continuité de A, Au_n converge vers Au. Autrement dit, la limite de la suite Au_n appartient bien à $R(A)$. De plus, $R(A)$ est dense dans V. En effet, si ce n'était pas le cas, on pourrait trouver un élément $w \in V$ non nul qui soit orthogonal à $R(A)$, on aurait alors $(Aw, w)_V = 0$, ce qui implique, d'après la coercivité de a, la nullité de w et aboutit à une contradiction. Il en résulte que $R(A) = V$, soit la surjectivité de l'opérateur A.

□

Remarque 7.10 *Il est important de noter que le théorème de Lax-Milgram est une généralisation du théorème de Riesz aux cas des formes bilinéaires non symétriques (espace de Hilbert réel) et des formes sesquilinéaires non hermitiennes (espace de Hilbert complexe). Autrement dit, dans le cas des formes bilinéaires symétriques (cas réel) et des formes sesquilinéaires hermitiennes (cas complexe), le théorème de Riesz est équivalent au théorème de Lax-Milgram. C'est en particulier le cas pour notre problème simplifié (7.6) ou le problème de Dirichlet (7.10). On notera en revanche que le théorème de Lax-Milgram est nécessaire dans le cas du problème de Neumann (7.11).*

7.4 Application aux problèmes aux limites elliptiques

Nous sommes maintenant en mesure, grâce au théorème de Lax-Milgram, de montrer que les différents problèmes introduits aux paragraphes 7.1 et 7.2 sont bien posés.

7.4.1 Problème simplifié dans l'espace libre

Considérons tout d'abord, pour $f \in L^2(\mathbb{R}^d)$, l'équation (7.3) posée dans $H^1(\mathbb{R}^d)$ (fonctions à valeurs réelles). En vertu du théorème 7.1, cette équation est équivalente à une formulation variationnelle de type (7.13) où

$$V = H^1(\mathbb{R}^d),\ a(u,v) = \int_{\mathbb{R}^d} (\nabla u \cdot \nabla v + u\, v)\ \mathrm{d}x \text{ et } \ell(v) = \int_{\mathbb{R}^d} f\, v\, \mathrm{d}x.$$

Il est clair que toutes les conditions du théorème de Max-Milgram sont réunies : V est un espace de Hilbert, $a(u,v)$ est une forme bilinéaire continue et coercice sur $V \times V$, s'agissant du produit scalaire de V, $\ell(v)$ est une forme linéaire continue sur V au vu de (7.5). On en déduit que la formulation variationnelle est bien posée, c'est-à-dire :

Théorème 7.11 *Pour tout $f \in L^2(\mathbb{R}^d)$, l'équation (7.3) a une solution unique u dans $H^1(\mathbb{R}^d)$ satisfaisant*

$$\|u\|_{H^1(\mathbb{R}^d)} \leq C\, \|f\|_{L^2(\mathbb{R}^d)},$$

la constante C étant indépendante de f.

Remarque 7.12 *Notons que le caractère bien posé de la formulation variationnelle ne nécessite pas que $f \in L^2(\mathbb{R}^d)$, il suffit que f soit dans le dual de $H^1(\mathbb{R}^d)$, soit $f \in H^{-1}(\mathbb{R}^d)$ (voir définition 6.6).*

7.4.2 Le problème de Dirichlet

Considérons maintenant le problème de Dirichlet homogène (7.1), dont l'inconnue u est une fonction à valeurs réelles et la donnée f appartient à $L^2(\Omega)$, le domaine Ω étant borné et de classe $\mathcal{C}^1$.

D'après le théorème 7.2, le problème (7.1) est équivalent à une formulation variationnelle de type (7.13) où

$$V = H_0^1(\Omega),\ a(u,v) = \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x \text{ et } \ell(v) = \int_\Omega f\, v \,\mathrm{d}x.$$

Avant tout, il nous faut vérifier que

(i) V est un espace de Hilbert : nous avons vu au chapitre précédent que $H_0^1(\Omega)$ est un espace de Hilbert pour la norme induite de $H^1(\Omega)$.

(ii) $a(u,v)$ est une forme bilinéaire continue sur $V \times V$: on l'a montré en (7.8).

(iii) $\ell(v)$ est une forme linéaire continue sur V : on l'a montré en (7.9).

Pour pouvoir appliquer le théorème de Lax-Milgram et affirmer que le problème (7.10) est bien posé, il suffit maintenant d'établir la coercivité de la forme bilinéaire a. Cela n'a rien d'évident : en effet, il s'agit de prouver l'existence d'une constante strictement positive α telle que

$$\int_\Omega |\nabla v|^2 \,\mathrm{d}x \geq \alpha \int_\Omega \left(|v|^2 + |\nabla v|^2\right) \,\mathrm{d}x,$$

pour tout $v \in H_0^1(\Omega)$.

Remarquons tout d'abord que ce résultat est faux pour un domaine Ω quelconque. Ainsi par exemple, supposons $\Omega = \mathbb{R}^d$ et soit $v \in \mathcal{D}(\Omega)$, alors la suite de fonctions

$$v_n(x) = \frac{1}{n^{d/2}} v\left(\frac{x}{n}\right)$$

contredit l'inégalité ci-dessus ; en effet

$$\int_\Omega |\nabla v_n|^2 \,\mathrm{d}x = \frac{1}{n^2} \int_\Omega |\nabla v|^2 \,\mathrm{d}x \text{ et } \int_\Omega |v_n|^2 \,\mathrm{d}x = \int_\Omega |v|^2 \,\mathrm{d}x.$$

En revanche, la coercivité de a est vérifiée dès que Ω est borné dans au moins une direction de l'espace (c'est le cas en particulier des domaines bornés). Cela résulte alors de l'inégalité de Poincaré qui fait l'objet du lemme suivant :

Lemme 7.13 (Inégalité de Poincaré) *Soit Ω un ouvert vérifiant : il existe R et $\alpha \in \mathbb{R}^d$, $\alpha \neq 0$, tels que*

$$\Omega \subset \left\{ x \in \mathbb{R}^d,\ -R < \sum_{i=1}^d \alpha_i x_i < R \right\}.$$

Alors il existe une constante C telle que

$$\int_\Omega |v|^2 \,\mathrm{d}x \leq C \int_\Omega |\nabla v|^2 \,\mathrm{d}x \quad \forall v \in H_0^1(\Omega). \tag{7.16}$$

Démonstration. Sans perte de généralité on peut supposer que $\alpha = (1, 0, \ldots, 0)$ (il suffit d'appliquer une rotation des axes), c'est-à-dire $\Omega \subset \{x \in \mathbb{R}^d,\ -R < x_1 < R\}$. Si $\psi \in \mathcal{D}(\Omega)$, alors, à condition de prolonger ψ par 0 à l'extérieur de Ω, on peut considérer que $\psi \in \mathcal{D}(\mathbb{R}^d)$ et l'on a

$$\psi(x_1, x_2, \ldots, x_d) = \int_{-\infty}^{x_1} \frac{\partial \psi}{\partial x_1}(t, x_2, \ldots, x_d)\, \mathrm{d}t,$$

pour tout $x \in \mathbb{R}^d$. En appppliquant l'inégalité de Cauchy-Schwarz, on montre alors que

$$|\psi(x_1, x_2, \ldots, x_d)|^2 \leq 2R \int_{\mathbb{R}} \left| \frac{\partial \psi}{\partial x_1}(t, x_2, \ldots, x_d) \right|^2 \mathrm{d}t.$$

En intégrant cette inégalité sur Ω, on obtient finalement

$$\int_\Omega |\psi|^2 \, \mathrm{d}x \leq 4R^2 \int_\Omega \left| \frac{\partial \psi}{\partial x_1} \right|^2 \mathrm{d}x.$$

A plus forte raison, on a donc établi l'inégalité (7.16) avec $C = 4R^2$ pour tout $\psi \in \mathcal{D}(\Omega)$. On peut alors conclure en remarquant que, par densité de $\mathcal{D}(\Omega)$ dans $H_0^1(\Omega)$, l'inégalité reste vraie pour tout $v \in H_0^1(\Omega)$. □

En conclusion, nous pouvons donc énoncer le résultat suivant :

Théorème 7.14 *Pour toute fonction $f \in L^2(\Omega)$, il existe une et une seule fonction $u \in H_0^1(\Omega)$ telle que $-\Delta u = f$ dans Ω (au sens des distributions), c'est-à-dire vérifiant le problème (7.1). De plus, il existe une constante C ne dépendant que du domaine Ω telle que*

$$\|u\|_{H^1(\Omega)} \leq C \|f\|_{L^2(\Omega)}.$$

Démonstration. Le domaine Ω est borné par hypothèse. L'inégalité de Poincaré est donc vérifiée, et la forme bilinéaire

$$a(u, v) = \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x$$

est donc coercive sur $H_0^1(\Omega)$: en effet, pour tout $v \in H_0^1(\Omega)$

$$a(v, v) = \int_\Omega |\nabla v|^2 \, \mathrm{d}x \geq \frac{1}{2C} \int_\Omega |v|^2 \, \mathrm{d}x + \frac{1}{2} \int_\Omega |\nabla v|^2 \, \mathrm{d}x \geq \min\left(\frac{1}{2}, \frac{1}{2C}\right) ||v||^2_{H^1(\Omega)}.$$

Les autres propriétés ayant déjà été vérifiées, on peut donc appliquer le théorème de Lax-Milgram à la formulation variationnelle (7.10) ce qui prouve l'existence d'une unique solution $u \in H_0^1(\Omega)$ à l'équation $-\Delta u = f$ dans Ω. □

7.4.3 Le problème de Neumann

Considérons maintenant le problème de Neumann (7.2), dont l'inconnue u est une fonction à valeurs complexes et dont les données f et g appartiennent respectivement à $L^2(\Omega)$ et $L^2(\partial\Omega)$, le domaine Ω étant borné et de classe $\mathcal{C}^1$. La formulation variationnelle correspondante est à nouveau de type (7.13) où

$$V = H^1(\Omega),\ a(u,v) = \int_\Omega (\nabla u \cdot \nabla \overline{v} + \nu\, u\, \overline{v})\ \mathrm{d}x \ \text{ et } \ \ell(v) = \int_\Omega f\, \overline{v}\, \mathrm{d}x + \int_{\partial\Omega} g\, \overline{v}\, \mathrm{d}\sigma.$$

A nouveau, il nous faut vérifier que

(i) V est un espace de Hilbert : c'est immédiat.

(ii) $a(u,v)$ est une forme sesquilinéaire continue sur $V \times V$: en effet, en appliquant l'inégalité de Cauchy-Schwarz, on obtient

$$|a(u,v)| \leq \left(\int_\Omega |\nabla u|^2\, \mathrm{d}x\right)^{1/2} \left(\int_\Omega |\nabla v|^2\, \mathrm{d}x\right)^{1/2} + |\nu| \left(\int_\Omega |u|^2\, \mathrm{d}x\right)^{1/2} \left(\int_\Omega |v|^2\, \mathrm{d}x\right)^{1/2},$$

d'où

$$|a(u,v)| \leq (1+|\nu|)\|u\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)},$$

pour tout couple $(u,v) \in V \times V$.

(iii) $\ell(v)$ est une forme linéaire continue sur V : en effet, en appliquant à nouveau l'inégalité de Cauchy-Schwarz, on vérifie que

$$|\ell(v)| \leq \left(\int_\Omega |f|^2\, \mathrm{d}x\right)^{1/2} \left(\int_\Omega |v|^2\, \mathrm{d}x\right)^{1/2} + \left(\int_{\partial\Omega} |g|^2\, \mathrm{d}\sigma\right)^{1/2} \left(\int_{\partial\Omega} |v|^2\, \mathrm{d}\sigma\right)^{1/2},$$

soit

$$|\ell(v)| \leq ||f||_{L^2(\Omega)}||v||_{L^2(\Omega)} + ||g||_{L^2(\partial\Omega)}||v||_{L^2(\partial\Omega)}.$$

On rappelle alors le résultat de continuité de l'application trace (voir le théorème 6.10) qui garantit l'existence d'une constante C telle que

$$||v||_{L^2(\partial\Omega)} \leq C\|v\|_{H^1(\Omega)}.$$

On a donc finalement

$$|\ell(\psi)| \leq C_\ell \|\psi\|_{H^1(\Omega)},$$

pour tout $v \in V$, où l'on a posé $C_\ell = \|f\|_{L^2(\Omega)} + C\|g\|_{L^2(\partial\Omega)}$.

Pour appliquer le théorème de Lax-Milgram nous allons maintenant établir la coercivité de la forme bilinéaire a.

Lemme 7.15 *Si $\nu \in \mathbb{C} \setminus \mathbb{R}^-$, alors la forme sesquilinéaire*

$$a(u,v) = \int_\Omega (\nabla u \cdot \nabla \overline{v} + \nu\, u\, \overline{v})\ \mathrm{d}x$$

est coercive sur $H^1(\Omega)$.

Démonstration. Remarquons tout d'abord que l'on peut écrire le nombre complexe ν sous la forme

$$\nu = \rho\, e^{i\theta} \text{ avec } -\pi < \theta < \pi.$$

On note $\sqrt{\nu}$ la racine complexe de ν suivante :

$$\sqrt{\nu} = \sqrt{\rho}\, e^{i\theta/2},$$

pour tout $v \in H^1(\Omega)$, on a alors

$$|a(v,v)| = \sqrt{\rho}\left|\frac{1}{\sqrt{\nu}}a(v,v)\right| \geq \sqrt{\rho}\ \mathfrak{Re}\left(\frac{1}{\sqrt{\nu}}a(v,v)\right).$$

Or

$$\mathfrak{Re}\left(\frac{1}{\sqrt{\nu}}a(v,v)\right) = \int_\Omega \mathfrak{Re}\left(\frac{1}{\sqrt{\nu}}\right)|\nabla v|^2\,\mathrm{d}x + \int_\Omega \mathfrak{Re}\left(\sqrt{\nu}\right)|v|^2\,\mathrm{d}x,$$

et il facile de voir que $\mathfrak{Re}\,(1/\sqrt{\nu})$ et $\mathfrak{Re}\,(\sqrt{\nu})$ sont deux nombres réels strictement positifs. Plus précisément, on obtient ainsi l'inégalité suivante :

$$|a(v,v)| \geq \min(1,\rho)\cos(\theta/2)\|v\|^2_{H^1(\Omega)},$$

pour tout $v \in H^1(\Omega)$. □

On peut conclure en énonçant le

Théorème 7.16 *Si $\nu \in \mathbb{C} \setminus \mathbb{R}^-$, le problème variationnel : trouver $u \in H^1(\Omega)$ satisfaisant (7.11) admet, pour toutes fonctions $f \in L^2(\Omega)$ et $g \in L^2(\partial\Omega)$, une et une seule solution $u \in H^1(\Omega)$, et il existe une constante C (indépendante des données f et g) telle que*

$$\|u\|_{H^1(\Omega)} \leq C\left(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\partial\Omega)}\right).$$

Si de plus la solution $u \in H^2(\Omega)$, alors elle est solution du problème (7.2).

Remarque 7.17 *Dans le cas où ν est égal à 0, on peut montrer que le problème variationnel précédent admet une solution, unique à une constante additive près, si et seulement si les données f et g satisfont la relation de compatibilité suivante :*

$$\int_\Omega f(x)\,\mathrm{d}x + \int_{\partial\Omega} g(x)\,\mathrm{d}\sigma = 0.$$

*Plus généralement, lorsque ν est un réel négatif, la question de l'existence et de l'unicité de la solution ne rentre plus dans le cadre du théorème de Lax-Milgram. En utilisant un résultat connu sous le nom d'*alternative de Fredholm*, on peut montrer que l'existence et l'unicité de la solution sont encore assurées sauf pour une suite de valeurs de ν qui correspondent à des fréquences propres du problème.*

Bibliographie

[1] **J.-M. Bony**, *Cours d'Analyse - Théorie des Distributions et Analyse de Fourier*, Editions de l'Ecole Polytechnique, 2001.

[2] **H. Brezis**, *Analyse Fonctionnelle. Théorie et Applications*, Masson, 1983.

[3] **W. Rudin**, *Real and Complex Analysis*, McGraw-Hill, 1966 (traduction française : *Analyse Réelle et Complexe - Cours et Exercices*, Dunod, 1988).

[4] **L. Schwartz**, *Méthodes Mathématiques pour les Sciences Physiques*, Masson, 1965.

[5] **M. Willem**, *Analyse Harmonique Réelle*, Hermann, 1995.

Index

anti-linéaire (forme), 90

Banach (espace de), 14
base hilbertienne, 97
Beppo–Levi (théorème de), 32
bien posé (problème), 114

$\mathcal{C}^0(D)$, 9
$\mathcal{C}^m(\Omega)$, $\mathcal{C}^m(\overline{\Omega})$, 10
$\mathcal{C}_0^\infty(\Omega)$, 49
Cauchy (suite de), 13
Cauchy-Schwarz (inégalité de), 81, 90
changement de variable, 37
coercive (forme bilinéaire), 121
complet (espace), 13
complété (d'un espace), 15
continuité
    d'une application, 11
    uniforme, 11
convergence
    d'une série, 18
    d'une suite, 11
    des distributions, 62
    dominée (théorème de), 32
    en moyenne, 10
    monotone (théorème de), 32
    normale, 18
    uniforme, 10
convexe (ensemble), 92
convolutifs (ensembles), 67
convolution
    dans $\mathcal{D}'(\mathbb{R}^d)$, 69
    dans $L^1(\mathbb{R}^d)$, 36, 80
    dans $L^1_{\text{loc}}(\mathbb{R}^d)$, 67
coordonnées
    polaires, 38
    sphériques, 39

$\mathcal{D}(\Omega)$, $\mathcal{D}'(\Omega)$, 49
$\mathcal{D}_K(\Omega)$, 49
$\mathcal{D}^0(\Omega)$, 25
définie (forme), 91
dense (ensemble), 15
dérivée
    d'une distribution, 53
    normale, 46
difféomorphisme, 37
Dirac (mesure de), 52, 54, 60, 63, 65, 71
Dirichlet (condition de), 112
distribution, 50
division (des distributions), 60
dualité (produit de), 50

équivalence (relation d'), 15
équivalentes (normes), 10

faible (formulation), 114
Fischer–Riesz (théorème de), 30
Fourier (transformation de)
    dans L^1, 74
    dans L^2, 87
frontière (d'un ouvert), 40
Fubini (théorème de), 34

gaussienne (fonction), 78
gradient, 46
Green (formule de), 45, 106

$H^m(\Omega)$, 101
$H_0^m(\Omega)$, 102
Heaviside (fonction de), 50
Helmholtz (équation de), 113
hermitienne (forme), 90

Hilbert (espace de), 91

intégrale
    de Lebesgue, 25
    de Riemann, 22, 31
    de surface, 41, 44
    impropre, 31
    semi-convergente, 31
intégration par parties, 45

jacobien, 37
jacobienne (matrice), 37

$\mathcal{L}^1(\Omega)$, 28
$L^1(\Omega)$, 29
$L^1_{\text{loc}}(\Omega)$, 51
$L^2(\Omega)$, 81
ℓ^2, 91
laplacien, 46
Lax-Milgram (théorème de), 121
Lebesgue (théorème de), 32
localement (intégrable), 51

mesurable (fonction), 26
Minkowski (inégalité de), 81, 91
multiplication (des distributions), 58
multipôle, 52, 54

négligeable (ensemble), 24
Neumann (condition de), 113
normale extérieure, 40, 43
norme, 10, 91

ordre (d'une distribution), 51
orthogonal (espace), 95

parallélogramme (identité du), 90
Parseval (identité de), 88
partition de l'unité, 44
Plancherel (identité de), 88
Poincaré (inégalité de), 124
Poisson (équation de), 112
presque partout, 23
produit scalaire, 91
projection, 93
prolongement (par densité), 17

restriction (d'une distribution), 64
Riemann-Lebesgue (théorème de), 74
Riesz (théorème de), 96

sauts (formule des), 55
semi-norme, 10, 51
séparable (espace), 97
séries, 18
sesquilinéaire (forme), 90, 120
Sobolev (espaces de), 101
Stokes (formule de), 40, 45
support
    d'une distribution, 64
    d'une fonction, 25

test (fonction), 48, 49
trace, 104

variationnelle (formulation), 114
vp$(1/x)$ (valeur principale), 52, 54, 60