

ISP/BUKAVU

Département de
Mathématique-Physique



Analyse fonctionnelle

AFM2111

Dr. Dieudonné Z. BALIKE, Ph.D.

Cours destiné aux étudiants de Master 1 Mathématiques
2024-2025

Contrat pédagogique

Identification de l'UE

Code : AFM2111

Intitulé : Analyse fonctionnelle.

Rattachement : Section Sciences Exactes, Département de Mathématique-Physique.

Destinataires : Étudiants de Master 1 Mathématiques.

Prérequis : Pour bien participer à ce cours, l'étudiant est supposé avoir des connaissances de Topologie générale, de Théorie de la mesure et de l'analyse réelle. Néanmoins, conscient que ces prérequis peuvent avoir été oubliés, je vais rappeler les notions fondamentales de Topologie et de Théorie de la mesure qui seront utilisées ici.

Présentation de l'animateur

Dieudonné Zirhumanana Balike est chargé des CMI et les TD seront assurés cette année par le Chef de Travaux Byamezi Munigwa Guillaume.

M. Balike a un doctorat en Mathématiques Appliquées de l'Université de Naples Federico II, en Italie.

Ses recherches portent sur les équations différentielles ordinaires et les équations différentielles aux dérivées partielles et leurs applications en Biologie, en Écologie, en Médecine, etc.

Objectifs de l'UE et Compétences à développer

Cette UE a pour objectifs, entre autres, de :

- Introduire les bases de l'Analyse fonctionnelle aux étudiants de Master en Mathématiques.
- Donner les fondements théoriques des équations aux dérivées partielles.

Chaque étudiant qui aura suivi l'intégralité de cette UE et qui l'aura révisée plusieurs fois sera à mesure de :

- Énoncer et démontrer les théorèmes, les propositions, les lemmes donnés dans ce cours ;

- Contextualiser et justifier chacun des théorèmes, propositions, lemmes, etc. donnés dans ce cours.
- Choisir les espaces fonctionnels adaptés pour un problème faisant intervenir les équations aux dérivées partielles.

Matériels du cours

Le cours est sur un support électronique et il est gratuitement distribué aux étudiants sous le format électronique. Le cours et les TD sont sur des supports différents. Des références (majoritairement en anglais et des sites internet seront aussi rendus disponibles pour les approfondissements).

Nous utiliserons le tableau et la craie et de temps en temps nous ferons la projection quand c'est nécessaire.

Déroulement de l'UE

L'UE a lieu au cinquième semestre .

Les CMI et les TD seront séparés comme ils sont animés par deux intervenants.

Mode d'évaluation

Des travaux en groupes et individuels seront faits pendant les CMI et les TD pour constituer les points du contrôle continu dont la pondération est de 30 points. L'examen sera à notes fermées et sera coté sur 30. Les questions de l'examen seront tirées des questions déjà dans les notes et de la partie théorique de l'UE (démonstration des théorèmes importants).

Tout étudiant absent à une évaluation se verra attribuée la cote zéro s'il échoue de présenter une justification délivrée par le département dans les 72 heures qui suivent l'évaluation. Lorsque la justification est présentée, une évaluation spéciale sera programmée par l'étudiant *selon la disponibilité de l'enseignant et avant l'examen*.

Un étudiant qui aura raté l'examen, sera automatiquement ajourné et présentera un examen de rattrapage.

Avertissements

Malgré mon effort de nettoyer ce manuscrit, il est fort possible qu'il subsiste des coquilles et des fautes. N'hésitez pas à me les signaler pour les corriger.

Ce cours peut vite devenir très difficile pour les étudiants qui n'ont pas des bases solides des prérequis donnés. Même pour ceux qui ont des bases nécessaires, il peut être indigeste si des efforts constants d'assimilation ne sont pas fournis.

Il est donc conseillé aux étudiants de lire, relire et plusieurs fois ce cours en *refaisant tous les calculs et démonstrations* faits en classe. Un jour manqué risque de gâcher l'apprentissage.

De grâce, ne passez pas un jour sans avoir refait au moins 3 fois les calculs déjà faits, même quand vous estimez que vous les connaissez déjà.

Table des matières

Table des matières	i
1 Rappels de Topologie et de Théorie de la Mesure	1
1.1 Espaces de Baire	4
1.2 Opérateurs linéaires et théorème de Banach-Steinhaus	6
1.3 Théorèmes de l'application ouverte et du graphe fermé	8
1.4 Rappels des résultats importants de la théorie de la mesure	10
1.4.1 Quelques définitions fondamentales	10
1.4.2 Théorèmes majeurs de la théorie de la mesure	10
1.4.3 Résumé sur les Espaces L^p	11
1.4.3.1 Définitions	11
1.4.3.2 Propriétés fondamentales des espaces L^p	12
1.5 Exercices	12
2 Théorème de Hahn-Banach	15
2.1 Forme analytique	15
2.2 Formes géométriques	16
2.3 Un exemple d'application : le théorème de Krein Milman	21
3 Dualité et Topologies faibles	25
3.1 Dual topologique	26
3.2 Topologie faible	30
3.3 Topologie faible*	33
3.4 Notion d'espace réflexif	34
4 Théorie de distributions (ou de Schwartz)	40
4.1 Définition et propriétés élémentaires	40
4.1.1 L'espace des distributions	40
4.1.2 Dérivation au sens des distributions	42
4.2 Quelques exemples	46
4.2.1 Mesures de Radon	46
4.2.2 Valeurs principales et parties finies	49
4.2.3 Dérivées des indicatrices et formule de saut	52
4.3 Localisation, convolution, régularisation	53
5 Application des distributions aux solutions singulières des EDP	61
5.1 Méthode des caractéristiques et explosion	61
5.1.1 La méthode des caractéristiques	62

5.2	Solutions faibles et critère d'unicité	64
5.2.1	Solutions au sens des distributions	64
5.2.2	Solutions entropiques	65
5.3	Approximation visqueuse	66
5.3.1	La transformation de Cole-Hopf	67
5.3.2	Le problème de Cauchy pour l'équation de Burgers	69
5.3.3	Convergence vers l'équation de Hopf	70
6	Transformée de Fourier	75
6.1	Transformation de Fourier des fonctions sommables	75
6.2	Transformation de Fourier sur la classe de Schwartz	78
6.3	Transformation de Fourier des distributions tempérées	81
6.4	Equations de convolution	85
7	Espaces de Hilbert	87
7.1	Définition et propriétés fondamentales	87
7.2	Théorèmes de Stampacchia et de Lax-Milgram	93
7.3	Bases hilbertiennes	97
8	Espaces de Sobolev	99
8.1	Définition et propriétés topologiques	100
8.2	Injections de Sobolev	102
8.3	Restriction, prolongement et traces	104
8.4	Compacité	109

Rappels de Topologie et de Théorie de la Mesure

Comme l'indique le titre, l'objectif de ce chapitre n'est pas de présenter de façon complète les bases de topologie ou de Théorie de la mesure mais de donner quelques éléments dont nous auront besoin par la suite pour asseoir la théorie que nous développons ici. On se réfère pour cela aux cours de Topologie générale et de Théorie de la mesure où les résultats sont énoncés dans un cadre très général. Voir aussi [3, 11, 12, 19, 20].

Nous rappelons uniquement quelques conséquences du théorème de Baire sur la continuité des applications linéaires. Mais avant cela, il est important de se remémorer les définitions suivantes :

Définition 1.0.1. (*Espace topologique*) Un **espace topologique** est un couple (X, τ) où :

- X est un ensemble non vide,
- τ est une famille de sous-ensembles de X (appelée **topologie sur X**) qui satisfait les trois axiomes suivants :
 1. **L'ensemble vide et l'ensemble total sont ouverts** : $\emptyset \in \tau$ et $X \in \tau$.
 2. **Stabilité par réunion arbitraire** : toute réunion (finie ou infinie) d'ouverts est un ouvert.
 3. **Stabilité par intersection finie** : toute intersection finie d'ouverts est un ouvert.

Définition 1.0.2. (*Ouvert*) Un sous-ensemble $U \subset X$ est dit **ouvert** s'il appartient à la topologie τ .

Définition 1.0.3. (*Fermé*) Un sous-ensemble $F \subset X$ est dit **fermé** si son complémentaire $X \setminus F$ est un ouvert.

Définition 1.0.4. (*Voisinage*) Soit (X, τ) un espace topologique et soit $x \in X$. Un **voisinage** de x est tout ensemble $V \subset X$ contenant un ouvert U tel que $x \in U$. Autrement dit, il existe un ouvert $U \in \tau$ vérifiant $x \in U \subset V$.

En particulier, un **voisinage ouvert** de x est simplement un ouvert U contenant x .

Définition 1.0.5. (*Intérieur d'un ensemble*) L'intérieur d'un ensemble $A \subset X$, noté $\text{Int}(A)$ ou $\overset{\circ}{A}$, est le plus grand ouvert contenu dans A , c'est-à-dire l'union de tous les ouverts inclus dans A .

Définition 1.0.6. (*Adhérence d'un ensemble*) On appelle *adhérence* (ou *fermeture*) de A noté $\overline{A}$ et rarement $\text{cl}(A)$ (ou $\text{Cl}(A)$), une notation spécifique pour les espaces topologiques) le plus petit fermé contenant A (c'est l'intersection de tous les fermés contenant A).

Définition 1.0.7. (*Espace topologique séparé*) Un espace topologique (X, τ) est dit **séparé** (ou de Hausdorff) si pour tous $x, y \in X$ avec $x \neq y$, il existe deux ouverts disjoints $U, V \in \tau$ tels que $x \in U$ et $y \in V$.

Définition 1.0.8. (*Système fondamental de voisinages*) Un **système fondamental de voisinages** d'un point $x \in X$ est une famille de voisinages $\mathcal{V}_x$ telle que tout voisinage de x contient un élément de $\mathcal{V}_x$.

Définition 1.0.9. (*Espace métrique*) Un **espace métrique** est un couple (X, d) , où X est un ensemble non vide et $d : X \times X \rightarrow \mathbb{R}$ est une fonction appelée **métrique** (ou *distance*) satisfaisant les propriétés suivantes pour tous $x, y, z \in X$:

1. **Séparation** : $d(x, y) = 0$ si et seulement si $x = y$.
2. **Symétrie** : $d(x, y) = d(y, x)$.
3. **Inégalité triangulaire** : $d(x, z) \leq d(x, y) + d(y, z)$.

Définition 1.0.10. (*Boule en espace métrique*). Soit (X, d) un espace métrique et soit $x \in X$. Pour tout $r > 0$, on définit la **boule ouverte** de centre x et de rayon r par :

$$B(x, r) = \{y \in X \mid d(x, y) < r\}.$$

De même, la **boule fermée** de centre x et de rayon r est donnée par :

$$\bar{B}(x, r) = \{y \in X \mid d(x, y) \leq r\}.$$

Définition 1.0.11. (*Espace métrique séparable*) Un espace métrique (X, d) est **séparable** s'il possède un sous-ensemble dense et dénombrable, c'est-à-dire un ensemble $D \subset X$ dénombrable tel que toute boule ouverte contient au moins un élément de D .

Définition 1.0.12. (*Partie grasse d'un espace métrique* :) Soit (X, d) un espace métrique. Une **partie grasse** de X est un sous-ensemble contenant un intérieur non vide, c'est-à-dire un ensemble $A \subset X$ tel que $\text{Int}(A) \neq \emptyset$.

Définition 1.0.13. (*Norme*) Une **norme** sur un espace vectoriel X sur $\mathbb{R}$ ou $\mathbb{C}$ est une application $\|\cdot\| : X \rightarrow \mathbb{R}$ qui satisfait les trois propriétés suivantes pour tous $x, y \in X$ et tout scalaire λ :

1. **Séparation (ou positivité)** :

$$\|x\| \geq 0 \quad \text{et} \quad \|x\| = 0 \iff x = 0.$$

2. **Homogénéité** :

$$\|\lambda x\| = |\lambda| \|x\|.$$

3. **Inégalité triangulaire** :

$$\|x + y\| \leq \|x\| + \|y\|.$$

Définition 1.0.14. (*Espace normé*) Un **espace normé** est un couple $(X, \|\cdot\|)$, où X est un espace vectoriel et $\|\cdot\|$ une norme sur X .

Définition 1.0.15. (*Boule en espace normé*).

Soit $(X, \|\cdot\|)$ un espace normé et soit $x \in X$. Pour tout $r > 0$, on définit :

- **La boule ouverte de centre x et de rayon r :**

$$B(x, r) = \{y \in X \mid \|y - x\| < r\}.$$

- **La boule fermée de centre x et de rayon r :**

$$\bar{B}(x, r) = \{y \in X \mid \|y - x\| \leq r\}.$$

Définition 1.0.16. Un espace topologique (X, τ) est dit **compact** si toute couverture ouverte de X admet une sous-couverture finie.

Autrement dit, si $\{U_i\}_{i \in I}$ est une famille d'ouverts de X telle que :

$$X \subseteq \bigcup_{i \in I} U_i,$$

alors il existe un sous-ensemble fini $\{i_1, i_2, \dots, i_n\} \subset I$ tel que :

$$X \subseteq U_{i_1} \cup U_{i_2} \cup \dots \cup U_{i_n}.$$

Exemple 1.0.17. — Le segment $[a, b]$ dans $\mathbb{R}$ muni de la topologie usuelle est compact d'après le théorème de Heine-Borel, car il est fermé et borné.

- L'intervalle $(0, 1)$ dans $\mathbb{R}$ n'est pas compact, car il peut être recouvert par une infinité d'intervalles ouverts sans qu'un nombre fini suffise.
- L'espace discret $X = \{1, 2, 3\}$ muni de la topologie discrète est compact, car toute couverture ouverte est triviale et admet une sous-couverture finie.

Définition 1.0.18. (*Suite de Cauchy*) Une suite $(x_n)_{n \in \mathbb{N}}$ dans (X, d) est dite **de Cauchy** si :

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall m, n \geq N, \quad d(x_m, x_n) < \varepsilon.$$

Définition 1.0.19. Un espace métrique (X, d) est **complet** si toute suite de Cauchy dans X converge vers un élément de X .

Exemple 1.0.20. — $(\mathbb{R}, d)$ est complet avec la distance usuelle.

- $(\mathbb{Q}, d)$ n'est pas complet car la suite des décimales successives de $\sqrt{2}$ est de Cauchy mais ne converge pas dans $\mathbb{Q}$.
- L'espace ℓ^2 des suites (x_n) telles que $\sum_{n=1}^{\infty} |x_n|^2 < \infty$ est un espace métrique complet.

Définition 1.0.21. Un espace topologique X est dit **localement compact** si, pour tout point $x \in X$, il existe un voisinage ouvert U de x tel que U soit compact dans X . Autrement dit, chaque point de X a un voisinage qui est compact.

Définition 1.0.22. Un **homéomorphisme** est une application bijective $f : X \rightarrow Y$ entre deux espaces topologiques $(X, \mathcal{T}_X)$ et $(Y, \mathcal{T}_Y)$ telle que f et f^{-1} sont toutes deux continues. Autrement dit :

- f est une bijection ;

- f est continue, c'est-à-dire que pour tout ouvert $U \subset Y$, $f^{-1}(U)$ est ouvert dans X ;
- f^{-1} est continue, c'est-à-dire que pour tout ouvert $V \subset X$, $f(V)$ est ouvert dans Y .

Deux espaces topologiques X et Y sont dits **homéomorphes** s'il existe un homéomorphisme entre eux.

Définition 1.0.23. Un **difféomorphisme** est une application bijective $f : M \rightarrow N$ entre deux variétés différentielles M et N telle que f et f^{-1} sont toutes deux de classe C^∞ (ou de classe C^k pour une certaine régularité $k \geq 1$). Autrement dit :

- f est une bijection ;
- f est C^k , c'est-à-dire que toutes les dérivées partielles jusqu'à l'ordre k existent et sont continues ;
- f^{-1} est C^k ;
- La différentielle df est inversible en tout point.

Deux variétés différentielles M et N sont dites **difféomorphes** s'il existe un difféomorphisme entre elles.

1.1 Espaces de Baire

Définition 1.1.1. Soit X un espace topologique.

- X est dit de Baire si et seulement si toute intersection dénombrable d'ouverts denses dans X est dense dans X (ou de façon équivalente, toute réunion dénombrable de fermés d'intérieurs vides est d'intérieur vide).
- Une partie $A \subset X$ est maigre si et seulement si A est incluse dans une réunion dénombrable de fermés d'intérieurs vides.
- Une propriété est vraie Baire presque partout si et seulement si elle est vraie sur le complémentaire d'une partie maigre, i.e. sur un ensemble gras (ou résiduel).

Théorème 1.1.2. (de Baire 1ère version). -Dans un espace métrique complet, la réunion d'une famille dénombrable de fermés d'intérieur vide est encore d'intérieur vide.

Théorème 1.1.3. (de Baire 2ème version) Si X est un espace métrique complet, l'intersection d'une famille dénombrable d'ouverts denses est encore dense dans X .

Nous allons démontrer la deuxième version seulement. Il s'agit de montrer que tout espace métrique complet est un espace de Baire. Rappelons qu'un sous ensemble A de X est dense dans $X \iff \overline{A} = X$. Autrement dit, tout point de X est une limite de points de A , ou encore, A rencontre toute boule ouverte non vide de X .

Démonstration: Soit $(O_n)_{n \geq 1}$ une suite d'ouverts denses, et soit O un ouvert non vide arbitraire de X . Notons

$$G = \bigcap_{n \geq 1} O_n.$$

On va montrer que :

$$O \cap G \neq \emptyset,$$

ce qui prouvera que G est dense dans X .

On notera respectivement $B_f(x, r)$ et $B_o(x, r)$ les boules fermée et ouverte de centre x et

de rayon r dans l'espace métrique X .

Soit $x_0 \in O$ et $r_0 > 0$ tels que :

$$B_f(x_0, r_0) \subset O.$$

Comme O_1 est dense, l'ouvert $O_1 \cap B_o(x_0, r_0)$ n'est pas vide. Il existe donc $x_1 \in O_1 \cap B_o(x_0, r_0)$ et $r_1 > 0$ tels que :

$$B_f(x_1, r_1) \subset O_1 \cap B_o(x_0, r_0) \subset O_1 \cap O,$$

et l'on peut prendre $r_1 \leq \frac{r_0}{2}$.

En continuant ainsi, on obtient une suite $(x_n) \subset O_n$ et des rayons $r_n > 0$ tels que :

$$B_f(x_{n+1}, r_{n+1}) \subset O_{n+1} \cap B_o(x_n, r_n) \subset O_{n+1} \cap O_n \cap \cdots \cap O_1 \cap O,$$

avec la propriété :

$$0 < r_{n+1} \leq \frac{r_n}{2}.$$

Alors la suite (x_n) est de Cauchy. Comme X est complet, elle possède une limite l .

De plus, pour tout $n, p \geq 1$, on a

$$x_{n+p} \in B_0(x_n, r_n).$$

Il s'en suit que

$$l \in B_f(x_n, r_n),$$

et comme $B_f(x_n, r_n) \subset O_n \cap \cdots \cap O_1 \cap O$, on obtient finalement :

$$l \in G \cap O.$$

□

Ce résultat permet de montrer par exemple que l'ensemble des points de continuité de la limite simple f d'une suite de fonctions continues (f_n) de $C^0([0, 1])$ est un ensemble gras de $[0, 1]$.

On utilise pour cela la propriété suivante : s'il existe une suite (F_n) de parties de l'espace de Baire X telle que $X \subset \bigcup_{n \in \mathbb{N}} F_n$, alors

$$\bigcup_{n \in \mathbb{N}} \overset{\circ}{F}_n \supset X \setminus \left(\bigcup_{n \in \mathbb{N}} \partial F_n \right) \quad \text{est dense dans } X.$$

Pour tout $\varepsilon > 0$, on pose

$$F_n^\varepsilon = \{x \in [0, 1] / \forall m \geq n, \quad |f_n(x) - f_m(x)| \leq \varepsilon\}.$$

Par définition, F_n^ε est fermé et $\bigcup_{n \in \mathbb{N}} F_n^\varepsilon = [0, 1]$.

La suite

$$O_p = \bigcup_n \overset{\circ}{F}_n^{1/p}$$

est alors une suite d'ouverts denses de $[0, 1]$, et le théorème de Baire montre que $G = \bigcap_p O_p$, est un ensemble gras.

Il suffit alors de vérifier que les éléments de G sont des points de continuité de f .

Remarque 1.1.4. - Attention à ne pas confondre les notions de " presque partout " au sens de Baire et au sens des mesures.

Ainsi, si μ est une mesure borélienne qui ne charge pas les points sur un espace métrique séparable X , il existe toujours une partie grasse de X qui est μ -négligeable.

Il suffit pour construire une telle partie de prendre une suite dense (x_i) , de poser pour tout $k \in \mathbb{N}^*$

$$V^k = \bigcup O_i^k \text{ où } O_i^k \text{ est un ouvert contenant } x_i \text{ tel que } \mu(O_i^k) \leq \frac{2^{-i}}{k}$$

puis de considérer $G = \bigcap_{k=1}^{\infty} V^k$.

1.2 Opérateurs linéaires et théorème de Banach-Steinhaus

Théorème 1.2.1. (de Banach-Steinhaus). - Soient E un espace de Banach et F un espace vectoriel normé. On considère une famille (T_α) d'applications linéaires continues de E dans F telles que

$$\forall x \in E, \quad \sup_{\alpha} \|T_\alpha x\|_F < +\infty$$

Alors il existe $C > 0$ tel que

$$\forall x \in E, \forall \alpha, \quad \|T_\alpha x\|_F \leq C \|x\|_E.$$

Démonstration: Posons :

$$\Phi_n = \{x \in E \mid \forall \alpha \in I, \|T_\alpha x\| \leq n\} = \bigcap_{\alpha \in I} \{x \in E \mid \|T_\alpha x\| \leq n\}$$

Chaque Φ_n est fermé dans E car les T_α sont continues. De plus, l'hypothèse implique que :

$$E = \bigcup_{n=1}^{\infty} \Phi_n.$$

Puisque E est un espace de Banach, le théorème de Baire assure l'existence de $N \geq 1$ tel que Φ_N possède une partie intérieure non vide. Ainsi, il existe $x_0 \in E$ et $r_0 > 0$ tels que :

$$B_E(x_0, r_0) \subset \Phi_N.$$

Alors, pour tout $y \in B_E(0, 1)$,

$$\|T_\alpha(x_0 + r_0 y)\|_F \leq N, \quad \forall \alpha \in I.$$

Ce qui entraîne :

$$\|T_\alpha y\|_F \leq \frac{1}{r_0} \left(N + \sup_{\alpha \in I} \|T_\alpha x_0\|_F \right), \quad \forall y \in B_E(0, 1).$$

Il en résulte que :

$$\sup_{\alpha \in I} \|T_\alpha\| < +\infty.$$

□

Corollaire 1.2.2. Soient E un espace de Banach et F un espace vectoriel normé. On considère une famille (T_α) d'applications linéaires continues de E dans F convergeant simplement vers T .

Alors T est linéaire continu de E dans F . De plus, si $x_n \rightarrow x$ dans E , alors $T_n(x_n) \rightarrow T(x)$ dans F .

Le théorème de Banach-Steinhaus peut en fait s'étendre à des espaces plus généraux, que nous utiliserons dans la suite pour munir l'espace des distributions d'une topologie.

Définition 1.2.3. Soit E un espace vectoriel. On appelle semi-norme sur E toute application $p : E \rightarrow \mathbb{R}^+$ sous-additive et positivement homogène, i.e. telle que

$$\begin{aligned} \forall x, y \in E, \quad p(x + y) &\leq p(x) + p(y) \\ \forall \lambda \in \mathbb{R}, \forall x \in E, \quad p(\lambda x) &= |\lambda|p(x) \end{aligned}$$

On dit qu'une famille $\mathcal{P}$ de semi-normes sépare les points si

$$p(x) = 0 \text{ pour tout } p \in \mathcal{P} \Rightarrow x = 0.$$

Les ouverts de la topologie associée à $\mathcal{P}$ sont les parties U de E telles que, pour tout $x \in U$, il existe un ensemble fini J et $r > 0$

$$B_J(x, r) \equiv \{y \in E / \forall j \in J, p_j(x - y) < r\} \subset U$$

Définition 1.2.4. Un espace vectoriel E est un pré-Fréchet si et seulement si il existe une famille dénombrable de semi-normes $(p_j)_{j \in \mathbb{N}^*}$ telles que, pour tout $x \in E$

$$\begin{cases} \text{si } p_j(x) = 0 \text{ pour tout } j, \text{ alors } x = 0, \\ \text{pour tout } j, \quad p_j(x) \leq p_{j+1}(x). \end{cases}$$

La topologie d'un pré-Fréchet est métrisable avec la distance (invariante par translation)

$$d(x, y) = \sum_{j=1}^{+\infty} 2^{-j} \min(1, p_j(x - y)).$$

Exercice Montrer que

$$d(x, y) = \sum_{j=1}^{+\infty} 2^{-j} \min(1, p_j(x - y))$$

est une distance.

Définition 1.2.5. Un espace de Fréchet est un pré-Fréchet complet.

L'ensemble des fonctions $C^\infty(K)$ sur un compact K est un espace de Fréchet, muni des semi-normes

$$p_j(f) = \sup_{|\alpha| \leq j} \sup_{x \in K} |D^\alpha f(x)|$$

Une "bonne" topologie sur $C_c^\infty(\Omega)$ (où Ω est un ouvert) est beaucoup plus subtile à définir. On y reviendra brièvement dans le chapitre sur les distributions.

La continuité d'une application entre deux pré-Fréchet peut s'exprimer à l'aide des semi-normes

Lemme 1.2.6. Soient $(E, (p_j)_{j \in \mathbb{N}^*})$ et $(F, (q_j)_{j \in \mathbb{N}^*})$ deux pré-Fréchet, et T une application linéaire de E dans F . Alors T est continue si et seulement si

$$\forall j \in \mathbb{N}^*, \exists C \geq 0, k \in \mathbb{N}^*, \quad q_j(Tx) \leq Cp_k(x). \quad (1.2.1)$$

Démonstration: Si T est continue, pour j fixé, il existe un voisinage U de 0 tel que

$$\forall x \in U, \quad q_j(Tx) \leq 1$$

Il existe alors $k \in \mathbb{N}^*$ et $\varepsilon > 0$ tels que

$$B_k^E(0, \varepsilon) \subset U$$

Par homogénéité, on en déduit la condition (1.2.1) (on peut considérer séparément les cas où $p_k(x) = 0$ et $p_k(x) > 0$).

Réciproquement, supposons que (1.2.1) soit satisfaite. Soit $x \in E$. Pour tout voisinage U de Tx dans F , il existe une boule $B_j^F(Tx, \varepsilon)$ incluse dans U , Il découle alors de (1.2.1) que pour tout y dans $B_k(x, \varepsilon/C)$

$$Ty \in B_j^F(Tx, \varepsilon) \subset U$$

ce qui montre que T est continue. □

On a alors la généralisation suivante du théorème de Banach-Steinhaus.

Théorème 1.2.7. Soient E un espace de Fréchet et F un pré-Fréchet. On considère une suite (T_n) d'applications linéaires continues de E dans F convergeant simplement vers T .

Alors T est linéaire continu de E dans F . De plus, si $x_n \rightarrow x$ dans E , alors $T_n(x_n) \rightarrow T(x)$ dans F .

1.3 Théorèmes de l'application ouverte et du graphe fermé

Théorème 1.3.1. (de l'application ouverte) Soient E et F deux espaces de Banach, et T un opérateur linéaire continu de E dans F .

- Si T est surjectif, alors T envoie les ouverts de E sur les ouverts de F .
- Si T est bijectif, T^{-1} est un opérateur linéaire continu de F dans E .

Démonstration: On cherche $\delta > 0$ tel que $B_F(0, \delta) \subset T(B_E(0, 1))$.

On commence par définir

$$F_n = \overline{T(B_E(0, n))} \text{ de sorte que } \bigcup_{n=1}^{+\infty} F_n = F$$

D'après le théorème de Baire, il existe donc n_0 tel que F_{n_0} est d'intérieur non vide

$$B(x_0, r_0) \subset F_{n_0}.$$

Par linéarité, on a aussi

$$B(-x_0, r_0) \subset F_{n_0}.$$

Si $\|y\|_F < r_0$,

$$y = \frac{1}{2}(y_0 + y) + \frac{1}{2}(-y_0 + y) \in \overline{T(B_E(0, n_0))}.$$

Autrement dit,

$$B_F\left(0, \frac{r_0}{n_0}\right) \subset \overline{T(B_E(0, 1))}.$$

Soit maintenant $y \in B_F\left(0, \frac{r_0}{10n_0}\right)$. D'après ce qui précède, il existe $x_0 \in B_E\left(0, \frac{1}{10}\right)$ tel que

$$\|Tx_0 - y\|_F \leq \frac{r_0}{10n_0} \times \frac{1}{2}.$$

Par récurrence, on construit une suite x_n telle que

$$\|x_n\|_E \leq \frac{1}{10} \frac{1}{2^n}, \quad \|(Tx_0 + \dots + Tx_n) - y\|_F \leq \frac{r_0}{10n_0} \times \frac{1}{2^{n+1}}$$

La série $\sum_n x_n$ est donc normalement convergente et, par continuité de T , sa somme x vérifie $Tx = y$.

On en déduit que

$$B_F\left(0, \frac{r_0}{10n_0}\right) \subset T(B_E(0, n_0))$$

ce qui montre que l'image de tout ouvert de E est ouverte. $\square$

Théorème 1.3.2. (du graphe fermé) Soient E et F deux espaces de Banach, et T un opérateur continu de E dans F .

Alors T est continu si et seulement si son graphe $G = \{(x, Tx)/x \in E\}$ est fermé dans $E \times F$.

Démonstration: On introduit sur E la norme du graphe

$$\|x\|_G = \|x\|_E + \|Tx\|_F.$$

- Si T est continu, son graphe est toujours fermé : si $(x, y) \in \bar{G}$, il existe une suite (x_n) de E telle que $(x_n, Tx_n) \rightarrow (x, y)$, et par continuité $y = Tx$.
- Supposons maintenant que $G \subset E \times F$ soit fermé. En particulier, G est un espace de Banach. On considère alors l'application

$$\Pi : (x, Tx) \in G \mapsto x \in E$$

qui est clairement continue et bijective. D'après le théorème de l'application ouverte, Π est un homéomorphisme, ce qui implique que T est continu. $\square$

Ces résultats peuvent aussi se généraliser au cas des espaces de Fréchet.

Théorème 1.3.3. *Soient E et F deux espaces de Fréchet, et T un opérateur continu de E dans F .*

- *Si $T(E)$ n'est pas maigre dans F , alors T est une application surjective et ouverte.*
- *Si T est bijectif, c'est un homéomorphisme.*

1.4 Rappels des résultats importants de la théorie de la mesure

1.4.1 Quelques définitions fondamentales

Définition 1.4.1. (Tribus ou σ -algèbres) Une **tribu** $\mathcal{A}$ sur X vérifie :

- $X \in \mathcal{A}$,
- Stabilité par complémentation et unions dénombrables.

Exemple 1.4.2. Tribu borélienne $\mathcal{B}(\mathbb{R})$.

Définition 1.4.3. Une **mesure** $\mu : \mathcal{A} \rightarrow [0, +\infty]$ vérifie :

- $\mu(\emptyset) = 0$,
- σ -additivité : $\mu(\bigcup A_n) = \sum \mu(A_n)$.

Définition 1.4.4. Pour $f : X \rightarrow \mathbb{R}$ mesurable :

$$\int f \, d\mu = \sup \left\{ \int \phi \, d\mu \mid \phi \text{ étagée}, 0 \leq \phi \leq f \right\}.$$

Exercice : Calculer $\int_{\mathbb{R}} e^{-x^2} \, d\delta_0(x)$. **Réponse :** 1.

1.4.2 Théorèmes majeurs de la théorie de la mesure

Théorème 1.4.5. (Théorème de Convergence Monotone de Beppo Levi) Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions mesurables positives croissantes sur un espace mesuré $(X, \mathcal{A}, \mu)$. Si $f = \lim_{n \rightarrow \infty} f_n$ existe, alors :

$$\int_X f_n \, d\mu \rightarrow \int_X f \, d\mu.$$

Démonstration : Par croissance de f_n , on applique la monotonie de l'intégrale et passe à la limite. $\square$

Lemme 1.4.6. (de Fatou) Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions mesurables positives sur X , alors :

$$\int_X \liminf_{n \rightarrow \infty} f_n d\mu \leq \liminf_{n \rightarrow \infty} \int_X f_n d\mu.$$

Démonstration : Appliqué via une approximation par des fonctions simples et le théorème de convergence monotone. $\square$

Théorème 1.4.7. (Théorème de Convergence Dominée de Lebesgue) Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions mesurables sur $(X, \mathcal{A}, \mu)$ qui converge presque partout vers f . Si une fonction intégrable g majore uniformément $|f_n|$, alors :

$$\lim_{n \rightarrow \infty} \int_X f_n d\mu = \int_X f d\mu.$$

Démonstration : On applique le lemme de Fatou aux fonctions $g - f_n$ et $g + f_n$, puis on utilise la linéarité de l'intégrale. $\square$

Théorème 1.4.8. (Théorème de Tonelli) Si f est une fonction mesurable positive sur $X \times Y$, alors :

$$\int_X \left(\int_Y f(x, y) d\nu(y) \right) d\mu(x) = \int_Y \left(\int_X f(x, y) d\mu(x) \right) d\nu(y).$$

Démonstration : Résultat basé sur une approximation par des fonctions simples et application du théorème de convergence monotone. $\square$

Théorème 1.4.9. (Théorème de Fubini) Si f est intégrable sur $X \times Y$, alors :

$$\int_{X \times Y} f d(\mu \times \nu) = \int_X \left(\int_Y f(x, y) d\nu(y) \right) d\mu(x) = \int_Y \left(\int_X f(x, y) d\mu(x) \right) d\nu(y).$$

Démonstration : Décomposition de f en parties positives et négatives et application du théorème de Tonelli. $\square$

1.4.3 Résumé sur les Espaces L^p

1.4.3.1 Définitions

Soit $(X, \mathcal{A}, \mu)$ un espace mesuré.

Définition 1.4.10 (Norme L^p et espace L^p). Pour $1 \leq p < \infty$, on définit l'espace $L^p(X, \mathcal{A}, \mu)$ par :

$$L^p(X, \mathcal{A}, \mu) = \{f \text{ mesurable} \mid \|f\|_p < \infty\},$$

où la norme L^p est donnée par :

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

Pour $p = \infty$, on définit :

$$L^\infty(X, \mathcal{A}, \mu) = \{f \text{ mesurable} \mid \|f\|_\infty < \infty\},$$

où la norme essentielle est :

$$\|f\|_\infty = \inf\{M \geq 0 \mid |f(x)| \leq M \text{ p.p.}\}.$$

1.4.3.2 Propriétés fondamentales des espaces L^p

Théorème 1.4.11 (Riesz-Fischer). *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré. Pour tout $1 \leq p < \infty$, l'espace $L^p(X, \mathcal{A}, \mu)$ muni de la norme*

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

est un espace de Banach

Autrement-dit, toute suite de Cauchy dans L^p converge vers un élément de L^p .

Démonstration: Soit (f_n) une suite de Cauchy définie dans L^p . Puisque (f_n) est de Cauchy en norme L^p , on peut montrer que la suite $f_n(x)$ possède une sous-suite convergente presque partout (grâce au théorème de Banach-Alaoglu et à une extraction diagonale). On définit alors une fonction limite f et on montre qu'elle appartient à L^p , c'est-à-dire que

$$\int_X |f(x)|^p d\mu(x) < \infty.$$

L'étape finale consiste à démontrer que $f_n \rightarrow f$ en norme L^p , c'est-à-dire :

$$\|f_n - f\|_p \rightarrow 0.$$

Ceci est obtenu grâce à une combinaison de la convergence dominée de Lebesgue et de la propriété de Cauchy.

Ainsi, l'espace L^p est complet, ce qui conclut la preuve du théorème de Riesz-Fischer.

Théorème 1.4.12 (Complétude). *L'espace $L^p(X, \mathcal{A}, \mu)$ est un espace de Banach. Si $p = 2$, c'est un espace de Hilbert avec le produit scalaire :*

$$\langle f, g \rangle = \int_X f(x)g(x) d\mu(x).$$

Théorème 1.4.13 (Inégalité de Hölder). *Soient $f \in L^p$ et $g \in L^q$ avec $\frac{1}{p} + \frac{1}{q} = 1$, alors :*

$$\int_X |fg| d\mu \leq \|f\|_p \|g\|_q.$$

Théorème 1.4.14 (Inégalité de Minkowski). *Pour $f, g \in L^p(X)$, on a :*

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

Proposition 1.4.15 (Inclusion entre espaces L^p). *Si $\mu(X) < \infty$ et $p < q$, alors $L^q \subset L^p$ et :*

$$\|f\|_p \leq \mu(X)^{\frac{1}{p} - \frac{1}{q}} \|f\|_q.$$

1.5 Exercices

Exercice 1. Montrer que $B_\infty = \{f \in L^2([0, 1]); |f| \leq 1 \text{ presque partout} \}$ est fermé dans $L^2([0, 1])$.

Exercice 2. Soit $\delta_0 : \mathcal{C}([0, 1]) \rightarrow \mathbb{K}$ la forme linéaire définie par $\delta_0(f) = f(0)$. Montrer qu'elle est continue pour la norme $\|\cdot\|_\infty$, mais pas pour la norme $\|\cdot\|_1$.

Exercice 3.

1. Montrer que c_0 est fermé dans ℓ_∞ .
2. Montrer que les espaces c_0 et $\ell_p, 1 \leq p \leq \infty$ sont complets.
3. Montrer que $(\mathcal{C}([0, 1]), \|\cdot\|_1)$ n'est pas complet.
4. On rappelle qu'une fonction $\varphi : I \rightarrow \mathbb{R}$ définie sur un intervalle I de $\mathbb{R}$ est convexe si $\varphi(tu + (1-t)v) \leq t\varphi(u) + (1-t)\varphi(v)$ pour tous $u, v \in I$ et $t \in [0, 1]$.

Soit $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ une fonction convexe strictement croissante telle que $\varphi(0) = 0$.

a) Montrer que φ est continue sur $\mathbb{R}_+$ et bijective.

On note ℓ_φ l'ensemble des suites de nombres complexes $x = (x_n)_{n \geq 1}$ pour lesquelles il existe $C > 0$ tel que $\sum_{n=1}^\infty \varphi(|x_n|/C) < +\infty$ et on pose :

$$\|x\|_\varphi = \inf \left\{ C > 0; \sum_{n=1}^\infty \varphi(|x_n|/C) \leq 1 \right\}$$

b) Montrer que ℓ_φ est un espace vectoriel et que $\|\cdot\|_\varphi$ est une norme sur ℓ_φ (on montrera que $\sum_{n=1}^\infty \varphi(|x_n|/\|x\|_\varphi) \leq 1$. On dit que ℓ_φ est un espace d'Orlicz.

c) Montrer que ℓ_φ est complet.

Exercice 4. Soit (X, d) un espace métrique. On note $\text{Lip}(X)$ l'ensemble des fonctions lipschitziennes, à valeurs réelles, sur X . Pour $f \in \text{Lip}(X)$, on note $\text{Lip}(f)$ la constante de Lipschitz de f , à savoir $\text{Lip}(f) = \sup \left\{ \frac{|f(x)-f(y)|}{d(x,y)}; x, y \in X, x \neq y \right\}$.

1. Montrer que $\text{Lip}(X)$ est un espace vectoriel.
2. Soit $a \in X$. Pour $f \in \text{Lip}(X)$, on pose $\|f\|_{\text{Lip},a} = |f(a)| + \text{Lip}(f)$.
 - a) Montrer que $\|\cdot\|_{\text{Lip},a}$ est une norme sur $\text{Lip}(X)$.
 - b) Montrer que si b est un autre point de X , alors le norme $\|\cdot\|_{\text{Lip},b}$ est équivalente à la norme $\|\cdot\|_{\text{Lip},a}$.
 - c) On suppose que X est un espace vectoriel normé (et que d est la distance associée à la norme de X). Montrer que toute forme linéaire continue sur X est dans $\text{Lip}(X)$, et que l'on a $\|\varphi\|_{\text{Lip},0} = \|\varphi\|_{X^*}$ pour toute $\varphi \in X^*$.
3. On revient au cas d'un espace métrique (X, d) quelconque. Montrer que $\text{Lip}(X)$ est complet pour n'importe quelle norme $\|f\|_{\text{Lip},a}$.

Exercice 5. Soit $\mathcal{C}^1([0, 1])$ l'espace vectoriel (réel) des fonctions $f : [0, 1] \rightarrow \mathbb{R}$ continûment dérivables. On pose, pour $f \in \mathcal{C}^1([0, 1])$: $\|f\| = \left([f(0)]^2 + \int_0^1 |f'(t)|^2 dt \right)^{1/2}$.

1. Montrer que l'on définit ainsi une norme sur $\mathcal{C}^1([0, 1])$.
2. Montrer que si la suite $(f_n)_{n \geq 1}$ converge pour cette norme, alors elle converge uniformément sur $[0, 1]$.
3. On pose $f_n(t) = t^n(1-t), n \geq 1$. Calculer $\|f_n\|$.

Exercice 6.

1. Soit $\mathcal{C}_b(\mathbb{R})$ l'espace des fonctions continues bornées sur $\mathbb{R}$ muni de la norme uniforme, i.e

$$\|f\|_\infty = \sup_{x \in \mathcal{C}_b(\mathbb{R})} |f(x)|.$$

On considère l'application linéaire $T : \mathcal{C}_b(\mathbb{R}) \rightarrow \mathcal{C}_b(\mathbb{R})$ définie par $(Tf)(x) = 3f(x) - 2f(x+4)$. Montrer que T est continue et calculer $\|T\|$.

2. On considère la forme linéaire $\varphi : c_0 \rightarrow \mathbb{R}$ définie par $\varphi(u) = \sum_{n=1}^{\infty} \frac{u_n}{2^n}$. Montrer qu'elle est continue et calculer sa norme. Montrer que $\sup_{\|x\|_{\infty} \leq 1} |\varphi(x)|$ n'est pas atteint.

Exercice 7. On identifie l'espace $M_d(\mathbb{R})$ des matrices réelles d'ordre d avec l'espace des applications linéaires (automatiquement continues) de $\mathbb{R}^d$ dans lui-même. Pour $1 \leq p \leq \infty$, on note $\|\cdot\|_{p,p}$ la norme sur $M_d(\mathbb{R})$ associée à la norme $\|\cdot\|_p$ sur $\mathbb{R}^d$.

1. Montrer que si $A = (a_{i,j})_{1 \leq i,j \leq d} \in M_d(\mathbb{R})$, alors $\|A\|_{1,1} = \max_{1 \leq j \leq d} \sum_{i=1}^d |a_{i,j}|$.
2. Déterminer de même $\|A\|_{\infty,\infty}$ en fonction des coefficients de A .
3. Pour $A \in M_d(\mathbb{R})$, on note A^* la transposée de A . Montrer que $\|A\|_{2,2} = \sup \left\{ \sqrt{\lambda}; \lambda \text{ valeur propre de } A^*A \right\}$ (utiliser une base orthonormée de vecteurs propres de A^*A).

Exercice 8. Montrer que sur tout espace normé de dimension infinie, il existe des formes linéaires non continues (on commencera par considérer une suite de vecteurs linéairement indépendants et on construira une forme linéaire non continue sur le sous-espace vectoriel qu'ils engendrent).

Exercice 8. Montrer qu'un espace vectoriel normé est complet si et seulement si toute série absolument convergente, c'est-à-dire telle que $\sum_{n=1}^{\infty} \|x_n\| < +\infty$ (on dit aussi série normalement convergente) est convergente (on montrera que de toute suite de Cauchy $(u_n)_{n \geq 1}$, on peut extraire une sous-suite $(u_{n_k})_{k \geq 1}$ telle que $\sum_{k=1}^{\infty} \|u_{n_{k+1}} - u_{n_k}\| < +\infty$).

Exercice 10. Soit X un espace de Banach, Y un espace vectoriel normé et $T : X \rightarrow Y$ une application linéaire continue. On suppose qu'il existe une constante $c > 0$ telle que $\|T(x)\| \geq c\|x\|$ pour tout $x \in X$. Montrer que $\text{im}(T)$ est fermée dans Y , et que T réalise un isomorphisme entre X et $\text{im}(T)$.

Exercice 11. Soit $(S, \mathcal{T}, m)$ un espace mesuré et $(f_n)_n$ une suite bornée dans $L^1(m)$: il existe un nombre $M > 0$ tel que $\|f_n\|_1 \leq M$ pour tout $n \geq 1$. On suppose que $(f_n)_n$ converge m -presque partout vers une fonction f .

1. Montrer que $f \in L^1(m)$.
2. Montrer que si $(f_n)_n$ converge pour la norme $\|\cdot\|_1$, c'est forcément vers f . Donner un exemple pour lequel $(f_n)_n$ ne converge pas pour la norme $\|\cdot\|_1$.
3. Montrer que, en utilisant le Théorème de convergence dominée, que :

$$\lim_{n \rightarrow \infty} \int_S (|f_n - f| - |f_n| + |f|) dm = 0$$

4. Dédire de la question 3) que, si l'on suppose de plus que $\|f_n\|_1 \xrightarrow{n \rightarrow \infty} \|f\|_1$, alors $\|f_n - f\|_1 \xrightarrow{n \rightarrow \infty} 0$.

Théorème de Hahn-Banach

Le théorème de Hahn-Banach donne l'existence de prolongements de formes linéaires satisfaisant à certaines conditions de norme.

- C'est un outil fondamental de l'analyse fonctionnelle car il permet d'introduire des topologies faibles (définies par dualité) pour lesquelles on a de bonnes propriétés de compacité, ce qui fera l'objet du prochain chapitre.
- Par son interprétation géométrique en termes d'hyperplans évitant un convexe fixé, il joue galement un rôle primordial dans l'étude de la géométrie des convexes, et plus généralement en analyse convexe [7].

La preuve repose d'une part sur un argument simple permettant de prolonger une forme linéaire définie sur un sous-espace de codimension 1, et d'autre part sur le lemme de Zorn (équivalent à l'axiome du choix) qui permet d'obtenir une récurrence transfinie.

2.1 Forme analytique

Théorème 2.1.1. *Soient E un espace vectoriel réel, et $p : E \rightarrow \mathbb{R}$ une application sous-additive, positivement homogène, i.e. telle que*

$$\forall x \in E, \quad \forall \alpha \geq 0, \quad p(\alpha x) = \alpha p(x) \tag{2.1}$$

$$\forall x, y \in E, \quad p(x + y) \leq p(x) + p(y) \tag{2.2}$$

Si F est un sous-espace de E , et f est une forme linéaire sur F telle que

$$\forall x \in F, \quad f(x) \leq p(x)$$

il existe un prolongement linéaire $\tilde{f}$ de f à E tel que

$$\forall x \in E, \quad \tilde{f}(x) \leq p(x).$$

La démonstration de ce théorème repose sur le lemme de Zorn que nous allons énoncer sans démonstration.

Démonstration: On considère l'ensemble $\mathcal{P}$ des prolongements de f constitué des paires (G, g) où G est un sous-espace de E contenant F et g une forme linéaire sur G tels que

$$G|_F = f \text{ et } \forall x \in G, \quad g(x) \leq p(x)$$

- On vérifie sans difficulté que
- $\mathcal{P} \neq \emptyset$ car $(F, f) \in \mathcal{P}$;
 - $\mathcal{P}$ est ordonné par la relation d'ordre partiel

$$(G_1, g_1) \prec (G_2, g_2) \text{ si } G_1 \subset G_2 \text{ et } g_2|_{G_1} = g_1;$$

— $\mathcal{P}$ est inductif (toute partie totalement ordonnée admet un majorant).

D'après le lemme de Zorn, il existe alors un élément maximal $(\bar{G}, \bar{g}) \in \mathcal{P}$. Si $\bar{G} \neq E$, il existe $x_0 \in E \setminus \bar{G}$. On cherche alors $\tilde{g}(x_0)$ de sorte à ce que l'application $\tilde{g}$ définie sur $\tilde{G} = \bar{G} + \mathbb{R}x_0$ par

$$\tilde{g}(x + \alpha x_0) = \bar{g}(x) + \alpha \tilde{g}(x_0),$$

satisfasse la majoration attendue : pour tout $x \in \tilde{G}$ et tout $t \in \mathbb{R}$

$$\bar{g}(x) + \alpha \tilde{g}(x_0) \leq p(x + \alpha x_0).$$

Cela implique en particulier que

$$\sup_{x \in E} (\bar{g}(x) - p(x - x_0)) \leq \tilde{g}(x_0) \leq \inf_{x \in E} (p(x + x_0) - \bar{g}(x)).$$

Comme $\tilde{g}(x) + \tilde{g}(x') \leq p(x + x') \leq p(x + x_0) + p(x' - x_0)$, cette condition est satisfaite, et on vérifie qu'elle suffit à obtenir la majoration requise pour $\tilde{g}$. En particulier $(\tilde{G}, \tilde{g}) \in \mathcal{P}$, ce qui contredit la maximalité de $(\bar{G}, \bar{g})$. $\square$

Sous des hypothèses supplémentaires sur l'espace E , on peut démontrer un résultat analogue sans utiliser le lemme de Zorn. C'est le cas par exemple si E est de dimension finie, si E est un espace de Hilbert (il suffit alors d'utiliser le théorème de projection sur les fermés), si E est un espace vectoriel normé séparable...

2.2 Formes géométriques

On rappelle qu'un espace vectoriel topologique est un espace vectoriel muni d'une topologie rendant continues l'addition et la multiplication par un scalaire.

Définition 2.2.1. *Un espace vectoriel topologique est dit localement convexe si et seulement si tout voisinage de l'origine contient un voisinage convexe.*

En associant à chaque voisinage convexe de l'origine une semi-norme (jauge de Minkowski), on a alors la caractérisation suivante :

Propriété 2.2.2. *Soit E un espace vectoriel topologique. E est localement convexe si et seulement si sa topologie peut être définie par une famille de semi-normes (p_α) telles que*

$$(\forall \alpha, \quad p_\alpha(x) = 0) \Rightarrow x = 0$$

Sans perte de généralité, on peut supposer de plus que cette famille est filtrante, c'est-à-dire que

$$\forall \alpha_1, \dots, \alpha_n, \quad \exists \beta, \quad p_\beta \geq p_{\alpha_1} + \dots + p_{\alpha_n}$$

Une partie A de E est alors ouverte si et seulement si

$$\forall x \in A, \quad \exists \alpha, \delta, \quad B_\alpha(x, \delta) \subset A.$$

Les espaces vectoriels normés et les espaces pré-Fréchet sont des espaces vectoriels topologiques localement convexes.

Par contre, l'espace des fonctions continues sur $[0, 1]$ muni de la distance

$$d(f, g) = \int_0^1 \sqrt{|f - g|}$$

est un espace vectoriel topologique non localement convexe.

Théorème 2.2.3. Soit E un espace vectoriel topologique localement convexe (de seminormes (p_α)). Si A et B sont deux parties convexes non vides disjointes, alors

- (1ère forme géométrique) si A est ouverte, il existe une forme linéaire $f : E \rightarrow \mathbb{R}$ continue non nulle telle que

$$\sup_A f \leq \inf_B f.$$

- (2ème forme géométrique) si A est compacte et B est fermée, il existe une forme linéaire $f : E \rightarrow \mathbb{R}$ continue non nulle telle que

$$\sup_A f < \inf_B f.$$

Notez qu'il s'agit d'une combinaison des deux théorèmes de Hahn-Banach.

En plus, il existe une autre formulation *plus géométrique* de ces théorèmes que voici :

Théorème 2.2.4. (Hahn-Banach, première forme géométrique) Soient $A \subset E$ et $B \subset E$ deux ensembles convexes, non vides et disjoints. On suppose que A est ouvert. Alors il existe un hyperplan fermé qui sépare A et B au sens large.

Théorème 2.2.5. (Hahn-Banach, deuxième forme géométrique). - Soient $A \subset E$ et $B \subset E$ deux ensembles convexes, non vides, disjoints. On suppose que A est fermé et que B est compact. Alors il existe un hyperplan fermé qui sépare A et B au sens strict.

Nous démontrerons les deux dernières formulations. Rappelons d'abord ce qui suit :

- Un hyperplan (affine) est un ensemble de la forme

$$H = \{x \in E; f(x) = \alpha\}$$

où f est une forme linéaire sur E , non identiquement nulle et $\alpha \in \mathbb{R}$. On dit que H est l'hyperplan d'équation $[f = \alpha]$.

— Soient $A \subset E$ et $B \subset E$. On dit que l'hyperplan H d'équation $[f = \alpha]$ sépare A et B au sens large si l'on a

$$f(x) \leq \alpha \quad \forall x \in A \quad \text{et} \quad f(x) \geq \alpha \quad \forall x \in B.$$

On dit que H sépare A et B au sens strict s'il existe $\varepsilon > 0$ tel que

$$f(x) \leq \alpha - \varepsilon \quad \forall x \in A \quad \text{et} \quad f(x) \geq \alpha + \varepsilon \quad \forall x \in B.$$

Géométriquement la séparation exprime que A et B se situent «de part et d'autre de H ».

La démonstration du théorème 2.2.4 repose sur les deux lemmes suivants :

Lemme 2.2.6. (*Jauge d'un convexe*). Soit $C \subset E$ un convexe ouvert avec $0 \in C$. Pour tout $x \in E$ on pose :

$$p(x) = \inf \{ \alpha > 0; \alpha^{-1}x \in C \} \quad (2.3)$$

(on dit que p est la jauge de C parfois appelée fonction de Minkowski).

Alors p vérifie 2.1–2.2 et

$$\text{il existe } M \text{ tel que } 0 \leq p(x) \leq M\|x\| \quad \forall x \in E, \quad (2.4)$$

$$C = \{x \in E; p(x) < 1\}. \quad (2.5)$$

Démonstration: Soit $r > 0$ tel que $B(0, r) \subset C$; il est clair que

$$p(x) \leq \frac{1}{r}\|x\| \quad \forall x \in E.$$

D'où 2.4.

La propriété 2.1 est évidente.

Prouvons 2.5. Supposons d'abord que $x \in C$; comme C est ouvert, $(1 + \varepsilon)x \in C$ pour $\varepsilon > 0$ assez petit. Donc $p(x) \leq \frac{1}{1+\varepsilon} < 1$. Inversement si $p(x) < 1$ il existe $0 < \alpha < 1$ tel que $\alpha^{-1}x \in C$ et donc $x = \alpha(\alpha^{-1}x) + (1 - \alpha)0 \in C$.

Prouvons 2.2. Soient $x, y \in E$ et soit $\varepsilon > 0$. D'après 2.1 et 2.5 on sait que $\frac{x}{p(x)+\varepsilon} \in C$ et $\frac{y}{p(y)+\varepsilon} \in C$. Donc $\frac{tx}{p(x)+\varepsilon} + \frac{(1-t)y}{p(y)+\varepsilon} \in C$ pour tout $t \in [0, 1]$. En particulier pour $t = \frac{p(x)+\varepsilon}{p(x)+p(y)+2\varepsilon}$ on obtient $\frac{x+y}{p(x)+p(y)+2\varepsilon} \in C$. On en déduit, grâce à 2.1 et 2.5, que $p(x+y) < p(x) + p(y) + 2\varepsilon$, $\forall \varepsilon > 0$. D'où 2.2. $\square$

Lemme 2.2.7. Soit $C \subset E$ un convexe ouvert non vide et soit $x_0 \in E$ avec $x_0 \notin C$. Alors il existe $f \in E^*$ tel que $f(x) < f(x_0) \forall x \in C$. En particulier l'hyperplan d'équation $[f = f(x_0)]$ sépare $\{x_0\}$ et C au sens large, où E^* est le dual de E .

Démonstration: Par translation on peut toujours supposer que $0 \in C$ et introduire la jauge de C (Lemme 2.2.6) notée p . On considère $G = \mathbb{R}x_0$ et la forme linéaire g définie sur G par

$$g(tx_0) = t, \quad t \in \mathbb{R}.$$

Il est clair que

$$g(x) \leq p(x) \quad \forall x \in G$$

(prendre $x = tx_0$ et distinguer les cas $t > 0$ et $t \leq 0$). Grâce au théorème de Hahn-Banach analytique, il existe une forme linéaire f sur E , qui prolonge g , et telle que

$$f(x) \leq p(x) \quad \forall x \in E.$$

En particulier on a $f(x_0) = 1$ et f est continue grâce à 2.4. D'autre part on déduit de 2.5 que $f(x) < 1$ pour tout $x \in C$. $\square$

Démonstration du théorème 2.2.4 On pose $C = A - B$ de sorte que C est convexe (vérification facile), C est ouvert (noter que $C = \bigcup_{y \in B} (A - y)$) et $0 \notin C$ (puisque $A \cap B = \emptyset$). D'après le lemme 2.2.7 il existe $f \in E^*$ (i.e une forme linéaire f) tel que

$$f(z) < 0 \quad \forall z \in C$$

c'est-à-dire

$$f(x) < f(y) \quad \forall x \in A, \quad \forall y \in B.$$

On fixe $\alpha \in \mathbb{R}$ avec

$$\sup_{x \in A} f(x) \leq \alpha \leq \inf_{y \in B} f(y)$$

et donc l'hyperplan d'équation $[f = \alpha]$ sépare au sens large A et B . $\square$

Démonstration du théorème 2.2.5 Pour $\varepsilon > 0$ on pose $A_\varepsilon = A + B(0, \varepsilon)$ et $B_\varepsilon = B + B(0, \varepsilon)$ de sorte que A_ε et B_ε sont convexes, ouverts et non vides. De plus, pour $\varepsilon > 0$ assez petit, A_ε et B_ε sont disjoints (sinon on pourrait trouver des suites $\varepsilon_n \rightarrow 0, x_n \in A$ et $y_n \in B$ telles que $\|x_n - y_n\| < 2\varepsilon_n$; on pourrait ensuite extraire une sous-suite $y_{n_k} \rightarrow y \in A \cap B$). D'après le théorème 2.2.4, il existe un hyperplan fermé d'équation $[f = \alpha]$ qui sépare A_ε et B_ε au sens large. On a donc

$$f(x + \varepsilon z) \leq \alpha \leq f(y + \varepsilon z) \quad \forall x \in A, \quad \forall y \in B, \quad \forall z \in B(0, 1).$$

Il en résulte que

$$f(x) + \varepsilon\|f\| \leq \alpha \leq f(y) - \varepsilon\|f\|. \quad \forall x \in A, \quad \forall y \in B.$$

On conclut que A et B sont séparés au sens strict par l'hyperplan $[f = \alpha]$ puisque $\|f\| \neq 0$. $\square$

1) Justification alternative de la disjonction de A_ε et B_ε

On suppose par l'absurde que $A_\varepsilon \cap B_\varepsilon \neq \emptyset$, c'est-à-dire qu'il existe un point $z \in E$ tel que $z \in A + B(0, \varepsilon)$ et $z \in B + B(0, \varepsilon)$.

Cela signifie qu'il existe $x \in A, y \in B$ et $u, v \in B(0, \varepsilon)$ tels que :

$$z = x + u = y + v.$$

En réarrangeant, on obtient :

$$x - y = v - u.$$

Or, par définition de la boule unité, on a $\|u\| < \varepsilon$ et $\|v\| < \varepsilon$, donc :

$$\|x - y\| = \|v - u\| \leq \|v\| + \|u\| < 2\varepsilon.$$

Si cela était vrai pour une suite $\varepsilon_n \rightarrow 0$, on obtiendrait une suite $x_n \in A$ et $y_n \in B$ avec $\|x_n - y_n\| < 2\varepsilon_n$. En passant à la limite, la compacité de B permet d'extraire une sous-suite $y_{n_k} \rightarrow y \in B$ et par définition de la fermeture, toute suite convergente dans A a sa limite dans A . Ainsi, $x_{n_k} \rightarrow x \in A$. Cela conduit à $x = y$, ce qui contredit la disjonction de A et B .

On en conclut que A_ε et B_ε restent disjoints pour ε suffisamment petit.

2) Justification de la dernière inégalité

D'après l'inégalité obtenue après application du théorème de séparation au sens large :

$$f(x + \varepsilon z) \leq \alpha \leq f(y + \varepsilon z), \quad \forall x \in A, \quad \forall y \in B, \quad \forall z \in B(0, 1).$$

En utilisant l'inégalité triangulaire et le fait que $|f(z)| \leq \|f\|\|z\|$, avec $\|z\| < 1$, on obtient :

$$f(x) + \varepsilon f(z) \leq \alpha \leq f(y) + \varepsilon f(z).$$

En prenant le sup sur $z \in B(0, 1)$ dans la première inégalité et le inf dans la seconde, on en déduit :

$$f(x) + \varepsilon\|f\| \leq \alpha \leq f(y) - \varepsilon\|f\|.$$

Ce qui prouve que l'hyperplan $[f = \alpha]$ sépare strictement A et B .

Remarque 2.2.8. Soient $A \subset E$ et $B \subset E$ deux ensembles convexes, non vides, disjoints. Sans hypothèse supplémentaire on ne peut pas toujours séparer A et B au sens large par un hyperplan fermé. On peut même construire un exemple où A et B sont deux convexes fermés, non vides, disjoints tels qu'il n'existe aucun hyperplan fermé séparant A et B au sens large.. Toutefois si E est un espace de dimension finie on peut toujours séparer au sens large deux convexes A et B , non vides, disjoints (sans hypothèse supplémentaire!).

Indiquons enfin un corollaire très utile lorsque l'on cherche à prouver qu'un sous-espace vectoriel est dense.

Corollaire 2.2.9. Soit $F \subset E$ un sous-espace vectoriel tel que $\overline{F} \neq E$. Alors il existe $f \in E^*, f \neq 0$ tel que

$$\langle f, x \rangle = 0 \quad \forall x \in F.$$

1) La notation $\overline{F}$ désigne l'adhérence du sous-espace F dans E . Cela signifie que $\overline{F}$ est l'ensemble de tous les points de E qui peuvent être approchés arbitrairement près par des éléments de F . Formellement, un point $x \in E$ appartient à $\overline{F}$ si et seulement si toute boule ouverte centrée en x contient au moins un élément de F .

2) L'affirmation repose sur le fait que l'application du théorème de séparation stricte garantit l'existence d'une forme linéaire $f \in E^*$ et d'un réel α tels que $\langle f, x \rangle < \alpha < \langle f, x_0 \rangle$ pour tout $x \in F$. En prenant un scalaire $\lambda \in \mathbb{R}$ et un vecteur $x \in F$, on a $\lambda \langle f, x \rangle < \alpha$. Si $\langle f, x \rangle \neq 0$, on peut choisir λ assez grand en valeur absolue pour que cette inégalité soit violée, ce qui est absurde. Autrement-dit, il est possible de prendre $\lambda \rightarrow \infty$ pour avoir $\lambda \langle f, x \rangle \geq 0$ et prendre $\lambda \rightarrow -\infty$ pour avoir $\lambda \langle f, x \rangle \leq 0$. Il s'ensuit que $\langle f, x \rangle = 0$ pour tout $x \in F$.

Démonstration: Soit $x_0 \in E, x_0 \notin \overline{F}$. On applique le théorème 2.2.5 avec $A = \overline{F}$ et $B = \{x_0\}$. Il existe donc $f \in E^*, f \neq 0$ tel que l'hyperplan d'équation $[f = \alpha]$ sépare au sens strict $\overline{F}$ et $\{x_0\}$. On a

$$\langle f, x \rangle < \alpha < \langle f, x_0 \rangle \quad \forall x \in F.$$

D'où il résulte que $\langle f, x \rangle = 0$, $\forall x \in F$, puisque $\lambda \langle f, x \rangle < \alpha$ pour tout $\lambda \in \mathbb{R}$. Soit $x_0 \in E$, $x_0 \notin \overline{F}$. On applique le théorème 2.2.5 avec $A = \overline{F}$ et $B = \{x_0\}$. Il existe donc $f \in E'$, $f \neq 0$ tel que l'hyperplan d'équation $[f = \alpha]$ sépare au sens strict $\overline{F}$ et $\{x_0\}$. On a

$$\langle f, x \rangle < \alpha < \langle f, x_0 \rangle \quad \forall x \in F.$$

D'où il résulte que $\langle f, x \rangle = 0$, $\forall x \in F$, puisque $\lambda \langle f, x \rangle < \alpha$ pour tout $\lambda \in \mathbb{R}$.

2.3 Un exemple d'application : le théorème de Krein Milman

Définition 2.3.1. Soit K une partie d'un espace vectoriel E . On dit que x_0 est un point extrémal de K si

$$(x_0 = \theta z + (1 - \theta)y \text{ avec } \theta \in]0, 1[\text{ et } z, y \in K) \Rightarrow x_0 = y = z.$$

Théorème 2.3.2. Soient E un espace vectoriel topologique localement convexe séparé et K une partie compacte de E . Alors K est inclus dans l'enveloppe convexe fermée de ses points extrémaux et coïncide avec elle si K est convexe.

Soit A un sous-ensemble d'un espace vectoriel E . L'enveloppe convexe de A , notée $\text{conv}(A)$, est le plus petit ensemble convexe contenant A . Elle est donnée par :

$$\text{conv}(A) = \left\{ \sum_{i=1}^n \lambda_i x_i \mid n \in \mathbb{N}, x_i \in A, \lambda_i \geq 0, \sum_{i=1}^n \lambda_i = 1 \right\}.$$

Cela signifie que $\text{conv}(A)$ est l'ensemble de toutes les combinaisons convexes d'éléments de A , où une combinaison convexe est une somme pondérée des points de A avec des coefficients positifs dont la somme est égale à 1.

Un *espace topologique localement convexe* est un espace vectoriel topologique (E, τ) sur un corps $\mathbb{K}$ (où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) dont la topologie τ est définie par une famille de semi-normes. Autrement dit, il existe une famille $(p_i)_{i \in I}$ de semi-normes sur E telle que la topologie τ soit la topologie initiale associée à cette famille, c'est-à-dire la plus fine des topologies rendant toutes les applications $x \mapsto p_i(x)$ continues.

Une caractérisation équivalente est qu'un espace vectoriel topologique est localement convexe si et seulement si son origine admet une base de voisinages formée d'ensembles convexes.

Soit (X, τ) un espace topologique et soit $x \in X$. Une *base de voisinages* ou *système fondamental de voisinages* de x est une famille $\mathcal{B}_x$ de voisinages de x telle que tout voisinage V de x contienne au moins un élément de $\mathcal{B}_x$, c'est-à-dire :

$$\forall V \in \tau, \quad x \in V \Rightarrow \exists B \in \mathcal{B}_x, \quad B \subseteq V.$$

Autrement dit, les éléments de $\mathcal{B}_x$ permettent de décrire complètement la structure locale de la topologie autour de x , dans le sens où il suffit de connaître $\mathcal{B}_x$ pour retrouver tous les voisinages de x par prise d'union.

Dans un espace vectoriel topologique, une *base de voisinages de l'origine* est particulièrement importante, car elle permet de définir la topologie sur tout l'espace en utilisant la

translation.

En particulier, dans un espace topologique localement convexe, il existe une base de voisinages de l'origine constituée d'ensembles convexes.

Démonstration: La démonstration de ce résultat se fait en deux étapes basées sur la version géométrique du théorème de Hahn-Banach : on commence par montrer que l'ensemble des points extrémaux $\mathcal{E}$ est non vide, puis on prouve que tout point de K est nécessairement dans l'adhérence de l'enveloppe convexe de $\mathcal{E}$.

- On considère l'ensemble $\mathcal{P}$ des parties extrémales de l'enveloppe convexe fermée $\overline{co(K)}$ de K , i.e. des parties A compactes vérifiant

$$\text{s'il existe } x, y \in \overline{co(K)}, \theta \in]0, 1[\text{ tels que } \theta x + (1 - \theta)y \in A, \quad \text{alors } x, y \in A$$

que l'on munit de la relation d'ordre $A \prec B$ si $B \subset A$.

Si $\tilde{\mathcal{P}} \subset \mathcal{P}$ est totalement ordonné, $B = \bigcap_{A \in \tilde{\mathcal{P}}} A$ est une partie extrémale non vide (intersection de fermés emboîtés dans un compact) majorant $\tilde{\mathcal{P}}$. Donc $\mathcal{P}$ est inductif. D'après le lemme de Zorn, $\mathcal{P}$ a alors un élément maximal M .

Si M contient deux points distincts x_0 et x_1 , d'après le théorème de Hahn-Banach, il existe une forme linéaire continue sur E telle que $f(x_0) < f(x_1)$. On définit alors $\tilde{M} = \{x \in M / f(x) = \inf_M f\}$. Comme M est compact et f est continue, $\tilde{M} \neq \emptyset$. De plus, $\tilde{M}$ est compact (fermé dans un compact) et extrémal

$$\begin{aligned} &\text{s'il existe } x, y \in \overline{co(K)}, \theta \in]0, 1[\text{ tels que } \theta x + (1 - \theta)y \in \tilde{M} \\ &\text{alors } \begin{cases} x, y \in M \text{ car } M \text{ est extrémal,} \\ \theta f(x) + (1 - \theta)f(y) = \inf_M f \quad \text{donc } x, y \in \tilde{M}. \end{cases} \end{aligned}$$

Ceci est absurde puisque M est extrémal et que $\tilde{M}$ est strictement inclus dans M . M est donc réduit à un point.

- Soit $\mathcal{E}$ l'ensemble des points extrémaux de $\overline{co(K)}$ (qui par définition appartiennent à K), et x_0 un point de $\overline{co(K)}$ qui n'appartient pas à l'adhérence de l'enveloppe convexe de $\mathcal{E}$.

D'après le théorème de Hahn-Banach, il existe une forme linéaire continue f sur E telle que

$$f(x_0) > \sup\{f(x) / x \in \overline{co(\mathcal{E})}\}.$$

Soit A l'ensemble des points de $\overline{co(K)}$ où f atteint son maximum. L'argument précédent montre que A est extrémale. L'ensemble des parties extrémales de $\overline{co(K)}$ incluses dans A admet un élément maximal réduit à un point $x_1 \in \mathcal{E}$:

$$f(x_1) \geq f(x_0) > \sup\{f(x) / x \in \overline{co(\mathcal{E})}\} \geq \sup_{\mathcal{E}} f$$

ce qui aboutit à une contradiction. □

Interprétation du théorème

Le théorème de Krein-Milman affirme que toute partie compacte et convexe d'un espace vectoriel topologique localement convexe séparé est entièrement déterminée par ses points extrémaux. Intuitivement, cela signifie que si un compact convexe possède des points extrémaux, alors l'enveloppe convexe fermée de ces points extrémaux est suffisante pour reconstruire l'ensemble lui-même.

Autrement dit, même si l'ensemble K peut paraître complexe, il est en réalité la “coque convexe fermée” de ses éléments les plus extrêmes, ce qui fournit une description géométrique plus simple et utile en analyse fonctionnelle et optimisation convexe.

Démonstration détaillée

La démonstration du théorème repose sur deux étapes principales :

1. L'existence d'un point extrémal dans K .
2. La densité de l'enveloppe convexe des points extrémaux dans K .

Existence d'un point extrémal

On considère l'ensemble $\mathcal{P}$ des parties extrémales compactes de K , muni de la relation d'ordre d'inclusion. D'après le lemme de Zorn, il existe une partie extrémale maximale M . Si M contient au moins deux points distincts x_0 et x_1 , alors, par le théorème de séparation de Hahn-Banach, il existe une forme linéaire continue f telle que $f(x_0) < f(x_1)$. On construit alors une sous-partie $\tilde{M} = \{x \in M \mid f(x) = \inf_M f\}$ qui est compacte, extrémale et strictement incluse dans M , ce qui contredit la maximalité de M . Ainsi, M est réduit à un seul point, qui est extrémal.

Densité de l'enveloppe convexe des points extrémaux

On suppose par l'absurde qu'il existe $x_0 \in K$ qui n'appartient pas à l'adhérence de l'enveloppe convexe des points extrémaux de K . Par le théorème de Hahn-Banach, il existe une forme linéaire continue f telle que

$$f(x_0) > \sup\{f(x) \mid x \in \overline{\text{co}(\mathcal{E})}\}.$$

En considérant les points où f atteint son maximum, on peut construire une contradiction en utilisant le même argument que précédemment. Cela prouve que x_0 est bien dans l'adhérence de l'enveloppe convexe des points extrémaux, ce qui conclut la démonstration.

Autre formulation et démonstration

Formulation alternative

Si K est un sous-ensemble compact et convexe d'un espace vectoriel topologique localement convexe séparé, alors K est l'adhérence de l'enveloppe convexe de son ensemble de points extrémaux.

Autre démonstration

L'idée est d'utiliser un raisonnement par compacité et séparation affine :

- Pour tout $x \in K$, si x n'est pas dans l'enveloppe convexe des points extrémaux, alors il peut être séparé par un hyperplan affine.
- En considérant des suites de points et en exploitant la compacité, on obtient une contradiction similaire.

Remarque 2.3.3. *Dans le cas des espaces de Fréchet, l'enveloppe convexe fermée d'un compact est précompacte donc compacte, et pour tout x dans l'enveloppe convexe, il existe une mesure de probabilité μ_x sur l'ensemble $\mathcal{E}$ des points extrémaux telle que $x = \int_{\mathcal{E}} y d\mu_x(y)$ (théorème de Choquet).*

Dualité et Topologies faibles

La notion de dualité est au coeur des méthodes d'analyse moderne, et jouera un rôle fondamental dans ce cours, puisqu'elle est à l'origine de la notion même de distributions. La présentation qui en est faite dans ce chapitre est assez abstraite, mais on verra par la suite beaucoup d'applications : dérivation de fonctions non continues, définition des solutions faibles d'équations aux dérivées partielles, démonstration des propriétés de la transformée de Fourier,...

C'est probablement l'un des objectifs les plus importants de ce cours d'utiliser ce nouveau point de vue pour étudier les espaces de fonctions.

Ici et dans la suite de ce cours on notera par E^* le dual de l'espace vectoriel E , c'est-à-dire l'espace vectoriel des applications (ou des formes) linéaires définies sur E . Notez que certains manuels préfèrent la notation E' suggérant que E^* est une notation anglosaxonne. Quelle est différence entre la topologie $\sigma(E, E^*)$ et $\sigma(E^*, E)$? Quelle est la définition exacte de chacune d'elles?

Soit E un espace vectoriel et E^* son dual algébrique, c'est-à-dire l'ensemble des formes linéaires sur E . On définit deux topologies faibles associées :

1. Topologie faible sur E :

La topologie $\sigma(E, E^*)$, appelée *topologie faible* sur E , est la topologie la plus faible pour laquelle toutes les formes linéaires de E^* sont continues. Autrement dit, la famille de semi-normes définissant cette topologie est donnée par :

$$p_f(x) = |f(x)|, \quad \text{pour tout } f \in E^*.$$

Les ouverts de cette topologie sont donc générés par les ensembles de la forme :

$$U = \{x \in E \mid |f_i(x) - a_i| < \varepsilon, \quad i = 1, \dots, n\},$$

où $f_1, \dots, f_n \in E^*$, $a_1, \dots, a_n \in \mathbb{R}$ et $\varepsilon > 0$.

2. Topologie faible-* sur E^* :

La topologie $\sigma(E^*, E)$, appelée *topologie faible* sur E^* ou tout simplement *topologie faible**, est la topologie la plus faible pour laquelle toutes les applications d'évaluation (cad application de E^* vers E) $x \mapsto f(x)$ pour $x \in E$ sont continues sur E^* . Cette topologie est engendrée par les semi-normes :

$$p_x(f) = |f(x)|, \quad \text{pour tout } x \in E.$$

Ainsi, les ouverts de cette topologie sont générés par les ensembles de la forme :

$$V = \{f \in E^* \mid |f(x_i) - b_i| < \delta, \quad i = 1, \dots, m\},$$

où $x_1, \dots, x_m \in E$, $b_1, \dots, b_m \in \mathbb{R}$ et $\delta > 0$.

Différence entre les deux topologies

- La topologie $\sigma(E, E^*)$ sur E est définie à partir des éléments de E^* , tandis que la topologie $\sigma(E^*, E)$ sur E^* est définie à partir des éléments de E .
- La convergence faible dans E signifie que les évaluations par les formes linéaires de E^* convergent, tandis que la convergence faible-* dans E^* signifie que les évaluations en chaque élément de E convergent.
- Si E est un espace de Banach, alors E^* est muni de sa topologie faible-* qui est strictement plus faible que la topologie forte induite par la norme de E^* . En résumé, $\sigma(E, E^*)$ est la topologie faible sur E et $\sigma(E^*, E)$ est la topologie faible-* sur E^* .

3.1 Dual topologique

Si E et F sont deux espaces vectoriels, et $\langle \cdot, \cdot \rangle$ une forme bilinéaire de $F \times E$ dans $\mathbb{R}$ telle que

$$\begin{aligned} (\forall f \in F, \langle f, x \rangle = 0) &\Rightarrow x = 0 \\ (\forall x \in E, \langle f, x \rangle = 0) &\Rightarrow f = 0 \end{aligned} \quad (3.1.1)$$

on peut définir une topologie d'espace vectoriel localement convexe sur E (respectivement sur F) en considérant toutes les semi-normes

$$p_B(x) = \sup_{f \in B} |\langle f, x \rangle| \quad \text{où } B \text{ est une partie finie de } F$$

(respectivement

$$q_A(f) = \sup_{x \in A} |\langle f, x \rangle| \quad \text{où } A \text{ est une partie finie de } E).$$

Si E est un espace vectoriel topologique localement convexe séparé, défini par ses ouverts $\tau(E)$, on peut considérer la forme bilinéaire

$$(x, f) \in E \times E^* \mapsto \langle f, x \rangle = f(x) \in \mathbb{R}.$$

D'après le théorème de Hahn-Banach, le crochet de dualité vérifie l'hypothèse (3.1.1)*. On peut ainsi définir une nouvelle topologie sur E , notée $\sigma(E, E^*)$, et une topologie sur E^* , notée $\sigma(E^*, E)$.

★ En effet, d'après le théorème de Hahn-Banach, toute forme linéaire non triviale sur un espace vectoriel topologique localement convexe séparé peut être prolongée sans augmenter sa norme. Cela implique que pour tout $x \neq 0$ dans E , il existe un fonctionnel linéaire $f \in E^*$ tel que $\langle f, x \rangle \neq 0$, ce qui montre que la première condition de (3.1.1) est satisfaite.

De même, si $f \neq 0$ dans E^* , alors par définition d'un fonctionnel non nul, il existe $x \in E$ tel que $f(x) = \langle f, x \rangle \neq 0$, ce qui prouve la seconde condition de (3.1.1). Ainsi, le crochet de dualité satisfait bien les hypothèses (3.1.1).

Théorème 3.1.1. Soit E est un espace vectoriel topologique localement convexe séparé (défini par ses semi-normes (p_α)).

Alors la topologie $\sigma(E, E^*)$ est plus pauvre que $\tau(E)$: elle a moins d'ouverts, plus de compacts, il existe moins de fonctions continues sur E à valeurs réelles,...

Démonstration: Soit A un ouvert pour $\sigma(E, E^*)$. Pour tout $x_0 \in A$, il existe une partie finie B de E^* et $r > 0$ tels que

$$p_B(x - x_0) < r \Rightarrow x \in A.$$

Comme B est finie, ses éléments sont uniformément continus : il existe $C > 0$ et α tels que

$$\forall x \in E, \forall f \in B, \quad | \langle f, x - x_0 \rangle | \leq C h_\alpha(x - x_0).$$

On en déduit que $B_\alpha(x_0, r/C) \subset A$, et donc que A est ouvert pour la topologie initiale.

□

Les formes linéaires de E^* sont continues par définition dans la topologie faible-*, car la topologie $\sigma(E, E^*)$ est précisément la topologie la plus faible rendant continues toutes les formes linéaires de E^* . Cependant, nous devons montrer que ces formes sont également continues pour la topologie initiale $\tau(E)$.

Comme E est un espace localement convexe séparé, sa topologie est définie par une famille de semi-normes (p_α) . Par définition de la continuité dans un espace localement convexe, une forme linéaire $f \in E^*$ est continue s'il existe une semi-norme h_α et une constante $C > 0$ telles que

$$\forall x \in E, \quad |f(x)| \leq C h_\alpha(x).$$

Cela signifie que chaque forme linéaire bornée est contrôlée par une semi-norme parmi celles qui définissent la topologie initiale.

Puisque B est un ensemble fini, nous pouvons choisir une semi-norme h_α commune pour tous les éléments de B . Ainsi, il existe une constante $C > 0$ et un indice α tels que, pour tout $f \in B$ et tout $x \in E$,

$$|f(x - x_0)| \leq C h_\alpha(x - x_0).$$

Cette inégalité exprime une forme d'uniforme continuité des éléments de B relativement à la topologie définie par les semi-normes. Cela garantit que la boule $B_\alpha(x_0, r/C)$ pour la semi-norme h_α est incluse dans l'ouvert A , ce qui prouve que A est également ouvert pour la topologie initiale $\tau(E)$.

L'application de cette propriété permet de conclure que la topologie faible-* est plus pauvre que la topologie initiale, c'est-à-dire qu'elle a moins d'ouverts.

En effet, si $x \in B_\alpha(x_0, r/C)$, alors $h_\alpha(x - x_0) < r/C$. En utilisant l'inégalité précédente, nous avons, pour tout $f \in B$,

$$|f(x - x_0)| \leq C h_\alpha(x - x_0) < C \frac{r}{C} = r.$$

Par conséquent, $p_B(x - x_0) = \sup_{f \in B} |f(x - x_0)| < r$, ce qui implique $x \in A$ d'après la définition de A . Ainsi, nous avons bien $B_\alpha(x_0, r/C) \subset A$.

L'application de cette propriété permet de conclure que la topologie faible-* est plus pauvre que la topologie initiale, c'est-à-dire qu'elle a moins d'ouverts.

Dans le cas d'un espace de Banach $(E, \|\cdot\|)$, E et E^* sont mutuellement en dualité. En effet, E^* est un espace de Banach pour la norme duale

$$\|f\|_{E^*} = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|_E}.$$

Définition 3.1.2. Soit $(E, \|\cdot\|)$ un espace de Banach.

- La topologie associée à la norme $\|\cdot\|$ est dite topologie forte sur E .
- La topologie $\sigma(E, E^*)$ est dite topologie faible sur E .
- La topologie $\sigma(E^*, E)$ est dite topologie faible sur E^* .

Si E est lui-même le dual d'un espace de Banach, on a alors trois topologies sur E .

Propriété 3.1.3. Soit $(E, \|\cdot\|)$ un espace de Banach.

- Si (x_n) converge fortement vers x , x_n converge faiblement vers x .
- Si (x_n) converge faiblement vers x , x_n est bornée et

$$\|x\| \leq \liminf_{n \rightarrow \infty} \|x_n\|.$$

Démonstration: La première propriété est une simple conséquence de la continuité de $f \in E^*$ (voir aussi le théorème précédent comparant les topologies forte et faible).

La seconde propriété est un corollaire du théorème de Banach-Steinhaus. L'inégalité est obtenue en utilisant la caractérisation suivante de la norme

$$\|x\|_E = \max_{\|f\|_{E^*} \leq 1} |f(x)|$$

qui résulte du théorème de Hahn-Banach. □

Démonstration détaillée

Première propriété : Si (x_n) converge fortement vers x , alors (x_n) converge faiblement vers x .

Soit (x_n) une suite convergeant fortement vers x dans E , c'est-à-dire,

$$\lim_{n \rightarrow \infty} \|x_n - x\| = 0.$$

Nous devons montrer que (x_n) converge faiblement vers x . Autrement dit, nous devons montrer que pour tout $f \in E^*$,

$$\lim_{n \rightarrow \infty} f(x_n) = f(x).$$

La continuité de la fonctionnelle linéaire $f \in E^*$ implique que pour tout $f \in E^*$,

$$\lim_{n \rightarrow \infty} f(x_n) = f\left(\lim_{n \rightarrow \infty} x_n\right) = f(x).$$

Cette propriété est simplement une conséquence directe de la continuité des fonctionnelles linéaires dans un espace de Banach. Par conséquent, la suite (x_n) converge faiblement vers x .

Définition : Une fonctionnelle est une fonction qui agit sur des fonctions. Autrement dit, c'est une application d'un espace de fonctions dans l'ensemble des réels (ou dans un autre corps). Formellement, si $\mathcal{F}$ est une fonctionnelle, on peut l'écrire comme suit :

$$\mathcal{F} : X \longrightarrow \mathbb{R}$$

où X est un espace de fonctions et $\mathcal{F}$ prend une fonction $f \in X$ et lui associe un nombre réel $\mathcal{F}(f)$.

Exemple classique de fonctionnelle : dans le cadre de l'analyse fonctionnelle, la fonctionnelle $\mathcal{F}(f) = \int_a^b f(x) dx$ associe à chaque fonction f de l'espace $L^2([a, b])$ l'intégrale de f sur l'intervalle $[a, b]$.

Deuxième propriété : Si (x_n) converge faiblement vers x et si (x_n) est bornée, alors

$$\|x\| \leq \liminf_{n \rightarrow \infty} \|x_n\|.$$

Soit (x_n) une suite convergeant faiblement vers x dans E , c'est-à-dire, pour tout $f \in E^*$,

$$\lim_{n \rightarrow \infty} f(x_n) = f(x).$$

De plus, supposons que la suite (x_n) est bornée. Cela signifie qu'il existe une constante $M > 0$ telle que

$$\|x_n\| \leq M \quad \text{pour tout } n.$$

Nous devons maintenant démontrer l'inégalité

$$\|x\| \leq \liminf_{n \rightarrow \infty} \|x_n\|.$$

Utilisons la caractérisation de la norme dans E . Par le théorème de Hahn-Banach, pour chaque $f \in E^*$ de norme 1, on a

$$\|x\| = \sup_{\|f\|_{E^*} \leq 1} |f(x)|.$$

Comme (x_n) converge faiblement vers x , pour chaque $f \in E^*$, on a

$$\lim_{n \rightarrow \infty} f(x_n) = f(x).$$

En particulier, pour chaque f tel que $\|f\|_{E^*} \leq 1$, on a

$$\lim_{n \rightarrow \infty} |f(x_n)| = |f(x)|.$$

De plus, la suite $(f(x_n))$ est bornée, car la suite (x_n) est bornée et les fonctionnelles de norme 1 sont continues et donc bornées sur les éléments de norme bornée.

En prenant la limite inférieure, on obtient

$$\|x\| = \sup_{\|f\|_{E^*} \leq 1} \liminf_{n \rightarrow \infty} |f(x_n)| \leq \liminf_{n \rightarrow \infty} \sup_{\|f\|_{E^*} \leq 1} |f(x_n)| = \liminf_{n \rightarrow \infty} \|x_n\|.$$

Ceci prouve l'inégalité souhaitée.

De façon analogue, on a

Propriété 3.1.4. Soit $(E, \|\cdot\|)$ un espace de Banach, et E^* son dual.

- Si (f_n) converge fortement vers f dans E^* , f_n converge faiblement* vers f .
- Si (f_n) converge faiblement* vers f dans E^* , f_n est bornée et

$$\|f\| \leq \liminf_{n \rightarrow \infty} \|f_n\|.$$

3.2 Topologie faible

Lorsque E est de dimension infinie, la topologie faible est distincte de la topologie forte : il existe toujours des fermés pour la topologie forte qui ne sont pas fermés pour la topologie faible (par exemple la sphère unité**).

* La sphère unité d'un espace vectoriel topologique localement convexe E est définie par

$$S = \{x \in E \mid \|x\| = 1\}.$$

Fermeture dans la topologie forte

La topologie forte sur E est définie par une famille de semi-normes $\{p_\alpha\}$. La boule unité fermée s'écrit

$$B = \{x \in E \mid \|x\| \leq 1\}.$$

Cet ensemble est fermé car si une suite $(x_n) \subset B$ converge fortement vers x , alors $\|x_n - x\| \rightarrow 0$, ce qui implique $\|x\| \leq 1$. La sphère unité est l'intersection de B avec l'ensemble $\{x \mid \|x\| = 1\}$, donc elle est fermée.

Non-fermeture dans la topologie faible

Dans la topologie faible $\sigma(E, E^*)$, la convergence d'une suite (x_n) vers x signifie que

$$f(x_n) \rightarrow f(x), \quad \forall f \in E^*.$$

Cependant, cette convergence n'impose pas nécessairement $\|x_n\| \rightarrow \|x\|$, car la norme n'est pas forcément une fonction continue dans la topologie faible. Par conséquent, il existe des suites $(x_n) \subset S$ qui convergent faiblement vers un point x avec $\|x\| < 1$, ce qui montre que S n'est pas fermé dans $\sigma(E, E^*)$.

Pour les sous-ensembles convexes toutefois on a le résultat suivant :

Théorème 3.2.1. *Soit E est un espace vectoriel topologique localement convexe séparé. Si $C \subset E$ est un convexe fermé pour $\tau(E)$, alors C est aussi fermé pour $\sigma(E, E^*)$.*

Démonstration: Soit $x_0 \notin C$. D'après le théorème de Hahn-Banach, on peut séparer $\{x_0\}$ et C

$$\exists f \in E^*, \quad \exists \alpha \in \mathbb{R}, \quad f(x_0) < \alpha \leq \inf_C f$$

En particulier,

$$\{x \in E \mid |f(x) - f(x_0)| < \alpha - f(x_0)\} \cap C = \emptyset$$

ce qui signifie que C est fermé (complémentaire d'un ouvert) pour la topologie $\sigma(E, E^*)$.

□

Démonstration détaillée :

Soit $x_0 \notin C$. Nous devons montrer que C est fermé pour la topologie faible- $\sigma(E, E^*)$, c'est-à-dire que si une suite (ou un filet) $(x_n) \subset C$ converge faiblement vers x_0 , alors $x_0 \in C$.

D'après l'hypothèse, E est un espace vectoriel topologique localement convexe séparé. Le théorème de séparation de Hahn-Banach garantit alors que tout convexe fermé C peut

être strictement séparé d'un point $x_0 \notin C$.

Autrement dit, il existe une forme linéaire continue $f \in E^*$ et un réel $\alpha \in \mathbb{R}$ tels que :

$$f(x_0) < \alpha \leq \inf_{x \in C} f(x).$$

Cela signifie que l'ensemble C est entièrement contenu dans le demi-espace fermé $\{x \in E \mid f(x) \geq \alpha\}$, tandis que x_0 est strictement en dehors.

Or, la topologie $\sigma(E, E^*)$ est la topologie initiale engendrée par les formes linéaires continues de E^* , c'est-à-dire la topologie la moins fine rendant toutes les applications $x \mapsto f(x)$ continues pour tout $f \in E^*$. Ainsi, l'ensemble

$$\{x \in E \mid f(x) < \alpha\}$$

est un ouvert pour $\sigma(E, E^*)$ contenant x_0 et disjoint de C . Cela prouve que C est fermé dans $\sigma(E, E^*)$, ce qui conclut la démonstration. $\square$

Corollaire 3.2.2. (Lemme de Mazur). - Soit E un espace de Banach. Si (x_n) converge faiblement vers x dans E , alors il existe une suite (y_n) avec chaque y_n combinaison convexe des $(x_k)_{k \geq n}$, qui converge fortement vers x .

Démonstration: On considère l'enveloppe convexe C_n des $(x_k)_{k \geq n}$, de sorte que x appartient à la fermeture faible de C_n . D'après le théorème 3.2.1, la fermeture forte de C_n est égale à la fermeture faible. $\square$

Démonstration détaillée : La convergence faible de (x_n) vers x signifie que pour tout $f \in E^*$, on a :

$$f(x_n) \rightarrow f(x).$$

Pour chaque entier n , considérons l'enveloppe convexe C_n de l'ensemble $\{x_k \mid k \geq n\}$, c'est-à-dire :

$$C_n = \left\{ \sum_{k \geq n} \lambda_k x_k \mid \lambda_k \geq 0, \sum_{k \geq n} \lambda_k = 1 \right\}.$$

Puisque $x_n \rightharpoonup x$, il en résulte que x appartient à la fermeture faible de C_n :

$$x \in \overline{C_n}^{\text{faible}}.$$

D'après le théorème 3.2.1, dans un espace de Banach, la fermeture forte d'un convexe fermé est égale à sa fermeture faible :

$$\overline{C_n}^{\text{forte}} = \overline{C_n}^{\text{faible}}.$$

Par conséquent, comme $x \in \overline{C_n}^{\text{faible}}$, il appartient aussi à $\overline{C_n}^{\text{forte}}$. Ainsi, il existe une suite (y_n) dans C_n , donc chaque y_n est une combinaison convexe des $(x_k)_{k \geq n}$, telle que :

$$y_n \rightarrow x \quad \text{fortement.}$$

Cela prouve l'énoncé. $\square$

Proposition 3.2.3. (sur l'équivalence continuité forte/faible). - Soient E, F deux espaces de Banach, E^*, F^* leurs duaux respectifs, et $T : E \rightarrow F$ une application linéaire.

Alors T est continue de $(E, \|\cdot\|_E)$ dans $(F, \|\cdot\|_F)$ si et seulement si T est continue de $(E, \sigma(E, E^*))$ dans $(F, \sigma(F, F^*))$.

Démonstration: • Supposons d'abord T continue de $(E, \|\cdot\|_E)$ dans $(F, \|\cdot\|_F)$. Pour tout $f \in F^*$, $f \circ T \in E^*$ et donc est continue pour $\sigma(E, E^*)$. On en déduit alors que T est continue de $(E, \sigma(E, E^*))$ dans $(F, \sigma(F, F^*))$.

En effet, $\sigma(F, F^*)$ est la topologie la moins fine contenant la famille

$$\Lambda = \{f^{-1}(\omega) | f \in E^* \text{ et } \omega \text{ ouvert de } \mathbb{R}\}.$$

Soit U un ouvert de $\sigma(F, F^*)$, on a

$$U = \bigcup_{j \in J} \bigcap_{i \in I_j} f_i^{-1}(\omega_i)$$

où chaque I_j est fini, $f_i \in F^*$ et ω_i est un ouvert de $\mathbb{R}$. On a alors

$$T^{-1}(U) = \bigcup_{j \in J} \bigcap_{i \in I_j} (f_i \circ T)^{-1}(\omega_i)$$

qui est bien ouvert puisque chaque $f_i \circ T$ est continue.

• Supposons maintenant que T est continue de $(E, \sigma(E, E^*))$ dans $(F, \sigma(F, F^*))$. Alors le graphe de T est fermé pour $\sigma(E \times F, E^* \times F^*)$ et donc aussi fortement fermé dans $E \times F$. Comme E et F sont des espaces de Banach, on déduit du théorème du graphe fermé, que T est continue de $(E, \|\cdot\|_E)$ dans $(F, \|\cdot\|_F)$. $\square$

Démonstration détaillée : • Supposons d'abord que T est continue de $(E, \|\cdot\|_E)$ dans $(F, \|\cdot\|_F)$.

Cela signifie qu'il existe une constante $C > 0$ telle que, pour tout $x \in E$,

$$\|T(x)\|_F \leq C\|x\|_E.$$

Soit $f \in F^*$. L'application linéaire $f \circ T : E \rightarrow \mathbb{R}$ est alors continue pour la norme de E , donc appartient à E^* . Cela signifie que $f \circ T$ est continue pour la topologie faible $\sigma(E, E^*)$.

Puisque la topologie $\sigma(F, F^*)$ est la plus faible rendant continues toutes les formes linéaires de F^* , cela implique que T est continue pour les topologies faibles $\sigma(E, E^*)$ et $\sigma(F, F^*)$. En d'autres termes, l'image réciproque de tout ouvert faible dans F par T est un ouvert faible dans E , ce qui prouve que T est continue de $(E, \sigma(E, E^*))$ dans $(F, \sigma(F, F^*))$.

• Supposons maintenant que T est continue de $(E, \sigma(E, E^*))$ dans $(F, \sigma(F, F^*))$.

Nous allons utiliser le théorème du graphe fermé. Ce théorème stipule que si un opérateur linéaire entre espaces de Banach a un graphe fermé, alors il est nécessairement continu pour la norme.

Définissons le graphe de T :

$$G(T) = \{(x, T(x)) \mid x \in E\} \subset E \times F.$$

Nous devons montrer que $G(T)$ est fermé dans l'espace produit $E \times F$ muni de la topologie produit forte.

Puisque T est continue pour les topologies faibles, le graphe $G(T)$ est fermé pour la topologie faible de $E \times F$, c'est-à-dire la topologie $\sigma(E \times F, E^* \times F^*)$. Or, dans un espace de Banach, une partie qui est fermée pour la topologie faible est aussi fermée pour la topologie forte (résultat classique en analyse fonctionnelle).

Ainsi, $G(T)$ est fermé dans $E \times F$ muni de la topologie forte. Le théorème du graphe fermé garantit alors que T est continue pour la norme, ce qui conclut la démonstration.

$\square$

3.3 Topologie faible*

Théorème 3.3.1. (de Banach-Alaoglu). - Soient E un espace de Banach et E^* son dual. Alors la boule unité $B = \{f \in E^* \mid \|f\|_{E^*} \leq 1\}$ est compacte pour la topologie faible*. De plus, si E est séparable, B est séquentiellement compacte pour la topologie faible*.

La démonstration utilise sur le théorème de Tychonov que nous donnons sans démonstration (elle est basée sur le lemme de Zorn).

Théorème 3.3.2. Le produit d'une famille quelconque de compacts est un compact pour la topologie produit.

Démonstration du théorème de Banach-Alaoglu : En utilisant la correspondance entre f et $(f(x))_{x \in E}$, on peut voir B comme une partie de

$$K = \prod_{x \in E} [-\|x\|, \|x\|].$$

Lorsqu'on considère un fonctionnel $f \in E^*$, on peut le voir comme une application de E dans $\mathbb{R}$ (ou $\mathbb{C}$ si l'espace est complexe). Ainsi, on peut identifier f avec la famille de valeurs $(f(x))_{x \in E}$.

L'ensemble B est alors constitué de toutes ces applications f qui sont linéaires et bornées par 1 en norme duale. Étant donné que chaque $f(x)$ satisfait l'inégalité $|f(x)| \leq \|f\| \|x\| \leq \|x\|$, il est naturel de voir B comme une partie du produit cartésien de segments bornés $\prod_{x \in E} [-\|x\|, \|x\|]$, qui est exactement l'ensemble K défini dans la démonstration.

D'après le théorème de Tychonov, un produit quelconque d'espaces topologiques compacts est compact pour la topologie produit, donc K est compact. Comme B est le sous-ensemble de K constitué des fonctions linéaires

$$B = \{f \in K \mid \forall x, y \in E, \forall \lambda \in \mathbb{R}, \quad f(\lambda x + y) = \lambda f(x) + f(y)\}$$

B est fermé dans K , donc compact. Dans le cas où E est séparable, il existe une suite (x_j) dense dans E . Si on se donne une suite (f_n) de B , par extraction diagonale, on peut construire une suite n_k telle que, pour tout j ,

$$f_{n_k}(x_j) \rightarrow g(x_j) \text{ quand } k \rightarrow \infty \text{ avec } |g(x_j)| \leq \|x_j\|.$$

L'extraction diagonale est une technique classique en analyse fonctionnelle pour construire une sous-suite convergente lorsque l'on dispose d'une famille infinie de suites.

Dans ce cas, on considère une suite (f_n) dans B . Pour chaque élément d'une partie dénombrable dense $\{x_j\}_{j \in \mathbb{N}} \subset E$, on extrait une sous-suite de (f_n) qui converge en x_j .

On construit alors la sous-suite (f_{n_k}) en procédant de manière inductive :

- On choisit une sous-suite $(f_n^{(1)})$ qui converge en x_1 ,
- Puis, on extrait une sous-suite $(f_n^{(2)})$ de $(f_n^{(1)})$ qui converge en x_2 ,
- En continuant ainsi, à l'étape k , on extrait une sous-suite $(f_n^{(k)})$ de $(f_n^{(k-1)})$ qui converge en x_k .

On définit ensuite $f_{n_k} = f_n^{(k)}$, ce qui assure la convergence en chaque x_j , et donc, par densité, une convergence faible* dans E^* .

Par densité et linéarité, on peut alors prolonger g en une fonction linéaire sur E telle que,

$$\forall x \in E, \quad |g(x)| \leq \|x\|$$

Par inégalité triangulaire, on peut alors montrer que f_{n_k} converge simplement vers g (convergence faible*).

Cette preuve utilise uniquement le fait que, si U est un voisinage de 0 ,

$$K = \{f \in E^* / \forall x \in U, \quad |f(x)| \leq 1\} \text{ est compact.}$$

En particulier, le théorème de Banach-Alaoglu est encore vrai dans les espaces vectoriels topologiques localement convexes séparés. $\square$

Remarque 3.3.3. Si E est séparable, la topologie faible* sur la boule est métrisable : si (x_n) est une suite dense dans $B_E(0, 1)$, on peut définir par exemple

$$\forall f, g \in B_{E^*}(0, 1), \quad d(f, g) = \sum_{n=0}^{+\infty} 2^{-n} |(f - g)(x_n)|.$$

DSK La remarque énonce le résultat suivant : si E est un espace de Banach séparable, alors la topologie faible-* sur la boule unité fermée $B_{E^*}(0, 1)$ de son dual E^* est métrisable. Plus précisément, si $(x_n)_{n \in \mathbb{N}}$ est une suite dense dans la boule unité $B_E(0, 1)$ de E , alors une distance compatible avec cette topologie peut être définie par :

$$d(f, g) = \sum_{n=0}^{+\infty} 2^{-n} |(f - g)(x_n)|.$$

La notation $\sum_0^{+\infty} 2^{-n} |(f - g)(x_n)|$ semble être une erreur de frappe ou une notation mal formatée. La version correcte est celle donnée ci-dessus, où :

- 2^{-n} est un poids assurant la convergence de la série,
- $(f - g)(x_n)$ désigne l'évaluation de la forme linéaire $f - g \in E^*$ au point $x_n \in E$.

Cette distance est bien définie car $|(f - g)(x_n)| \leq \|f - g\|_{E^*} \|x_n\|_E \leq \|f - g\|_{E^*}$, et la série converge normalement. Elle induit la topologie faible-* sur $B_{E^*}(0, 1)$ car une suite (f_k) converge vers f pour d si et seulement si $f_k(x_n) \rightarrow f(x_n)$ pour tout n , ce qui équivaut à la convergence faible-* par densité de (x_n) dans $B_E(0, 1)$.

Si K est un espace métrique compact, $E = C(K)$ muni de la norme uniforme est un espace de Banach séparable. Son dual E^* est l'espace des mesures signées sur K (théorème de représentation de Riesz, voir chapitre 8).

D'après le théorème de Banach-Alaoglu, si (μ_n) est une suite de mesures de probabilités, il existe alors une probabilité μ sur K telle que, à extraction d'une sous-suite près, $\mu_n \rightharpoonup \mu$ faiblement*.

3.4 Notion d'espace réflexif

Soit E un espace de Banach. On voit facilement que E s'identifie à un sous-espace vectoriel de E^{**} grâce à l'injection

$$j : x \in E \mapsto j(x) \in E^{**} \text{ où } \langle j(x), f \rangle = \langle f, x \rangle \text{ pour tout } f \in E^*,$$

qui est une isométrie d'espace vectoriel normé.

Soit E un espace de Banach. Définissons l'application $j : E \rightarrow E^{**}$ par

$$j(x)(f) = f(x), \quad \forall f \in E^*.$$

Il est clair que $j(x)$ est un élément de E^{**} , car il s'agit d'une application linéaire continue sur E^* . L'application j est donc bien définie.

Montrons que j est une isométrie d'espace vectoriel normé, c'est-à-dire que $\|j(x)\| = \|x\|$ pour tout $x \in E$. Par définition de la norme sur E^{**} , nous avons

$$\|j(x)\| = \sup_{\|f\| \leq 1} |j(x)(f)| = \sup_{\|f\| \leq 1} |f(x)|.$$

Or, la norme de x dans E est définie par

$$\|x\| = \sup_{\|f\| \leq 1} |f(x)|.$$

Ainsi, nous obtenons $\|j(x)\| = \|x\|$, ce qui montre que j est une isométrie.

Enfin, j est injective car si $j(x) = 0$, alors pour tout $f \in E^*$, $f(x) = 0$, ce qui implique que $x = 0$ puisque E^* sépare les points de E . Ainsi, j est une isométrie sur son image, et E s'identifie naturellement à un sous-espace de E^{**} .

Définition 3.4.1. Soit E un espace de Banach. On dit que E est réflexif s'il est le dual de son dual, i.e. si l'application j définie précédemment est surjective.

Pour un espace réflexif, on a donc $\sigma(E, E^*) = \sigma(E^{**}, E^*)$ et la boule unité est faiblement compacte. Il s'agit en fait d'une caractérisation des espaces réflexifs :

Théorème 3.4.2. (de Kakutani). - Soit E un espace de Banach. Alors E est réflexif si et seulement si sa boule unité est compacte pour $\sigma(E, E^*)$.

La preuve de ce résultat est un peu technique car elle nécessite de jongler avec les différentes topologies. Elle repose sur les deux lemmes suivants.

Lemme 3.4.3. (de Helly). - Soit E un espace de Banach. On se donne $f_1, \dots, f_n$ dans E^* et $\alpha_1, \dots, \alpha_n$ des réels. Alors les deux propriétés suivantes sont équivalentes :

(i) pour tout $\varepsilon > 0$, il existe $x_\varepsilon \in B_E$ tel que

$$\forall i = 1, \dots, n, \quad |f_i(x_\varepsilon) - \alpha_i| < \varepsilon;$$

(ii) pour tout $\beta_1, \dots, \beta_n$ réels, on a

$$\left| \sum_{i=1}^n \beta_i \alpha_i \right| \leq \left\| \sum_{i=1}^n \beta_i f_i \right\|.$$

Démonstration: Supposons d'abord que pour tout $\varepsilon > 0$, il existe $x_\varepsilon \in B_E$ tel que

$$\forall i = 1, \dots, n, \quad |f_i(x_\varepsilon) - \alpha_i| < \varepsilon.$$

Soient $\beta_1, \dots, \beta_n$ des réels quelconques. On a

$$\left| \sum_{i=1}^n \beta_i \alpha_i \right| \leq \left\| \sum_{i=1}^n \beta_i f_i(x_\varepsilon) \right\| + \varepsilon \sum_{i=1}^n |\beta_i| \leq \left\| \sum_{i=1}^n \beta_i f_i \right\| + \varepsilon \sum_{i=1}^n |\beta_i|$$

d'où l'on déduit (ii) en faisant tendre $\varepsilon \rightarrow 0$.

En multipliant par β_i et en sommant, on obtient :

$$\left| \sum_{i=1}^n \beta_i \alpha_i - \sum_{i=1}^n \beta_i f_i(x_\varepsilon) \right| \leq \sum_{i=1}^n |\beta_i| |\alpha_i - f_i(x_\varepsilon)|. \quad (3.1)$$

Par l'hypothèse (2), on majore :

$$\sum_{i=1}^n |\beta_i| |\alpha_i - f_i(x_\varepsilon)| < \varepsilon \sum_{i=1}^n |\beta_i|. \quad (3.2)$$

On en déduit :

$$\left| \sum_{i=1}^n \beta_i \alpha_i \right| \leq \left| \sum_{i=1}^n \beta_i f_i(x_\varepsilon) \right| + \varepsilon \sum_{i=1}^n |\beta_i|. \quad (3.3)$$

Enfin, en utilisant la norme sur E^* , on a directement :

$$\left| \sum_{i=1}^n \beta_i f_i(x_\varepsilon) \right| \leq \left\| \sum_{i=1}^n \beta_i f_i \right\| \cdot \|x_\varepsilon\| \leq \left\| \sum_{i=1}^n \beta_i f_i \right\|, \quad (3.4)$$

car $x_\varepsilon \in B_E$, donc $\|x_\varepsilon\| \leq 1$. Cela conclut la démonstration de l'inégalité annoncée.

On établit ensuite la réciproque par contraposition. On définit $F = (f_1, \dots, f_n)$. Par définition, $\overline{F(B_E)}$ est un convexe fermé de $\mathbb{R}^n$. D'après le théorème de Hahn-Banach, si $\alpha = (\alpha_1, \dots, \alpha_n) \notin \overline{F(B_E)}$, il existe une forme linéaire sur $\mathbb{R}^n$ qui sépare strictement α et $\overline{F(B_E)}$, autrement dit il existe $\beta_1, \dots, \beta_n$ et γ tels que :

$$\forall x \in B_E, \quad \sum_{i=1}^n \beta_i f_i(x) < \gamma < \sum_{i=1}^n \beta_i \alpha_i$$

ce qui implique en particulier que

$$\left\| \sum_{i=1}^n \beta_i f_i \right\| < \left| \sum_{i=1}^n \beta_i \alpha_i \right|. \square$$

Lemme 3.4.4. (de Goldstine). - Soit E un espace de Banach. On définit

$$j : x \in E \mapsto j(x) \in E^{**} \text{ où } \langle j(x), f \rangle = \langle f, x \rangle \text{ pour tout } f \in E^*,$$

Alors $j(B_E)$ est dense dans $B_{E^{**}}$ pour la topologie faible* $\sigma(E^{**}, E^*)$.

Démonstration: Soit $\eta \in B_{E^{**}}$ et V un voisinage de η pour la topologie $\sigma(E^{**}, E^*)$. Il s'agit de montrer que $V \cap j(B_E) \neq \emptyset$.

Sans perte de généralité, on peut supposer qu'il existe $\varepsilon > 0, f_1, \dots, f_n$ dans E^* tels que

$$V = \{\xi \in E^{**} \mid |(\xi - \eta)(f_i)| < \varepsilon \text{ pour tout } i = 1, \dots, n\}.$$

Si on pose $\alpha_i = \eta(f_i)$, on a

$$\forall \beta_1, \dots, \beta_n \in \mathbb{R}^n, \quad \left| \sum_{i=1}^n \beta_i \alpha_i \right| = \left| \eta \left(\sum_{i=1}^n \beta_i f_i \right) \right| \leq \left\| \sum_{i=1}^n \beta_i f_i \right\|.$$

D'après le lemme d'Helly, il existe alors $x_\varepsilon \in B_E$ tel que

$$\forall i = 1, \dots, n, \quad |f_i(x_\varepsilon) - \alpha_i| < \varepsilon,$$

ce qui signifie exactement que $j(x_\varepsilon) \in V$, et donc que $V \cap j(B_E) \neq \emptyset$. $\square$

Démonstration du théorème de Kakutani. - Supposons d'abord que E est réflexif, de sorte que l'isométrie

$$j : x \in E \mapsto j(x) \in E^{**} \text{ où } \langle j(x), f \rangle = \langle f, x \rangle \text{ pour tout } f \in E^*,$$

est surjective : $j(B_E) = B_{E^{**}}$. D'après le théorème de Banach-Alaoglu, $B_{E^{**}}$ est compact pour la topologie $\sigma(E^{**}, E^*)$.

Comme j^{-1} est une isométrie, pour tout $f \in E^*$, $f \circ j^{-1}$ est continue pour $\sigma(E^{**}, E^*)$. Cela implique que j^{-1} est continue de $(E^{**}, \sigma(E^{**}, E^*))$ vers $(E, \sigma(E, E^*))$. La boule B_E est donc faiblement compacte dans E .

Réciproquement, si B_E est faiblement compacte, comme j est une isométrie de E sur E^{**} , elle est continue de $(E, \sigma(E, E^*))$ dans $(E^{**}, \sigma(E^{**}, E^*))$ d'après le Corollaire 3.2.2. Comme $\sigma(E^{**}, E^*)$ est moins fine que $\sigma(E^{**}, E^{***})$, j est aussi continue de $(E, \sigma(E, E^*))$ dans $(E^{**}, \sigma(E^{**}, E^*))$. Cela implique que $j(B_E)$ est compact pour $\sigma(E^{**}, E^*)$.

Le lemme de Goldstine montre alors que $B_{E^{**}} = B_E$, et donc $E^{**} = E$. $\square$

1) Pourquoi est-il affirmé que "Comme j^{-1} est une isométrie, pour tout $f \in E^*$, $f \circ j^{-1}$ est continue pour $\sigma(E^{**}, E^*)$ " ?

Comme j^{-1} est une isométrie, c'est-à-dire une bijection linéaire continue de $j(E)$ (muni de la topologie induite par $\sigma(E^{**}, E^*)$) vers E (muni de la topologie faible $\sigma(E, E^*)$), on peut considérer la composition $f \circ j^{-1}$ pour tout $f \in E^*$.

Or, pour tout $x^{**} \in j(E)$, $f \circ j^{-1}(x^{**}) = \langle f, j^{-1}(x^{**}) \rangle = \langle x^{**}, f \rangle$, ce qui montre que $f \circ j^{-1}$ coïncide avec l'évaluation canonique $\langle x^{**}, f \rangle$.

Ainsi, $f \circ j^{-1}$ est simplement la restriction à $j(E)$ de la fonctionnelle linéaire $x^{**} \mapsto \langle x^{**}, f \rangle$, qui est $\sigma(E^{**}, E^*)$ -continue.

Donc, $f \circ j^{-1}$ est $\sigma(E^{**}, E^*)$ -continue sur $j(E)$, ce qui justifie l'affirmation.

2) Pourquoi est-il affirmé que "Cela implique que j^{-1} est continue de $(E^{**}, \sigma(E^{**}, E^*))$ vers $(E, \sigma(E, E^*))$." ? Soit $x_\alpha^{**} \rightarrow x^{**}$ pour la topologie $\sigma(E^{**}, E^*)$.

Alors pour tout $f \in E^*$, on a $\langle x_\alpha^{**}, f \rangle \rightarrow \langle x^{**}, f \rangle$. Comme $f \circ j^{-1}(x^{**}) = \langle x^{**}, f \rangle$, cela signifie que $f(j^{-1}(x_\alpha^{**})) \rightarrow f(j^{-1}(x^{**}))$ pour tout $f \in E^*$, donc $j^{-1}(x_\alpha^{**}) \rightarrow j^{-1}(x^{**})$ faiblement dans E , i.e. pour la topologie $\sigma(E, E^*)$.

Autrement dit, la convergence dans $\sigma(E^{**}, E^*)$ entraîne la convergence dans $\sigma(E, E^*)$ par j^{-1} , donc j^{-1} est continue de $(E^{**}, \sigma(E^{**}, E^*))$ vers $(E, \sigma(E, E^*))$.

Une condition suffisante pour qu'un espace de Banach soit réflexif est qu'il soit uniformément convexe.

Définition 3.4.5. *Un espace vectoriel normé E est dit uniformément convexe si on a la propriété*

$$\forall \varepsilon > 0, \quad \exists \delta > 0, \quad \left(\|x\| \leq 1, \quad \|y\| \leq 1 \text{ et } \left\| \frac{1}{2}(x+y) \right\| \geq 1 - \delta \right) \Rightarrow \|x - y\| \leq \varepsilon.$$

Théorème 3.4.6. *(de Milman-Pettis). - Tout espace de Banach uniformément convexe est réflexif.*

Démonstration: On doit montrer que $j(E) = E^{**}$, ou de façon équivalente que $j(B_E) = B_{E^{**}}$. Comme $j(B_E)$ est fermée, il suffit par homogénéité de montrer que $j(B_E)$ est dense (pour la topologie forte) dans la sphère

$$S = \{\eta \in E^{**} \mid \|\eta\|_{E^{**}} = 1\}.$$

On se donne alors $\eta \in S$ et $\varepsilon > 0$, et on va montrer qu'il existe $x \in B_E$ tel que $\|j(x) - \eta\| \leq \varepsilon$.

Comme E est uniformément convexe, il existe $\delta > 0$ tel que

$$\forall x, y \in B_E, \quad \|x - y\| > \varepsilon \Rightarrow \|x + y\| \leq 2(1 - \delta).$$

On choisit alors $f \in E^*$ telle que

$$\|f\|_{E^*} = 1 \text{ et } \eta(f) \geq 1 - \frac{\delta}{2}.$$

D'après le lemme de Goldstine, $j(B_E)$ est dense dans $B_{E^{**}}$ pour la topologie faible* $\sigma(E^{**}, E^*)$: il existe donc $x \in B_E$ tel que

$$|\eta(f) - f(x)| < \frac{\delta}{2}$$

- si $\|j(x) - \eta\|_{E^{**}} \leq \varepsilon$, on a la propriété voulue.
- si $\|j(x) - \eta\|_{E^{**}} > \varepsilon$, comme $E^{**} \setminus B_{E^{**}}(j(x), \varepsilon)$ est ouvert pour la topologie $\sigma(E^{**}, E^*)$, il résulte du lemme de Goldstine qu'il existe $y \in B_E$ tel que

$$\|j(y) - j(x)\| = \|x - y\| > \varepsilon \text{ et } |\eta(f) - f(y)| < \frac{\delta}{2}$$

On a alors

$$\frac{1}{2}\|x + y\| \geq (1 - \delta)$$

puisque

$$1 - \frac{\delta}{2} \leq \eta(f) < \frac{1}{2}f(x+y) + \frac{\delta}{2} \leq \frac{1}{2}\|x+y\| + \frac{\delta}{2}$$

ce qui aboutit à une contradiction. □

Pour tout ouvert $\Omega \subset \mathbb{R}^N$, on en rappelle sans démonstration quelques propriétés fondamentales des espaces $L^p(\Omega)$ résumés au paragraphe 1.4.3.

Propriété 3.4.7. Soit Ω un ouvert de $\mathbb{R}^N$. Si $p \in [1, \infty[$,

- l'espace $L^p(\Omega)$ est complet ;
- l'espace $L^p(\Omega)$ est séparable ;
 $-C_c^\infty(\Omega)$ (et a fortiori $C_c(\Omega)$) est dense dans $L^p(\Omega)$.
 Un autre résultat important est que ces espaces sont réflexifs si $p \in]1, +\infty[$, et que les bornés sont donc faiblement compacts.

Propriété 3.4.8. Soit Ω un ouvert de $\mathbb{R}^N$. Si $p \in]1, \infty[$, $L^p(\Omega)$ est uniformément convexe, donc réflexif.

Démonstration: Un espace normé $(X, \|\cdot\|)$ est uniformément convexe si, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que, pour tous $x, y \in X$ avec $\|x\| = \|y\| = 1$ et $\|x - y\| \geq \varepsilon$, on ait :

$$\left\| \frac{x+y}{2} \right\| \leq 1 - \delta.$$

L'uniforme convexité de $L^p(\Omega)$ découle de l'inégalité de Clarkson, qui affirme que pour tout $u, v \in L^p(\Omega)$, on a :

- Si $p \geq 2$:

$$\|u+v\|_p^p + \|u-v\|_p^p \leq (\|u\|_p + \|v\|_p)^p + \|\|u\|_p - \|v\|_p\|^p.$$

- Si $1 < p \leq 2$:

$$\|u+v\|_p^p + \|u-v\|_p^p \geq (\|u\|_p + \|v\|_p)^p + \|\|u\|_p - \|v\|_p\|^p.$$

Prenons $u, v \in L^p(\Omega)$ de norme 1, et supposons que $\|u - v\|_p \geq \varepsilon$. L'inégalité de Clarkson pour $p \geq 2$ donne :

$$\left\| \frac{u+v}{2} \right\|_p^p + \left\| \frac{u-v}{2} \right\|_p^p \leq 1.$$

Puisque $\|u - v\|_p \geq \varepsilon$, alors $\left\| \frac{u-v}{2} \right\|_p^p \geq (\varepsilon/2)^p$. Donc,

$$\left\| \frac{u+v}{2} \right\|_p^p \leq 1 - (\varepsilon/2)^p.$$

D'où,

$$\left\| \frac{u+v}{2} \right\|_p \leq (1 - (\varepsilon/2)^p)^{1/p}.$$

Par un développement limité,

$$(1 - (\varepsilon/2)^p)^{1/p} \approx 1 - \frac{1}{p}(\varepsilon/2)^p.$$

On pose $\delta = \frac{1}{p}(\varepsilon/2)^p$, et on conclut que $L^p(\Omega)$ est uniformément convexe.

Un espace uniformément convexe est automatiquement réflexif. Donc, $L^p(\Omega)$ est réflexif pour $1 < p < \infty$.

□

Théorie de distributions (ou de Schwartz)

L'espace des distributions est, en un sens convenable, le "plus petit espace" contenant les fonctions continues, et où la dérivation est partout définie. La théorie des distributions, élaborée par Schwartz à la fin des années quarante, est l'aboutissement d'un processus s'étalant sur plus d'un demi-siècle et comprenant, entre autres, le calcul symbolique de l'ingénieur Heaviside (1893), le formalisme introduit par le physicien Dirac (1926), les "parties finies" d'intégrales divergentes de Hadamard (1932) et les dérivées généralisées de Sobolev (1936). On trouvera une présentation complète de cette théorie par exemple dans [13], [18].

L'idée de départ consiste à changer de point de vue pour décrire les fonctions : plutôt que de considérer une fonction f de variable réelle comme la collection de ses valeurs $f(x)$ où x parcourt $\mathbb{R}$, on décrit f par la collection de ses moyennes pondérées $\int f' \varphi(x) dx$ où φ parcourt $C_c^\infty(\mathbb{R})$. La formule d'intégration par parties

$$\int_{\mathbb{R}} f' \varphi(x) dx = - \int_{\mathbb{R}} f \varphi'(x) dx$$

permet alors de définir l'objet f' .

4.1 Définition et propriétés élémentaires

4.1.1 L'espace des distributions

Définition 4.1.1. Une distribution T sur un ouvert $\Omega \subset \mathbb{R}^N$ est une forme linéaire sur $C_c^\infty(\Omega)$

$$\varphi \in C_c^\infty(\Omega) \mapsto \langle T, \varphi \rangle \in \mathbb{R},$$

qui vérifie la propriété de continuité : pour tout K compact de Ω , il existe $p \in \mathbb{N}$ et $C > 0$ tels que

$$\forall \varphi \in C_c^\infty(\Omega) \text{ à support dans } K, \quad |\langle T, \varphi \rangle| \leq C \sup_{|\alpha| \leq p} \|\partial^\alpha \varphi\|_\infty. \quad (4.1.1)$$

Si on peut choisir p indépendamment du compact K , on dit que la distribution T est d'ordre fini p .

L'espace $\mathcal{D}'(\Omega)$ des distributions sur Ω est un espace vectoriel. C'est en fait le dual topologique de $C_c^\infty(\Omega)$.

- Pour chaque compact $K \subset \Omega$, l'ensemble des fonctions $C^\infty(\Omega)$ à support dans K muni des semi-normes

$$p_m(\varphi) = \sup_{|\alpha| \leq m} \|\partial^\alpha \varphi\|_\infty$$

est un espace de Fréchet, donc un espace métrisable.

- On définit alors la topologie de $C_c^\infty(\Omega) = \cup_{K \subset \subset \Omega} C^\infty(K)$ comme limite inductive des $C^\infty(K)$, i.e. comme la topologie la plus fine rendant continues les injections de $C^\infty(K)$ dans $C_c^\infty(\Omega)$;
- On vérifie que la condition (4.1.1) exprime bien la continuité pour cette topologie.

Lemme 4.1.2. Soient Ω un ouvert de $\mathbb{R}^N$, $T \in \mathcal{D}'(\Omega)$ et (φ_n) une suite de $C_c^\infty(\Omega)$ convergeant vers φ dans $C_c^\infty(\Omega)$. Alors

$$\lim_{n \rightarrow \infty} \langle T, \varphi_n \rangle = \langle T, \varphi \rangle.$$

Démonstration: Si (φ_n) converge vers φ dans $C_c^\infty(\Omega)$, alors par définition

- il existe un compact $K \subset \Omega$ tel que

$$\forall n \in \mathbb{N}, \quad \forall x \notin K, \quad \varphi_n(x) = 0$$

(en particulier φ est à support dans K),

- pour tout $\alpha \in \mathbb{N}^N$, la suite $(\partial^\alpha \varphi_n)$ converge vers $\partial^\alpha \varphi$ uniformément sur K .

Puisque T est une distribution, il existe $C > 0$ et $p \in \mathbb{N}$ (dépendant uniquement de K) tels que

$$|\langle T, \varphi_n - \varphi \rangle| \leq C \sup_{|\alpha| \leq p} \|\partial^\alpha (\varphi_n - \varphi)\|_\infty.$$

ce qui donne la convergence attendue. □

$L_{loc}^1(\Omega)$ désigne l'espace des fonctions localement intégrables sur Ω , c'est-à-dire l'ensemble des fonctions $f : \Omega \rightarrow \mathbb{R}$ telles que pour tout compact $K \subset \Omega$, on ait :

$$\int_K |f(x)| dx < \infty.$$

Le théorème d'annulation sur $L_{loc}^1(\Omega)$ affirme que si une fonction $f \in L_{loc}^1(\Omega)$ vérifie

$$\int_\Omega f(x)g(x)dx = 0$$

pour toute fonction test $g \in C_c^\infty(\Omega)$, alors $f = 0$ presque partout dans Ω . Autrement dit, une fonction localement intégrable qui s'annule sur toutes les fonctions tests est nulle en

tant qu'élément de l'espace des distributions.

Enoncé du théorème : Soit $\Omega \subset \mathbb{R}^n$ un ouvert. Si $u \in L^1_{loc}(\Omega)$ vérifie

$$\int_{\Omega} u \varphi dx = 0 \quad \text{pour tout } \varphi \in C_c^\infty(\Omega),$$

alors $u = 0$ presque partout dans Ω .

Le théorème d'annulation sur $L^1_{loc}(\Omega)$

$$\left(\forall v \in C_c^\infty(\Omega), \quad \int_{\Omega} u v dx = 0 \right) \Rightarrow u = 0 \text{ presque partout}$$

permet d'identifier $L^1_{loc}(\Omega)$ à un sous-espace de $\mathcal{D}'(\Omega)$. A chaque $f \in L^1_{loc}(\Omega)$, on associe en effet (de manière injective) la distribution T_f définie par

$$\forall \varphi \in C_c^\infty(\Omega), \quad \langle T_f, \varphi \rangle = \int_{\Omega} f \varphi dx$$

de sorte que

$$\forall \varphi \in C_c^\infty(\Omega) \text{ à support dans } K, \quad |\langle T_f, \varphi \rangle| \leq \|f\|_{L^1(K)} \|\varphi\|_{\infty}$$

Définition 4.1.3. Soient Ω un ouvert de $\mathbb{R}^N$, et (T_n) une suite de $\mathcal{D}'(\Omega)$. On dit que (T_n) converge vers T dans $\mathcal{D}'(\Omega)$ si

$$\forall \varphi \in C_c^\infty(\Omega), \quad \langle T_n, \varphi \rangle \rightarrow \langle T, \varphi \rangle \text{ quand } n \rightarrow \infty.$$

La convergence au sens des distributions est une convergence très faible : elle est impliquée par exemple par la convergence dans $L^1_{loc}(\Omega)$, mais aussi par la convergence faible dans $L^1_{loc}(\Omega)$.

4.1.2 Dérivation au sens des distributions

Définition 4.1.4. Soient Ω un ouvert de $\mathbb{R}^N$, $T \in \mathcal{D}'(\Omega)$ et $\alpha \in \mathbb{N}^N$. La dérivée $\partial^\alpha T$ d'ordre α de la distribution T est la distribution définie par

$$\forall \varphi \in C_c^\infty(\Omega), \quad \langle \partial^\alpha T, \varphi \rangle = (-1)^{|\alpha|} \langle T, \partial^\alpha \varphi \rangle.$$

Il est facile de vérifier que la dérivée distribution coïncide avec la dérivée classique, en utilisant la formule de Green.

Soit $u \in C^1(\Omega)$. On considère T_u , la distribution associée à u , définie par

$$\langle T_u, \varphi \rangle = \int_{\Omega} u(x) \varphi(x) dx, \quad \forall \varphi \in C_c^\infty(\Omega).$$

D'après la définition de la dérivée au sens des distributions, on a

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

En explicitant T_u , il vient

$$\langle \partial^\alpha T_u, \varphi \rangle = (-1)^{|\alpha|} \int_{\Omega} u(x) \partial^\alpha \varphi(x) dx.$$

Utilisons maintenant la formule de Green (ou l'intégration par parties) : pour toute fonction test $\varphi \in C_c^\infty(\Omega)$,

$$\int_{\Omega} u(x) \partial^\alpha \varphi(x) dx = (-1)^{|\alpha|} \int_{\Omega} \partial^\alpha u(x) \varphi(x) dx.$$

Ainsi,

$$\langle \partial^\alpha T_u, \varphi \rangle = \int_{\Omega} \partial^\alpha u(x) \varphi(x) dx.$$

On reconnaît alors la distribution associée à $\partial^\alpha u$, notée $T_{\partial^\alpha u}$, qui est définie par

$$\langle T_{\partial^\alpha u}, \varphi \rangle = \int_{\Omega} \partial^\alpha u(x) \varphi(x) dx.$$

Donc, par unicité de la représentation des distributions par intégrale contre une fonction test, on conclut que

$$\partial^\alpha T_u = T_{\partial^\alpha u}.$$

Cela montre que la dérivée au sens des distributions coïncide avec la dérivée classique lorsque u est suffisamment régulière.

En utilisant cette remarque très simple et le lemme de fermeture suivant, on peut alors obtenir les règles de dérivation des distributions.

Lemme 4.1.5. *Soient Ω un ouvert de $\mathbb{R}^N$, et (T_n) une suite de $\mathcal{D}'(\Omega)$ convergeant vers T dans $\mathcal{D}'(\Omega)$. Alors, pour tout $\alpha \in \mathbb{N}^N$,*

$$\partial^\alpha T_n \rightarrow \partial^\alpha T \text{ au sens des distributions.}$$

Démonstration: On a

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

et ceci pour toute fonction $\varphi \in C_c^\infty(\Omega)$. □

Propriété 4.1.6. - *Soient Ω un ouvert de $\mathbb{R}^N$ et $T \in \mathcal{D}'(\Omega)$.*

— *Pour toute $f \in C^\infty(\Omega)$, $fT \in \mathcal{D}'(\Omega)$ et*

$$\forall \alpha \in \mathbb{N}^N, \quad \partial^\alpha (fT) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \partial^{\alpha-\beta} f \partial^\beta T$$

avec les notations usuelles pour les multi-indices et les coefficients binomiaux.

— *Pour tout difféomorphisme $\phi \in C^\infty(\omega, \Omega)$, $T \circ \phi \in \mathcal{D}'(\omega)$ et*

$$\partial_j (T \circ \phi) = \sum_k (\partial_j \phi_k) (\partial_k T \circ \phi).$$

Démonstration: Les règles de calcul s'obtiennent à partir du calcul différentiel usuel par dualité.

— On définit la distribution produit fT par

$$\forall \varphi \in C_c^\infty(\Omega), \quad \langle fT, \varphi \rangle = \langle T, f\varphi \rangle$$

de sorte que, pour tout compact K ,

$$\begin{aligned} |\langle fT, \varphi \rangle| &\leq C \sup_{|\alpha| \leq p} \|\partial^\alpha(f\varphi)\|_\infty \\ &\leq \tilde{C} \sup_{|\alpha| \leq p} \|\partial^\alpha \varphi\|_\infty \end{aligned}$$

pour toute fonction $\varphi \in C^\infty$ à support dans K , où $\tilde{C}$ ne dépend que de C et de f (par la formule de Leibniz).

La constante $\tilde{C}$ est obtenue en utilisant la formule de Leibniz pour la dérivation d'un produit de fonctions. Plus précisément, si l'on écrit $f\varphi$ comme un produit, alors ses dérivées successives peuvent être exprimées sous forme de combinaisons linéaires des dérivées de f et φ , chacune étant bornée par une constante dépendant de f . Ainsi, $\tilde{C}$ est une constante qui dépend uniquement des dérivées de f jusqu'à un certain ordre et de la constante C associée à la distribution T .

On a alors

$$\begin{aligned} \langle \partial_j(fT), \varphi \rangle &= -\langle fT, \partial_j \varphi \rangle = -\langle T, f\partial_j \varphi \rangle \\ &= -\langle T, \partial_j(f\varphi) - \varphi \partial_j f \rangle \\ &= \langle \partial_j T, f\varphi \rangle + \langle (\partial_j f) T, \varphi \rangle \end{aligned}$$

d'où $\partial_j(fT) = (\partial_j f) T + f(\partial_j T)$ dans $\mathcal{D}'(\Omega)$. Par récurrence sur $|\alpha|$, on peut établir la formule de Leibniz annoncée.

L'égalité

$$-\langle fT, \partial_j \varphi \rangle = -\langle T, f\partial_j \varphi \rangle$$

est une conséquence de la bilinéarité de l'action d'une distribution sur une fonction test. En effet, pour une distribution T et des fonctions test f, φ , nous avons :

$$\langle fT, \partial_j \varphi \rangle = \langle T, f\partial_j \varphi \rangle.$$

Cela résulte du fait que la multiplication d'une distribution par une fonction régulière est définie par :

$$\langle fT, \varphi \rangle = \langle T, f\varphi \rangle.$$

Ainsi, en appliquant cette propriété avec la fonction test $\partial_j \varphi$, on obtient directement l'égalité souhaitée.

L'opérateur ∂_j désigne ici la dérivation partielle par rapport à la j -ième variable. Dans le cadre des distributions, $\partial_j T$ est défini par dualité via l'identité :

$$\forall \varphi \in C_c^\infty(\Omega), \quad \langle \partial_j T, \varphi \rangle = -\langle T, \partial_j \varphi \rangle.$$

Il s'agit donc d'une extension de la dérivation classique aux distributions.

— On définit l'image inverse $T \circ \phi$ de T par le changement de variable ϕ par

$$\forall \varphi \in C_c^\infty(\omega), \quad \langle T \circ \phi, \varphi \rangle = \langle T, |\det D\phi^{-1}| \varphi \circ \phi^{-1} \rangle.$$

Le déterminant $\det D\phi^{-1}$ apparaît naturellement lorsqu'on effectue un changement de variable dans une intégrale. Soit $\phi : \omega \rightarrow \Omega$ un difféomorphisme. Pour toute fonction $\psi \in C_c^\infty(\Omega)$, le changement de variable $x = \phi(y)$ implique :

$$\int_{\Omega} \psi(x) dx = \int_{\omega} \psi(\phi(y)) |\det D\phi(y)| dy.$$

Par conséquent, en définissant $T \circ \phi$ de manière à respecter cette transformation, on introduit naturellement le facteur $|\det D\phi^{-1}|$ pour assurer l'invariance sous le changement de variable.

En utilisant les règles de dérivation des fonctions composées, la formule de Leibniz et le fait que $|\det D\phi^{-1}|$ est dans $C^\infty(\Omega)$, on obtient alors que, pour tout compact K ,

$$\begin{aligned} \langle T \circ \phi, \varphi \rangle &\leq C \sup_{|\alpha| \leq p} \|\partial^\alpha (|\det D\phi^{-1}| \varphi \circ \phi^{-1})\|_\infty \\ &\leq \tilde{C} \sup_{|\alpha| \leq p} \|(\partial^\alpha \varphi) \circ \phi^{-1}\|_{L^\infty(K)} \\ &\leq \tilde{C} \sup_{|\alpha| \leq p} \|(\partial^\alpha \varphi)\|_{L^\infty(\phi^{-1}(K))} \end{aligned}$$

pour toute fonction $\varphi \in C^\infty$ à support dans $\phi^{-1}(K)$.

On a alors

$$\langle \partial_j(T \circ \phi), \varphi \rangle = -\langle T, |\det D\phi^{-1}| (\partial_j \varphi) \circ \phi^{-1} \rangle.$$

Par ailleurs

$$\begin{aligned} \langle \sum_k (\partial_k T \circ \phi) \partial_j \phi_k, \varphi \rangle &= \sum_k \langle \partial_k T, |\det D\phi^{-1}| (\partial_j \phi_k \varphi) \circ \phi^{-1} \rangle \\ &= -\sum_k \langle T, \partial_k (|\det D\phi^{-1}| (\partial_j \phi_k \circ \phi^{-1}) (\varphi \circ \phi^{-1})) \rangle. \end{aligned}$$

Or, pour tout $g \in C_c^\infty(\Omega)$,

$$\begin{aligned} \int_{\omega} \partial_j(g \circ \phi) dy &= \sum_k \int_{\omega} \partial_j \phi_k (\partial_k g \circ \phi) dy \\ &= \sum_k \int_{\Omega} (\partial_j \phi_k \circ \phi^{-1}) \partial_k g |\det D\phi^{-1}| dx = 0 \end{aligned}$$

d'où l'on déduit que

$$\partial_k (|\det D\phi^{-1}| (\partial_j \phi_k \circ \phi^{-1})) = 0.$$

La dernière égalité s'obtient par un changement de variables $y = \phi(x)$, ce qui donne

$$\int_{\omega} \partial_j(g \circ \phi)(y) dy = \int_{\Omega} \partial_j(g \circ \phi)(\phi(x)) |\det D\phi(x)| dx.$$

Or,

$$\partial_j(g \circ \phi)(\phi(x)) = \sum_k \partial_j \phi_k(x) \cdot \partial_k g(\phi(x)),$$

donc

$$\int_{\omega} \partial_j(g \circ \phi)(y) dy = \sum_k \int_{\Omega} \partial_j \phi_k(x) \cdot \partial_k g(\phi(x)) \cdot |\det D\phi(x)| dx.$$

En posant $x = \phi^{-1}(y)$, on obtient

$$\sum_k \int_{\Omega} (\partial_j \phi_k \circ \phi^{-1})(y) \cdot \partial_k g(y) \cdot |\det D\phi^{-1}(y)| dy.$$

Finalement, comme $g \in C_c^\infty(\Omega)$, on a

$$\int_{\omega} \partial_j (g \circ \phi)(y) dy = 0.$$

On obtient finalement

$$\begin{aligned} \left\langle \sum_k (\partial_k T \circ \phi) \partial_j \phi_k, \varphi \right\rangle &= - \sum_k \langle T, |\det D\phi^{-1}| (\partial_j \phi_k \circ \phi^{-1}) \partial_k (\varphi \circ \phi^{-1}) \rangle \\ &= - \langle T, |\det D\phi^{-1}| (\partial_j \varphi) \circ \phi^{-1} \rangle, \end{aligned}$$

ce qui conclut la preuve. $\square$

4.2 Quelques exemples

4.2.1 Mesures de Radon

L'ensemble $\mathcal{M}_+(\Omega)$ des mesures positives localement finies sur Ω s'identifie à un sous-ensemble de $\mathcal{D}'(\Omega)$ via l'application

$$\mu \in \mathcal{M}_+(\Omega) \mapsto T_\mu \in \mathcal{D}'(\Omega)$$

définie par

$$\forall \varphi \in C_c^\infty(\Omega), \quad \langle T_\mu, \varphi \rangle = \int_{\Omega} \varphi d\mu.$$

Cette application est en effet injective puisque $C_c^\infty(\Omega)$ est dense dans $C_c(\Omega)$.

Il est facile de vérifier que les distributions ainsi obtenues sont d'ordre 0 : pour tout compact $K \subset \Omega$,

$$\forall \varphi \in C^\infty \text{ à support dans } K, \quad |\langle T_\mu, \varphi \rangle| \leq \mu(K) \|\varphi\|_\infty.$$

L'application $\mu \mapsto T_\mu$ est bien définie, et nous devons montrer que les distributions obtenues sont d'ordre 0. Par définition, une distribution $T \in \mathcal{D}'(\Omega)$ est d'ordre 0 si pour tout compact $K \subset \Omega$, il existe une constante C_K telle que

$$\forall \varphi \in C_c^\infty(K), \quad |\langle T, \varphi \rangle| \leq C_K \|\varphi\|_\infty.$$

Dans notre cas, pour toute mesure positive localement finie $\mu \in \mathcal{M}_+(\Omega)$ et tout $\varphi \in C_c^\infty(K)$,

$$|\langle T_\mu, \varphi \rangle| = \left| \int_{\Omega} \varphi d\mu \right| \leq \int_K |\varphi| d\mu \leq \mu(K) \|\varphi\|_\infty.$$

Ainsi, on peut choisir $C_K = \mu(K)$, ce qui prouve que T_μ est une distribution d'ordre 0. La norme $\|\varphi\|_\infty$ désigne la norme uniforme de φ , définie par

$$\|\varphi\|_\infty = \sup_{x \in \Omega} |\varphi(x)|.$$

Elle représente donc la borne supérieure des valeurs absolues de φ sur Ω .

L'espace $L^\infty(\Omega)$ est l'ensemble des classes de fonctions mesurables bornées sur Ω , muni de la norme essentielle suprême définie par

$$\|f\|_{L^\infty} = \text{ess sup}_{x \in \Omega} |f(x)| = \inf\{M \geq 0 \mid |f(x)| \leq M \text{ presque partout}\}.$$

Contrairement à $\|\varphi\|_\infty$ qui est défini pour des fonctions continues, cette norme tient compte des ensembles de mesure nulle.

On a en fait la caractérisation suivante :

Théorème 4.2.1. *Soit Ω un ouvert de $\mathbb{R}^N$. Si T est une distribution positive sur Ω , i.e.*

$$\forall \varphi \in C^\infty(\Omega, \mathbb{R}^+), \quad \langle T, \varphi \rangle \geq 0$$

alors T est une mesure de Radon positive.

Démonstration: On commence par vérifier que T se prolonge en une forme linéaire positive sur $C_c(\Omega)$.

- Pour tout compact $K \subset \Omega$, on peut construire (par régularisation de $\mathbf{1}_{\{x/d(x,K) \leq \delta\}}$) une fonction $\psi \in C_c^\infty(\Omega)$ telle que

$$\psi(\Omega) \subset [0, 1], \text{ et } \psi|_K \equiv 1.$$

Pour toute fonction $\varphi \in C^\infty$ à support dans K , on a alors

$$\forall x \in \Omega, \quad -\|\varphi\|_\infty \psi(x) \leq \varphi(x) \leq \|\varphi\|_\infty \psi(x).$$

Par positivité de T , on en déduit

$$|\langle T, \varphi \rangle| \leq \|\varphi\|_\infty \langle T, \psi \rangle$$

ce qui signifie que T est continue pour la norme $L^\infty(K)$.

T est continue pour la norme $L^\infty(K)$ car, pour toute fonction $\varphi \in C^\infty$ à support dans un compact $K \subset \Omega$, on a l'inégalité suivante :

$$|\langle T, \varphi \rangle| \leq \|\varphi\|_\infty \langle T, \psi \rangle,$$

où $\psi \in C_c^\infty(\Omega)$ est une fonction positive telle que $\psi = 1$ sur K . Cela signifie que la forme linéaire T est bornée sur l'espace C^∞ à support dans K pour la norme $\|\cdot\|_\infty$, ce qui implique sa continuité par rapport à cette norme. Il est important de bien distinguer les notions. Lorsque l'on écrit que $|\langle T, \varphi \rangle| \leq \|\varphi\|_\infty \langle T, \psi \rangle$ pour toute $\varphi \in C^\infty$ à support dans un compact K , on montre que la forme linéaire T est **bornée** sur $C_c^\infty(K)$ pour la norme $\|\cdot\|_\infty$. Or, cette inégalité implique que T est **continue** par rapport à la norme $\|\cdot\|_\infty$ restreinte à $C_c^\infty(K)$. Autrement dit, **ce n'est pas une continuité sur tout l'espace**

$L^\infty(K)$, mais uniquement sur le **sous-espace** $C_c^\infty(K)$ **muni de la norme** L^∞ . Il ne s'agit donc pas d'un prolongement en tant que forme linéaire sur $L^\infty(K)$, mais bien d'une **continuité sur** $C_c^\infty(K)$ **pour la norme** L^∞ , ce qui suffit pour construire ensuite un prolongement à $C_c(\Omega)$ via densité.

Donc la bonne formulation serait :

La forme linéaire T est continue sur $C_c^\infty(K)$ muni de la norme $\|\cdot\|_\infty$, ce qui signifie que T est bornée sur cet espace.

et non une continuité sur l'ensemble de $L^\infty(K)$. Sur un compact $K \subset \Omega$, pour toute fonction $\varphi \in C_c^\infty(K)$, on peut considérer deux normes :

- La norme $\|\varphi\|_\infty := \sup_{x \in K} |\varphi(x)|$, qui est la norme uniforme usuelle sur les fonctions continues.
- La norme $L^\infty(K)$, définie comme la norme essentielle : $\|\varphi\|_{L^\infty(K)} := \text{ess sup}_{x \in K} |\varphi(x)|$, c'est-à-dire le plus petit M tel que $|\varphi(x)| \leq M$ presque partout.

Or, puisque toute fonction de $C_c^\infty(K)$ est continue, donc mesurable et bornée, et que les points de discontinuité sont inexistantes, on a :

$$\|\varphi\|_\infty = \|\varphi\|_{L^\infty(K)}.$$

Autrement dit, sur l'espace $C_c^\infty(K)$ (et plus généralement sur $C^0(K)$), **les normes** $\|\cdot\|_\infty$ et $\|\cdot\|_{L^\infty(K)}$ **coïncident**. Par conséquent, la continuité d'une forme linéaire T sur $C_c^\infty(K)$ pour la norme $\|\cdot\|_\infty$ est **équivalente** à sa continuité pour la norme L^∞ sur ce même espace. Donc, pour répondre directement à votre question :

Oui, sur l'espace $C_c^\infty(K)$, **la continuité pour la norme** $\|\cdot\|_\infty$ **entraîne bien la continuité pour la norme** $L^\infty(K)$, **car les deux normes coïncident sur cet espace.**

Pour une fonction mesurable et bornée u , on a toujours l'inégalité suivante :

$$\|u\|_{L^\infty} \leq \|u\|_\infty.$$

Voici pourquoi :

- $\|u\|_\infty := \sup_{x \in K} |u(x)|$ est le **supremum pointwise**, donc sensible à la valeur de $u(x)$ en chaque point.
- $\|u\|_{L^\infty}$ est le **esssup**, c'est-à-dire le plus petit M tel que $|u(x)| \leq M$ pour presque tout x . Autrement dit, on ignore les ensembles de mesure nulle.

Par conséquent, s'il existe des valeurs extrêmes atteintes seulement sur un ensemble de mesure nulle, elles sont comptées dans $\|u\|_\infty$, mais pas dans $\|u\|_{L^\infty}$. D'où :

$$\|u\|_{L^\infty} \leq \|u\|_\infty.$$

Et dans le cas où u est continue (en particulier pour $u \in C_c^\infty$), l'ensemble des points où la valeur maximale est atteinte est non négligeable (souvent même atteint), donc :

$$\|u\|_{L^\infty} = \|u\|_\infty.$$

Pour une fonction mesurable et bornée u , on a toujours l'inégalité suivante :

$$\|u\|_{L^\infty} \leq \|u\|_\infty.$$

Voici pourquoi :

- $\|u\|_\infty := \sup_{x \in K} |u(x)|$ est le **supremum pointwise**, donc sensible à la valeur de $u(x)$ en chaque point.
- $\|u\|_{L^\infty}$ est le **esssup**, c'est-à-dire le plus petit M tel que $|u(x)| \leq M$ pour presque tout x . Autrement dit, on ignore les ensembles de mesure nulle.

Par conséquent, s'il existe des valeurs extrêmes atteintes seulement sur un ensemble de mesure nulle, elles sont comptées dans $\|u\|_\infty$, mais pas dans $\|u\|_{L^\infty}$. D'où :

$$\|u\|_{L^\infty} \leq \|u\|_\infty.$$

Et dans le cas où u est continue (en particulier pour $u \in C_c^\infty$), l'ensemble des points où la valeur maximale est atteinte est non négligeable (souvent même atteint), donc :

$$\|u\|_{L^\infty} = \|u\|_\infty.$$

Conclusion : En général, $\|u\|_{L^\infty} \leq \|u\|_\infty$, et l'égalité a lieu si u est continue sur un compact.

- Pour toute fonction $g \in C_c(\Omega)$, g est la limite d'une suite (g_n) de $C_c^\infty(\Omega)$ uniformément supportée dans un compact K (régularisation par convolution). On pose alors

$$\langle T, g \rangle = \lim_{n \rightarrow \infty} \langle T, g_n \rangle$$

qui est bien définie car $(\langle T, g_n \rangle)$ est de Cauchy.

Le théorème de représentation de Riesz permet alors de conclure que T s'identifie à une mesure de Radon. $\square$ On a utilisé ici une propriété classique d'approximation des fonctions continues à support compact par des fonctions test régulières. Pour toute fonction $g \in C_c(\Omega)$, il existe une suite $(g_n) \subset C_c^\infty(\Omega)$ telle que $g_n \rightarrow g$ uniformément et tous les g_n ont leur support dans un même compact K (cela est obtenu par régularisation, typiquement en convoluant g avec un noyau mollifiant). Étant donné que T est continue sur $C_c^\infty(\Omega)$ pour la topologie de la convergence uniforme sur les compacts (comme montré précédemment), on peut définir :

$$\langle T, g \rangle := \lim_{n \rightarrow \infty} \langle T, g_n \rangle,$$

et cette limite est bien définie car la suite $(\langle T, g_n \rangle)$ est de Cauchy, ce qui découle de l'inégalité :

$$|\langle T, g_n - g_m \rangle| \leq \|g_n - g_m\|_\infty \langle T, \psi \rangle,$$

où ψ est une fonction majorante comme précédemment. Cela assure la convergence de la suite dans $\mathbb{R}$.

Attention : toute distribution ne s'écrit pas comme la différence de deux distributions positives. Par exemple, δ' - qui est d'ordre 1 - n'est pas la différence de deux distributions positives.

4.2.2 Valeurs principales et parties finies

Définition 4.2.2. Soit T_n la distribution sur $\mathbb{R}$ associée à la fonction intégrable

$$f_n(x) = \frac{1}{x} \mathbf{1}_{|x| \geq \frac{1}{n}}.$$

La valeur principale de $1/x$ est la distribution $vp\left(\frac{1}{x}\right)$ d'ordre 1 sur $\mathbb{R}$, obtenue en passant à la limite dans la suite (T_n) . C'est la dérivée-distribution de $\log|x|$.

Démonstration: Pour toute fonction $\varphi \in C_c^\infty(\mathbb{R})$ à support dans $[-a, a]$

$$\langle T_n, \varphi \rangle = \int_{\frac{1}{n} \leq |x| \leq a} \frac{\varphi(x)}{x} dx = \int_{\frac{1}{n} \leq |x| \leq a} \frac{\varphi(x) - \varphi(0)}{x} dx$$

par parité.

Comme la fonction $x \mapsto (\varphi(x) - \varphi(0))/x$ se prolonge par continuité en 0, elle est intégrable sur $[-a, a]$, et on a

$$\lim_{n \rightarrow \infty} \langle T_n, \varphi \rangle = \int_{|x| \leq a} \frac{\varphi(x) - \varphi(0)}{x} dx.$$

Nous prouvons que

$$\int_{\frac{1}{n} \leq |x| \leq a} \frac{\varphi(x)}{x} dx = \int_{\frac{1}{n} \leq |x| \leq a} \frac{\varphi(x) - \varphi(0)}{x} dx.$$

En ajoutant et retranchant $\varphi(0)$ dans le numérateur, nous avons

$$\frac{\varphi(x)}{x} = \frac{\varphi(x) - \varphi(0) + \varphi(0)}{x} = \frac{\varphi(x) - \varphi(0)}{x} + \frac{\varphi(0)}{x}.$$

En intégrant sur l'intervalle donné,

$$\int_{\frac{1}{n} \leq |x| \leq a} \frac{\varphi(x)}{x} dx = \int_{\frac{1}{n} \leq |x| \leq a} \frac{\varphi(x) - \varphi(0)}{x} dx + \varphi(0) \int_{\frac{1}{n} \leq |x| \leq a} \frac{dx}{x}.$$

Or, le dernier terme est nul car l'intégrale de $1/x$ donne $\log|x|$, qui est antisymétrique autour de 0, et donc son intégrale s'annule. Cela prouve l'égalité.

En utilisant à nouveau la parité, on obtient que la limite est indépendante de a . Par exemple,

$$\left\langle vp\left(\frac{1}{x}\right), \varphi \right\rangle = \int_{|x| \leq 1} \frac{\varphi(x) - \varphi(0)}{x} dx + \int_{|x| > 1} \frac{\varphi(x)}{x} dx$$

L'égalité

$$\left\langle vp\left(\frac{1}{x}\right), \varphi \right\rangle = \int_{|x| \leq 1} \frac{\varphi(x) - \varphi(0)}{x} dx + \int_{|x| > 1} \frac{\varphi(x)}{x} dx$$

s'obtient en décomposant l'intégrale sur deux intervalles :

- Pour $|x| \leq 1$, on utilise l'écriture avec $\varphi(x) - \varphi(0)$ comme démontré dans la question (1).
- Pour $|x| > 1$, l'intégrale est bien définie et ne nécessite pas de correction. Ainsi, on obtient directement la somme des deux termes annoncés.

- Pour toute fonction $\varphi \in C_c^\infty(\mathbb{R})$ à support dans $[-a, a]$, l'inégalité des accroissements finis donne

$$\left| \left\langle vp \left(\frac{1}{x} \right), \varphi \right\rangle \right| \leq 2a \|\varphi'\|_{\infty}.$$

Nous utilisons l'inégalité des accroissements finis :

$$\varphi(x) - \varphi(0) = x\varphi'(\xi), \quad \text{pour un } \xi \text{ entre } 0 \text{ et } x.$$

Ainsi,

$$\left| \frac{\varphi(x) - \varphi(0)}{x} \right| = |\varphi'(\xi)| \leq \|\varphi'\|_{\infty}.$$

En majorant l'intégrale correspondante,

$$\left| \int_{|x| \leq a} \frac{\varphi(x) - \varphi(0)}{x} dx \right| \leq \int_{|x| \leq a} |\varphi'(\xi)| dx \leq 2a \|\varphi'\|_{\infty},$$

on obtient l'inégalité souhaitée.

On en déduit que $vp \left(\frac{1}{x} \right)$ est bien une distribution, d'ordre au plus 1. En particulier, on peut la prolonger en une forme linéaire continue sur $C_c^1(K)$ pour tout compact $K \subset \mathbb{R}$. Elle n'est pas d'ordre 0 comme on le voit facilement en choisissant $\varphi_k = \rho_k * \mathbf{1}_{[0,1]}$, elle est donc exactement d'ordre 1.

— Reste à prouver l'identité $vp \left(\frac{1}{x} \right) = (\log |x|)'$. Pour toute fonction $\varphi \in C_c^{\infty}(\mathbb{R})$,

$$\langle (\log |x|)', \varphi \rangle = -\langle \log |x|, \varphi' \rangle = -\int \log |x| \varphi'(x) dx.$$

Comme $x \mapsto \log |x|$ est dans $L_{loc}^1(\mathbb{R})$, on a

$$\int \log |x| \varphi'(x) dx = \lim_{n \rightarrow 0} \int_{|x| \geq 1/n} \log |x| \varphi'(x) dx.$$

Par intégration par parties,

$$\int_{|x| \geq 1/n} \log |x| \varphi'(x) dx = - \int_{|x| \geq 1/n} \frac{\varphi(x)}{x} dx + \log n \left(\varphi \left(\frac{1}{n} \right) - \varphi \left(-\frac{1}{n} \right) \right).$$

Le second terme du membre de droite tend vers 0 quand $n \rightarrow \infty$ car φ est dérivable en 0, donc

$$\int \log |x| \varphi'(x) dx = - \lim_{n \rightarrow 0} \int_{|x| \geq 1/n} \frac{\varphi(x)}{x} dx = - \left\langle vp \left(\frac{1}{x} \right), \varphi \right\rangle,$$

ce qui est l'identité annoncée. $\square$

On a supposé que φ est dérivable en 0, donc en particulier φ est continue en 0 et l'on peut écrire le développement limité de φ en 0 :

$$\varphi(x) = \varphi(0) + x\varphi'(0) + o(x) \quad \text{quand } x \rightarrow 0.$$

En particulier, on a :

$$\varphi \left(\frac{1}{n} \right) = \varphi(0) + \frac{1}{n} \varphi'(0) + o \left(\frac{1}{n} \right), \quad \varphi \left(-\frac{1}{n} \right) = \varphi(0) - \frac{1}{n} \varphi'(0) + o \left(\frac{1}{n} \right).$$

Alors,

$$\varphi\left(\frac{1}{n}\right) - \varphi\left(-\frac{1}{n}\right) = 2 \cdot \frac{1}{n} \varphi'(0) + o\left(\frac{1}{n}\right).$$

Par conséquent,

$$\log n \left(\varphi\left(\frac{1}{n}\right) - \varphi\left(-\frac{1}{n}\right) \right) = \log n \left(2 \cdot \frac{1}{n} \varphi'(0) + o\left(\frac{1}{n}\right) \right).$$

Or $\log n/n \rightarrow 0$ quand $n \rightarrow \infty$, donc le terme ci-dessus tend bien vers 0 lorsque $n \rightarrow \infty$. De façon similaire, on définit les parties finies $\text{Pf}(x_+^\alpha)$ pour $\alpha \in]-2, -1[$ par dérivation de $x_+^{\alpha+1}/(\alpha+1)$.

Définition 4.2.3. La partie finie de x_+^α est la distribution $\text{Pf}(x_+^\alpha)$ d'ordre 1, définie par

$$\langle \text{Pf}(x_+^\alpha), \varphi \rangle = - \int_0^{+\infty} \frac{x^{\alpha+1}}{\alpha+1} \varphi'(x) dx = \lim_{n \rightarrow \infty} \int_{x \geq \frac{1}{n}} x^\alpha \varphi(x) dx.$$

4.2.3 Dérivées des indicatrices et formule de saut

- Soit K un compact de classe C^1 de $\mathbb{R}^N$. En utilisant une partition de l'unité et des changements de carte locaux, on peut montrer la formule de Stokes

$$\partial_{x_i}(\mathbf{1}_K) = -\nu_i d\sigma,$$

où ν est le vecteur unitaire normal à ∂K orienté vers l'extérieur et $d\sigma$ est la mesure de Lebesgue superficielle de ∂K . Cette formule n'est rien d'autre que la généralisation de la formule d'intégration par parties

$$\frac{d}{dx} \mathbf{1}_{[a,b]} = \delta_a - \delta_b.$$

Proposition 4.2.4. Soient Ω un ouvert de $\mathbb{R}^N$, K un compact de classe C^1 de Ω et $f \in C(\Omega)$ telle que $\partial_{x_i} f \in L_{loc}^1(\Omega)$. Alors

$$\partial_{x_i}(f \mathbf{1}_K) = (\partial_{x_i} f) \mathbf{1}_K - \nu_i f d\sigma$$

où ν est le vecteur unitaire normal à ∂K orienté vers l'extérieur et $d\sigma$ est la mesure de Lebesgue superficielle de ∂K .

Démonstration: On commence par se ramener au cas d'une fonction à support compact. Par régularisation de $\mathbf{1}_{\{x/d(x,K) \leq \delta\}}$ avec $\delta \leq \frac{1}{2}d(K, \partial\Omega)$, on peut construire une fonction $\psi \in C_c^\infty(\Omega)$ vérifiant

$$\psi(\Omega) \subset [0, 1] \quad \text{et} \quad \psi \equiv 1 \text{ sur un voisinage de } K.$$

Si on pose $g = \psi f$, $g \in C_c(\Omega)$ et $\partial_{x_i} g \in L^1(\Omega)$. Comme $f \equiv g$ sur un voisinage de K , il suffit de prouver la formule des sauts pour g .

On se ramène alors au cas où la fonction est de classe C^∞ par régularisation. On se donne une suite régularisante (ρ_n) définie par

$\forall x \in \mathbb{R}^N, \quad \rho_n(x) = n^N \rho(nx)$ où $\rho \in C^\infty(\mathbb{R}^N, [0, 1])$ à support dans la boule unité,

et on pose $g_n = \rho_n * g$. La formule de Stokes et la règle de dérivation du produit montrent que

$$\partial_{x_i}(g_n \mathbf{1}_K) = g_n(\partial_{x_i} \mathbf{1}_K) + (\partial_{x_i} g_n) \mathbf{1}_K = -\nu_i g_n d\sigma + (\partial_{x_i} g_n) \mathbf{1}_K.$$

Par le lemme de fermeture,

$$\partial_{x_i} g_n \rightarrow \partial_{x_i} g \text{ dans } \mathcal{D}'(\Omega), \quad \text{et } \partial_{x_i}(g_n \mathbf{1}_K) \rightarrow \partial_{x_i}(g \mathbf{1}_K) \text{ dans } \mathcal{D}'(\Omega).$$

Les résultats classiques sur la convolution montrent que

$$g_n \rightarrow g \text{ uniformément sur } \Omega, \quad \text{et } \partial_{x_i} g_n \rightarrow \partial_{x_i} g \text{ dans } L^1(\Omega).$$

On obtient alors la formule des sauts pour g . $\square$

4.3 Localisation, convolution, régularisation

Comme pour les fonctions, le procédé de régularisation des distributions le plus standard est la convolution. Dans le cas où les distributions sont définies sur un ouvert Ω , cette régularisation est partielle car elle ne va pas jusqu'au bord. On va donc définir la notion de support d'une distribution, mais pour simplifier la présentation, on exposera la suite des résultats dans le cas où $\Omega = \mathbb{R}^N$.

Définition 4.3.1. Soit Ω un ouvert de $\mathbb{R}^N$ et T une distribution sur Ω . Le support de T est le complémentaire du plus grand ouvert V tel que $T|_V \equiv 0$.

$$\text{Supp}(T) = \Omega \setminus \{x \in \Omega / \exists \omega \text{ voisinage de } x, \quad T|_\omega = 0\}.$$

On note $\mathcal{E}'(\Omega)$ l'ensemble des distributions à support compact.

On a la caractérisation suivante des distributions dont le support est réduit à un point :

Propriété 4.3.2. Soit $a \in \mathbb{R}^N$. Les distributions sur $\mathbb{R}^N$ dont le support est $\{a\}$ sont les combinaisons linéaires des dérivées de la masse de Dirac en a .

La distribution δ_a , appelée masse de Dirac en $a \in \mathbb{R}^n$, ou fonction de Dirac en un point $a \in \mathbb{R}^n$, est définie par :

$$\langle \delta_a, \varphi \rangle = \varphi(a), \quad \text{pour tout } \varphi \in \mathcal{D}(\mathbb{R}^n).$$

Sa dérivée d'ordre multi-indice α est définie par :

$$\langle \partial^\alpha \delta_a, \varphi \rangle = (-1)^{|\alpha|} \partial^\alpha \varphi(a), \quad \text{pour tout } \varphi \in \mathcal{D}(\mathbb{R}^n).$$

Soit $a \in \mathbb{R}$. La fonction caractéristique en a est définie par :

$$\chi_a(x) = \begin{cases} 1 & \text{si } x = a, \\ 0 & \text{sinon.} \end{cases}$$

La fonction χ_a n'est pas dérivable en aucun point, car elle est discontinue en a et constante ailleurs. En tant que distribution, sa dérivée au sens des distributions est donnée par :

$$\langle \chi'_a, \varphi \rangle = -\langle \chi_a, \varphi' \rangle = -\varphi'(a), \quad \text{pour tout } \varphi \in \mathcal{D}(\mathbb{R}).$$

Ainsi, la dérivée au sens des distributions de χ_a est la distribution $-\delta'_a$, la dérivée de la distribution de Dirac en a :

$$\chi'_a = -\delta'_a.$$

Démonstration: On commence par montrer que les distributions à support compact sont toujours d'ordre fini. Soient $T \in \mathcal{E}'(\Omega)$ et K un voisinage compact de $\text{Supp}(T)$. On peut construire (par régularisation d'une fonction caractéristique par exemple) une fonction $\psi \in C_c^\infty(\Omega)$ à support dans K telle que

$$\psi(\Omega) \subset [0, 1] \text{ et } \psi|_{\text{Supp}(T)} \equiv 1.$$

On a alors $T = T\psi$. Donc, pour toute fonction $\varphi \in C_c^\infty(\Omega)$, comme $\psi\varphi$ est à support dans K ,

$$\begin{aligned} |\langle T, \varphi \rangle| &= |\langle T, \psi\varphi \rangle| \leq C_K \sup_{|\alpha| \leq p_K} \|\partial^\alpha(\psi\varphi)\|_\infty \\ &\leq C \sup_{|\alpha| \leq p_K} \|\partial^\alpha \varphi\|_\infty \end{aligned}$$

par la formule de Leibniz. T est donc d'ordre p_K .

Soit maintenant T une distribution dont le support est réduit à $\{a\}$, disons d'ordre p . On se donne une fonction $\psi \in C_c^\infty(\Omega)$ à support dans $B(a, \delta)$ telle que

$$\psi(\Omega) \subset [0, 1] \text{ et } \psi \equiv 1 \text{ sur un voisinage de } a$$

et on pose $\psi_n(x) = \psi(a + n(x - a))$.

La formule de Taylor donne, pour toute fonction $\varphi \in C_c^\infty(\Omega)$

$$\varphi(x) = \psi(x) \left(\sum_{|\alpha| \leq p} \frac{(x-a)^\alpha}{\alpha!} \partial^\alpha \varphi(a) \right) + \tilde{\varphi}(x)$$

avec

$$\forall \alpha \text{ tel que } |\alpha| \leq p, \quad \partial^\alpha \tilde{\varphi}(a) = 0.$$

Soit $\varphi \in C_c^\infty(\Omega)$ et $p \in \mathbb{N}$. On fixe un point $a \in \Omega$. La formule de Taylor d'ordre p en a pour une fonction φ donne :

$$\varphi(x) = \sum_{|\alpha| \leq p} \frac{(x-a)^\alpha}{\alpha!} \partial^\alpha \varphi(a) + R_p(x),$$

où le reste $R_p(x)$ est une fonction lisse vérifiant :

$$\forall |\beta| \leq p, \quad \partial^\beta R_p(a) = 0.$$

On introduit maintenant une fonction $\psi \in C_c^\infty(\Omega)$ telle que $\psi \equiv 1$ dans un voisinage de a . On écrit :

$$\varphi(x) = \psi(x)\varphi(x) = \psi(x) \left(\sum_{|\alpha| \leq p} \frac{(x-a)^\alpha}{\alpha!} \partial^\alpha \varphi(a) + R_p(x) \right).$$

En posant :

$$\tilde{\varphi}(x) := \psi(x)R_p(x),$$

on obtient :

$$\varphi(x) = \psi(x) \left(\sum_{|\alpha| \leq p} \frac{(x-a)^\alpha}{\alpha!} \partial^\alpha \varphi(a) \right) + \tilde{\varphi}(x).$$

Enfin, comme $\psi \equiv 1$ au voisinage de a et R_p s'annule avec toutes ses dérivées jusqu'à l'ordre p en a , on a bien :

$$\forall \alpha \text{ tel que } |\alpha| \leq p, \quad \partial^\alpha \tilde{\varphi}(a) = 0.$$

Par définition du support de T ,

$$\langle T, \tilde{\varphi}(1 - \psi_n) \rangle = 0.$$

De plus, comme T est d'ordre p ,

$$|\langle T, \tilde{\varphi}\psi_n \rangle| \leq C \sup_{|\alpha| \leq p} \|\partial^\alpha (\tilde{\varphi}\psi_n)\|_\infty \leq \tilde{C}/n$$

en utilisant la formule de Leibniz et la formule de Taylor

$$\forall x \in B(a, \delta/n), \quad |\partial^\beta \tilde{\varphi}(x)| \leq C \left(\frac{1}{n} \right)^{p+1-|\beta|}.$$

On en déduit que $\langle T, \tilde{\varphi} \rangle = 0$.

Soit $\varphi \in C_c^\infty(\Omega)$ et écrivons son développement de Taylor d'ordre p autour du point a :

$$\varphi(x) = \sum_{|\alpha| \leq p} \frac{(x-a)^\alpha}{\alpha!} \partial^\alpha \varphi(a) + R_p(x),$$

où le reste $R_p(x)$ s'écrit, pour tout x proche de a , sous la forme :

$$R_p(x) = \sum_{|\alpha|=p+1} \frac{(x-a)^\alpha}{\alpha!} \int_0^1 (1-t)^p \partial^\alpha \varphi(a + t(x-a)) dt.$$

En posant $\tilde{\varphi}(x) := \varphi(x) - \sum_{|\alpha| \leq p} \frac{(x-a)^\alpha}{\alpha!} \partial^\alpha \varphi(a)$, on a donc :

$$\tilde{\varphi}(x) = R_p(x),$$

et en particulier, pour tout multi-indice β tel que $|\beta| \leq p$:

$$\partial^\beta \tilde{\varphi}(x) = \sum_{|\alpha|=p+1} \partial^\beta \left(\frac{(x-a)^\alpha}{\alpha!} \int_0^1 (1-t)^p \partial^\alpha \varphi(a + t(x-a)) dt \right).$$

En utilisant le fait que φ est lisse et à support compact, on a que les dérivées de φ sont bornées par une constante C sur un compact contenant a , et donc :

$$|\partial^\beta \tilde{\varphi}(x)| \leq C |x-a|^{p+1-|\beta|}.$$

Or, pour $x \in B(a, \delta/n)$, on a $|x - a| \leq \delta/n$, donc :

$$|\partial^\beta \tilde{\varphi}(x)| \leq C \left(\frac{\delta}{n} \right)^{p+1-|\beta|},$$

ce qui donne bien :

$$\forall x \in B(a, \delta/n), \quad |\partial^\beta \tilde{\varphi}(x)| \leq C \left(\frac{1}{n} \right)^{p+1-|\beta|}.$$

Autrement dit, T est une combinaison linéaire de δ_a et de ses dérivées jusqu'à l'ordre p .
□

Définition 4.3.3. Soient $T \in \mathcal{D}'(\mathbb{R}^N)$ et $\varphi \in C_c^\infty(\mathbb{R}^N)$. La convolution de T par φ est définie sur $\mathbb{R}^N$ par

$$T * \varphi(x) = \langle T, \tau_x \hat{\varphi} \rangle \text{ où } \hat{\varphi}(y) = \varphi(-y).$$

$\tau_x \hat{\varphi} = \hat{\varphi}(y - x)$ est l'opérateur de translation.

Propriété 4.3.4. Soient $T \in \mathcal{D}'(\mathbb{R}^N)$ et $\varphi \in C_c^\infty(\mathbb{R}^N)$. Alors

$$\text{Supp}(T * \varphi) \subset \text{Supp } T + \text{Supp } \varphi.$$

De plus, $T * \varphi \in C^\infty(\mathbb{R}^N)$ et on a

$$\partial^\alpha(T * \varphi) = T * (\partial^\alpha \varphi) = (\partial^\alpha T) * \varphi.$$

Démonstration: Soit $x \notin \text{Supp } T + \text{Supp } \varphi$, on a

$$\text{Supp } T \cap \text{Supp } (\tau_x \hat{\varphi}) = \text{Supp } T \cap (x - \text{Supp } \varphi) = \emptyset,$$

ce qui implique que $T * \varphi(x) = \langle T, \tau_x \hat{\varphi} \rangle = 0$.

Soient $x \in \mathbb{R}^N$ et $\alpha \in \mathbb{N}^N$ avec $|\alpha| = 1$. Pour tout $\varepsilon > 0$,

$$\frac{1}{\varepsilon}(T * \varphi(x + \varepsilon\alpha) - T * \varphi(x)) = \frac{1}{\varepsilon} \langle T, (\tau_{x+\varepsilon\alpha} \hat{\varphi} - \tau_x \hat{\varphi}) \rangle.$$

Or on peut montrer que $\frac{1}{\varepsilon}(\tau_{x+\varepsilon\alpha} \hat{\varphi} - \tau_x \hat{\varphi})$ converge vers $\tau_x(\widehat{\partial^\alpha \varphi})$ dans $C_c^\infty(\mathbb{R}^N)$ quand $\varepsilon \rightarrow 0$:

— pour tout $\varepsilon \in]0, 1]$,

$$\text{Supp} \left(\frac{1}{\varepsilon}(\tau_{x+\varepsilon\alpha} \hat{\varphi} - \tau_x \hat{\varphi}) \right) \subset \text{Supp}(\hat{\varphi}) + B(x, 1),$$

— pour tout $\beta \in \mathbb{N}^N$

$$\partial^\beta \left(\frac{1}{\varepsilon}(\tau_{x+\varepsilon\alpha} \hat{\varphi} - \tau_x \hat{\varphi}) - \tau_x(\widehat{\partial^\alpha \varphi}) \right) \rightarrow 0$$

uniformément quand $\varepsilon \rightarrow 0$.

On a alors

$$\frac{1}{\varepsilon}(T * \varphi(x + \varepsilon\alpha) - T * \varphi(x)) \rightarrow \left\langle T, \tau_x(\widehat{\partial^\alpha \varphi}) \right\rangle.$$

Par définition de la dérivation des distributions,

$$\begin{aligned} \langle \partial^\alpha T, \tau_x \hat{\varphi} \rangle &= -\langle T, \partial^\alpha \tau_x \hat{\varphi} \rangle \\ &= -\langle T, \tau_x \partial^\alpha \hat{\varphi} \rangle = \left\langle T, \tau_x \widehat{\partial^\alpha \varphi} \right\rangle \end{aligned}$$

ce qui donne la dernière identité.

On conclut alors par récurrence sur la longueur de α . □

1) Justifions l'égalité

$$\text{Supp } T \cap \text{Supp } (\tau_x \hat{\varphi}) = \text{Supp } T \cap (x - \text{Supp } \varphi).$$

Rappelons que $\hat{\varphi}(y) = \varphi(-y)$, donc $\tau_x \hat{\varphi}(y) = \hat{\varphi}(y - x) = \varphi(x - y)$, ce qui signifie que $\tau_x \hat{\varphi}$ est la fonction $y \mapsto \varphi(x - y)$. Le support de $\tau_x \hat{\varphi}$ est alors donné par

$$\text{Supp}(\tau_x \hat{\varphi}) = \{y \in \mathbb{R}^N \mid \varphi(x - y) \neq 0\} = x - \text{Supp}(\varphi).$$

Ainsi, on a bien

$$\text{Supp } T \cap \text{Supp } (\tau_x \hat{\varphi}) = \text{Supp } T \cap (x - \text{Supp } \varphi).$$

2) Par définition du produit de convolution d'une distribution T avec une fonction test $\varphi \in C_c^\infty(\mathbb{R}^N)$, on a

$$(T * \varphi)(x) = \langle T, \tau_x \hat{\varphi} \rangle.$$

Donc :

$$\begin{aligned} \frac{1}{\varepsilon}(T * \varphi(x + \varepsilon\alpha) - T * \varphi(x)) &= \frac{1}{\varepsilon}(\langle T, \tau_{x+\varepsilon\alpha} \hat{\varphi} \rangle - \langle T, \tau_x \hat{\varphi} \rangle) \\ &= \left\langle T, \frac{1}{\varepsilon}(\tau_{x+\varepsilon\alpha} \hat{\varphi} - \tau_x \hat{\varphi}) \right\rangle. \end{aligned}$$

3) On veut montrer que

$$\frac{1}{\varepsilon}(\tau_{x+\varepsilon\alpha} \hat{\varphi} - \tau_x \hat{\varphi}) \rightarrow \tau_x(\widehat{\partial^\alpha \varphi}) \quad \text{dans } C_c^\infty(\mathbb{R}^N).$$

Soit $f_\varepsilon := \frac{1}{\varepsilon}(\tau_{x+\varepsilon\alpha} \hat{\varphi} - \tau_x \hat{\varphi})$. On considère le développement de Taylor :

$$\hat{\varphi}(y - \varepsilon\alpha) = \hat{\varphi}(y) - \varepsilon \partial^\alpha \hat{\varphi}(y) + \varepsilon R_\varepsilon(y),$$

avec $R_\varepsilon \rightarrow 0$ uniformément (ainsi que toutes ses dérivées) sur les compacts quand $\varepsilon \rightarrow 0$. Alors,

$$f_\varepsilon(y) = \frac{1}{\varepsilon}(\hat{\varphi}(y - \varepsilon\alpha) - \hat{\varphi}(y)) = -\partial^\alpha \hat{\varphi}(y) + R_\varepsilon(y),$$

donc

$$f_\varepsilon(y) = -\partial^\alpha \hat{\varphi}(y) + o(1) \quad \text{dans } C_c^\infty.$$

Or, $\partial^\alpha \hat{\varphi}(y) = (-1)^{|\alpha|} \widehat{\partial^\alpha \varphi}(y)$, donc :

$$f_\varepsilon(y) \rightarrow (-1)^{|\alpha|+1} \widehat{\partial^\alpha \varphi}(y) = \tau_x \widehat{\partial^\alpha \varphi}(y - x).$$

Finalement :

$$f_\varepsilon \rightarrow \tau_x \widehat{\partial^\alpha \varphi} \quad \text{dans } C_c^\infty(\mathbb{R}^N).$$

4) Soit $f_\varepsilon = \frac{1}{\varepsilon}(\tau_{x+\varepsilon\alpha}\hat{\varphi} - \tau_x\hat{\varphi})$. On sait que

$$\text{Supp}(\tau_x\hat{\varphi}) = x - \text{Supp } \varphi,$$

donc

$$\text{Supp}(f_\varepsilon) \subset (x + \varepsilon\alpha - \text{Supp } \varphi) \cup (x - \text{Supp } \varphi).$$

Les deux ensembles sont inclus dans $x - \text{Supp } \varphi + B(0, \varepsilon|\alpha|) \subset x - \text{Supp } \varphi + B(0, 1)$ pour $\varepsilon \in (0, 1]$. Comme $-\text{Supp } \varphi = \text{Supp } \hat{\varphi}$, on a donc :

$$\text{Supp}(f_\varepsilon) \subset \text{Supp}(\hat{\varphi}) + B(x, 1).$$

Théorème 4.3.5. *L'ensemble des fonctions lisses $C^\infty(\mathbb{R}^N)$ est dense dans $\mathcal{D}'(\mathbb{R}^N)$.*

Démonstration: Soit (ρ_n) une suite régularisante définie par

$$\forall x \in \mathbb{R}^N, \quad \rho_n(x) = n^N \rho(nx) \text{ où } \rho \in C^\infty(\mathbb{R}^N, [0, 1]) \text{ à support dans la boule unité.}$$

Pour toute fonction $\varphi \in C_c^\infty(\mathbb{R}^N)$, $\varphi * \hat{\rho}_n \rightarrow \varphi$ dans $C_c^\infty(\mathbb{R}^N)$ puisque

— pour tout $n \geq 1$,

$$\text{Supp}(\varphi * \hat{\rho}_n) \subset \text{Supp}(\varphi) + B(0, 1),$$

— pour tout $\beta \in \mathbb{N}^N$

$$\partial^\beta(\varphi * \hat{\rho}_n) = (\partial^\beta \varphi) * \hat{\rho}_n \rightarrow \partial^\beta \varphi$$

uniformément quand $n \rightarrow \infty$.

On a alors

$$\begin{aligned} \langle T * \rho_n, \varphi \rangle &= \int_{\mathbb{R}^N} T * \rho_n(x) \varphi(x) dx = \int_{\mathbb{R}^N} \langle T, \tau_x \hat{\rho}_n \rangle \varphi(x) dx \\ &= \left\langle T, \int_{\mathbb{R}^N} \varphi(x) \tau_x \hat{\rho}_n dx \right\rangle = \langle T, \varphi * \hat{\rho}_n \rangle \rightarrow \langle T, \varphi \rangle \end{aligned}$$

en utilisant par exemple des sommes de Riemann pour justifier l'interversion du crochet de distribution et de l'intégrale.

$(T * \rho_n)_{n \in \mathbb{N}}$ est donc une suite de $C^\infty(\mathbb{R}^N)$ qui converge vers T au sens des distributions. $\square$

En fait, de façon plus générale, on peut définir la convolution de deux distributions, pourvu que l'une au moins soit à support compact.

Définition 4.3.6. Soient $T \in \mathcal{D}'(\mathbb{R}^N)$ et $S \in \mathcal{E}'(\mathbb{R}^N)$. La convolution de T par S est la distribution définie par

$$\forall \varphi \in C_c^\infty(\mathbb{R}^N), \quad \langle T * S, \varphi \rangle = \langle T, \hat{S} * \varphi \rangle$$

où $\hat{S}$ est l'image inverse de S par le changement de variable $x \mapsto -x$.

Avec cette définition, on a les propriétés usuelles de la convolution :

Propriété 4.3.7. Soient $T \in \mathcal{D}'(\mathbb{R}^N)$ et $S, S_1, S_2 \in \mathcal{E}'(\mathbb{R}^N)$.

$$\begin{aligned} \text{Supp}(T * S) &\subset \text{Supp}(T) + \text{Supp}(S), \quad \text{ordre}(T * S) \leq \text{ordre}(T) + \text{ordre}(S) \\ \forall \alpha \in \mathbb{N}^N, \quad \partial^\alpha(T * S) &= \partial^\alpha T * S = T * \partial^\alpha S \\ S_1 * S_2 &= S_2 * S_1, \quad (T * S_1) * S_2 = T * (S_1 * S_2) \end{aligned}$$

Démonstration: Puisque

$$\text{Supp}(\hat{S} * \varphi) \subset \text{Supp}(\hat{S}) + \text{Supp}(\varphi)$$

$\hat{S} * \varphi$ est une fonction C^∞ à support compact, et $T * S$ est bien définie.

De plus,

$$\begin{aligned} \text{si } \text{Supp}(\varphi) \cap (\text{Supp}(T) + \text{Supp}(S)) &= \emptyset, \quad \text{alors } \text{Supp}(\hat{S} * \varphi) \cap \text{Supp}(T) = \emptyset \\ \text{et } \langle T * S, \varphi \rangle &= \langle T, \hat{S} * \varphi \rangle = 0. \end{aligned}$$

On obtient ainsi la condition sur les supports convolutifs.

Si T est d'ordre fini p

$$\begin{aligned} |\langle T * S, \varphi \rangle| &= |\langle T, \hat{S} * \varphi \rangle| \leq C \sup_{|\alpha| \leq p} \sup_{x \in K + \text{Supp}(\hat{S})} \left| \partial^\alpha (\hat{S} * \varphi)(x) \right| \\ &\leq C \sup_{\substack{|\alpha| \leq p \\ |\beta| \leq p'}} \sup_{x \in K} \left| \partial^{\alpha+\beta} \varphi(x) \right| \end{aligned}$$

ce qui montre que l'ordre de $T * S$ est inférieur à $p + p'$.

Pour la dérivation, en utilisant la dualité, on a les identités

$$\begin{aligned} \langle \partial^\alpha(T * S), \varphi \rangle &= (-1)^{|\alpha|} \langle T * S, \partial^\alpha \varphi \rangle = (-1)^{|\alpha|} \langle T, \hat{S} * \partial^\alpha \varphi \rangle \\ &= (-1)^{|\alpha|} \langle T, \partial^\alpha (\hat{S} * \varphi) \rangle = \langle \partial^\alpha T * S, \varphi \rangle \\ &= (-1)^{|\alpha|} \langle T, \partial^\alpha \hat{S} * \varphi \rangle = \langle T * \partial^\alpha S, \varphi \rangle. \end{aligned}$$

Les propriétés de réflexivité et d'associativité se prouvent par régularisation

$$\langle S_1 * S_2, \varphi \rangle = \langle S_1, \hat{S}_2 * \varphi \rangle = \lim_{n \rightarrow \infty} \langle S_1 * \rho_n, \hat{S}_2 * \varphi \rangle$$

car $\hat{S}_2 * \varphi \in C_c^\infty(\mathbb{R}^N)$ et $S_1 * \rho_n \rightarrow S_1$ dans $\mathcal{D}'(\mathbb{R}^N)$.

Comme précédemment, en utilisant des sommes de Riemann, on intervertit le crochet et l'intégrale

$$\begin{aligned} \langle S_1 * \rho_n, \hat{S}_2 * \varphi \rangle &= \int S_1 * \rho_n(x) \langle \hat{S}_2, \tau_x \hat{\varphi} \rangle dx = \left\langle \hat{S}_2, \int S_1 * \rho_n(x) \tau_x \hat{\varphi} dx \right\rangle \\ &= \left\langle \hat{S}_2, S_1 * (\rho_n * \hat{\varphi}) \right\rangle = \left\langle S_2, \hat{S}_1 * (\hat{\rho}_n * \varphi) \right\rangle \\ &= \langle S_2 * S_1, \hat{\rho}_n * \varphi \rangle \rightarrow \langle S_2 * S_1, \varphi \rangle \end{aligned}$$

ce qui donne la réflexivité.

De la même façon, on a

$$\begin{aligned} T * (S_1 * S_2) &= \lim_{\substack{n \rightarrow \infty \\ m \rightarrow \infty}} T * ((S_1 * \rho_n) * (S_2 * \rho_m)) \\ &= \lim_{\substack{n \rightarrow \infty \\ m \rightarrow \infty}} (T * (S_1 * \rho_n)) * (S_2 * \rho_m) = (T * S_1) * S_2 \end{aligned}$$

car

$$\forall \varphi_1, \varphi_2 \in C_c^\infty(\mathbb{R}^N), \quad T * (\varphi_1 * \varphi_2) = (T * \varphi_1) * \varphi_2.$$

Cela montre l'associativité. □

Les deux résultats qui précèdent reposent essentiellement sur les généralisations suivantes des théorèmes de Fubini et de dérivation sous le signe somme.

- **Intégration sous le crochet.** Soient $S \in \mathcal{E}'(\mathbb{R}^N)$ et $\varphi \in C^\infty(\mathbb{R}^{N+q})$. Pour tout compact $K \subset \mathbb{R}^q$,

$$\left\langle S, \int_K \varphi(\cdot, y) dy \right\rangle = \int_K \langle S, \varphi(\cdot, y) \rangle dy.$$

- **Dérivation sous le crochet.** Soient $S \in \mathcal{E}'(\mathbb{R}^N)$ et $\varphi \in C^\infty(\mathbb{R}^{N+q})$. La fonction $\phi : y \mapsto \langle S, \varphi(\cdot, y) \rangle$ est de classe C^∞ et

$$\partial^\alpha \langle S, \varphi(\cdot, y) \rangle = \langle S, \partial_y^\alpha \varphi(\cdot, y) \rangle$$

Application des distributions aux solutions singulières des EDP

On va présenter maintenant une application très simple de la notion de distribution à la résolution d'équations aux dérivées partielles présentant des singularités en temps fini. La résolution se fera par approximation, puis par passage à la limite : cela permettra donc de manipuler les différentes notions de convergence introduites aux chapitres précédents.

L'équation aux dérivées partielles que nous allons considérer ici n'a pas d'application physique directe, mais c'est un prototype de système hyperbolique de lois de conservation, comme il en apparaît naturellement par exemple en mécanique des fluides ou en élasticité [17], [4].

L'équation de Hopf est une équation scalaire unidimensionnelle, ce qui signifie que l'inconnue u est réelle et dépend du temps $t \in \mathbb{R}^+$ et d'une seule variable d'espace $x \in \mathbb{R}$. Elle s'écrit

$$\partial_t u + \partial_x \left(\frac{1}{2} u^2 \right) = 0 \tag{5.0.1}$$

ce qui est équivalent pour des solutions classiques régulières à

$$\partial_t u + u \partial_x u = 0. \tag{5.0.2}$$

Cette dernière équation est aussi appelée équation de Burgers non visqueuse.

5.1 Méthode des caractéristiques et explosion

En utilisant la forme (5.0.2) de l'équation, on peut calculer explicitement la solution en fonction de la donnée initiale

$$u|_{t=0} = u_0,$$

au moins pour des temps petits. En effet, on dispose d'une méthode très générale pour la résolution des équations de transport, dite méthode des caractéristiques.

5.1.1 La méthode des caractéristiques

Proposition 5.1.1. *Les solutions de l'équation de transport*

$$\partial_t v + a(t, x) \partial_x v = 0, \quad v|_{t=0} = v_0 \quad (5.1.1)$$

peuvent s'écrire simplement à partir des solutions de l'équation différentielle ordinaire

$$\frac{dX}{dt} = a(t, X), \quad X(0, x_0) = x_0. \quad (5.1.2)$$

On a en effet

$$v(t, X(t, x_0)) = v_0(x_0)$$

Si $X_t : x \mapsto X(t, x)$ est bijective, alors

$$v(t, x) = v_0(X_t^{-1}(x))$$

Dans le cas d'un champ de convection constant a , le mouvement est uniforme

$$v(t, x) = v_0(x - at)$$

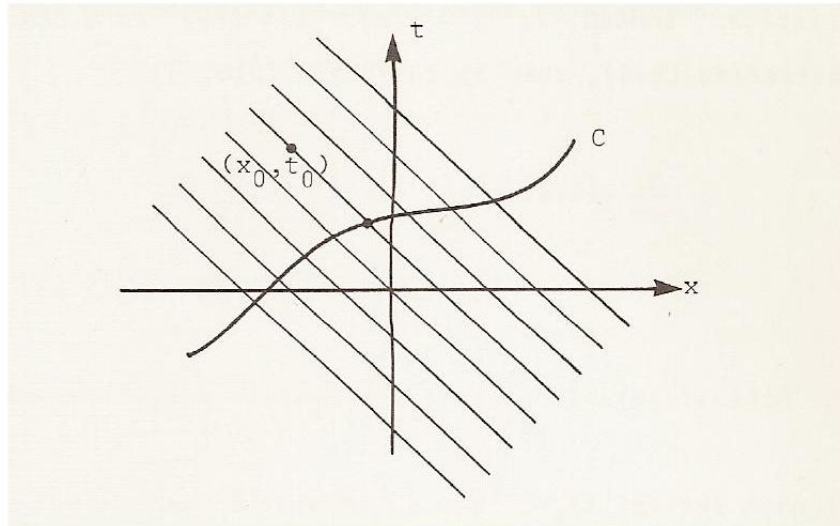


Figure 1. Les caractéristiques sont des lignes droites.

Sous des hypothèses de régularité sur le champ a , le théorème de Cauchy-Lipschitz assure que les trajectoires de (5.1.2) sont localement bien définies et uniques, de sorte que X_t est un difféomorphisme.

Théorème 5.1.2. *(de Cauchy-Lipschitz). - Soit E un espace de Banach réel, U un ouvert de $\mathbb{R} \times E$, et $a : U \rightarrow E$ une application lipschitzienne. Alors*

- *Pour tout $(t_0, x_0) \in U$, il existe une unique solution maximale $X_{(t_0, x_0)} : I_{(t_0, x_0)} \rightarrow E$ de l'équation différentielle*

$$\frac{dX}{dt} = a(t, X(t))$$

- L'application $x \mapsto X_{(t_0, x)}(s)$ est continue au voisinage de x_0 pour tout $s \in I_{(t_0, x_0)}$.
- Si a est de classe C^k , toute solution de l'équation différentielle est de classe C^{k+1} .

Dans le cas de l'équation de Hopf, le champ de vitesses u est transporté par lui-même. On a donc

$$\begin{aligned} \frac{dX}{dt} &= u(t, X), & X(0, x_0) &= x_0, \\ u(t, X(t, x)) &= u_0(x). \end{aligned}$$

tant que X est un difféomorphisme, i.e. un changement de variables régulier. On rappelle que le fait que X soit un difféomorphisme dépend de la régularité de u (par le théorème de Cauchy-Lipschitz).

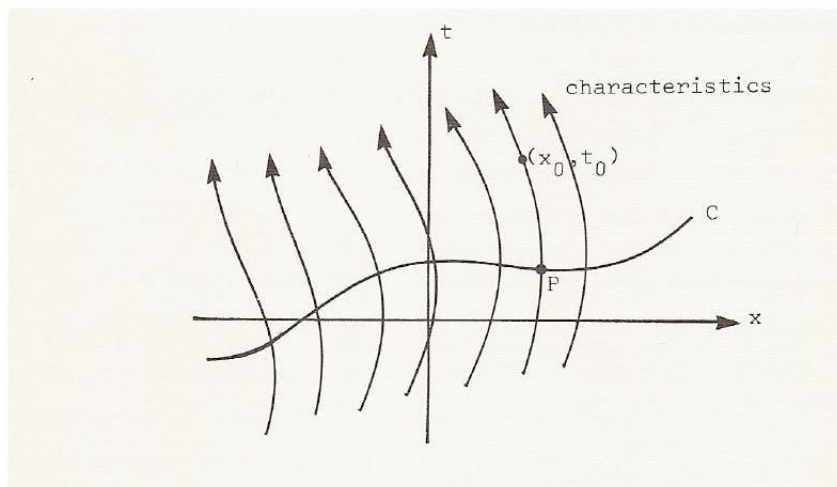


Figure 2. Les caractéristiques définissent un difféomorphisme sur $\mathbb{R}$

5.1.2. Apparition de singularité

En différentiant l'équation des caractéristiques par rapport à x , on obtient

$$\frac{\partial}{\partial t} \frac{\partial X}{\partial x} = \frac{\partial}{\partial x} (u(t, X)) = u'_0(x)$$

En intégrant par rapport au temps t , on en déduit que

$$\frac{\partial X}{\partial x} = 1 + u'_0(x)t.$$

Au temps

$$t_0 = \frac{1}{\sup_{x \in \mathbb{R}} (-u'_0(x))_+}$$

la fonction $\frac{\partial X}{\partial x}$ s'annule au moins en un point, de sorte que X n'est plus un difféomorphisme.

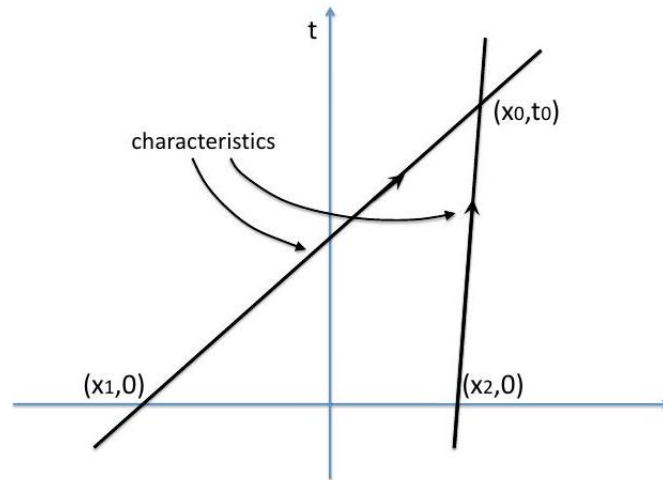


Figure 3. Croisement des caractéristiques

En effet, au point (t_0, x_0) la solution u devient multivaluée et donc $u(t_0, x_0)$ n'est pas défini. De plus, il y a un saut

$$\lim_{x \rightarrow x_0^-} u(t_0, x) \neq \lim_{x \rightarrow x_0^+} u(t_0, x).$$

En d'autres termes, t_0 est le temps d'explosion, il correspond à l'apparition d'une singularité.

Il est alors naturel de se demander si les solutions peuvent être définies en un sens plus faible pour des temps au-delà de t_0 .

5.2 Solutions faibles et critère d'unicité

5.2.1 Solutions au sens des distributions

Définition 5.2.1. On appelle solution au sens des distributions de l'équation de Hopf (5.0.1) avec donnée initiale u_0 une fonction $u \in L^\infty(\mathbb{R}^+ \times \mathbb{R})$ (définie et bornée presque partout) telle que, pour tout $\phi \in C_c^\infty(\mathbb{R}^+ \times \mathbb{R})$,

$$\iint \left(u \partial_t \phi + \frac{1}{2} u^2 \partial_x \phi \right) dx dt = - \int u_0 \phi|_{t=0} dx.$$

Attention : la distribution u^2 est bien définie car on a imposé que u soit une fonction (définie et bornée presque partout) : u^2 est donc une fonction définie et bornée presque partout.

Par contre, au sens des distributions, les deux formulations (5.0.1) et (5.0.2) de l'équation de Hopf ne sont pas équivalentes : on utilisera toujours la forme conservative (5.0.1), qui est la seule à être bien définie !

Avec cette notion de solution, on n'a plus d'unicité ! Si on part par exemple de la fonction d'Heaviside $u_0 = \mathbf{1}_{\mathbb{R}^+}$, on peut vérifier que les deux fonctions u_1 et u_2 définies par

$$u_1(t, x) = H\left(x - \frac{1}{2}t\right)$$

$$u_2(t, x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{x}{t} & \text{if } 0 < x < t \\ 1 & \text{if } x > t \end{cases}$$

sont des solutions au sens des distributions de l'équation de Hopf (5.0.1), i.e.

$$\iint \left(u \partial_t \phi + \frac{1}{2} u^2 \partial_x \phi \right) dx dt = - \int u_0 \phi|_{t=0} dx.$$

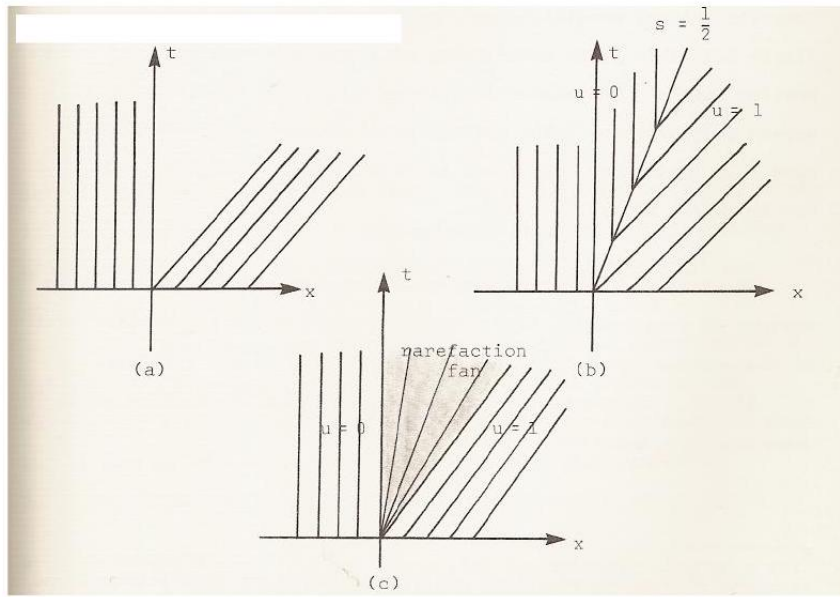


Figure 4. Non unicité des solutions au sens des distributions

Remarque 5.2.2. *La méthode des caractéristiques ne prédit pas le comportement de la solution dans la zone de non unicité.*

5.2.2 Solutions entropiques

Les solutions physiquement pertinentes sont sélectionnées en imposant des conditions supplémentaires, qui devraient en particulier garantir l'unicité.

On considère le cas d'un choc, c'est-à-dire d'une discontinuité se propageant à la vitesse s

$$u(t, x) = u_- + (u_+ - u_-) \mathbf{1}_{\{x \geq st\}}$$

et satisfaisant la condition de saut (dite condition de Rankine-Hugoniot)

$$-s[u]_-^+ + \frac{1}{2} [u^2]_-^+ = 0.$$

Par chaque point de la ligne de choc, il passe deux caractéristiques (définies par 5.1.2), une de chaque côté du choc.

— Soit chacune de ces caractéristiques peut être suivie en arrière jusqu'au temps initial,

— soit les deux caractéristiques sont définies après le choc.

Les chocs de la deuxième espèce, appelés chocs de raréfaction, ne sont pas admissibles physiquement car ils ne sont pas déterminés par la donnée initiale (principe de causalité). Seuls les chocs de la première espèce, appelés chocs entropiques, doivent être retenus.

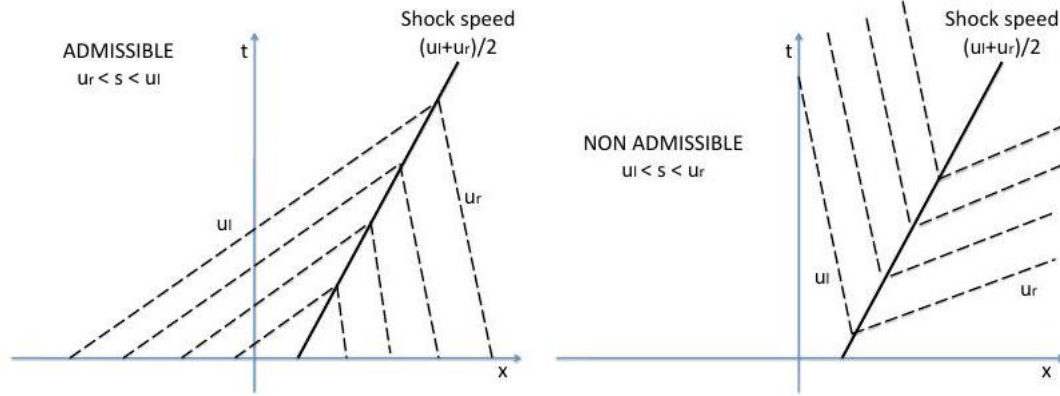


Figure 5. Condition d'entropie

Pour les chocs entropiques, on a

$$\partial_x u \leq 0 \text{ au sens des distributions}$$

Pour les solutions régulières (par exemple pour la solution u_2 définie au paragraphe précédent), on a

$$\partial_x u(t, x) = \partial_x u_0(X_t^{-1}(x)) = \left(\frac{u'_0}{1 + tu'_0} \right) (X_t^{-1}(x)) \leq \frac{1}{t}.$$

On va donc ajouter une condition de signe sur la partie singulière de $\partial_x u$.

Avec cette contrainte supplémentaire, on va retrouver l'unicité des solutions. Autrement dit, on s'attend à ce que le système

$$\left. \begin{aligned} \partial_t u + \frac{1}{2} \partial_x u^2 &= 0 \\ t \partial_x u &\leq 1 \end{aligned} \right\} \text{ au sens des distributions} \quad (5.2.1)$$

capture la dynamique de façon satisfaisante, au-delà de l'apparition des singularités.

5.3 Approximation visqueuse

Pour montrer l'existence de solutions au sens des distributions pour l'équation de Hopf, l'idée est de construire des solutions pour un modèle approché, puis de passer à la limite. Il existe trois procédures très classiques : l'approximation par discrétisation et résolution de problèmes de Riemann élémentaires (méthode de Glimm), l'approximation par des modèles cinétiques (méthode de relaxation), et l'approximation par des équations paraboliques (méthode de viscosité). On a choisi ici de présenter cette dernière approche. L'équation de Burgers

$$\partial_t u + u \partial_x u = \frac{\varepsilon}{2} \partial_{xx}^2 u$$

est une approximation visqueuse de l'équation de Hopf (par analogie avec les modèles visqueux de la dynamique des gaz). On s'attend typiquement à ce que

- l'effet régularisant de l'équation de la chaleur (qui sera étudié de façon plus systématique dans la partie suivante du cours) permette de contrôler la nonlinéarité,
- les chocs soient lissés selon des profils réguliers à l'échelle $\sqrt{\varepsilon}$.

On va montrer en effet que, pour tout $\varepsilon > 0$ fixé, il existe une unique solution globale (classique) à l'équation de Burgers, puis que cette famille converge quand $\varepsilon \rightarrow 0$ vers l'unique solution entropique de l'équation de Hopf.

5.3.1 La transformation de Cole-Hopf

La résolution de l'équation de Burgers est en fait explicite, elle repose sur un changement de variable très astucieux, appelé transformation de Cole-Hopf :

Lemme 5.3.1. (*Transformation de Cole-Hopf*). - Soit ϕ une solution (classique) positive de l'équation de la chaleur

$$\partial_t \phi - \frac{\varepsilon}{2} \partial_{xx}^2 \phi = 0.$$

On définit $u = \partial_x U$ où

$$\lambda U(t, x) = -\log \phi(t, x).$$

Alors u est une solution de l'équation de Burgers

$$\partial_t u + \frac{1}{2}(\varepsilon \lambda) \partial_x (u^2) - \frac{\varepsilon}{2} \partial_{xx}^2 u = 0.$$

Démonstration: La preuve repose sur des calculs élémentaires. A partir de la formule $\phi(t, x) = \exp(-\lambda U(t, x))$, on obtient

$$\partial_t \phi - \frac{\varepsilon}{2} \partial_{xx}^2 \phi = -\lambda \exp(-\lambda U) \left(\partial_t U + \frac{1}{2}(\varepsilon \lambda) (\partial_x U)^2 - \frac{\varepsilon}{2} \partial_{xx}^2 U \right) = 0$$

ce qui montre que

$$\partial_t U + \frac{1}{2}(\varepsilon \lambda) (\partial_x U)^2 - \frac{\varepsilon}{2} \partial_{xx}^2 U = 0.$$

En différentiant par rapport à x , on obtient finalement

$$\partial_t u + \frac{1}{2}(\varepsilon \lambda) \partial_x (u^2) - \frac{\varepsilon}{2} \partial_{xx}^2 u = 0$$

ce qui signifie que u est solution de l'équation de Burgers pourvu que $\varepsilon \lambda = 1$. □

La stratégie qu'on va utiliser pour résoudre l'équation de Burgers consiste alors

- à calculer la donnée initiale ϕ_0 pour l'équation de la chaleur associée

$$\phi_0 = \exp\left(-\frac{1}{\varepsilon} U_0\right) \text{ avec } U_0(x) = \int_{-\infty}^x u_0(y) dy$$

qui est bien définie dès que u_0 est régulière et rapidement décroissante à l'infini par exemple ;

— à résoudre l'équation de la chaleur avec donnée initiale ϕ_0 .

On vérifiera en particulier que ϕ reste positive :

$$\forall t \geq 0, \forall x \in \mathbb{R}, \quad 0 \leq \phi(t, x) \leq \|\phi_0\|_\infty$$

— à retrouver la solution de l'équation de Burgers par changement de variable inverse

$$u = -\varepsilon \frac{\partial_x \phi}{\phi}.$$

Comme l'équation de la chaleur est linéaire à coefficients constants, on en obtient toutes les solutions par convolution à partir de la fonction de Green, solution fondamentale de l'équation ayant la masse de Dirac comme donnée initiale. Nous verrons dans la suite de ce cours des méthodes systématiques pour étudier ce type d'équations, dite paraboliques, mais ici la situation est suffisamment simple pour que toutes les propriétés de ϕ puissent être obtenues à partir des formules explicites :

Lemme 5.3.2. (*Fonction de Green*). La solution ϕ de l'équation de la chaleur

$$\partial_t \phi - \frac{\varepsilon}{2} \partial_{xx}^2 \phi = 0.$$

avec donnée initiale $\phi_0 = \exp\left(-\frac{1}{\varepsilon} U_0\right)$ s'écrit

$$\phi(t, x) = \int \frac{1}{\sqrt{\varepsilon t}} \exp\left(-\frac{1}{\varepsilon} U_0(x - y) - \left(\frac{y}{\sqrt{2\varepsilon t}}\right)^2\right) dy.$$

En particulier, elle est C^∞ , positive et vérifie le principe du maximum

$$\forall t \geq 0, \quad \|\phi(t)\|_\infty \leq \|\phi_0\|_\infty.$$

Démonstration: Soit G la solution au sens des distributions de l'équation de la chaleur ayant la masse de Dirac comme donnée initiale

$$\partial_t G - \frac{\varepsilon}{2} \partial_{xx}^2 G = 0, \quad G|_{t=0} = \delta_0$$

On vérifie facilement, par définition de la convolution sur les distributions, que $\phi = G * \phi_0$ est bien une fonction régulière et qu'elle satisfait l'équation de la chaleur au sens des distributions. Comme la dérivation-distribution coïncide avec la dérivation classique pour les fonctions régulières, on obtient ainsi une solution classique de l'équation de la chaleur.

— A cause de l'invariance par changement d'échelle de l'équation de la chaleur et de la forme particulière de la donnée initiale, il est naturel de chercher G sous la forme d'une solution auto-similaire du type

$$G(t, x) = \frac{1}{\sqrt{\varepsilon t}} g\left(\frac{x}{\sqrt{\varepsilon t}}\right).$$

En effet, cela assure que $G(t)$ converge vers la masse de Dirac (au sens des distributions) quand $t \rightarrow 0$.

En insérant cet Ansatz dans l'équation de la chaleur, on obtient l'équation différentielle ordinaire :

$$g(\xi) + \xi g'(\xi) + g''(\xi) = 0.$$

Une solution de cette équation différentielle est la distribution gaussienne

$$g(\xi) = \exp\left(-\frac{1}{2}\xi^2\right).$$

On a alors la formule explicite suivante

$$\begin{aligned}\phi(t, x) &= \int \frac{1}{\sqrt{\varepsilon t}} \exp\left(-\frac{1}{\varepsilon}U_0(y) - \left(\frac{x-y}{\sqrt{2\varepsilon t}}\right)^2\right) dy \\ &= \int \frac{1}{\sqrt{\varepsilon t}} \exp\left(-\frac{1}{\varepsilon}U_0(x-y) - \left(\frac{y}{\sqrt{2\varepsilon t}}\right)^2\right) dy.\end{aligned}\tag{5.3.1}$$

— De la première formule, on déduit que $\phi(t) \in C^\infty(\mathbb{R})$ pour tout $t > 0$ (quelle que soit la régularité de la donnée initiale) : on dit que l'équation de la chaleur a un effet régularisant (caractéristique des équations paraboliques).

L'intégrande admet en effet des dérivées par rapport à x à tout ordre k : pour tout $\varepsilon > 0$ et $t > 0$ fixés, et tout $x \in [-R, R]$

$$\left| \partial_x^k \left(\frac{1}{\sqrt{\varepsilon t}} \exp\left(-\frac{1}{\varepsilon}U_0(y) - \left(\frac{x-y}{\sqrt{2\varepsilon t}}\right)^2\right) \right) \right| \leq C(\varepsilon, t, R) \exp\left(-\left(\frac{y}{\sqrt{4\varepsilon t}}\right)^2\right).$$

On peut alors échanger l'ordre de l'intégration et de la dérivation (théorème de convergence dominée de Lebesgue). $\square$

5.3.2 Le problème de Cauchy pour l'équation de Burgers

Les deux lemmes du paragraphe précédent permettent de mettre en oeuvre la stratégie proposée, et de montrer l'existence et l'unicité d'une solution pour l'équation de Burgers :

Proposition 5.3.3. (*Existence et unicité*). - Soit u_0 une fonction bornée sur $\mathbb{R}^+$. L'équation de Burgers

$$\partial_t u + \frac{1}{2}\partial_x u^2 - \frac{\varepsilon}{2}\partial_{xx}^2 u = 0 \quad u|_{t=0} = u_0,$$

admet une unique solution u_ε bornée sur $\mathbb{R}^+ \times \mathbb{R}$, telle que $u_\varepsilon(t)$ est de classe C^∞ pour tout $t > 0$.

Démonstration: - A partir de la seconde formule définissant ϕ dans (5.3.1) et en utilisant le théorème de convergence dominée, on obtient

$$-\varepsilon \partial_x \phi(t, x) = \int \frac{1}{\sqrt{\varepsilon t}} u_0(x-y) \exp\left(-\frac{1}{\varepsilon}U_0(x-y) - \left(\frac{y}{\sqrt{2\varepsilon t}}\right)^2\right) dy$$

— Par inversion de la transformation de Cole-Hopf, on a alors

$$u(t, x) = -\varepsilon \frac{\partial_x \phi}{\phi}(t, x) = \frac{\int u_0(y) \exp\left(-\frac{1}{\varepsilon} U_0(y) - \left(\frac{x-y}{\sqrt{2\varepsilon t}}\right)^2\right) dy}{\int \exp\left(-\frac{1}{\varepsilon} U_0(y) - \left(\frac{x-y}{\sqrt{2\varepsilon t}}\right)^2\right) dy}$$

de sorte que pour tout $t > 0$, on a le principe du maximum

$$|u(x, t)| \leq \|u_0\|_{\infty}.$$

En combinant le théorème de dérivation sous l'intégrale, et les règles de calcul usuelles pour les dérivées de quotients, on a alors que $u(t)$ est de classe $C^\infty(\mathbb{R}^+)$ pour tout temps $t > 0$.

- Reste à établir l'unicité dans la classe des fonctions bornées. La transformation de Cole-Hopf est clairement injective, et elle envoie les fonctions bornées sur les fonctions à croissance au plus exponentielle. Il suffit donc d'établir l'unicité pour l'équation de la chaleur dans cette classe de fonctions. On présente ici une preuve élémentaire de ce résultat, mais on verra des méthodes plus systématiques dans la suite du cours.

Comme l'équation de la chaleur est une équation linéaire, il suffit de prouver que la solution nulle est l'unique solution qui a 0 pour donnée initiale. On considère donc une solution à croissance au plus exponentielle de

$$\partial_t \psi - \frac{\varepsilon}{2} \partial_{xx}^2 \psi = 0, \quad \psi|_{t=0} = 0.$$

Comme la convolution commute avec les opérateurs différentiels, quitte à régulariser $\psi(t)$, on peut supposer que $\psi(t) \in C_b^\infty(\mathbb{R})$.

En multipliant l'équation par $\exp(-\mu|x|)\psi$ et en intégrant par parties, on obtient

$$\frac{1}{2} \partial_t \int \exp(-\mu|x|) \psi^2(t, x) dx + \frac{\varepsilon}{2} \int \exp(-\mu|x|) (\partial_x \psi)^2(t, x) dx + \frac{\varepsilon \mu}{2} \int \text{sign}(x) \exp(-\mu|x|) \psi \partial_x \psi(t, x) dx = 0.$$

En utilisant l'inégalité de Cauchy-Schwarz, on en déduit que

$$\frac{1}{2} \partial_t \int \exp(-\mu|x|) \psi^2(t, x) dx + \frac{\varepsilon}{4} \int \exp(-\mu|x|) (\partial_x \psi)^2(t, x) dx \leq \frac{\varepsilon \mu^2}{4} \int \exp(-\mu|x|) \psi^2(t, x) dx.$$

L'inégalité de Gronwall montre alors que

$$\int \exp(-\mu|x|) \psi^2(t, x) dx \leq \left(\int \exp(-\mu|x|) \psi^2(0, x) dx \right) \exp\left(\frac{\varepsilon \mu^2}{2} t\right) \equiv 0.$$

Ceci prouve que ψ est identiquement nulle. □

5.3.3 Convergence vers l'équation de Hopf

Pour tout ε fixé, on note u_ε l'unique solution de l'équation de Burgers avec viscosité ε et donnée initiale u_0 . On s'intéresse maintenant à la limite quand $\varepsilon \rightarrow 0$ dans laquelle on s'attend à obtenir une solution de l'équation de Hopf. Plus précisément on va montrer

Proposition 5.3.4. *Soit u_ε l'unique solution de l'équation de Burgers avec viscosité ε et donnée initiale continue u_0 . Alors, quand $\varepsilon \rightarrow 0$, (u_ε) converge (en un sens faible) vers une fonction u définie globalement en temps, et bornée presque partout. De plus,*

- les points de discontinuité de u sont dénombrables ;
- u est une solution globale de l'équation de Hopf (au sens des distributions) ;
- u vérifie la condition d'entropie

$$\partial_x u(t, \cdot) \leq \frac{1}{t}.$$

La preuve repose sur la formule explicite obtenue pour u_ε :

$$u_\varepsilon(t, x) = \int u_0(y) d\mu_\varepsilon(t, x, y)$$

où la mesure de probabilité μ_ε est définie par

$$d\mu_\varepsilon(t, x, y) = \frac{\exp\left(-\frac{1}{\varepsilon}U_0(y) - \left(\frac{x-y}{\sqrt{2\varepsilon t}}\right)^2\right) dy}{\int \exp\left(-\frac{1}{\varepsilon}U_0(y) - \left(\frac{x-y}{\sqrt{2\varepsilon t}}\right)^2\right) dy}.$$

On va montrer que, dans la limite $\varepsilon \rightarrow 0$, la mesure μ_ε va se concentrer sur les points minimaux de la phase

$$\Psi(t, x, y) = U_0(y) + \frac{1}{2t}(x - y)^2$$

Comme $\Psi(t, x, \cdot)$ est régulière (lipschitzienne) et tend vers l'infini en $\pm\infty$, l'ensemble des points où elle atteint son minimum est un compact

$$I(t, x) \subset \left\{ y \in \mathbb{R} / u_0(y) - \frac{1}{t}(x - y) = 0 \right\}.$$

On définit alors

$$y_-(t, x) = \min I(t, x), \quad y_+(t, x) = \max I(t, x).$$

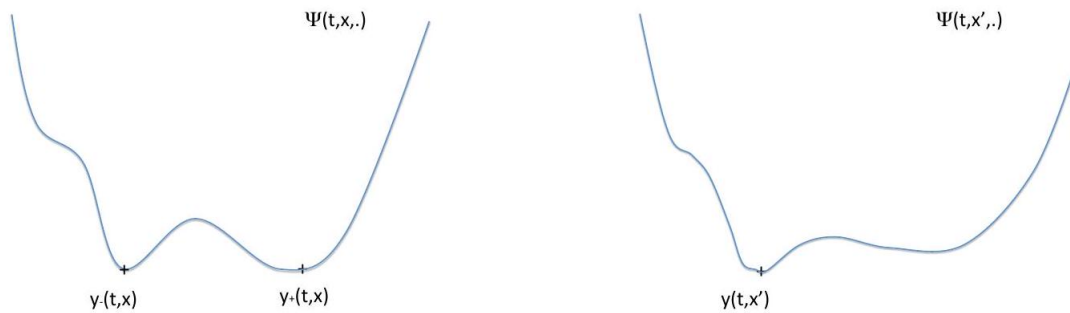


Figure 6. Représentation graphique de $I(t, x)$

La première étape consiste donc à étudier les fonctions $y_\pm$.

Lemme 5.3.5. *Pour tout $t > 0$, les fonctions $y_+(t)$ et $y_-(t)$ sont croissantes, et coïncident en dehors d'un ensemble dénombrable de points.*

Démonstration: Soient $x_1 < x_2$ et $y < y_+(t, x_1)$. Par définition de $y_+(t, x_1)$, on a

$$\begin{aligned} & \Psi(t, x_2, y) - \Psi(t, x_2, y_+(t, x_1)) \\ &= [\Psi(t, x_2, y) - \Psi(t, x_1, y)] + [(\Psi(t, x_1, y) - \Psi(t, x_1, y_+(t, x_1))) + [\Psi(t, x_1, y_+(t, x_1)) - \Psi(t, x_2, y_+(t, x_1))]] \\ &= \frac{1}{2t} (x_2 - x_1) (x_1 + x_2 - 2y) + \left[(\Psi(t, x_1, y) - \Psi(t, x_1, y_+(t, x_1))) \right] - \frac{1}{2t} (x_2 - x_1) (x_1 + x_2 - 2y_+(t, x_1)) \\ &\geq \frac{1}{t} (x_2 - x_1) (y_+(t, x_1) - y) > 0. \end{aligned}$$

On en déduit que

$$y_+(t, x_2) \geq y_-(t, x_2) \geq y_+(t, x_1) \geq y_-(t, x_1).$$

Les fonctions $x \mapsto y_{\pm}(t, x)$ sont donc croissantes. En particulier, elles ont un ensemble dénombrable de points de discontinuité, qu'on appelle S_t .

Comme la fonction $x \mapsto \min_{y \in \mathbb{R}} \Psi(t, x, y)$ est continue, on a

$$\lim_{x \rightarrow x_0} y_-(t, x) \in I(t, x_0), \quad \lim_{x \rightarrow x_0} y_+(t, x) \in I(t, x_0).$$

En utilisant la monotonie, on a de plus que

$$\begin{aligned} \lim_{x \rightarrow x_0} y_-(t, x) &\leq y_-(t, x_0) = \min I(t, x_0), \\ \lim_{x \rightarrow x_0+} y_+(t, x) &\geq y_+(t, x_0) = \max I(t, x_0) \end{aligned}$$

de sorte que

$$\begin{aligned} \lim_{x \rightarrow x_0-} y_-(t, x) &= y_-(t, x_0), \\ \lim_{x \rightarrow x_0+} y_+(t, x) &= y_+(t, x_0). \end{aligned}$$

Autrement dit,

- $y_-(t)$ est continue à gauche ;
 - $-y_+(t)$ est continue à droite.
- On en déduit que

$$\begin{aligned} y_+(t, x) &\leq y_-(t, x+0) \leq y_+(t, x+0) = y_+(t, x) \\ y_-(t, x) &= y_-(t, x-0) \leq y_+(t, x-0) \leq y_-(t, x). \end{aligned}$$

En dehors de l'ensemble dénombrable S_t , on a donc $y_+(t, x) = y_-(t, x)$ et les deux fonctions sont continues. $\square$

La deuxième étape consiste alors à appliquer un résultat de "phase stationnaire" pour caractériser la limite u de la suite (u_ε) .

Lemme 5.3.6. *Pour tout $t > 0$, la suite (u_ε) converge ponctuellement sur $\mathbb{R} \setminus S_t$ vers la fonction u définie par*

$$u(t, x) = u_0(y_{\pm}(t, x)) = \frac{1}{t} (x - y_{\pm}(t, x)).$$

Démonstration: - Comme la fonction de phase Ψ se comporte comme $y^2/2t$ quand $y \rightarrow \pm\infty$, on peut se restreindre à l'étude de l'intégrale sur un domaine borné : pour R assez grand,

$\int \exp\left(-\frac{1}{\varepsilon}(\Psi(t, x, y) - \Psi_{\min}(t, x))\right) dy = \int_{-R}^R \exp\left(-\frac{1}{\varepsilon}(\Psi(t, x, y) - \Psi_{\min}(t, x))\right) dy + o(\exp(-R^2/4\varepsilon t))$.
On décompose alors l'intégrale en deux termes, en fonction de la taille de $\Psi(t, x, y) - \Psi_{\min}(t, x)$. D'une part, on a

$$\int_{-R}^R \mathbf{1}_{\Psi(t, x, y) - \Psi_{\min}(t, x) > \delta} \exp\left(-\frac{1}{\varepsilon}(\Psi(t, x, y) - \Psi_{\min}(t, x))\right) dy \leq 2R \exp\left(-\frac{\delta}{\varepsilon}\right).$$

D'autre part, comme $\Psi(t, x, \cdot)$ est continue, il existe ω tel que

$$\forall y \text{ tel que } |y - y(t, x)| \leq \omega(t, x), \quad \Psi(t, x, y) - \Psi_{\min}(t, x) < \delta/2$$

de sorte que

$$\int_{-R}^R \mathbf{1}_{\Psi(t, x, y) - \Psi_{\min}(t, x) < \delta} \exp\left(-\frac{1}{\varepsilon}(\Psi(t, x, y) - \Psi_{\min}(t, x))\right) dy \geq 2\omega \exp\left(-\frac{\delta}{2\varepsilon}\right).$$

En particulier, quand $\varepsilon \rightarrow 0$,

$\int \exp\left(-\frac{1}{\varepsilon}(\Psi(t, x, y) - \Psi_{\min}(t, x))\right) dy \sim \int \mathbf{1}_{\Psi(t, x, y) - \Psi_{\min}(t, x) < \delta} \exp\left(-\frac{1}{\varepsilon}(\Psi(t, x, y) - \Psi_{\min}(t, x))\right) dy$.
Cela signifie que le support de la mesure limite μ est inclus dans l'ensemble

$$\{y / \Psi(t, x, y) - \Psi_{\min}(t, x) < \delta\}$$

et ceci pour tout $\delta > 0$.

La mesure limite μ est donc une mesure de probabilité, supportée par $\{y_{\pm}(t, x)\}$. On en déduit la formule de Lax-Oleinik

$$u(t, x) = u_0(y_{\pm}(t, x)) = \frac{1}{t}(x - y_{\pm}(t, x)). \quad \square$$

Démonstration de la Proposition 5.3.4.

— On commence par prouver que u est une solution de l'équation de Hopf au sens des distributions.

Le lemme précédent montre que u_{ε} converge presque partout vers u . D'après le principe du maximum, $|u_{\varepsilon}|$ est uniformément bornée (par $\max |u_0|$). Le théorème de convergence dominée permet donc de conclure que

$$u_{\varepsilon} \rightarrow u \text{ dans } L^2, \text{ et donc au sens des distributions.}$$

On peut alors passer à la limite dans la formulation faible de l'équation de Burgers

$$\int_0^{+\infty} \int \left(u_{\varepsilon} \partial_t \varphi + \frac{1}{2} u_{\varepsilon}^2 \partial_x \varphi + \frac{\varepsilon}{2} u_{\varepsilon} \partial_{xx} \varphi \right) dt dx = - \int u_0 \varphi|_{t=0} dx$$

ce qui donne

$$\int_0^{+\infty} \int \left(u \partial_t \varphi + \frac{1}{2} u^2 \partial_x \varphi \right) dt dx = - \int u_0 \varphi|_{t=0} dx$$

qui est la formulation faible de l'équation de Hopf.

- On vérifie ensuite que u coïncide avec l'unique solution classique de l'équation de Hopf tant que cette dernière est définie.

L'explosion de la solution classique a lieu au temps de première intersection des caractéristiques

$$t_0 = \frac{1}{\sup_{x \in \mathbb{R}} (-u'_0(x))_+}$$

ce qui signifie que pour tout $t \in [0, t_0[$

$y \mapsto y + tu_0(y)$ est une fonction continue bijective (croissante) sur $\mathbb{R}$.

Autrement dit,

$\frac{\partial}{\partial y} \Psi(t, x, y) = \frac{1}{t} (y - x + tu_0(y))$ est une fonction continue bijective (croissante) sur $\mathbb{R}$.

En particulier, pour tout $t \in [0, t_0[$ et tout $x \in \mathbb{R}$, il existe un unique $y = y(t, x)$ tel que

$$\begin{aligned} \frac{\partial}{\partial y} \Psi(t, x, y) &= u_0(y) + \frac{1}{t}(y - x) = 0 \\ \Psi(t, x, y(t, x)) &= \min_{y \in \mathbb{R}} \Psi(t, x, y) \end{aligned}$$

L'ensemble S_t est donc vide, et la formule de Lax-Oleinik donne

$$u(t, x) = \frac{1}{t}(x - y(t, x)) = u_0(y(t, x))$$

ce qui est exactement la formule des caractéristiques.

- Pour finir, on va montrer que la solution u de l'équation de Hopf (obtenue à partir des solutions de l'équation de Burgers dans la limite de viscosité évanescence) satisfait de plus la condition d'entropie

$$\partial_x u(t, \cdot) \leq \frac{1}{t}$$

ce qui signifie que u décroît sur les discontinuités admissibles.

On sait que l'application $x \mapsto y_+(t, x)$ est continue en dehors de l'ensemble dénombrable S_t et localement bornée. C'est donc une distribution, et elle admet une dérivée au sens des distributions. Comme $x \mapsto y_+(t, x)$ est croissante, on a de plus

$$\partial_x y_+ \geq 0 \text{ au sens des distributions.}$$

En utilisant la formule de Lax-Oleinik, on a alors

$$t\partial_x u = 1 - \partial_x y_+ \leq 1 \text{ au sens des distributions,}$$

ce qui conclut la preuve. □

Transformée de Fourier

L'analyse de Fourier consiste à décomposer une fonction et plus généralement une distribution en une "superposition" de fonctions oscillantes simples (fonctions exponentielles complexes). Cette analyse en fréquences est un outil très puissant pour l'étude d'un certain nombre de questions, notamment les phénomènes oscillants (par définition), les phénomènes régis par des équations aux dérivées partielles linéaires à coefficients constants (telles que l'équation de la chaleur), les problèmes de régularité [10, 9, 15, 14, 16]. L'idée fondamentale est que l'opération de convolution est transformée en produit, ou encore que la régularité est transformée en décroissance à l'infini.

6.1 Transformation de Fourier des fonctions sommables

Définition 6.1.1. Soit $f \in L^1_{loc}(\mathbb{R}^N)$. La transformée de Fourier de f est la fonction $\mathcal{F}f$ définie sur $\mathbb{R}^N$ par

$$\mathcal{F}f(\xi) = \int_{\mathbb{R}^N} \exp(-ix \cdot \xi) f(x) dx.$$

Propriété 6.1.2. Soient $f, g \in L^1(\mathbb{R}^N)$. Alors

- (i) $\mathcal{F}f$ est continue sur $\mathbb{R}^N$ et $\|\mathcal{F}f\|_\infty \leq \|f\|_{L^1}$;
- (ii) on a l'identité $\mathcal{F}(f * g) = (\mathcal{F}f) \times (\mathcal{F}g)$.

Démonstration : - Ces deux propriétés découlent très simplement de la théorie de l'intégration.

— La continuité de $\mathcal{F}f$ repose sur le théorème de convergence dominée. En effet, on a

$$\forall \xi \in \mathbb{R}^N, \quad |\mathcal{F}f(\xi)| \leq \int_{\mathbb{R}^N} |\exp(-ix \cdot \xi)| |f(x)| dx \leq \|f\|_{L^1}$$

Comme, pour tout $x \in \mathbb{R}^N$,

$$\lim_{\xi \rightarrow \xi_0} \exp(-ix \cdot \xi) = \exp(-ix \cdot \xi_0),$$

on a alors

$$\lim_{\xi \rightarrow \xi_0} \mathcal{F}f(\xi) = \mathcal{F}f(\xi_0).$$

— Pour obtenir l'identité sur la convolution, on utilise le théorème de Fubini. Par définition,

$$\mathcal{F}(f * g)(\xi) = \int_{\mathbb{R}^N} \exp(-ix \cdot \xi) \left(\int_{\mathbb{R}^N} f(x-y)g(y)dy \right) dx.$$

Comme la fonction à intégrer est sommable sur $\mathbb{R}^{2N}$, on a

$$\mathcal{F}(f * g)(\xi) = \int_{\mathbb{R}^{2N}} \exp(-ix \cdot \xi) f(x-y)g(y)dx dy$$

d'où, par le changement de variable $z = x - y$,

$$\mathcal{F}(f * g)(\xi) = \int_{\mathbb{R}^{2N}} \exp(-iz \cdot \xi) f(z) \exp(-iy \cdot \xi) g(y) dx dy = \mathcal{F}f(\xi) \mathcal{F}g(\xi)$$

ce qui prouve que la convolution est transformée en produit. $\square$

L'ensemble des fonctions gaussiennes est conservé par transformation de Fourier. En effet, si on note $f : x \in \mathbb{R} \mapsto \exp\left(-\frac{x^2}{2}\right)$, en utilisant le théorème de convergence dominée, on montre que $\mathcal{F}f$ est de classe C^1 sur $\mathbb{R}$ et que

$$\begin{aligned} (\mathcal{F}f)'(\xi) &= \int_{\mathbb{R}} (-ix) \exp(-ix \cdot \xi) \exp\left(-\frac{x^2}{2}\right) dx = i \int_{\mathbb{R}} \exp(-ix \cdot \xi) \frac{d}{dx} \left(\exp\left(-\frac{x^2}{2}\right) \right) dx \\ &= -\xi \int_{\mathbb{R}} \exp(-ix \cdot \xi) \exp\left(-\frac{x^2}{2}\right) dx = -\xi \mathcal{F}f(\xi) \end{aligned}$$

On en déduit que

$$\mathcal{F}f(\xi) = \mathcal{F}f(0) \exp\left(-\frac{\xi^2}{2}\right) \text{ avec } \mathcal{F}f(0) = \int_{\mathbb{R}} \exp\left(-\frac{x^2}{2}\right) dx = \sqrt{2\pi}.$$

Cette propriété permet d'obtenir un procédé de régularisation (par convolution), qui est bien adapté à la transformation de Fourier. Il est utilisé par exemple pour montrer le

Théorème 6.1.3. (Théorème d'inversion de Fourier). - Soit $f \in L^1(\mathbb{R}^N)$ telle que $\mathcal{F}f \in L^1(\mathbb{R}^N)$. Alors on a l'identité

$$f(x) = \frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} \exp(ix \cdot \xi) \mathcal{F}f(\xi) d\xi = \frac{1}{(2\pi)^N} \overline{\mathcal{F}}(\mathcal{F}f)(x) \text{ presque partout sur } \mathbb{R}^N.$$

Démonstration :- La valeur au point x de la fonction figurant au membre de droite peut s'écrire

$$\frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} \exp(ix \cdot \xi) \left(\int_{\mathbb{R}^N} \exp(-iy \cdot \xi) f(y) dy \right) d\xi$$

mais il est impossible d'appliquer le théorème de Fubini, la fonction à intégrer n'étant pas sommable dans $\mathbb{R}^{2N}$. On la multiplie alors par la gaussienne $\exp\left(-\varepsilon^2 \frac{\xi^2}{2}\right)$ qui tend vers 1 quand $\varepsilon \rightarrow 0$, et on évalue l'intégrale

$$I_\varepsilon(x) = \frac{1}{(2\pi)^N} \int_{\mathbb{R}^{2N}} \exp(i(x-y) \cdot \xi) \exp\left(-\varepsilon^2 \frac{\xi^2}{2}\right) f(y) dy d\xi$$

de deux façons différentes.

— En intégrant d'abord par rapport à y , on obtient

$$I_\varepsilon(x) = \frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} \exp(ix \cdot \xi) \exp\left(-\varepsilon^2 \frac{\xi^2}{2}\right) \mathcal{F}f(\xi) d\xi.$$

Le théorème de convergence dominée de Lebesgue montre alors que

$$\lim_{\varepsilon \rightarrow 0} I_\varepsilon(x) = \frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} \exp(ix \cdot \xi) \mathcal{F}f(\xi) d\xi$$

— En intégrant d'abord par rapport à ξ , on obtient - en utilisant l'invariance de la gaussienne par transformation de Fourier -

$$\begin{aligned} I_\varepsilon(x) &= \frac{1}{(2\pi\varepsilon)^N} \int_{\mathbb{R}^N} f(y) \int_{\mathbb{R}^N} \exp\left(i\frac{1}{\varepsilon}(x-y) \cdot \varepsilon\xi\right) \exp\left(-\varepsilon^2 \frac{\xi^2}{2}\right) \varepsilon^N d\xi \\ &= \frac{1}{(\varepsilon\sqrt{2\pi})^N} \int_{\mathbb{R}^N} f(y) \exp\left(-\frac{(x-y)^2}{2\varepsilon}\right) dy. \end{aligned}$$

La fonction ρ_ε définie par

$$\rho_\varepsilon(z) = \frac{1}{\varepsilon^N} \rho\left(\frac{z}{\varepsilon}\right) \text{ où } \rho(z) = \frac{1}{(2\pi)^{N/2}} \exp\left(-\frac{z^2}{2}\right)$$

est positive, régulière (de classe C^∞) et d'intégrale 1. Autrement dit, ρ_ε est une suite régularisante qui converge (en mesure) vers la masse de Dirac. On a alors

$$\lim_{\varepsilon \rightarrow 0} I_\varepsilon(x) = f(x).$$

La formule d'inversion est alors obtenue par unicité de la limite. □

On considère le terme

$$I_\varepsilon(x) = \frac{1}{(2\pi)^N} \int_{\mathbb{R}^{2N}} \exp(i(x-y) \cdot \xi) \exp\left(-\varepsilon^2 \frac{|\xi|^2}{2}\right) f(y) dy d\xi.$$

En intégrant d'abord par rapport à ξ , on reconnaît une transformée de Fourier d'une gaussienne translatée. En effet, on peut écrire :

$$I_\varepsilon(x) = \frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} f(y) \left(\int_{\mathbb{R}^N} \exp(i(x-y) \cdot \xi) \exp\left(-\varepsilon^2 \frac{|\xi|^2}{2}\right) d\xi \right) dy.$$

On effectue le changement de variable $\eta = \varepsilon\xi$, d'où $\xi = \eta/\varepsilon$ et $d\xi = \varepsilon^{-N} d\eta$, ce qui donne :

$$\int_{\mathbb{R}^N} \exp(i(x-y) \cdot \xi) \exp\left(-\varepsilon^2 \frac{|\xi|^2}{2}\right) d\xi = \varepsilon^{-N} \int_{\mathbb{R}^N} \exp\left(i\frac{(x-y) \cdot \eta}{\varepsilon}\right) \exp\left(-\frac{|\eta|^2}{2}\right) d\eta.$$

Cette intégrale est exactement la transformée de Fourier inverse de la gaussienne $\exp(-|\eta|^2/2)$, évaluée en $(x-y)/\varepsilon$, ce qui donne une autre gaussienne :

$$(2\pi)^{N/2} \exp\left(-\frac{|x-y|^2}{2\varepsilon^2}\right).$$

Ainsi, en reportant dans l'expression de $I_\varepsilon(x)$, on obtient :

$$I_\varepsilon(x) = \frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} f(y) \cdot (2\pi)^{N/2} \varepsilon^{-N} \exp\left(-\frac{|x-y|^2}{2\varepsilon^2}\right) dy,$$

soit

$$I_\varepsilon(x) = \frac{1}{(\varepsilon\sqrt{2\pi})^N} \int_{\mathbb{R}^N} f(y) \exp\left(-\frac{|x-y|^2}{2\varepsilon^2}\right) dy,$$

ce qui est une convolution avec la gaussienne ρ_ε centrée en x , qui est une suite régularisante convergeant vers la masse de Dirac en x quand $\varepsilon \rightarrow 0$. Le facteur ε^{-N} provient du changement de variable dans une intégrale de dimension N . On part de l'intégrale

$$\int_{\mathbb{R}^N} \exp(i(x-y) \cdot \xi) \exp\left(-\varepsilon^2 \frac{|\xi|^2}{2}\right) d\xi,$$

et on effectue le changement de variable $\eta = \varepsilon\xi$, donc $\xi = \eta/\varepsilon$. Le jacobien de ce changement de variable dans $\mathbb{R}^N$ est :

$$d\xi = \varepsilon^{-N} d\eta,$$

car dans un changement de variable linéaire $x = ay$ dans $\mathbb{R}^N$, on a $dx = |a|^N dy$ (ici $a = 1/\varepsilon$). En remplaçant dans l'intégrale, on obtient :

$$\int_{\mathbb{R}^N} \exp\left(i(x-y) \cdot \frac{\eta}{\varepsilon}\right) \exp\left(-\frac{|\eta|^2}{2}\right) \varepsilon^{-N} d\eta.$$

D'où l'apparition du facteur ε^{-N} dans l'intégrale. Il est donc crucial de se rappeler que lorsqu'on change de variable dans une intégrale multiple, le jacobien affecte le volume élémentaire, ce qui explique la puissance $-N$ (et non simplement -1 comme dans le cas d'une variable réelle).

A posteriori, on voit que si f et $\mathcal{F}f$ sont dans $L^1(\mathbb{R}^N)$, f est égale (presque partout) à une fonction continue. Cela écarte des fonctions d'un usage tout-à-fait courant en mathématiques et en physique.

Plancherel a donc étendu la transformée de Fourier aux fonctions de carré sommable, et Schwartz aux distributions tempérées.

6.2 Transformation de Fourier sur la classe de Schwartz

Définition 6.2.1. La classe de Schwartz $\mathcal{S}(\mathbb{R}^N)$ est l'ensemble des fonctions $\varphi \in C^\infty(\mathbb{R}^N)$ à décroissance rapide, c'est-à-dire telles que

$$\forall p \in \mathbb{N}, \quad \exists C_p > 0, \quad \sup_{|\alpha| \leq p, |\beta| \leq p} \|x^\alpha \partial^\beta \varphi(x)\|_\infty \leq C_p.$$

Soit $\mathbb{R}^n$ un espace euclidien de dimension $n \geq 1$. La classe de Schwartz $\mathcal{S}(\mathbb{R}^n)$ est l'ensemble des fonctions $f : \mathbb{R}^n \rightarrow \mathbb{C}$ indéfiniment dérivables telles que, pour tout multi-indice $\alpha, \beta \in \mathbb{N}^n$, la fonction $x \mapsto x^\alpha \partial^\beta f(x)$ est bornée. Formellement,

$$\mathcal{S}(\mathbb{R}^n) = \left\{ f \in C^\infty(\mathbb{R}^n) \mid \sup_{x \in \mathbb{R}^n} |x^\alpha \partial^\beta f(x)| < \infty, \forall \alpha, \beta \in \mathbb{N}^n \right\}.$$

On munit $\mathcal{S}(\mathbb{R}^n)$ d'une topologie définie par la famille de semi-normes

$$\|f\|_{\alpha, \beta} = \sup_{x \in \mathbb{R}^n} |x^\alpha \partial^\beta f(x)|,$$

pour tout $\alpha, \beta \in \mathbb{N}^n$. Ainsi, $\mathcal{S}(\mathbb{R}^n)$ est un espace vectoriel topologique de Fréchet. L'espace des distributions tempérées, noté $\mathcal{S}'(\mathbb{R}^n)$, est l'ensemble des formes linéaires continues sur $\mathcal{S}(\mathbb{R}^n)$, c'est-à-dire,

$$\mathcal{S}'(\mathbb{R}^n) = \{T : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathbb{C} \mid T \text{ linéaire et continue}\}.$$

L'espace $\mathcal{S}'(\mathbb{R}^n)$ est strictement plus grand que $\mathcal{S}(\mathbb{R}^n)$: par exemple, la distribution de Dirac δ appartient à $\mathcal{S}'(\mathbb{R}^n)$ mais pas à $\mathcal{S}(\mathbb{R}^n)$. Ainsi, $\mathcal{S}(\mathbb{R}^n) \subsetneq \mathcal{S}'(\mathbb{R}^n)$, et les deux ensembles ne sont pas égaux.

Propriété 6.2.2. *La classe de Schwartz $\mathcal{S}(\mathbb{R}^N)$ est stable par dérivation, et par multiplication par les polynômes. Les fonctions $C_c^\infty(\mathbb{R}^N)$ sont denses dans $\mathcal{S}(\mathbb{R}^N)$.*

Démonstration. - On se donne $\varphi \in \mathcal{S}(\mathbb{R}^N)$.

— La stabilité par dérivation et par multiplication par les polynômes se vérifie par un simple calcul :

Pour tout $i \in \{1, \dots, N\}$, $\partial_i \varphi \in C^\infty(\mathbb{R}^N)$ et

$$\forall p \in \mathbb{N}, \quad \sup_{\substack{|\alpha| \leq p \\ |\beta| \leq p}} \|x^\alpha \partial^\beta \partial_i \varphi\|_\infty \leq \sup_{\substack{|\alpha| \leq p+1 \\ |\beta| \leq p+1}} \|x^\alpha \partial^\beta \varphi\|_\infty < +\infty.$$

Autrement dit, $\partial_i \varphi \in \mathcal{S}(\mathbb{R}^N)$.

Pour tout $P \in \mathbb{R}[X]$, $P\varphi \in C^\infty(\mathbb{R}^N)$ et, grâce à la formule de Leibniz,

$$\forall p \in \mathbb{N}, \quad \sup_{\substack{|\alpha| \leq p \\ |\beta| \leq p}} \|x^\alpha \partial^\beta (P\varphi)\|_\infty \leq C \sup_{\substack{|\alpha| \leq p+k \\ |\beta| \leq p+k}} \|x^\alpha \partial^\beta \varphi\|_\infty < +\infty,$$

où k est le degré de P . Autrement dit, $(P\varphi) \in \mathcal{S}(\mathbb{R}^N)$.

— L'approximation par des fonctions $C_c^\infty(\mathbb{R}^N)$ s'obtient par troncature.

On se donne $\psi \in C_c^\infty(\mathbb{R}^N, [0, 1])$ telle que $\psi \equiv 1$ sur $B(0, 1)$. On définit alors

$$\psi_j(x) = \psi\left(\frac{x}{j}\right) \text{ et } \varphi_j = \varphi \psi_j$$

de sorte que $\varphi_j \in C_c^\infty(\mathbb{R}^N)$ et $\varphi_j \equiv \varphi$ sur $B(0, j)$.

D'après la formule de Leibniz,

$$\partial^\beta (\varphi - \varphi_j)(x) = \partial^\beta \varphi (1 - \psi_j) - \sum_{\substack{\gamma \leq \beta \\ |\gamma| \geq 1}} C_{\beta}^{\gamma} j^{-|\gamma|} (\partial^\gamma \psi)_j \partial^{\beta-\gamma} \varphi,$$

de sorte que

$$\sup_{|\alpha| \leq p, |\beta| \leq p} \|x^\alpha \partial^\beta (\varphi - \varphi_j)\|_\infty \leq \frac{C}{j} \sup_{|\alpha| \leq p+1, |\gamma| \leq p+1} \|x^\alpha \partial^\gamma \varphi\|_\infty,$$

tend vers 0 quand $j \rightarrow \infty$. Ceci prouve que $\varphi_j \rightarrow \varphi$ dans $\mathcal{S}(\mathbb{R}^N)$. □

L'exemple typique de fonction de $\mathcal{S}(\mathbb{R})$ est la gaussienne, dont on a vu précédemment qu'elle est stable par transformation de Fourier. De façon plus générale, la classe de Schwartz est stable par transformation de Fourier.

Théorème 6.2.3. *La transformation de Fourier $\mathcal{F}$ est un isomorphisme de $\mathcal{S}(\mathbb{R}^N)$ sur lui-même, d'inverse $(2\pi)^{-N}\overline{\mathcal{F}}$.*

$\overline{\mathcal{F}}$ désigne la transformée de Fourier inverse, définie par $\overline{\mathcal{F}}(f)(x) = \int_{\mathbb{R}^N} e^{ix \cdot \xi} f(\xi) d\xi$.

Démonstration. - On commence par prouver que la transformation de Fourier échange régularité et décroissance à l'infini.

Soit $\varphi \in \mathcal{S}(\mathbb{R}^N)$.

$$\begin{aligned} \partial_{\xi_j}(\exp(-ix \cdot \xi)\varphi(x)) &= -ix_j \exp(-ix \cdot \xi)\varphi(x) \\ |-ix_j \exp(-ix \cdot \xi)\varphi(x)| &\leq |x_j \varphi(x)| \in L^1(\mathbb{R}^N). \end{aligned}$$

Par le théorème de convergence dominée, on a alors

$$\partial_{\xi_j}(\mathcal{F}\varphi) = \mathcal{F}(-ix_j\varphi)$$

Par récurrence, on en déduit que

$$\partial^\alpha(\mathcal{F}\varphi) = (-i)^{|\alpha|} \mathcal{F}(x^\alpha\varphi).$$

D'autre part, par intégration par parties, on a

$$\int \exp(-ix_j\xi_j) \partial_j\varphi(x) dx_j = i\xi_j \int \exp(-ix_j\xi_j) \varphi(x) dx_j$$

$$|\partial_j\varphi|, |\varphi| \in L^1(\mathbb{R}^N).$$

Par le théorème de Fubini, on a alors

$$\mathcal{F}(\partial_j\varphi) = i\xi_j\mathcal{F}\varphi$$

Par récurrence, on en déduit que

$$\mathcal{F}(\partial^\beta\varphi) = (i\xi)^\beta \mathcal{F}\varphi.$$

La transformation de Fourier applique donc $\mathcal{S}(\mathbb{R}^N)$ dans lui-même. En effet, en utilisant les propriétés précédentes et la formule de Leibniz, on a

$$\begin{aligned} \sup_{|\alpha| \leq p, |\beta| \leq p} \|\xi^\alpha \partial^\beta \mathcal{F}\varphi\|_\infty &\leq \sup_{|\alpha| \leq p, |\beta| \leq p} \|\partial^\alpha (x^\beta \varphi)\|_{L^1} \\ &\leq C \sup_{|\alpha| \leq p, |\beta| \leq p} \|x^\beta \partial^\alpha \varphi\|_{L^1} \\ &\leq C \sup_{|\alpha| \leq p, |\beta| \leq p+N+1} \|x^\beta \partial^\alpha \varphi\|_\infty \end{aligned}$$

puisque $(1 + |x|)^{-(N+1)}$ est intégrable sur $\mathbb{R}^N$. Cela signifie que, si $\varphi \in \mathcal{S}(\mathbb{R}^N)$, $\mathcal{F}\varphi \in \mathcal{S}(\mathbb{R}^N)$.

Le théorème d'inversion de Fourier, applicable lorsqu'on sait que φ et $\mathcal{F}\varphi$ sont dans $\mathcal{S}(\mathbb{R}^N) \subset L^1(\mathbb{R}^N)$, assure alors que $(2\pi)^{-N}\mathcal{F}\overline{\mathcal{F}} = (2\pi)^{-N}\overline{\mathcal{F}}\mathcal{F} = Id$. $\square$

6.3 Transformation de Fourier des distributions tempérées

Définition 6.3.1. Une distribution u est dite tempérée ($u \in \mathcal{S}'(\mathbb{R}^N)$) s'il existe $p \in \mathbb{N}$ et $C > 0$ tels que

$$\forall \varphi \in C_c^\infty(\mathbb{R}^N), \quad |\langle u, \varphi \rangle| \leq C \sup_{|\alpha| \leq p, |\beta| \leq p} \|x^\alpha \partial^\beta \varphi\|_\infty.$$

Une telle distribution se prolonge de manière unique en une forme linéaire sur $\mathcal{S}(\mathbb{R}^N)$. Soit (u_n) une suite de $\mathcal{S}'(\mathbb{R}^N)$. On dit que (u_n) converge vers u dans $\mathcal{S}'(\mathbb{R}^N)$ si

$$\forall \varphi \in \mathcal{S}(\mathbb{R}^N), \quad \langle u_n, \varphi \rangle \rightarrow \langle u, \varphi \rangle \text{ quand } n \rightarrow \infty.$$

Les espaces de Lebesgue $L^p(\mathbb{R}^N)$ s'identifient à des sous-ensembles de $\mathcal{S}'(\mathbb{R}^N)$. L'espace $\mathcal{E}'(\mathbb{R}^N)$ est un sous-espace de $\mathcal{S}'(\mathbb{R}^N)$ car toute distribution à support compact est nécessairement d'ordre fini.

Propriété 6.3.2. L'espace $\mathcal{S}'(\mathbb{R}^N)$ est stable par dérivation, et par multiplication par les fonctions $f \in C^\infty(\mathbb{R}^N)$ à croissance lente, i.e. telles que

$$\forall \beta \in \mathbb{N}^N, \quad \exists C_\beta, m_\beta, \quad |\partial^\beta f(x)| \leq C_\beta (1 + |x|)^{m_\beta}.$$

Démonstration. - On se donne $u \in \mathcal{S}'(\mathbb{R}^N)$.

Pour tout $i \in \{1, \dots, N\}$,

$$\begin{aligned} \forall \varphi \in C_c^\infty(\mathbb{R}^N), \quad |\langle \partial_i u, \varphi \rangle| &= |\langle u, \partial_i \varphi \rangle| \leq C \sup_{|\alpha| \leq p, |\beta| \leq p} \|x^\alpha \partial^\beta \partial_i \varphi\|_\infty \\ &\leq \sup_{|\alpha| \leq p+1, |\beta| \leq p+1} \|x^\alpha \partial^\beta \varphi\|_\infty \end{aligned}$$

ce qui prouve que $\partial_i u \in \mathcal{S}'(\mathbb{R}^N)$.

Pour toute fonction f à croissance lente,

$$\forall \varphi \in C_c^\infty(\mathbb{R}^N), \quad |\langle uf, \varphi \rangle| = |\langle u, f\varphi \rangle| \leq C \sup_{|\alpha| \leq p, |\beta| \leq p} \|x^\alpha \partial^\beta (f\varphi)\|_\infty$$

La formule de Leibniz donne

$$\partial^\beta (f\varphi) = \sum_{\gamma \leq \beta} C_\beta^\gamma \partial^\gamma f \partial^{\beta-\gamma} \varphi$$

Si on désigne par M_p le plus grand des m_β intervenant dans la croissance de $\partial^\beta f$ pour $|\beta| \leq p$, on a alors

$$\sup_{|\alpha| \leq p, |\beta| \leq p} \|x^\alpha \partial^\beta (f\varphi)\|_\infty \leq C \sup_{|\alpha| \leq p+M_p, |\beta| \leq p+M_p} \|x^\alpha \partial^\beta \varphi\|_\infty$$

d'où l'on déduit que

$$|\langle uf, \varphi \rangle| \leq C \sup_{|\alpha| \leq p+M_p, |\beta| \leq p+M_p} \|x^\alpha \partial^\beta \varphi\|_\infty.$$

Finalement $uf \in \mathcal{S}'(\mathbb{R}^N)$. □

Par dualité, on peut alors définir la transformation de Fourier sur les distributions tempérées.

Définition 6.3.3. Soit $u \in \mathcal{S}'(\mathbb{R}^N)$. La transformée de Fourier de u est la distribution $\mathcal{F}u$ définie par

$$\forall \varphi \in \mathcal{S}(\mathbb{R}^N), \quad \langle \mathcal{F}u, \varphi \rangle = \langle u, \mathcal{F}\varphi \rangle.$$

La distribution $\mathcal{F}u$ ainsi définie est une distribution tempérée puisqu'on a la majoration

$$\sup_{|\alpha| \leq p, |\beta| \leq p} \|\xi^\alpha \partial^\beta \mathcal{F}\varphi\|_\infty \leq C \sup_{|\alpha| \leq p+N+1, |\beta| \leq p+N+1} \|\xi^\beta \partial^\alpha \mathcal{F}\varphi\|_\infty$$

Théorème 6.3.4. La transformation de Fourier $\mathcal{F}$ est un isomorphisme de $\mathcal{S}'(\mathbb{R}^N)$ sur lui-même, d'inverse $(2\pi)^{-N}\overline{\mathcal{F}}$.

Démonstration – Par définition, la transformation de Fourier applique $\mathcal{S}'(\mathbb{R}^N)$ dans lui-même.

Comme $(2\pi)^{-N}\mathcal{F}\overline{\mathcal{F}} = (2\pi)^{-N}\overline{\mathcal{F}}\mathcal{F} = Id$ sur $\mathcal{S}(\mathbb{R}^N)$, on a pour tout $u \in \mathcal{S}'(\mathbb{R}^N)$, et tout $\varphi \in \mathcal{S}(\mathbb{R}^N)$

$$\begin{aligned} \langle \mathcal{F}\overline{\mathcal{F}}u, \varphi \rangle &= \langle u, \overline{\mathcal{F}}\mathcal{F}\varphi \rangle = (2\pi)^N \langle u, \varphi \rangle, \\ \langle \overline{\mathcal{F}}\mathcal{F}u, \varphi \rangle &= \langle u, \mathcal{F}\overline{\mathcal{F}}\varphi \rangle = (2\pi)^N \langle u, \varphi \rangle, \end{aligned}$$

d'où l'on déduit que $(2\pi)^{-N}\mathcal{F}\overline{\mathcal{F}} = (2\pi)^{-N}\overline{\mathcal{F}}\mathcal{F} = Id$ sur $\mathcal{S}'(\mathbb{R}^N)$. □

Comme conséquence directe de ce résultat, on obtient la formule de Plancherel pour les fonctions de carré sommable :

Corollaire 6.3.5. - La transformation de Fourier $\mathcal{F}$ est un isomorphisme de $L^2(\mathbb{R}^N)$ sur lui-même, d'inverse $(2\pi)^{-N}\overline{\mathcal{F}}$.

Démonstration. – On commence par montrer que $(2\pi)^{-N/2}\mathcal{F}$ est une isométrie de $\mathcal{S}(\mathbb{R}^N)$ sur lui-même, lorsque celui-ci est muni de la norme L^2 . Pour toutes fonctions $f, g \in \mathcal{S}(\mathbb{R}^N)$, on a

$$\begin{aligned} \int f(x)\overline{g}(x)dx &= \frac{1}{(2\pi)^N} \int f(x)(\mathcal{F}\overline{\mathcal{F}}\overline{g})(x)dx \\ &= \frac{1}{(2\pi)^N} \int \mathcal{F}f(\xi)(\overline{\mathcal{F}}\overline{g})(\xi)d\xi \\ &= \frac{1}{(2\pi)^N} \int \mathcal{F}f(\xi)\overline{\mathcal{F}g}(\xi)d\xi \end{aligned}$$

ce qui montre que $(2\pi)^{-N/2}\mathcal{F}$ conserve le produit scalaire, et donc la norme L^2 pour les éléments de $\mathcal{S}(\mathbb{R}^N)$.

L'ensemble $\mathcal{S}(\mathbb{R}^N)$ (qui contient $C_c^\infty(\mathbb{R}^N)$) est dense dans $L^2(\mathbb{R}^N)$. Pour tout $u \in L^2(\mathbb{R}^N)$, il existe donc une suite (u_j) de $\mathcal{S}(\mathbb{R}^N)$ telle que

$$u_j \rightarrow u \text{ dans } L^2(\mathbb{R}^N)$$

La suite (u_j) est de Cauchy, et donc - par isométrie - la suite $(\mathcal{F}u_j)$ est aussi de Cauchy dans $L^2(\mathbb{R}^N)$. Comme $L^2(\mathbb{R}^N)$ est complet, on a alors

$$\mathcal{F}u_j \rightarrow g \text{ dans } L^2(\mathbb{R}^N).$$

En identifiant $L^2(\mathbb{R}^N)$ à un sous-espace de $\mathcal{S}'(\mathbb{R}^N)$, on obtient

$$\mathcal{F}u_j \rightarrow \mathcal{F}u \text{ dans } \mathcal{S}'(\mathbb{R}^N)$$

d'où l'on déduit que $\mathcal{F}u = g \in L^2(\mathbb{R}^N)$.

Par continuité de la norme, on a de plus $(2\pi)^{-N/2} \|\mathcal{F}u\|_{L^2(\mathbb{R}^N)} = \|u\|_{L^2(\mathbb{R}^N)}$.

En appliquant le même raisonnement à $(2\pi)^{-N/2} \overline{\mathcal{F}}$, on conclut qu'on a deux isométries de $L^2(\mathbb{R}^N)$ dans lui-même, dont les composées à droite et à gauche sont égales à l'identité. Cela termine la démonstration du théorème. $\square$

La définition de la transformée de Fourier d'une distribution ne constitue pas un moyen de calcul effectif. On dispose de formules explicites lorsque u est une fonction sommable. A partir de résultats déjà connus, on peut en déduire d'autres par inversion de Fourier, passage à la limite, dérivation, convolution...

Propriété 6.3.6. Soit $u \in \mathcal{S}'(\mathbb{R}^N)$.

- (i) Si u possède la symétrie hermitienne, $\mathcal{F}u$ est réelle.
- (ii) Si u est paire (resp. impaire), $\mathcal{F}u$ est paire (resp. impaire).
- (iii) Si τ_a désigne la translation de a , $\mathcal{F}(\tau_a u)(\xi) = \exp(-ia \cdot \xi) \mathcal{F}u(\xi)$.
- (iv) Si h_λ désigne la dilatation de rapport λ , $\mathcal{F}(h_\lambda u) = |\lambda|^{-N} h_{\lambda^{-1}}(\mathcal{F}u)$.
- (v) Si $v \in \mathcal{E}'(\mathbb{R}^N)$, $\mathcal{F}(u * v) = (\mathcal{F}u)(\mathcal{F}v)$.

Démonstration. – Les propriétés (i) – (iv) se montrent par dualité, en utilisant la formule du changement de variable. Il suffit donc de les vérifier pour les fonctions $\varphi \in \mathcal{S}(\mathbb{R}^N)$.

Si φ possède la symétrie hermitienne $\varphi(-x) = \bar{\varphi}(x)$

$$\begin{aligned} \mathcal{F}\varphi(\xi) &= \int \exp(-ix \cdot \xi) \varphi(x) dx \\ &= \frac{1}{2} \int (\exp(-ix \cdot \xi) \varphi(x) + \exp(ix \cdot \xi) \varphi(-x)) dx = \operatorname{Re} \left(\int \exp(-ix \cdot \xi) \varphi(x) dx \right) \end{aligned}$$

Pour la conservation de la parité, il suffit de remarquer que

$$\mathcal{F}\varphi(-\xi) = \int \exp(ix \cdot \xi) \varphi(x) dx = \int \exp(-ix \cdot \xi) \varphi(-x) dx$$

De la même façon, on a

$$\begin{aligned} \mathcal{F}(\tau_a \varphi)(\xi) &= \int \exp(-ix \cdot \xi) \varphi(x - a) dx \\ &= \int \exp(-i(y + a) \cdot \xi) \varphi(y) dy \\ &= \exp(-ia \cdot \xi) \mathcal{F}\varphi(\xi) \\ \mathcal{F}(h_\lambda \varphi)(\xi) &= \int \exp(-ix \cdot \xi) \varphi(\lambda x) dx \\ &= \frac{1}{|\lambda|^N} \int \exp(-iy \cdot \frac{\xi}{\lambda}) \varphi(y) dy \\ &= |\lambda|^{-N} h_{\lambda^{-1}}(\mathcal{F}\varphi)(\xi). \end{aligned}$$

Reste à montrer l'échange de la convolution et de la multiplication.

- On commence pour cela par prouver que pour toute distribution $v \in \mathcal{E}'(\mathbb{R}^N)$, $\mathcal{F}v$ est une fonction à croissance lente et que $\mathcal{F}v(\xi) = \langle v, \exp(-ix \cdot \xi) \rangle$.

D'après le théorème de dérivation sous le crochet, la fonction $\xi \mapsto \langle v, \exp(-ix \cdot \xi) \rangle$ est de classe C^∞ et

$$\partial^\alpha \langle v, \exp(-ix \cdot \xi) \rangle = \langle v, (-ix)^\alpha \exp(-ix \cdot \xi) \rangle.$$

Si K est un voisinage compact du support de v et si p est l'ordre de v ,

$$|\partial^\alpha \langle v, \exp(-ix \cdot \xi) \rangle| \leq C \sup_{|\beta| \leq p} \|\partial^\beta (x^\alpha \exp(-ix \cdot \xi))\|_\infty \leq C(1 + |\xi|^p)$$

où C ne dépend que de K et de α . La fonction $\xi \mapsto \langle v, \exp(-ix \cdot \xi) \rangle$ est donc à croissance lente.

Pour $\varphi \in C_c^\infty(\mathbb{R}^N)$, d'après le théorème d'intégration sous le crochet, on a

$$\int_{\mathbb{R}^N} \langle v, \exp(-ix \cdot \xi) \rangle \varphi(\xi) d\xi = \left\langle v, \int \exp(-ix \cdot \xi) \varphi(\xi) d\xi \right\rangle = \langle v, \mathcal{F}\varphi \rangle = \langle \mathcal{F}v, \varphi \rangle$$

ce qui montre l'identité $\mathcal{F}v(\xi) = \langle v, \exp(-ix \cdot \xi) \rangle$.

- Soient alors $u, v \in \mathcal{E}'(\mathbb{R}^N)$. D'après ce qui précède, $\mathcal{F}u, \mathcal{F}v$ et $\mathcal{F}(u * v)$ sont des fonctions C^∞ à croissance lente, et on a - en notant $e_\xi : x \mapsto \exp(-ix \cdot \xi)$ -

$$\begin{aligned} \mathcal{F}(u * v)(\xi) &= \langle u * v, e_\xi \rangle = \langle u, \hat{v} * e_\xi \rangle \\ \hat{v} * e_\xi(x) &= \langle \hat{v}, \widehat{\tau_x e_\xi} \rangle = \langle v, \tau_x e_\xi \rangle \\ &= \exp(-ix \cdot \xi) \langle v, e_\xi \rangle = \exp(-ix \cdot \xi) \mathcal{F}v(\xi) \end{aligned}$$

En reportant dans la première identité, on obtient

$$\mathcal{F}(u * v)(\xi) = \mathcal{F}u(\xi) \mathcal{F}v(\xi)$$

Si $u \in \mathcal{E}'(\mathbb{R}^N)$ et $v \in \mathcal{S}'(\mathbb{R}^N)$, on approche v par une suite (v_j) d'éléments de $\mathcal{E}'(\mathbb{R}^N)$ (obtenue par troncature) convergeant dans $\mathcal{S}'(\mathbb{R}^N)$. On a alors

$$\mathcal{F}(u * v_j)(\xi) = \mathcal{F}u(\xi) \mathcal{F}v_j(\xi)$$

Comme $\mathcal{F}u$ est à croissance lente, et que $\mathcal{F}v_j \rightarrow \mathcal{F}v$ dans $\mathcal{S}'(\mathbb{R}^N)$, on a

$$\mathcal{F}u \mathcal{F}v_j \rightarrow \mathcal{F}u \mathcal{F}v \text{ dans } \mathcal{S}'(\mathbb{R}^N).$$

Par conséquent, $u * v_j$ converge dans $\mathcal{S}'(\mathbb{R}^N)$. Comme par ailleurs on sait que $u * v_j \rightarrow u * v$ dans $\mathcal{D}'(\mathbb{R}^N)$, par unicité de la limite on a

$$\mathcal{F}(u * v_j) \rightarrow \mathcal{F}(u * v) \text{ dans } \mathcal{S}'(\mathbb{R}^N).$$

On a bien l'identité annoncée. □

6.4 Equations de convolution

Comme on l'a déjà mentionné au chapitre précédent, la transformation de Fourier est un outil très puissant pour résoudre les équations aux dérivées partielles linéaires à coefficients constants.

De façon plus générale, on s'intéresse ici aux équations de convolution, i.e. aux équations de la forme

$$T * S = R \text{ où } S \in \mathcal{E}'(\mathbb{R}^N), R \in \mathcal{S}'(\mathbb{R}^N) \text{ sont des distributions données.} \quad (6.4.1)$$

Quand $R = \delta$, on dit que T est une solution élémentaire de l'équation de convolution de noyau S .

D'après la propriété 6.3.6 (v), l'équation (6.4.1) est équivalente à

$$\mathcal{F}(S)\mathcal{F}(T) = \mathcal{F}(R)$$

puisque la transformation de Fourier est bijective sur $\mathcal{S}'(\mathbb{R}^N)$.

Lorsque S est une distribution supportée en $\{0\}$, donc de la forme

$$S = \sum_{|\alpha| \leq m} c_\alpha \delta^{(\alpha)}$$

on a

$$T * S = \sum_{|\alpha| \leq m} c_\alpha T * \delta^{(\alpha)} = \sum_{|\alpha| \leq m} c_\alpha \partial^\alpha T.$$

L'équation de convolution (6.4.1) est équivalente à l'équation aux dérivées partielles

$$\sum_{|\alpha| \leq m} c_\alpha \partial^\alpha T = R.$$

Proposition 6.4.1. - Soit $P(\partial) = \sum_{|\alpha| \leq m} c_\alpha \partial^\alpha$ un opérateur différentiel à coefficients constants. L'équation

$$P(\partial)T = 0$$

possède une solution tempérée non nulle si et seulement si le polynôme $P(i\xi)$ possède une racine $\xi_0 \in \mathbb{R}^N$.

Démonstration. - Pour $T \in \mathcal{S}'(\mathbb{R}^N)$, l'équation $P(\partial)T = 0$ équivaut à

$$P(i\xi)\mathcal{F}T = 0.$$

On a donc $\mathcal{F}T \equiv 0$ (et $T \equiv 0$) si le polynôme $P(i\xi)$ ne s'annule jamais.

Réciproquement, s'il existe $\xi_0 \in \mathbb{R}^N$ tel que $P(i\xi_0) = 0$, alors la distribution tempérée

$$T(x) = \exp(i\xi_0 \cdot x) \text{ est telle que } \mathcal{F}T = (2\pi)^{-N} \delta_{\xi_0}.$$

Elle vérifie donc $P(\partial)T = 0$. □

On vérifie ainsi que les distributions tempérées harmoniques sont nécessairement des polynômes. En effet, si

$$\mathcal{F}(\Delta T) = -|\xi|^2 \mathcal{F}T = 0$$

le support de $\mathcal{F}T$ est $\{0\}$, ce qui implique que $\mathcal{F}T$ est une combinaison linéaire de la masse de Dirac et de ses dérivées. Autrement dit,

$$T(x) = \sum_{|\alpha| \leq m} c_\alpha \frac{(-ix)^\alpha}{(2\pi)^N}.$$

Attention : dès que $N \geq 2$, il existe des polynômes harmoniques de degré arbitraire, puisque $T = (x_1 + ix_2)^k$ vérifie $\Delta T = 0$.

Espaces de Hilbert

Dans cette partie du cours, on va s'intéresser aux propriétés spectrales des opérateurs linéaires définis sur des espaces de Banach. Ces propriétés peuvent être précisées dans le cas où les espaces ont une structure hilbertienne, i.e. dans le cas où la norme est construite sur un produit scalaire.

On va donc commencer par rappeler quelques résultats importants concernant les espaces de Hilbert, et on renvoie à la littérature existante pour plus de détails.

7.1 Définition et propriétés fondamentales

Définition 7.1.1. *Un espace de Hilbert est un espace vectoriel H muni d'un produit scalaire (i.e. une forme bilinéaire de $H \times H$ dans $\mathbb{R}$, symétrique et définie positive), et qui est complet pour la norme associée*

$$\forall u \in H, \quad |u| = (u, u)^{1/2}.$$

L'identité du parallélogramme

$$\left| \frac{a+b}{2} \right|^2 + \left| \frac{a-b}{2} \right|^2 = \frac{1}{2} (|a|^2 + |b|^2)$$

montre que les espaces de Hilbert sont uniformément convexes, donc réflexifs (théorème de Milman Pettis). Les résultats de topologie générale permettent alors d'obtenir des propriétés très fortes sur la dualité ou les projections sur les convexes fermés.

Exemple fondamental : $L^2(\Omega)$ muni du produit scalaire

$$(u, v) = \int_{\Omega} u(x)v(x)dx$$

est un espace de Hilbert ; l'espace de Sobolev H^1 que nous rencontrerons au chapitre 8 est un espace de Hilbert «modélé» sur L^2 .

On va en donner ici des démonstrations élémentaires constructives, basées sur l'utilisation du produit scalaire.

Définition 7.1.2. *Soit H un espace de Hilbert, et $K \subset H$ un sous-ensemble non vide. On appelle projection de $f \in H$ sur K tout élément $u \in K$ tel que*

$$\|f - u\| = \inf_{v \in K} \|f - v\|.$$

Lorsque K est un convexe fermé non vide, cette projection existe et est unique.

Théorème 7.1.3. (*Projection sur un convexe fermé*). — Soient H un espace de Hilbert, et $K \subset H$ un convexe fermé non vide.

Alors pour tout $f \in H$, il existe une unique projection de f sur K , i.e. un unique $u \in K$ tel que $|f - u| = \min_{v \in K} |f - v|$.

— La projection $u \in K$ de f sur K est caractérisée par l'inégalité

$$\forall v \in K, \quad (f - u, v - u) \leq 0.$$

En particulier l'opérateur de projection $P_K : f \mapsto u$ est contractant.

— Si K est un sous-espace vectoriel, on a l'égalité

$$\forall v \in K, \quad (f - u, v - u) = 0$$

et l'opérateur de projection $P_K : f \mapsto u$ est linéaire.

Démonstration. - Soit (u_n) une suite minimisante pour la distance $d = \inf_{v \in K} |f - v|$

$$u_n \in K, \quad d_n = |f - u_n| \rightarrow d.$$

En appliquant l'identité du parallélogramme avec $a = f - u_n$ et $b = f - u_m$, on obtient

$$\left| f - \frac{u_n + u_m}{2} \right|^2 + \left| \frac{u_n - u_m}{2} \right|^2 = \frac{1}{2} (d_n^2 + d_m^2).$$

L'égalité

$$\left| f - \frac{u_n + u_m}{2} \right|^2 + \left| \frac{u_n - u_m}{2} \right|^2 = \frac{1}{2} (d_n^2 + d_m^2)$$

provient de l'identité du parallélogramme dans un espace de Hilbert. En effet, pour tout $a, b \in H$,

$$\left| \frac{a+b}{2} \right|^2 + \left| \frac{a-b}{2} \right|^2 = \frac{1}{2} (|a|^2 + |b|^2).$$

On applique cette identité avec $a = f - u_n$ et $b = f - u_m$, ce qui donne l'égalité souhaitée. Les quantités d_n et d_m sont définies par $d_n = |f - u_n|$ et $d_m = |f - u_m|$. Ce sont les distances entre f et les points $u_n, u_m \in K$, qui forment une suite minimisante pour la distance de f à K .

Comme $(u_n + u_m)/2 \in K$, sa distance à f est supérieure à d . Donc

$$\left| \frac{u_n - u_m}{2} \right|^2 = \frac{1}{2} (d_n^2 + d_m^2) - d^2 \rightarrow 0 \text{ quand } n, m \rightarrow \infty.$$

La suite (u_n) est de Cauchy et converge donc vers $u \in K$.

— Si u réalise la distance de f à K , on a pour tout $v \in K$

$$\forall t \in [0, 1], \quad w = (1 - t)u + tv \in K$$

et donc

$$|f - u|^2 \leq |f - w|^2 \leq |(f - u) - t(v - u)|^2 = |f - u|^2 - 2t(f - u, v - u) + t^2|v - u|^2.$$

En faisant tendre $t \rightarrow 0$, on obtient alors

$$(f - u, v - u) \leq 0.$$

On utilise ici le fait que $w = (1 - t)u + tv = u + t(v - u)$, donc

$$f - w = f - u - t(v - u) = (f - u) - t(v - u).$$

Alors

$$|f - w|^2 = |(f - u) - t(v - u)|^2 = |f - u|^2 - 2t(f - u, v - u) + t^2|v - u|^2.$$

En utilisant que $|f - u|^2 \leq |f - w|^2$ (par minimalité de u), on obtient

$$|f - u|^2 \leq |f - w|^2 \leq |f - u|^2 - 2t(f - u, v - u) + t^2|v - u|^2.$$

En faisant tendre $t \rightarrow 0$, le terme en t^2 devient négligeable, et on obtient que $-2t(f - u, v - u) \geq 0$ pour tout $t > 0$, donc $(f - u, v - u) \leq 0$.

Dans l'inégalité

$$|f - w|^2 = |(f - u) - t(v - u)|^2 = |f - u|^2 - 2t(f - u, v - u) + t^2|v - u|^2,$$

on a utilisé l'identité de polarisation :

$$|a + tb|^2 = |a|^2 + 2t(a, b) + t^2|b|^2.$$

Ici $a = f - u$ et $b = -(v - u)$, d'où le signe négatif devant le terme linéaire. Ensuite, puisque $|f - u|^2 \leq |f - w|^2$ par minimalité de u , on obtient :

$$0 \leq |f - w|^2 - |f - u|^2 = -2t(f - u, v - u) + t^2|v - u|^2.$$

En divisant par t (positif) et en faisant tendre $t \rightarrow 0$, le terme en t tend vers 0, et on obtient :

$$-2(f - u, v - u) \geq 0 \quad \Rightarrow \quad (f - u, v - u) \leq 0.$$

Inversement, si pour tout $v \in K$ on a $(f - u, v - u) \leq 0$, alors

$$|u - f|^2 - |v - f|^2 = 2(f - u, v - u) - |v - u|^2 \leq 0$$

donc u réalise la distance de f à K .

Cette caractérisation de u montre en particulier que la projection $P_K f$ est unique : si pour tout $v \in K$

$$(f - u_1, v - u_1) \leq 0 \text{ et } (f - u_2, v - u_2) \leq 0$$

alors

$$(u_1 - u_2, u_1 - u_2) \leq 0.$$

En supposant que u_1 et u_2 sont deux projections, on a pour tout $v \in K$:

$$(f - u_1, v - u_1) \leq 0 \quad \text{et} \quad (f - u_2, v - u_2) \leq 0.$$

En particulier, on prend $v = u_2$ dans la première inégalité, ce qui donne :

$$(f - u_1, u_2 - u_1) \leq 0.$$

Et $v = u_1$ dans la deuxième inégalité :

$$(f - u_2, u_1 - u_2) \leq 0.$$

En additionnant ces deux inégalités :

$$(f - u_1, u_2 - u_1) + (f - u_2, u_1 - u_2) \leq 0.$$

Or $(f - u_1, u_2 - u_1) + (f - u_2, u_1 - u_2) = (f - u_1 - f + u_2, u_2 - u_1) = (u_2 - u_1, u_2 - u_1) = |u_1 - u_2|^2$. Donc :

$$|u_1 - u_2|^2 \leq 0 \quad \Rightarrow \quad u_1 = u_2.$$

Revenons à l'expression suivante utilisée dans la démonstration :

$$(f - u_1, u_2 - u_1) + (f - u_2, u_1 - u_2)$$

Observons que :

$$u_1 - u_2 = -(u_2 - u_1)$$

Donc :

$$(f - u_2, u_1 - u_2) = -(f - u_2, u_2 - u_1)$$

Ainsi, on peut écrire :

$$(f - u_1, u_2 - u_1) + (f - u_2, u_1 - u_2) = (f - u_1, u_2 - u_1) - (f - u_2, u_2 - u_1)$$

On factorise alors :

$$= [(f - u_1) - (f - u_2), u_2 - u_1] = (u_2 - u_1, u_2 - u_1) = \|u_2 - u_1\|^2$$

Donc l'étape :

$$(f - u_1, u_2 - u_1) + (f - u_2, u_1 - u_2) = (f - u_1 - f + u_2, u_2 - u_1)$$

est bien **justifiée**, mais uniquement parce que les deux produits scalaires sont pris avec des vecteurs **opposés**, ce qui permet la simplification correcte. Autrement dit, ce n'est pas une application directe d'une identité comme :

$$(u, v) + (w, v) = (u + w, v)$$

qui est fausse, mais plutôt un **regroupement soigné** tenant compte des signes des vecteurs impliqués.

Elle montre aussi que le projecteur P_K est contractant : si pour tout $v \in K$

$$(f_1 - u_1, v - u_1) \leq 0 \text{ et } (f_2 - u_2, v - u_2) \leq 0$$

alors

$$(u_1 - u_2, u_1 - u_2) \leq (f_1 - f_2, u_1 - u_2) \leq |f_1 - f_2| |u_1 - u_2|.$$

On a pour tout $v \in K$:

$$(f_1 - u_1, v - u_1) \leq 0 \text{ et } (f_2 - u_2, v - u_2) \leq 0.$$

On applique la première avec $v = u_2$ et la seconde avec $v = u_1$:

$$(f_1 - u_1, u_2 - u_1) \leq 0 \text{ et } (f_2 - u_2, u_1 - u_2) \leq 0.$$

On somme les deux inégalités :

$$(f_1 - u_1, u_2 - u_1) + (f_2 - u_2, u_1 - u_2) \leq 0.$$

Cela donne :

$$(f_1 - f_2 - u_1 + u_2, u_2 - u_1) \leq 0,$$

c'est-à-dire :

$$(f_1 - f_2, u_2 - u_1) \leq |u_1 - u_2|^2.$$

En réécrivant :

$$|u_1 - u_2|^2 \leq (f_1 - f_2, u_1 - u_2) \leq |f_1 - f_2| |u_1 - u_2|,$$

d'où on déduit que $|u_1 - u_2| \leq |f_1 - f_2|$.

— L'inégalité précédente devient une égalité dans le cas où K est un sous-espace vectoriel. En effet, si pour $v \in K$ et tout $t \in \mathbb{R}$,

$$(f - u, tv - u) \leq 0$$

on a nécessairement

$$(f - u, v) = 0$$

Cette identité montre en particulier que le projecteur est linéaire. □

Soit K un sous-espace vectoriel de l'espace de Hilbert considéré, et soit $u \in K$ une fonction telle que pour tout $v \in K$ et tout $t \in \mathbb{R}$, on ait

$$(f - u, tv - u) \leq 0.$$

En particulier, cette inégalité est satisfaite pour tout $t \in \mathbb{R}$ fixé.

Posons la fonction $\varphi(t) = (f - u, tv - u)$, que l'on développe comme suit :

$$\varphi(t) = (f - u, tv - u) = t(f - u, v) - (f - u, u).$$

L'inégalité donnée devient donc :

$$\varphi(t) = t(f - u, v) - (f - u, u) \leq 0, \quad \forall t \in \mathbb{R}.$$

Supposons que $(f - u, v) \neq 0$. Alors, la fonction affine $\varphi(t)$ est strictement croissante ou décroissante selon le signe de $(f - u, v)$, et tend vers $\pm\infty$ lorsque $t \rightarrow \pm\infty$. Dans ce cas, il existe un t tel que $\varphi(t) > 0$, ce qui contredirait l'hypothèse $\varphi(t) \leq 0$ pour tout t . Donc, on doit nécessairement avoir :

$$(f - u, v) = 0,$$

ce qui montre que l'inégalité devient une égalité dans le cas où K est un sous-espace vectoriel.

Théorème 7.1.4. (Théorème de Riesz-Fréchet). - Soient H un espace de Hilbert, et H^* son dual. Pour tout $\varphi \in H^*$, il existe un unique $f \in H$ tel que

$$\forall v \in H, \quad \langle \varphi, v \rangle = (f, v)$$

En particulier, $\|\varphi\|_{H^*} = |f|$.

Démonstration. - Soit $M = \varphi^{-1}(0)$. Par définition, M est un sous-espace fermé de H .

Si $M = H$, $\varphi \equiv 0$ et on prend $f = 0$.

$M = \varphi^{-1}(0)$ est un sous-espace fermé de H car φ est une forme linéaire continue (c'est-à-dire un élément de H^* , le dual continu de H), et le noyau d'une application linéaire continue entre espaces de Banach (ici, de Hilbert) est toujours un fermé. En effet, l'ensemble $\{v \in H : \varphi(v) = 0\}$ est le préimage de l'ensemble fermé $\{0\}$ par l'application continue φ , donc fermé.

Sinon, il existe un élément h de H tel que $\langle \varphi, h \rangle \neq 0$, autrement dit tel que $h \notin M$. On pose alors

$$g = \frac{h - P_M h}{|h - P_M h|} \text{ de sorte que } |g| = 1 \text{ et } \forall w \in M, (g, w) = 0.$$

Dans la construction $g = \frac{h - P_M h}{|h - P_M h|}$, on choisit un vecteur orthogonal à M de norme 1. En effet, $P_M h$ est la projection orthogonale de h sur M , donc $h - P_M h$ est orthogonal à M . En normalisant ce vecteur, on obtient g de norme 1 (i.e., $|g| = 1$), et par construction, $(g, w) = 0$ pour tout $w \in M$.

Tout $v \in H$ se décompose alors sous la forme

$$v = \frac{\langle \varphi, v \rangle}{\langle \varphi, g \rangle} g + w \text{ avec } w \in M$$

Tout vecteur $v \in H$ peut se décomposer de manière unique selon la somme orthogonale $H = M \oplus M^\perp$. Comme $g \in M^\perp$, la composante de v sur $M^\perp$ est colinéaire à g . On écrit donc $v = \lambda g + w$ avec $w \in M$. En appliquant φ à cette décomposition, on obtient $\langle \varphi, v \rangle = \lambda \langle \varphi, g \rangle$, d'où $\lambda = \frac{\langle \varphi, v \rangle}{\langle \varphi, g \rangle}$, ce qui donne la forme souhaitée. On a alors

$$(g | v) = \frac{\langle \varphi, v \rangle}{\langle \varphi, g \rangle} |g|^2 = \frac{\langle \varphi, v \rangle}{\langle \varphi, g \rangle}.$$

La notation $(g | v)$ désigne le produit scalaire dans l'espace de Hilbert H , c'est-à-dire (g, v) dans la notation usuelle. Cette notation est équivalente et parfois utilisée pour insister sur le rôle d'un produit hermitien dans le cas complexe.

On obtient donc la représentation attendue en posant $f = \langle \varphi, g \rangle g$. □

Soit $f = \langle \varphi, g \rangle g$. Montrons que pour tout $v \in H$, on a $\langle \varphi, v \rangle = (f, v)$.

Tout d'abord, tout élément $v \in H$ se décompose de manière unique comme

$$v = \frac{\langle \varphi, v \rangle}{\langle \varphi, g \rangle} g + w, \quad \text{avec } w \in M = \ker(\varphi),$$

c'est-à-dire $\langle \varphi, w \rangle = 0$.

Alors,

$$(f, v) = \left(\langle \varphi, g \rangle g, \frac{\langle \varphi, v \rangle}{\langle \varphi, g \rangle} g + w \right) = \langle \varphi, g \rangle \cdot \frac{\langle \varphi, v \rangle}{\langle \varphi, g \rangle} \cdot (g, g) + \langle \varphi, g \rangle \cdot (g, w).$$

Or, $|g| = 1$ donc $(g, g) = 1$, et comme $w \in M$, on a $(g, w) = 0$ (car $g \perp M$). Il vient donc :

$$(f, v) = \langle \varphi, v \rangle.$$

Ainsi, $f = \langle \varphi, g \rangle g$ est bien l'élément de H tel que $\langle \varphi, v \rangle = (f, v)$ pour tout $v \in H$. Cela achève la démonstration.

Le théorème de Riesz-Fréchet montre que toute forme linéaire continue sur H peut se représenter à l'aide du produit scalaire, ce qui permet d'identifier H et H^* . On fera très souvent cette identification, mais pas toujours.

Typiquement, si V et H sont deux espaces de Hilbert tels que V est un sous-espace vectoriel de H dense dans H , on a

$$V \subset H \text{ et } H^* \subset V^*$$

On ne peut alors pas identifier simultanément H et H^* , et V et V^* . En général, on identifie l'espace pivot H et son dual H^* , de sorte qu'on a

$$V \subset H = H^* \subset V^*$$

7.2 Théorèmes de Stampacchia et de Lax-Milgram

Définition 7.2.1. On dit qu'une forme bilinéaire $a : H \times H \rightarrow \mathbb{R}$ est

— continue s'il existe $C > 0$ telle que

$$\forall u, v \in H, \quad |a(u, v)| \leq C|u||v|$$

— coercive s'il existe $\alpha > 0$ telle que

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

Théorème 7.2.2. (Théorème de Stampacchia). - Soient H un espace de Hilbert, K un convexe fermé non vide, et $a : H \times H \rightarrow \mathbb{R}$ une forme bilinéaire continue coercive.

Pour tout $\varphi \in H^*$, il existe un unique $u \in K$ tel que

$$\forall v \in K, \quad a(u, v - u) \geq \langle \varphi, v - u \rangle$$

De plus, si a est symétrique, $u \in K$ est caractérisé par la propriété

$$\frac{1}{2}a(u, u) - \langle \varphi, u \rangle = \min_{v \in K} \left(\frac{1}{2}a(v, v) - \langle \varphi, v \rangle \right).$$

Ce théorème affirme qu'étant donné une forme bilinéaire continue et coercive $a(\cdot, \cdot)$ sur un espace de Hilbert H , et un convexe fermé $K \subset H$, alors pour tout élément $f \in H'$ (le dual de H), il existe une unique fonction $u \in K$ telle que :

$$a(u, v - u) \geq \langle f, v - u \rangle \quad \text{pour tout } v \in K.$$

Cette formulation s'interprète comme une généralisation des problèmes aux limites elliptiques du second ordre en présence de contraintes. Le théorème est particulièrement utile pour l'étude de problèmes comme l'obstacle (où une fonction est contrainte à rester au-dessus d'un certain seuil) et plus généralement pour tout problème où les solutions sont naturellement restreintes à un ensemble de configurations admissibles.

Démonstration. - La preuve de ce résultat repose sur le théorème de point fixe de Picard que nous rappelons ici sans démonstration.

Théorème 7.2.3. (*Théorème de point fixe de Banach - méthode des approximations successives de Picard*). Soit X un espace métrique complet et soit $S : X \rightarrow X$ une application telle que $d(Sv_1, Sv_2) < kd(v_1, v_2)$ avec $\forall v_1, v_2 \in X$, $k < 1$. Alors S admet un point fixe unique, $u = Su$.

D'après le théorème de Riesz-Fréchet, il existe un unique $f \in H$ tel que

$$\forall v \in H, \quad \langle \varphi, v \rangle = (f, v).$$

Comme, pour tout $u \in H$ fixé, l'application $v \mapsto a(u, v)$ est une forme linéaire continue sur H , il existe un élément Au de H tel que

$$\forall v \in H, \quad a(u, v) = (Au, v).$$

Comme a est bilinéaire, continue et coercive, A est un opérateur linéaire sur H , et on a pour tout $u \in H$.

Pour prouver qu'il existe un élément $Au \in H$ tel que $\forall v \in H$, $a(u, v) = (Au, v)$, on utilise le théorème de Riesz-Fréchet.

Fixons $u \in H$. Alors $v \mapsto a(u, v)$ est une forme linéaire continue sur H (car a est bilinéaire continue), donc il existe un unique élément $Au \in H$ tel que

$$\forall v \in H, \quad a(u, v) = (Au, v)$$

Cette application $u \mapsto Au$ définit un opérateur linéaire sur H , appelé l'opérateur associé à a .

$$|Au| \leq C|u|, \quad (Au, u) \geq \alpha|u|^2$$

On cherche alors $u \in K$ tel que

$$\forall v \in K, \quad (Au, v - u) \geq (f, v - u).$$

Soit $\rho > 0$. D'après le théorème 7.1.3, l'inégalité

$$\forall v \in K, \quad (\rho(f - Au) + u - u, v - u) \leq 0$$

est équivalente au fait que

$$u = Su \text{ où } Su = P_K(\rho f - \rho Au + u).$$

Le théorème 7.1.3 affirme que $u = P_K(f)$ si et seulement si $\forall v \in K, (f - u, v - u) \leq 0$. Dans la démonstration du théorème de Stampacchia, on cherche u tel que

$$\forall v \in K, (Au, v - u) \geq (f, v - u) \iff (f - Au, v - u) \leq 0$$

En introduisant un paramètre $\rho > 0$ et en ajoutant $u - u = 0$ à gauche, cette inégalité devient

$$(\rho(f - Au) + u - u, v - u) \leq 0 \iff (\rho(f - Au), v - u) \leq 0$$

Ceci est exactement la condition de projection dans K du point $\rho f - \rho Au + u$. Donc, d'après le théorème 8.1.3, cela revient à écrire

$$u = P_K(\rho f - \rho Au + u) =: Su$$

Or

$$\begin{aligned} |Su_1 - Su_2|^2 &\leq |(u_1 - u_2) - \rho A(u_1 - u_2)|^2 \\ &\leq |u_1 - u_2|^2 - 2\rho(A(u_1 - u_2), u_1 - u_2) + \rho^2(A(u_1 - u_2), A(u_1 - u_2)) \\ &\leq |u_1 - u_2|^2 (1 - 2\rho\alpha + \rho^2 C^2) \end{aligned}$$

Justifions les inégalités :

$$\begin{aligned} |Su_1 - Su_2|^2 &= |P_K(\rho f - \rho Au_1 + u_1) - P_K(\rho f - \rho Au_2 + u_2)|^2 \\ &\leq |\rho f - \rho Au_1 + u_1 - (\rho f - \rho Au_2 + u_2)|^2 \quad (\text{par la propriété 1-Lipschitz de la projection}) \\ &= |(u_1 - u_2) - \rho A(u_1 - u_2)|^2 \end{aligned}$$

On développe alors ce carré scalaire :

$$\begin{aligned} |(u_1 - u_2) - \rho A(u_1 - u_2)|^2 &= |u_1 - u_2|^2 - 2\rho(A(u_1 - u_2), u_1 - u_2) + \rho^2|A(u_1 - u_2)|^2 \\ &\leq |u_1 - u_2|^2 - 2\rho\alpha|u_1 - u_2|^2 + \rho^2 C^2|u_1 - u_2|^2 \\ &= |u_1 - u_2|^2 (1 - 2\rho\alpha + \rho^2 C^2) \end{aligned}$$

où on a utilisé la coercivité $(Au, u) \geq \alpha\|u\|^2$ et la continuité $\|Au\| \leq C\|u\|$ de l'opérateur A . En choisissant $\rho < 2\alpha/C^2$, on obtient que S est contractant. D'après le théorème de Picard, S a donc un unique point fixe, qui est la solution de l'équation

$$\forall v \in K, \quad a(u, v - u) \geq \langle \varphi, v - u \rangle$$

Dans le cas où la forme bilinéaire est symétrique, a définit un nouveau produit scalaire sur H . De plus, comme a est continue et coercive, les deux normes sont équivalentes. En appliquant le théorème de Riesz-Fréchet, on obtient alors l'existence d'un unique $g \in H$ tel que

$$\forall v \in H, \quad \langle \varphi, v \rangle = a(g, v).$$

L'équation

$$\forall v \in K, \quad a(u, v - u) \geq \langle \varphi, v - u \rangle$$

est alors équivalente à $u = \tilde{P}_K g$ où $\tilde{P}_K$ est la projection au sens du produit scalaire défini par a . D'après le théorème 8.1.3, u est alors la solution du problème de minimisation

$$a(g - u, g - u) = \min_{v \in K} a(g - v, g - v)$$

ou de façon équivalente

$$\frac{1}{2}a(u, u) - a(g, u) = \min_{v \in K} \left(\frac{1}{2}a(v, v) - a(g, v) \right)$$

On part du premier problème de minimisation :

$$a(g - u, g - u) = \min_{v \in K} a(g - v, g - v).$$

Développons le terme $a(g - v, g - v)$ en utilisant la bilinéarité symétrique de $a(\cdot, \cdot)$:

$$a(g - v, g - v) = a(g, g) - 2a(g, v) + a(v, v).$$

Comme $a(g, g)$ est constant (indépendant de v), minimiser $a(g - v, g - v)$ revient à minimiser :

$$a(v, v) - 2a(g, v).$$

Cela équivaut à minimiser la fonctionnelle suivante :

$$\frac{1}{2}a(v, v) - a(g, v),$$

car le facteur $\frac{1}{2}$ devant $a(v, v)$ n'affecte pas la position du minimum. Ainsi, on a :

$$\frac{1}{2}a(u, u) - a(g, u) = \min_{v \in K} \left(\frac{1}{2}a(v, v) - a(g, v) \right).$$

On conclut en utilisant la définition de g . □

Le théorème de Stampacchia est le point de départ de la théorie des inéquations variationnelles ; cette théorie a de nombreuses applications en mécanique, en physique et en contrôle optimal.

Pour la résolution des équations aux dérivées partielles elliptiques par exemple, l'approche variationnelle est la suivante :

- on définit la notion de solution faible en utilisant les espaces de Sobolev ;
- on établit l'existence et l'unicité d'une solution faible en appliquant un corollaire du théorème de Stampacchia, dit théorème de Lax-Milgram ;
- on montre alors des estimations de régularité pour la solution faible et on en déduit que c'est en fait une solution classique.

Corollaire 7.2.4. (*Théorème de Lax-Milgram*). - Soient H un espace de Hilbert, et $a : H \times H \rightarrow \mathbb{R}$ une forme bilinéaire continue coercive.

Pour tout $\varphi \in H^*$, il existe un unique $u \in H$ tel que

$$\forall v \in H, \quad a(u, v) = \langle \varphi, v \rangle. \quad (7.2.1)$$

De plus, si a est symétrique, u est caractérisé par la propriété

$$\frac{1}{2}a(u, u) - \langle \varphi, u \rangle = \min_{v \in H} \left(\frac{1}{2}a(v, v) - \langle \varphi, v \rangle \right) \quad (7.2.2)$$

Démonstration. Puisque a est une forme bilinéaire continue et coercive sur l'espace de Hilbert H , on peut lui associer un opérateur linéaire $A : H \rightarrow H^*$ défini par :

$$(Au)(v) = a(u, v) \quad \text{pour tout } v \in H.$$

Alors A est un isomorphisme de H dans H^* :

- A est linéaire par bilinéarité de a ;
- A est continu car :

$$\|Au\|_{H^*} = \sup_{\|v\|_H=1} |a(u, v)| \leq M\|u\|_H;$$

- A est injectif : si $Au = 0$, alors $a(u, v) = 0$ pour tout $v \in H$, donc en particulier $a(u, u) = 0$, ce qui contredit la coercivité sauf si $u = 0$;
- A est surjectif par le théorème de Banach sur l'isomorphisme (puisque H est de Hilbert, H^* est aussi de Hilbert).

Par conséquent, pour toute forme linéaire continue $L \in H^*$, il existe un unique $u \in H$ tel que $Au = L$, c'est-à-dire :

$$a(u, v) = L(v) \quad \text{pour tout } v \in H.$$

Ce qui conclut la démonstration. □

La relation entre l'équation (7.2.1) et le problème de minimisation (7.2.2) est l'expression mathématique du principe de moindre action.

7.3 Bases hilbertiennes

Définition 7.3.1. Soient H un espace de Hilbert, et $(E_n)_{n \geq 1}$ une suite de sous-espaces fermés de H . On dit que H est somme hilbertienne des E_n ($H = \oplus_n E_n$) si

- les E_n sont deux-à-deux orthogonaux

$$\forall u \in E_m, \forall v \in E_n \text{ avec } n \neq m, \quad (u, v) = 0;$$

- l'espace vectoriel engendré par les E_n est dense dans H .

Proposition 7.3.2. (Identité de Bessel-Parseval). - Soient H un espace de Hilbert, et $(E_n)_{n \geq 1}$ une suite de sous-espaces fermés de H tels que $H = \oplus_n E_n$.

Soient $u \in H$ et $u_n = P_{E_n} u$. Alors

$$u = \sum_{n=1}^{\infty} u_n \text{ et } \|u\|^2 = \sum_{n=1}^{\infty} \|u_n\|^2.$$

Démonstration. - On pose

$$S_k = \sum_{n=1}^k P_{E_n}$$

de sorte que S_k est un opérateur linéaire continu de H dans H .

Pour tout $u \in H$, le théorème 7.1.3 montre que

$$(u, P_{E_n} u) = |u_n|^2$$

et donc par sommation

$$(u, S_k u) = \sum_{n=1}^k |u_n|^2 = |S_k u|^2. \quad (7.3.1)$$

On en déduit, par l'inégalité de Cauchy-Schwarz, que

$$|S_k u| \leq |u|.$$

Soit F l'espace vectoriel engendré par les (E_n) . Pour tout $\varepsilon > 0$, il existe $\bar{u} \in F$ tel que

$$|u - \bar{u}| \leq \varepsilon.$$

Pour k assez grand, on a $\bar{u} = S_k \bar{u}$, d'où l'on déduit que

$$|u - S_k u| \leq |u - \bar{u}| + |S_k \bar{u} - S_k u| \leq 2|u - \bar{u}| \leq 2\varepsilon.$$

On a donc $u = \lim_{k \rightarrow \infty} \sum_{n=1}^k u_n$. En passant à la limite dans (7.3.1), on obtient l'égalité de Bessel-Parseval. $\square$

Définition 7.3.3. Soit H un espace de Hilbert. On appelle base hilbertienne de H une suite $(e_n)_{n \geq 1}$ telle que

$$-\forall m, n, \quad (e_n, e_m) = \delta_{n,m};$$

— l'espace vectoriel engendré par les (e_n) est dense dans H .

D'après la proposition précédente, tout élément $u \in H$ peut alors se décomposer

$$u = \sum_{n=1}^{\infty} (e_n, u) e_n \text{ avec } \sum_{n=1}^{\infty} |(u, e_n)|^2 = |u|^2$$

Théorème 7.3.4. Tout espace de Hilbert séparable admet une base hilbertienne.

Démonstration. - Soit (u_n) un ensemble dénombrable dense de H . On note F_k l'espace vectoriel engendré par $(u_n)_{1 \leq n \leq k}$. Les (F_k) forment alors une suite croissante de sous-espaces de dimension finie telle que $\bigcup_{k=1}^{\infty} F_k$ est dense dans H .

Le procédé d'orthonormalisation de Schmidt permet alors de construire une base hilbertienne.

Espaces de Sobolev

Les espaces de Sobolev constituent un outil fondamental pour l'étude des problèmes d'analyse non linéaire [1]. Ce sont en effet des espaces de Banach, plus faciles à manipuler que des distributions quelconques.

Ils sont construits à partir des espaces de Lebesgue $L^p(\Omega)$. Pour $s \in \mathbb{N}$,

- l'espace $W^{s,p}(\Omega)$ est le sous-espace de $L^p(\Omega)$ constitué des fonctions dont les dérivées jusqu'à l'ordre s sont dans $L^p(\Omega)$;
- l'espace $W^{-s,p}(\Omega)$ est un espace de distributions d'ordre s , défini par dualité.

En particulier, sous certaines conditions sur s et p , on va pouvoir définir des produits, et plus généralement des quantités non linéaires.

Dans cette partie du cours, on se limitera à l'étude des espaces de Sobolev construits sur $L^2(\mathbb{R}^N)$, ce qui permet d'avoir une caractérisation très simple en termes de la transformée de Fourier.

Motivation

Considérons le problème suivant. Étant donné $f \in C([a, b])$ trouver une fonction $u(x)$ vérifiant

$$\begin{cases} -u'' + u = f & \text{sur } [a, b] \\ u(a) = u(b) = 0 \end{cases} \quad (1)$$

Une solution classique - ou solution forte - du problème (1) est une fonction de classe C^2 sur $[a, b]$ vérifiant (1) au sens usuel. Bien entendu (1) peut être résolu explicitement par un calcul très simple, mais nous ignorerons cet aspect des choses afin d'illustrer la méthode sur cet exemple élémentaire.

On multiplie (1) par $\varphi \in C^1([a, b])$ et on intègre par parties ; il vient

$$\int_a^b u' \varphi' + \int_a^b u \varphi = \int_a^b f \varphi \quad \forall \varphi \in C^1([a, b]), \quad \varphi(a) = \varphi(b) = 0. \quad (2)$$

On notera que (2) a un sens dès que $u \in C^1([a, b])$ (contrairement à (1) qui suppose u deux fois dérivable) ; en fait il suffirait même d'avoir $u, u' \in L^1(a, b), u'$ en un sens à préciser.

Disons (provisoirement) qu'une fonction u de classe C^1 qui vérifie (2) est une solution faible de (1).

Le programme suivant décrit les grandes lignes de l'approche variationnelle en théorie des équations aux dérivées partielles :

Etape A. - On précise la notion de solution faible ; celle-ci fait intervenir les espaces de Sobolev qui sont les outils de base.

Etape B. - On établit l'existence et l'unicité d'une solution faible par la méthode variationnelle, via le théorème de Lax-Milgram.

Etape C. - On prouve que la solution faible est de classe C^2 (par exemple) : c'est un résultat de régularité.

Etape D. - Retour aux solutions classiques. On montre qu'une solution faible de classe C^2 est une solution classique.

L'étape D est très simple. En effet supposons que $u \in C^2([a, b])$, $u(a) = u(b) = 0$ et u vérifie (2). En intégrant (2) par parties on obtient

$$\int_a^b (-u'' + u - f) \varphi = 0 \quad \forall \varphi \in C^1([a, b]), \quad \varphi(a) = \varphi(b) = 0$$

et a fortiori

$$\int_a^b (-u'' + u - f) \varphi = 0 \quad \forall \varphi \in C_c^1([a, b]).$$

Or $C_c^1([a, b])$ est dense dans $L^2(a, b)$ et donc $-u'' + u = f$ p.p. (en fait partout puisque $u \in C^2$).

Pur bien approfondir le contenu de ce chapitre, je recommande les chapitres 7 et 8 de l'excellent livre de Haim Brezis [3].

8.1 Définition et propriétés topologiques

Définition 8.1.1. Soit $s \in \mathbb{R}$. La distribution tempérée u appartient à l'espace de Sobolev $H^s(\mathbb{R}^N)$ si et seulement si

$$\mathcal{F}u \in L_{loc}^2(\mathbb{R}^N) \quad \text{et} \quad \mathcal{F}u \in L^2(\mathbb{R}^N, (1 + |\xi|^2)^s d\xi).$$

Pour tout $s \in \mathbb{R}$, l'espace de Sobolev $H^s(\mathbb{R}^N)$, muni de la norme

$$\|u\|_{H^s}^2 = \int |\mathcal{F}u(\xi)|^2 (1 + |\xi|^2)^s d\xi$$

est un espace de Banach (et même un espace de Hilbert comme on le verra dans la suite). Le seul point à vérifier est la complétude. Par définition de la norme, si (u_n) est une suite de Cauchy de $H^s(\mathbb{R}^N)$, la suite $(\mathcal{F}u_n)$ est une suite de Cauchy dans $L^2(\mathbb{R}^N, (1 + |\xi|^2)^s d\xi)$, et il existe donc $\tilde{u} \in L^2(\mathbb{R}^N, (1 + |\xi|^2)^s d\xi)$ tel que

$$\lim_{N \rightarrow \infty} \|\mathcal{F}u_n - \tilde{u}\|_{L^2(\mathbb{R}^N, (1 + |\xi|^2)^s d\xi)} = 0. \quad (8.1.1)$$

En particulier, $(\mathcal{F}u_n)$ converge vers $\tilde{u}$ dans $\mathcal{S}'(\mathbb{R}^N)$. Comme la transformée de Fourier est un isomorphisme de $\mathcal{S}'(\mathbb{R}^N)$, (u_n) tend vers $\mathcal{F}^{-1}\tilde{u}$ dans $\mathcal{S}'(\mathbb{R}^N)$ et donc dans $H^s(\mathbb{R}^N)$ par (8.1.1).

On note que la transformée de Fourier est un isomorphisme entre $H^s(\mathbb{R}^N)$ et $L^2(\mathbb{R}^N, (1 + |\xi|^2)^s d\xi)$.

Propriété 8.1.2. – Pour tout $s \in \mathbb{N}$, l'espace $H^s(\mathbb{R}^N)$ est l'espace des fonctions de $L^2(\mathbb{R}^N)$ dont toutes les dérivées au sens des distributions jusque'à l'ordre s sont dans L^2

$$\|u\|_{H^s}^2 \sim \sum_{|\alpha| \leq s} \|\partial^\alpha u\|_{L^2}^2$$

— Pour tout $s \in \mathbb{R}$, $\mathcal{S}(\mathbb{R}^N)$ est dense dans $H^s(\mathbb{R}^N)$.

Démonstration. – Comme la transformée de Fourier est, à constante près, une isométrie de $L^2(\mathbb{R}^N)$, on a

$$\xi^\alpha \mathcal{F}u \in L^2(\mathbb{R}^N) \quad \text{si et seulement si} \quad \partial^\alpha u \in L^2(\mathbb{R}^N)$$

De plus, pour tout s fini, il existe $C_s > 0$ tel que

$$\forall \xi \in \mathbb{R}^N, \quad C_s^{-1} \left(1 + \sum_{|\alpha| \leq s} |\xi^\alpha|^2 \right) \leq (1 + |\xi|^2)^s \leq C_s \left(1 + \sum_{|\alpha| \leq s} |\xi^\alpha|^2 \right)$$

On en déduit que $u \in H^s(\mathbb{R}^N)$ si et seulement si $\partial^\alpha u \in L^2(\mathbb{R}^N)$ pour tout α tel que $|\alpha| \leq s$, et on a l'équivalence des normes.

Pour tout $u \in H^s(\mathbb{R}^N)$, $(1 + |\xi|^2)^{s/2} \mathcal{F}u$ est dans $L^2(\mathbb{R}^N)$ et peut donc être approché (par convolution) par une suite de fonctions de $\mathcal{S}(\mathbb{R}^N)$. Par stabilité de $\mathcal{S}(\mathbb{R}^N)$ par transformation de Fourier et par multiplication par les fonctions à croissance lente, on en déduit que u peut être approché par une suite de fonctions de $\mathcal{S}(\mathbb{R}^N)$. $\square$

Pour $s \in \mathbb{R}$, le dual de $H^s(\mathbb{R}^N)$ s'identifie avec $H^{-s}(\mathbb{R}^N)$.

Théorème 8.1.3. – Si $u \in (H^s(\mathbb{R}^N))'$, il existe un unique $v \in H^{-s}(\mathbb{R}^N)$ tel que

$$\forall \varphi \in H^s(\mathbb{R}^N), \quad \langle u, \varphi \rangle = \int \mathcal{F}v \overline{\mathcal{F}\varphi}(\xi) d\xi$$

Démonstration. – Le point important de la preuve est la formule d'inversion de Fourier qui assure que

$$\forall \varphi \in \mathcal{S}(\mathbb{R}^N), \quad \langle u, \varphi \rangle = \langle \mathcal{F}u, \mathcal{F}\varphi \rangle$$

On peut alors montrer que $(1 + |\xi|^2)^{-s/2} \mathcal{F}u$ définit une forme linéaire continue sur $L^2(\mathbb{R}^N)$. En effet, pour toute fonction $\psi \in L^2(\mathbb{R}^N)$, la formule précédente donne

$$\begin{aligned} \left| \left\langle (1 + |\xi|^2)^{-s/2} \mathcal{F}u, \mathcal{F}\psi \right\rangle \right| &= \left| \left\langle \mathcal{F}u, (1 + |\xi|^2)^{s/2} \mathcal{F}\psi \right\rangle \right| \leq \|u\|_{(H^s)'} \left\| \mathcal{F}^{-1} \left((1 + |\xi|^2)^{s/2} \psi \right) \right\|_{H^s} \\ &\leq C \|u\|_{(H^s)'} \|\mathcal{F}\psi\|_{L^2} \leq C \|u\|_{(H^s)'} \|\psi\|_{L^2} \end{aligned}$$

Le théorème de représentation de $L^2(\mathbb{R}^N)$ montre alors qu'il existe un unique $\tilde{w} \in L^2(\mathbb{R}^N)$ tel que

$$\left\langle (1 + |\xi|^2)^{-s/2} \mathcal{F}u, \psi \right\rangle = \int_{\mathbb{R}^N} \tilde{w}(\xi) \psi(\xi) d\xi.$$

On en déduit que

$$\langle \mathcal{F}u, \psi \rangle = \int_{\mathbb{R}^N} (1 + |\xi|^2)^{s/2} \tilde{w}(\xi) \psi(\xi) d\xi.$$

En posant $v = \mathcal{F}^{-1} \left((1 + |\xi|^2)^{s/2} \tilde{w} \right)$, on obtient la représentation attendue. $\square$

8.2 Injections de Sobolev

L'objectif de cette section est de montrer des propriétés d'inclusion des espaces de Sobolev dans les espaces de Lebesgue. On va montrer en quelque sorte qu'on peut gagner de l'intégrabilité quitte à perdre de la régularité.

Théorème 8.2.1. (*Théorème de l'injection de Sobolev*) Si $s > N/2$, l'espace $H^s(\mathbb{R}^N)$ s'injecte continûment dans $C(\mathbb{R}^N)$

$$\forall \psi \in H^s(\mathbb{R}^N), \quad \|\psi\|_{C(\mathbb{R}^N)} \leq C \|\psi\|_{H^s(\mathbb{R}^N)}$$

Si $s < N/2$, l'espace $H^s(\mathbb{R}^N)$ s'injecte continûment dans $L^p(\mathbb{R}^N)$ avec $p = 2N/(N - 2s)$

$$\forall \psi \in H^s(\mathbb{R}^N), \quad \|\psi\|_{L^p(\mathbb{R}^N)} \leq C \|\psi\|_{H^s(\mathbb{R}^N)}$$

Démonstration. - Le premier point du théorème est très facile à prouver : il repose sur le fait que la transformée de Fourier (inverse) d'une fonction $\mathcal{F}\varphi \in L^1(\mathbb{R}^N)$ est une fonction continue, avec

$$\|\varphi\|_{\infty} \leq (2\pi)^{-N} \|\mathcal{F}\varphi\|_{L^1(\mathbb{R}^N)}.$$

Si $u \in H^s(\mathbb{R}^N)$ avec $s > N/2$, l'inégalité de Cauchy-Schwarz donne

$$\|\mathcal{F}u\|_{L^1(\mathbb{R}^N)} \leq \left\| (1 + |\xi|^2)^{s/2} \mathcal{F}u \right\|_{L^2(\mathbb{R}^N)} \left\| (1 + |\xi|^2)^{-s/2} \right\|_{L^2(\mathbb{R}^N)} \leq C \|u\|_{H^s(\mathbb{R}^N)},$$

ce qui montre que u est continue et qu'on a

$$\|u\|_{C(\mathbb{R}^N)} \leq C \|u\|_{H^s(\mathbb{R}^N)}$$

Le second point de la preuve est plus délicat. Une façon intuitive de comprendre le choix de l'indice p est l'argument suivant d'invariance par changement d'échelle. Soit u une fonction définie sur $\mathbb{R}^N$ et u_λ la fonction définie par $u_\lambda(x) = u(\lambda x)$. On a

$$\|u_\lambda\|_{L^p(\mathbb{R}^N)} = \lambda^{-N/p} \|u\|_{L^p(\mathbb{R}^N)}$$

et

$$\int |\xi|^{2s} |\mathcal{F}u_\lambda(\xi)|^2 d\xi = \lambda^{-2N} \int |\xi|^{2s} |\mathcal{F}u(\lambda^{-1}\xi)|^2 d\xi = \lambda^{-N+2s} \|u\|_{\dot{H}^s(\mathbb{R}^N)}^2.$$

Les normes L^p et $\dot{H}^s$ ont le même scaling, ce qui signifie qu'elles ont le même comportement vis-à-vis des changements d'unité. Cela a donc du sens de les comparer.

On se donne $u \in \mathcal{S}(\mathbb{R}^N)$. Quitte à multiplier u par une constante, on peut supposer sans perte de généralité que $\|u\|_{\dot{H}^s} = 1$. D'après le théorème de Fubini, on a

$$\|u\|_{L^p(\mathbb{R}^N)}^p = p \int_0^\infty \lambda^{p-1} \mu(|u| > \lambda) d\lambda.$$

On décompose alors u en une composante basse fréquence et une composante haute fréquence :

$$u_1 = \mathcal{F}^{-1}(\mathbf{1}_{|\xi| \leq A} \mathcal{F}u), \quad u_2 = \mathcal{F}^{-1}(\mathbf{1}_{|\xi| > A} \mathcal{F}u).$$

Comme la transformée de Fourier de u_1 est à support compact, la fonction u_1 est bornée, et plus précisément on a

$$\begin{aligned} \|u_1\|_\infty &\leq (2\pi)^{-N} \|\mathcal{F}u_1\|_{L^1(\mathbb{R}^N)} \leq (2\pi)^{-N} \int_{B(0,A)} |\xi|^{-s} |\xi|^s |\mathcal{F}u(\xi)| d\xi \\ &\leq C_{s,N} A^{N/2-s} \|u\|_{H^s(\mathbb{R}^N)} \leq C_{s,N} A^{N/2-s}. \end{aligned} \quad (8.2.1)$$

L'inégalité triangulaire montre que

$$\{|u| > \lambda\} \subset \{|u_1| > \lambda/2\} \cup \{|u_2| > \lambda/2\}.$$

Si on choisit

$$A = (\lambda/4C_{s,N})^{2/(N-2s)}$$

l'estimation (8.2.1) montre que

$$\mu(\{|u_1| > \lambda/2\}) = 0.$$

On a alors

$$\|u\|_{L^p(\mathbb{R}^N)}^p \leq p \int_0^\infty \lambda^{p-1} \mu(|u_2| > \lambda/2) d\lambda.$$

L'inégalité de Bienaymé-Tchebichev donne

$$\mu(|u_2| > \lambda/2) \leq \frac{4}{\lambda^2} \|u_2\|_{L^2(\mathbb{R}^N)}^2$$

de sorte que

$$\|u\|_{L^p(\mathbb{R}^N)}^p \leq 4p \int_0^\infty \lambda^{p-3} \|u_2\|_{L^2(\mathbb{R}^N)}^2 d\lambda.$$

Comme la transformée de Fourier est, à constante près, une isométrie de L^2 , on a

$$\|u_2\|_{L^2(\mathbb{R}^N)}^2 = (2\pi)^{-N} \int \mathbf{1}_{|\xi| \geq A} |\mathcal{F}u(\xi)|^2 d\xi.$$

Par définition de A ,

$$|\xi| \leq A \text{ si et seulement si } \lambda \leq C|\xi|^{N/2-s}.$$

En appliquant le théorème de Fubini, on obtient alors

$$\begin{aligned} \|u\|_{L^p(\mathbb{R}^N)}^p &\leq \frac{4p}{(2\pi)^N} \int_0^\infty \lambda^{p-3} \int \mathbf{1}_{|\xi| \geq A} |\mathcal{F}u(\xi)|^2 d\xi d\lambda \\ &\leq C_{p,N} \int |\xi|^{(N-2s)(p-2)/2} |\mathcal{F}u(\xi)|^2 d\xi. \end{aligned}$$

Comme $(N/2 - s)(p - 2) = 2s$, ceci prouve le théorème, par densité de $\mathcal{S}(\mathbb{R}^N)$ dans $H^s(\mathbb{R}^N)$. $\square$

Par dualité, on obtient que l'espace $L^p(\mathbb{R}^N)$ pour $p \in]1, 2[$ s'injecte continûment dans $H^{-s}(\mathbb{R}^N)$ avec $s = N\left(\frac{1}{p} - \frac{1}{2}\right)$. On a en effet

$$\begin{aligned} \|u\|_{H^{-s}(\mathbb{R}^N)} &\leq C \sup_{\|\psi\|_{H^s} \leq 1} |\langle u, \psi \rangle| \\ &\leq C' \sup_{\|\psi\|_{L^{p'}} \leq 1} |\langle u, \psi \rangle| \end{aligned}$$

par injection de $L^{p'}(\mathbb{R}^N)$ dans $H^s(\mathbb{R}^N)$. On conclut alors avec l'inégalité de Hölder.

8.3 Restriction, prolongement et traces

Pour étendre les résultats des paragraphes précédents au cas de domaines plus généraux, on doit comprendre ce qui se passe au voisinage des bords. Pour simplifier, on se restreindra ici au cas où le bord est un hyperplan. Le cas des domaines réguliers s'en déduit en paramétrant le bord localement par des graphes. Dans ce qui suit, on note $\mathbb{R}_+^N$ le demi-espace ouvert de $\mathbb{R}^N$

$$\mathbb{R}_+^N = \{x = (x_1, \dots, x_N) \in \mathbb{R}^N / x_N > 0\}$$

Proposition 8.3.1. - *L'application de restriction*

$$f \in H^1(\mathbb{R}^N) \mapsto f|_{x_N > 0} \in H^1(\mathbb{R}_+^N)$$

est surjective, et il existe une application linéaire $P : H^1(\mathbb{R}_+^N) \rightarrow H^1(\mathbb{R}^N)$ de prolongement telle que

$$\|P\| \leq \sqrt{2}, \quad P(f)|_{x_N > 0} = f.$$

Démonstration. - Pour $f \in L^2(\mathbb{R}_+^N)$, on définit la fonction $P(f) \in L^2(\mathbb{R}^N)$ par symétrie en posant

$$P(f)(x', x_N) = f(x', |x_N|).$$

Pour $g \in C^\infty(\mathbb{R}^N)$ vérifiant $g|_{x_N > 0} \in H^1(\mathbb{R}_+^N)$, la formule des sauts implique que les dérivées distributions de $P(f)$ vérifient

$$\partial_j P(f) = P(\partial_j f) \text{ si } j < N, \text{ et } \partial_N P(f) = \text{sign}(x_N) P(\partial_N f)$$

donc $P(f) \in H^1(\mathbb{R}^N)$ et

$$\|Pf\|_{H^1(\mathbb{R}^N)} \leq \sqrt{2} \|f\|_{H^1(\mathbb{R}_+^N)}. \quad (8.3.1)$$

Pour $f \in H^1(\mathbb{R}_+^N)$ arbitraire, on commence par montrer qu'il existe une suite (φ_n) de fonctions $C^\infty(\mathbb{R}^N)$ telles que $\varphi_n|_{x_N > 0} \in H^1(\mathbb{R}_+^N)$ et

$$\lim_{n \rightarrow \infty} \|f - \varphi_n\|_{H^1(\mathbb{R}_+^N)} = 0$$

Pour cela, on utilise une approximation de l'identité (ρ_n) définie par

$$\forall x \in \mathbb{R}^N, \quad \rho_n(x) = n^N \rho(nx)$$

où $\rho \in C^\infty(\mathbb{R}^N, \mathbb{R}^+)$ est à support dans la boule unité, d'intégrale 1 et telle que

$$\text{Supp}(\rho) \subset \left\{ -\frac{3}{4} \leq x_N \leq -\frac{1}{4} \right\}.$$

On pose alors $\varphi_n = \rho_n * (f \mathbf{1}_{x_N > 0})$, de sorte que les fonctions φ_n appartiennent à $C^\infty(\mathbb{R}^N)$ et convergent vers $f \mathbf{1}_{x_N > 0}$ dans $L^2(\mathbb{R}^N)$ quand $n \rightarrow \infty$. De plus, on a

$$\forall x \in \mathbb{R}^N, \quad \partial_j \varphi_n(x) = \langle \partial_j (f \mathbf{1}_{x_N > 0}), \tau_x \hat{\rho}_n \rangle$$

Pour $x_N > 0$, comme le support de la fonction test est inclus dans $\{y/y_N \geq x_N + \frac{1}{4n}\}$, on a

$$\forall x \in \mathbb{R}^N, \quad \partial_j \varphi_n(x) = \langle (\partial_j f) \mathbf{1}_{x_N > 0}, \tau_x \hat{\rho}_n \rangle = ((\partial_j f) \mathbf{1}_{x_N > 0}) * \rho_n.$$

Comme $((\partial_j f) \mathbf{1}_{x_N > 0}) * \rho_n$ converge vers $(\partial_j f) \mathbf{1}_{x_N > 0}$ dans $L^2(\mathbb{R}^N)$, on a le résultat d'approximation annoncé.

L'estimation de continuité (7.3.1) montre alors que la suite $P\varphi_n|_{x_N > 0}$ est de Cauchy dans $H^1(\mathbb{R}^N)$, donc converge vers $h \in H^1(\mathbb{R}^N)$ vérifiant

$$\|h\|_{H^1(\mathbb{R}^N)} \leq \sqrt{2} \|f\|_{H^1(\mathbb{R}_+^N)}.$$

Comme $P\varphi_n|_{x_N > 0}$ converge vers $P\varphi$ dans $L^2(\mathbb{R}^N)_+$, on a bien $h = Pf$, ce qui conclut la preuve. $\square$

Le théorème d'injection de Sobolev montre qu'en dimension 1, l'espace $H^s(\mathbb{R})$ pour $s > \frac{1}{2}$ s'injecte dans l'ensemble des fonctions continues : on peut donc définir ponctuellement la valeur d'une fonction $H^s(\mathbb{R})$. Il est alors naturel de se demander si en dimension quelconque $N \geq 1$, on peut définir la section d'une fonction $H^s(\mathbb{R}^N)$ pour $s > \frac{1}{2}$ sur des ensembles de codimension 1, typiquement des hyperplans.

On introduit donc l'espace $C_b^0(\mathbb{R}, H^\sigma(\mathbb{R}^{N-1}))$ des fonctions continues bornées sur $\mathbb{R}$ à valeurs dans $H^\sigma(\mathbb{R}^{N-1})$. Muni de la norme

$$\|u\|_{L^\infty(H^\sigma)} = \sup_{y \in \mathbb{R}} \|u(y)\|_{H^\sigma(\mathbb{R}^{N-1})}$$

c'est clairement un espace de Banach. De plus, $C_b^0(\mathbb{R}, H^\sigma(\mathbb{R}^{N-1}))$ s'injecte dans $\mathcal{S}'(\mathbb{R}^N)$ en posant

$$\forall \varphi \in C_c^\infty(\mathbb{R}^N), \quad \langle u, \varphi \rangle = \int_{-\infty}^{\infty} \langle u(x_N), \varphi(\cdot, x_N) \rangle dx_N.$$

On a en effet pour $q = \max(-\sigma, 0)$ et $p > \frac{N-1}{2}$,

$$\begin{aligned} \left| \int_{-\infty}^{\infty} \langle u(x_N), \varphi(\cdot, x_N) \rangle dx_N \right| &\leq \|u\|_{L^\infty(H^\sigma)} \int_{-\infty}^{\infty} \|\varphi(\cdot, x_N)\|_{H^{-\sigma}(\mathbb{R}^{N-1})} dx_N \\ &\leq C \|u\|_{L^\infty(H^\sigma)} \int_{-\infty}^{\infty} \sup_{|\alpha| \leq p, |\beta| \leq q} \|x^\alpha \partial^\beta \varphi(\cdot, x_N)\|_{L^\infty(\mathbb{R}^{N-1})} dx_N \\ &\leq \tilde{C} \|u\|_{L^\infty(H^\sigma)} \sup_{|\alpha| \leq p+2, |\beta| \leq q} \|x^\alpha \partial^\beta \varphi\|_\infty. \end{aligned}$$

L'injectivité s'obtient en considérant par exemple la famille de fonctions test

$$\varphi_{n,z}(x) = \psi(x') \rho_n(z - x_N) \text{ où } \psi \in C_c^\infty(\mathbb{R}^{N-1}) \text{ et } (\rho_n) \text{ approximation de l'unité}$$

de sorte que

$$(\langle u, \varphi_{n,z} \rangle = 0 \quad \forall n, z, \psi) \Rightarrow (\langle u(z), \psi \rangle = 0 \quad \forall z, \psi) \Rightarrow (u(z) = 0 \quad \forall z).$$

En utilisant cette identification, on obtient le résultat attendu sur les sections :

Proposition 8.3.2. - Soient $s > \frac{1}{2}$ et $f \in H^s(\mathbb{R}^N)$. Il existe un unique $u \in C_b^0(\mathbb{R}, H^{s-1/2}(\mathbb{R}^{N-1}))$ tel que

$$\forall \varphi \in C_c^\infty(\mathbb{R}^N), \quad \langle f, \varphi \rangle = \int_{-\infty}^{\infty} \langle u(x_N), \varphi(\cdot, x_N) \rangle$$

et on a

$$\|u\|_{L^\infty(\mathbb{R}, H^{s-1/2}(\mathbb{R}^{N-1}))} \leq C \|f\|_{H^s(\mathbb{R}^N)}.$$

Démonstration. - Soit $f \in H^s(\mathbb{R}^N)$. Par définition,

$$a(\xi) = (1 + |\xi|^2)^{s/2} |\mathcal{F}f(\xi)| \text{ appartient à } L^2(\mathbb{R}^N).$$

Le théorème de Fubini montre alors qu'il existe une partie mesurable $Z \subset \mathbb{R}^{N-1}$ de mesure nulle telle que

$$\forall \xi' \notin Z, \quad \xi_N \mapsto a(\xi', \xi_N) \text{ appartient à } L^2(\mathbb{R}).$$

Comme la fonction $\xi_N \mapsto (1 + |\xi'|^2 + |\xi_N|^2)^{-s/2}$ appartient à $L^2(\mathbb{R})$,

$$\forall \xi' \notin Z, \quad \xi_N \mapsto \mathcal{F}f(\xi', \xi_N) \text{ appartient à } L^1(\mathbb{R})$$

On peut alors définir la transformation de Fourier partielle de $\mathcal{F}f$ par rapport à la variable ξ_N . On pose donc

$$g(x_N, \xi') = \int \exp(ix_N \xi_N) \mathcal{F}f(\xi', \xi_N) d\xi_N$$

Pour $\xi' \notin Z$,

$$\begin{aligned} |g(x_N, \xi')|^2 &\leq \left(\int |\mathcal{F}f(\xi', \xi_N)| d\xi_N \right)^2 \leq \left(\int (1 + |\xi'|^2 + |\xi_N|^2)^{-s} d\xi_N \right) \left(\int |a(\xi', \xi_N)|^2 d\xi_N \right) \\ &\leq C (1 + |\xi'|^2)^{-s+\frac{1}{2}} \int |a(\xi', \xi_N)|^2 d\xi_N \end{aligned}$$

d'où l'on déduit que

$$\int (1 + |\xi'|^2)^{s-\frac{1}{2}} |g(x_N, \xi')|^2 d\xi' \leq C \|f\|_{H^s(\mathbb{R}^N)}^2$$

Le même calcul montre que

$$\int (1 + |\xi'|^2)^{s-\frac{1}{2}} |g(x_N, \xi') - g(y_N, \xi')|^2 d\xi' \leq C \int |\exp(ix_N \xi_N) - \exp(iy_N \xi_N)|^2 |a(\xi)|^2 d\xi$$

donc, en appliquant le théorème de convergence dominée, on obtient la continuité de la fonction

$$x_N \in \mathbb{R} \mapsto g(x_N, \cdot) \in L^2 \left(\mathbb{R}^{N-1}, \left(1 + |\xi'|^2\right)^{s-\frac{1}{2}} d\xi' \right).$$

Comme la transformation de Fourier partielle $\mathcal{F}'$ par rapport à la variable ξ' est une isométrie de $H^{s-1/2}(\mathbb{R}^{N-1})$ sur $E = L^2 \left(\mathbb{R}^{N-1}, \left(1 + |\xi'|^2\right)^{s-1/2} d\xi' \right)$, on peut alors définir

$$u(x_N) = \frac{1}{2\pi} (\mathcal{F}')^{-1} g(x_N) \quad (8.3.2)$$

On a bien $u \in C_b^0(\mathbb{R}, H^{s-1/2}(\mathbb{R}^{N-1}))$ avec

$$\|u\|_{L^\infty(\mathbb{R}, H^{s-1/2}(\mathbb{R}^{N-1}))} = \frac{1}{2\pi} \|g\|_{L^\infty(\mathbb{R}, E)} \leq C \|f\|_{H^s(\mathbb{R}^N)}$$

et, par le théorème de Fubini,

$$\begin{aligned} \int_{-\infty}^{\infty} \langle u(x_N), \varphi(\cdot, x_N) \rangle &= \frac{1}{(2\pi)^{N-1}} \int \frac{1}{2\pi} g(x_N, \xi') \mathcal{F}' \varphi(x_N, -\xi') d\xi' dx_N \\ &= \frac{1}{(2\pi)^N} \int \mathcal{F} f(\xi) \mathcal{F} \varphi(-\xi) d\xi = \langle f, \varphi \rangle \end{aligned}$$

ce qui conclut la preuve. $\square$

En composant le prolongement $\varphi \in H^1(\mathbb{R}_+^N) \mapsto P\varphi \in H^1(\mathbb{R}^N)$ défini dans la proposition 7.3.1, et la section $\varphi \in H^1(\mathbb{R}^N) \mapsto \varphi|_{x_N=0} \in H^{1/2}(\mathbb{R}^{N-1})$ définie dans la proposition 7.3.2, on peut alors définir l'application trace qui à une fonction de $\varphi \in H^1(\mathbb{R}_+^N)$ associe sa "valeur au bord" $\varphi|_{x_N=0} \in H^{1/2}(\mathbb{R}^{N-1})$.

Plus précisément, on va montrer qu'on a la suite exacte

$$0 \rightarrow H_0^1(\mathbb{R}_+^N) \subset H^1(\mathbb{R}_+^N) \rightarrow H^{1/2}(\mathbb{R}^{N-1}) \rightarrow 0$$

l'image de chaque application de la suite étant le noyau de l'application suivante.

Théorème 8.3.3. *L'application trace*

$$f \in C_c^\infty(\mathbb{R}^N) \mapsto f|_{x_N=0} \in H^{1/2}(\mathbb{R}^{N-1})$$

se prolonge de manière unique en une application γ continue de $H^1(\mathbb{R}_+^N)$ dans $H^{1/2}(\mathbb{R}^{N-1})$. L'image de γ est $H^{1/2}(\mathbb{R}^{N-1})$, et son noyau est $H_0^1(\mathbb{R}_+^N)$.

Démonstration. - On construit l'application trace de la façon suivante :

$$\gamma(f) = u|_{x_N=0}$$

où u est l'unique élément de $C_b^0(\mathbb{R}, H^{1/2}(\mathbb{R}^{N-1}))$ tel que

$$\forall \varphi \in C_c^\infty(\mathbb{R}^N), \quad \langle Pf, \varphi \rangle = \int_{-\infty}^{\infty} \langle u(x_N), \varphi(\cdot, x_N) \rangle$$

et $Pf \in H^1(\mathbb{R}^N)$ est l'extension de f définie dans la proposition 8.3.1.

D'après les propositions 8.3.1 et 8.3.2, γ est continue.

De plus, si $\tilde{f} \in H^1(\mathbb{R}^N)$ vérifie $\tilde{f}\mathbf{1}_{x_N>0} = f\mathbf{1}_{x_N>0}$ et si $\tilde{f} = v$ avec $v \in C_b^0(\mathbb{R}, H^{1/2}(\mathbb{R}^{N-1}))$, on a

$$u(x_N) = v(x_N) \text{ pour } x_N > 0$$

et donc aussi pour $x_N = 0$ par continuité. Donc $\gamma(f)$ ne dépend pas du prolongement P choisi. En particulier, pour $\varphi \in C_c^\infty(\mathbb{R}^N)$, on a

$$\gamma(\varphi) = \varphi|_{x_N=0}.$$

Vérifions que le noyau de γ est l'espace $H_0^1(\mathbb{R}_+^N)$, i.e. l'adhérence pour la topologie de $H^1(\mathbb{R}_+^N)$ de $C_c^\infty(\mathbb{R}_+^N)$.

Comme $\gamma(\varphi) = \varphi|_{x_N=0} = 0$ pour les fonctions $\varphi \in C_c^\infty(\mathbb{R}_+^N)$, le noyau de γ contient $H_0^1(\mathbb{R}_+^N)$.

Réciproquement, soit $f \in H^1(\mathbb{R}_+^N)$ telle que $\gamma(f) = 0$. On vérifie facilement que la fonction de $L^2(\mathbb{R}^N)$ définie par $\bar{f} = f\mathbf{1}_{x_N \geq 0}$ appartient à $C_b^0(\mathbb{R}, H^{1/2}(\mathbb{R}^{N-1}))$. En effet, d'après les propositions 7.3.1 et 7.3.2, $\bar{f} \in C_b^0(\mathbb{R}^+, H^{1/2}(\mathbb{R}^{N-1}))$ et sa trace en 0 est nulle par hypothèse. Le prolongement sur $\mathbb{R}_*^-$ par 0 est donc continu.

On va alors montrer que $\bar{f}$ appartient à $H^1(\mathbb{R}^N)$. Calculons pour cela ses dérivées au sens des distributions. Soit $\psi \in C^\infty(\mathbb{R}, [0, 1])$ telle que

$$\psi \equiv 0 \text{ sur } (-\infty, 1] \text{ et } \psi \equiv 1 \text{ sur } [2, +\infty).$$

On a pour $\varphi \in C_c^\infty(\mathbb{R}^N)$,

$$\begin{aligned} \langle \bar{f}, \partial_j \varphi \rangle &= -\lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^N} \bar{f} \psi\left(\frac{x_N}{\varepsilon}\right) \partial_j \varphi dx = \int_{\mathbb{R}^N} \partial_j f \varphi dx \text{ si } j \neq N, \\ \langle \bar{f}, \partial_N \varphi \rangle &= \int_{\mathbb{R}^N} \partial_N f \varphi dx + \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^N} f \varphi \frac{1}{\varepsilon} \psi'\left(\frac{x_N}{\varepsilon}\right) dx_N \rightarrow 0 \end{aligned}$$

Comme $\bar{f} \in C_b^0(\mathbb{R}, H^{1/2}(\mathbb{R}^{N-1}))$, on a

$$\left| \int_{\mathbb{R}^N} f \varphi \frac{1}{\varepsilon} \psi'\left(\frac{x_N}{\varepsilon}\right) dx_N \right| \leq C \int_{\varepsilon}^{2\varepsilon} \|\bar{f}(x_N)\|_{H^{1/2}} \|\varphi(x_N)\|_{H^{-1/2}} \frac{1}{\varepsilon} \left| \psi'\left(\frac{x_N}{\varepsilon}\right) \right| dx_N \rightarrow 0$$

quand $\varepsilon \rightarrow 0$. On a donc

$$\partial_j \bar{f} = \mathbf{1}_{x_N \geq 0} \partial_j f \in L^2(\mathbb{R}^N)$$

ce qui prouve que $\bar{f} \in H^1(\mathbb{R}^N)$, et que

$$\psi\left(\frac{x_N}{\varepsilon}\right) \bar{f} \rightarrow \bar{f} \text{ dans } H^1(\mathbb{R}^N)$$

Comme $C_c^\infty(\mathbb{R}^N)$ est dense dans $H^1(\mathbb{R}^N)$, il existe une suite (φ_n) de fonctions $C^\infty(\mathbb{R}^N)$ telles que

$$\lim_{n \rightarrow \infty} \|\bar{f} - \varphi_n\|_{H^1(\mathbb{R}^N)} = 0.$$

La suite $\psi\left(\frac{x_N}{\varepsilon}\right) \varphi_n$ est alors une suite de $C_c^\infty(\mathbb{R}^N)$ à support dans $\mathbb{R}_+^N$ qui approche $\psi\left(\frac{x_N}{\varepsilon}\right) \bar{f}$. Cela signifie que $\psi\left(\frac{x_N}{\varepsilon}\right) f \in H_0^1(\mathbb{R}_+^N)$, et par conséquent que $f \in H_0^1(\mathbb{R}_+^N)$.

Vérifions enfin la surjectivité de l'application trace. Soit $g \in H^{1/2}(\mathbb{R}^{N-1})$ et $\mathcal{F}'g$ sa

transformée de Fourier. On choisit $\theta \in C_c^\infty(\mathbb{R})$ tel que $\int \theta(t)dt = 1$ et on définit $f \in \mathcal{S}'(\mathbb{R}^N)$ par

$$\mathcal{F}f(\xi) = 2\pi (\mathcal{F}'g)(\xi') \theta\left(\frac{\xi_N}{\sqrt{1+|\xi'|^2}}\right) \frac{1}{\sqrt{1+|\xi'|^2}}.$$

On a

$$\int \left(1 + |\xi'|^2 + |\xi_N|^2\right) |\mathcal{F}f(\xi)|^2 d\xi = 4\pi^2 \left(\int (1+t^2) \theta^2(t) dt\right) \left(\int \sqrt{1+|\xi'|^2} |\mathcal{F}'g(\xi')|^2 d\xi'\right).$$

De plus, en utilisant la formule (7.3.2), on obtient bien que $g = \gamma(f)$. $\square$

8.4 Compacité

En combinant les résultats des deux paragraphes précédents, on obtient de nouveaux critères de compacité forte dans les espaces de Lebesgue sur des ouverts réguliers et bornés.

Théorème 8.4.1. (Théorème de Rellich-Kondrachov). - Soit Ω un ouvert borné de $\mathbb{R}^N$ de classe C^1 (ou le produit de N intervalles ouverts bornés).

- si $N > 2$ et $1 \leq q < 2N/(N-2)$, l'injection de $H^1(\Omega)$ dans $L^q(\Omega)$ est compacte.
- si $N = 1$, l'injection de $H^1(\Omega)$ dans $C^0(\bar{\Omega})$ est compacte.

Démonstration. - Le théorème de Rellich-Kondrachov se déduit du théorème d'injection de Sobolev et des théorèmes de compacité forte (théorème d'Ascoli dans $C^0(\bar{\Omega})$, théorème de Riesz dans les espaces de Lebesgue).

La première étape consiste à montrer qu'on peut étendre les injections de Sobolev pour les espaces $H^1(\Omega)$ où Ω un ouvert de $\mathbb{R}^N$ dont la frontière est bornée et de classe C^1 (ou le produit de N intervalles ouverts bornés).

- Si Ω est un ouvert de classe C^1 dont la frontière est bornée, il existe un nombre fini K d'ouverts bornés $U_j \subset \mathbb{R}^N$, d'ouverts $\omega_j \subset \mathbb{R}^{N-1}$, de fonctions $\alpha_j \in C^1(\omega_j)$ et de réels δ_j tels que $\partial\Omega \subset \bigcup_{j=1}^K U_j$

$$U_j = \{(y', \alpha_j(y') + t) / y' \in \omega_j, |t| < \delta_j\} \text{ dans un repère orthonormé}$$

$$\text{et } \Omega \cap U_j = \{(y', \alpha_j(y') + t) / y' \in \omega_j, 0 < t < \delta_j\}.$$

Le théorème des partitions de l'unité implique alors l'existence de fonctions $\psi_0, \dots, \psi_K \in C^\infty(\mathbb{R}^N, [0, 1])$ telles que

$$\text{Supp}(\psi_0) \subset \mathbb{R}^N \setminus \partial\Omega, \quad \text{Supp}(\psi_j) \subset U_j, \text{ et } \sum_{j=0}^K \psi_j \equiv 1.$$

Pour tout $u \in H^s(\Omega)$, on définit alors sur chacun des U_j le prolongement de u par réflexion

$$\forall t \in]-\delta_j, \delta_j[, \quad u_j(y', \alpha_j(y') + t) = (u\psi_j)(y', \alpha_j(y') + |t|)$$

On pose finalement

$$Pu = u\psi_0 + \sum_{i=1}^K u_i\psi_i.$$

Les résultats du paragraphe précédent sur les prolongements par réflexion, couplés avec les règles de calcul des dérivées (changement de variable, produit) montrent que $u_j \psi_j \in H^1(\Omega)$ avec

$$\|u_j\|_{H^1(\mathbb{R}^N)} \leq C_j \|u\|_{H^1(\Omega)}$$

d'où l'on déduit que

$$\|Pu\|_{H^1(\mathbb{R}^N)} \leq C \|u\|_{H^1(\Omega)}$$

De plus, on a bien $Pu|_\Omega = u$.

- Si Ω est le produit de N intervalles ouverts bornés, il suffit d'utiliser un nombre fini de prolongements par réflexion et une troncature. On conclut de la même façon que $Pu|_\Omega = u$ et

$$\|Pu\|_{H^1(\mathbb{R}^N)} \leq C \|u\|_{H^1(\Omega)}$$

Cas $N > 2$. D'après le théorème d'injection de Sobolev, pour tout $u \in H^1(\Omega)$,

$$\|u\|_{L^p(\Omega)} \leq \|Pu\|_{L^p(\mathbb{R}^N)} \leq C \|Pu\|_{H^1(\mathbb{R}^N)} \leq \tilde{C} \|u\|_{H^1(\Omega)}$$

où $p = 2N/(N-2)$. Soit alors (u_n) une suite bornée de $H^1(\Omega)$.

- De l'inégalité de Hölder, on déduit que, pour $q < p$

$$\|u_n\|_{L^q(\Omega)} \leq \|u_n\|_{L^p(\Omega)} |\Omega|^{\frac{1}{q} - \frac{1}{p}}.$$

- Soient ω un ouvert relativement compact de Ω et $|y| < d(\omega, \partial\Omega)$. Pour $v \in C_c^\infty(\bar{\Omega})$, on a

$$\begin{aligned} \|\tau_y v - v\|_{L^1(\omega)} &\leq |y| \int_\omega dx \int_0^1 dt |\nabla v(x - ty)| \\ &\leq |y| \|\nabla v\|_{L^2(\Omega)} |\Omega|^{1/2} \end{aligned}$$

et donc par interpolation

$$\|\tau_y v - v\|_{L^q(\omega)} \leq (2\|v\|_{L^p(\Omega)})^\lambda (|y| \|\nabla v\|_{L^2(\Omega)} |\Omega|^{1/2})^{1-\lambda}$$

où $\frac{1}{q} = (1-\lambda) + \frac{\lambda}{p}$.

Par densité de $C_c^\infty(\bar{\Omega})$ dans $H^1(\Omega)$, cette inégalité s'étend à toutes les fonctions de $H^1(\Omega)$.

On a alors

$$\|\tau_y u_n - u_n\|_{L^q(\omega)} \leq C \|u_n\|_{H^1(\Omega)} |y|^{(\frac{1}{q} - \frac{1}{p})/(1-\frac{1}{p})}.$$

- Pour tout $\varepsilon > 0$, il existe $\Omega_k = \{x \in \Omega / d(x, \partial\Omega) > \frac{1}{k}\}$ tel que, pour tout n ,

$$\|u_n\|_{L^q(\Omega \setminus \Omega_k)} \leq \|u_n\|_{L^p(\Omega)} |\Omega \setminus \Omega_k|^{\frac{1}{q} - \frac{1}{p}} \leq \varepsilon.$$

Les trois hypothèses du théorème de Riesz sont satisfaites : (u_n) est donc précompacte dans $L^q(\Omega)$.

Cas $N = 1$. D'après le théorème d'injection de Sobolev, pour tout $u \in H^1(\Omega)$,

$$\|u\|_{C(\bar{\Omega})} \leq \|Pu\|_{C(\mathbb{R})} \leq C \|Pu\|_{H^1(\mathbb{R})} \leq \tilde{C} \|u\|_{H^1(\Omega)}.$$

Soit alors (u_n) une suite bornée de $H^1(\Omega)$

- la suite (u_n) est bornée dans $C(\bar{\Omega})$.

— Pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que, pour tout n , et pour tout $|y| \leq \delta$,

$$\|\tau_y u_n - u_n\|_\infty \leq C|y|^{1/2} \|\nabla u_n\|_{L^2(\Omega)} \leq \varepsilon$$

Les deux hypothèses du théorème d'Ascoli sont satisfaites : (u_n) est donc précompacte dans $C(\bar{\Omega})$. $\square$

Bibliographie

- [1] Robert A. Adams. Sobolev spaces. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.
- [2] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. Functions of bounded variation and free discontinuity problems. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2000.
- [3] Haïm Brezis. Analyse fonctionnelle. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1983. Théorie et applications. [Theory and applications].
- [4] Constantine M. Dafermos. Hyperbolic conservation laws in continuum physics, volume 325 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, third edition, 2010.
- [5] Guy David and Stephen Semmes. Analysis of and on uniformly rectifiable sets, volume 38 of Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 1993.
- [6] Nelson Dunford and Jacob T. Schwartz. Linear operators. Part I. Wiley Classics Library. John Wiley & Sons Inc., New York, 1988. General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.
- [7] Ivar Ekeland and Roger Témam. Convex analysis and variational problems, volume 28 of Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition, 1998. Translated from the French.
- [8] Lawrence C. Evans and Ronald F. Gariepy. Measure theory and fine properties of functions. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.
- [9] Loukas Grafakos. Classical Fourier analysis, volume 249 of Graduate Texts in Mathematics. Springer, New York, second edition, 2008.
- [10] Loukas Grafakos. Modern Fourier analysis, volume 250 of Graduate Texts in Mathematics. Springer, New York, second edition, 2008.
- [11] Serge Lang. Real analysis. Addison-Wesley Publishing Company Advanced Book Program, Reading, MA, second edition, 1983.
- [12] Walter Rudin. Functional analysis. International Series in Pure and Applied Mathematics. McGraw-Hill Inc., New York, second edition, 1991.
- [13] Laurent Schwartz. Analyse. Hermann, Paris, 1970. Deuxième partie : Topologie générale et analyse fonctionnelle, Collection Enseignement des Sciences, No. 11.
- [14] Elias M. Stein. Singular integrals and differentiability properties of functions. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.
- [15] Elias M. Stein. Harmonic analysis : real-variable methods, orthogonality, and oscillatory integrals, volume 43 of Princeton Mathematical Series. Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.
- [16] Michael E. Taylor. Pseudodifferential operators, volume 34 of Princeton Mathemati-

cal Series. Princeton University Press, Princeton, N.J., 1981.

[17] G. B. Whitham. Linear and nonlinear waves. Pure and Applied Mathematics (New York). John Wiley & Sons Inc., New York, 1998. Reprint of the 1974 original, A Wiley-Interscience Publication.

[18] Claude Zuily. Distributions et équations aux dérivées partielles. Collection Méthodes. [Methods Collection]. Hermann, Paris, second edition, 1986. Exercices corrigés. [Exercises with corrections].

[19] Dudley, Richard M. Real analysis and probability . Chapman and Hall/CRC, 2018.

[20] Folland, Gerald B. Real analysis : modern techniques and their applications . Vol. 40. John Wiley & Sons, 1998.