

Analyse fonctionnelle et équations aux dérivées partielles

P. Cardaliaguet

Les démonstrations comportant le signe (*) sont à connaître.

Bibliographie : “Analyse fonctionnelle” H. Brézis, Masson.

Table des matières

1	Les espaces de Hilbert et le théorème de Lax-Milgram	5
1.1	Espaces vectoriels normés	5
1.1.1	Normes sur un espace vectoriel	5
1.1.2	Espaces complets	6
1.1.3	Applications linéaires continues	7
1.1.4	Produits d'EVN	9
1.2	Espaces de Hilbert	10
1.2.1	Produit scalaire et norme associée	10
1.2.2	Le théorème de projection	12
1.2.3	Orthogonalité	14
1.2.4	Bases hilbertiennes	16
1.2.5	Dualité	18
1.2.6	Le théorème de Lax-Milgram	20
2	Espaces de Sobolev et équations elliptiques linéaires.	22
2.1	Espaces de Sobolev sur un intervalle	22
2.1.1	Définitions et exemples	22
2.1.2	Approximation	25
2.1.3	L'espace $W_0^{1,p}(I)$	27
2.2	Application aux équations elliptiques en dimension 1	28
2.2.1	Existence et unicité d'une solution faible	29
2.2.2	Régularité de la solution	30
2.3	Application au problème avec conditions au bord de type Neumann	31
2.4	Espaces de Sobolev en dimension supérieure - formulation faible	32
2.4.1	Définitions	32
2.4.2	Exemples de formulation faible	34
3	Distributions tempérées et transformée de Fourier	35
3.1	Classe de Schwartz	35
3.2	Distributions tempérées	38
3.3	Applications à l'équation de Laplace	41

A	Brefs rappels d'intégration	43
A.1	Quelques résultats fondamentaux en intégration	43
A.2	Les espaces L^p	44
A.3	Intégration sur un espace produit	45
A.4	Un peu de calcul intégral	47
A.5	Produit de convolution	47
A.6	Régularisation	48

Introduction

De nombreux modèles en physique, chimie, biologie et économie sont régis par des équations : ce peut être des équations différentielles ordinaires, mais aussi des équations aux dérivées partielles. L'objet de ce cours est d'introduire quelques techniques simples d'analyse des équations aux dérivées partielles.

Equations différentielles ordinaires

Les équations différentielles ordinaires servent à modéliser des systèmes qui évoluent avec le temps. L'inconnue y est une fonction—scalaire ou vectorielle—qui ne dépend que de la variable temporelle.

Voici quelques exemples :

- Modèle logistique : $x(t)$ est la population à l'instant t .

$$x'(t) = x(t)(\alpha - x(t))$$

α = taux de reproduction.

- Modèle proie-prédateur (Lotka-Volterra) si $x(t)$ est la population de proies et $y(t)$ la population des prédateurs,

$$\begin{cases} x'(t) = x(t)(\alpha - y(t)) \\ y'(t) = -y(t)(\beta - x(t)) \end{cases}$$

- Plus généralement,

$$\begin{cases} x'(t) = f(t, x(t)) \\ x(t_0) = x_0 \end{cases}$$

où $x(t) = (x_1(t), \dots, x_N(t)) \in \mathbb{R}^N$, t est le temps et $f : \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^N$ champs de vecteurs. Enfin, t_0 est l'instant initial du système, $x_0 \in \mathbb{R}^N$ la position initiale.

Méthodes de résolution :

1. Calcul explicite (modèle logistique) ou semi-explicite (courbes intégrales)
2. Calcul numérique : nécessite en général une bonne connaissance des conditions d'existence, d'unicité, de stabilité des solutions.

Pour l'EDO donnée ci-dessus, on montre l'existence et l'unicité en utilisant un théorème de point-fixe : supposons par exemple que f soit continue et lipschitzienne en espace uniformément en temps :

$$\|f(t, x) - f(t, y)\| \leq C\|x - y\| \quad \forall x, y \in \mathbb{R}^N, t \in \mathbb{R}$$

Posons $X = \mathcal{C}^0([t_0 - \delta, t_0 + \delta])$ (où $\delta > 0$). Alors l'application $\Phi : X \rightarrow X$ définie par

$$\Phi(x)(t) = x_0 + \int_{t_0}^t f(s, x(s)) ds \quad \forall t \in [t_0 - \delta, t_0 + \delta]$$

est contractante pour $\delta > 0$ suffisamment petit. Comme X est un espace de Banach pour la norme $\|\cdot\|_\infty$, on en déduit que Φ possède un unique point fixe.

Exemples d'équations aux dérivées partielles

Dans les équations aux dérivées partielles, l'inconnue est une fonction qui dépend de plusieurs variables et l'équation lie les dérivées partielles de l'inconnue.

- Equation de transport : un fluide (1 dimensionnel) avance avec la vitesse c . A l'instant initial on ajoute un peu de sel (par exemple) dans ce fluide. Si $u(t, x)$ est la concentration de sel (gramme/litre par ex.), alors

$$\frac{\partial u}{\partial t}(t, x) + c \frac{\partial u}{\partial x}(t, x) = 0$$

On peut se fixer aussi la condition initiale $u(0, x) = u_0(x)$ (concentration initiale de solvant).

- Equation de la chaleur : si $u(t, x)$ est la température dans un fil infini de section nulle, alors la chaleur se propage suivant l'équation :

$$\frac{\partial u}{\partial t}(t, x) - \frac{\partial^2 u}{\partial x^2}(t, x) = 0$$

On peut se fixer aussi la condition initiale $u(0, x) = u_0(x)$ (température initiale).

- Equation des ondes : la hauteur $u(t, x)$ à l'instant t et à la position x d'une corde vibrante évolue suivant l'équation :

$$\frac{\partial^2 u}{\partial t^2}(t, x) + \frac{\partial^2 u}{\partial x^2}(t, x) = 0$$

On peut se fixer aussi la condition initiale $u(0, x) = u_0(x)$ (Position initiale de la corde).

- Equation de Black-Scholes : le prix $C(t, x)$ d'une option d'achat dépend du temps t et de la valeur x de l'option suivant l'équation :

$$\frac{\partial C}{\partial t}(t, x) + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 C}{\partial x^2}(t, x) + rx \frac{\partial C}{\partial x}(t, x) - rC(t, x) = 0 \quad x > 0, 0 < t < T$$

où T est la maturité de l'option, σ est la volatilité de l'actif sous-jacent et $r > 0$ de taux de l'actif sans risque. L'équation précédente est complétée par une condition terminale

$$C(T, x) = \max\{0, x - K\}$$

où K est le prix d'exercice.

Méthodes de résolution :

- Formules explicites : c'est le cas pour les modèles exposés ci-dessus.
 - Pour l'équation de transport, par exemple, on note que toute fonction réelle ϕ , la fonction $u(t, x) = \phi(x - ct)$ vérifie l'équation. Il suffit alors de prendre $\phi = u_0$.
 - Pour l'équation de la chaleur, il existe une *formule intégrale* :

$$u(t, x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{(x-y)^2}{2t}} u_0(y) dy \quad t > 0, x \in \mathbb{R}$$

qui fonctionne dès que u_0 est suffisamment régulière.

- Pour l'équation de Black-Scholes, formule assez lourde.
- Par contre, il n'existe pas de solution explicite pour la plupart des modèles plus complexes, comme les options européennes avec des taux et volatilités non constantes, des modèles avec des dividendes. Ce n'est pas le cas non plus pour les options américaines.
- Malheureusement, les techniques de résolution sont beaucoup plus complexes que pour les EDO. Il n'y a pas de méthode générale, qui marche pour toutes les équations.

Nous verrons dans ce cours des techniques hilbertiennes (espaces de Sobolev), ainsi qu'une approche très générale, fonctionnant pour les EDP linéaires : la théorie des distributions.

1 Les espaces de Hilbert et le théorème de Lax-Milgram

1.1 Espaces vectoriels normés

Dans tout le cours, \mathbb{K} désigne soit l'ensemble des nombres réels \mathbb{R} , soit l'ensemble des nombres complexes \mathbb{C} .

1.1.1 Normes sur un espace vectoriel

Définition 1.1 Soit E un espace vectoriel sur \mathbb{K} . On dit qu'une application $\|\cdot\| : E \rightarrow \mathbb{R}$ est une norme si

- $\|x\| \geq 0$ pour tout $x \in E$,
- $\forall x \in E, \|x\| = 0 \Rightarrow x = 0_E$,
- (positive homogénéité) $\forall x \in E, \forall \lambda \in \mathbb{K}, \|\lambda x\| = |\lambda| \|x\|$
- (inégalité triangulaire) $\forall x, y \in E, \|x + y\| \leq \|x\| + \|y\|$

Remarque 1.2 On dit que $(E, \|\cdot\|)$ est un espace vectoriel normé (EVN). La norme $\|\cdot\|$ définit naturellement une notion de distance sur E : on mesure la distance entre deux points x et y de E par $\|x - y\|$. Rappelons que cette distance induit les notions d'ouvert, fermé, compact, voisinage, etc...

Voici quelques exemples classiques d'EVN. D'autres seront étudiés en TD.

1. \mathbb{R}^N , muni d'une des normes suivantes, est un EVN : pour $p \in [1, +\infty[$,

$$\|x\|_p = \left(\sum_{i=1}^N |x_i|^p \right)^{1/p}, \quad \|x\|_\infty = \max_{i=1, \dots, N} |x_i|, \quad \text{où } x = (x_1, \dots, x_N) \in \mathbb{R}^N$$

2. Soit X un ensemble. L'espace vectoriel E des applications bornées de X dans \mathbb{K} peut être muni de la norme suivante :

$$\|f\|_\infty = \sup_{x \in X} |f(x)| \quad \forall f \in E.$$

3. Soit K est un sous-ensemble compact de \mathbb{R}^N . L'espace vectoriel E des applications continues de K dans \mathbb{K} peut être muni de la norme suivante :

$$\|f\|_\infty = \max_{x \in K} |f(x)| \quad \forall f \in E.$$

(rappelons que, puisque $x \rightarrow |f(x)|$ est continue, le maximum est atteint).

4. Soit ℓ^1 espace vectoriel des séries réelles absolument convergentes. Alors ℓ^1 peut-être muni de la norme

$$\|x\|_1 = \sum_{i=1}^{\infty} |x_i| \quad \forall x = (x_1, x_2, \dots) \in \ell^1$$

Définition 1.3 On dit que deux normes $\|\cdot\|_1$ et $\|\cdot\|_2$ sur un espace vectoriel E sont équivalentes s'il existe deux constantes $C_1, C_2 > 0$ telle que

$$C_1 \|x\|_1 \leq \|x\|_2 \leq C_2 \|x\|_1 \quad \forall x \in E.$$

On rappelle que “être équivalent à” est une relation d’équivalence (d’où la terminologie).

Théorème 1.4 *Si E est de dimension finie, toutes les normes sur E sont équivalentes.*

Lorsque E est de dimension infinie, ce résultat est toujours faux :

Théorème 1.5 *Soit E un espace vectoriel. Si toutes les normes sur E sont équivalentes, alors E est de dimension finie.*

En fait, en dimension infinie, il est rare que deux normes soient équivalentes.

1.1.2 Espaces complets

Définition 1.6 *Soit $(E, \|\cdot\|)$ un EVN sur \mathbb{K} . On dit qu’une suite (x_n) d’éléments de E est de Cauchy si*

$$\forall \epsilon > 0, \exists n_0 \in \mathbb{N}, \forall n \geq n_0, \forall p \geq 0 \quad \|x_n - x_{n+p}\| \leq \epsilon.$$

Rappelons que toute suite convergente est de Cauchy. Lorsque $(E, \|\cdot\|)$ est un EVN quelconque, la réciproque n’est pas forcément vraie.

Définition 1.7 *On dit qu’un espace vectoriel normé E , muni de la norme $\|\cdot\|$ est complet pour cette norme, si toute suite de Cauchy (pour cette norme) d’éléments de E converge. On dit aussi que E est un espace de Banach.*

- Remarque 1.8**
1. Si deux normes sont équivalentes sur un EV E et si E est complet pour l’une des normes, autre E est complet pour l’autre (exercice).
 2. Si E est de dimension finie, alors E est complet (pour toute norme).
 3. En dimension infinie, il est souvent essentiel pour les applications de travailler avec un espace complet (cf. la suite du cours).

Exemples :

1. Si X un ensemble et $(E, \|\cdot\|_\infty)$ est l’espace vectoriel des applications *bornées* de X dans \mathbb{K} être muni de la norme

$$\|f\|_\infty = \sup_{x \in X} |f(x)| \quad \forall f \in E,$$

alors E est un espace de Banach.

2. De même, si K est un sous-ensemble compact de \mathbb{R}^N et $(E, \|\cdot\|_\infty)$ est l’espace vectoriel des applications *continues* de K dans \mathbb{K} muni de la norme

$$\|f\|_\infty = \max_{x \in K} |f(x)| \quad \forall f \in E,$$

alors E est un espace de Banach.

Voici une condition nécessaire et suffisante pour être complet dans un sous-espace complet :

Proposition 1.9 *Soit $(E, \|\cdot\|)$ un EVN complet et $K \subset E$ non vide. Alors K est complet, si et seulement si, K est fermé.*

Preuve (*): Supposons K complet. Soit (x_n) une suite d'éléments de K qui admet une limite $x \in E$. Comme (x_n) converge, (x_n) est une suite de Cauchy (exercice). Comme K est complet, (x_n) possède une limite \bar{x} dans K . Or la limite d'une suite est unique, ce qui prouve que $x = \bar{x} \in K$. Donc K est fermé.

Supposons maintenant que K est fermé. Soit (x_n) une suite de Cauchy de K . Alors (x_n) est une suite de Cauchy dans E , qui est complet. Donc (x_n) admet une limite $x \in E$. Comme K est fermé et (x_n) est une suite d'éléments de K , la limite x est aussi dans K . Donc (x_n) possède une limite dans K , et K est complet. \square

1.1.3 Applications linéaires continues

Soit $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés sur \mathbb{K} .

Théorème 1.10 Soit $L : E \rightarrow F$ une **application linéaire**. Les assertions suivantes sont équivalentes :

- (i) L est continue sur E ,
- (ii) L est continue en 0_E ,
- (iii) il existe une constante K telle que

$$\|L(x)\|_F \leq K\|x\|_E \quad \forall x \in E$$

- (iv) L est lipschitzienne sur E (c'est-à-dire, il existe une constante $K \geq 0$ telle que $\|L(x) - L(y)\|_F \leq K\|x - y\|_E$ pour tout $x, y \in E$).

Preuve (*): Il est clair que (i) \Rightarrow (ii) et (iv) \Rightarrow (i). L'implication (iii) \Rightarrow (iv) est aussi très facile : soit K la constante de (iii). Alors, pour tout $x, y \in E$, on a par linéarité de L ,

$$\|L(x) - L(y)\|_F = \|L(x - y)\|_F \leq K\|x - y\|_E \quad (\text{par (iii)}).$$

D'où (iv).

Le seul point sur lequel il faut un peu travailler est (ii) \Rightarrow (iii) : comme, d'après (ii), L est continue en 0_E , pour $\epsilon = 1 > 0$ il existe une constante $\eta > 0$ telle que, si $\|y - 0_E\|_E \leq \eta$, alors $\|L(y) - L(0_E)\|_F \leq \epsilon = 1$. Cela se réécrit $\|L(y)\|_F \leq 1$ si $\|y\|_E \leq \eta$, puisque L est linéaire, et donc $L(0_E) = 0$. Soit maintenant $x \in E$ avec $x \neq 0_E$. Notons que $y = \eta \frac{x}{\|x\|_E}$ vérifie $\|y\|_E \leq \eta$, et donc $\|L(y)\|_F \leq 1$. On multiplie cette dernière inégalité par $\|x\|_E/\eta$ pour obtenir, par positive homogénéité de la norme puis linéarité de L :

$$(\|x\|_E/\eta)\|L(y)\|_F = \|(\|x\|_E/\eta)L(y)\|_F = \|L((\|x\|_E/\eta)y)\|_F = \|L(x)\|_F \leq \frac{\|x\|_E}{\eta}.$$

Cette inégalité étant évidente pour $x = 0_E$, il existe donc une constante $K = 1/\eta$ pour laquelle l'inégalité de (iii) a lieu. \square

Soit $\mathcal{L}(E, F)$ l'ensemble de applications *linéaires continues* de E dans F . Notons que $\mathcal{L}(E, F)$ est un \mathbb{K} -espace vectoriel.

Proposition 1.11 L 'espace vectoriel $\mathcal{L}(E, F)$ est muni de la norme

$$\|T\|_{\mathcal{L}(E, F)} = \sup_{x \neq 0_E} \frac{\|T(x)\|_F}{\|x\|_E}$$

Remarques :

1. Lorsque E et F sont de dimension finie, on retrouve la notion de norme matricielle.
2. On montre facilement (exercice) que

$$\|T\|_{\mathcal{L}(E,F)} = \sup \{ \|T(x)\|_F \mid x \in E, \|x\|_E \leq 1 \}$$

3. Par la suite nous travaillerons fréquemment avec l'espace $E^* := \mathcal{L}(E, \mathbb{K})$. Cet espace s'appelle le dual topologique de E .

Preuve de la proposition (*): Il est clair que $\mathcal{L}(E, F)$ est un espace vectoriel. Montrons que $\|\cdot\|_{\mathcal{L}(E,F)}$ est une norme.

- il est clair que $\|T\|_{\mathcal{L}(E,F)} \geq 0$ pour tout T ,
- Supposons que $\|T\|_{\mathcal{L}(E,F)} = 0$. Alors on a, pour tout $x \in E$, $\|T(x)\|_F \leq \|T\|_{\mathcal{L}(E,F)} = 0$, soit $T(x) = 0_F$. Donc T est l'application linéaire nulle.
- (positive homogénéité) Soit $T \in \mathcal{L}(E, F)$ et $\lambda \in \mathbb{K}$. Comme, pour tout $x \in E$, on a $\|\lambda T(x)\|_F = |\lambda| \|T(x)\|_F$, on en déduit que

$$\|\lambda T\|_{\mathcal{L}(E,F)} = \sup_{x \neq 0_E} |\lambda| \frac{\|T(x)\|_F}{\|x\|_E} = |\lambda| \sup_{x \neq 0_E} \frac{\|T(x)\|_F}{\|x\|_E} = |\lambda| \|T\|_{\mathcal{L}(E,F)}.$$

- (inégalité triangulaire) soient $T_1, T_2 \in \mathcal{L}(E, F)$. On utilise la définition équivalente donnée dans la remarque. Pour tout $x \in E$, avec $\|x\|_E \leq 1$. On a

$$\|(T_1 + T_2)(x)\|_F = \|T_1(x) + T_2(x)\|_F \leq \|T_1(x)\|_F + \|T_2(x)\|_F \leq \|T_1\|_{\mathcal{L}(E,F)} + \|T_2\|_{\mathcal{L}(E,F)}$$

En prenant le supremum sur x , avec $\|x\|_E \leq 1$, on obtient :

$$\|T_1 + T_2\|_{\mathcal{L}(E,F)} \leq \|T_1\|_{\mathcal{L}(E,F)} + \|T_2\|_{\mathcal{L}(E,F)}$$

□

Voici quelques propriétés élémentaires de cette norme : si $T \in \mathcal{L}(E, F)$, alors

- (i) $\|T(x)\|_F \leq \|T\|_{\mathcal{L}(E,F)} \|x\|_E \quad \forall x \in E$
- (ii) en particulier, $\|T(x) - T(y)\|_F \leq \|T\|_{\mathcal{L}(E,F)} \|x - y\|_E \quad \forall x, y \in E$

Théorème 1.12 Si F est un espace de Banach, alors $\mathcal{L}(E, F)$ est également un espace de Banach.

Remarque 1.13 En particulier le dual d'un EVN est toujours complet : rappelons que le dual E^* d'un EVN E est l'ensemble des formes linéaires continues sur E :

$$E^* = \{T : E \rightarrow \mathbb{R}, T \text{ linéaire continue}\}$$

Preuve : Supposons que (T_n) soit une suite de Cauchy dans $\mathcal{L}(E, F)$. Montrons d'abord que, pour tout $x \in E$, la suite $(T_n(x))$ est une suite de Cauchy dans F . Si $x = 0_E$, alors $T_n(x) = 0_F$, et le résultat est évident. Supposons maintenant que $x \neq 0$. Fixons $\epsilon > 0$. Comme (T_n) est de Cauchy, il existe $n_0 \in \mathbb{N}$ tel que

$$\forall n \geq n_0, \forall p \geq 0, \quad \|T_n - T_{n+p}\|_{\mathcal{L}(E,F)} \leq \frac{\epsilon}{\|x\|_E}$$

Par conséquent,

$$\forall n \geq n_0, \forall p \geq 0, \quad \|T_n(x) - T_{n+p}(x)\|_F \leq \|T_n - T_{n+p}\|_{\mathcal{L}(E,F)} \|x\|_E \leq \frac{\epsilon}{\|x\|_E} \|x\|_E = \epsilon$$

Donc la suite $(T_n(x))$ est de Cauchy dans l'espace complet E : elle admet une limite notée $T(x)$.

Comme les T_n sont linéaires, on voit facilement que T l'est aussi. Montrons que T est continue. Pour cela, on note que, puisque la suite (T_n) est de Cauchy, la suite de nombres réels $(\|T_n\|_{\mathcal{L}(E,F)})$ l'est aussi, puisque

$$\left| \|T_n\|_{\mathcal{L}(E,F)} - \|T_{n+p}\|_{\mathcal{L}(E,F)} \right| \leq \|T_n - T_{n+p}\|_{\mathcal{L}(E,F)} \quad \forall n, p \geq 0.$$

Donc, comme \mathbb{R} est complet, cette suite $(\|T_n\|_{\mathcal{L}(E,F)})$ converge et, en particulier, est bornée par une constante M . On a alors

$$\|T_n(x)\|_F \leq \|T_n\|_{\mathcal{L}(E,F)} \|x\|_E \leq M \|x\|_E \quad \forall x \in E, \forall n \in \mathbb{N}.$$

On fait tendre n vers $+\infty$, ce qui donne, puisque $T_n(x) \rightarrow T(x)$ et la norme $\|\cdot\|_E$ est continue,

$$\|T(x)\|_F \leq M \|x\|_E \quad \forall x \in E.$$

Cela montre que T est continue.

Montrons finalement que T_n tend vers T pour la norme $\|\cdot\|_{\mathcal{L}(E,F)}$: fixons $\epsilon > 0$ et soit $n_0 \in \mathbb{N}$ tel que

$$\forall n \geq n_0, \forall p \geq 0, \quad \|T_n - T_{n+p}\|_{\mathcal{L}(E,F)} \leq \epsilon$$

(un tel n_0 existe puisque (T_n) est de Cauchy). On a alors

$$\|T_n(x) - T_{n+p}(x)\|_F \leq \epsilon \quad \forall n \geq n_0, p \geq 0, x \in E \text{ avec } \|x\|_E \leq 1$$

On fait tendre p vers $+\infty$ dans l'inégalité ci-dessus : $T_{n+p}(x)$ tend vers $T(x)$, ce qui donne

$$\|T_n(x) - T(x)\|_F \leq \epsilon \quad \forall n \geq n_0, x \in E \text{ avec } \|x\|_E \leq 1$$

Donc

$$\|T_n - T\|_{\mathcal{L}(E,F)} = \sup \{ \|T_n(x) - T(x)\|_F \mid x \in E, \|x\|_E \leq 1 \} \leq \epsilon \quad \forall n \geq n_0.$$

En conclusion, la suite de Cauchy (T_n) tend vers T : cela prouve que $\mathcal{L}(E, F)$ est complet. \square

1.1.4 Produits d'EVN

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux EVN. On munit (le plus souvent) le produit $E \times F$ d'une des normes équivalentes suivantes :

$$\|(x, y)\|_1 = \|x\|_E + \|y\|_F, \quad \|(x, y)\|_\infty = \max\{\|x\|_E, \|y\|_F\},$$

$$\|(x, y)\|_2 = (\|x\|_E^2 + \|y\|_F^2)^{\frac{1}{2}} \quad \forall (x, y) \in E \times F.$$

(le fait que ces normes sont équivalentes vient juste du fait que, sur \mathbb{R}^2 , les normes $\|\cdot\|_1$, $\|\cdot\|_\infty$ et $\|\cdot\|_2$ sont équivalentes).

Proposition 1.14 *Si E et F sont des espaces de Banach, alors $E \times F$ (muni d'une des normes ci-dessus), l'est également.*

Preuve : exercice. \square

Définition 1.15 (Applications bilinéaires) *Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ et $(G, \|\cdot\|_G)$ trois EVN. On dit que l'application $T : E \times F \rightarrow G$ est bilinéaire si l'application $x \rightarrow T(x, y)$ est linéaire pour tout $y \in F$ et l'application $y \rightarrow T(x, y)$ est linéaire pour tout $x \in E$.*

Remarque : Une application bilinéaire n'est linéaire... que si elle est nulle (car $T(\lambda x, \lambda y) = \lambda^2 T(x, y)$).

Proposition 1.16 Soit $T : E \times F \rightarrow G$ est une application bilinéaire. Alors T est continue sur $E \times F$ si et seulement si il existe une constante C telle que

$$\|T(x, y)\|_G \leq C \|x\|_E \|y\|_F \quad \forall (x, y) \in E \times F .$$

Preuve : exercice. □

1.2 Espaces de Hilbert

1.2.1 Produit scalaire et norme associée

Commençons par le cas des espaces vectoriels réels.

Définition 1.17 (Produit scalaire réel) Soit H un \mathbb{R} -espace vectoriel. On appelle **produit scalaire** sur H toute application $B : H \times H \rightarrow \mathbb{R}$ bilinéaire, symétrique, définie positive, i.e.,

(i) B est une forme bilinéaire sur H ,

(ii) B est symétrique, i.e., $B(x, y) = B(y, x) \quad \forall (x, y) \in H \times H$,

(iii) B est définie positive : $B(x, x) \geq 0 \quad \forall x \in H$ et si $B(x, x) = 0$, alors $x = 0_H$.

Définition 1.18 (Produit scalaire complexe) Soit H un \mathbb{C} -espace vectoriel. On appelle **produit scalaire** sur H toute application $B : H \times H \rightarrow \mathbb{C}$ sesquilinéaire, hermitienne, définie positive, i.e.,

(i) B est une forme sesquilinéaire sur H , i.e.,

a) pour tout $y \in H$, l'application $x \rightarrow B(x, y)$ (de H dans \mathbb{C}) est linéaire,

b) pour tout $x \in H$, l'application $y \rightarrow B(x, y)$ (de H dans \mathbb{R}) est anti-linéaire ($B(x, \lambda y + z) = \bar{\lambda} B(x, y) + B(x, z)$ pour tout $x, y, z \in H$, $\lambda \in \mathbb{C}$),

(ii) B est hermitienne, i.e., $B(x, y) = \overline{B(y, x)} \quad \forall (x, y) \in H \times H$,

(iii) B est définie positive : $B(x, x) \geq 0 \quad \forall x \in H$ et si $B(x, x) = 0$, alors $x = 0_H$.

Remarque: si B est hermitienne, alors $B(x, x)$ est réel puisque $B(x, x) = \overline{B(x, x)}$.

Le plus souvent un produit scalaire est noté $\langle \cdot, \cdot \rangle$ au lieu de B .

Proposition 1.19 Soit H un espace vectoriel réel ou complexe muni d'un produit scalaire $\langle \cdot, \cdot \rangle$. Alors, si on pose

$$\|x\| = (\langle x, x \rangle)^{\frac{1}{2}} \quad \forall x \in H \tag{1}$$

on a :

(i)

$$\|x + y\|^2 = \|x\|^2 + 2\operatorname{Re}(\langle x, y \rangle) + \|y\|^2 \quad \forall x, y \in H .$$

(ii) (Cauchy-Schwarz)

$$|\langle x, y \rangle| \leq \|x\| \|y\| \quad \forall x, y \in H .$$

(iii) (Identité du parallélogramme)

$$2(\|x\|^2 + \|y\|^2) = \|x + y\|^2 + \|x - y\|^2 \quad \forall x, y \in H.$$

(iv) $\|\cdot\|$ définit une norme sur H .

Preuve (dans le cas complexe par exemple) :

(i) Si $x, y \in H$, alors

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \|x\|^2 + \langle x, y \rangle + \overline{\langle x, y \rangle} + \|y\|^2 \\ &= \|x\|^2 + 2\mathcal{R}e(\langle x, y \rangle) + \|y\|^2 \end{aligned}$$

(ii) (Cauchy-Schwarz) soient $x, y \in H$. Si $\langle x, y \rangle = 0$, le résultat est évident. Sinon, soit θ un argument du nombre complexe $\langle x, y \rangle$. On a alors, pour tout $\lambda \in \mathbb{R}$,

$$\|x + \lambda e^{i\theta} y\|^2 = \|x\|^2 + 2\lambda \mathcal{R}e(e^{-i\theta} \langle x, y \rangle) + |\lambda|^2 \|y\|^2 = \|x\|^2 + 2\lambda |\langle x, y \rangle| + |\lambda|^2 \|y\|^2.$$

Comme le polynôme à coefficients réels $\lambda \rightarrow \|x\|^2 + 2\lambda |\langle x, y \rangle| + \lambda^2 \|y\|^2$ ne prend pas de valeur négative, son discriminant est négatif ou nul : $4|\langle x, y \rangle|^2 - 4\|x\|^2 \|y\|^2 \leq 0$, ce qui donne l'inégalité annoncée.

(iii) (Identité du parallélogramme) si $x, y \in H$, on a

$$\|x + y\|^2 + \|x - y\|^2 = \|x\|^2 + 2\mathcal{R}e(\langle x, y \rangle) + \|y\|^2 + \|x\|^2 - 2\mathcal{R}e(\langle x, y \rangle) + \|y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

(iv) $\|\cdot\|$ définit une norme sur H . Nous ne montrons que l'inégalité triangulaire, le reste étant laissé en exercice :

$$\begin{aligned} \|x + y\|^2 &= \|x\|^2 + 2\mathcal{R}e(\langle x, y \rangle) + \|y\|^2 \leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 \quad (\text{par Cauchy-Schwarz}) \\ &\leq (\|x\| + \|y\|)^2 \end{aligned}$$

□

Exemples

- sur \mathbb{R}^N , le produit scalaire usuel est défini par :

$$\langle x, y \rangle = \sum_{i=1}^N x_i y_i \quad \forall x = (x_1, \dots, x_N), y = (y_1, \dots, y_N) \in \mathbb{R}^N.$$

- de même sur \mathbb{C}^N , le produit scalaire usuel est défini par :

$$\langle x, y \rangle = \sum_{i=1}^N x_i \bar{y}_i \quad \forall x = (x_1, \dots, x_N), y = (y_1, \dots, y_N) \in \mathbb{C}^N.$$

- L'espace $L^2_\mu(X, \mathbb{R})$ est muni du produit scalaire

$$\langle f, g \rangle = \int_X f g d\mu \quad \forall f, g \in L^2,$$

- En particulier, le prototype des espaces de Hilbert est l'espace $\ell^2(\mathbb{R})$ (ou $\ell^2(\mathbb{C})$) des suites réelles (resp. complexes) $(x_i)_{i \in \mathbb{N}}$ de carré sommable : $\sum_{i=1}^{\infty} |x_i|^2 < +\infty$. Cet espace est muni du produit scalaire

$$\langle x, y \rangle = \sum_{i=1}^{\infty} x_i y_i \quad \forall x = (x_i)_{i \in \mathbb{N}}, y = (y_i)_{i \in \mathbb{N}} \in \ell^2.$$

(Dans le cas complexe, le produit scalaire est $\langle x, y \rangle = \sum_{i=1}^{\infty} x_i \overline{y_i}$).

Définition 1.20 Soit H un espace vectoriel muni d'un produit scalaire $\langle \cdot, \cdot \rangle$. On dit que H est un espace de Hilbert si H , muni de la norme associée au produit scalaire $\langle \cdot, \cdot \rangle$ (par (1)) est un espace complet.

Remarque: Tout espace de Hilbert est un espace de Banach. La réciproque est fautive : en général une norme quelconque ne provient pas d'un produit scalaire (exercice : montrer que, si $H = \mathbb{R}^N$, il n'existe pas de produit scalaire B tel que $B(x, x) = \|x\|_{\infty}^2$).

1.2.2 Le théorème de projection

Rappelons d'abord quelques notions élémentaires sur les ensembles convexes.

Définition 1.21 (Ensemble convexe) Soit E un espace vectoriel. On dit qu'un sous-ensemble C de E est convexe si

$$\forall x, y \in C, \forall \lambda \in [0, 1], \lambda x + (1 - \lambda)y \in C.$$

Par exemple, un sous-espace vectoriel de E est toujours convexe.

Définition 1.22 (Fonction convexe) Soit E un espace vectoriel et C un sous-ensemble convexe de E . Une application $f : C \rightarrow \mathbb{R}$ est convexe si

$$\forall x, y \in C, \forall \lambda \in [0, 1], f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Voici quelques propriétés classiques des ensembles convexes (leur preuve est un bon exercice) :

1. Une intersection quelconque de convexes est convexe.
2. Si $f : E \rightarrow \mathbb{R}$ est convexe, alors, pour tout $c \in \mathbb{R}$ l'ensemble $\{x \in E \mid f(x) \leq c\}$ est convexe.

Le théorème suivant est le résultat une des propriétés les plus importantes des espaces de Hilbert.

Théorème 1.23 (de projection (cas réel)) Si H est un espace vectoriel muni d'un produit scalaire $\langle \cdot, \cdot \rangle$ et F est un sous-ensemble convexe complet non vide de H , pour tout $x \in H$ il existe un unique point $\Pi_F(x) \in F$ (appelé le projeté de x sur F) tel que

$$\|x - \Pi_F(x)\| = \min_{y \in F} \|y - x\|.$$

De plus, $\Pi_F(x)$ est caractérisé par l'inégalité suivante :

$$\langle x - \Pi_F(x), y - \Pi_F(x) \rangle \leq 0 \quad \forall y \in F.$$

Remarques 1.24 1. Le théorème s'applique en particulier, lorsque H est un espace de Hilbert et F est un sous-ensemble convexe fermé non vide de H : en effet, un sous-ensemble fermé d'un espace complet est complet.

2. Dans la cas où H est complexe et $\langle \cdot, \cdot \rangle$ un produit scalaire complexe, le théorème s'applique sous les mêmes hypothèses et la caractérisation prend la forme suivante :

$$\operatorname{Re}(\langle x - \Pi_F(x), y - \Pi_F(x) \rangle) \leq 0 \quad \forall y \in F .$$

Il faut connaître la preuve de la partie “unicité” du théorème.

Preuve : Commençons par l'unicité (*): soient z_1 et z_2 deux projections du points x sur F . Notons que, par convexité de F , le point $(z_1 + z_2)/2$ appartient aussi à F . Par définition du projeté, on a donc

$$\left\| \frac{z_1 + z_2}{2} - x \right\| \geq \|z_1 - x\| = \|z_2 - x\| .$$

Appliquons maintenant l'identité du parallélogramme aux vecteurs $z_1 - x$ et $z_2 - x$, on a

$$\|z_1 + z_2 - 2x\|^2 + \|z_1 - z_2\|^2 = 2(\|z_1 - x\|^2 + \|z_2 - x\|^2) \leq 4\left\| \frac{z_1 + z_2}{2} - x \right\|^2$$

Donc $\|z_1 - z_2\|^2 \leq 0$, ce qui prouve que $z_1 = z_2$.

Prouvons maintenant l'existence. Comme F est non vide, il existe un point $\bar{z} \in F$ et donc $\bar{m} := \inf_{z \in F} \|z - x\|$ est bien défini. Soit (z_n) une suite minimisante du problème : $z_n \in F$ et $\lim_{n \rightarrow +\infty} \|z_n - x\| = \bar{m}$. L'objectif est de montrer que (z_n) est une suite de Cauchy de F . Soit $\epsilon > 0$ et $N \geq 0$ tel que $\|z_n - x\| \leq \bar{m} + \epsilon$ pour tout $n \geq N$. Fixons $n, m \geq N$. On applique l'égalité du parallélogramme à $z_n - x$ et $z_m - x$.

$$\|z_n + z_m - 2x\|^2 + \|z_n - z_m\|^2 = 2(\|z_n - x\|^2 + \|z_m - x\|^2) \leq 4(\bar{m} + \epsilon)^2$$

Or, comme le point $\frac{z_n + z_m}{2}$ appartient à F par convexité de F , on a $\left\| \frac{z_n + z_m}{2} - x \right\| \geq \bar{m}$. L'inégalité ci-dessus se réécrit donc

$$4\bar{m}^2 + \|z_n - z_m\|^2 \leq 4(\bar{m} + \epsilon)^2 ,$$

ce qui donne

$$\|z_n - z_m\|^2 \leq 4(\bar{m} + \epsilon)^2 - 4\bar{m}^2 = 4\epsilon(2\bar{m} + \epsilon) .$$

Comme le membre de droite est arbitrairement petit lorsque $\epsilon \rightarrow 0$, la suite (z_n) est bien de Cauchy. Or F est complet, donc cette suite converge vers une limite $\bar{z} \in F$. La norme étant continue, on a finalement,

$$\bar{m} = \lim_{n \rightarrow +\infty} \|x - z_n\| = \|x - \bar{z}\|$$

et \bar{z} est la projection de x sur F .

Montrons maintenant la caractérisation. Vérifions d'abord que $\Pi_F(x)$ satisfait l'inégalité : fixons $z \in F$ et, pour $t \in]0, 1]$, considérons le point $(1 - t)\Pi_F(x) + tz$, qui appartient à F puisque F est convexe. On a donc, par définition du projeté,

$$\|\Pi_F(x) - x\|^2 \leq \|(1 - t)\Pi_F(x) + tz - x\|^2 = \|\Pi_F(x) - x\|^2 + 2t\langle \Pi_F(x) - x, z - \Pi_F(x) \rangle + t^2\|z - \Pi_F(x)\|^2 .$$

On simplifie par $\|\Pi_F(x) - x\|^2$, on divise par $t > 0$ et on fait tendre t vers 0 pour obtenir $0 \leq 2\langle \Pi_F(x) - x, z - \Pi_F(x) \rangle$. Ceci est l'inégalité voulue.

Inversement, soit $\bar{z} \in F$ un point vérifiant l'inégalité de caractérisation. Montrons que \bar{z} est le projeté de x sur F . On a, pour tout $z \in F$,

$$\|z - x\|^2 = \|z - \bar{z} + \bar{z} - x\|^2 = \|z - \bar{z}\|^2 + 2\langle z - \bar{z}, \bar{z} - x \rangle + \|\bar{z} - x\|^2$$

Comme $\|z - \bar{z}\|^2 \geq 0$ et $\langle z - \bar{z}, \bar{z} - x \rangle \geq 0$ par inégalité de caractérisation, on obtient $\|z - x\|^2 \geq \|\bar{z} - x\|^2$, ce qui prouve que \bar{z} est le projeté de x sur F . □

Lorsque l'ensemble F est un sous-espace vectoriel fermé de l'espace de Hilbert H , le théorème de projection prend la forme suivante :

Théorème 1.25 (Projection sur un SEV) *On suppose que H est un espace de Hilbert et que F est un sous-espace vectoriel fermé de H . Alors*

(i) *l'application $x \rightarrow \Pi_F(x)$ est linéaire continue, de norme 1 si F n'est pas réduit à $\{0_H\}$.*

(ii) *c'est une projection : $\Pi_F \circ \Pi_F = \Pi_F$*

(iii) *on a l'égalité*

$$\langle x - \Pi_F(x), y \rangle = 0 \quad \forall y \in F .$$

(iv) *cette égalité caractérise $\Pi_F(x)$.*

$\Pi_F(x)$ s'appelle la projection orthogonale de x sur F .

Preuve (*): Montrons d'abord la caractérisation : vu le théorème de projection (cas convexe), il suffit de vérifier que, lorsque F est un espace vectoriel, la condition

$$(*) \quad \langle x - \bar{z}, y - \bar{z} \rangle \leq 0 \quad \forall y \in F .$$

est équivalente à la condition

$$(**) \quad \langle x - \bar{z}, y \rangle = 0 \quad \forall y \in F .$$

Supposons d'abord (*). Soit $y \in F$. Alors, comme F est un espace vectoriel et $\bar{z} \in F$, $y + \bar{z}$ appartient aussi à F . Donc, par (*), $\langle x - \bar{z}, y + \bar{z} - \bar{z} \rangle = \langle x - \bar{z}, y \rangle \leq 0$. Ceci est vrai aussi pour $-y$, qui appartient aussi à F : $\langle x - \bar{z}, -y \rangle \leq 0$. D'où (**).

Inversement, si \bar{z} satisfait (**), alors pour tout $y \in F$, on a $y - \bar{z} \in F$ par linéarité de F , et donc, par (**), $\langle x - \bar{z}, y - \bar{z} \rangle = 0$. D'où (*).

Une fois connue la caractérisation du projeté, la linéarité de l'application $x \rightarrow \Pi_F(x)$ est évidente : en effet, soient $x_1, x_2 \in H$, $\lambda \in \mathbb{R}$. On a alors, pour tout $y \in H$,

$$\langle x_1 + \lambda x_2 - (\Pi_F(x_1) + \lambda \Pi_F(x_2)), y \rangle = \langle x_1 - \Pi_F(x_1), y \rangle + \lambda \langle x_2 - \Pi_F(x_2), y \rangle = 0 ,$$

Donc $\Pi_F(x_1 + \lambda x_2) = \Pi_F(x_1) + \lambda \Pi_F(x_2)$ d'après la caractérisation, ce qui prouve la linéarité de Π_F .

Montrons maintenant que $\|\Pi_F\| = 1$. Soit $z \in F$, avec $z \neq 0$. On a $\Pi_F(z) = z$, et donc $\|\Pi_F\| \geq \|\Pi_F(z)\|/\|z\| = 1$. Inversement, si $x \in H$, alors d'après la caractérisation, $\langle x - \Pi_F(x), y \rangle = 0$ pour tout $y \in F$. Donc, en prenant $y = \Pi_F(x)$,

$$\|\Pi_F(x)\|^2 = \langle \Pi_F(x), \Pi_F(x) \rangle = \langle x, \Pi_F(x) \rangle \leq \|x\| \|\Pi_F(x)\|$$

ce qui prouve que $\|\Pi_F(x)\| \leq \|x\|$. Donc $\|\Pi_F\| = 1$.

Finalement, $\Pi_F \circ \Pi_F = \Pi_F$ par définition de la projection. □

1.2.3 Orthogonalité

Soit $(H, \langle \cdot, \cdot \rangle)$ un espace de Hilbert. On dit que deux vecteurs $x, y \in H$ sont orthogonaux si $\langle x, y \rangle = 0$. On note alors $x \perp y$.

Définition 1.26 (Orthogonal d'une partie) *Soit A un sous-ensemble non vide de H . L'orthogonal de A est le sous-ensemble A^\perp de H défini par*

$$A^\perp = \{x \in H \mid \langle x, y \rangle = 0 \quad \forall y \in A\}$$

Remarque 1.27 Si F est un sous-espace vectoriel fermé de H , alors, d'après la caractérisation de la projection, on a $x - \Pi_F(x) \in F^\perp$ pour tout $x \in F$.

Proposition 1.28 Soit A un sous-ensemble non vide de H .

- (i) A^\perp est toujours un sous-espace vectoriel fermé de H .
- (ii) $H^\perp = \{0_H\}$ et $\{0_H\}^\perp = H$.
- (iii) Si $A \subset B$, alors $B^\perp \subset A^\perp$.
- (iv) $(\overline{\text{Vect}(A)})^\perp = A^\perp$ (où $\overline{\text{Vect}(A)}$ est la fermeture de $\text{Vect}(A)$).

Preuve (*): Pour fixer les idées, on travaille dans le cas réel. (i) Montrons que A^\perp est un sous-espace vectoriel : si $z_1, z_2 \in A^\perp$ et $\lambda \in \mathbb{R}$, alors, pour tout $y \in A$, $\langle z_1 + \lambda z_2, y \rangle = \langle z_1, y \rangle + \lambda \langle z_2, y \rangle = 0$. Donc $z_1 + \lambda z_2 \in A^\perp$, ce qui prouve que A^\perp est un sous-espace vectoriel. Montrons maintenant que A^\perp est fermé : si (y_n) est une suite d'éléments de A^\perp qui tend vers $z \in F$, on a, pour tout $y \in A$, $\langle z, y \rangle = \lim_{n \rightarrow +\infty} \langle y_n, y \rangle = 0$. Cela montre que $z \in A^\perp$, qui est donc fermé.

(ii) Si $z \in H^\perp$, alors, comme $z \in H$, $\|z\|^2 = \langle z, z \rangle = 0$, ce qui prouve que $z = 0$. Donc $H^\perp = \{0_H\}$. Il est clair que $\{0_H\}^\perp = H$.

(iii) est une application directe de la définition.

(iv) Comme $A \subset \overline{\text{Vect}(A)}$, l'inclusion $(\overline{\text{Vect}(A)})^\perp \subset A^\perp$ est immédiate. Soit maintenant $z \in A^\perp$. Alors, si y est une combinaison linéaire d'éléments de A , c'est-à-dire si $y = \sum_{i=1}^n \lambda_i y_i$ avec $y_i \in A$ et $\lambda_i \in \mathbb{R}$, on a

$$\langle z, \sum_{i=1}^n \lambda_i y_i \rangle = \sum_{i=1}^n \lambda_i \langle z, y_i \rangle = 0 \quad \text{puisque } \langle z, y_i \rangle = 0 \text{ car } z \in A^\perp.$$

Donc $z \in (\overline{\text{Vect}(A)})^\perp$. D'autre part, si $y \in \overline{\text{Vect}(A)}$, il existe une suite (y_n) d'éléments de $\text{Vect}(A)$ telle que $y_n \rightarrow y$. Or $\langle z, y_n \rangle = 0$ et le produit scalaire est continu. Donc $\langle z, y \rangle = 0$, ce qui prouve que $z \in (\overline{\text{Vect}(A)})^\perp$. \square

Proposition 1.29 Soit F un sous-ensemble de H .

- (i) F est un sous-espace vectoriel fermé de H si et seulement si $(F^\perp)^\perp = F$.
- (ii) Dans ce cas, $H = F \oplus F^\perp$ et $\text{id}_H = \Pi_F + \Pi_{F^\perp}$.

Preuve (*): (i) Supposons que F est un sous-espace vectoriel fermé de H . Soit $z \in F$ et $y \in F^\perp$. Alors $\langle z, y \rangle = 0$, et donc $F \subset (F^\perp)^\perp$. Inversement, soit $z \in (F^\perp)^\perp$. Soit $\Pi_F(z)$ le projeté de z sur F . Par caractérisation, on a $\langle z - \Pi_F(z), y \rangle = 0$ pour tout $y \in F$, et donc $z - \Pi_F(z) \in F^\perp$. Mais $z \in (F^\perp)^\perp$, et donc $0 = \langle z - \Pi_F(z), z \rangle$. Rappelons que l'application linéaire continue Π_F est de norme 1, ce qui prouve que $\|\Pi_F(z)\| \leq \|z\|$. D'où

$$\|z\|^2 = \langle z, z \rangle = \langle \Pi_F(z), z \rangle \leq \|\Pi_F(z)\| \|z\| \leq \|z\|^2.$$

Il y a donc une égalité dans Cauchy-Schwarz, ce qui prouve que $\Pi_F(z) = \lambda z$ avec $\lambda \geq 0$. De l'égalité $0 = \langle z - \Pi_F(z), z \rangle$ on tire facilement que $\lambda = 1$, ce qui prouve que $z = \Pi_F(z)$, c'est-à-dire que $z \in F$.

Supposons maintenant que $(F^\perp)^\perp = F$. Alors, comme $(F^\perp)^\perp = F$ est un sous-espace vectoriel fermé de H , F est fermé.

(ii) On a, pour tout $x \in H$, $x = \Pi_F(x) + x - \Pi_F(x)$ avec $\Pi_F(x) \in F$ et $x - \Pi_F(x) \in F^\perp$. Donc $H = F + F^\perp$. Or $F \cap F^\perp = \{0_H\}$ car si $z \in F \cap F^\perp$, alors $\|z\|^2 = \langle z, z \rangle = 0$. D'où $H = F \oplus F^\perp$. L'égalité $x = \Pi_F(x) + x - \Pi_F(x)$ pour tout $x \in H$ montre aussi que $\text{id}_H = \Pi_F + \Pi_{F^\perp}$. \square

1.2.4 Bases hilbertiennes

Soit $(H, \langle \cdot, \cdot \rangle)$ un espace de Hilbert. On dit qu'une famille A de H est *totale* si l'adhérence de l'espace vectoriel engendré par A est égal à H tout entier : $\overline{\text{Vect}(A)} = H$.

Proposition 1.30 *A est totale si et seulement si $A^\perp = \{0_H\}$.*

Noter que c'est l'implication " $A^\perp = \{0_H\} \Rightarrow A$ est totale" qui est utile en pratique.

Preuve (*): Si A est totale, alors

$$A^\perp = (\overline{\text{Vect}(A)})^\perp = H^\perp = \{0_H\}$$

Supposons maintenant que $A^\perp = \{0_H\}$. Alors

$$\overline{\text{Vect}(A)} = \left((\overline{\text{Vect}(A)})^\perp \right)^\perp = \left(A^\perp \right)^\perp = (\{0_H\})^\perp = H.$$

□

Définition 1.31 (Base hilbertienne) Soit $(e_n)_{n \in \mathbb{N}}$ une suite d'éléments de H . On dit que cette famille est une *base hilbertienne* de H si elle est totale et orthonormée :

$$(\text{Vect}(\{e_n, n \in \mathbb{N}\}))^\perp = \{0_H\} \quad \text{et} \quad \langle e_n, e_m \rangle = \begin{cases} 0 & \text{si } n \neq m \\ 1 & \text{si } n = m \end{cases} \quad \forall n, m \in \mathbb{N}.$$

Définition 1.32 (Espace séparable) On dit qu'un EVN $(E, \|\cdot\|)$ est *séparable* si E contient une famille dénombrable dense.

Exemples et contre-exemple : Une grande partie des espaces utilisés en analyse fonctionnelle sont séparables.

1. si $a < b$, $(\mathcal{C}^0([a, b]), \|\cdot\|_\infty)$ est séparable : en effet $\mathbb{Q}[X]$ est une famille dénombrable dense (c'est une conséquence directe du théorème de Stone-Weierstrass, qui dit que toute fonction continue sur $[a, b]$ est limite uniforme d'une suite de polynômes).
2. les espaces ℓ^p (pour $1 \leq p < +\infty$) sont séparables. En effet si on note \mathcal{D}_0 l'ensemble des suites nulles à partir d'un certain rang et à coefficients dans \mathbb{Q} , alors \mathcal{D}_0 est dénombrable et dense dans ℓ^p (voir les exercices).
3. Si I est un intervalle non vide de \mathbb{R} et $1 \leq p < +\infty$, alors l'espace $L^p(I)$ est séparable. (admis)
4. Par contre ni ℓ^∞ , ni $L^\infty([a, b])$ ne sont séparables. Montrons-le pour ℓ^∞ : soit $\Phi : \mathcal{P}(\mathbb{N}) \rightarrow \ell^\infty$ l'application qui à une partie E de \mathbb{N} associe la suite

$$\Phi(E) = (\Phi(E)_i)_{i \in \mathbb{N}} \quad \text{où } \Phi(E)_i = \begin{cases} 1 & \text{si } i \in E \\ 0 & \text{sinon} \end{cases}$$

Alors, par définition,

$$\|\Phi(E) - \Phi(E')\|_\infty = 1 \quad \forall E, E' \in \mathcal{P}(\mathbb{N}) \text{ avec } E \neq E'. \quad (2)$$

Montrons maintenant que cela implique que ℓ^∞ n'est pas séparable. En effet, raisonnons par l'absurde en supposant qu'il existe une suite dénombrable dense (x^n) dans ℓ^∞ . A tout $E \in \mathcal{P}(E)$ il existerait un plus petit entier n_E tel que $\|x^{n_E} - \Phi(E)\|_\infty \leq 1/4$ puisque (x^n)

est dense dans ℓ^∞ . Notons que l'application $E \rightarrow n_E$ est injective, puisque si $n_E = n'_E$, alors $x^{n_E} = x^{n'_E}$ et

$$\|\Phi(E) - \Phi(E')\|_\infty \leq \|\Phi(E) - x^{n_E}\|_\infty + \|x^{n'_E} - \Phi(E')\|_\infty \leq \frac{1}{2},$$

ce qui implique que $E = E'$ d'après (2). Mais comme $\mathcal{P}(\mathbb{N})$ est non dénombrable tandis que \mathbb{N} est dénombrable, il ne peut exister d'injection de $\mathcal{P}(\mathbb{N})$ dans \mathbb{N} . Une contradiction.

Théorème 1.33 *Tout espace de Hilbert séparable possède au moins une base hilbertienne.*

Il faut connaître la preuve de ce résultat : elle fournit une procédure constructive d'une base hilbertienne (procédé d'orthonormalisation de Schmidt).

Preuve (*): Soit $(x_n)_{n \in \mathbb{N}}$ une suite dense dans H . On note F_n l'espace vectoriel engendré par $\{x_1, \dots, x_n\}$. Notons que $\bigcup_n F_n$ est dense dans H , puisque cet espace vectoriel contient tous les x_n . Nous allons construire explicitement une suite $(e_n)_{n \in \mathbb{N}}$ et une suite strictement croissante d'indices $(k_n)_{n \in \mathbb{N}}$ telles que $\{e_1, \dots, e_n\}$ est une base orthonormée de F_{k_n} .

Sans perte de généralité on peut supposer que $x_1 \neq 0$. On pose alors $e_1 = x_1/\|x_1\|$ et $\{e_1\}$ est une base orthonormée de F_1 . Supposons construits $\{e_1, \dots, e_n\}$ et k_n . Soit k_{n+1} le plus petit indice $k > k_n$ tel que $F_k \neq F_{k_n}$. Alors la famille $\{e_1, \dots, e_n, x_{k_{n+1}}\}$ est une base de $F_{k_{n+1}}$. Posons

$$e'_{n+1} = x_{k_{n+1}} - \sum_{i=1}^n \langle x_{k_{n+1}}, e_i \rangle e_i.$$

Notons que $\langle e'_{n+1}, e_j \rangle = 0$ pour tout $j \leq n$ et que $e'_{n+1} \neq 0$ puisque $x_{k_{n+1}}$ n'est pas combinaison linéaire des e_1, \dots, e_n . On pose alors $e_{n+1} = e'_{n+1}/\|e'_{n+1}\|$. On montre facilement que $\{e_1, \dots, e_{n+1}\}$ est une base orthonormée de $F_{k_{n+1}}$. On conclut à l'existence de la famille $\{e_n\}_{n \in \mathbb{N}}$ par récurrence.

Par construction, $\{e_n\}_{n \in \mathbb{N}}$ est une famille orthonormée. Elle est également totale puisque $\text{Vect}(e_n, n \in \mathbb{N}) = \bigcup_n F_n$ et que ce dernier espace est dense dans H . \square

Théorème 1.34 *Soit H un espace de Hilbert séparable, muni d'une base hilbertienne $(e_n)_{n \in \mathbb{N}}$. Alors l'application*

$$\begin{aligned} \Phi : H &\rightarrow \ell^2 \\ x &\mapsto (\langle x, e_n \rangle)_{n \in \mathbb{N}} \end{aligned}$$

est une application linéaire bijective et isométrique. En particulier, on a les égalités

- (Bessel) $x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$
- (Parseval) $\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2$

Preuve (*): L'application Φ est clairement linéaire. Montrons qu'elle est continue : en effet, notons F_n l'espace vectoriel engendré par $\{e_1, \dots, e_n\}$. Alors, si $x \in H$, alors on montre aisément que $\Pi_{F_n}(x) = \sum_{i=1}^n \langle x, e_i \rangle e_i$ et on sait que $x = \Pi_{F_n}(x) + \Pi_{F_n^\perp}(x)$, où les deux vecteurs sont orthogonaux. Donc, par Pythagore,

$$\|x\|^2 = \|\Pi_{F_n}(x)\|^2 + \|\Pi_{F_n^\perp}(x)\|^2 \geq \|\Pi_{F_n}(x)\|^2 = \sum_{i=1}^n \langle x, e_i \rangle^2.$$

Lorsque $n \rightarrow +\infty$, on a alors

$$\|\Phi(x)\|^2 = \sum_i \langle x, e_i \rangle^2 \leq \|x\|^2. \quad (3)$$

Donc Φ est continue avec $\|\Phi\| \leq 1$.

Montrons maintenant qu'en fait l'égalité a lieu dans (3). En effet comme la famille $\{e_n\}$ est totale, il existe une suite (y_k) de $Vect(x_n, n \in \mathbb{N})$ qui converge vers x . Par définition de F_n , pour tout k , il existe n_k telle que $y_k \in F_{n_k}$. Comme la suite des espaces vectoriels (F_n) est croissante, on peut supposer sans perte de généralité que la suite n_k est strictement croissante, et donc tend vers $+\infty$. Alors, par définition de la projection, on a

$$\|x\|^2 - \sum_{i=1}^{n_k} \langle x, e_i \rangle^2 = \|x\|^2 - \|\Pi_{F_{n_k}}(x)\|^2 = \|x - x_{n_k}\|^2 \leq \|x - y_{n_k}\|^2 \rightarrow 0,$$

ce qui prouve l'égalité de Bessel. Cette égalité affirme en particulier que Φ est une isométrie. En passant, nous avons également montré l'égalité de Parseval :

$$\lim_{n \rightarrow +\infty} \sum_{i=1}^n \langle x, e_i \rangle e_i = \lim_{n \rightarrow +\infty} x_n = x.$$

Montrons maintenant que Φ est bijective. D'abord Φ est clairement injective, car si $\Phi(x) = 0$, alors

$$\|x\| = \|\Phi(x)\| = \|0\| = 0.$$

Donc $x = 0$. Montrons finalement que Φ est surjective. Soit $y = (y_n)_{n \in \mathbb{N}}$ un élément de ℓ^2 . Posons $x_n = \sum_{i=1}^n y_i e_i$. Alors la suite (x_n) est de Cauchy dans H puisque

$$\|x_{n+p} - x_n\|^2 = \sum_{i=n+1}^{n+p} |y_i|^2$$

et que la série $\sum_i |y_i|^2$ converge. Donc, comme H est complet, (x_n) converge vers un certain $x \in H$. Notons que, pour tout $i \in \mathbb{N}$,

$$\langle x, e_i \rangle = \lim_{n \rightarrow +\infty} \langle x_n, e_i \rangle = \lim_{n \rightarrow +\infty} y_i = y_i.$$

Donc $\Phi(x) = y$, ce qui prouve que Φ est surjective. □

Exemples fondamentaux :

1. Si $(\ell^2, \langle \cdot, \cdot \rangle_2)$ est l'espace de Hilbert des suites de carré sommable, la famille $(e^n)_{n \in \mathbb{N}}$ définie par $e_i^n = \delta_{in}$ est une base hilbertienne de ℓ^2 .

2. La théorie des séries de Fourier nous apprend que la famille $\left\{ \frac{1}{\sqrt{2\pi}} e^{inx} \right\}_{n \in \mathbb{Z}}$ est une base hilbertienne de $L^2((0, 2\pi), \mathbb{C})$.

En particulier, toute fonction de $L^2((0, 2\pi), \mathbb{C})$ est limite, dans L^2 , d'une suite de polynômes trigonométriques.

3. Lorsque l'on revient aux espaces réels, on a : La famille $\left\{ \frac{1}{\sqrt{2\pi}} \cos(nx) \right\}_{n \in \mathbb{N}} \cup \left\{ \frac{1}{\sqrt{2\pi}} \sin(nx) \right\}_{n \in \mathbb{N}^*}$ est une base hilbertienne de $L^2((0, 2\pi), \mathbb{R})$.

1.2.5 Dualité

Rappelons que le dual (topologique) E^* d'un EVN E est l'ensemble des applications linéaires continues de E dans \mathbb{R} . Pour les espaces de Hilbert, on peut identifier E^* et E , au sens suivant :

Théorème 1.35 (de représentation de Riesz) Soit $(H, \langle \cdot, \cdot \rangle)$ un espace de Hilbert (réel ou complexe). Pour tout $f \in H^*$, il existe un unique élément $\bar{x} \in H$ tel que

$$f(y) = \langle y, \bar{x} \rangle \quad \forall y \in H. \quad (4)$$

De plus l'application $f \mapsto \bar{x}$ de H^* dans H est une application (anti)linéaire continue bijective (et d'inverse continu).

Preuve : Soit $f \in H^*$. Si f est l'application nulle, alors $\bar{x} = 0_H$ est bien l'unique vecteur de H vérifiant (4).

Supposons maintenant que $f \neq 0$. Avant de commencer la démonstration, notons que, si \bar{x} existe, alors $\bar{x} \in (Ker(f))^\perp$ puisque, pour tout $y \in Ker(f)$, $\langle \bar{x}, y \rangle = f(y) = 0$. De plus, $f(\bar{x}) = \langle \bar{x}, \bar{x} \rangle = \|\bar{x}\|^2$. L'objet des calculs suivants est précisément de construire un tel \bar{x} .

Comme $f \neq 0$, il existe $x \in H$ tel que $f(x) \neq 0$. Soit z la projection orthogonale de x sur $Ker(f)$ (qui est un sous-espace vectoriel fermé de H). Comme $f(x) \neq 0$, on a $z \neq x$. Posons $\bar{x} = (x - z)f(x)/\|x - z\|^2$. Notons d'abord que

$$f(\bar{x}) = \frac{f(x)}{\|x - z\|^2} (f(x) - f(z)) = \frac{(f(x))^2}{\|x - z\|^2} = \|\bar{x}\|^2 \neq 0.$$

Par définition de la projection orthogonale, on a également

$$\langle y, \bar{x} \rangle = \frac{f(x)}{\|x - z\|^2} \langle y, x - z \rangle = 0 \quad \forall y \in Ker(f).$$

Soit maintenant $y \in H$ quelconque. Alors $y - f(y)\bar{x}/f(\bar{x})$ appartient à $Ker(f)$. Donc

$$\langle y - f(y)\bar{x}/f(\bar{x}), \bar{x} \rangle = 0,$$

ce qui prouve que

$$\langle y, \bar{x} \rangle = f(y) \langle \bar{x}, \bar{x} / f(\bar{x}) \rangle = f(y) \frac{\|\bar{x}\|^2}{f(\bar{x})} = f(y).$$

Il existe donc $\bar{x} \in H$ tel que $\langle y, \bar{x} \rangle = f(y)$ pour tout $y \in H$.

Montrons l'unicité de \bar{x} : si $\bar{x}_1 \in H$ est tel que $\langle y, \bar{x}_1 \rangle = f(y)$ pour tout $y \in H$, alors

$$\langle y, \bar{x} - \bar{x}_1 \rangle = 0 \quad \forall y \in H,$$

ce qui implique que $\bar{x} - \bar{x}_1 = 0_H$ et donc que $\bar{x} = \bar{x}_1$.

Finalement, l'application $\Phi : H \rightarrow H^*$ définie par $\Phi(x)(y) = \langle y, x \rangle$ est (anti)linéaire, continue car

$$\begin{aligned} \|\Phi(x)\| &= \sup\{|\Phi(x)(y)| \mid y \in H, \|y\| \leq 1\} \\ &= \sup\{|\langle x, y \rangle| \mid y \in H, \|y\| \leq 1\} \\ &\leq \|x\| \end{aligned} \quad (5)$$

(par Cauchy-Schwarz). Nous avons montré ci-dessus que Φ est bijective. Notons aussi que Φ^{-1} est continue car, si $f \in H^*$ et \bar{x} vérifie (4), on a

$$\|\bar{x}\|_H^2 = \langle \bar{x}, \bar{x} \rangle = f(\bar{x}) \leq \|f\|_{H^*} \|\bar{x}\|_H$$

D'où

$$\|\bar{x}\|_H \leq \|f\|_{H^*} = \|\Phi(\bar{x})\|_{H^*} \quad \forall \bar{x} \in H. \quad (6)$$

En terme de Φ , les inégalités (5) et (6) donnent

$$\|\bar{x}\|_H = \|\Phi(\bar{x})\|_{H^*} \quad \forall \bar{x} \in H,$$

et donc Φ est une isométrie de H dans H^* . □

1.2.6 Le théorème de Lax-Milgram

Un des outils de base pour démontrer l'existence de solution pour certaines équations aux dérivées partielles est le théorème de Lax-Milgram. Pour l'énoncer, nous aurons besoin de la notion suivante :

Définition 1.36 Soit H un espace de Hilbert réel et $a : H \times H \rightarrow \mathbb{R}$ une forme bilinéaire continue. On dit que a est coercive, s'il existe une constante $\alpha > 0$ telle que

$$a(x, x) \geq \alpha \|x\|^2 \quad \forall x \in H .$$

Bien noter que la constant α doit être strictement positive, et ne dépend que de a (et pas de x).

Théorème 1.37 (de Lax-Milgram) Soit H un espace de Hilbert réel et $a : H \times H \rightarrow \mathbb{R}$ une forme bilinéaire continue coercive. Alors, pour tout $f \in H^*$ il existe un unique élément $\bar{x} \in H$ tel que

$$a(\bar{x}, y) = f(y) \quad \forall y \in H .$$

De plus, si a est symétrique, \bar{x} est l'unique minimum de la fonctionnelle

$$\Phi(x) = \frac{1}{2}a(x, x) - f(x)$$

Remarque (*): (Preuve connaître) Lorsque a est symétrique, le théorème possède une démonstration très rapide : en effet, a définit un produit scalaire dont la norme associée, notée $\|\cdot\|_a$, est équivalente à la norme initiale de H . En particulier, H , muni du produit scalaire a , est un espace de Hilbert et f reste une forme linéaire continue sur H . Le théorème de représentation de Riesz affirme alors qu'il existe un unique point $\bar{x} \in H$ tel que

$$a(\bar{x}, y) = f(y) \quad \forall y \in H .$$

Preuve du théorème de Lax-Milgram : La preuve de l'unicité est à connaître (*): Supposons qu'il existe \bar{x}_1 et \bar{x}_2 dans H tels que

$$a(\bar{x}_1, y) = f(y) = a(\bar{x}_2, y) \quad \forall y \in H .$$

Alors on a $a(\bar{x}_1 - \bar{x}_2, y) = 0$ pour tout y , et donc, en prenant $y = \bar{x}_1 - \bar{x}_2$ et en utilisant la coercivité de a ,

$$0 = a(\bar{x}_1 - \bar{x}_2, \bar{x}_1 - \bar{x}_2) \geq \alpha \|\bar{x}_1 - \bar{x}_2\|^2 .$$

Comme $\alpha > 0$, ceci prouve que $\bar{x}_1 = \bar{x}_2$.

Montrons maintenant l'existence : comme a est continue, pour tout $x \in H$ fixé, l'application $y \rightarrow a(x, y)$ est une forme linéaire continue. D'après le théorème de représentation de Riesz, il existe donc pour tout $x \in H$ un unique élément $A(x) \in H$ tel que

$$\langle A(x), y \rangle = a(x, y) \quad \forall y \in H .$$

On montre sans difficulté que $A : H \rightarrow H$ est linéaire et continue. Montrons que A est une bijection.

A est injective : Comme A est linéaire, il suffit de montrer que son noyau est réduit à $\{0_H\}$. Soit $x \in H$ tel que $A(x) = 0$. Alors

$$0 = \langle A(x), x \rangle = a(x, x) \geq \alpha \|x\|^2$$

par coercivité. Donc $x = 0_H$ et A est injective.

Pour prouver que A est surjective, il suffit de montrer que l'image de A est fermée et dense, ce qui implique que $Im(A) = H$. Dans ce but, notons que, si $x \in H \setminus \{0\}$ et si $y = A(x)$, alors, par hypothèse de coercivité,

$$\alpha \|x\|^2 \leq a(x, x) = \langle A(x), x \rangle = \langle y, x \rangle \leq \|y\| \|x\|$$

Donc, comme $x \neq 0$, on a $\|x\| \leq \|y\|/\alpha$.

Im(A) est fermé : Soit $y_n \in Im(A)$ telle que $y_n \rightarrow y$. Par définition de $Im(A)$ il existe $x_n \in H$ tel que $A(x_n) = y_n$. Notons que (x_n) est une suite de Cauchy : en effet, comme (y_n) converge, (y_n) est une suite de Cauchy. Donc, pour tout $\epsilon > 0$, il existe $N \geq 0$ tel que

$$\|y_n - y_{n+p}\| \leq \alpha \epsilon \quad \forall n \geq N, \forall p \geq 0.$$

Mais alors, comme par linéarité, $A(x_n - x_{n+p}) = (y_n - y_{n+p})$, on a

$$\|x_n - x_{n+p}\| \leq \frac{\|y_n - y_{n+p}\|}{\alpha} \leq \frac{\alpha \epsilon}{\alpha} = \epsilon \quad \forall n \geq N, \forall p \geq 0.$$

Donc (x_n) est de Cauchy. Comme H est complet, (x_n) converge vers un vecteur $x \in H$, et, par continuité de A , on a $A(x) = y$. Cela prouve que $Im(A)$ est fermé.

Im(A) est dense : comme $Im(A)$ est un espace vectoriel, il suffit d'établir que $Im(A)^\perp = \{0_H\}$. Soit $x \in Im(A)^\perp$. Alors $\langle A(y), x \rangle = 0$ pour tout $y \in H$. En particulier, pour $y = x$, cela donne

$$0 = \langle A(x), x \rangle = a(x, x) \geq \alpha \|x\|^2$$

par coercivité. Donc $x = 0_H$ et $Im(A)$ est dense. En conclusion, A est surjective.

En conclusion, A est une application linéaire bijective de H dans H . Si $f \in H^*$, le théorème de Riesz affirme qu'il existe $x_0 \in H$ tel que $f(y) = \langle y, x_0 \rangle$. Mais par bijectivité de A il existe un unique $\bar{x} \in H$ tel que $A(\bar{x}) = x_0$. D'où

$$a(\bar{x}, y) = \langle A(\bar{x}), y \rangle = \langle x_0, y \rangle = f(y) \quad \forall y \in H.$$

On suppose finalement que a est symétrique. Soit $\bar{x} \in H$ tel que $a(\bar{x}, y) = f(y)$ pour tout $y \in H$. Alors, pour tout $x \in H$, on a

$$\begin{aligned} \Phi(x) &= \frac{1}{2} a(x, x) - f(x) = \frac{1}{2} [a(\bar{x}, \bar{x}) + 2a(\bar{x}, x - \bar{x}) + a(x - \bar{x}, x - \bar{x})] - f(\bar{x}) - f(x - \bar{x}) \\ &= \Phi(\bar{x}) + a(\bar{x}, x - \bar{x}) - f(x - \bar{x}) + \frac{1}{2} a(x - \bar{x}, x - \bar{x}) \\ &\geq \Phi(\bar{x}) + 0 + \frac{\alpha}{2} \|x - \bar{x}\|^2 \end{aligned}$$

ce qui prouve que \bar{x} est l'unique point de minimum de Φ . □

2 Espaces de Sobolev et équations elliptiques linéaires.

La plus grande partie du chapitre se situe en dimension d'espace égale à 1 ; la dimension supérieure n'est abordée que dans la dernière partie.

2.1 Espaces de Sobolev sur un intervalle

2.1.1 Définitions et exemples

Soit I un intervalle ouvert non vide de \mathbb{R} et $p \in [1, +\infty]$. On notera \bar{I} l'adhérence de I .

On cherche à décrire des fonctions qui n'ont pas de dérivée au sens usuel, mais qui vérifient quand même des relations d'intégration par parties. Pour cela, nous aurons besoin des remarques suivantes :

Lemme 2.1 Soit $w \in L^p(I)$.

1. Si $\int_I w(x)\phi(x)dx = 0$ pour tout $\phi \in \mathcal{C}_c^1(I)$, alors $w = 0$ p.p. dans I .
2. Si $\int_I w(x)\phi'(x)dx = 0$ pour tout $\phi \in \mathcal{C}_c^1(I)$, alors w est constante : il existe un réel C tel que $w = C$ p.p. dans I .

Remarque : Bien noter que la réciproque des assertions ci-dessus est évidente.

Preuve :

1. Fixons $a < b$, avec $a, b \in I$ et montrons que $w = 0$ p.p. sur $[a, b]$. Notons que $w \in L^1([a, b])$ puisque $w \in L^p([a, b])$. L'étape-clé consiste à montrer qu'il existe une suite de fonctions $\phi_n \in \mathcal{C}_c^1(I)$, bornées par 1, telles que ϕ_n tend vers $\text{sign}(w(x))$ pour presque tout $x \in [a, b]$. En effet, comme $\mathcal{C}_c^1([a, b])$ est dense dans $L^1([a, b])$, il existe une suite de fonction $\psi_n \in \mathcal{C}_c^1(I)$ qui converge vers $\text{sign}(w)$ dans $L^1([a, b])$ et donc, à une sous-suite près encore notée (ψ_n) , presque partout sur $[a, b]$. Soit $\theta : \mathbb{R} \rightarrow [-1, 1]$ une fonction \mathcal{C}^∞ , croissante, telle que $\theta(0) = 0$, $\theta(-1) = -1$ et $\theta(1) = 1$. Alors la suite de fonctions $(\phi_n = \theta \circ \psi_n)$ vérifie les conditions demandées.

Par hypothèse, on a $\int_I w(x)\phi_n(x)dx = 0$. Comme $|w(x)\phi_n(x)| \leq |w(x)|$ et $(w(x)\phi_n(x))$ tend vers $w(x)\text{sign}(w(x)) = |w(x)|$ dans $[a, b]$ et vers 0 en dehors de $[a, b]$, on a, par convergence dominée, $\int_{[a,b]} |w(x)|dx = 0$, soit $w = 0$ p.p. dans I .

2. Soit maintenant $\phi \in \mathcal{C}_c^1(I)$ et $\eta \in \mathcal{C}_c^1(I)$ avec $\int_I \eta = 1$. Notons qu'il existe $\psi \in \mathcal{C}_c^1(I)$ tel que $\psi'(x) = \phi(x) - (\int_I \phi(y)dy)\eta(x)$, puisque la fonction $x \rightarrow \phi(x) - (\int_I \phi(y)dy)\eta(x)$ est à support compact et d'intégrale nulle. On applique l'hypothèse à ψ pour obtenir

$$\int_I w(x) \left(\phi(x) - \left(\int_I \phi(y)dy \right) \eta(x) \right) dx = 0$$

Comme

$$\int_I w(x) \left(\int_I \phi(y)dy \right) \eta(x) dx = \int_I \phi(x) \left(\int_I w(y)\eta(y)dy \right) dx$$

par Fubini et changement de variable, on a

$$\int_I \phi(x) \left(w(x) - \left(\int_I w(y)\eta(y)dy \right) \right) dx = 0$$

Comme ceci est vrai pour toute fonction $\phi \in \mathcal{C}_c^1(I)$, on en déduit que $w(x) - (\int_I w(y)\eta(y)dy) = 0$ p.p., et donc que w est égal presque partout à la constante $(\int_I w(y)\eta(y)dy)$.

□

Définition 2.2 On dit qu'une fonction u est dans l'espace de Sobolev $W^{1,p}(I)$ si $u \in L^p(I)$ et s'il existe une fonction $w \in L^p(I)$ telle que

$$\int_I u(x)\phi'(x)dx = - \int_I w(x)\phi(x)dx \quad \forall \phi \in \mathcal{C}_c^1(I).$$

L'espace le plus utilisé est lorsque $p = 2$, et on note plutôt $H^1(I) := W^{1,2}(I)$.

Notons que, si I est borné, les fonctions de classe \mathcal{C}^1 dans \bar{I} appartiennent à tous les espaces $W^{1,p}(I)$. Si I est non borné, c'est le cas de la restriction à I de toutes les fonctions de classe \mathcal{C}^1 et à support compact dans \mathbb{R} .

Remarque : La fonction w est définie de façon unique. En effet, si w_1 et w_2 vérifient toutes deux la relation

$$\int_I u(x)\phi'(x)dx = - \int_I w_1(x)\phi(x)dx = - \int_I w_2(x)\phi(x)dx \quad \forall \phi \in \mathcal{C}_c^1(I),$$

alors

$$\int_I (w_1(x) - w_2(x))\phi(x)dx = 0 \quad \forall \phi \in \mathcal{C}_c^1(I).$$

On en déduit (grâce au Lemme 2.1) que $w_1 = w_2$ p.p.. On notera $u' = w$ dans toute la suite. La relation caractérisant u' est donc

$$\int_I u(x)\phi'(x)dx = - \int_I u'(x)\phi(x)dx \quad \forall \phi \in \mathcal{C}_c^1(I).$$

Lemme 2.3 Si u appartient à $W^{1,p}(I)$, alors u possède un représentant continu dans \bar{I} et

$$u(y) - u(x) = \int_x^y u'(t)dt \quad \forall x, y \in I.$$

A partir de maintenant, nous choisirons systématiquement le représentant continu d'une fonction $u \in W^{1,p}(I)$.

Preuve *: Pour fixer les idées, on suppose que $I =]a, b[$. Posons $\xi(x) = \int_a^x u'(t)dt$ et fixons $\phi \in \mathcal{C}_c^1(I)$. On a par Fubini

$$\int_I \xi(x)\phi'(x)dx = \int_I \int_a^x u'(t)\phi'(x) dt dx = \int_I u'(t) \int_t^b \phi'(x)dx = - \int_I u'(t)\phi(t)dt = \int_I u(x)\phi'(x)dx.$$

Donc $\int_I (\xi(x) - u(x))\phi'(x)dx = 0$ pour toute fonction $\phi \in \mathcal{C}_c^1(I)$. Le lemme 2.1 affirme alors qu'il existe une constante C telle que

$$u(x) = C + \xi(x) = C + \int_a^x u'(t)dt \quad p.p. x \in I.$$

Montrons maintenant que la fonction ξ est continue dans \bar{I} . Comme $u' \in L^p(I)$, on a $u' \in L^1([a, b])$. Alors si la suite (x_n) tend vers x dans $[a, b]$, on a

$$\xi(x_n) = \int_{[a, b]} u'(t)\mathbf{1}_{[a, x_n]}(t)dt$$

La suite de fonction $(u'(t)\mathbf{1}_{[a,x_n]}(t))$ tend p.p. vers la fonction $u'(t)\mathbf{1}_{[a,x]}(t)$ tandis que $|u'\mathbf{1}_{[a,x_n]}| \leq |u'|$ où $|u'|$ est dans $L^1([a,b])$. Donc, par théorème de convergence dominée, on a

$$\lim_{n \rightarrow +\infty} \xi(x_n) = \int_{[a,b]} u'(t)\mathbf{1}_{[a,x]}(t)dt = \xi(x).$$

Ceci prouve que ξ est continue et, comme $u = C + \xi$ p.p., que u possède un représentant continu. De plus, par définition de ξ , on a

$$\xi(y) - \xi(x) = \int_x^y u'(t)dt \quad \forall x, y \in I.$$

□

En fait, lorsque $p > 1$, la régularité des fonctions u de $W^{1,p}(I)$ peut être quantifiée :

Lemme 2.4 *Si $p > 1$ et $u \in W^{1,p}(I)$, alors*

$$|u(y) - u(x)| \leq \|u'\|_p |y - x|^{1-1/p} \quad \forall x, y \in \bar{I}.$$

Preuve *: Fixons $x, y \in I$ avec, pour fixer les idées, $x < y$. En utilisant le résultat précédent ainsi que l'inégalité de Hölder (avec les fonctions $u' \in L^p$ et $v = \mathbf{1}_{[x,y]}$) on obtient, en posant $q = p/(p-1)$ (i.e., $1/q = 1 - 1/p$, avec convention $1/\infty = 0$),

$$|u(y) - u(x)| = \left| \int_I \mathbf{1}_{[x,y]}(t)u'(t)dt \right| \leq \left(\int_I \mathbf{1}_{[x,y]}^q(t) \right)^{1/q} \left(\int_I |u'(t)|^p \right)^{1/p} = |y - x|^{1-1/p} \|u'\|_p$$

□

Théorème 2.5 *L'espace $W^{1,p}(I)$ est un espace de Banach lorsqu'on le munit de la norme*

$$\|u\|_{W^{1,p}(I)} = (\|u\|_p^p + \|u'\|_p^p)^{1/p}$$

De plus, si $p = 2$, l'espace $H^1(I) = W^{1,2}(I)$ est un espace de Hilbert pour le produit scalaire

$$\langle u, v \rangle = \int_I u(t)\overline{v(t)}dt + \int_I u'(t)\overline{v'(t)}dt$$

Preuve* : Le fait que $W^{1,p}(I)$ est un espace vectoriel et que $\|\cdot\|_{W^{1,p}(I)}$ est une norme sur cet espace est laissé en exercice (facile... à condition de se souvenir que $(a, b) \rightarrow (|a|^p + |b|^p)^{1/p}$ est une norme dans \mathbb{R}^2).

Montrons que $W^{1,p}(I)$ est complet pour cette norme. Soit (u_n) une suite de Cauchy de $W^{1,p}(I)$. Alors, comme

$$\|u\|_p \leq \|u\|_{W^{1,p}} \quad \text{et} \quad \|u'\|_p \leq \|u\|_{W^{1,p}} \quad \forall u \in W^{1,p},$$

les suites (u_n) et (u'_n) sont de Cauchy dans L^p . Or L^p est complet, donc ces suites convergent vers une limite $u \in L^p$ et $w \in L^p$ respectivement. Reste à montrer que $u \in W^{1,p}$ et que (u_n) tend vers u dans $W^{1,p}$. Par définition de $W^{1,p}(I)$, on a

$$\int_I u_n(x)\phi'(x)dx = - \int_I u'_n(x)\phi(x)dx \quad \forall \phi \in C_c^1(I).$$

On passe facilement à la limite dans chacune des expressions, ce qui donne

$$\int_I u(x)\phi'(x)dx = - \int_I w(x)\phi(x)dx \quad \forall \phi \in C_c^1(I).$$

Comme u et w sont dans $L^p(I)$, cela prouve que $u \in W^{1,p}(I)$ et que $u' = w$. La convergence de (u_n) vers u dans $W^{1,p}(I)$ est alors une conséquence immédiate de la convergence de (u_n) et (u'_n) vers u et $w = u'$ dans $L^p(I)$. \square

On admettra par la suite que si $u \in W^{1,p}(I)$, alors u admet une dérivée (usuelle) égale à $u'(x)$ en presque tout point x de I . Donc, en pratique, pour montrer qu'une fonction donnée appartient à $W^{1,p}(I)$, il suffit

1. de vérifier que $u \in L^p(I)$,
2. de prouver que la dérivée (usuelle) $u'(x)$ de u existe en presque tout point x de I et que $u' \in L^p(I)$,
3. et enfin de vérifier qu'il existe $x_0 \in I$ et $\alpha \in \mathbb{R}$ tels que $u(x) = \alpha + \int_{x_0}^x u'(t)dt$.

Attention, le dernier point est essentiel. Par exemple, si $u(x)$ vaut 0 sur $[0, 1/2]$ et 1 sur $[1/2, 1]$, alors $u'(x)$ existe et vaut 0 pour tout $x \neq 1/2$, mais bien sûr, u n'est pas une primitive de 0.

Si $u \in W^{1,p}(I)$ et $x \in I$, alors u est une fonction continue et la quantité $u(x)$ est définie sans ambiguïté. On peut aller un peu plus loin lorsque $p > 1$ (et en particulier lorsque $p = 2$) :

Théorème 2.6 (Fonction évaluation) *Pour tout $x \in \bar{I}$, l'application $e_x : W^{1,p}(I) \rightarrow \mathbb{R}$, qui à $u \in W^{1,p}(I)$ associe $e_0(u) := u(x)$, est linéaire continue.*

Preuve : L'application est clairement linéaire. Montrons qu'il existe une constante, qui ne dépend que de I et de p , telle que $|e_x(u)| \leq K \|u\|_{W^{1,p}}$. Notons d le diamètre de I (on pose, par exemple, $d = 1$ si le diamètre est infini). Comme I a un intérieur non vide, d est strictement positif. De plus, comme I est un intervalle, soit $[x, x + d/2[\subset \bar{I}$, soit $[x - d/2, x] \subset \bar{I}$. On suppose pour fixer les idées qu'on est dans le premier cas et on pose $I_0 := [x, x + d/2[$. Alors, pour tout $y \in I_0$, on a

$$|u(y)| \geq |u(x)| - |u(y) - u(x)| \geq |u(x)| - |x - y|^{1/q} \|u'\|_{L^p}$$

d'après le lemme 2.4. Distinguons deux cas : si $|u(x)| \leq 2|d/2|^{1/q} \|u'\|_{L^p}$, alors on a une estimation de $|u(x)|$. Sinon, $|u(x)| > 2|d/2|^{1/q} \|u'\|_{L^p}$ et donc

$$|u(y)| \geq |u(x)| - |d/2|^{1/q} \|u'\|_{L^p} \geq |u(x)| - |u(x)|/2 = |u(x)|/2.$$

Mais alors, $\int_x^{x+d/2} |u(y)|^p \geq (d/2) \frac{|u(x)|^p}{2^p}$, et donc $|u(x)| \leq (2/d)^{1/p} \|u\|_p$.

On a donc prouvé que

$$|e_x(u)| = |u(x)| \leq \max\{2|d/2|^{1/q} \|u'\|_{L^p}, (2/d)^{1/p} \|u\|_p\} \leq K \|u\|_{W^{1,p}}$$

avec $K = \max\{2|d/2|^{1/q}, (2/d)^{1/p}\}$. \square

2.1.2 Approximation

Commençons par le cas où $I = \mathbb{R}$.

Théorème 2.7 *On suppose que $p \in [1, +\infty[$. Alors l'espace $C^\infty(\mathbb{R}) \cap W^{1,p}(\mathbb{R})$ est dense dans $W^{1,p}(\mathbb{R})$.*

Remarque 2.8 Comme souvent avec les résultats d'approximation, le résultat est faux pour $p = +\infty$. En fait on peut montrer aussi que $C_c^\infty(\mathbb{R})$ est aussi dense dans $W^{1,p}(\mathbb{R})$.

Preuve *: La preuve se fait par convolution. Soit ϕ un noyau de convolution "standard" : ϕ est de classe \mathcal{C}^∞ , à support dans $[-1, 1]$, avec $\phi \geq 0$ et $\int_{\mathbb{R}} \phi = 1$. On pose $\phi_\epsilon(x) = \phi(x/\epsilon)/\epsilon$. Alors ϕ_ϵ est à support dans $[-\epsilon, \epsilon]$ et $\int_{\mathbb{R}} \phi_\epsilon = 1$.

Soit $u \in W^{1,p}(\mathbb{R})$ et notons $u_\epsilon = \phi_\epsilon \star u$. Rappelons que u_ϵ est de classe \mathcal{C}^∞ et converge vers u dans $L^p(\mathbb{R})$.

Nous montrons maintenant que $u'_\epsilon = \phi_\epsilon \star u'$, ce qui prouve que u'_ϵ appartient à $L^p(\mathbb{R})$ (puisque c'est le cas de u') et converge aussi vers u' dans $L^p(\mathbb{R})$. En effet, rappelons que $u'_\epsilon = \phi'_\epsilon \star u$. Or, pour tout $x \in \mathbb{R}$, la fonction $t \rightarrow \phi_\epsilon(x-t)$ est de classe $\mathcal{C}^\infty(\mathbb{R})$ et, par définition de la dérivée au sens des distributions, on a

$$(\phi'_\epsilon \star u)(x) = \int_{\mathbb{R}} \phi'_\epsilon(x-t)u(t)dt = - \int_{\mathbb{R}} u(t) \frac{d}{dt}(\phi(x-t))dt = \int_{\mathbb{R}} u'(t)\phi_\epsilon(x-t)dt = (\phi_\epsilon \star u')(x).$$

Comme u_ϵ et u'_ϵ convergent respectivement vers u et u' dans $L^p(\mathbb{R})$ lorsque $\epsilon \rightarrow 0$, on a que u_ϵ tend vers u dans $W^{1,p}(\mathbb{R})$. \square

On note $\mathcal{C}^\infty([a, b])$ la restriction à $[a, b]$ des fonctions $\mathcal{C}^\infty(\mathbb{R})$.

Théorème 2.9 *On suppose que $I =]a, b[$ est un intervalle ouvert borné. Alors l'espace $\mathcal{C}^\infty([a, b])$ est dense dans $W^{1,p}([a, b])$ pour tout $p \in [1, +\infty[$. Plus précisément, pour tout $u \in W^{1,p}(I)$, il existe une suite de fonction (u_n) qui converge uniformément vers u dans $[a, b]$ et telle que (u'_n) converge vers u' dans $L^p(I)$.*

Remarque 2.10 1. Comme précédemment, le résultat est faux pour $p = +\infty$.

2. Il n'est pas vrai non plus que $\mathcal{C}^\infty_c(]a, b[)$ est dense dans $W^{1,p}([a, b])$. Nous verrons plus loin que l'adhérence de $\mathcal{C}^\infty_c(]a, b[)$ pour la norme de $W^{1,p}([a, b])$ est l'ensemble $W_0^{1,p}([a, b])$ des fonctions de $W^{1,p}([a, b])$ qui s'annulent en a et en b .

Preuve * : La preuve se fait en utilisant le théorème précédent et un argument de prolongement des fonctions de $W^{1,p}([a, b])$ à $W^{1,p}(\mathbb{R})$. Fixons $u \in W^{1,p}([a, b])$ et $\epsilon > 0$. On affirme qu'il existe $w \in W^{1,p}(\mathbb{R})$ tel que $u = w$ dans $[a, b]$. En effet, si on pose

$$w(x) = \begin{cases} u(x) & \text{dans } [a, b] \\ (x-a+1)u(a) & \text{dans } [a-1, a] \\ (b+1-x)u(b) & \text{dans } [b, b+1] \\ 0 & \text{ailleurs} \end{cases}$$

alors w est dans $W^{1,p}(\mathbb{R})$ avec $w'(x) = u'(x)$ p.p. dans $[a, b]$, $w'(x) = u(a)$ dans $]a-1, a[$ et $w'(x) = -u(b)$ dans $[b, b+1]$.

Comme $\mathcal{C}^\infty(\mathbb{R}) \cap W^{1,p}(\mathbb{R})$ est dense dans $W^{1,p}(\mathbb{R})$, il existe $w_\epsilon \in \mathcal{C}^\infty(\mathbb{R}) \cap W^{1,p}(\mathbb{R})$ tel que $\|w - w_\epsilon\|_{W^{1,p}(\mathbb{R})} \leq \epsilon$. Rappelons que w_ϵ est obtenu par simple convolution de la fonction continue à support compact w , et donc w_ϵ converge uniformément vers w dans \mathbb{R} tandis que (w'_ϵ) tend vers w' dans $L^p(\mathbb{R})$. Définissons u_ϵ comme la restriction de w_ϵ à $[a, b]$. Alors $u_\epsilon \in W^{1,p}([a, b])$, u_ϵ converge uniformément vers u dans $[a, b]$ et

$$\|u - u_\epsilon\|_{W^{1,p}([a,b])} = \|u - u_\epsilon\|_{L^p([a,b])} + \|u' - u'_\epsilon\|_{L^p([a,b])} \leq \|w - w_\epsilon\|_{L^p(\mathbb{R})} + \|w' - w'_\epsilon\|_{L^p(\mathbb{R})} \leq \epsilon$$

puisque $w = u$ et $w_\epsilon = u_\epsilon$ dans $[a, b]$. \square

Une application des techniques précédentes est la formule d'intégration par parties :

Proposition 2.11 *On suppose que $I =]a, b[$ avec $a < b$ réels et $p \geq 1$. Si $u, v \in W^{1,p}(I)$, alors $uv \in W^{1,p}(I)$ et $(uv)' = u'v + uv'$. De plus, on a*

$$\int_a^b u(x)v'(x)dx = [u(x)v(x)]_a^b - \int_a^b u'(x)v(x)dx$$

Preuve : Comme u et v sont dans $W^{1,p}(I)$, u et v sont bornés et donc $uv' + u'v \in L^p(I)$. Le reste de la preuve se fait en supposant que $p = 1$ sans perte de généralité.

Régularisons u et v par des fonctions (u_n) et (v_n) de classe \mathcal{C}^∞ qui convergent dans $W^{1,1}(I)$ et uniformément sur $[a, b]$ vers u et v respectivement. Comme, pour tout $\phi \in \mathcal{C}_c^1(I)$, on a

$$\int_a^b u_n v_n \phi' = - \int_a^b (u_n v_n)' \phi = - \int_a^b (u_n' v_n + u_n v_n') \phi$$

on obtient, en passant à la limite,

$$\int_a^b uv \phi' = - \int_a^b (u'v + uv') \phi$$

ce qui prouve que $(u'v + uv')$ est la dérivée au sens des distributions de uv . Or uv est borné et $(u'v + uv')$ est dans L^1 , ce qui prouve que $uv \in W^{1,1}(I)$ (dans le cas où $p \in [1, +\infty]$, le même argument montre que $uv \in W^{1,p}(I)$).

Montrons finalement la formule d'intégration par parties : elle est vraie pour u_n et v_n :

$$\int_a^b u_n(x) v_n'(x) dx = [u_n(x) v_n(x)]_a^b - \int_a^b u_n'(x) v_n(x) dx$$

Comme (u_n) (resp. (v_n)) converge uniformément vers u (resp. v) et u_n' (resp. v_n') converge vers u' (resp. v') dans L^1 , on peut passer à la limite dans l'égalité ci-dessus pour obtenir le résultat. \square

2.1.3 L'espace $W_0^{1,p}(I)$

On suppose ici que $I =]a, b[$. On appelle $W_0^{1,p}(I)$ le sous-ensemble de $W^{1,p}(]a, b[)$ constitué des éléments $u \in W^{1,p}(]a, b[)$ tels que $u(a) = u(b) = 0$.

Proposition 2.12 $W_0^{1,p}(]a, b[)$ est un sous-espace vectoriel fermé de $W^{1,p}(]a, b[)$. En particulier, muni de la distance de $W^{1,p}(]a, b[)$, c'est un espace de Banach.

L'espace le plus utilisé est lorsque $p = 2$, et on note plutôt $W_0^{1,2}(I) = H_0^1(I)$; c'est un espace de Hilbert.

Preuve * : Lorsque $p > 1$, le résultat est évident, puisque l'application $e_a : W^{1,p}(]a, b[) \rightarrow \mathbb{R}$ et $e_b : W^{1,p}(]a, b[) \rightarrow \mathbb{R}$, qui à u associent $u(a)$ et $u(b)$, sont linéaires continues. Donc $W_0^{1,p}(]a, b[)$ est fermé, comme intersection du noyau de e_a et de celui de e_b (qui sont fermés par continuité).

Lorsque $p = 1$, il suffit de prendre la définition : soit (u_n) une suite de $W_0^{1,1}(I)$ qui tend dans $W^{1,1}$ vers une fonction $u \in W^{1,1}(I)$. Alors on sait que, pour tout $x \in I$, $u_n(x) = \int_a^x u_n'(t) dt$. Comme, en particulier, (u_n) tend vers u dans $L^1(I)$, il existe une sous-suite (u_{n_k}) qui tend vers u presque partout. Si $x \in I$ est un tel élément, on a $u_{n_k}(x) \rightarrow u(x)$ tandis que $(\int_a^x u_{n_k}'(t) dt)$ tend vers $\int_a^x u'(t) dt$ par convergence L^1 de (u_n') vers u' . Donc $u(x) = \int_a^x u'(t) dt$ pour presque tout $x \in I$ et, par continuité de u , pour tout $x \in I$. En particulier, $u(a) = 0$. On montre de même que $u(b) = 0$, ce qui conclut la preuve de la fermeture de $W_0^{1,1}(I)$. \square

Théorème 2.13 Soit $p \in [1, +\infty[$. Alors l'espace $\mathcal{C}_c^\infty(]a, b[)$ est dense dans $W_0^{1,p}(]a, b[)$.

Idée de la preuve : Quitte à faire une translation, on peut supposer que l'intervalle $[a, b]$ est symétrique par rapport à 0, i.e. $b > 0$ et $a = -b$. Soit $u \in W_0^{1,p}([-b, b])$ et $\delta > 0$ fixé. Soit $\epsilon \in]0, 1[$ et $u_\epsilon(x) = u(x/(1-\epsilon))$ si $x \in]-b(1-\epsilon), b(1-\epsilon)[$ et $u_\epsilon(x) = 0$ si $x \in [-b, b] \setminus]-b(1-\epsilon), b(1-\epsilon)[$. Alors il est facile de voir que $u_\epsilon \in W_0^{1,p}([-b, b])$ avec $u_\epsilon'(x) = u'(x/(1-\epsilon))/(1-\epsilon)$ si $x \in]-b(1-\epsilon), b(1-\epsilon)[$

et $u'_\epsilon(x) = 0$ si $x \in [-b, b] \setminus]-b(1-\epsilon), b(1-\epsilon)[$. De plus, on peut montrer que (u_ϵ) tend vers u dans $W^{1,p}([-b, b])$ lorsque $\epsilon \rightarrow 0$: la convergence de (u_ϵ) vers u dans L^p est directe (en fait uniforme) ; la partie plus délicate est la convergence de (u'_ϵ) vers u' , qui se montre comme pour la continuité des translations dans $L^1(\mathbb{R})$ (en utilisant la densité des fonctions continues dans L^p).

En particulier, il existe $\epsilon > 0$ tel que $\|u - u_\epsilon\|_{W^{1,p}} \leq \delta/2$. Maintenant u_ϵ est à support contenu dans $]-b(1-\epsilon), b(1-\epsilon)[$ et peut être approché par régularisation par convolution par une fonction $v \in C_c^\infty(\mathbb{R})$ dont le support est contenu dans l'intervalle un peu plus grand $]-b(1-\epsilon/2), b(1-\epsilon/2)[$. En particulier, on peut prendre $v \in C_c^\infty(]-b, b[)$ et $\|v - u_\epsilon\|_{W^{1,p}} \leq \delta/2$. Alors $\|u - v\|_{W^{1,p}} \leq \delta$. \square

Théorème 2.14 (Inégalité de Poincaré) *Soit $p \geq 1$. Il existe une constante C , qui ne dépend que de $b - a$ et de p , telle que*

$$\|u\|_{W^{1,p}(]a,b[)} \leq C \|u'\|_{L^p(]a,b[)} \quad \forall u \in W_0^{1,p}(]a,b[)$$

Remarque 2.15 La conclusion de ce résultat est fautive pour $W^{1,p}(]a,b[)$ (penser aux fonctions constantes!). De façon à peine plus subtile, elle est fautive pour $W^{1,p}(\mathbb{R})$ (alors que les fonctions de $W^{1,p}(\mathbb{R})$ tendent vers 0 à l'infini : voir exercice) : cela se voit par un argument d'échelle. En effet, soit u n'importe quel élément de $W^{1,p}(\mathbb{R})$ et posons $u_\epsilon(x) = u(\epsilon x)$. Alors $u'_\epsilon(x) = \epsilon u'(\epsilon x)$, et donc $\|u'_\epsilon\|_p = \epsilon^{1-1/p} \|u'\|_p$, tandis que $\|u_\epsilon\|_p = \epsilon^{-1/p} \|u\|_p$. Si $u \not\equiv 0$, on ne peut espérer avoir une constante $C > 0$ telle que $\epsilon^{-1/p} \|u\|_p = \|u_\epsilon\|_p \leq C \|u'_\epsilon\|_p = \epsilon^{1-1/p} \|u'\|_p$ pour tout $\epsilon > 0$.

Preuve : Vu la définition de la norme sur $W^{1,p}(]a,b[)$, il suffit de montrer l'existence d'une constante C telle que, pour tout $u \in W_0^{1,p}(]a,b[)$, on a $\|u\|_p \leq C \|u'\|_p$. Or, comme $u(a) = 0$, on a pour tout $x \in [a, b]$,

$$|u(x)| = |u(x) - u(a)| = \left| \int_a^x u'(t) dt \right| \leq \|u'\|_1 \leq (b-a)^{1-1/p} \|u'\|_p$$

(où la dernière inégalité est obtenue par inégalité de Hölder). Donc

$$\|u\|_p \leq (b-a)^{1/p} \|u\|_\infty \leq (b-a) \|u'\|_p .$$

\square

2.2 Application aux équations elliptiques en dimension 1

Soit $I :=]a, b[$ un intervalle de \mathbb{R} , avec $a < b$. On s'intéresse ici aux équations de la forme

$$(P) \quad -u''(x) + c(x)u(x) = f(x) \text{ dans }]a, b[, \quad u(a) = u(b) = 0 ,$$

où c et f sont des applications continues de $[a, b]$ dans \mathbb{R} , **avec** $c \geq 0$. Ce problème est le prototype d'équation elliptique (mais une caricature d'EDP, puisqu'on peut trouver les solutions en résolvant une EDO ! Attention cependant, ce n'est pas une EDO habituelle : en effet, dans la théorie classique des EDO, on fixe une condition initiale du type $(u(a), u'(a))$, l'EDO étant d'ordre 2. Ici on fixe une condition initiale $u(a)$ et une condition finale $u(b)$).

La méthode de résolution est la suivante :

1. on écrit le problème sous une forme faible (dans laquelle la dérivée seconde n'apparaît pas).
2. on utilise le théorème de Lax-Milgram pour montrer que le problème sous sa formulation faible possède une solution,
3. on montre que cette solution est régulière et solution classique du problème initial.

2.2.1 Existence et unicité d'une solution faible

Proposition 2.16 *Si u une solution de classe \mathcal{C}^2 du problème (P), alors*

$$\int_I (u'(x)\phi'(x) + c(x)u(x)\phi(x)) dx = \int_I f(x)\phi(x)dx \quad \forall \phi \in \mathcal{C}_c^\infty(I) \quad (7)$$

Preuve* : On multiplie l'équation par $\phi \in \mathcal{C}_c^\infty(I)$ et on intègre :

$$\int_I (-u''(x) + c(x)u(x))\phi(x)dx = \int_I f(x)\phi(x)dx ,$$

ce qui donne, par intégration par parties dans le premier terme :

$$[-u'(x)\phi(x)]_a^b + \int_I (u'(x)\phi'(x) + c(x)u(x)\phi(x))dx = \int_I f(x)\phi(x)dx ,$$

Or $\phi(a) = \phi(b) = 0$ puisque ϕ est à support compact. D'où le résultat. \square

Afin de pouvoir utiliser les résultats d'analyse fonctionnelle, on va chercher la solution u dans un espace a priori plus large que \mathcal{C}^2 . Comme l'expression (7) ne comporte que des dérivées d'ordre 1 et qu'on cherche une fonction qui s'annule en a et b , il est naturel de travailler avec $H_0^1(I)$. Notons que l'expression (7) a un sens si u et ϕ sont dans $H_0^1(I)$: dans ce cas, u' et ϕ' étant dans L^2 , $u'\phi'$ est dans L^1 ; de plus u et ϕ sont continues, ce qui garantit que $cu\phi \in L^1$ et $f\phi \in L^1$, l'intervalle $I =]a, b[$ étant borné. On parle alors de solution faible de (P) :

Définition 2.17 (Formulation faible) *On dit que $u \in H_0^1(I)$ est une solution faible de (P) si*

$$\int_I (u'(x)v'(x) + c(x)u(x)v(x)) dx = \int_I f(x)v(x)dx \quad \forall v \in H_0^1(I)$$

Au vu de la formulation faible, on pose

$$a(u, v) = \int_I (u'(x)v'(x) + c(x)u(x)v(x)) dx \quad \forall u, v \in H_0^1(I) .$$

Lemme 2.18 *On suppose que $c \in L^\infty([a, b])$. Alors a est une forme bilinéaire continue sur $H_0^1(I) \times H_0^1(I)$.*

Preuve* : Comme $u', v' \in L^2$, le produit $u'v'$ est dans L^1 . D'autre part, comme c est dans L^∞ , le produit cuv est dans L^1 (en fait dans L^∞ car les fonctions u et v sont continues sur $[a, b]$, donc bornées). En conséquence, l'intégrale définissant $a(u, v)$ est bien définie. Comme a est clairement bilinéaire, il suffit de montrer qu'il existe une constante $C > 0$ telle que

$$|a(u, v)| \leq C \|u\|_{H^1(I)} \|v\|_{H^1(I)} \quad \forall u, v \in H_0^1(I) .$$

D'après Cauchy-Schwarz, on a

$$\left| \int_I u'v' \right| \leq \|u'\|_2 \|v'\|_2 \leq \|u\|_{H^1(I)} \|v\|_{H^1(I)} .$$

D'autre part,

$$\begin{aligned} \left| \int_I c(x)u(x)v(x) dx \right| &\leq \int_I |c(x)u(x)v(x)| dx \leq \|c\|_\infty \int_I |u(x)v(x)| dx \\ &\leq \|c\|_\infty \|u\|_2 \|v\|_2 \leq \|c\|_\infty \|u\|_{H^1(I)} \|v\|_{H^1(I)} \end{aligned}$$

Donc

$$|a(u, v)| \leq (1 + \|c\|_\infty) \|u\|_{H^1(I)} \|v\|_{H^1(I)} .$$

\square

Lemme 2.19 *On suppose maintenant que $c \in L^\infty([a, b])$, avec $c \geq 0$ p.p.. Alors la forme bilinéaire a est coercive.*

Preuve* : On utilise ici l'inégalité de Poincaré, qui affirme l'existence d'une constante $C > 0$ telle que

$$\|u\|_{H^1(I)} \leq C\|u'\|_2 \quad \forall u \in H_0^1(I).$$

Pour tout $u \in H_0^1(I)$, on a, puisque $c \geq 0$,

$$a(u, u) = \int_I ((u'(x))^2 + c(x)(u(x))^2) dx \geq \|u'\|_2^2 \geq (1/C)\|u\|_{H^1(I)}$$

□

Proposition 2.20 *On suppose que $c \in L^\infty([a, b])$, avec $c \geq 0$ p.p. et que $f \in L^2([a, b])$. Alors il existe un unique solution faible $u \in H_0^1(I)$ de (P) :*

$$(*) \quad \int_I (u'(x)v'(x) + c(x)u(x)v(x)) dx = \int_I f(x)u(x) dx \quad \forall v \in H_0^1(I).$$

Preuve* : Posons $F(v) = \int_I f(x)v(x) dx$ pour $v \in H_0^1(I)$. Notons que F est une forme linéaire continue puisque F est bien défini (car f et v sont bornés), F est clairement linéaire et

$$|F(v)| \leq \int_I |f(x)| |v(x)| dx \leq \|f\|_2 \|v\|_2 \leq \|f\|_2 \|v\|_{H^1(I)}$$

pour tout $v \in H_0^1(I)$.

La relation (*) se réécrit :

$$(**) \quad a(u, v) = F(v) \quad \forall v \in H_0^1(I).$$

Comme F est une forme linéaire continue sur $H_0^1(I)$ et que a est une forme bilinéaire continue coercive sur $H_0^1(I) \times H_0^1(I)$, où $H_0^1(I)$ est un espace de Hilbert, le théorème de Lax-Milgram affirme qu'il existe un unique élément $u \in H_0^1(I)$ pour lequel (**) est vrai. □

2.2.2 Régularité de la solution

Dans cette partie, nous complétons le programme, en montrant que la solution faible construite auparavant est en fait une solution au sens classique. En étape préliminaire, on a

Lemme 2.21 *Sous les hypothèses de la proposition 2.20, si $u \in H_0^1(I)$ est solution faible du problème, alors $u' \in H^1(I)$ et $(u')'(x) = c(x)u(x) - f(x)$ presque partout.*

Preuve* : On utilise la définition de l'espace de Sobolev $H^1(I)$ et le fait que u est solution faible pour déduire que

$$\int_I u'(x)\phi'(x) dx = a(u, \phi) - \int_I c(x)u(x)\phi(x) dx = \int_I (f(x) - c(x)u(x))\phi(x) dx \quad \forall \phi \in C_c^1(I).$$

Or $f - cu$ est dans $L^2(I)$, puisque c est dans L^∞ , u est continue, et f dans L^2 . Comme $u' \in L^2$ par hypothèse, on en déduit que u' est dans $H^1(I)$ avec $u' = -(f - cu) = cu - f$ p.p. □

Nous pouvons maintenant compléter la démarche :

Théorème 2.22 *On suppose que c et f sont continues dans $[a, b]$ et que $c \geq 0$ dans $[a, b]$. Alors le problème (P) possède une unique solution de classe C^2 .*

Preuve* : Si u est une solution de classe \mathcal{C}^2 de (P), alors u est solution faible d'après la proposition 2.16 et donc il y a au plus une telle solution.

D'autre part, si $u \in H_0^1(I)$ est solution faible du problème, alors nous avons vu ci-dessus que $u' \in H^1(I)$ et $(u')'(x) = c(x)u(x) - f(x)$ presque partout. Or la fonction $x \rightarrow c(x)u(x) - f(x)$ est continue, ce qui montre que $(u)'$ possède un représentant continu. Donc u est de classe \mathcal{C}^2 et satisfait $-u'' + cu = f$ dans I . Finalement, $u(a) = u(b) = 0$ puisque u appartient à $H_0^1(I)$. \square

2.3 Application au problème avec conditions au bord de type Neumann

La démarche décrite précédemment se généralise à de nombreux autres problèmes. On se contente de décrire celui avec les conditions au bord de type Neumann. Soit $I :=]a, b[$. On s'intéresse maintenant au problème

$$(N) \quad -u''(x) + c(x)u(x) = f(x) \text{ dans }]a, b[, \quad u'(a) = u'(b) = 0,$$

où c et f sont des applications continues de $[a, b]$ dans \mathbb{R} , **avec** $c > 0$ **dans** $[a, b]$. La différence avec le problème précédent se lit dans les conditions aux limites $u'(a) = u'(b) = 0$ (qui remplace les conditions de type Dirichlet $u(a) = u(b) = 0$).

Pour trouver la notion de solution faible, on suppose que la solution u est de classe \mathcal{C}^2 et on multiplie par $\phi \in C^\infty$ (pas nécessairement nulle au bord de I) : on a alors

$$\int_I (u'(x)\phi'(x) + c(x)u(x)\phi(x)) dx = \int_I f(x)\phi(x) dx \quad \forall \phi \in C^\infty(I)$$

car le terme $[-u'(x)\phi(x)]_a^b$ dans l'intégration par parties est nul grâce aux conditions au bord $u'(a) = u'(b) = 0$.

Cela conduit à la formulation faible :

Définition 2.23 On dit que $u \in H^1(I)$ est une solution faible de (N) si

$$\int_I (u'(x)v'(x) + c(x)u(x)v(x)) dx = \int_I f(x)v(x) dx \quad \forall v \in H^1(I)$$

Comme précédemment, on introduit la forme bilinéaire

$$a(u, v) = \int_I (u'(x)v'(x) + c(x)u(x)v(x)) dx \quad \forall u, v \in H^1(I).$$

Lemme 2.24 On suppose que c est dans L^∞ et qu'il existe une constante $c_0 > 0$ telle que $c(x) \geq c_0$ p.p. sur $[a, b]$. Alors la forme bilinéaire a est continue et coercive sur $H^1(I)$.

Noter que l'hypothèse sur c est bien sûr satisfaite dès que c est continue et positive sur $[a, b]$.

Preuve* : La continuité de a se montre comme précédemment. La coercivité vient de l'hypothèse $c > 0$ qui, par continuité de c devient $c(x) \geq c_0$ pour tout $x \in I$, où $c_0 > 0$. D'où, pour tout $u \in H^1(I)$,

$$a(u, u) = \int_I ((u'(x))^2 + c(x)(u(x))^2) dx \geq \int_I ((u'(x))^2 + c_0(u(x))^2) dx \geq (1 \wedge c_0) \|u\|_{H^1}^2.$$

\square

Comme l'application $v \rightarrow \int_I f v$ est continue sur $H^1(I)$ (cela se montre comme précédemment), le théorème de Lax-Milgram affirme que le problème (N) possède une unique solution faible : il existe un unique $u \in H^1(I)$ tel que

$$a(u, v) = \int_I f v \quad \forall v \in H^1(I).$$

Reste à montrer que u est une solution classique.

Lemme 2.25 *Si u est la solution faible de (N), alors u' appartient à $H_0^1(I)$.*

Preuve * : Le fait que $u' \in H^1(I)$ et que $(u')' = f - cu$ se montre comme précédemment. Nous montrons maintenant que $u'(a) = u'(b) = 0$. On choisit comme fonction test une fonction qui vaut 0 en b et 1 en a (par exemple $v(x) = (b-x)/(b-a)$) : alors par définition de la solution faible, puis intégrations par parties

$$\begin{aligned} \int_I f(x)v(x)dx &= \int_I (u'(x)v'(x) + c(x)u(x)v(x)) dx \\ &= [u'(x)v(x)]_a^b - \int_I ((f(x) - c(x)u(x))v'(x) + c(x)u(x)v(x)) dx \end{aligned}$$

Le choix de v implique que $u'(a) = 0$. On fait de même avec une fonction test v qui vaut 1 en b et 0 en a pour obtenir $u'(b) = 0$. \square

En conclusion, on a prouvé :

Théorème 2.26 *Le problème (P) possède une unique solution classique.*

2.4 Espaces de Sobolev en dimension supérieure - formulation faible

Les espaces de Sobolev en dimension supérieure se définissent de façon similaire à ce que nous avons fait en dimension un. Cependant, nombre de propriétés ne se conservent pas : notamment il n'est plus vrai que les fonctions de $W^{1,p}$ soient continues, ni même bornées, pour toutes les valeurs de p .

2.4.1 Définitions

Soit Ω est un ouvert non vide de \mathbb{R}^N . Pour définir $W^{1,p}(\Omega)$, on part de la formule d'intégration par parties : si $u \in C^1(\Omega)$, $v \in C_c^1(\Omega)$, alors, pour tout $i = 1, \dots, N$, on a

$$\int_{\Omega} u(x) \frac{\partial v}{\partial x_i}(x) dx = - \int_{\Omega} \frac{\partial u}{\partial x_i}(x) v(x) dx$$

C'est une conséquence directe de la formule d'intégration par parties dans \mathbb{R} et du théorème de Fubini.

Définition 2.27 *Soit $p \in [1, +\infty]$. On dit que $u \in W^{1,p}(\Omega)$ si $u \in L^p(\Omega)$ et s'il existe $g_1, \dots, g_N \in L^p(\Omega)$ tels que*

$$\int_{\Omega} u(x) \frac{\partial v}{\partial x_i}(x) dx = - \int_{\Omega} g_i(x) v(x) dx \quad \forall v \in C_c^1(\Omega).$$

Les fonctions g_i sont définies de façon unique et s'appellent les dérivées aux sens des distributions de u . On utilise par abus de notation l'écriture standard $\frac{\partial u}{\partial x_i} = g_i$.

Notons que, si Ω est borné, les fonctions de classe C^1 dans $\bar{\Omega}$ appartiennent à tous les espaces $W^{1,p}(\Omega)$. Si Ω n'est pas borné, c'est le cas de la restriction à Ω de toutes les fonctions de classe C^1 et à support compact dans \mathbb{R}^N . L'unicité des dérivées faibles se montre comme en dimension 1.

Contrairement au cas de la dimension 1 d'espace, les fonctions de l'espace de Sobolev $u \in W^{1,p}(\Omega)$ ne sont pas continues en général (sauf si $p > N$, où N est la dimension de l'espace ambiant : c'est l'inégalité de Morrey (cf. Brézis)).

Par exemple, supposons que Ω est la boule unité de \mathbb{R}^N et $u(x) = \|x\|^{-\alpha}$ (avec $\alpha > 0$). Notons que $u \in L^p(\Omega)$ dès que $N - 1 > \alpha p$ et, comme $\frac{\partial u}{\partial x_i}(x) = -\alpha \|x\|^{-\alpha-2} x_i$ (au sens usuel) pour $x \neq 0$, on a

$\frac{\partial u}{\partial x_i} \in L^p(\Omega)$ lorsque $N - 1 > (\alpha + 1)p$. On montre aisément que $-\alpha\|x\|^{-\alpha-2}x_1, \dots, -\alpha\|x\|^{-\alpha-2}x_N$ sont les dérivées au sens distribution de u (il suffit d'approcher u par la suite de fonctions $u_n(x) := (\|x\|^2 + 1/n)^{-\alpha/2}$ qui sont de classe C^∞ et de passer à la limite dans la formule d'intégration par parties pour u_n). Si $p = 1$ et $N \geq 3$ par exemple, on peut prendre $\alpha = N - 5/2$ et la fonction (discontinue, non bornée) u appartient à $W^{1,1}(\Omega)$. En dimension $N = 2$, on montre de même que la fonction $u(x) = \ln(\|x\|)$ est dans $W^{1,p}(\Omega)$ si $p \in [1, 2[$.

Proposition 2.28 *L'espace $W^{1,p}(\Omega)$ est un espace de Banach lorsqu'on le muni de la norme (pour $p \neq 2$)*

$$\|u\|_{W^{1,p}} = \|u\|_{L^p} + \sum_{i=1}^N \left\| \frac{\partial u}{\partial x_i} \right\|_{L^p}$$

Lorsque $p = 2$, on pose $H^1(\Omega) := W^{1,2}(\Omega)$ et on choisit plutôt

$$\|u\|_{H^1} = \left(\|u\|_{L^2}^2 + \sum_{i=1}^N \left\| \frac{\partial u}{\partial x_i} \right\|_{L^2}^2 \right)^{1/2}$$

L'espace $H^1(\Omega)$ est un espace de Hilbert.

La preuve est identique à celle de la dimension 1.

Le fait que les fonctions de $u \in W^{1,p}(\Omega)$ n'ont pas de représentant continu implique qu'on ne peut pas définir de valeur ponctuelle à $u(x)$ pour tout x . En particulier, l'espace $W_0^{1,p}(\Omega)$ (les fonctions de $W^{1,p}(\Omega)$ qui, au moins heuristiquement, sont nulles sur le bord de Ω) doit être défini de façon indirecte. Nous avons vu en dimension 1 que $W_0^{1,p}(]a, b[)$ est l'adhérence des fonctions $C_c^\infty(]a, b[)$ pour la norme $W^{1,p}(]a, b[)$: nous utilisons cette idée pour définir $W_0^{1,p}(\Omega)$.

Définition 2.29 $W_0^{1,p}(\Omega)$ est l'adhérence, pour la norme $W^{1,p}$, de l'espace $C_c^\infty(\Omega)$ (ou, de façon équivalente, de $C_c^1(\Omega)$). Pour $p = 2$, on pose $H_0^1(\Omega) := W_0^{1,2}(\Omega)$.

Par densité, $W_0^{1,p}(\Omega)$ est aussi l'adhérence, pour la norme $W^{1,p}$, de l'espace $C_c^1(\Omega)$. On interprète $W_0^{1,p}(\Omega)$ comme l'ensemble des fonctions de $W^{1,p}(\Omega)$ qui sont nulles sur le bord de Ω . Par exemple, si Ω est un ouvert borné et si u est une fonction constante, alors u appartient à $W_0^{1,p}(\Omega)$, si et seulement si, u est identiquement nulle (ce qui est rassurant, à défaut d'être convaincant). En effet, si (u_n) est une suite de fonctions de $C_c^\infty(\Omega)$ qui converge vers u pour la norme $W^{1,p}$, par intégration par parties classique, et en prenant $v_i(x) = x_i$,

$$\int_{\Omega} u_n(x) dx = \int_{\Omega} u_n(x) \frac{\partial v_i}{\partial x_i}(x) dx = - \int_{\Omega} \frac{\partial u_n}{\partial x_i}(x) v_i(x) dx$$

Or $u_n \rightarrow u$ dans L^p (et donc dans L^1 puisque Ω est borné) et, de même, $\frac{\partial u_n}{\partial x_i}$ tend vers $0 = \frac{\partial u}{\partial x_i}$ dans L^p et donc dans L^1 . Comme v_i est bornée sur Ω (toujours parce que Ω est borné), on a finalement $\int_{\Omega} u(x) dx = 0$. Comme u est constante, cela implique que u est nulle.

Remarque 2.30 Comme, par définition, $W_0^{1,p}(\Omega)$ est fermé dans $W^{1,p}(\Omega)$ et que $W^{1,p}(\Omega)$ est complet, $W_0^{1,p}(\Omega)$ est lui-même un espace de Banach, tandis que $H_0^1(\Omega)$ est un espace de Hilbert.

2.4.2 Exemples de formulation faible

Une des applications les plus simples est la résolution d'équations de la forme

$$(E) \quad -\Delta u + u = f \quad \text{dans } \Omega, \quad u = 0 \text{ dans } \partial\Omega$$

où Ω est un ouvert borné de \mathbb{R}^N , $\Delta u = \sum_{i=1}^N \frac{\partial^2 u}{\partial x_i^2}$ (le laplacien) et $f \in L^2(\Omega)$.

Afin de tenir compte du fait qu'on cherche une fonction u qui soit nulle au bord de Ω , il est naturel de travailler dans l'espace $H_0^1(\Omega)$. Comme en dimension 1, on cherche une formulation faible : on multiplie l'équation par une fonction $v \in \mathcal{C}_c^\infty(\Omega)$ et on intègre pour obtenir

$$\int_{\Omega} v(x) \left(-\sum_{i=1}^N \frac{\partial^2 u}{\partial x_i^2}(x) \right) + v(x)u(x) \, dx = \int_{\Omega} f(x)v(x) \, dx .$$

Par intégration par parties, on a

$$\int_{\Omega} v(x) \left(-\sum_{i=1}^N \frac{\partial^2 u}{\partial x_i^2}(x) \right) \, dx = -\sum_{i=1}^N \int_{\Omega} v(x) \frac{\partial^2 u}{\partial x_i^2}(x) \, dx = \sum_{i=1}^N \int_{\Omega} \frac{\partial u}{\partial x_i}(x) \frac{\partial v}{\partial x_i}(x) \, dx$$

D'où

$$\sum_{i=1}^N \int_{\Omega} \frac{\partial u}{\partial x_i}(x) \frac{\partial v}{\partial x_i}(x) \, dx + \int_{\Omega} v(x)u(x) \, dx = \int_{\Omega} f(x)v(x) \, dx \quad \forall v \in \mathcal{C}_c^\infty(\Omega) .$$

Comme cette relation a encore un sens si $v \in H_0^1(\Omega)$, on introduit la notion de la formulation faible :

Définition 2.31 *On dit que u est une solution faible de l'équation (E) si $u \in H_0^1(\Omega)$ et si*

$$\sum_{i=1}^N \int_{\Omega} \frac{\partial u}{\partial x_i}(x) \frac{\partial v}{\partial x_i}(x) \, dx + \int_{\Omega} v(x)u(x) \, dx = \int_{\Omega} f(x)v(x) \, dx \quad \forall v \in H_0^1(\Omega) .$$

Le résultat suivant est alors immédiat.

Théorème 2.32 *Il existe une unique solution faible à l'équation (E).*

Preuve* : La formulation faible se réécrit plus simplement comme

$$(Ef) \quad \langle u, v \rangle_{H^1} = \int_{\Omega} f(x)v(x) \, dx \quad \forall v \in H_0^1(\Omega) .$$

Comme l'application $v \rightarrow \int_{\Omega} f v$ est linéaire continue sur $H_0^1(\Omega)$ (par une application directe de Cauchy-Schwarz), le théorème de représentation de Riesz affirme qu'il existe un unique élément $u \in H_0^1(\Omega)$ vérifiant (Ef). Il existe donc une unique solution faible de l'équation. \square

3 Distributions tempérées et transformée de Fourier

Les distributions sont des objets mathématiques qui sont définis par leur action sur des fonctions très régulières : les fonctions de la classe de Schwartz. On pourra alors transposer les actions qui l'on fait usuellement sur ces fonctions régulières aux distributions “par dualité”.

3.1 Classe de Schwartz

On s'intéresse dans cette partie aux fonctions très régulières sur \mathbb{R}^N et qui tendent vers 0 à l'infini plus rapidement que n'importe quelle puissance de $1/\|x\|$. Le prototype de ces fonctions dans \mathbb{R} (i.e., en dimension $N = 1$) est la fonction $x \rightarrow e^{-\frac{x^2}{2}}$. Dans \mathbb{R}^N , c'est la fonction $x \rightarrow e^{-\|x\|^2/2}$.

Nous aurons besoin de la notation suivante : soit $\alpha = (\alpha_1, \dots, \alpha_N) \in \mathbb{N}^N$ un multi-indice et $u \in C^\infty(\mathbb{R}^N)$, on pose

$$\partial^\alpha u(x) = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \dots \partial x_N^{\alpha_N}}(x)$$

On notera $|\alpha| = \alpha_1 + \dots + \alpha_N$ l'ordre de dérivation. Nous parlerons aussi de polynômes de plusieurs variables, en définissant le monôme $x^\alpha = x_1^{\alpha_1} \dots x_N^{\alpha_N}$: un polynôme sur \mathbb{R}^N sera juste une combinaison linéaire finie de tels monômes.

Définition 3.1 On dit que ϕ appartient à la classe de Schwartz $\mathcal{S}(\mathbb{R}^N)$ si $\phi : \mathbb{R}^N \rightarrow \mathbb{C}$ est de classe C^∞ sur \mathbb{R}^N et si, pour tout entiers naturel n et k ,

$$N_{n,k}(\phi) := \sup_{x \in \mathbb{R}^N, \alpha \in \mathbb{N}^N, |\alpha|=k} ((1 + \|x\|^2)^n |\partial^\alpha \phi(x)|) < +\infty$$

On dit qu'une suite de fonctions (ϕ_p) de la classe de Schwartz converge vers $\phi \in \mathcal{S}(\mathbb{R}^N)$ si

$$\lim_{p \rightarrow +\infty} N_{n,k}(\phi - \phi_p) = 0 \quad \forall n, k \in \mathbb{N}.$$

Remarques :

1. Par exemple, les fonctions $C_c^\infty(\mathbb{R}^N)$ appartiennent à $\mathcal{S}(\mathbb{R}^N)$. Notons que le seul polynôme de $\mathcal{S}(\mathbb{R}^N)$ est le polynôme nul.
2. Nous utiliserons à de multiples reprises le fait que les $N_{k,n}$ sont positivement homogènes et vérifient l'inégalité triangulaire :

$$N_{n,k}(\lambda\phi) = |\lambda| N_{n,k}(\phi) \text{ et } N_{n,k}(\phi + \psi) \leq N_{n,k}(\phi) + N_{n,k}(\psi) \quad \forall \phi, \psi \in \mathcal{S}(\mathbb{R}^N), \forall \lambda \in \mathbb{C}.$$

En particulier $\mathcal{S}(\mathbb{R}^N)$ est un espace vectoriel sur \mathbb{C} .

3. Même si les $N_{n,k}$ sont en fait des normes sur $\mathcal{S}(\mathbb{R}^N)$ (car si $N_{n,k}(\phi) = 0$ alors ϕ est un polynôme de degré au plus k ; or le seul polynôme de $\mathcal{S}(\mathbb{R}^N)$ est le polynôme nul, donc $\phi = 0$), la notion de convergence dans $\mathcal{S}(\mathbb{R}^N)$ introduite plus haut ne se réduit pas à la convergence dans une seule de ces normes. Bien garder en tête que $\mathcal{S}(\mathbb{R}^N)$ n'est pas un espace vectoriel normé.

Proposition 3.2 On a $\mathcal{S}(\mathbb{R}^N) \subset L^p(\mathbb{R}^N)$ pour tout $p \in [1, +\infty]$. Plus précisément, pour tout $p \in [1, +\infty]$ et pour tout $n > N/(2p)$, il existe une constante $C_{n,p} \geq 0$ telle que

$$\|\phi\|_p \leq C_{n,p} N_{n,0}(\phi) \quad \forall \phi \in \mathcal{S}(\mathbb{R}^N).$$

Preuve (*) : En effet, soit $p \in [1, +\infty[$ et $\phi \in \mathcal{S}(\mathbb{R}^N)$. Pour $n > N/(2p)$, on a $N_{n,0}(\phi) := \sup_{x \in \mathbb{R}^N} (\|x\|^2 + 1)^n |\phi(x)| < +\infty$. En particulier,

$$\int_{\mathbb{R}^N} |\phi(x)|^p dx \leq (N_{n,0}(\phi))^p \int_{\mathbb{R}^N} \frac{dx}{(\|x\|^2 + 1)^{np}} = (N_{n,0}(\phi))^p C_N \int_0^{+\infty} \frac{r^{N-1}}{(r^2 + 1)^{np}} dr < +\infty$$

puisque $N - 1 - 2np < -1$ (on utilise ici que $n > N/(2p)$). Pour $p = +\infty$, notons que, par définition de $N_{n,0}$, on a $|\phi(x)| \leq N_{n,0}(\phi)$ pour tout $x \in \mathbb{R}^N$ et pour tout $n \in \mathbb{N}$. Donc $\phi \in L^\infty(\mathbb{R}^N)$. \square

Une des propriétés importantes de la classe de Schwartz est la stabilité par rapport à la dérivation et la multiplication par un polynôme :

Proposition 3.3 1. Si $\phi \in \mathcal{S}(\mathbb{R}^N)$, alors pour tout $\alpha \in \mathbb{N}^N$, $\partial^\alpha \phi \in \mathcal{S}(\mathbb{R}^N)$. De plus, si (ϕ_p) tend vers ϕ dans $\mathcal{S}(\mathbb{R}^N)$, alors $\partial^\alpha \phi_p$ tend vers $\partial^\alpha \phi$ dans $\mathcal{S}(\mathbb{R}^N)$.

2. Si P est un polynôme sur \mathbb{R}^N et $\phi \in \mathcal{S}(\mathbb{R}^N)$, alors l'application $x \rightarrow (P\phi)(x) := P(x)\phi(x)$ appartient à $\mathcal{S}(\mathbb{R}^N)$. De plus, si (ϕ_n) tend vers ϕ dans $\mathcal{S}(\mathbb{R}^N)$, alors $P\phi_n$ tend vers $P\phi$ dans $\mathcal{S}(\mathbb{R}^N)$.

Preuve:

1. (*) Notons que, pour tout n, k ,

$$N_{n,k}(\partial^\alpha \phi) := \sup_{x \in \mathbb{R}^N, |\alpha'| \leq k} \left((1 + \|x\|^2)^n |\partial^{\alpha'}(\partial^\alpha \phi)(x)| \right) \leq N_{n,k+|\alpha|}(\phi) < +\infty$$

De plus, si (ϕ_p) tend vers ϕ dans $\mathcal{S}(\mathbb{R}^N)$, alors, pour tout n, k ,

$$N_{n,k}(\partial^\alpha \phi_p - \partial^\alpha \phi) \leq N_{n,k+|\alpha|}(\phi - \phi_p) \rightarrow 0.$$

Donc $(\partial^\alpha \phi_p)$ tend vers $\partial^\alpha \phi$ dans $\mathcal{S}(\mathbb{R}^N)$.

2. Pour simplifier l'exposé, on ne fait la preuve qu'en dimension $N = 1$. Comme $\mathcal{S}(\mathbb{R})$ est un espace vectoriel, il suffit de montrer le résultat pour un monôme de la forme $P(x) = x^\alpha$, $\alpha \in \mathbb{N}^*$. Par formule de Leibnitz, on a, pour tout k ,

$$(P\phi)^{(k)} = \sum_{r=0}^k C_k^r \phi^{(r)} P^{(k-r)}$$

où $P^{(k-r)}(x) = \frac{\alpha!}{(\alpha - k + r)!} x^{\alpha - k + r}$ si $k - r \leq \alpha$, 0 sinon. Donc, pour tout n, k ,

$$\begin{aligned} N_{n,k}(P\phi) &\leq \sum_{r=\max\{0, k-\alpha\}}^k C_k^r \frac{\alpha!}{(\alpha - k + r)!} \sup_{x \in \mathbb{R}} \left((1 + x^2)^n |x|^{\alpha - k + r} |\phi^{(r)}(x)| \right) \\ &\leq \sum_{r=\max\{0, k-\alpha\}}^k C_k^r \frac{\alpha!}{(\alpha - k + r)!} N_{r, n + \alpha - k + r}(\phi) < +\infty \end{aligned}$$

De même, si (ϕ_p) dans vers ϕ dans $\mathcal{S}(\mathbb{R})$, alors, quand $p \rightarrow +\infty$,

$$N_{n,k}(P\phi - P\phi_p) \leq \sum_{r=\max\{0, k-\alpha\}}^k C_k^r \frac{\alpha!}{(\alpha - k + r)!} N_{r, n + \alpha - k + r}(\phi - \phi_p) \rightarrow 0.$$

Donc $(P\phi_p)$ tend vers $P\phi$ dans $\mathcal{S}(\mathbb{R})$.

□

Une autre propriété importante de la classe de Schwartz est la stabilité par transformée de Fourier : rappelons que, si $\phi \in L^1(\mathbb{R}^N)$, alors la transformée de Fourier de ϕ est la fonction continue et bornée

$$\mathcal{F}(\phi)(x) = \int_{\mathbb{R}^N} e^{i\langle x, y \rangle} \phi(y) dy \quad \forall x \in \mathbb{R}^N,$$

tandis que la transformée de Fourier inverse de ϕ est donnée par

$$\overline{\mathcal{F}}(\phi)(x) = \frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} e^{-i\langle x, y \rangle} \phi(y) dy \quad \forall x \in \mathbb{R}^N.$$

Notons $\overline{\mathcal{F}}(f)(x) = \frac{1}{(2\pi)^N} \mathcal{F}(f)(-x) = \frac{1}{(2\pi)^N} \mathcal{F}(f^\sharp)(x)$ pour tout $x \in \mathbb{R}^N$, où $f^\sharp(x) = f(-x)$. Dans ce qui suit, on utilise les notations suivantes : $(-ix)^\alpha := (-i)^{|\alpha|} x^\alpha$ et $(ix)^\alpha := i^{|\alpha|} x^\alpha$.

Théorème 3.4 *La transformée de Fourier \mathcal{F} est une application linéaire continue bijective de $\mathcal{S}(\mathbb{R}^N)$ dans lui-même, et a pour inverse $\overline{\mathcal{F}}$:*

$$\overline{\mathcal{F}}\mathcal{F}(\phi) = \mathcal{F}\overline{\mathcal{F}}(\phi) = \phi \quad \forall \phi \in \mathcal{S}(\mathbb{R}^N).$$

De plus, pour tout multi-indice $\alpha \in \mathbb{N}^N$, on a

$$\mathcal{F}(\partial^\alpha \phi)(x) = (-ix)^\alpha \mathcal{F}(\phi)(x) \quad \forall x \in \mathbb{R}^N \quad (8)$$

tandis que

$$\mathcal{F}((ix)^\alpha \phi)(x) = \partial^\alpha \mathcal{F}(\phi)(x) \quad \forall x \in \mathbb{R}^N. \quad (9)$$

Enfin, on peut remplacer \mathcal{F} par $\overline{\mathcal{F}}$ dans les énoncés ci-dessus.

Preuve : A nouveau, pour simplifier les notations, on travaille en dimension $N = 1$. On rappelle que \mathcal{F} est une bijection de $L^2(\mathbb{R})$ dans lui-même, d'inverse $\overline{\mathcal{F}}$. D'autre part, il est bien connu que, si $f \in \mathcal{C}^1(\mathbb{R})$ est tel que $f \in L^1(\mathbb{R})$ et $f' \in L^1(\mathbb{R})$, alors $\mathcal{F}(f')(x) = (-ix)\mathcal{F}(f)(x)$ pour tout $x \in \mathbb{R}$. Comme $\mathcal{F}(f')$ est bornée, cela prouve que $(1 + |x|)\mathcal{F}(f)$ est bornée. Inversement, si $(1 + |x|)f \in L^1(\mathbb{R})$, alors $\mathcal{F}(f) \in \mathcal{C}^1(\mathbb{R})$ avec $\mathcal{F}(f)' = \mathcal{F}((ix)f)$.

Lorsque $\phi \in \mathcal{S}(\mathbb{R})$, alors on montre par récurrence que, comme ϕ est de classe \mathcal{C}^∞ et que toutes ses dérivées sont dans $L^1(\mathbb{R})$, la fonction $(1 + |x|)^n \mathcal{F}(\phi)$ est bornée pour tout $n \in \mathbb{N}$. D'autre part, comme $(1 + x^2)^n \phi$ est borné pour tout n , la fonction $\mathcal{F}(\phi)$ est de classe \mathcal{C}^∞ . On montre également par récurrence les relations (8) et (9). Quant à la preuve, un peu plus calculatoire, de la continuité de \mathcal{F} , elle vient de la définition de la convergence dans $\mathcal{S}(\mathbb{R}^N)$ combinée à (8) et de (9). □

Proposition 3.5 *Si $f, g \in \mathcal{S}(\mathbb{R}^N)$, alors*

$$\int_{\mathbb{R}^N} \mathcal{F}(f)g = \int_{\mathbb{R}^N} f\mathcal{F}(g)$$

Preuve (*) : C'est une conséquence directe du théorème de Fubini : comme $(x, t) \rightarrow |f(t)e^{i\langle t, x \rangle} g(x)|$ est dans $L^1(\mathbb{R}^{2N})$ (car f et g sont dans $L^1(\mathbb{R}^N)$), on a

$$\int_{\mathbb{R}^N} \mathcal{F}(f)g = \int_{\mathbb{R}^N} \left(\int_{\mathbb{R}^N} f(t)e^{i\langle t, x \rangle} dt \right) g(x) dx = \int_{\mathbb{R}^N} f(t) \left(\int_{\mathbb{R}^N} e^{i\langle t, x \rangle} g(x) dx \right) dt = \int_{\mathbb{R}^N} f\mathcal{F}(g).$$

□

Proposition 3.6 (Stabilité par convolution) *L'application qui, à tout couple $(\phi, \psi) \in \mathcal{S}(\mathbb{R}^N) \times \mathcal{S}(\mathbb{R}^N)$ associe $\phi \star \psi$ est bilinéaire et continue de $\mathcal{S}(\mathbb{R}^N) \times \mathcal{S}(\mathbb{R}^N)$ dans $\mathcal{S}(\mathbb{R}^N)$. De plus, si $\phi, \psi \in \mathcal{S}(\mathbb{R}^N)$, alors*

$$\mathcal{F}(\phi \star \psi) = \mathcal{F}(\phi)\mathcal{F}(\psi), \quad \mathcal{F}(\phi\psi) = (2\pi)^{-N}\mathcal{F}(\phi) \star \mathcal{F}(\psi)$$

Preuve : Montrons d'abord les deux dernières égalités : on a, pour tout $t \in \mathbb{R}^N$ et en utilisant le théorème de Fubini

$$\begin{aligned} \mathcal{F}(\phi \star \psi)(t) &= \int_{\mathbb{R}^{2N}} \phi(x-y)\psi(y)e^{i\langle t, x \rangle} dydx = \int_{\mathbb{R}^N} \psi(y)e^{i\langle t, y \rangle} \left(\int_{\mathbb{R}^N} \phi(x-y)e^{i\langle t, x-y \rangle} dx \right) dy \\ &= \int_{\mathbb{R}^N} \psi(y)e^{i\langle t, y \rangle} \left(\int_{\mathbb{R}^N} \phi(z)e^{i\langle t, z \rangle} dz \right) dy = \int_{\mathbb{R}^N} \psi(y)e^{i\langle t, y \rangle} \mathcal{F}(\phi)(t) dy \\ &= \mathcal{F}(\phi)(t)\mathcal{F}(\psi)(t). \end{aligned}$$

(où on a effectué le changement de variable $z = x - y$ dans la troisième égalité).

Lorsque l'on fait exactement le même calcul avec $\overline{\mathcal{F}}$ à la place de \mathcal{F} , on obtient :

$$\overline{\mathcal{F}}(\phi \star \psi) = (2\pi)^N \overline{\mathcal{F}}(\phi)\overline{\mathcal{F}}(\psi)$$

Pour montrer la seconde égalité de la proposition, prenons l'image $\overline{\mathcal{F}}$ de l'égalité $\mathcal{F}(\phi \star \psi) = \mathcal{F}(\phi)\mathcal{F}(\psi)$. On a, d'après la remarque précédente (utilisée pour $\mathcal{F}(\phi)$ et $\mathcal{F}(\psi)$ à la place de ϕ et ψ) :

$$\overline{\mathcal{F}}((2\pi)^{-N}\mathcal{F}(\phi) \star \mathcal{F}(\psi)) = \overline{\mathcal{F}}\mathcal{F}(\phi) \overline{\mathcal{F}}\mathcal{F}(\psi) = \phi\psi$$

On reprend l'image par \mathcal{F} de l'égalité précédente pour obtenir $(2\pi)^{-N}\mathcal{F}(\phi) \star \mathcal{F}(\psi) = \mathcal{F}(\phi\psi)$.

Montrons finalement que $\phi \star \psi \in \mathcal{S}(\mathbb{R}^N)$. Comme $\phi, \psi \in \mathcal{S}(\mathbb{R}^N)$, on a aussi $\mathcal{F}(\phi), \mathcal{F}(\psi) \in \mathcal{S}(\mathbb{R}^N)$. Donc $\mathcal{F}(\phi)\mathcal{F}(\psi) \in \mathcal{S}(\mathbb{R}^N)$ (ceci se montre facilement en utilisant la formule de Leibnitz). Comme $\mathcal{F}(\phi \star \psi) = \mathcal{F}(\phi)\mathcal{F}(\psi)$, on a $\phi \star \psi = \overline{\mathcal{F}}\mathcal{F}(\phi \star \psi) \in \mathcal{S}(\mathbb{R}^N)$, puisque $\mathcal{S}(\mathbb{R}^N)$ est stable pour $\overline{\mathcal{F}}$. La continuité se prouve de même. \square

3.2 Distributions tempérées

Définition 3.7 *Une distribution tempérée est une application linéaire continue de $\mathcal{S}(\mathbb{R}^N)$ dans \mathbb{C} . L'espace des distributions tempérées est noté (fort logiquement) $\mathcal{S}'(\mathbb{R}^N)$.*

Remarques :

1. L'ensemble $\mathcal{S}'(\mathbb{R}^N)$ des distributions tempérées forme un espace vectoriel sur \mathbb{C} .
2. On peut démontrer (mais nous ne le ferons pas) que, si T est une forme linéaire sur $\mathcal{S}(\mathbb{R}^N)$, alors $T \in \mathcal{S}'(\mathbb{R}^N)$ (i.e., T est continue), si et seulement si, il existe deux entiers m et n et une constante $C > 0$ tels que

$$|T(\phi)| \leq C \sum_{k \leq m} N_{n,k}(\phi) \quad \forall \phi \in \mathcal{S}(\mathbb{R}^N).$$

Dans ce cas, on dit que T est **d'ordre** m si m est le plus petit indice pour lequel l'inégalité ci-dessus a lieu.

3. Si $T \in \mathcal{S}'(\mathbb{R}^N)$, le **support de** T est le plus petit fermé $K \subset \mathbb{R}^N$ tel que

$$\forall \phi \in \mathcal{S}(\mathbb{R}^N) \cap C_c^\infty(\mathbb{R}^N \setminus K), \quad T(\phi) = 0.$$

Exemples

1. Si $f \in L^p(\mathbb{R}^N)$, avec $p \in [1, +\infty]$, alors

$$T_f(\phi) = \int_{\mathbb{R}^N} f(x)\phi(x)dx \quad \forall \phi \in \mathcal{S}(\mathbb{R}^N)$$

définit une distribution tempérée d'ordre 0 (grâce à la proposition 3.2 et l'inégalité de Hölder). Son support est celui de f .

2. **La masse de Dirac :** L'application

$$T(\phi) = \phi(0) \quad \forall \phi \in \mathcal{S}(\mathbb{R}^N)$$

définit aussi une distribution tempérée d'ordre 0. Notons qu'on peut l'identifier à la masse de Dirac δ_0 en 0 :

$$T(\phi) = \int_{\mathbb{R}^N} \phi(x)d\delta_0(x)$$

Par abus de langage, nous dirons que $T = \delta_0$ et verrons δ_0 comme une distribution tempérée particulière.

3. **Valeur principale :** On suppose que $N = 1$. Pour tout $\phi \in \mathcal{S}(\mathbb{R})$, la limite

$$T(\phi) = \lim_{\epsilon \rightarrow 0^+} \int_{\mathbb{R} \setminus]-\epsilon, \epsilon[} \frac{\phi(x)}{x} dx$$

existe. Cette limite définit une distribution tempérée d'ordre 1, habituellement notée vp .

Preuve : La fonction $x \rightarrow \frac{\phi(x)}{x}$ est clairement intégrable sur $\mathbb{R} \setminus]-\epsilon, \epsilon[$. Notons que l'application $x \rightarrow \frac{\phi(x) - \phi(0)}{x}$ est continue en 0, avec pour limite $\phi'(0)$ lorsque $x \rightarrow 0$. Donc, pour $\epsilon \in]0, 1[$,

$$\int_{[-1,1] \setminus]-\epsilon, \epsilon[} \frac{\phi(x)}{x} dx = \int_{[-1,1] \setminus]-\epsilon, \epsilon[} \frac{\phi(x) - \phi(0)}{x} dx + \int_{[-1,1] \setminus]-\epsilon, \epsilon[} \frac{\phi(0)}{x} dx$$

où le premier terme à droite possède une limite lorsque $\epsilon \rightarrow 0^+$ tandis que le second est nul car $1/x$ est impaire. En particulier,

$$\left| \int_{[-1,1] \setminus]-\epsilon, \epsilon[} \frac{\phi(x)}{x} dx \right| \leq \|\phi'\|_\infty .$$

tandis que

$$\left| \int_{\mathbb{R} \setminus [-1,1]} \frac{\phi(x)}{x} dx \right| \leq \|\phi\|_{L^1(\mathbb{R})}$$

Ceci montre que T est continue et d'ordre au plus 1. Pour montrer que T est d'ordre 1, supposons au contraire que T soit d'ordre 0. Alors il existe $n \in \mathbb{N}$ et $C > 0$ tels que

$$|T(\phi)| \leq C \sup_{x \in \mathbb{R}} |(1+x^2)^n \phi(x)| \quad \forall \phi \in \mathcal{S}(\mathbb{R}^N) .$$

Pour $p \in \mathbb{N}^* > 0$, on définit ϕ_p comme étant la fonction de $\mathcal{S}(\mathbb{R})$ donnée par $\phi_p(x) = \text{th}(px)\psi(x)$ où ψ est une fonction $C_c^\infty(\mathbb{R})$, qui vaut 1 sur $[-1,1]$, 0 en dehors de $[-2,2]$, positive sur \mathbb{R} . Alors ϕ_p est positive sur \mathbb{R}^+ et négative sur \mathbb{R}^- et on a, puisque $|\text{th}(x)| \leq 1$ sur \mathbb{R} ,

$$|T(\phi_p)| \leq C \sup_{x \in \mathbb{R}} |(1+x^2)^n \phi_p(x)| \leq C'$$

où C' est une constante indépendante de p . Alors, comme $\phi_p(x)/x$ est positif sur \mathbb{R}^* , on a, par Fatou,

$$C' \geq \liminf_{p \rightarrow +\infty} T(\phi_p) \geq \liminf_{p \rightarrow +\infty} \int_0^1 \frac{\text{th}(px)}{x} dx \geq \int_0^1 \frac{dx}{x} = +\infty,$$

ce qui est impossible. \square

Proposition 3.8 Soit P un polynôme de \mathbb{R}^N et $T \in \mathcal{S}'(\mathbb{R}^N)$. Alors l'application PT définie par

$$(PT)(\phi) = T(P\phi) \quad \forall \phi \in \mathcal{S}(\mathbb{R}^N)$$

est une distribution tempérée.

Preuve (*) : Nous avons vu ci-dessus que $P\phi \in \mathcal{S}(\mathbb{R}^N)$ pour tout $\phi \in \mathcal{S}(\mathbb{R}^N)$. De plus l'application $\phi \rightarrow P\phi$ est continue de $\mathcal{S}(\mathbb{R}^N)$ dans lui-même. Ceci prouve que PT est continue de $\mathcal{S}(\mathbb{R}^N)$ dans \mathbb{R} . \square

Proposition 3.9 (Dérivée d'une distribution tempérée) Si $T \in \mathcal{S}'(\mathbb{R}^N)$ et $\alpha \in \mathbb{N}^N$ un multi-indice. La distribution tempérée $\partial^\alpha T$ est définie par

$$\partial^\alpha T(\phi) = (-1)^{|\alpha|} T(\partial^\alpha \phi) \quad \forall \phi \in \mathcal{S}(\mathbb{R}^N)$$

Preuve (*) : La preuve est identique à la démonstration de la proposition précédente. \square

Exemple : Si $N = 1$ et $T = \delta_0$, alors

$$T'(\phi) = -\phi'(0) \quad \forall \phi \in \mathcal{S}(\mathbb{R})$$

Définition 3.10 (Transformée de Fourier d'une distribution) Pour tout $T \in \mathcal{S}'(\mathbb{R}^N)$, l'application $\mathcal{F}T$ définie par

$$\mathcal{F}T(\phi) = T(\mathcal{F}(\phi)) \quad \forall \phi \in \mathcal{S}(\mathbb{R}^N)$$

est une distribution tempérée. De plus, pour tout multi-indice $\alpha \in \mathbb{N}^N$,

$$\mathcal{F}(\partial^\alpha T) = (-ix)^\alpha \mathcal{F}T \quad \text{et} \quad \mathcal{F}((ix)^\alpha T) = (\mathcal{F}T)^{(\alpha)}$$

Preuve (*) : C'est une conséquence directe des propriétés des fonctions de la classe de Schwartz. \square

Exemples : On suppose ici que $N = 1$ pour fixer les idées.

1. La distribution $\mathcal{F}\delta_0$ est donnée par définition par

$$\mathcal{F}(\delta_0)(\phi) = \mathcal{F}(\phi)(0) = \int_{\mathbb{R}} \phi(x) dx = T_1(\phi).$$

Donc on peut assimiler la distribution $\mathcal{F}(\delta_0)$ à la fonction constante égale à 1 : $\mathcal{F}(\delta_0) = 1$.

2. Inversement, si $T = 1$, i.e., T est la distribution

$$T(\phi) = \int_{\mathbb{R}} \phi(x) dx \quad \forall \phi \in \mathcal{S}(\mathbb{R}),$$

alors $\mathcal{F}(T) = 2\pi\delta_0$. En effet, comme $\mathcal{F}(\delta_0) = 1$, on a $\delta_0 = \overline{\mathcal{F}}\mathcal{F}(\delta_0) = \overline{\mathcal{F}}(T)$, où

$$\begin{aligned} \overline{\mathcal{F}}(T)(\phi) &= T(\overline{\mathcal{F}}(\phi)) = \int_{\mathbb{R}} \overline{\mathcal{F}}(\phi)(x) dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathcal{F}(\phi)(-x) dx = \frac{1}{2\pi} \mathcal{F}(\phi)(x) dx = \frac{1}{2\pi} \mathcal{F}(T)(\phi) \end{aligned}$$

On peut également définir la convolution d'une distribution pondérée avec une fonction de la classe de Schwartz : pour cela, notons que, pour tout $\psi \in \mathcal{S}(\mathbb{R})$ et pour tout $x \in \mathbb{R}$, la fonction $y \rightarrow \psi(x - y)$ est également dans $\mathcal{S}(\mathbb{R})$.

Proposition 3.11 *Soit $T \in \mathcal{S}'(\mathbb{R})$ et $\psi \in \mathcal{S}(\mathbb{R})$, on pose*

$$(T \star \psi)(x) = T(\psi(x - \cdot)) \quad \forall x \in \mathbb{R} .$$

Alors nous admettrons que $T \star \psi$ est aussi la distribution tempérée donnée par

$$(T \star \psi)(\phi) = T(\psi^\sharp \star \phi) \quad \forall \phi \in \mathcal{S}(\mathbb{R}) ,$$

où $\psi^\sharp(x) = \psi(-x)$. De plus

$$\mathcal{F}(T \star \psi) = \mathcal{F}(\psi)\mathcal{F}(T) . \tag{10}$$

Exemple : Si on prend $T = \delta_0$, alors

$$(\delta_0 \star \psi)(x) = \delta_0(\psi(x - \cdot)) = \psi(x)$$

Donc $\delta_0 \star \psi = \psi$ pour tout $\psi \in \mathcal{S}(\mathbb{R}^N)$.

Preuve de l'égalité (10) (*) : On rappelle (cf. Proposition 3.6) que $\mathcal{F}(a) \star \mathcal{F}(b) = (2\pi)^N \mathcal{F}(ab)$ et $\mathcal{F}(a) = (2\pi)^N \overline{\mathcal{F}(a)}$ pour tout $a, b \in \mathcal{S}(\mathbb{R}^N)$. Donc pour tout $\phi \in \mathcal{S}(\mathbb{R}^N)$, on a

$$\begin{aligned} \mathcal{F}(T \star \psi)(\phi) &= (T \star \psi)(\mathcal{F}(\phi)) = T(\psi^\sharp \star \mathcal{F}(\phi)) = T(\mathcal{F}\overline{\mathcal{F}}(\psi^\sharp) \star \mathcal{F}(\phi)) \\ &= T\left((2\pi)^N \mathcal{F}\left(\overline{\mathcal{F}}(\psi^\sharp)\phi\right)\right) = \mathcal{F}(T)\left((2\pi)^N \overline{\mathcal{F}}(\psi^\sharp)\phi\right) = \mathcal{F}(T)(\mathcal{F}(\psi)\phi) \\ &= (\mathcal{F}(\psi)\mathcal{F}(T))(\phi) , \end{aligned}$$

ce qui prouve l'égalité $\mathcal{F}(T \star \psi) = \mathcal{F}(\psi)\mathcal{F}(T)$. □

3.3 Applications à l'équation de Laplace

Soit $f \in \mathcal{S}(\mathbb{R}^N)$. On cherche à résoudre l'équation de Laplace

$$(L) \quad -\Delta u = f \quad \text{dans } \mathbb{R}^N$$

avec $u(x) \rightarrow 0$ lorsque $\|x\| \rightarrow +\infty$.

Définition 3.12 *On appelle solution fondamentale de l'équation de Laplace une distribution tempérée $E \in \mathcal{S}'(\mathbb{R}^N)$ telle que*

$$-\Delta E = \delta_0$$

au sens des distributions.

Un des intérêts de la solution fondamentale est de permettre de résoudre (L) directement :

Proposition 3.13 *Si E est une (la) solution fondamentale de l'équation de Laplace et $f \in \mathcal{S}(\mathbb{R}^N)$, alors $u = E \star f$ vérifie $-\Delta u = f$.*

Preuve (*) : En effet, on a

$$-\Delta(E \star f) = (-\Delta E) \star f = \delta_0 \star f = f .$$

□

La relation $-\Delta E = \delta_0$ s'écrit en Fourier : $\mathcal{F}(-\Delta E) = \mathcal{F}(\delta_0) = 1$. Or

$$\mathcal{F}(\Delta E) = \sum_{k=1}^N (-ix_k)^2 \mathcal{F}(E) = -\|x\|^2 \mathcal{F}(E),$$

d'où $\|x\|^2 \mathcal{F}(E) = 1$. En particulier, si $\mathcal{F}(E) = 1/\|x\|^2$, on obtient une solution de notre problème. En dimension $N \geq 3$, la fonction $x \rightarrow 1/\|x\|^2$ est localement intégrable et tend vers 0 à l'infini. Donc elle définit une distribution T par la formule

$$T(\phi) = \int_{\mathbb{R}^N} \frac{\phi(x)}{\|x\|^2} dx \quad \forall \phi \in \mathcal{S}(\mathbb{R}^N).$$

Par inversion de Fourier, on a alors $E = \overline{\mathcal{F}}\mathcal{F}(E) = \overline{\mathcal{F}}(T)$. On peut vérifier (mais cela demande un peu de calcul) que $\overline{\mathcal{F}}(T) = c_N \|x\|^{2-N}$ où c_N est une constante qui dépend de la dimension.

Proposition 3.14 *En dimension $N = 2$, la distribution distribution tempérée définie par*

$$E(\phi) = -\frac{1}{2\pi} \int_{\mathbb{R}^2} \ln(\|x\|) \phi(x) dx$$

est une solution fondamentale de l'équation de Laplace.

Preuve : Posons $\tilde{\phi}(r, \theta) = \phi(r \cos(\theta), r \sin(\theta))$ pour $(r, \theta) \in]0, +\infty[\times]0, 2\pi[$. Alors

$$\Delta \phi = \frac{\partial^2 \tilde{\phi}}{\partial r^2} + \frac{1}{r} \frac{\partial \tilde{\phi}}{\partial r} + \frac{1}{r^2} \frac{\partial^2 \tilde{\phi}}{\partial \theta^2}$$

Comme $x \rightarrow \ln(\|x\|) \Delta \phi(x)$ est dans $L^1(\mathbb{R}^2)$, on a, par passage en coordonnées polaires,

$$\int_{\mathbb{R}^2} \ln(\|x\|) \Delta \phi(x) dx = \lim_{\epsilon \rightarrow 0^+} \int_{\|x\| > \epsilon} \ln(\|x\|) \Delta \phi(x) dx$$

où

$$\int_{\|x\| > \epsilon} \ln(\|x\|) \Delta \phi(x) dx = \int_0^{2\pi} \int_{\epsilon}^{+\infty} r \ln(r) \left(\frac{\partial^2 \tilde{\phi}}{\partial r^2} + \frac{1}{r} \frac{\partial \tilde{\phi}}{\partial r} + \frac{1}{r^2} \frac{\partial^2 \tilde{\phi}}{\partial \theta^2} \right)$$

Notons que $\int_0^{2\pi} \frac{\partial^2 \tilde{\phi}}{\partial \theta^2} d\theta = 0$ pour tout $r > 0$ et que, pour tout $\theta \in]0, 2\pi[$,

$$\begin{aligned} \int_{\epsilon}^{+\infty} r \ln(r) \frac{\partial^2 \tilde{\phi}}{\partial r^2} &= \left[r \ln(r) \frac{\partial \tilde{\phi}}{\partial r} \right]_{\epsilon}^{+\infty} - \int_{\epsilon}^{+\infty} (1 + \ln(r)) \frac{\partial \tilde{\phi}}{\partial r} dr \\ &= -\epsilon \ln(\epsilon) \frac{\partial \tilde{\phi}}{\partial r}(\epsilon, \theta) + \tilde{\phi}(\epsilon, \theta) - \int_{\epsilon}^{+\infty} \ln(r) \frac{\partial \tilde{\phi}}{\partial r} dr \end{aligned}$$

car ϕ et toutes ses dérivées sont à décroissance rapide. Donc

$$\int_{\|x\| > \epsilon} \ln(\|x\|) \Delta \phi(x) dx = \int_0^{2\pi} \left(-\epsilon \ln(\epsilon) \frac{\partial \tilde{\phi}}{\partial r}(\epsilon, \theta) + \tilde{\phi}(\epsilon, \theta) \right) d\theta \rightarrow 2\pi \phi(0) = 2\pi \delta_0(\phi)$$

lorsque $\epsilon \rightarrow 0$. Donc on a bien $-\Delta E = \delta_0$. □

A Brefs rappels d'intégration

A.1 Quelques résultats fondamentaux en intégration

Soit (X, \mathcal{A}, μ) un espace mesuré : X est un ensemble, \mathcal{A} une tribu sur X et μ une mesure sur \mathcal{A} . Un exemple typique est le triplet $(\mathbb{R}, \mathcal{B}, \lambda)$, où \mathcal{B} est la tribu borélienne de \mathbb{R} (i.e., la plus petite tribu contenant les ouverts de \mathbb{R}) et où la mesure λ est la mesure de Lebesgue de \mathbb{R} (i.e., l'unique mesure sur \mathcal{B} vérifiant $\lambda(]a, b]) = b - a$ pour tout couple de réels $a < b$).

Rappelons qu'une fonction $f : X \rightarrow \mathbb{R}$ est dite *mesurable* si

$$f^{-1}(S) \in \mathcal{A} \quad \forall S \in \mathcal{B}^1$$

ou, de façon équivalente,

$$\forall a \in \mathbb{R}, \quad f^{-1}(]-\infty, a]) \in \mathcal{A}.$$

La mesurabilité est une propriété stable par addition, par multiplication, par passage au sup, et par limite simple : si f et g sont mesurables et λ est un scalaire, alors $f + g$, fg , λf , $\sup\{f, g\}$ le sont, et si (f_n) est une suite de fonctions mesurables qui converge simplement vers une fonction f , alors f est également mesurable.

Rappelons également que, si $f : X \rightarrow \mathbb{R}$ est une fonction mesurable et positive, alors $\int_X f(x)d\mu(x)$ est une quantité bien définie, qui appartient à $[0, +\infty]$. Lorsque cette quantité est finie, on dit que f est intégrable. Plus généralement, si $f : X \rightarrow \mathbb{R}$, on dit encore que f est intégrable si $|f|$ est intégrable (noter que $|f| = \max\{f, -f\}$, donc $|f|$ est encore mesurable). Dans ce cas

$$\int_X f(x)d\mu(x) := \int_X f^+(x)d\mu(x) - \int_X f^-(x)d\mu(x) \quad \text{où } f^+ = \max\{f, 0\} \text{ et } f^- = \max\{-f, 0\}.$$

Rappelons enfin qu'une propriété $\mathcal{P}(x)$ définie pour $x \in X$ est vraie μ -presque partout s'il existe un ensemble de mesure nulle N telle que $\mathcal{P}(x)$ est vraie pour tout $x \in X \setminus N$.

Voici quelques inégalités classiques qui permettent de majorer des intégrales :

Proposition A.1 *Soit (X, \mathcal{A}, μ) un espace mesuré.*

- (inégalité triangulaire) si f est intégrable, alors

$$\left| \int_X f d\mu \right| \leq \int_X |f| d\mu$$

- (inégalité de Jensen) Soit $\mu(X) = 1$ et f intégrable. On suppose que $a < f < b$ μ -p.p. (où $-\infty \leq a < b \leq +\infty$). Soit $\Phi :]a, b[\rightarrow \mathbb{R}$ une fonction convexe. Alors

$$\Phi\left(\int_X f d\mu\right) \leq \int_X \Phi(f) d\mu$$

où le membre de droite est bien défini et appartient à $] -\infty, +\infty]$.

Voici trois résultats de convergence qu'il faut très bien connaître :

Proposition A.2 (Convergence monotone) *Si (f_n) est une suite croissante de fonctions mesurables, positives, alors*

$$\lim_{n \rightarrow +\infty} \int_X f_n(x) d\mu(x) = \int_X \lim_{n \rightarrow +\infty} f_n(x) d\mu(x).$$

Lemme A.3 (de Fatou) Si (f_n) est une suite de fonctions mesurables positives, alors

$$\int_X \liminf_{n \rightarrow +\infty} f_n(x) d\mu(x) \leq \liminf_{n \rightarrow +\infty} \int_X f_n(x) d\mu(x)$$

Théorème A.4 (Convergence dominée) Si (f_n) une suite de fonctions mesurables qui converge presque partout vers une fonction f et pour laquelle il existe une fonction intégrable g telle que

$$|f_n(x)| \leq g(x) \quad \text{pour } \mu\text{-presque tout } x \in X, \text{ pour tout } n \in \mathbb{N}.$$

Alors f est intégrable et

$$\lim_n \int_X f_n d\mu = \int_X f d\mu.$$

Remarque A.5 De plus, la convergence de (f_n) vers f a lieu au sens L^1 :

$$\lim_{n \rightarrow +\infty} \int_X |f - f_n| d\mu = 0.$$

A.2 Les espaces L^p

Soit (X, \mathcal{A}, μ) un espace mesuré et p un réel supérieur ou égal à 1. L'espace $\mathcal{L}^p(X, \mu)$ est l'ensemble des fonctions mesurables $f : X \rightarrow \mathbb{R}$ telles que l'intégrale $\int_X |f(x)|^p d\mu(x)$ est finie. Sur $\mathcal{L}^p(X, \mu)$, on définit la relation d'équivalence $f \sim g$ si $f = g$ μ -p.p. On note $L^p(X, \mu)$ l'ensemble des classes d'équivalences de \sim . L'idée est que l'on peut manipuler les éléments de $L^p(X, \mu)$ à peu près comme ceux de $\mathcal{L}^p(X, \mu)$: en particulier, si $f, g \in L^p(X, \mu)$, on peut définir $f + g$ en prenant la classe d'équivalence de n'importe somme $\tilde{f} + \tilde{g}$ où \tilde{f} et \tilde{g} sont des représentants de f et g (exercice).

Lorsque $p = +\infty$, on définit $L^\infty(X, \mu)$ l'ensemble des (classes d'équivalence de) fonctions qui sont essentiellement bornées : $f \in L^\infty(X, \mu)$ si f est mesurable et s'il existe une constante $C \geq 0$ telle que $|f| \leq C$ μ -p.p. La norme $\|f\|_\infty$ est alors la plus petite constante C pour laquelle cette inégalité est vérifiée.

Soit $p \in]1, +\infty[$. On appelle exposant conjugué de p le nombre réel p' tel que

$$\frac{1}{p} + \frac{1}{p'} = 1 \quad \text{i.e., } p' = \frac{p}{p-1}.$$

En particulier, $p' > 1$. Si par exemple $p = 2$, alors $p' = 2$. Lorsque $p = 1$, on pose par convention $p' = +\infty$, tandis que lorsque $p = +\infty$, on pose $p' = 1$.

Lemme A.6 (Inégalité de Hölder) Soit $p \in [1, +\infty]$ et $f \in L^p(X, \mu)$ et $g \in L^{p'}(X, \mu)$, alors $fg \in L^1(X, \mu)$ et

$$\int_X fg d\mu \leq \left(\int_X |f|^p d\mu \right)^{\frac{1}{p}} \left(\int_X |g|^{p'} d\mu \right)^{\frac{1}{p'}}$$

Remarque : Une conséquence très importante de l'inégalité de Hölder, est la suite d'inclusion suivantes :

$$\text{si } \mu(X) < +\infty, \text{ alors } L^\infty(X) \subset L^p(X) \subset L^r(X) \subset L^1(X) \quad \forall 1 \leq r \leq p \leq +\infty.$$

Par contre, aucune de ces inclusions n'est vraie si $\mu(X) = \infty$.

Pour montrer ces inclusions (*), il suffit de prendre $r < p$ et, si $u \in L^p(X)$, on peut appliquer l'inégalité de Hölder aux fonctions $f = |u|^r$ et $g = 1$ avec le coefficient $\theta = p/r > 1$ et $\theta' = \theta/(\theta-1)$:

$$\int_X |u|^r d\mu \leq \left(\int_X |u|^{r\theta} d\mu \right)^{1/\theta} \left(\int_X 1^{\theta'} d\mu \right)^{1/\theta'} = (\mu(X))^{1/\theta'} \left(\int_X |u|^p d\mu \right)^{1/\theta'} < +\infty$$

Donc $u \in L^r(X)$. □

Pour $f \in L^p(X, \mu)$, on pose

$$\|f\|_p = \left(\int_X |f(x)|^p d\mu(x) \right)^{1/p}$$

Rappelons l'inégalité de Minkowski : si $f, g \in L^p(X, \mu)$, alors $f + g \in L^p(X, \mu)$ et

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p .$$

En particulier, $\|\cdot\|_p$ est une norme sur $L^p(X, \mu)$.

Théorème A.7 (Riesz-Fischer) *L'espace $L^p(X, \mu)$, muni de la norme $\|\cdot\|_p$, est un espace de Banach. Lorsque $p = 2$, l'espace $L^2(X, \mu)$ est un espace de Hilbert lorsqu'on le munit du produit scalaire*

$$\langle f, g \rangle = \int_X f(x) \overline{g(x)} d\mu(x) \quad \forall f, g \in L^2(X, \mu) .$$

Rappelons que la convergence dans L^p n'implique pas en général la convergence ponctuelle, ni même la convergence presque partout. Par contre, si (f_n) converge vers f dans L^p (pour $p \in [1, +\infty]$), alors il existe une sous-suite (f_{n_k}) qui converge presque partout vers f .

En effet, supposons que $p < +\infty$ (pour $p = +\infty$ c'est évident). Comme (f_n) est une suite de Cauchy dans L^p , il existe une sous-suite (f_{n_k}) telle que

$$\|f_{n_{k+1}} - f_{n_k}\|_p \leq \frac{1}{2^k} \quad \forall k \geq 0 .$$

Donc, si on pose $g_n = \sum_{k=0}^n |f_{n_{k+1}} - f_{n_k}|$, on a $\|g_n\|_p \leq 1$. Par convergence monotone, cela implique que la limite ponctuelle g de la suite croissante (g_n) vérifie également $\|g\|_p \leq 1$. En particulier $X_\infty := \{x \in X, g(x) = +\infty\}$ est de mesure nulle. Pour tout $x \in X \setminus X_\infty$, on a $\sum_k |f_{n_{k+1}}(x) - f_{n_k}(x)| < +\infty$, et donc la série $\sum_k (f_{n_{k+1}}(x) - f_{n_k}(x))$ est absolument convergente, donc convergente. On en déduit que la limite $\bar{f}(x)$ de la suite $(f_{n_k}(x))$ existe pour presque tout x . Mais, d'après le lemme de Fatou, on a

$$\int_X |f(x) - \bar{f}(x)|^p d\mu(x) \leq \liminf_{k \rightarrow +\infty} \int_X |f(x) - f_{n_k}(x)|^p d\mu(x) = 0$$

où la dernière égalité vient du fait que (f_{n_k}) tend vers f dans L^p . On en déduit que $\bar{f} = f$ p.p., et donc que la suite (f_{n_k}) tend vers f p.p. □

A.3 Intégration sur un espace produit

Soient $(X_1, \mathcal{A}_1, \mu_1)$ et $(X_2, \mathcal{A}_2, \mu_2)$ deux espace mesurés. On appelle tribu produit de \mathcal{A}_1 et \mathcal{A}_2 la tribu engendrée par le produit $\mathcal{A}_1 \times \mathcal{A}_2$ (qui n'est pas une tribu en général). On note cette tribu produit $\mathcal{A}_1 \otimes \mathcal{A}_2$.

On montre qu'il existe une unique mesure μ sur $X_1 \times X_2$ telle que

$$\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2) \quad \forall (A_1, A_2) \in \mathcal{A}_1 \times \mathcal{A}_2 .$$

Cette mesure est notée $\mu_1 \otimes \mu_2$ et est appelée la mesure produit de μ_1 et μ_2 .

Théorème A.8 (Fubini 1—pour les fonctions positives) Soit $f : X_1 \times X_2 \rightarrow \mathbb{R}$ une application mesurable par rapport à la tribu produit $\mathcal{A}_1 \otimes \mathcal{A}_2$ et positive. Alors

1. L'application $h_1(x_1) := \int_{X_2} f(x_1, x_2) d\mu_2(x_2)$ est mesurable et

$$\int_{X_1} h_1(x_1) d\mu_1(x_1) = \int_{X_1 \times X_2} f(x_1, x_2) d(\mu_1 \otimes \mu_2)(x_1, x_2).$$

2. L'application $h_2(x_2) := \int_{X_1} f(x_1, x_2) d\mu_1(x_1)$ est mesurable et

$$\int_{X_2} h_2(x_2) d\mu_2(x_2) = \int_{X_1 \times X_2} f(x_1, x_2) d(\mu_1 \otimes \mu_2)(x_1, x_2).$$

3. En particulier,

$$\begin{aligned} & \int_{X_1 \times X_2} f d(\mu_1 \otimes \mu_2) \\ &= \int_{X_1} \left(\int_{X_2} f(x_1, x_2) d\mu_2(x_2) \right) d\mu_1(x_1) = \int_{X_2} \left(\int_{X_1} f(x_1, x_2) d\mu_1(x_1) \right) d\mu_2(x_2). \end{aligned}$$

Toutes les intégrales ci-dessus sont bien définies, et appartiennent à $[0, +\infty]$.

Remarque A.9 En pratique, le résultat ci-dessus permet de montrer qu'une fonction $f = f(x_1, x_2)$ est intégrable. Lorsque c'est le cas, on peut alors appliquer le théorème de Fubini 2 :

Théorème A.10 (Fubini 2—pour les fonctions intégrables) Soit $f : X_1 \times X_2 \rightarrow \mathbb{R}$ une application intégrable par rapport à la mesure produit $\mu_1 \otimes \mu_2$. Alors

1. L'application $h_1(x_1) := \int_{X_2} f(x_1, x_2) d\mu_2(x_2)$ est définie pour μ_1 -presque tout x_1 , est intégrable, et

$$\int_{X_1} h_1(x_1) d\mu_1(x_1) = \int_{X_1 \times X_2} f d(\mu_1 \otimes \mu_2).$$

2. L'application $h_2(x_2) := \int_{X_1} f(x_1, x_2) d\mu_1(x_1)$ est définie pour μ_2 -presque tout x_2 , est intégrable, et

$$\int_{X_2} h_2(x_2) d\mu_2(x_2) = \int_{X_1 \times X_2} f d(\mu_1 \otimes \mu_2).$$

3. En particulier,

$$\begin{aligned} & \int_{X_1 \times X_2} f d(\mu_1 \otimes \mu_2) \\ &= \int_{X_1} \left(\int_{X_2} f(x_1, x_2) d\mu_2(x_2) \right) d\mu_1(x_1) = \int_{X_2} \left(\int_{X_1} f(x_1, x_2) d\mu_1(x_1) \right) d\mu_2(x_2). \end{aligned}$$

Remarque A.11 Un exemple particulièrement simple d'application est lorsque f_1 et f_2 sont intégrables par rapport à μ_1 et μ_2 respectivement. Alors la fonction $f(x_1, x_2) = f_1(x_1)f_2(x_2)$ est intégrable par rapport à la mesure produit $\mu_1 \otimes \mu_2$ et

$$\int_{X_1 \times X_2} f(x_1, x_2) d(\mu_1 \otimes \mu_2)(x_1, x_2) = \left(\int_{X_1} f_1(x_1) d\mu_1(x_1) \right) \left(\int_{X_2} f_2(x_2) d\mu_2(x_2) \right).$$

A.4 Un peu de calcul intégral

Nous aurons besoin de faire un tout petit peu de calcul intégral dans \mathbb{R}^N (pour $N \geq 2$) et d'utiliser le passage en coordonnées polaires. Pour cela, on gardera en tête que, si $u : \mathbb{R}^N \rightarrow \mathbb{R}$ est une fonction radiale, i.e., s'il existe une fonction $\tilde{u} : [0, +\infty[\rightarrow \mathbb{R}$ telle que $u(x) = \tilde{u}(\|x\|)$, alors $u \in L^1(\mathbb{R}^N)$, si et seulement si, la fonction $r \rightarrow r^{N-1}\tilde{u}(r)$ appartient à $L^1([0, +\infty[)$. Dans ce cas, on a

$$\int_{\mathbb{R}^N} u(x) dx = C_N \int_0^{+\infty} r^{N-1} \tilde{u}(r) dr$$

où C_N est le volume (pour la mesure de Lebesgue) de la boule unité de \mathbb{R}^N .

Dans le cas particulier où $N = 2$, si $u : \mathbb{R}^2 \rightarrow \mathbb{R}$, on pose $\tilde{u}(r, \theta) = u(r \cos(\theta), r \sin(\theta))$ pour $(r, \theta) \in [0, +\infty[\times [0, 2\pi[$. Alors on a par changement de variables que $u \in L^1(\mathbb{R}^2)$, si et seulement si l'application $(r, \theta) \rightarrow r\tilde{u}(r, \theta)$ est dans $L^1([0, +\infty[\times [0, 2\pi[)$. Dans ce cas

$$\int_{\mathbb{R}^2} u(x) dx = \int_0^{+\infty} \int_0^{2\pi} r\tilde{u}(r, \theta) dr d\theta$$

A.5 Produit de convolution

On travaille ici dans \mathbb{R}^N muni de la tribu borélienne et de la mesure de Lebesgue. Soient f et g deux applications de \mathbb{R}^N dans \mathbb{R}^N , positives. On appelle produit de convolution de f et g , noté $f \star g$, l'application

$$(f \star g)(x) = \int_{\mathbb{R}^N} f(y)g(x-y) dy = \int_{\mathbb{R}^N} f(x-y)g(y) dy,$$

lorsque ces quantités sont bien définies et qu'on peut appliquer le théorème de Fubini.

Proposition A.12 *Voici trois cas où le produit de convolution est bien défini.*

- Si f et g sont intégrables, alors $(f \star g)(x)$ est défini pour presque tout $x \in \mathbb{R}^N$, $f \star g$ est également intégrable et

$$\int_{\mathbb{R}^N} (f \star g)(x) dx = \left(\int_{\mathbb{R}^N} f(x) dx \right) \left(\int_{\mathbb{R}^N} g(x) dx \right).$$

- Si $f \in L^p(\mathbb{R}^N)$ et $g \in L^{p'}(\mathbb{R}^N)$ (avec $p \in [1, +\infty[$ et p' l'exposant conjugué de p), alors $f \star g(x)$ est défini pour presque tout $x \in \mathbb{R}^N$ et $f \star g \in L^\infty(\mathbb{R}^N)$ avec $\|f \star g\|_\infty \leq \|f\|_p \|g\|_{p'}$.
- Si $f \in L^p(\mathbb{R}^N)$ et $g \in L^1(\mathbb{R}^N)$ (avec $p \in [1, +\infty[$), alors $(f \star g)(x)$ est défini pour presque tout $x \in \mathbb{R}^N$, $f \star g \in L^p(\mathbb{R}^N)$ et $\|f \star g\|_p \leq \|f\|_p \|g\|_1$.

Remarque A.13 En particulier, si f et g sont des densités de probabilité, i.e., $f \geq 0$ et $g \geq 0$ p.p. avec

$$\int_{\mathbb{R}} f(x) dx = \int_{\mathbb{R}} g(x) dx = 1,$$

alors $f \star g$ est également une densité de probabilité. Rappelons que si X et Y sont des variables aléatoires réelles de densité respectives f et g , et si X et Y sont indépendantes, alors la variable aléatoire $X + Y$ a pour densité $f \star g$.

Preuve de la proposition (*): On suppose d'abord que $f, g \in L^1(\mathbb{R}^N)$. Notons d'abord que l'application $(x, t) \rightarrow f(y)g(x - y)$ est bien intégrable sur $\mathbb{R}^N \times \mathbb{R}^N$ (on admet la mesurabilité). En effet, en utilisant le théorème de Fubini 1 puis un changement de variable, on a

$$\begin{aligned} \int_{\mathbb{R}^N \times \mathbb{R}^N} |f(y)g(x - y)| dx dy &= \int_{\mathbb{R}^N} |f(y)| \left(\int_{\mathbb{R}^N} |g(x - y)| dx \right) dy = \int_{\mathbb{R}^N} |f(y)| \left(\int_{\mathbb{R}^N} |g(x)| dx \right) dy \\ &= \|g\|_1 \int_{\mathbb{R}^N} |f(y)| dy = \|g\|_1 \|f\|_1 < +\infty . \end{aligned}$$

Donc $(x, t) \rightarrow f(y)g(x - y)$ est intégrable sur $\mathbb{R}^N \times \mathbb{R}^N$, et on a, d'après le théorème de Fubini 2, que l'intégrale $\int_{\mathbb{R}^N} f(y)g(x - y)$ est définie pour presque tout x (ce qui définit $(f \star g)(x)$ pour presque tout x) et

$$\begin{aligned} \int_{\mathbb{R}^N} (f \star g)(x) dx &= \int_{\mathbb{R}^N} \left(\int_{\mathbb{R}^N} f(y)g(x - y) dy \right) dx = \int_{\mathbb{R}^N} f(y) \left(\int_{\mathbb{R}^N} g(x - y) dx \right) dy \\ &= \int_{\mathbb{R}^N} f(y) \left(\int_{\mathbb{R}^N} g(x) dx \right) dy = \left(\int_{\mathbb{R}^N} g(x) dx \right) \left(\int_{\mathbb{R}^N} f(y) dy \right) . \end{aligned}$$

Supposons maintenant que $f \in L^p(\mathbb{R}^N)$ tandis que $g \in L^{p'}(\mathbb{R}^N)$. Alors pour tout $x \in \mathbb{R}^N$, on a par inégalité de Hölder,

$$\int_{\mathbb{R}^N} |f(y)g(x - y)| dy \leq \left(\int_{\mathbb{R}^N} |f(y)|^p \right)^{1/p} \left(\int_{\mathbb{R}^N} |g(x - y)|^{p'} \right)^{1/p'} = \|f\|_p \|g\|_{p'} < +\infty .$$

Donc la fonction $y \rightarrow f(y)g(x - y)$ est intégrable, avec

$$\left| \int_{\mathbb{R}^N} f(y)g(x - y) dy \right| \leq \int_{\mathbb{R}^N} |f(y)g(x - y)| dy \leq \|f\|_p \|g\|_{p'} .$$

Supposons pour finir que $f \in L^p(\mathbb{R}^N)$ et $g \in L^1(\mathbb{R}^N)$ (avec $p \in [1, +\infty]$). Par Hölder, on a

$$\begin{aligned} \int_{\mathbb{R}^N} |f(x - y)| |g(y)| dy &= \int_{\mathbb{R}^N} |f(x - y)| |g(y)|^{1/p} |g(y)|^{1/p'} dy \\ &\leq \left(\int_{\mathbb{R}^N} |f(x - y)|^p |g(y)| dy \right)^{1/p} \left(\int_{\mathbb{R}^N} |g(y)| dy \right)^{1/p'} , \end{aligned}$$

où la fonction $|f|^p$ est dans $L^1(\mathbb{R}^N)$. Donc

$$\begin{aligned} \|f \star g\|_p^p &\leq \int_{\mathbb{R}^N} \left(\int_{\mathbb{R}^N} |f(x - y)| |g(y)| dy \right)^p dx \leq \|g\|_1^{p-1} \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} |f(x - y)|^p |g(y)| dy \\ &\leq \|g\|_1^{p-1} \| |f|^p \|_1 \|g\|_1 = \|f\|_p \|g\|_1 \end{aligned}$$

d'après la première partie du théorème. □

A.6 Régularisation

Fonction continue à support compact : soit $]a, b[$ un intervalle ouvert de \mathbb{R} , avec $-\infty \leq a < b \leq +\infty$. On dit qu'une fonction mesurable $f :]a, b[\rightarrow \mathbb{R}$ est à support compact s'il existe un sous-intervalle $[b, c]$ fermé et borné de $]a, b[$ en dehors duquel la fonction f est nulle p.p..

Plus généralement, si Ω est un ouvert de \mathbb{R}^N , on dit qu'une fonction mesurable $f : \Omega \rightarrow \mathbb{R}$ est à support compact s'il existe un compact $K \subset \Omega$ tel que f est nulle p.p. en dehors de K . On appelle support de f le plus petit compact (au sens de l'inclusion) en dehors duquel f est nulle p.p..

Rappelons qu'il existe des fonctions de classe \mathcal{C}^∞ non identiquement nulles et à support compact dans \mathbb{R} : pour construire une telle fonction, on remarque d'abord que la fonction

$$\phi(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ e^{-1/x} & \text{si } x > 0 \end{cases}$$

est de classe \mathcal{C}^∞ sur \mathbb{R} , est strictement positive dans $]0, +\infty[$ et nulle dans $] - \infty, 0]$. On pose alors $f(x) = \phi(x)\phi(1-x)$. La fonction f est de classe \mathcal{C}^∞ , positive sur \mathbb{R} . De plus $f(x) > 0$ si et seulement si $x > 0$ et $1-x > 0$: donc f est à support compact et son support est le compact $[0, 1]$.

Un tel exemple se généralise aisément à \mathbb{R}^N : il suffit de prendre par exemple $\tilde{f}(x) = f(\|x\|^2)$ où f est la fonction construite ci-dessus et $\|x\|$ est la norme euclidienne dans \mathbb{R}^N . La fonction \tilde{f} est de classe \mathcal{C}^∞ et a pour support la boule unité de \mathbb{R}^N qui est compacte.

Si Ω est un ouvert de \mathbb{R}^N et $k \geq 0$, on note $\mathcal{C}_c^k(\Omega)$ l'ensemble des fonctions de classe \mathcal{C}^k qui sont à support compact dans Ω .

Fonctions localement intégrables : soit Ω un sous-ensemble de \mathbb{R}^N . On dit qu'une fonction mesurable $f : \Omega \rightarrow \mathbb{R}$ est localement intégrable si la restriction de f à tout compact de \mathbb{R}^N est intégrable. On note $f \in L_{loc}^1(\Omega)$.

Par exemple, les constantes sont localement intégrables dans \mathbb{R}^N (mais seule la constante nulle est intégrable). Notons aussi que, si Ω est un ouvert de \mathbb{R}^N , $f \in L_{loc}^1(\Omega)$ et $g \in \mathcal{C}_c^0(\Omega)$, alors le produit fg est intégrable dans Ω .

Proposition A.14 *Si $f \in L_{loc}^1(\mathbb{R})$ et $g \in \mathcal{C}_c^k(\mathbb{R})$, alors $f \star g$ est bien défini et de classe \mathcal{C}^k dans \mathbb{R} . De plus, $(f \star g)^{(r)} = f \star g^{(r)}$ pour tout $r = 1, \dots, k$.*

De même, si $f \in L_{loc}^1(\mathbb{R}^N)$ et $g \in \mathcal{C}_c^k(\mathbb{R}^N)$, alors $f \star g$ est bien défini et de classe \mathcal{C}^k dans \mathbb{R}^N . De plus,

$$\frac{\partial^r (f \star g)}{\partial x_1^{\alpha_1} \dots \partial x_N^{\alpha_N}} = f \star \left(\frac{\partial^r g}{\partial x_1^{\alpha_1} \dots \partial x_N^{\alpha_N}} \right)$$

pour tout multi-indice $(\alpha_1, \dots, \alpha_N) \in \mathbb{N}^N$ avec $\alpha_1 + \dots + \alpha_N = r \leq k$.

Preuve * : On ne fait la preuve qu'en dimension $N = 1$, la preuve dans le cas général étant identique. Par argument de récurrence, il suffit aussi de faire la démonstration pour $k = 1$. Comme g est à support compact, il existe un intervalle $[c, d]$ en dehors duquel g est nulle. Pour $x \in \mathbb{R}$, $h \in \mathbb{R}$, $|h| < 1$ et $h \neq 0$, on a

$$\begin{aligned} \frac{(f \star g)(x+h) - (f \star g)(x)}{h} &= \int_{\mathbb{R}} f(y) \left(\frac{g(x+h-t) - g(x-t)}{h} \right) dt \\ &= \int_{x-1-d}^{x+1-c} f(y) \left(\frac{g(x+h-t) - g(x-t)}{h} \right) dt \end{aligned}$$

Or, si $h_n \rightarrow 0$, on a que la suite $\left(f(y) \left(\frac{g(x+h_n-t) - g(x-t)}{h_n} \right) \right)$ tend vers $f(y)g'(x-t)$ pour presque tout $t \in [x-1-d, x+1-c]$ et

$$\left| f(y) \left(\frac{g(x+h_n-t) - g(x-t)}{h_n} \right) \right| \leq |f(y)| \|g'\|_\infty \text{ p.p.t } t \in [x-1-d, x+1-c].$$

Or f est intégrable sur le compact $[x-1-d, x+1-c]$ et on peut donc appliquer le théorème de convergence dominée :

$$\lim_{n \rightarrow +\infty} \frac{(f \star g)(x+h_n) - (f \star g)(x)}{h_n} = \int_{x-1-d}^{x+1-c} f(y)g'(x-t)dt = (f \star g')(x).$$

Comme ceci est vrai pour toute suite (h_n) qui tend vers 0, on a montré que $f \star g$ est dérivable sur \mathbb{R} , avec $(f \star g)' = f \star g'$. Par un argument de convergence dominée, on a aussi que l'application $x \rightarrow (f \star g')(x)$ est continue, ce qui prouve que $f \star g$ est de classe \mathcal{C}^1 . \square

Rappelons un résultat essentiel de densité :

Théorème A.15 Soit Ω un ouvert non vide de \mathbb{R}^N et $p \in [1, +\infty[$. Alors $\mathcal{C}_c^0(\Omega)$ est dense dans $L^p(\Omega)$ au sens où : pour tout $f \in L^p(\Omega)$ et pour $\epsilon > 0$, il existe $f_\epsilon \in \mathcal{C}_c^0(\Omega)$ tel que $\|f - f_\epsilon\|_p \leq \epsilon$.

Ce résultat, qui repose sur la construction de l'intégrale de Lebesgue, ne sera pas montré ici.

Approximation par convolution : Soit $\phi : \mathbb{R}^N \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^∞ , à support compact inclus dans la boule unité $B(0, 1)$, avec $\phi \geq 0$ et $\int_{\mathbb{R}^N} \phi(x) dx = 1$. Pour tout $\epsilon > 0$, on pose $\phi_\epsilon(x) = \epsilon^{-N} \phi(\frac{x}{\epsilon})$. On note que ϕ_ϵ a un support inclus dans la boule $B(0, \epsilon)$ et que $\int_{\mathbb{R}^N} \phi_\epsilon(x) dx = 1$.

Théorème A.16 Soit $u \in L^p(\mathbb{R}^N)$ et $u_\epsilon = \phi_\epsilon \star u$. Alors u_ϵ tend vers u dans $L^p(\mathbb{R}^N)$ lorsque $\epsilon \rightarrow 0$.

Remarques A.17 1. Si u a un support contenu dans un compact K , alors u_ϵ est également à support compact, avec un support contenu dans l'ensemble $K_\epsilon := \{x \in \mathbb{R}^N, \exists y \in K, \|x - y\| \leq \epsilon\}$ (par exemple, si $N = 1$ et $K = [a, b]$, alors le support de u_ϵ est contenu dans $[a - \epsilon, b + \epsilon]$).

2. Si u est continue et à support compact, la convergence de u_ϵ vers u est uniforme sur \mathbb{R}^N .

Preuve * : On suppose d'abord que $u \in \mathcal{C}_c^0(\mathbb{R}^N)$. Alors, comme $\int_{\mathbb{R}^N} \phi_\epsilon = 1$, on a

$$u(x) - u_\epsilon(x) = \int_{\mathbb{R}^N} (u(x) - u(x - z)) \phi_\epsilon(z) dz$$

Comme u est continue et a un support compact, u est uniformément continue : pour tout $\delta > 0$, il existe $\eta > 0$ tel que $|u(x) - u(y)| \leq \delta$ dès que $\|x - y\| \leq \eta$. Pour tout $\epsilon \in]0, \eta[$, on a, puisque le support de ϕ_ϵ est contenu dans la boule $B(0, \epsilon)$ et $\phi_\epsilon \geq 0$,

$$\begin{aligned} |u(x) - u_\epsilon(x)| &= \left| \int_{B(0, \epsilon)} (u(x) - u(x - z)) \phi_\epsilon(z) dz \right| \\ &\leq \int_{B(0, \epsilon)} |u(x) - u(x - z)| \phi_\epsilon(z) dz \leq \delta \int_{\mathbb{R}^N} \phi_\epsilon(z) dz = \delta. \end{aligned}$$

Donc u_ϵ tend uniformément vers u sur \mathbb{R}^N . Comme, pour $\epsilon \in]0, 1]$, u_ϵ et u ont un support contenu dans un même compact K (cf. la remarque ci-dessus), u_ϵ tend aussi vers u dans $L^p(\mathbb{R}^N)$ par convergence dominée.

Nous traitons maintenant le cas général : soit $u \in L^p(\mathbb{R}^N)$. On veut montrer que $\|u - u_\epsilon\|_p \rightarrow 0$ lorsque $\epsilon \rightarrow 0$. Fixons $\delta > 0$. Alors il existe $w \in \mathcal{C}_c^0(\Omega)$ tel que $\|u - w\|_p \leq \delta/3$. Posons $w_\epsilon = \phi_\epsilon \star w$. Alors, comme $(w - u) \in L^p$ tandis que $\phi_\epsilon \in L^1$, on a

$$\|w_\epsilon - u_\epsilon\|_p = \|(w - u) \star \phi_\epsilon\|_p \leq \|w - u\|_p \|\phi_\epsilon\|_1 = \|w - u\|_p \leq \delta/3,$$

puisque $\phi_\epsilon \geq 0$ et $\int \phi_\epsilon = 1$. D'autre part, comme w_ϵ tend vers w dans L^p , il existe $\epsilon_0 > 0$ tel que pour $\epsilon \in]0, \epsilon_0[$, $\|w_\epsilon - w\|_p \leq \delta/3$. Alors, pour $\epsilon \in]0, \epsilon_0[$, on a

$$\|u - u_\epsilon\|_p \leq \|u - w\|_p + \|w - w_\epsilon\|_p + \|w_\epsilon - u_\epsilon\|_p \leq \delta.$$

\square