Distribution and Interpolation Theory

Alexis Michelat*

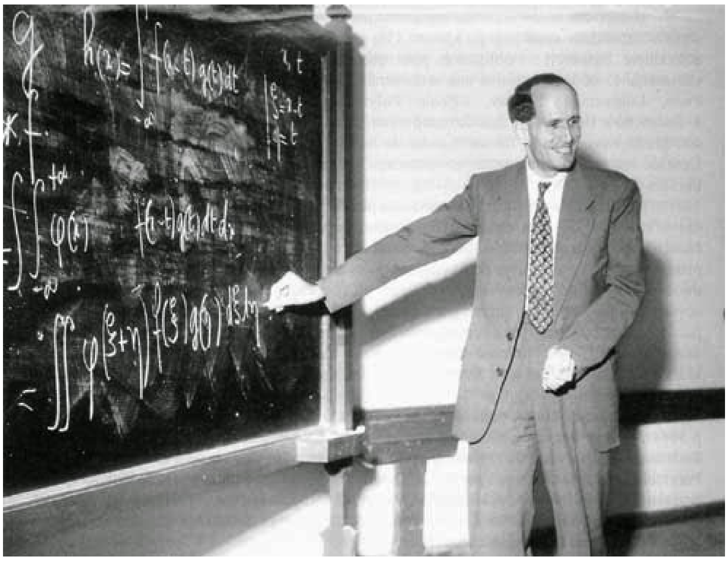

Figure 1: Laurent Schwartz

EPFL MATH-502

Version of February 28, 2025

*EPFL B, Station 8, CH-1015 Lausanne, Switzerland alexis.michelat@epfl.ch

Contents

Introduction

The theory of distributions, introduced by Laurent Schwartz in 1945, gives the modern suitable framework to think of partial differential equations, and has had a considerable success ever since. We will not cover the most subtle aspects of this enormous theory, but mention suitable references to study the most immediate generalisation, such as the Sobolev spaces. Distributions permit to give meaning to expressions that have a very low regularity. Distributions are generalised functions that should rather be thought of generalised measures, and they notably permitted to obtain very deep geometric results after Federer and Fleming introduced, following previous contributions of de Rham and Whitney ([39]), the notion of normal and integral currents ([16], [15]). Indeed, one can think of currents as generalised surfaces that have good compactness properties. One easily solves generalisation of the Plateau problem or obtains minimal surfaces in any dimension and codimension thanks to a very fine theory of Almgren and Pitts ([33]). If existence is relatively easy, regularity is extremely challenging (especially in higher codimension, see [5] and [6]), but extensions of this theory by Codá Marques and Neves permitted to solve the Willmore conjecture in codimension 1 ([30], [31]) and other conjectures ([29]). Let us mention *en passant* that distribution theory permits one to think of Willmore surfaces in a very weak setting ([34]), which has many applications to calculus of variation and geometry. The best references on geometric distributions or current remains Federer's treatise ([15]), and Whitney's classical reference ([40]).

A function belongs to a Sobolev space if its "weak derivatives" belong to some L^p space. Sobolev's major advance was to show that a distribution whose derivative belongs to some L^p space $(1 \leq p \leq \infty)$ was in fact a function from some L^q space (where q depends on p and the ambient dimension). More than half of the lectures will be centred on these spaces and the obtention of their basic properties. Another classical idea going back to the work of Hadmard on fractional derivatives will allow us to introduce Sobolev spaces of fractional order, such that some s-derivative of a given function (where $0 < s < \infty$ is not necessarily integer) belongs to some L^p spaces. Contrary to what one may be thinking, those spaces are nothing but arbitrary, and we will notably mention the classical link between harmonic functions of finite Dirichlet energy on the (unit) disk $\mathbb{D} = \mathbb{C} \cap \{z : |z| < 1\}$ and functions $f : S^1 \to \mathbb{R}$ whose Fourier coefficients $\{c_n\}_{n\in\mathbb{Z}}$ belong to $H^{\frac{1}{2}}$, *i.e.* satisfy

$$\sum_{n\in\mathbb{Z}} |n||c_n|^2 < \infty.$$

We note that those ideas were not especially new as Schwartz recognised it himself (he also claimed that his discovery would have very likely been made in the decade following his work), and Sobolev spaces had already been used implicitly in the work of many, including the one of Leray on Navier-Stokes equations.

In these lecture notes, our main references will be Schwartz's original treatise ([35]), and the classical books of Brezis ([11], [12]), Adams and Fournier ([1]), Gilbarg and Trudinger ([17]), Evans ([14]), and Katznelson ([28]) amongst others. On distributions, we will mainly follow Bony ([10]), Schwartz ([35]), and Edwards ([13]). Another useful references for rather advanced topics is Tartar's lecture notes ([38]). We will not refer to much more advanced references like Hörmander ([23, 24, 25, 26]) or Alinhac-Gérard ([4]). This course is an *introductory* lecture on distribution theory that mainly focuses on basic properties of Sobolev spaces: topics such as micro-local analysis (including pseudo-differential operators, wave front set, wavelets, etc) will not even be mentioned.

The course will be divided into three parts. We will first spend a couple of lectures reminding the attendee about fundamental results of functional analysis—the course Functional I (MATH-302) is not a mandatory prerequisite—define distributions, give their fundamental properties, introduce the notion of Fourier transform of tempered distribution, before moving on to Sobolev spaces, and end the lecture with applications and fractional Sobolev spaces, that will motivate the concept of interpolation spaces, though one may define those spaces directly in practice.

You are welcome to use the above email address to tell me about typos in those notes!

Chapter 1

Topology and Functional Spaces

1.1 Basic definitions

We assume the reader familiar with the basic notions of topology, and only recall a few basic definitions.

Definition 1.1.1. Let X be an arbitrary set. We say that $\mathscr{T} \subset \mathscr{P}(X)$ is a topology if the following properties are verified:

1. If $\{U_i\}_{i\in I} \subset \mathscr{T}$ is an arbitrary family of elements of $\mathscr{T}$, then $\displaystyle\bigcup_{i\in I} U_i \in \mathscr{T}$ **(stability by arbitrary union)**
2. If $U_1, \cdots, U_n \in \mathscr{T}$, then $\displaystyle\bigcap_{i=1}^{n} U_i \in \mathscr{T}$ **(stability by finite intersection).**

Elements of $\mathscr{T}$ are called *open sets*, and complements of open sets are called *closed sets*. We say that such a couple $(X, \mathscr{T})$ is a *topological space*.

Remark 1.1.2. Notice that a set may be closed and open. Taking an empty union and empty intersection, we deduce that both $\varnothing$ and X are open sets, which implies by definition that they are closed too.

On a non-empty set X, there are always at least two topologies: the *trivial topology* given by $\mathscr{T} = \{\varnothing, X\}$, and the *discrete topology* given by $\mathscr{T} = \mathscr{P}(X)$.

We will need of the notion of basis of topology later.

Definition-Proposition 1.1.3. *Let $\mathscr{T}_0 = \{U_i\}_{i\in I}$ be a non-empty collection of sets of a non-empty set X. The smallest topology $\mathscr{T}$ that contains $\mathscr{T}_0$ is given by the following construction. Let $\mathscr{T}_1$ be the family of finite intersection of $\mathscr{T}_0$. Then, $\mathscr{T}$ is given by*

$$\mathscr{T} = \mathscr{P}(X) \cap \left\{ W : W = \bigcup_{j\in J} V_j, V_j \in \mathscr{T}_1 \text{ for all } j \in J \right\}. \tag{1.1.1}$$

Proof. Notice that an arbitrary intersection of topologies is a topology. Indeed, let $\{\mathscr{T}_j\}_{j\in J}$ be a family of topologies, and consider $\mathscr{T} = \displaystyle\bigcap_{j\in J} \mathscr{T}_j$, and $\{U_i\}_{i\in I} \subset \mathscr{T}$. In particular, we have $\displaystyle\bigcup_{i\in I} U_i \in \mathscr{T}_j$ for all $j \in J$, which implies that $\displaystyle\bigcup_{i\in I} U_i \in \mathscr{T}$. Therefore, $\mathscr{T}$ is well-defined and given by

$$\mathscr{T}' = \bigcap_{\mathscr{T}'' \text{ topology } \mathscr{T}_0 \subset \mathscr{T}''} \mathscr{T}'',$$

which is a topology by the above proof. Now, we need to show that $\mathscr{T} = \mathscr{T}'$. Notice that we trivially have $\mathscr{T} \subset \mathscr{T}'$ by using both defining properties of topologies. Therefore, we need only show that $\mathscr{T}$ is a topology to conclude the proof. By construction, $\mathscr{T}$ is stable by arbitrary unions, so we only have to check that $\mathscr{T}$ is stable under finite intersection. Let $W_1, \cdots, W_n \in \mathscr{T}$. Then, there exists sets $J_1, \cdots, J_n$ and $V_{i,j_i} \in \mathscr{T}_1$ ($j_i \in J_i$, $1 \leq i \leq n$) such that

$$W_i = \bigcup_{j_i \in J_i} V_{i,j_i}.$$

Furthermore, we have $V_{i,j_i} = U^1_{i,j_i} \cap \cdots \cap U^{k_i}_{i,j_i}$ for some $U^k_{i,j_i} \in \mathscr{T}$. Finally, we deduce that

$$W_1 \cap \cdots \cap W_n = \bigcap_{i=1}^{n} \bigcup_{j_i \in J_i} \bigcap_{k=1}^{k_i} U^k_{i,j_i}.$$

Let $x \in W_1 \cap \cdots \cap W_n$. Then, for all $1 \leq i \leq n$, there exists $j_i \in J_i$ such that $x \in U^1_{i,j_i} \cap \cdots \cap U^{k_i}_{i,j_i}$. In particular, we have

$$x \in \bigcap_{i=1}^{n} \left(\bigcap_{k=1}^{k_i} U^k_{i,j_i} \right),$$

and

$$W_1 \cap \cdots \cap W_n \in W = \bigcup_{(j_1, \cdots, j_n) \in J_1 \times \cdots \times J_n} \bigcap_{i=1}^{n} \left(\bigcap_{k=1}^{k_i} U^k_{i,j_i} \right) \in \mathscr{T}_1.$$

Likewise, if $x \in W$, then there exists $(j_1, \cdots, j_n) \in J_1 \times \cdots \times J_n$ such that

$$x \in \bigcap_{i=1}^{n} \left(\bigcap_{k=1}^{k_i} U^k_{i,j_i} \right).$$

A fortiori, we have

$$x \in \bigcap_{i=1}^{n} \bigcup_{j_i \in J_i} \bigcap_{k=1}^{k_i} U^k_{i,j_i} = W_1 \cap \cdots \cap W_n,$$

which proves that $W = W_1 \cap \cdots \cap W_n \in \mathscr{T}_1$ and that $\mathscr{T}$ is a topology on X. □

Let us also recall the fundamental notion of neighbourhood.

Definition 1.1.4. Let $(X, \mathscr{T})$ be a topological space. We say that a (non-empty) set N is a neighbourhood of a point $x \in X$ if there exists an open set U containing x such that $U \subset N$.

Finally, we also need the basic notion of interior, closure and frontier of a set.

Definition 1.1.5. Let $(X, \mathscr{T})$ be a topological space. Let $A \subset X$. Its *interior*, denoted by $\mathrm{int}(A)$ or $\mathring{A}$, is the largest open set contained in A, given explicitly by

$$\mathrm{int}(A) = \bigcup_{U \subset A, U \in \mathscr{T}} U,$$

whilst the *closure* of A, denoted by $\mathrm{clos}(A)$ or $\overline{A}$, is the smallest closed set containing A, given explicitly by

$$\mathrm{clos}(A) = \bigcap_{F \supset A, X \setminus F \in \mathscr{T}} F.$$

The frontier (or boundary) of A is given by $\partial A = \overline{A} \setminus \mathrm{int}(A)$.

The defining properties of a topology trivially imply that those notions are well-defined for the arbitrary intersection of closed sets is closed. Those definitions show that arbitrary unions are in general needed to perform basic operations that mimic the classical notions in Euclidean spaces and manifolds.

The following notion will prove crucial in many a proof of those lectures. Indeed, proofs are typically much easier for smooth or more regular functions, and when those functions are dense in a given (Banach) space of functions, a standard argument typically allows one to extend the proof from smooth functions to arbitrary functions in the said Banach space.

Definition 1.1.6. We say that a subset $A \subset X$ of a topological space $(X, \mathscr{T})$ is dense if $\overline{A} = X$.

We say that X is separable if it admits a countable dense set.

Finally, recall the notion of continuity.

Definition 1.1.7. Let $(X, \mathscr{T})$, $(Y, \mathscr{S})$ be two topological spaces. We say that a map $f : X \to Y$ is continuous if for all open set $V \in \mathscr{S}$, we have $f^{-1}(V) \in \mathscr{T}$.

We can finally move to the familiar concept of metric spaces (all spaces encountered in this lecture are metrisable).

Definition 1.1.8. Let X be an arbitrary set. We say that a map $d : X \times X \to \mathbb{R}_+$ is a metric if the following three properties are satisfied

1. $d(x, y) = 0$ if and only if $y = x$ **(definiteness)**.
2. $d(x, y) = d(y, x)$ for all $x, y \in X$ **(symmetry)**.
3. $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in X$ **(triangle inequality)**.

If d is a metric on X, the *open ball* of centre $x \in X$ and radius $r > 0$ is defined by $B(x, r) = X \cap \{y : d(x, y) < r\}$, and the closed ball by $\overline{B}(x, r) = X \cap \{y : d(x, y) \leq r\}$.

Definition 1.1.9. A metric space (X, d) is a topological space whose basis of open sets is given by the sets of all open balls $\{B(x, r)\}_{x \in X, r > 0}$.

Remark 1.1.10. Notice that metric spaces are always separated. It is quite unfortunate choice of terminology, for the closed ball in an arbitrary metric is not always closed. However, the closed ball is always closed in a normed space.

Theorem 1.1.11. *Let (X, d) and (Y, h) be two metric spaces. Then $f : X \to Y$ is continuous if and only if f is sequentially continuous,* i.e. *for all $x \in X$ and for all sequences $\{x_n\}_{n \in \mathbb{N}}$ such that $x_n \underset{n\to\infty}{\longrightarrow} x$, we have $f(x_n) \underset{n\to\infty}{\longrightarrow} f(x) \in Y$.*

We can now move on to the definition of normed space, Banach space, and Hilbert space.

Definition 1.1.12. **1**. Let X be a vector space on a field $\mathbb{K}$ (where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$). We say that a map $\|\cdot\|_X : X \to \mathbb{R}$ is a norm if the following associated map $d : X \to X \to \mathbb{R}_+$, such that $d_X(x, y) = \|x - y\|_X$ is a distance on X, and for all $\lambda \in \mathbb{K}$, we have

$$\|\lambda\, x\|_X = |\lambda|\, \|x\|_X\,.$$

The metric space (X, d_X) is called a *normed space* and denoted (abusively) $(X, \|\cdot\|_X)$.

2. We say that $(X, \|\cdot\|_X)$ is a Banach space if the metric space $(X, \|\cdot\|_X)$ is a complete metric space.

In the following, $\mathbb{K}$ will denote either $\mathbb{R}$ or $\mathbb{C}$.

Remarks 1.1.13. **1**. Notice that we have by the triangle inequality for all $x, y \in X$

$$\|x + y\| = d(x, -y) \leq d(x, 0) + d(0, -y) = \|x\| + \|y\|\,.$$

2. In reality, there are no abuses of notations for the distance associated to a norm is defined bi-univocally.

We can now move to the definition of Hilbert spaces. We first need to remind the definition of scalar product.

Definition 1.1.14. Let E be a vector space on $\mathbb{K}$. A scalar product $\langle\,\cdot\,,\,\cdot\,\rangle : E\times E\to\mathbb{R}$ is a positive-definite symmetric bilinear functional. In other words, it satisfies the following properties:

1. $\langle x,x\rangle>0$ for all $x\in E\setminus\{0\}$ **(positive-definiteness)**.
2. $\langle x,y\rangle=\overline{\langle y,x\rangle}$ for all $x,y\in E$ **(conjugate symmetry)**.
3. $\langle\lambda\,u+v,w\rangle=\lambda\langle u,w\rangle+\langle v,w\rangle$ for all $u,v,w\in E$ and $\lambda\in\mathbb{K}$ **(linearity in the first variable)**.

Remark 1.1.15. Since $\langle\,\cdot\,,\,\cdot\,\rangle$ is symmetric, we need only check the linearity in the first variable.

Definition 1.1.16. We say that a Banach space $(H,\|\cdot\|)$ is a Hilbert space if the following functional

$$\langle x,y\rangle=\frac{1}{4}\left(\|x+y\|^2-\|x-y\|^2\right),\qquad x,y\in H$$

is a scalar product on H.

We do not recall here the useful properties of Hilbert spaces (Riesz-Fréchet representation theorem, Hilbertian basis, and spectral decomposition that will not play a role immediately).

Remark 1.1.17. It may seem that we are replacing a definition by a theorem, but the polarisation formula shows that it is a trivially equivalent definition.

Before mentioning the notion of dual space of a normed space and weak topology, let us recall a statement of the Hahn-Banach theorem (see [11]).

Theorem 1.1.18 (Hahn-Banach)**.** *Let X be a real vector space and $N:X\to\mathbb{R}$ be a sub-linear homogenous map of degree* 1, i.e. *a map such that*

1. $N(\lambda\,x)=\lambda\,N(x)$ *for all* $x\in X$ *and* $\lambda>0$.
2. $N(x+y)\leq N(x)+N(y)$ *for all* $x,y\in X$.

Let $Y\subset X$ be a sub-vector space, and $f:Y\to\mathbb{R}$ be a linear map such that $f\leq N_{|Y}$. Then, there exists an extension $\overline{f}:X\to\mathbb{R}$—i.e. *such that $\overline{f}_{|Y}=f$—such that $\overline{f}\leq N$ on X.*

The proof uses the axiom of choice, and more precisely, the equivalent formulation known as the Zorn's lemma.* First introduce the following definitions.

Definition 1.1.19. (i) A partial order on a set X is a binary relation $\leq$ on $X\times X$ that satisfies the following properties:

1. $x\leq x$ for all $x\in X$ **(reflexivity)**.
2. For all $x,y\in X$, if $x\leq y$ and $y\leq x$, then $x=y$ **(anti-symmetry)**.
3. For all x,y,z, if $x\leq y$ and $y\leq z$, then $x\leq z$ **(transitivity)**.

(ii) We say that a subset $Y\subset X$ is totally ordered (by $\leq$) if for all $x,y\in Y$, we have either $x\leq y$, or $y\leq x$—in which case, we say that $\leq$ is a *total order* (on Y).
(iii) We say that an element $x\in X$ is an upper bound of Y is $y\leq x$ for all $y\in X$.
(iv) Finally, we say that $x\in X$ is a maximal element if for all $y\in X$ such that $x\leq y$, we have $y=x$.

*Another equivalent statement for the axiom of choice is *Zermelo's Theorem*, that asserts that any set can be *well-ordered*. This terminology is rather poorly chosen for what is called either a lemma or a theorem is nothing else than an axiom. However, more than a century of usage will not be erased easily.

Lemma 1.1.20 (Zorn's lemma)**.** *Let $(X, \leq)$ be a non-empty inductive set,* i.e. *a set such that every totally ordered subset admits an upper bound. Then, X admits a maximal element.*

We can finally prove the Hahn-Banach theorem.

Proof. (Of Theorem 1.1.18)

Step 1. Finite-dimensional case.

The theorem is true in finite dimension without the axiom of choice, so let us first prove that a linear map $f : \mathbb{R}^k \subset \mathbb{R}^n \to \mathbb{R}$ (where $k < n$) always admits an extension $\overline{f}$ to $\mathbb{R}^{k+1}$ satisfying $\overline{f} \leq N$ on $\mathbb{R}^{k+1}$. Seeing $\mathbb{R}^k$ as $\mathbb{R}^k \times \{0\} \subset \mathbb{R}^n$, we extend f by $\overline{f} : \mathbb{R}^k \times \mathbb{R} \to \mathbb{R}$ by

$$\overline{f}(x, t) = f(x) + \alpha\, t \qquad \text{for all } (x, t) \in \mathbb{R}^k \times \mathbb{R},$$

for some $\alpha \in \mathbb{R}$ to be determined later. For all $(x, t) \in \mathbb{R}^{k+1}$, we must have

$$f(x) + \alpha\, t \leq N(x, t),$$

where we identify by abuse of notation (x, t) with $(x, , t, 0) \in \mathbb{R}^n$. For $t > 0$, by homogeneity of N, the inequality is equivalent to

$$\left(f(x) + \alpha\, t \leq tN(t^{-1}x, 1)\right) \Longleftrightarrow \left(f(y) + \alpha \leq N(y, 1) \quad (y = t^{-1}x)\right),$$

and for $t < 0$, we get the condition

$$f(y) - \alpha \leq N(y, -1).$$

Therefore, α must satisfy

$$\sup_{y \in \mathbb{R}^k} \left(f(y) - N(y, -1)\right) \leq \alpha \leq \inf_{z \in \mathbb{R}^k} (-f(z) + N(z, 1)).$$

Such an α always exists for $f(y) - N(y, -1) \leq -f(z) + N(z, 1)$ for all $y, z \in \mathbb{R}^k$. Indeed, we have by linearity of f

$$f(y) + f(z) = f(y + z) \leq N(y + z) = N(y + z, -1 + 1) \leq N(y, -1) + N(z, 1),$$

which concludes the proof of this step. Notice that an immediate induction gives an extension of f to $\mathbb{R}^n$.

Step 2. General case.

Let E be the set of extensions $g : D(g) \to \mathbb{R}$ of f (where $D(g) \supset Y$ is the domain of g) such that $g \leq N|_{D(g)}$. We introduce the partial order relation $\leq$ on E as follows:

$$(g_1 \leq g_2) \Longleftrightarrow \left(D(g_1) \subset D(g_2) \text{ and } g_2 = g_1 \text{ on } D(g_1)\right).$$

The set E is not empty since $f \in E$. Furthermore if $F \subset E$ is totally ordered, writing $F = \{g_i\}_{i \in I}$, we see that $g : \bigcup_{i \in I} D(g_i) \to \mathbb{R}$ such that $g = g_i$ on $D(g_i)$ is a well-defined function and an upper bound of F. Therefore, E is inductive, and admits a maximal element that we will denote by f_0. By **Step 1**, if $D(f_0) \neq X$, f_0 admits an extension $\overline{f}_0 : D(\overline{f}_0) \to \mathbb{R}$ such that $D(\overline{f}_0)/D(f_0) \simeq \mathbb{R}$ has codimension 1. In particular, it would imply that f_0 is not a maximal element, a contradiction. Therefore, $D(f_0) = X$ and $\overline{f} = f_0$ is an extension of f satisfying all expected properties. □

Remark 1.1.21. Notice that we do not use the finite-dimension of the ambient space $\mathbb{R}^n$ in **Step 1**, and this why we can apply it to the (potentially) infinite-dimensional case of **Step 2**.

We now let in the rest of this chapter $(X, \|\cdot\|)$ be a normed space. The dual space X' (or X^*) is the space of continuous linear forms $f : X \to \mathbb{R}$ equipped with the following dual norm

$$\|f\|_{X'} = \sup_{\substack{x \in X \\ \|x\| \leq 1}} |f(x)|. \tag{1.1.2}$$

From Hahn-Banach theorem, we deduce the following corollary.

Corollary 1.1.22. *Let $Y \subset X$ be a sub-vector space, and $f : Y \to \mathbb{R}$ be a continuous linear form. Then, there exists an extension $\overline{f} : X \to \mathbb{R}$ such that $\|\overline{f}\|_{X'} = \|f\|_{Y'}$.*

Proof. Take $N(x) = \|f\|_{Y'} \|x\|$. □

Corollary 1.1.23. *For all $x \in X$, there exists $f \in X'$ such that $\|f\|_{X'} = \|x\|_X$ and $f(x) = \|x\|_X^2$.*

Proof. Apply Corollary 1.1.22 to $f_0 : \mathbb{R}\, x \to \mathbb{R}, t \mapsto \|x\|_X^2\, t$. □

Corollary 1.1.24. *For all $x \in X$, we have*

$$\|x\|_X = \sup_{\substack{f \in X' \\ \|f\|_{X'} \leq 1}} |f(x)| = \max_{\substack{f \in X' \\ \|f\|_{X'} \leq 1}} |f(x)|. \tag{1.1.3}$$

Proof. The inequality $|f(x)| \leq \|f\|_{X'} \|x\|_X$ and Corollary 1.1.23 imply the result immediately. □

We will not mention other the geometric forms of Hahn-Banach theorem (see [11]), but we will need the following very useful result in the rest of the lecture.

Theorem 1.1.25. *Let $Y \subset X$ be a sub-vector space such that $\overline{Y} \neq X$. Then, there exists $f \in X' \setminus \{0\}$ such that $f_{|Y} = 0$.*

1.2 The Three Fundamental Theorem of Linear Operators in Banach Spaces

First recall the Baire lemma.

Lemma 1.2.1 (Baire)**.** *Let (X, d) be a complete metric space. Let $\{F_n\}_{n\in\mathbb{N}} \subset X$ a sequence of closed spaces of empty interior,* i.e. *such that* $\mathrm{int}(F_n) = \varnothing$ *for all $n \in \mathbb{N}$. Then, $\displaystyle\bigcup_{n\in\mathbb{N}} F_n$ has empty interior too.*

Let Y be a normed vector space. We denote by $\mathscr{L}(X, Y)$ the space of continuous linear operators $X \to Y$, equipped with the following norm

$$\|T\|_{\mathscr{L}(X,Y)} = \sup_{\substack{x \in X \\ \|x\|_X \leq 1}} \|T(x)\|_Y .$$

We skip the standard proof by induction.

Theorem 1.2.2 (Banach-Steinhaus, or Principle of Uniform Boundedness)**.** *Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$, be two Banach spaces, and $\{T_i\}_{i\in I} \subset \mathscr{L}(X, Y)$ be a family of continuous linear operators from X into Y. Assume that for all $x \in X$, we have*

$$\sup_{i\in I} \|T_i(x)\|_Y < \infty. \tag{1.2.1}$$

Then, we have

$$\sup_{i\in I} \|T_i\|_{\mathscr{L}(X,Y)} < \infty. \tag{1.2.2}$$

Proof. For all $n \in \mathbb{N}$, let $F_n = X \cap \{x : \forall i \in I, \|T_i(x)\| \leq n\}$. Then F_n is an intersection of closed sets, therefore, a closed set. Furthermore, we have $\displaystyle\bigcup_{n\in\mathbb{N}} F_n = X$. Therefore, by Baire's lemma, we deduce that there exists $N \in \mathbb{N}$ such that $\mathrm{int}(F_N) \neq \varnothing$. In particular, there exists an open ball $B(x_0, r)$ in F_N, and we deduce that

$$\forall i \in I,\ \|T_i(x - x_0)\|_Y \leq N \qquad \text{for all } x \in B(x_0, r).$$

By linearity, we deduce that

$$\forall i \in I,\ \|T_i(x)\|_Y \leq \frac{1}{r}\left(N + \|T_i(x_0)\|_Y\right)\|x\|_X \leq C\,\|x\|_X\,,$$

using (1.2.1) with $x = x_0$. □

Let us list a few corollaries.

Corollary 1.2.3. *Let X and Y be two Banach spaces. Let $\{T_n\}_{n\in\mathbb{N}} \subset \mathscr{L}(X,Y)$ be a sequence of linear continuous operators from X to Y, such that for all $x \in X$, the sequence $\{T_n(x)\}_{n\in\mathbb{N}} \subset Y$ converges to a limit denoted by $T(x) \in Y$. Then, the following properties are satisfied:*

1. $\sup_{n\in\mathbb{N}} \|T_n\|_{\mathscr{L}(X,Y)} < \infty$.
2. $T \in \mathscr{L}(X,Y)$.
3. $\|T\|_{\mathscr{L}(X,Y)} \leq \liminf_{n\to\infty} \|T_n\|_{\mathscr{L}(X,Y)}$.

Proof. The first point **1**. follows from Theorem 1.2.2. In particular, there exists a constant $C < \infty$ such that

$$\sup_{n\in\mathbb{N}} \|T_n(x)\|_Y \leq C\,\|x\|_X \quad \text{for all } x \in X.$$

In particular, we have

$$\|T(x)\|_Y \leq C\,\|x\|_X \quad \text{for all } x \in X.$$

By linearity of T_n, we deduce that T is linear, which proves **2**. Finally, the inequality

$$\|T_n(x)\| \leq \|T_n\|_{\mathscr{L}(X,Y)}\,\|x\|_X \quad \text{for all } x \in X$$

implies the last point **3**. □

Corollary 1.2.4. *Let X be a Banach space and $A \subset X$ an arbitrary subset. Assume that A is weakly bounded,* i.e. *for all $f \in X'$, the set $f(A) \subset \mathbb{R}$ is bounded. Then, A is strongly bounded in X.*

Proof. Let $\{T_a\}_{a\in A} \subset \mathscr{L}(X',\mathbb{R})$ be defined by $T_a(f) = f(a)$ for all $f \in X'$. Then, we have

$$\sup_{a\in A} \|T_a(f)\| < \infty \quad \text{for all } f \in X'.$$

Therefore, by Theorem 1.2.2, we have

$$\sup_{a\in A} \|T_a\|_{\mathscr{L}(X',\mathbb{R})} < \infty.$$

In particular, we have

$$|f(a)| \leq C\,\|f\|_{X'} \ \text{ for all } f \in X'.$$

Using Corollary 1.1.23, we deduce that $\|a\| \leq C$ for all $a \in A$, which concludes the proof. □

The dual statement is given by the following.

Corollary 1.2.5. *Let X be a Banach space and $F \subset X'$. Assume that for all $x \in X$, the set $F(x) = \mathbb{R} \cap \{y : y = f(x)$ for some $f \in F\}$ is bounded. Then, F is bounded.*

Proof. The proof is almost identical, using the family $\{T_f = f\}_{f\in F}$. □

The second fundamental theorem of Banach is the following.

Theorem 1.2.6 (Open Mapping Theorem). *Let X and Y be two Banach spaces, and $T \in \mathscr{L}(X,Y)$ be a surjective linear continuous operator. Then, there exists $r > 0$ such that*

$$B_Y(0,r) \subset T(B_X(0,1)).$$

Proof. **Step 1.** We show that $B_Y(0,r) \subset \overline{T(B_X(0,1)}$.

Since T is surjective, we have $\displaystyle\bigcup_{n\in\mathbb{N}} \overline{T(B_X(0,n))} = Y$. Therefore, Baire's lemma implies that there exists $N \in \mathbb{N}^*$ such that $\mathrm{int}(\overline{T(B_X(0,N))}) \neq \varnothing$, and by linearity $(T(B_X(0,N)) = NT(B_X(0,1)))$, we deduce that $\overline{T(B_X(0,N))}$ has non-empty interior. Therefore, there exists $y_0 \in Y$ and $r > 0$ such that $B_Y(y_0,r) \subset \overline{T(B_X(0,1)}$. Therefore, for all $y \in B_Y(0,r)$, there exists $\{x_n(y)\}_{n\in\mathbb{N}} \subset B_X(0,1)$ such that

$$y_0 + y = \lim_{n\to\infty} T(x_n(y)).$$

Therefore, we have by linearity

$$y = \lim_{n\to\infty} T\left(\frac{x_n(y) - x_n(-y)}{2}\right),$$

and since $\frac{1}{2}(x_n(y) - x_n(-y)) \in B_X(0,1)$ by convexity, we deduce that $y \in \overline{T(B_X(0,1))}$, which shows that $B_Y(0,r) \subset \overline{T(B_X(0,1))}$.

Step 2. We show that $B_Y(0,r) \subset T(B_X(0,1))$.

Let $y \in B_Y(0,r)$, and fix some $0 < \delta < 1$ to be determined later. Then, there exists $\varepsilon > 0$ such that $(1+\varepsilon)y \in B_Y(0,r)$. Therefore, by the previous step, there exists $x_0 \in B_X(0,1)$ such that

$$\|(1+\varepsilon)y - T(x_0)\| < \frac{\delta r}{2}.$$

Therefore, we get

$$\left\|y - T((1+\varepsilon)^{-1}x_0)\right\| < \frac{\delta}{(1+\varepsilon)}\frac{r}{2} = \eta\frac{r}{2},$$

where $\eta = \frac{\delta}{1+\varepsilon}$ to simplify. We deduce that there exists $x_1 \in B_X(0,1)$ such that

$$\left\|2\eta^{-1}(y - T((1+\varepsilon)^{-1}x_0)) - T(x_1)\right\| < \eta\frac{r}{2},$$

or

$$\left\|y - T\left((1+\varepsilon)^{-1}\left(x_0 + \frac{\delta}{2}x_1\right)\right)\right\| < \eta^2\frac{r}{4}.$$

By immediate induction, we deduce that there exists $\{x_n\}_{n\in\mathbb{N}} \subset B_X(0,1)$ such that for all $n \in \mathbb{N}$, we have

$$\left\|y - T\left((1+\varepsilon)^{-1}\left(x_0 + \frac{\delta}{2}x_1 + \cdots + \left(\frac{\delta}{2}\right)^n x_n\right)\right)\right\| < \left(\frac{\eta}{2}\right)^{n+1} r \underset{n\to\infty}{\longrightarrow} 0.$$

Furthermore, defining $x_n' = (1+\varepsilon)^{-1}\left(\frac{\delta}{2}\right)^n x_n$, we have $\{x_n'\}_{n\in\mathbb{N}} \subset B_X(0, \frac{1}{1+\varepsilon})$, and

$$\sum_{n\in\mathbb{N}} \|x_n'\|_X \leq \frac{1}{1+\varepsilon}\left(1 + \sum_{n=1}^{\infty}\left(\frac{\delta}{2}\right)^n\right) = \frac{1}{1+\varepsilon}\left(1 + \frac{\delta}{1-\delta}\right). \tag{1.2.3}$$

Therefore, we deduce that $\sum_{n\in\mathbb{N}} x_n'$ converges absolutely in $B_X(0,1)$ (provided that $\delta > 0$ is small enough), which implies as X is a Banach space that

$$\sum_{k=0}^{n} x_k' \underset{n\to\infty}{\longrightarrow} x \in \overline{B}_X\left(0, \frac{1+\frac{\varepsilon}{2}}{1+\varepsilon}\right) \subset B_X(0,1),$$

taking $0 < \delta < \frac{\varepsilon}{2+\varepsilon}$ ($\delta = \frac{\varepsilon}{3}$ would work). By continuity of T, we finally deduce that $T(x) = y$, which concludes the proof of the theorem. □

Finally, we give the third theorem of Banach.

Theorem 1.2.7 (Closed Graph Theorem)**.** *Let X, Y be two Banach spaces, and $T : X \to Y$ be a linear map. Assume that the graph of T,*

$$\mathscr{G}(T) = X \times Y \cap \{(x, y) : y = T(x)\}$$

is a closed set of $X \times Y$. Then T is continuous.

Proof. Consider on X the norm $|||x||| = \|x\|_X + \|T(x)\|_Y$. Since $\mathscr{G}(T)$ is closed, $(X, |||\cdot|||)$ is a Banach space. Indeed, let $\{x_n\}_{n\in\mathbb{N}} \subset (X, |||\cdot|||)$ be a Cauchy sequence. Then, we have

$$\limsup_{n,m\to\infty} |||x_n - x_m||| = 0.$$

Equivalently, we have

$$\limsup_{n,m\to\infty} \|x_n - x_m\|_X = 0 \qquad \text{et} \qquad \limsup_{n,m\to\infty} \|T(x_n) - T(x_m)\|_Y = 0.$$

In particular, $\{x_n\}_{n\in\mathbb{N}} \subset X$ and $\{y_n = T(x_n)\}_{n\in\mathbb{N}} \subset Y$. As X and Y are Banach spaces, we deduce that there exists $x_\infty \in X$ and $y_\infty \in Y$ such that $x_n \underset{n\to\infty}{\longrightarrow} x_\infty$ in X and $y_n \underset{n\to\infty}{\longrightarrow} y_\infty$ in Y. By hypothesis, as $\mathscr{G}(T)$ is closed, we deduce that $(x_\infty, y_\infty) \in \mathscr{G}(T)$, which shows that $T(x_\infty) = y_\infty$. Therefore, we have

$$|||x_n - x_\infty||| \underset{n\to\infty}{\longrightarrow} 0$$

and this shows that $(X, |||\cdot|||)$ is a Banach space.

Furthermore, we have $\|\cdot\|_X \leq |||\cdot|||$. By the open map theorem applied to the identity map $\mathrm{Id}(X, |||\cdot|||) \to (X, \|\cdot\|)$, we deduce that there exists $r > 0$ such that $B_{\|\cdot\|}(0, r) \subset \mathrm{Id}\left(B_{|||\cdot|||}(0, 1)\right) = B_{|||\cdot|||}(0, 1)$. In other words, for all $x \in X$ such that $\|x\| < r$, we have $|||x||| < 1$. In other words, we have

$$\sup_{\|x\|\leq r} |||x||| \leq 1,$$

which is equivalent thanks to the homogeneity of the norm that

$$\sup_{\|x\|\leq 1} |||x||| \leq \frac{1}{r}.$$

In particular, we have

$$\sup_{\|x\|_X\leq 1} \|T(x)\|_Y \leq \frac{1}{r}$$

which shows that $\|T\|_{\mathscr{L}(X,Y)} \leq \dfrac{1}{r} < \infty$. □

The argument in the second part of the proof works in a more general setting.

Corollary 1.2.8. *Let X and Y be two Banach spaces, and let $T \in \mathscr{L}(X, Y)$ be a bijective linear operator. Then, the inverse $T^{-1} : Y \to X$ is continuous.*

Proof. By the Open Mapping Theorem (Theorem 1.2.6), we deduce that there exists $r > 0$ such that

$$r\,\|x\|_X \leq \|T(x)\|_Y \quad \text{for all } x \in B_X(0, 1),$$

which shows that $\left\|T^{-1}\right\|_{\mathscr{L}(Y,X)} \leq \dfrac{1}{r}$. □

1.3 Weak Topology

1.3.1 General Definition

Let X be a set and $\{Y_i\}_{i\in I}$ be a family of topological spaces. For all $i \in I$, we fix some map $\varphi_i : X \to Y$. The *weak topology* on X is with topology that makes all maps $\varphi_i : X \to Y_i$ continuous. Notice that this is well-defined by Definition 1.1.3), and the associated pre-topology is given by $\mathscr{T}_0 = \left\{\varphi_i^{-1}(V_i) : V_i \text{ open subset of } Y_i\right\}$.

Proposition 1.3.1. *Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence of X. Then $x_n \underset{n\to\infty}{\rightharpoonup} x$ for the weak topology if and only if $\varphi_i(x_n) \underset{n\to\infty}{\longrightarrow} \varphi_i(x) \in Y_i$ for all $i \in I$.*

Proof. The first implication is trivial for each map $\varphi_i : X \to Y_i$ is continuous with respect to the weak topology. Conversely, let U be a neighbourhood of x. By the construction of Definition 1.1.3, we can assume that

$$U = \bigcap_{j=1}^{n} \varphi_{i_j}^{-1}(V_{i_j}),$$

where V_{i_j} is an open set of Y_{i_j} (by hypothesis, V_{i_j} is also a neighbourhood of $\varphi_{i_j}(x)$). For all $1 \leq j \leq n$, there exists $N_j \in \mathbb{N}$ such that $\varphi_{i_j}(x_n) \in V_{i_j}$ for al $n \geq N_j$. In particular, taking $N = \max\{N_1, \cdots, N_n\}$, we deduce that for all $n \geq N$, $x_n \in U$, which concludes the proof. □

1.3.2 Weak Topology on a Banach Space

Let X be a Banach space, and $f \in X'$. Let $\varphi_f : X \to \mathbb{R}$ be defined by $\varphi_f(x) = f(x)$ for all $x \in X$. Then, the weak topology $\sigma(X, X')$ on X is the weak topology associated to the family of maps $\{\varphi_f\}_{f\in X'}$. To emphasise the duality, we will sometimes write $f(x) = \langle f, x\rangle$.

We will denote the weak convergence of $\{x_n\}_{n\in\mathbb{N}} \subset X$ to some element $x \in X$ in the weak topology by the half-arrow $\rightharpoonup$. Notice that by what precedes (Proposition 1.3.1), we have

$$\left(x_n \underset{n\to\infty}{\rightharpoonup} x\right) \Longleftrightarrow \left(f(x_n) \underset{n\to\infty}{\longrightarrow} f(x) \in \mathbb{R} \text{ for all } f \in X'\right).$$

Let us list some basic properties of the weak topology.

Proposition 1.3.2. *The weak topology $\sigma(X, X')$ is separated.*

Proof. The proof follows from the geometric version of Hahn-Banach theorem, and will be omitted. □

Proposition 1.3.3. *Let $\{x_n\}_{n\in\mathbb{N}} \subset X$. The following properties are verified:*

1. *The sequence $\{x_n\}_{n\in\mathbb{N}}$ weakly converges to some element $x \in X$ if and only if $f(x_n) \underset{n\to\infty}{\longrightarrow} f(x) \in \mathbb{R}$ for all $f \in X'$.*

2. *If $x_n \underset{n\to\infty}{\longrightarrow} x$ strongly, then $x_n \underset{n\to\infty}{\rightharpoonup} x$ weakly.*

3. *If $x_n \underset{n\to\infty}{\rightharpoonup} x$ weakly, then $\{\|x_n\|\}_{n\in\mathbb{N}} \subset \mathbb{R}_+$ is bounded and*

$$\|x\| \leq \liminf_{n\to\infty} \|x_n\|. \tag{1.3.1}$$

4. *If $x_n \underset{n\to\infty}{\rightharpoonup} x$ weakly, and $\{f_n\}_{n\in\mathbb{N}} \subset X'$ converges towards some element $f \in X'$, then $f_n(x_n) \underset{n\to\infty}{\longrightarrow} f(x)$.*

Proof. The first property **1**. is trivial by definition of the weak topology and Proposition 1.3.1. The second **2**. follows from the inequality $|f(x_n) - f(x)| \leq \|f\|_{X'} \|x_n - x\|_X$.

Let us prove **3**. now. We apply Corollary 1.2.5. We need to check that for all $f \in X'$, $\{f(x_n)\}_{n\in\mathbb{N}} \subset \mathbb{R}$ is bounded, which is trivially satisfied since $f(x_n) \underset{n\to\infty}{\longrightarrow} f(x)$ by definition of the weak convergence. Furthermore, for all $n \in \mathbb{N}$, we have

$$|f(x_n)| \leq \|f\|_{X'} \|x_n\|_X ,$$

which implies that

$$|f(x)| \leq \|f\|_{X'} \liminf_{n\to\infty} \|x_n\|_X .$$

Finally, Corollary 1.1.24 implies that

$$\|x\|_X = \max_{\substack{f\in X' \\ \|f\|_{X'}\leq 1}} |f(x)| \leq \liminf_{n\to\infty} \|x_n\|_X .$$

The last property **4**. follows immediately by the triangle inequality:

$$|f_n(x_n) - f(x)| \leq |f_n(x_n) - f(x_n)| + |f(x_n) - f(x)| \leq \|f_n - f\|_{X'} \|x_n\|_X + \|f\|_{X'} \|x_n - x\|_X ,$$

which implies the claim by the previous property **3**.. □

We end this section by a few remarks on the weak topology.

Remarks 1.3.4. The weak topology has many surprising properties:

1. The adherence of the unit sphere $S = X \cap \{x : \|x\|_X = 1\}$ for the weak topology is the closed ball $\overline{B} = X \cap \{x : \|x\|_X \leq 1\}$. We will see that in a reflexive space (to be defined in Definition 1.4.1), $\overline{B}$ is a *compact* set for the weak topology, although this set is never compact for the strong topology in infinite dimension. This is why the weak topology is so important: it allows one to solve partial differential equations thanks to a compactness argument.
2. The interior of $B = X \cap \{x : \|x\|_X < 1\}$ for the weak topology is empty.
3. In infinite dimension, the weak topology is never metrisable. This is why it is futile to define it using convergence of sequences, although for most applications, one need only look at sequences.
4. In infinite dimension, there are sequences that converge weakly but do not converge strongly.

1.4 Weak $*$ Topology

Let X be a Banach space, X' its dual space, and $X'' = (X')'$ the dual space of X' (also called *bidual* of X). We endow it with the following norm

$$\|\varphi\|_{X''} = \sup_{\substack{f\in X' \\ \|f\|_{X'}\leq 1}} |\varphi(f)|. \tag{1.4.1}$$

There is a canonical injection $J : X \to X''$, defined as follows. Let $x \in X$ and $J(x) : X' \to \mathbb{R}, f \mapsto \langle J(x), f\rangle = f(x)$. Then $J(x) \in X''$. Furthermore, we immediately check that J defines a linear map $X \to X''$, which is an isometry for

$$\|J(x)\|_{X''} = \sup_{\substack{f\in X' \\ \|f\|_{X'}\leq 1}} |\langle J(x), f\rangle| = \sup_{\substack{f\in X' \\ \|f\|_{X'}\leq 1}} |f(x)| = \|x\|_X , \tag{1.4.2}$$

using Corollary 1.1.23.

Therefore, X is isometric to a subset of X''. This allows us to introduce a fundamental notion that will prove fundamental in the following (and explain all the pathologies of spaces such as L^1 and L^∞).

Definition 1.4.1 (Reflexive spaces)**.** We say that a Banach space is reflexive if the isometric injection $J : X \hookrightarrow X''$ is surjective, *i.e.* $J(X) = X''$.

Common examples of reflexive spaces are the L^p spaces (on a locally compact group, say) for exponents $1 < p < \infty$.

Before listing the major properties of reflexive spaces, we now define the weak $*$ topology $\sigma(X', X)$ on X'.

Definition 1.4.2. The weak $*$ topology[†] is the smallest topology that makes all maps $J(x) : X' \to \mathbb{R}$ continuous, where $x \in X$. We denote it $\sigma(X', X)$. We denote by $\overset{*}{\rightharpoonup}$ the convergence for sequences of X'.

Let us give a few basic properties of the weak topology.

Proposition 1.4.3. *The weak $*$ topology on X' is separated.*

Proof. Let $f, g \in X'$ such that $f \neq g$. Then, there exists $x \in X$ such that $f(x) \neq g(x)$. Assume without loss of generality that $f(x) < g(x)$, and let $\alpha \in \mathbb{R}$ such that

$$f(x) < \alpha < g(x),$$

then $J(x)^{-1}(]-\infty, \alpha[)$ and $J(x)^{-1}(]\alpha, \infty[)$ are disjoint open (for the weak $*$ topology) subset of X' that respectively contain f and g. ☐

We now list the basic properties of the weak $*$ topology (the proof is almost identical to the one of Proposition 1.3.3, and we omit it).

Proposition 1.4.4. *Let $\{f_n\}_{n\in\mathbb{N}} \subset X'$. Then the following properties are satisfied.*

1. *The sequence $\{f_n\}_{n\in\mathbb{N}}$ converges to $f \in X'$ if and only if $f_n(x) \underset{n\to\infty}{\longrightarrow} f(x)$ for all $x \in X$.*
2. *If $f_n \underset{n\to\infty}{\longrightarrow} f \in X'$ strongly, then $f_n \underset{n\to\infty}{\rightharpoonup} f$ weakly for the weak topology $\sigma(X', X'')$. If $f_n \underset{n\to\infty}{\rightharpoonup} f \in X'$ for the weak topology $\sigma(X', X'')$, then $f_n \underset{n\to\infty}{\rightharpoonup} f$ weakly for the weak $*$ topology $\sigma(X', X)$.*
3. *If $f_n \underset{n\to\infty}{\overset{*}{\rightharpoonup}} f$, then $\{\|f_n\|_{X'}\}_{n\in\mathbb{N}} \subset \mathbb{R}_+$ is bounded and*
$$\|f\|_{X'} \leq \liminf_{n\to\infty} \|f_n\|_{X'}. \tag{1.4.3}$$
4. *If $f_n \underset{n\to\infty}{\overset{*}{\rightharpoonup}} f$, and $x_n \underset{n\to\infty}{\longrightarrow} x$ strongly, then $f_n(x_n) \underset{n\to\infty}{\longrightarrow} f(x)$.*

We end this section by a fundamental compactness theorem that justifies the introduction.

Theorem 1.4.5 (Banach-Alaoglu-Bourbaki)**.** *The unit closed ball $B = X' \cap \{f : \|f\|_{X'} \leq 1\}$ is compact for the weak $*$ topology $\sigma(X', X)$.*

Proof. The proof is an easy application of Tychonoff's theorem (the arbitrary product of compact set is compact). Notice that this "theorem" is equivalent to the axiom of choice, so it was not very limiting to use Hahn-Banach theorem previously considering that the compactness of the unit ball for the weak $*$ topology is needed in many applications.

Now, let $Y = \mathbb{R}^X$ equipped with the product topology. Let $\Phi : X' \to Y$ defined by

$$\Phi(f) = \{f(x)\}_{x\in X} \quad \text{for all } f \in X'.$$

By definition of the product topology, since each canonical projection $\pi_x \circ \Phi = J(x) : X' \to \mathbb{R}$ is continuous ($x \in X$), we deduce that Φ is a continuous map. Here, we denoted $\pi_x : Y = \mathbb{R}^X \to \mathbb{R}$ the canonical projection on the "x factor." Furthermore, note that Φ is injective since for all given elements

[†]One pronounces *weak star topology.*

$f, g \in X'$, the equality $f = g$ holds if and only if $f(x) = g(x)$ for all $x \in X$. Now, consider the inverse map $\Phi^{-1} : \Phi(X') \to X'$. We need only prove that for all $x \in X$, the map $y \mapsto \langle \Phi^{-1}(y), x\rangle$ is continuous, but it is trivial since $\langle \Phi^{-1}(y), x\rangle = \pi_x(y)$.

Now, we observe that

$$\Phi(\overline{B}) = Y \cap \{y : |\pi_x(y)| \leq \|x\|, \pi_{x+x'}(y) = \pi_x(y) + \pi_{x'}(y), \\ \pi_{\lambda\, x}(y) = \lambda\, \pi_x(y) \text{ for all } x, x' \in X \text{ and } \lambda \in \mathbb{R}\}.$$

Notice that the set $A_1 = Y \cap \{y : |\pi_x(y)| \leq \|x\| \text{ for all } x \in X\} = \prod_{x\in X}[-\|x\|, \|x\|]$ is compact by Tychonoff's theorem, whilst

$$A_2 = Y \cap \{y : \pi_{x+x'}(y) = \pi_x(y) + \pi_{x'}(y), \pi_{\lambda x}(y) = \lambda\, \pi_x(y) \text{ for all } x, x' \in X \text{ and } \lambda \in \mathbb{R}\}$$

is closed as intersection of closed sets. Therefore, we deduce that $\Phi(B) = A_1 \cap A_2$ is compact. □

1.5 Reflexive Spaces

Recall that by the Definition 1.4.1, a Banach space is reflexive if the canonical (isometric) injection $J : X \to X''$ is surjective. The major theorem is the following result of Kakutani.

Theorem 1.5.1 (Kakutani)**.** *Let X be a Banach space. Then, X is reflexive if and only if the unit closed ball $\overline{B} = X \cap \{x : \|x\|_X \leq 1\}$ is compact for the weak topology $\sigma(X, X')$.*

We omit the (rather technical) proof.

Remark 1.5.2. We see that for a reflexive space, the weak $*$ topology is useless. However, for a non-reflexive space that is the dual of a Banach space (as L^∞), the weak $*$ topology furnishes a topology for which the unit ball is compact, which has fundamental applications to calculus of variations and partial differential equations.

We also mention the following theorem that is not trivial, contrary to what one may think.

Theorem 1.5.3. *A Banach space is reflexive if and only if its dual space is reflexive.*

1.6 Separable Spaces

We have the following results.

Theorem 1.6.1. *Let X be a Banach space such that X' is separable. Then, X is separable.*

Remark 1.6.2. L^∞, the dual of L^1, is not separable, although L^1 (as all Lebesgue spaces L^p for $1 \leq p < \infty$) is separable (provided that we consider the space L^1 on an open subset of $\mathbb{R}^d$ for example).

Theorem 1.6.3. *Let X be a Banach space. Then X is reflexive and separable if and only if X' is reflexive and separable.*

We assume the reader familiar with Lebesgue spaces (since they are special cases of Sobolev spaces) and do not recall here the basic results such as the Hölder's inequality (we will see generalisations of it), the convergence theorems of Lebesgue or Fatou, or the inequality for convolutions that will all be treated in the more general setting of Lorentz and Orlicz spaces.

Chapter 2

Distributions

2.1 Basic Notations

From now on, we let $d \geq 1$ and consider a (connected) open subset $\Omega \subset \mathbb{R}^d$. For all $\alpha = (\alpha_1, \cdots, \alpha_d) \in \mathbb{R}^d$, we define the $|\alpha|$-order operator (where $|\alpha| = \alpha_1 + \cdots + \alpha_d$) by

$$D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}.$$

Then the Taylor expansion shows that for any smooth function $f : \Omega \to \mathbb{R}$, and for all $m \in \mathbb{N}$, and $x_0 \in \Omega$ we have for all $x \in \Omega$ such that
$[x_0, x] = \Omega \cap \{y : y = tx + (1-t)x_0 \ \text{for some} \ t \in [0,1]\} \subset \Omega$

$$f(x) = \sum_{|\alpha| \leq m} \frac{D^\alpha f(x_0)}{\alpha!}(x - x_0)^\alpha + \sum_{|\beta| = m+1} \frac{m+1}{\beta!} \int_0^1 D^\beta f(x_0 + t(x - x_0))(1-t)^m dt \tag{2.1.1}$$

where $\alpha! = \alpha_1! \cdots \alpha_d!$, and $y^\alpha = y_1^{\alpha_1} \cdots y_d^{\alpha_d}$ for all $y \in \mathbb{R}^d$.

2.2 Topological Vector Spaces

Distributions are defined as dual spaces of generalisations of Banach spaces for which the topology is defined by an (uncountable) family of semi-norms. Those spaces are locally convex topological vector spaces (see [13], Chapter 1) that are *not* metrisable (a metrisable locally convex topological vector space is called a Fréchet space). On the other hand, tempered distributions are defined as dual space of Fréchet spaces. A bad definition of locally convex topological vector spaces would be to start with a countable family of semi-norms, for it would not explain why such a definition is natural. We will therefore follow the order of Edwards ([13]).

In recent analysis textbooks on distributions, the topological foundations are barely mentioned, and only sequential convergence of distributions is mentioned. Even if sequences suffice in most applications, we find it helpful to give suitable background to those interested to continue further in the theoretical direction, that answer many a non-trivial question, such as the PhD thesis ([19]) of Alexandre Grothendieck (made under the direction of Laurent Schwartz) that solved most problems in the field, and notably gave rise to the notion of *nuclear spaces*. Grothendieck showed that there is another natural notion of tensor product of locally convex topological vector spaces, much to the dismay of Schwartz! However, for a sub-class of locally convex topological vector spaces, the tensor product is unique, and Grothendieck called them *nuclear*. Thankfully, all spaces of distributions $\mathscr{D}$, $\mathscr{D}'$, $\mathscr{E}$, $\mathscr{E}'$, $\mathscr{S}$, $\mathscr{S}'$ and the space of holomorphic functions on a complex manifold are all nuclear. However, one can endow those spaces with various topologies, and one should give a rationale why one notion is preferred to the other. Spaces of general distributions (contrary to tempered distributions) are *not* metrisable, and one therefore needs to give a satisfying definition of topology and not restrict to sequences without further explanations. Up to this day, Grothendieck's PhD remains his most quoted work. Out of the enormous amount of results it contains (even restricting to Banach or Hilbert spaces), we cannot resist to mention Grothendieck's inequality, that has deep applications to various fields, and whose optimal constants are still unknown.

Needless to say, this theorem is only mentioned for cultural reasons. Consider it as **hors-piste** (off-piste).

Theorem 2.2.1 (Grothendieck, 1955)**.** *For all $N \in \mathbb{N}$, and $1 \leq p < \infty$, define for all $x \in \mathbb{K}^N$ (where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$)*

$$\|x\|_p = \left(\sum_{i=1}^{N} |x_i|^p\right)^{\frac{1}{p}},$$

and for $p = \infty$

$$\|x\|_\infty = \sup_{1\leq i\leq N} |x_i| = \max_{1\leq i\leq N} |x_i|.$$

For all matrix $A \in M_{m,n}(\mathbb{K})$, and for all $1 \leq p, q \leq \infty$, define the operator norm of A as a map $(\mathbb{K}^n, l^p) \to (\mathbb{K}^m, l^q)$ is defined by

$$\|A\|_{p,q} = \sup_{\substack{x\in\mathbb{K}^n \\ \|x\|_p\leq 1}} \|A(x)\|_q .$$

Then, the following statement holds. There exists a universal constant $K_G^{\mathbb{K}} < \infty$ such that for all $m, n \geq 1$ and $A \in M_{m,n}(\mathbb{K})$, for all *Hilbert space $(H, \|\cdot\|_H)$,*

$$\sup_{\substack{(u,v)\in H^m\times H^n \\ \|u_i\|_H, \|v_j\|_H\leq 1}} \left|\sum_{i=1}^{m}\sum_{j=1}^{n} A_{i,j}\langle u_i, v_j\rangle_H\right| \leq K_G^{\mathbb{K}} \|A\|_{\infty,1} . \tag{2.2.1}$$

Remark 2.2.2. The $\|\cdot\|_{\infty,1}$ norm of A is alternatively given by

$$\|A\|_{\infty,1} = \sup_{\substack{(x,y)\in\mathbb{K}^m\times\mathbb{K}^n \\ \|x\|_\infty, \|y\|_\infty\leq 1}} \left|\sum_{i=1}^{m}\sum_{j=1}^{n} A_{i,j}x_i x_j\right| = \max_{\substack{(x,y)\in\mathbb{K}^m\times\mathbb{K}^n \\ \|x\|_\infty, \|y\|_\infty\leq 1}} \sum_{i=1}^{m}\sum_{j=1}^{n} A_{i,j}x_i x_j.$$

Therefore, the inequality can be recast in the more elegant way:

$$\max_{\substack{(u,v)\in H^m\times H^n \\ \|u_i\|_H, \|v_j\|_H\leq 1}} \sum_{i=1}^{m}\sum_{j=1}^{n} A_{i,j}\langle u_i, v_j\rangle_H \leq K_G^{\mathbb{K}} \left(\max_{\substack{(x,y)\in\mathbb{K}^m\times\mathbb{K}^n \\ \|x\|_\infty, \|y\|_\infty\leq 1}} \sum_{i=1}^{m}\sum_{j=1}^{n} A_{i,j}x_i x_j\right). \tag{2.2.2}$$

The main surprising feature of this theorem is that the constant is independent of $m, n \geq 1$ (refining the theorem, it means that the Grothendieck constant can seen be as a function of m, n, which, by Grothendieck's result, stays bounded as $m, n \to \infty$).

The exact values of the real and complex Grothendieck constants are unknown, but we have the bounds:

$$\frac{\pi}{2} < K_G^{\mathbb{R}} < \frac{\pi}{2\log(1+\sqrt{2})}$$

and

$$\frac{4}{\pi} < K_G^{\mathbb{C}} < e^{1-\gamma},$$

where γ is the famous Euler constant:

$$\gamma = \lim_{n\to\infty}\left(\sum_{k=1}^{n}\frac{1}{k} - \log(n)\right).$$

There are more precise bounds, but they are either complicated or false (using either elliptic function, or unpublished and only appearing as a private communication from 1984* in [20]). See the survey of Gilles Pisier for more information on this topic ([32]).

Definition 2.2.3. A gauge on X is real-valued function $G : X \to \mathbb{R}$ such that

1. $G(\lambda\, x) = \lambda\, G(x)$ for all $x \in X$ and $\lambda \geq 0$.
2. $G(x + y) \leq G(x) + G(y)$ for all $x, y \in X$.

Definition 2.2.4. Let X be a vector space over $\mathbb{K}$, where $\mathbb{K}$ is $\mathbb{R}$ or $\mathbb{C}$. Then, a semi-norm on X is a map $N : X \to \mathbb{R}_+$ such that

1. $N(\lambda x) = |\lambda| N(x)$ for all $x \in X, \lambda \in \mathbb{K}$.
2. $N(x + y) \leq N(x) + N(y)$ for all $x, y \in X$.

The only difference with a norm is that we do not require that $N(x) = 0$ if and only if $x = 0$.

In the rest of this section, we fix a field $\mathbb{K}$ that is either $\mathbb{R}$ or $\mathbb{C}$.

We can now move on to the definition of topological vector spaces.

Remark 2.2.5. Taking $\lambda = 2$ and $x = 0$, we deduce that $N(0) = 0$. In particular, for all $x \in X$, we have $0 = N(0) = N(x - x) \leq N(x) + N(-x) = 2N(x)$. Therefore, we see that the requirement that $N \geq 0$ is not necessary in the definition of a norm.

Definition 2.2.6 (Topological Vector Space)**.** A topological vector space on $\mathbb{K}$ is a vector space X on $\mathbb{K}$ equipped with a topology $\mathscr{T}$ such that the maps:

1. $X \times X \to X, (x, y) \mapsto x + y$;
2. $\mathbb{K} \times X \to X, (\lambda, x) \mapsto \lambda\, x$

are continuous.

Example 2.2.7. Although not commonly studied in first courses on integration due to their pathologies and surprising properties (such as the reverse Minkowski inequality for $p < 1$ and positive functions), L^p spaces of exponent $0 < p < 1$ do come up in some applications. Let $0 < p < \infty$, and consider the space

$$\mathscr{L}^p([0,1]) = \mathscr{M}([0,1],\mathbb{C}) \cap \left\{ f : \int_0^1 |f(t)|^p dt < \infty \right\},$$

where $\mathscr{M}([0,1],\mathbb{C})$ is the set of measurable functions from the unit interval $[0,1]$ into the complex plane $\mathbb{C}$. Then the following semi-metric (on L^p, since it is not a semi-metric on $\mathscr{L}^p$)

$$d_p(f,g) = \left(\int_0^1 |f(t) - g(t)|^p dt\right)^{\frac{1}{p}}$$

endows $\mathscr{L}^p$ with a topological vector space structure. When $p \geq 1$, the topology is induced by the semi-norm $N(f) = \|f\|_{\mathrm{L}^p([0,1])}$ (here, we did not take the quotient by the equivalence relation that identifies two functions that differ on a negligible set, so $\mathscr{L}^p$ is not a normed space, an *a fortiori* not a Banach space).

Recall that a semi-metric is a map $d : X \times X \to \mathbb{R}_+$ that satisfies all metric axioms but the triangle inequality (see Definition 1.1.8).

*What will people think of this *private communication* in 2084?

2.3 Locally Convex Topological Vector Spaces

We will not study the general properties of topological vector spaces for our main interest will be the smaller close of locally convex topological vector spaces.

Definition 2.3.1 (Locally convex topological vector spaces)**.** A locally convex topology vector space is a topological vector space that admits a basis of neighbourhoods of 0 made of convex sets.

Remarks 2.3.2.

1. L^p spaces with exponent $0 < p < 1$ are *not* locally convex, whilst L^p spaces for $1 \leq p \leq \infty$ is locally convex.
2. The space of continuous functions on a compact set equipped with the uniform convergence topology is locally convex.

Finally, we come to the statement that permits to define workable topology on locally convex topological vector spaces.

We first need a few elementary results after introduction some basic topological definitions.

Definition 2.3.3 (Absorbent and Balanced Sets)**.** A subset $A \subset X$ is called absorbent if for all $x \in X$, there exists $\lambda > 0$ such that $\lambda x \in X$.

A subset $A \subset X$ is called *balanced* if $\lambda x \in A$ for all $x \in A$ and $\lambda \in \mathbb{K}$ such that $|\lambda| \leq 1$.

Lemma 2.3.4. *Let X be a topological vector spaces. Then, there exist either a neighbourhood base at* 0 *comprised of sets that are all balanced and open,* or *balanced and closed.*

Proof. We only prove the existence of the neighbourhood of balanced open sets.

By the definition of a general topology, there exists a base of neighbourhoods at 0 comprised of open sets. Let U be an arbitrary neighbourhood of 0. By the second axiom of topological vector spaces (Definition 2.2.6), there exists $\varepsilon > 0$ and an open neighbourhood V of 0 such that for all $x \in V$ and $|\lambda| \leq \varepsilon$, we have λU. Let W be the balanced enveloped of εV (the smallest balanced set containing εV; since the arbitrary intersection of balanced sets is balanced, this is well-defined). Since V is open, we deduce that W is open. Indeed, the balanced envelope $W = \mathscr{B}(V)$ of V is explicitly given by

$$\mathscr{B}(V) = X \cap \{y = \lambda\, x : x \in V, \lambda \in \mathbb{K}, |\lambda| \leq 1\} = V \cup \bigcup_{\substack{\lambda \in \mathbb{K}^* \\ |\lambda| \leq 1}} \lambda\, V,$$

which is an arbitrary union of open sets thanks to the second axiom of topological vector spaces. Therefore, we deduce that for all $\lambda \in \mathbb{K}^*$ and for all open set Z, the set $\lambda\, Z$ is open.

Furthermore, we have $W \subset U$ by definition of V, which concludes the proof. □

Now, by Lemma 2.3.4, we deduce that there exists a neighbourhood base of 0 made of closed, convex, and balanced sets. Indeed, let U be a neighbourhood of 0, and let U_1 be a closed and balanced neighbourhood of 0 such that $U_1 \subset U$. Then, there exists a convex neighbourhood $U_2 \subset U_1$ of 0, and a closed and balanced neighbourhood $U_3 \subset U_2$ of 0. Finally, taking the closure of the convex envelope of U_3 yield a closed set $V \subset U$ that is a neighbourhood of 0.

Letting now $\{U_i\}_{i \in I}$ be a neighbourhood base of 0 made of closed, convex and balanced subsets of X. We will need the concept of Minkowski gauge and a few technical results.

Theorem 2.3.5. *Let C be a convex and absorbent subset of X which contains* 0*. Define the following real-valued function $G : X \to \mathbb{R}$ by*

$$G(x) = \inf \mathbb{R}_+^* \cap \left\{\lambda : \lambda^{-1} x \in C\right\}.$$

Then, G is a positive gauge on X such that

$$G^{-1}(]-\infty, 1[) \subset C \subset G^{-1}(]-\infty, 1]). \tag{2.3.1}$$

We say that G is the Minkowsky gauge of C.

Proof. Since C is absorbent, the set $\mathbb{R}_+^* \cap \{\lambda : \lambda^{-1}x \in C\}$ is non-empty, which shows that G is well-defined and satisfies $G(x) \geq 0$. Now, for all $\lambda > 0$, we have $G(\lambda x) = \inf \mathbb{R}_+^* \cap \{\mu : \mu^{-1}(\lambda x) \in C\} = \lambda G(x)$ by an immediate change of variable. Now, by convexity of C, if $\lambda, \mu > 0$ are such that $\lambda^{-1}x \in C$ and $\mu^{-1}y \in C$, we have

$$(\lambda+\mu)^{-1}(x+y) = \frac{\lambda}{\lambda+\mu}(\lambda^{-1}x) + \frac{\mu}{\lambda+\mu}(\mu^{-1}y) \in C.$$

Taking the infimum in λ and μ, we deduce that $G(x+y) \leq G(x) + G(y)$.

Finally, we need only prove the first inclusion, for the other one is trivial ($x = 1^{-1}x \in G^{-1}(]-\infty, 1])$ for all $x \in C$). If $G(x) < 1$, there exists $0 < \lambda < 1$ such that $\lambda^{-1}x \in C$. Since $0 \in C$ and C is convex, we have $x = \lambda(\lambda^{-1}x) + (1-\lambda)0 \in C$. This concludes the proof of the theorem. □

We now introduce a new definition to state a useful corollary.

Definition 2.3.6. We say that a subset $A \subset X$ is *open* (resp. *closed*) *in rays* if for all $x \in X$, the set $I_x = \mathbb{R}_+^* \cap \{\lambda : \lambda^{-1}x \in A\}$ is open (resp. closed) relative to the interval $I = \mathbb{R}_+^* = (0, \infty)$.

Corollary 2.3.7. *Let C be a convex and absorbent subset of X containing 0, and let G be its Minkowski gauge. If C is open (resp. closed) in rays, then $C = G^{-1}(]-\infty, 1[)$ (resp. $C = G^{-1}(]-\infty, 1])$). In either case, C is balanced if and only if G is a semi-norm on X.*

Proof. We only treat the case of C open in rays. By (2.3.1), we need only prove that $C \subset G^{-1}(]-\infty, 1[)$. Let $x \in C$. Then $1 \in I_x$, and since I_x is open in rays, we deduce that there exists $0 < \lambda < 1$ such that $\lambda \in I_x$. In particular, we have $G(x) \leq \lambda < 1$.

If G is a semi-norm, then $C = G^{-1}(]-\infty, 1[)$ is trivially balanced, for $G(\lambda x) = |\lambda| G(x) \leq G(x) < 1$ for all $x \in C$ and $|\lambda| \leq 1$. Conversely, if C is balanced, for all $\varepsilon > 0$ fixed and $x \in X$, we have $(G(x)+\varepsilon)^{-1}x \in C$, which implies in particular that for all $\lambda \in \mathbb{K}^*$, we have $(|\lambda|(G(x)+\varepsilon))^{-1}(\lambda x) \in C$, so that $G(\lambda x) \leq |\lambda|(G(x)+\varepsilon)$. Letting $\varepsilon \to 0$, we deduce that

$$G(\lambda x) \leq |\lambda| G(x).$$

Replacing $\lambda \in \mathbb{K}$ by λ^{-1}, and x by λx, we get

$$G(x) \leq |\lambda|^{-1} G(\lambda x),$$

which shows that $G(\lambda x) = |\lambda| G(x)$. □

Thanks to the previous results, we deduce that for all $i \in I$, if N_i is the Minkowski gauge of U_i, then N_i is a semi-norm, and $U_i = N_i^{-1}(]-\infty, 1])$. Therefore, we have proved that the family of semi-norms N_i on X gives a neighbourhood base at 0. Notice that the family $\{N_i^{-1}(]-\infty, 1[)\}_{i \in I}$ yields an open basis of neighbourhood of 0.

Conversely, we immediately check that an *arbitrary* family of semi-norms on X endows X with a structure of locally convex topological vector space. Such a family is called a defining family.

Summarising the previous discussion, we now state the following theorem.

Theorem 2.3.8. *Let X be a locally convex topological vector space. Then, there exists a family of semi-norms $\{N_i\}_{i \in I}$ on X such that $\{N_i^{-1}(]-\infty, 1[)\}_{i \in I}$ (resp. $\{N_i^{-1}(]-\infty, 1])\}_{i \in I}$) is a neighbourhood base at 0 of open (resp. closed) convex, and balanced subsets of X.*

Furthermore, X is separated if and only if

$$\sup_{i \in I} N_i(x) > 0 \quad \text{for all } x \in X \setminus \{0\}. \tag{2.3.2}$$

Proof. We have already proven the first part, so we need only prove the second part. Since X is a vector space, X is separated if and only if for all $x_0 \in X$, there exists two open sets U_1 and U_2 such that $0 \in U_1$, $x_0 \in U_2$ and $U_1 \cap U_2 = \varnothing$. Assume that X is separated. Without loss of generality, we can assume that $U_1 = X \cap \{x : N_1(x) < 1\}$ and $U_2 = X \cap \{x : N_2(x_0 - x) < 1\}$, where N_1 and

N_2 are semi-norms. Since U_1 and U_2 have empty intersection, we deduce in particular that $0 \notin U_2$, *i.e.* $N_2(x_0) = N_2(x_0 - 0) \geq 1$, which shows the second implication. If there exists $i \in I$ such that $N_i(x_0) = \varepsilon > 0$. Therefore, $U_1 \cap \left\{x : N_i(x) < \frac{\varepsilon}{2}\right\}$ and $U_1 = X \cap \left\{x : N_i(x) > \frac{\varepsilon}{2}\right\}$ are disjoint open neighbourhoods of 0 and x_0 respectively. □

2.4 Dual Space of a Topological Vector Space

On a topological vector space X, we define—as in the case of normed spaces—the space X' as the space of continuous linear forms $X \to \mathbb{R}$. If X, Y are topological vector spaces, we let $\mathscr{L}(X, Y)$ be the vector space of continuous linear maps from X to Y.

Finally, we remark that the definition we gave of weak topology for the dual of a normed space also makes sense for the dual of a topological vector space, and that allow us to define the natural topology on distributions. Furthermore, provided that X is locally convex, if $\{N_i\}_{i\in I}$ is a family of semi-norms defining the topology of X, we can define a strong topology on X', denoted by $\beta(X', X)$, such that for all sequence $\{f\}_{n\in\mathbb{N}} \in X'$, we say that $\{f_n\}_{n\in\mathbb{N}}$ converges to $f \in X'$ if and only if for all *bounded* subset $B \subset X$, we have

$$\sup_{i\in I}\sup_{x\in B} N_i(f_n(x) - f(x)) = 0.$$

Both topologies $\sigma(X', X)$ and $\beta(X', X)$ endow X with the topology of a locally convex topological space, as one immediately verifies.

Remark 2.4.1. Contrary to the familiar case of Banach spaces where dual spaces generally offer a lot of information, the dual of a locally convex topological vector space might be reduced to $\{0\}$, as the example of $L^p([0,1])$ shows when $0 < p < 1$.

2.5 Distributions

2.5.1 General Definition and Topologies

Finally, we can move on to the definition of distributions.

Let Ω be an open set of $\mathbb{R}^d$, and define the space space of smooth functions of compact support by

$$\mathscr{D}(\Omega) = C_c^\infty(\Omega) = C^\infty(\Omega) \cap \{\varphi : \mathrm{supp}\,(\varphi) \subset\subset \Omega\}$$

We can also see it as the inductive limit of the spaces

$$\mathscr{D}_K(\Omega) = C^\infty(\Omega) \cap \{\varphi : \mathrm{supp}(\varphi) \subset K\},$$

where $K \subset \Omega$ ranges all compact subsets of Ω, that we will define below. Notice that

$$\mathscr{D}(\Omega) = \bigcup_{\substack{K\subset\Omega \\ K \text{ compact}}} \mathscr{D}_K(\Omega)$$

For all $m \in \mathbb{N}$ and compact subset $K \subset \mathscr{D}(\Omega)$, define the semi-norm on $\mathscr{D}(\Omega)$ by

$$\|\varphi\|_{m,K} = \sup_{|\alpha|\leq m} \|D^\alpha \varphi\|_{\mathrm{L}^\infty(K)}.$$

This topology on $\mathscr{D}(\Omega)$ gives the latter space the If K is fixed, this family of semi-norms endows $\mathscr{D}_K(\Omega)$ with a Fréchet space structure, where the metric is given by

$$d_K(\varphi, \psi) = \sum_{m\in\mathbb{N}} \frac{1}{2^m} \frac{\|\varphi - \psi\|_{m,K}}{1 + \|\varphi - \psi\|_{m,K}}.$$

Now, if $\{K_n\}_{n\in\mathbb{N}}$ is an exhaustive sequence of compact sets of Ω, then the topological vector space $(\mathscr{D}(\Omega), \{\|\cdot\|_{m\in\mathbb{N}}\})$ can be equipped with a distance:

$$d(\varphi,\psi) = \sum_{m\in\mathbb{N}}\sum_{n\in\mathbb{N}} \frac{1}{2^{m+n}} \frac{\|\varphi-\psi\|_{m,K_n}}{1+\|\varphi-\psi\|_{m,K_n}},$$

making it a Fréchet space. However, the dual of such a space is *not* the space of distributions $\mathscr{D}'(\Omega)$, but the space of distributions with compact support $\mathscr{E}'(\Omega)$, and one checks that $\mathscr{E}(\Omega) = (\mathscr{D}(\Omega), \{\|\cdot\|_{m\in\mathbb{N}}\}) = C^\infty(\Omega)$, equipped with the compact-open topology: the topology of uniform convergence on all compact for each $D^\alpha\varphi$ ($\alpha \in \mathbb{N}^d$).

The topology on $\mathscr{D}(\Omega)$ is given by an *inductive limiting* procedure.

Definition 2.5.1. Let $\{X_i\}_{i\in I}$ a family of topological vector spaces, X be a vector space, $f_i : X_i \to X$ be a linear map. Then, the inductive (or final) topology on X is the biggest *locally convex* topology on X such that each map f_i is continuous, and we write $X = \varinjlim X_i$.

The same definition would make sense in the category of sets instead of the category (locally convex) topological vector spaces.

Remark 2.5.2.

1. We see that this topology is defined in a very similar way than the weak topology, although the base and target spaces are reverted—we also take the *finest topology*, not the coarsest one, but in the category of locally convex topologies (otherwise, we would end up taking the discrete topology on X).
2. Explicitly, a subset $U \subset X$ is open if and only if $f_i^{-1}(U)$ is open for all $i \in I$.
3. Although the notation $X = \varinjlim X_i$ might seem frightening, in practice, it poses no issues to check if a set is open for the inductive limit topology.

Now, take as previously an exhaustive sequence of compact sets $\{K_n\}_{n\in\mathbb{N}} \subset \Omega$. Then, as previously, we have

$$\mathscr{D}(\Omega) = \bigcup_{n\in\mathbb{N}} \mathscr{D}_{K_n}(\Omega),$$

and we endow $\mathscr{D}(\Omega)$ with the inductive limit topology associated to the canonical injections $\iota_n : K_n \hookrightarrow \Omega$. The space $\mathscr{D}(\Omega)$ will always be equipped with this topology that makes it an inductive limit of Fréchet spaces. In practice, the topology can be characterised as follows. For all $\{\varphi_n\}_{n\in\mathbb{N}}$, we have $\varphi_n \underset{n\to\infty}{\longrightarrow} \varphi \in \mathscr{D}(\Omega)$ if and only if there exists a compact set $K \subset \Omega$ such that $\mathrm{supp}(\varphi_n) \subset K$ for all $n \in \mathbb{N}$, and for all $m \in \mathbb{N}$, we have

$$\lim_{n\to\infty} \|\varphi_n - \varphi\|_{m,K} = 0.$$

This result can be showed by exhibiting a suitable family of semi-norms that define the inductive topology on $\mathscr{D}(\Omega)$. We first select a sequence $\{\Omega_n\}_{n\in\mathbb{N}}$ of open and relatively compact subsets of Ω such that $\Omega_0 = \varnothing$, $\overline{\Omega_n} \subset \Omega_{n+1}$ and $\displaystyle\bigcup_{n\in\mathbb{N}} \Omega_n = \Omega$. Then, for all decreasing sequence $\boldsymbol{\varepsilon} = \{\varepsilon_n\}_{n\in\mathbb{N}} \subset (0,\infty)$ converging to 0, and for all increasing sequence $\boldsymbol{m} = \{m_n\}_{n\in\mathbb{N}} \subset \mathbb{N}$ diverging to ∞, we define the semi-norm

$$\|\varphi\|_{\boldsymbol{\varepsilon},\boldsymbol{m}} = \sup_{n\in\mathbb{N}} \left(\frac{1}{\varepsilon_n} \sup_{|\alpha|\leq m_n} \|D^\alpha\varphi\|_{\mathrm{L}^\infty(\Omega\setminus\overline{\Omega_n})}\right).$$

First, we need to shows that for all $(\boldsymbol{\varepsilon},\boldsymbol{m})$ as above, $\|\cdot\|_{\boldsymbol{\varepsilon},\boldsymbol{m}}$ is continuous. By the definition of the inductive limit topology, we need only show that the restriction of $\|\cdot\|_{\boldsymbol{\varepsilon},\boldsymbol{m}}$ to any compact $K \subset \Omega$ is continuous. However, for all compact $K \subset \Omega$, there exists $N \in \mathbb{N}$ such that $K \cap \Omega \setminus \overline{\Omega_n} = \varnothing$ for all $n \geq N$. In particular, we have for all $\varphi \in \mathscr{D}_K(\Omega)$

$$\|\varphi\|_{\boldsymbol{\varepsilon},\boldsymbol{m}} = \sup_{1\leq n\leq N} \frac{1}{\varepsilon_n} \sup_{|\alpha|\leq m_n} \|D^\alpha\varphi\|_{\mathrm{L}^\infty(K)} = \varepsilon_N^{-1} \sup_{|\alpha|\leq m_N} \|D^\alpha\varphi\|_{\mathrm{L}^\infty(K)} = \varepsilon_N^{-1} \|\varphi\|_{K,m_N}.$$

Conversely, if U is a closed convex convex neighbourhood of 0 for the inductive limit topology, then we can find $\|\cdot\|_{\boldsymbol{\varepsilon},\boldsymbol{m}}$ such that $U \supset \mathscr{D}(\Omega) \cap \left\{\varphi : \|\varphi\|_{\boldsymbol{\varepsilon},\boldsymbol{m}} \leq 1\right\}$. By definition of the inductive limit topology, for all $n \in \mathbb{N}$, there exists $m_n \in \mathbb{N}$ and $\delta_n > 0$ such that $\|\varphi\|_{\Omega_{n+2},m_n} \leq \delta_n$ implies that $\varphi \in U$. Furthermore, we can assume without loss of generality that $\{\delta_n\}_{n\in\mathbb{N}} \subset (0,\infty)$ is decreasing and $\boldsymbol{m} = \{m_n\}_{n\in\mathbb{N}} \subset \mathbb{N}$ is increasing. Let $\{\beta_n\}_{n\in\mathbb{N}}$ be a smooth partition of unity, such that $\mathrm{supp}(\beta_n) \subset \Omega_{n+2} \setminus \overline{\Omega_n}$ and $\displaystyle\sum_{n\in\mathbb{N}} \beta_n = 1$. Then, for all $\varphi \in \mathscr{D}(\Omega)$, we have

$$\varphi = \sum_{n\in\mathbb{N}} \frac{1}{2^{n+1}}(2^{n+1}\beta_n\varphi),$$

and by convexity of φ we deduce that $\varphi \in U$ provided that $2^{n+1}\beta_n\varphi$ belongs to U. Furthermore, for all $n \in \mathbb{N}$, there exists $0 < C_n < \infty$ such that

$$\left\|2^{n+1}\beta_n\varphi\right\|_{m_n,\Omega_{n+2}} \leq C_n \|\varphi\|_{m_n\Omega_{n+2}}.$$

We can also assume that $\{C_n\}_{n\in\mathbb{N}} \subset (0,\infty)$ is an increasing sequence. In particular, taking $\varepsilon_n = C_n^{-1}\delta_n$, we have $\mathscr{D}(\Omega) \cap \left\{\varphi : \|\varphi\|_{\boldsymbol{\varepsilon},\boldsymbol{m}} \leq 1\right\} \subset U$, which concludes the proof of the following result.

Theorem 2.5.3. *Let Ω be an open subset of $\mathbb{R}^d$, and let $\mathscr{D}(\Omega) = C_c^\infty(\Omega)$ the space of smooth functions of compact support on Ω, and for all compact set $K \subset \Omega$, let $\mathscr{D}_K(\Omega) = C^\infty(\Omega) \cap \{\varphi : \mathrm{supp}(\varphi) \subset K\} \subset \mathscr{D}(\Omega)$, that we endow with the topology of uniform convergence of all derivatives, that makes it a Fréchet spaces with defining semi-norms given by $\left\{\|\cdot\|_{m,K}\right\}_{m\in\mathbb{N}}$, where for all $m \in \mathbb{N}$,*

$$\|\varphi\|_{m,K} = \sup_{|\alpha|\leq m} \|D^\alpha\varphi\|_{\mathrm{L}^\infty(K)} \quad \text{for all } \varphi \in \mathscr{D}_K(\Omega).$$

Then, the inductive topology on $\mathscr{D}(\Omega)$ is the inductive topology associated to all canonical injections $f_K : \mathscr{D}_K(\Omega) \to \mathscr{D}(\Omega)$, and we note

$$\mathscr{D}(\Omega) = \varinjlim \mathscr{D}_K(\Omega).$$

The resulting space is a separated *locally convex topological vector space. For all exhaustive sequence of relatively compact subsets $\{\Omega_n\}_{n\in\mathbb{N}} \subset \Omega$ such that $\Omega_0 = \varnothing$ and $\overline{\Omega_n} \subset \Omega_{n+1}$ for all $n \in \mathbb{N}$, a defining family of semi-norms is given by $\|\cdot\|_{\boldsymbol{\varepsilon},\boldsymbol{m}}$, where $\boldsymbol{\varepsilon} = \{\varepsilon_n\}_{n\in\mathbb{N}} \subset (0,\infty)$ ranges over all decreasing sequences converging to 0, and $\boldsymbol{m} = \{m_n\}_{n\in\mathbb{N}}$ of ranges over all increasing sequences diverging to ∞, and for all $\varphi \in \mathscr{D}(\Omega)$, we have*

$$\|\varphi\|_{\boldsymbol{\varepsilon},\boldsymbol{m}} = \sup_{n\in\mathbb{N}} \left(\frac{1}{\varepsilon_n} \sup_{|\alpha|\leq m_n} \|D^\alpha\varphi\|_{\mathrm{L}^\infty(\Omega\setminus\overline{\Omega_n})}\right). \tag{2.5.1}$$

Remark 2.5.4. We stress out that this topology is *not* metrisable and does *not* have a countable base of neighbourhoods at 0. It is not countable for the family of decreasing sequences converging to 0 has the cardinal of the continuum.

Finally, we can define the space of distributions as follows.

Definition 2.5.5. The topological dual space of $\mathscr{D}(\Omega)$ is called the space of distributions, and we denote it by $\mathscr{D}'(\Omega)$. We endow it with the weak topology $\sigma(\mathscr{D}(\Omega), \mathscr{D}'(\Omega))$.

We spoke about *weak topology* although we meant *weak topology.* However, since $\mathscr{D}'(\Omega)$ is a reflexive space, the weak topology and the weak $*$ topology coincide, and this terminology is rather standard.

Remark 2.5.6. Analogously, we define complex-valued, or vector-valued definition by taking the product spaces of distributions.

Thanks to Proposition 1.4.4, we deduce that a sequence $\{T_n\}_{n\in\mathbb{N}} \subset \mathscr{D}'(\Omega)$ converges to an element $T \in \mathscr{D}'(\Omega)$ if and only if

$$T_n(\varphi) \underset{n\to\infty}{\longrightarrow} T(\varphi) \quad \text{for all } \varphi \in \mathscr{D}(\Omega).$$

Definition 2.5.7 (Support)**.** By definition of the inductive topology, we deduce that a linear form $T : \mathscr{D}(\Omega) \to \mathbb{R}$ belongs to $\mathscr{D}'(\Omega)$ if and only if for all compact $K \subset \Omega$, there exists $C_K < \infty$ and $m_K \in \mathbb{N}$ such that

$$|T(\varphi)| \leq C_K \|\varphi\|_{m_K,K} = C_K \sup_{|\alpha| \leq m_K} \|D^\alpha \varphi\|_{\mathrm{L}^\infty(K)}\,. \tag{2.5.2}$$

The smallest such constant C_K is denoted $\|T\|_K$, and the smallest integer $m_K \in \mathbb{N}$ such that (2.5.2) holds is called the order of T on K, denoted by $\mathrm{ord}_K(T)$. If

$$m = \sup_{\substack{K \subset \Omega \\ K \text{ compact}}} \mathrm{ord}_K(T)$$

is finite, we say that T is a distribution of order $\mathrm{ord}(T) = m$. We denote by $\mathscr{D}'^m(\Omega)$ the space of distributions of order $m \in \mathbb{N}$.

Remark 2.5.8. Analogously, we define complex-valued, or vector-valued definition by taking the product spaces of distributions. Notice that distributions of support 0 are Radon measures.

Examples 2.5.9. **1**. If $f \in L^1_{\mathrm{loc}}(\Omega)$, then the distribution $T = f$ defined by integration such that

$$T(\varphi) = \int_\Omega f\,\varphi dx \quad \text{for all } \varphi \in \mathscr{D}(\Omega)$$

is a distribution of order 0, with $\|T\|_K = \|f\|_{\mathrm{L}^1(K)}$. More generally, if $T = \mu$ is a real Radon measure, then

$$T(\varphi) = \int_\Omega \varphi\, d\mu$$

is also a distribution of order 0, such that $\|T\|_K = \mu(K)$. An important example is the Dirac mass at $x_0 \in \Omega$, given by

$$\delta_{x_0}(\varphi) = \varphi(x_0).$$

2. The Dirac mass δ_a such that $\delta_a(\varphi) = \varphi(a)$ $(a \in \Omega)$ is a very important distribution (a measure, in fact), that will have a crucial importance in several theorems for reasons that will be made clear by convolution and Fourier transform.

3. Anticipating on the next section, for all $a \in \mathbb{R}$ define for all $n \in \mathbb{N}$ the distribution $\delta_a^{(n)} \in \mathscr{D}'(\mathbb{R})$ by $\delta_a^{(n)}(\varphi) = (-1)^n \varphi^{(n)}(a)$. Then, the following distribution

$$T = \sum_{n \in \mathbb{N}} \delta_n^{(n)}$$

has infinite order. Indeed, we see easily that for all $n \in \mathbb{N}$, the restriction of T to $B(0, n + \frac{1}{2})$ has order n.

4. The principal value integral (at 0) of a function $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ such that $f \in L^1_{\mathrm{loc}}(\mathbb{R} \setminus \{0\})$ is defined by

$$\mathrm{p.v.} f(\varphi) = \lim_{\varepsilon \to 0} \int_{\mathbb{R} \setminus [-\varepsilon, \varepsilon]} f(x)\varphi(x) dx.$$

Under suitable conditions on f, $\mathrm{p.v.}f$ is a well-defined distribution, known as Cauchy principal value. Take $f(x) = \frac{1}{x}$. Then, by oddness of f, for all $0 < \varepsilon < R < \infty$, we have

$$\left\langle \mathrm{p.v.}\frac{1}{x}, f \right\rangle = \int_{\mathbb{R} \setminus [-\varepsilon, \varepsilon]} \varphi(x) \frac{dx}{x} = \int_{[-R, -\varepsilon] \cup [\varepsilon, R]} (\varphi(x) - \varphi(0)) \frac{dx}{x} + \int_{\mathbb{R} \setminus [-R, R]} \varphi(x) dx,$$

and since φ is of class C^1, the function $\frac{\varphi(x)-\varphi(0)}{x}$ is bounded at 0, and for all $0 < R < \infty$, we have

$$\left\langle \text{p.v.}\frac{1}{x}, f\right\rangle = \int_{-\infty}^{-R}\varphi(x)\frac{dx}{x} + \int_{-R}^{R}(\varphi(x)-\varphi(0))\frac{dx}{x} + \int_{R}^{\infty}\varphi(x)\frac{dx}{x}.$$

Taking $R > 0$ large enough such that $\text{supp}(\varphi) \subset [-R, R]$, we deduce by Fubini's theorem that

$$\begin{aligned}\left\langle \text{p.v.}\frac{1}{x}, \varphi\right\rangle &= \int_{-R}^{R}(\varphi(x)-\varphi(0))\frac{dx}{x}\\ &= -\int_{-R}^{0}\left(\int_{x}^{0}\varphi'(t)dt\right)\frac{dx}{x} + \int_{0}^{R}\left(\int_{0}^{x}\varphi'(t)dt\right)\frac{dx}{x}. \qquad (2.5.3)\end{aligned}$$

We first compute

$$\begin{aligned}\int_{-R}^{0}\left(\int_{x}^{0}\varphi'(t)dt\right)\frac{dx}{x} &= \int_{-R}^{0}\left(\int_{-R}^{0}\varphi'(t)\mathbf{1}_{\{x\leq t\leq 0\}}dt\right)\frac{dx}{x} = \int_{-R}^{0}\varphi'(t)\left(\int_{-R}^{t}\frac{dx}{x}\right)dt\\ &= \int_{-R}^{0}\varphi'(t)\log\left(\frac{t}{-R}\right)dt. \qquad (2.5.4)\end{aligned}$$

Likewise, we have

$$\int_{0}^{R}\left(\int_{0}^{x}\varphi'(t)dt\right)\frac{dx}{x} = \int_{0}^{R}\varphi'(t)\log\left(\frac{R}{t}\right)dt, \qquad (2.5.5)$$

which implies that

$$\begin{aligned}\left\langle \text{p.v.}\frac{1}{x}, \varphi\right\rangle &= -\int_{-R}^{R}\varphi'(t)\log\left(\frac{|t|}{R}\right)dt = \int_{-R}^{R}\varphi'(t)\log|t|dt + \log(R)\int_{-R}^{R}\varphi'(t)dt\\ &= -\int_{\mathbb{R}}\varphi'(t)\log|t|dt, \qquad (2.5.6)\end{aligned}$$

since $\text{supp}(\varphi) \subset [-R, R]$. This expression easily shows that p.v.$\frac{1}{x}$ has order 1, and that the distributional derivative (as defined in the next section) of $x \mapsto \log|x|$ is p.v.$\frac{1}{x}$. Indeed, for all $\varphi \in \mathscr{D}(\mathbb{R})$, we have

$$\left|\int_{\mathbb{R}}\varphi'(t)\log|t|dt\right| \leq \left(\int_{\text{supp}(\varphi')}\log|t|dt\right)\|\varphi'\|_{L^\infty(\mathbb{R})},$$

which shows that p.v.$\frac{1}{x}$ has order at most 1. If this distribution had order 0, it would extend to a Radon measure. We will therefore exhibit a bounded sequence $\{\varphi_n\}_{n\in\mathbb{N}}$ in $C_c(\mathbb{R})$ such that $\langle \text{p.v.}\frac{1}{x}, \varphi_n\rangle$ diverges. Now, let $\{\varphi_n\}_{n\in\mathbb{N}} \subset C_c^0(\mathbb{R})$ such that φ_n is odd, $\text{supp}(\varphi_n) \subset [-2, 2]$, $\varphi_n = \varphi_0$ on $[-2, 2] \setminus [-1, 1]$,

$$\begin{cases}\varphi_n(x) = -1 & \text{for all } -1 \leq x \leq -\dfrac{1}{n}\\ \varphi_n(x) = nx & \text{for all } -\dfrac{1}{n} \leq x \leq \dfrac{1}{n}\\ \varphi_n(x) = 1 & \text{for all } \dfrac{1}{n} \leq x \leq 1.\end{cases}$$

This sequence is bounded in $C_c(\mathbb{R})$ since $\text{supp}(\varphi_n) \subset [-2, 2]$, and $|\varphi_n| \leq 1$. However, we have

$$\begin{aligned}\left\langle \text{p.v.}\frac{1}{x}, \varphi_n\right\rangle &= \lim_{\varepsilon\to 0}\left(2\int_{1}^{2}\varphi_0(x)\frac{dx}{x} + 2\int_{\frac{1}{n}}^{1}\frac{dx}{|x|} + 2\int_{\varepsilon}^{\frac{1}{n}}n\,dx\right)\\ &= 2\int_{1}^{2}\varphi_0(x)\frac{dx}{x} + 2 + 2\log(n) \underset{n\to\infty}{\longrightarrow} \infty.\end{aligned}$$

Therefore, we deduce that p.v.$\frac{1}{x}$ is a distribution of order exactly 1. By introducing $\varphi(x) - \varphi(-x)$, give an alternative proof of the above results.

The first basic property of distributions is the multiplication by smooth functions. Recall that $\mathscr{E}(\Omega) = C^\infty(\Omega)$ equipped with the compact-open topology (which makes it a Fréchet space).

Definition 2.5.10. For all $T \in \mathscr{D}'(\Omega)$ and $f \in \mathscr{E}(\Omega)$, the product $S = f\,T$ defined by

$$\langle f\,T, \varphi\rangle = \langle T, f\,\varphi\rangle \quad \text{for all } \varphi \in \mathscr{D}(\Omega)$$

is a distribution such that for all compact $K \subset \Omega$, we have

$$\mathrm{ord}_K(f\,T) \leq \mathrm{ord}_K(\varphi). \tag{2.5.7}$$

Remark 2.5.11. That $f\,T \in \mathscr{D}'(\Omega)$ follows immediately since $f\varphi \in \mathscr{D}(\Omega)$ for all $(f, \varphi) \in \mathscr{E}(\Omega) \times \mathscr{D}(\Omega)$, and the property of order is trivial by Leibniz formula.

Example 2.5.12. We have $x \cdot \left(\mathrm{p.v.}\frac{1}{x}\right) = 1$. Indeed, for all $\varphi \in \mathscr{D}(\Omega)$,

$$\left\langle x \cdot \left(\mathrm{p.v.}\frac{1}{x}\right), \varphi\right\rangle = \left\langle \mathrm{p.v.}\frac{1}{x}, x\,\varphi\right\rangle = \lim_{\varepsilon \to 0} \int_{\mathbb{R}\setminus[-\varepsilon,\varepsilon]} \Big(x\,\varphi(x)\Big)\frac{dx}{x} = \int_{\mathbb{R}} \varphi(x)dx = \langle 1, \varphi\rangle\,.$$

We saw above that $\mathrm{p.v.}\frac{1}{x}$ was a derivative of a Radon measure (of a function, more precisely). This fact is general to distributions of finite support, but we will prove it below (in Theorem 2.5.32) once we define the notion of support.

The previous example also poses the question of division of distributions. Obviously, the division by a non-zero function is always possible, and we will concentrate on isolated zeroes. This question has immediate applications to partial different equations of constant coefficients.

2.5.2 Division of Distributions by a Function with Isolated Singularities

Assume that $d = 1$. Then, the problem is reduced to the following one. Let $m \in \mathbb{N}^*$, and $T \in \mathscr{D}'(\mathbb{R})$ (the problem is local, so we can consider distributions on the whole space). We look for a distribution $S \in \mathscr{D}'(\mathbb{R})$ such that

$$x^m S = T.$$

By induction, it suffices to treat the case $m = 1$. The existence follows by Hahn-Banach theorem. Indeed, if S is a solution, then for all $\varphi \in \mathscr{D}(\mathbb{R})$ such that $\varphi(0) = 0$, we have $\psi = x^{-1}\varphi \in \mathscr{D}(\mathbb{R})$, and $S(\varphi) = T(\psi)$. Therefore, we define S by this relation on the hyperplane $H = \mathscr{D}(\mathbb{R}) \cap \{\varphi : \varphi(0) = 0\}$. By the Hahn-Banach theorem, S admits a continuous extension to $\mathscr{D}(\mathbb{R})$, such that

$$xS = T.$$

If S_1 and S_2 are two solution, then $X = S_1 - S_2$ satisfies

$$xX = 0.$$

Therefore, $X(\varphi) = 0$ for all $\varphi \in H$, which implies that $X = c\,\delta_0$ for some $c \in \mathbb{C}$. If $m > 0$, we have $X = \displaystyle\sum_{|\alpha| \leq m} c_\alpha D^\alpha \delta_0$ ($c_\alpha \in \mathbb{C}$). Therefore, all solutions differ by Dirac masses (and their derivatives) at 0.

This approach using a weak form of axiom of choice† is not very satisfying, and we will see below the direct method of Hadamard's *finite part* to deal with radial weights.

2.5.3 Fine Properties of Distribution Spaces

In order to understand the strong topology on $\beta(\mathscr{D}(\Omega), \mathscr{D}'(\Omega))$ on $\mathscr{D}'(\Omega)$, we need to characterise the bounded sets of $\mathscr{D}(\Omega)$. This characterisation and its proof are surprisingly easy.

†Hahn-Banach theorem is strictly weaker than the axiom of choice, but Hahn-Banach and Krein-Milman theorem imply the axiom of choice [8].

Theorem 2.5.13. *A subset $B \subset \mathscr{D}(\Omega)$ is bounded in the inductive limit topology if and only if it is contained and bounded in $\mathscr{D}_K(\Omega)$ for some compact $K \subset \Omega$.*

Proof. By definition of the inductive limit topology, any bounded set in $\mathscr{D}_K(\Omega)$ is bounded in $\mathscr{D}(\Omega)$. Conversely, let B be a bounded set of $\mathscr{D}(\Omega)$. If $B \notin \mathscr{D}_K(\Omega)$ for some compact $K \subset \Omega$, then for all $n \in \mathbb{N}$, there exists $\varphi_n \in B$ such that and $x_n \in \Omega \setminus \overline{\Omega_n}$ such that $\varphi_n(x_n) \neq 0$. Defining $\varepsilon_n = n^{-1}|\varphi_n(x_n)|$, the semi-norm $\|\cdot\|_{\boldsymbol{\varepsilon},\{n\}_{n\in\mathbb{N}}}$ would be unbounded on B, since

$$\frac{1}{\varepsilon_n}\sup_{\varphi\in B}\sup_{|\alpha|\leq n}\|D^\alpha\varphi\|_{\mathrm{L}^\infty(\Omega\setminus\overline{\Omega_n})} \geq \frac{1}{\varepsilon_n}|\varphi_n(x_n)| = n \underset{n\to\infty}{\longrightarrow} \infty.$$

This concludes the proof of the theorem. □

Remark 2.5.14. This theorem permits to make precise what the convergence in $\beta(\mathscr{D}(\Omega), \mathscr{D}'(\Omega))$ means. See [13].

Theorem 2.5.15. *The space $\mathscr{D}(\Omega)$ is complete.*

Proof. We only show the sequential completeness (the proof would be virtually unchanged for general completeness, but force one to introduce uniform structures and filters). Let $\{\varphi_n\}_{n\in\mathbb{N}}$ be a Cauchy sequence. Then, for all $(\boldsymbol{\varepsilon}, \boldsymbol{m})$ there exists $N \in \mathbb{N}$ such that

$$\|\varphi_n - \varphi_m\|_{\boldsymbol{\varepsilon},\boldsymbol{m}} \leq 1 \quad \text{for all } m, n \geq N.$$

In particular, we deduce that $D^\alpha\varphi_n$ converges locally uniformly for all $|\alpha| \leq m$ towards some function $D^\alpha\varphi$, and we need only show that φ has compact support. Taking $\boldsymbol{m} = \{0\}_{n\in\mathbb{N}}$, we deduce that for all $\boldsymbol{\varepsilon}$,

$$|\varphi_n(x) - \varphi_m(x)| \leq \varepsilon_l \quad \text{for all } x \in \Omega \setminus \overline{\Omega_l}$$

for all $m, n \geq N$. In particular, we have $|\varphi_n(x) - \varphi(x)| \leq \varepsilon_l$ for all $x \in \Omega \setminus \overline{\Omega_n}$. Since φ_N has compact support, we deduce that $|\varphi(x)| \leq \varepsilon_l$ for all $x \in \Omega \setminus \overline{\Omega_l}$ for l large enough. Therefore, φ has compact support choosing if $\varphi(x_l) \neq 0$ for some $x_l \in \Omega \setminus \overline{\Omega_l}$ the value $\varepsilon_l = \frac{1}{2}|\varphi(x_l)|$. □

Remark 2.5.16. The general proof works almost identically, but needs to use the notion of Cauchy filters that will not be used elsewhere. Notice that in a topological vector space X, if $\{U_i\}_{i\in I}$ is any neighbourhood base at 0, then a uniform structure is given by

$$\Delta = \{\Delta(U_i)\}_{i\in I},$$

where for all subset $U \subset X$, we have

$$\Delta(U) = X \times X \cap \{(x, y) : x - y \in U\}.$$

Continuing in this direction, we may state alternatives of the three Banach theorems in the case of linear maps on locally convex topological vector spaces. However, in applications, we study either maps between Sobolev-type spaces (that are Banach), or pseudo-differential operators (not introduced in these lectures) that are outside of this setting.

We will not study in more details the the properties of $\mathscr{D}'(\Omega)$, and simply mention that this space is (in a precise sense) reflexive, so that linear forms on $\mathscr{D}'(\Omega)$ are induced by functions in $\mathscr{D}(\Omega)$ (see [13], Chapter 8).

2.5.4 Differentiation of Distributions

The fundamental idea of Schwartz (1945) is to show that by duality, one can define differentiation of distributions, and that this operation is continuous with respective to either topology—weak or strong—on $\mathscr{D}'(\Omega)$.

Definition 2.5.17. For all multi-index $\alpha \in \mathbb{R}^d$, and $T \in \mathscr{D}'(\Omega)$, we define $D^\alpha T \in \mathscr{D}'(\Omega)$ to be the distribution satisfying

$$D^\alpha T(\varphi) = (-1)^{|\alpha|} T(D^\alpha \varphi).$$

It satisfies $\mathrm{ord}_K(D^\alpha T) \leq \mathrm{ord}_K(T) + |\alpha|$ for all compact $K \subset \Omega$.

The continuity of this operation for the weak topology is trivial for

$$|T(D^\alpha \varphi)| \leq \|T\|_K \, \|D^\alpha \varphi\|_{m,K} \leq \|\varphi\|_{m+|\alpha|,K}$$

holds for all compact subset $K \subset \Omega$.

In early work, Schwartz had not introduced the minus sign ([36]), but the sign convention is the one consistent with integration by parts.

Of course, if $T = f \in C^\infty(\Omega)$, integrating by parts, we deduce that

$$\left\langle \frac{\partial}{\partial x_i} T, \varphi \right\rangle = -\left\langle T, \frac{\partial}{\partial x_i} \varphi \right\rangle = -\int_\Omega f \, \partial_{x_i} \varphi \, dx = \int_\Omega \varphi \, \partial_{x_i} f \, dx,$$

so that $\partial_{x_i} T = \partial_{x_i} f$. Sobolev spaces, which will make for half of those lectures, are sets of distributions whose weak derivatives belong to some L^p space (see. Thinking about partial differential equation (energy functionals), it becomes apparent why Sobolev spaces are the natural settings to solve equations, and their good properties allows one to use (say) calculus of variation in order to build solutions.

Examples 2.5.18. Let $H = \mathbf{1}_{\mathbb{R}_+}$ be the Heaviside function. Then, we have $H' = \delta_0$ in the sense of distributions. Indeed, for all $\varphi \in \mathscr{D}(\mathbb{R})$, we have

$$\langle H', \varphi \rangle = -\langle H, \varphi' \rangle = -\int_{\mathbb{R}} H(x)\varphi'(x)dx = -\int_0^\infty \varphi'(x)dx = \varphi(0) = \langle \delta_0, \varphi \rangle \,.$$

We saw in the first example that for C^1 functions by arcs, the usual derivative and the distributional derivative. This is a general fact, the formula of "jumps" allows on to quantify the difference (both quantities only differ up to Dirac masses).

Theorem 2.5.19. *Let $I \subset \mathbb{R}$ an open interval, and $f : I \to \mathbb{R}$ be a C^1 function by arcs,* i.e. *a function such that there exists* $\inf I < a_1 < \cdots < a_n < \sup I$ *such that $f|_{(a_i,a_{i+1})}$, $f_{(\inf I, a_1)}$ and $f_{(a_n, \sup I)}$ are functions of class C^1. Then, we have*

$$f' = \sum_{i=1}^{n-1} f' \mathbf{1}_{(a_i,a_{i+1})} + f' \mathbf{1}_{(\inf I, a_1)} + f' \mathbf{1}_{(a_n, \sup I)} + \sum_{i=1}^{n} \left(f'(a_i^+) - f'(a_i^-) \right) \delta_{a_i},$$

where $f'(a_i^\pm) = \lim\limits_{x \to a_i^\pm} f'(x)$ for all $i \in \{1, \cdots, n\}$.

Proof. The proof goes exactly as in Examples 2.5.18 and is omitted. □

Remark 2.5.20. This formula has generalisations to higher dimension, but would force us to introduce notions of differential geometry, that we consider to be outside the scope of those lectures.

The basic theorem about differentiation shows that the solution to an elliptic equation is generally unique in $\mathscr{D}'(\mathbb{R}^d)$. There are deep theorems that involved Sobolev spaces—to be introduced in the next chapter—and we will simply mention elementary results related to continuous functions and first order derivatives.

Theorem 2.5.21. *Let $T \in \mathscr{D}'(\mathbb{R}^d)$ be such that $\nabla T = 0$. Then, there exists $\vec{C}_0 \in \mathbb{C}^n$ such that $T = \vec{C}_0$.*

Proof. Showing by induction that $\partial_{x_i}T = 0$ implies that T is independent of x_i,we need only show the result for $d = 1$. Assume that $T' = 0$, and separating real and complex part, assume without loss of generality that T is real-valued. For all $\varphi \in \mathscr{D}(\mathbb{R})$, we have $\varphi = \psi'$ for some $\psi \in \mathscr{D}(\mathbb{R})$ if and only if

$$\int_{\mathbb{R}} \varphi(x)dx = 0. \tag{2.5.8}$$

Denote by H the hyperplane of such functions. Indeed, provided that (2.5.8) holds, we deduce that the following function

$$\psi(x) = \int_{-\infty}^{x} \varphi(y)dy$$

belongs to $\mathscr{D}(\mathbb{R})$ since φ has compact support. Conversely, if $\varphi = \psi'$ with $\psi \in \mathscr{D}(\mathbb{R})$, we have

$$\int_{-\infty}^{x} \varphi(y)dy = \psi(x)$$

And since ψ has compact support, there exists $r \in \mathbb{R}$ such that $\psi(x) = 0$ for all $x \geq r$, which shows that (2.5.8) holds in particular. Now, let $\theta \in \mathscr{D}(\mathbb{R})$ such that

$$\int_{\mathbb{R}} \theta(x)dx = 1.$$

For all $\varphi \in \mathscr{D}(\mathbb{R})$, we have

$$\psi = \varphi - \theta \int_{\mathbb{R}} \varphi \, d\mathscr{L}^1 \in H,$$

which implies since $T' = 0$ that

$$0 = T\left(\psi\right) = T(\varphi) - T(\theta) \int_{\mathbb{R}} \varphi \, d\mathscr{L}^1,$$

or

$$T = c_0 = T(\theta).$$

This concludes the proof of the theorem. □

The theorem is also true on a simply connected domain, since the result is local. More generally, we have the following result.

Theorem 2.5.22. *Let $X \in C^0(\mathbb{R}^d, \mathbb{R}^d)$ be a continuous vector-field. Assume that $\nabla T = X$. Then, T is a C^1 function such that $\nabla T = X$ classically.*

Proof. We only treat the case $X \in C^1(\mathbb{R}^d)$ (see [35] for the general case). The condition $\nabla T = X$ shows that in the distributional sense, the following identities hold

$$\partial_{x_i} X_j = \partial_{x_j} X_i \qquad \text{for all } 1 \leq i, j \leq d. \tag{2.5.9}$$

Since $\nabla X \in C^0(\mathbb{R}^d)$, the identities (2.5.9) hold for C^0 functions, and we deduce by the Poincaré lemma that there exists $f \in C^2(\mathbb{R}^d)$ such that $\nabla f = X$. Therefore, we have $\nabla(T - f) = 0$ in $\mathscr{D}'(\mathbb{R}^d)$, which shows that

$$T = f + \vec{C}_0$$

for some $\vec{C}_0 \in \mathbb{R}^d$. □

Remark 2.5.23. This theorem immediately generalises to the case $X \in C^{k,\alpha}(\mathbb{R}^d, \mathbb{R}^d)$ for some $k \in \mathbb{N}$ and $0 \leq \alpha \leq 1$, since the equations become classical, and on simply connected domains. When we introduce Sobolev spaces, we will be able to prove a Sobolev spaces version of the Poincaré lemma.

2.5.5 Restriction, Support, and localisation of a Distribution

Definition 2.5.24 (**Restriction**)**.** Let $\Omega' \subset \Omega$ be an open subset, and $T \in \mathscr{D}'(\Omega)$. Then, for all $\varphi \in \mathscr{D}(\Omega')$, the extension by 0 of φ to Ω is an element of $\mathscr{D}(\Omega)$ denoted by $\widetilde{\varphi}$, and we define by duality the restriction $T_{\Omega'} \in \mathscr{D}'(\Omega')$ by

$$\langle T_{\Omega'}, \varphi\rangle = \langle T, \widetilde{\varphi}\rangle \quad \text{for all } \varphi \in \mathscr{D}(\Omega').$$

Then, $T_{\Omega'} \in \mathscr{D}'(\Omega')$, and for all compact $K \subset \Omega'$, we have $\mathrm{ord}_K(T_{\Omega'}) \leq \mathrm{ord}_K(T)$.

The definition is consistent with the one of restriction of functions.

Definition 2.5.25. If $\Psi : \Omega \to \Omega'$ is a diffeomorphism between two open sets of $\mathbb{R}^d$, and $T \in \mathscr{D}'(\Omega)$, we define $T \circ \Psi \in \mathscr{D}'(\Omega')$ by duality as

$$\langle T \circ \Psi, \varphi\rangle = \langle T, |\det d\Psi^{-1}|\varphi \circ \Psi^{-1}\rangle \quad \text{for all } \varphi \in \mathscr{D}(\Omega').$$

Remark 2.5.26. Notice that for a function $T = f \in L^1_{\mathrm{loc}}(\Omega)$, the change of variable formula shows that

$$\int_\Omega f(\Psi(x))\varphi(x)dx = \int_{\Omega'} f(y)\varphi(\Psi^{-1}(y))|\det d\Psi^{-1}(y)|dy.$$

The case of conformal transformations in $\mathbb{R}^d$ is one of the most important one. If $\tau_a : \mathbb{R}^d \to \mathbb{R}^d, x \mapsto x+a$ is the translation by $a \in \mathbb{R}^d$, we define for all $T \in \mathscr{D}'(\mathbb{R}^d)$

$$\langle \tau_a T, \varphi\rangle = \langle T \circ \tau_a, \varphi\rangle = \left\langle T, \varphi \circ \tau_a^{-1}\right\rangle = \langle T, \tau_{-a}\varphi\rangle\,.$$

For all rotation $R \in O(d)$, we define

$$\langle RT, \varphi\rangle = \langle T, R^{-1}\varphi\rangle,$$

and for all $\lambda \in \mathbb{C} \setminus \{0\}$, we define

$$\langle T_\lambda, \varphi\rangle = |\lambda|^{-d}\langle T, \varphi_{\lambda^{-1}}\rangle,$$

where $\varphi_{\lambda^{-1}}(x) = \varphi(\lambda^{-1}(x))$.

Proposition 2.5.27 (Principle of Localisation)**.** *Let $\Omega \subset \mathbb{R}^d$ be an open subset of $\mathbb{R}^d$, and $\{\Omega_i\}_{i\in I}$ be an open covering of Ω. Suppose that $\{T_i\}_{i\in I} \in \prod_{i\in I}\mathscr{D}'(\Omega_i)$ and that for all $i, j \in I$, we have $T_i|\Omega_i \cap \Omega_j = T_j|\Omega_i \cap \Omega_j$. Then, there exists a unique distribution $T \in \mathscr{D}'(\Omega)$ such that $T_{\Omega_i} = T_i$ for all $i \in I$.*

Proof. Let $\left\{\Omega_j'\right\}_{j\in J}$ be a locally finite open covering such that Ω_j' is relatively compact in Ω, and $\Omega_j' \subset \Omega_i$ for some $i \in I$. Then, we define the distribution $T_j' = T_i|\Omega_j'$ if $\Omega_j' \subset \Omega_i$. The compatibility hypothesis shows that this is well-defined. Then, let $\{\chi_j\}_{j\in J}$ be a partition of unity, *i.e.* such that $\chi_j \in \mathscr{D}(\Omega)$ for all $j \in J$, $0 \leq \chi_j \leq 1$, $\mathrm{supp}(\chi_j) \subset \Omega_j'$ for all $j \in J$, and $\displaystyle\sum_{j\in J}\chi_j = 1$. For all $\varphi \in \mathscr{D}(\Omega)$, we have $\chi_j\varphi \in \mathscr{D}(\Omega_j')$, which shows that if T as in the theorem exists, we have

$$T(\varphi) = \sum_{j\in J} T_j'(\chi_j\varphi). \tag{2.5.10}$$

Therefore, T is uniquely determined by the family $\{X_i\}_{i\in I}$. Conversely, if $\mathrm{supp}(\varphi) \subset \Omega_i$, since $T_i|\Omega_j' = T_j'|\Omega_i \cap \Omega_j'$ for all $(i,j) \in I \times J$, we have

$$T_i(\varphi) = \sum_{j\in J} T_i(\chi_j\varphi) = \sum_{j\in J} T_j'(\chi_j\varphi) = T(\varphi),$$

which shows that $T_i = T|\Omega_i$ for all $i \in I$, and concludes the proof. □

Corollary 2.5.28. *The support of $T \in \mathscr{D}'(\Omega)$ is the complementary of the union of all open sets $\Omega_i \subset \Omega$ such that $T|\Omega_i = 0$. We denote this closed set by* $\mathrm{supp}(T) \subset \Omega$. *We say that T has compact support if* $\mathrm{supp}(T)$ *is a compact subset of* Ω.

Remark 2.5.29. The previous Proposition 2.5.27 shows that the definition is unambiguous.

Let us list without proofs a few easy properties of the support of distributions.

Proposition 2.5.30. **1**. *For all $a \in \mathbb{R}^d$ and $\alpha \in \mathbb{N}^d$, we have* $\mathrm{supp}(D^\alpha \delta_a) = \{a\}$.

2. $\mathrm{supp}(fT) \subset \mathrm{supp}(f) \cap \mathrm{supp}(T)$ *for all $f \in C^\infty(\Omega)$ and $T \in \mathscr{D}'(\Omega)$.*

Let us show that distributions with compact support, are in one-to-one correspondence with the dual of $\mathscr{E}(\Omega) = C^\infty(\Omega)$, which will justify the notation $T \in \mathscr{E}'(\Omega)$.

Theorem 2.5.31. *Let $\mathscr{E}'(\Omega)$ be the dual space of $\mathscr{E}(\Omega) = C^\infty(\Omega)$. Then, $T \in \mathscr{E}'(\Omega)$ if and only if $T \in \mathscr{D}'(\Omega)$ and T has compact support.*

Proof. Let L be a continuous linear form on $C^\infty(\Omega)$. Then, by density of $C_c^\infty(\Omega)$ in $C^\infty(\Omega)$, we deduce that L is entirely determined by its restrictions L_0 on $\mathscr{D}(\Omega)$. Since L_0 is continuous, there exists a compact set $K \subset \Omega$ and $m \geq 0$ such that

$$|L(f)| \leq C \left\|\varphi\right\|_{m,K} \quad \text{for all } \; f \in C^\infty(\Omega).$$

In particular, we have

$$|L_0(\varphi)| \leq C \left\|\varphi\right\|_{m,K} \quad \text{for all } \; \varphi \in \mathscr{D}(\Omega).$$

Therefore, we have $\mathrm{supp}(L_0) \subset K$, and $\mathrm{ord}(L_0) \leq m$. Conversely, if $T \in \mathscr{D}'(\Omega)$ has compact support, and $\chi \in \mathscr{D}(\Omega)$ is such that $\chi = 1$ on an relatively compact open neighbourhood Ω' of $\mathrm{supp}(T)$, then T extends to a continuous linear form on $C^\infty(\Omega)$ by the formula

$$T(f) = T(\chi f) \quad \text{for all } \; f \in C^\infty(\Omega).$$

Furthermore, T has support order since $\mathrm{ord}(T) \leq \mathrm{ord}_{\overline{\Omega'}}(T) < \infty$ by the definition of continuity in the distribution topology. ☐

We can finally prove the structure theorem for distributions of finite order.

Theorem 2.5.32. *Let $T \in \mathscr{D}'^{m}(\Omega)$ be a distribution of finite order $m \in \mathbb{N}$. Then, there exists Radon measures μ_α ($\alpha \in \mathbb{N}^d$) such that*

$$T = \sum_{|\alpha| \leq m} D^\alpha \mu_\alpha.$$

Furthermore, if T has compact support K, for all open neighbourhood U of K we can assume that $\mathrm{supp}(\mu_\alpha) \subset U$. *Conversely, any finite sum of derivatives of Radon measure is an element of $\mathscr{D}'^{m}(\Omega)$.*

Proof. The converse statement is trivial.

By hypothesis, we deduce that for all sequence $\{\varphi_n\}_{n\in\mathbb{N}} \subset \mathscr{D}(\Omega)$ such that $\|D^\alpha \varphi_n\|_{\mathrm{L}^\infty(\Omega)} \underset{n\to\infty}{\longrightarrow} 0$, we have $T(\varphi_n) \underset{n\to\infty}{\longrightarrow} 0$. We now associate to each $\varphi \in \mathscr{D}(\Omega)$ the $N(d,m)$-uple $\Phi = \{D^\alpha \varphi\}_{|\alpha|\leq m} \subset \mathscr{K}(\Omega)$, the space of continuous functions with compact support. The previous discussion shows that the map $\Phi \mapsto T(\varphi)$ is a continuous linear form defined on the sub-vector spaces $\mathscr{K}(\Omega)^{N(d,m)}$ of elements Φ given by $\Phi = \{D^\alpha \varphi\}_{|\alpha|\leq m}$. Therefore, we deduce by the Hahn-Banach theorem that this linear form admits an extension (denoted by L) on $\mathscr{K}(\Omega)^{N(d,m)}$. By the structure theorem of the dual space of $\mathscr{K}(\Omega)$, we deduce that there exits Radon measures $\{\nu_\alpha\}_{|\alpha|\leq m}$ such that

$$L(\{f_\alpha\}_{|\alpha|\leq m}) = \sum_{|\alpha|\leq m} \int_\Omega f_\alpha d\nu_\alpha.$$

Taking $\{f_\alpha\}_{|\alpha|\leq m} = \{D^\alpha \varphi\}_{|\alpha|\leq m}$ for some $\varphi \in \mathscr{D}(\Omega)$, we deduce that

$$T(\varphi) = \sum_{|\alpha|\leq m} \int_\Omega D^\alpha \varphi \, d\nu_\alpha = \sum_{|\alpha|\leq m} \left\langle (-1)^{|\alpha|}\nu_\alpha, \varphi \right\rangle,$$

which concludes the proof with $\mu_\alpha = (-1)^{|\alpha|}\nu_\alpha$.

We omit the part on compact support, that follows easily by the introduction of a suitable cut-off function. □

2.6 Convolution of Distributions

2.6.1 First definitions

As previously, we want to generalise the notion of convolution to the case of distributions. Notice that whenever $f, g, h \in \mathscr{D}(\mathbb{R}^d, \mathbb{R})$, we have

$$\begin{aligned}
\langle f * g, h\rangle &= \int_{\mathbb{R}^d} (f*g)(x)h(x)dx = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(y)g(x-y)dy \right) h(x)dx \\
&= \int_{\mathbb{R}^d} f(y) \left(\int_{\mathbb{R}^d} g(x-y)h(x) \right) dy = \int_{\mathbb{R}^d} f(y)(g*h)(-y)dy \\
&= \int_{\mathbb{R}^d} f(y)\widetilde{g*h}(y)dy = \left\langle f, \widetilde{(g*h)} \right\rangle,
\end{aligned}$$

where $\widetilde{\varphi}(x) = \varphi(-x)$ for all $\varphi \in \mathscr{D}(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$. We can rewrite this expression as follows:

$$\langle f_x, \langle g_y, \varphi(x+y)\rangle\rangle = \langle g_y, \langle f_x, \varphi(x+y)\rangle\rangle .$$

Indeed, we have

$$\langle g_y, \varphi(x+y)\rangle = \int_{\mathbb{R}^d} g(y)\varphi(x+y)dy,$$

so that by Fubini theorem

$$\begin{aligned}
\langle f_x, \langle g_y, \varphi(x+y)\rangle\rangle &= \int_{\mathbb{R}^d} f(x) \left(\int_{\mathbb{R}^d} g(y)\varphi(x+y)dy \right) dx \\
&= \int_{\mathbb{R}^d} f(x) \left(\int_{\mathbb{R}^d} g(z-x)\varphi(z)dz \right) dx = \int_{\mathbb{R}^d} (f*g)(z)\varphi(z)dz,
\end{aligned}$$

where we made the change of variables $z = x + y$. The second computation is analogous.

Therefore, we want to define the convolution of two distributions $T_1, T_2 \in \mathscr{D}'(\mathbb{R}^d)$ as the distribution $T = T_1 * T_2$ which is the common value of $\langle T_{1,x}, \langle T_{2,y}, \varphi(x+y)\rangle\rangle$ and $\langle T_{2,y}, \langle T_{2,x}, \varphi(x+y)\rangle\rangle$. For general distributions, $\langle T_{2,y}, \varphi(x+y)\rangle \notin \mathscr{D}(\mathbb{R}^d)$, so the expression cannot be defined. However, if T_1 (or T_2) has compact support, then $x \mapsto \langle T_{2,y}, \varphi(x+y)\rangle$ is a $C^\infty(\mathbb{R}^d)$ function, and the first duality bracket makes sense, whilst the second duality bracket makes sense for $y \mapsto \langle T_{1,x}, \varphi(x+y)\rangle \in \mathscr{D}(\mathbb{R}^d)$. Let us show that both expression are equal if T_1 or T_2 has compact support, and let us denote them by $S_1(\varphi)$ and $S_2(\varphi)$. First assume that $T_1 \in \mathscr{D}'^m(\mathbb{R}^d)$. Then, the structure Theorem 2.5.32 shows that there exists a family of Radon measures $\{\mu_\alpha\}_{|\alpha|\leq m}$ with compact support on $\mathbb{R}^d$ such that

$$T_1 = \sum_{|\alpha|\leq m} D^\alpha \mu_\alpha.$$

By linearity, it suffices to treat the case $T_1 = D^\alpha \mu$, which is equivalent to showing that $S_1(D^\alpha \varphi)$ and $S_2(D^\alpha\varphi)$ coincide, where T is replaced by μ. Finally, we need only treat the case $T_1 = \mu$. We therefore need to show that

$$\int_{\mathbb{R}^d} \langle T_{2,y}, \varphi(x+y)\rangle \, d\mu(x) = \left\langle T_{2,y}, \int_{\mathbb{R}^d} \varphi(x+y)d\mu(x) \right\rangle.$$

If T_2 has compact support, a similar reduction shows that we can assume that $T_2 = \nu$ is a Radon measure, in which case the identity $S_1(\varphi) = S_2(\varphi)$ is nothing else than Fubini theorem. If T_2 does not have compact support, we let $\{\chi_n\}_{n\in\mathbb{N}} \subset \mathscr{D}(\mathbb{R}^d)$ be a sequence such that $\chi_n = 1$ on $B(0,n)$. Then,

$$\langle (\chi_n T_2)_y, \varphi(x+y)\rangle \underset{n\to\infty}{\longrightarrow} \langle T_{2,y}, \varphi(x+y)\rangle$$

uniformly as x stays within a compact domain of $\mathbb{R}^d$. Therefore, we have

$$\int_{\mathbb{R}^d} \langle (\chi_n T_2)_y, \varphi(x+y)\rangle \, d\mu(x) \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}^d} \langle T_{2,y}, \varphi(x+y)\rangle \, d\mu(y)$$

and since $y \mapsto \int_{\mathbb{R}^d} \varphi(x+y) d\mu(x) \in \mathscr{D}(\mathbb{R}^d)$, we also trivially have

$$\left\langle (\chi_n T_2)_y, \int_{\mathbb{R}^d} \varphi(x+y) d\mu(x) \right\rangle \underset{n\to\infty}{\longrightarrow} \left\langle T_{2,y}, \int_{\mathbb{R}^d} \varphi(x+y) d\mu(x) \right\rangle.$$

Since the equality between the left-hand sides of holds for all $n \in \mathbb{N}$, this concludes the proof.

Therefore, the equality is proved for $(T_1, T_2) \in \mathscr{D}'(\mathbb{R}^d) \times \mathscr{E}'(\mathbb{R}^d)$ and $(T_1, T_2) \in \mathscr{E}'(\mathbb{R}^d) \times \mathscr{D}'(\mathbb{R}^d)$. By induction, we can define the convolution of an arbitrary finite number of distributions, provided that their supports are all compact but one at most.

Proposition 2.6.1. *For all $(S, T, U) \in \mathscr{E}'(\mathbb{R}^d) \times \mathscr{D}'(\mathbb{R}^d) \times \mathscr{E}'(\mathbb{R}^d)$, we have*

1. $S * R = T * S$.
2. $(S * T) * U = S * (T * U)$.
3. $\delta_0 * T = T * \delta_0 = T$.
4. $D^\alpha(S * T) = D^\alpha S * T = S * D^\alpha T$.
5. $\mathrm{supp}(S * T) \subset \mathrm{supp}(S) + \mathrm{supp}(T)$.

Proof. We need only check the first property in the case of two Radon measures μ and ν on $\mathbb{R}^d$. Notice that by Fubini's theorem, we have

$$\begin{aligned}(\mu * \nu)(\varphi) &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \varphi(x+y) d\nu(y) \right) d\mu(x) = \int_{\mathbb{R}^d \times \mathbb{R}^d} \varphi(x+y) d(\mu \times \nu)(x,y) \\ &= \int_{\mathbb{R}^d \times \mathbb{R}^d} \varphi(y+x) d(\nu \times \mu)(x,y) = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \varphi(x+y) d\mu(y) \right) d\nu(x) = (\nu * \mu)(\varphi).\end{aligned}$$

The proof of **2**. is analogous and we omit it, while the proof of **3**. is trivial.

For the proof of **4**., we have in the case of two Radon measures ν and ν

$$\begin{aligned}\langle D^\alpha(\mu * \nu), \varphi \rangle &= (-1)^{|\alpha|} \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} D_y^\alpha \varphi(x,y) d\nu(y) \right) d\mu(x) = \langle \mu_x, \langle D^\alpha \nu_y, \varphi(x+y) \rangle \rangle \\ &= \langle (\mu * D^\alpha \nu), \varphi \rangle,\end{aligned}$$

and we prove similarly that $D^\alpha(\mu * \nu) = D^\alpha \mu * \nu$. The general proof is similar.

Finally, we prove the statement on supports. Since S has compact support, the following function

$$y \mapsto \langle S_x, \varphi(x+y) \rangle$$

is an element of $\mathscr{D}(\mathbb{R}^d)$. Indeed, it vanishes whenever $\mathrm{supp}\,(x \mapsto \varphi(x+y)) = y + \mathrm{supp}(\varphi)$ and $\mathrm{supp}(S)$ have vanishing intersection. In other words, $\mathrm{supp}(y \mapsto \langle S_x, \varphi(x+y) \rangle) \subset \mathrm{supp}\,(S) - \mathrm{supp}\,(\varphi)$. Therefore, $\langle T_y, \langle S_x, \varphi(x+y) \rangle \rangle$ vanishes provided that $\mathrm{supp}(T) \cap (\mathrm{supp}(S) - \mathrm{supp}(\varphi)) = \varnothing$, or $\mathrm{supp}\,(\varphi) \cap (\mathrm{supp}(S) + \mathrm{supp}(T)) = \varnothing$. Therefore, we have $\mathrm{supp}(S * T) \subset \mathrm{supp}(S) + \mathrm{supp}(T)$. ☐

Remark 2.6.2. We spoke above of tensor product of distributions, and this is how Schwartz historically (see [35], Chapitre 6) defined the convolution. However, it is not necessary to introduce this notion in order to define the convolution, and we omit the discussion of tensor products completely.

We see in particular that $D^\alpha T = D^\alpha \delta_0 * T$, so that all (linear) partial differential equations may be seen as convolution equations. We will see in the next section a way to transform a large class of such convolution equations to the afore-mentioned problem of division of distributions.

The reader has probably already shown that the solution of the Laplace equation $\Delta u = f$ (for f sufficiently smooth) in $\mathbb{R}^d$ ($d \geq 2$) is given by

$$u = G_d * f,$$

where G_d is the fundamental solution of the Laplacian, given by

$$G_d(x) = \begin{cases} -\dfrac{1}{(d-2)\beta(d)} \dfrac{1}{|x|^{d-2}} & \text{for } d \geq 3 \\ \dfrac{1}{2\pi} \log|x| \qquad \text{for } d = 2, \end{cases}$$

where $\beta(d) = \mathscr{H}^{d-1}(S^{d-1})$. The usual proof using integration by parts shows that $\Delta G_d = \delta_0$, but we will check this fact below by a different method.

2.6.2 A First Extension: Distributions of Convolutive Supports

Now, we extend the notion of convolution to the case of distributions having *convolutive supports.*

Definition 2.6.3 (Convolutive Supports)**.** Let A and B be two closed subsets of $\mathbb{R}^d$. We say that (A, B) is convolutive (or that A and B are convolutive) if for all $0 < R < \infty$, there exists $\rho(R) < \infty$ such that

$$A \times B \cap \{(a, b) : |a + b| < R\} \subset B(0, \rho(R)) \times B(0, \rho(R)).$$

More generally, a family $\{F_i\}_{i \in I}$ of closed sets is convolutive if for all subset $J \subset I$, and for all $0 < R < \infty$, there exists $\rho(R) < \infty$ such that

$$\prod_{j \in J} F_j \cap \left\{ (x_j)_{j \in J} : \left| \sum_{j \in J} x_j \right| < R \right\} \subset \prod_{j \in J} B(0, \rho(R))$$

Proposition 2.6.4. *If A and B are convolutive sets, $A + B$ is a closed subset of $\mathbb{R}^d$.*

Proof. Let $\{x_n = a_n + b_n\}_{n \in \mathbb{N}} \subset A + B$ be a convergent sequence (call the limit $x \in \mathbb{R}^d$). Then $\{x_n\}_{n \in \mathbb{N}}$ is bounded, and by the convolutiveness property, we deduce that both $\{a_n\}_{n \in \mathbb{N}}$ and $\{b_n\}_{n \in \mathbb{N}}$ are bounded sequences. Therefore, up to a subsequence, we have $a_n \underset{n \to \infty}{\longrightarrow} a \in A$ and $b_n \underset{n \to \infty}{\longrightarrow} b \in B$ since A and B are closed. By unicity of the limit, we deduce that $x = a + b \in A + B$. QED. ☐

Theorem 2.6.5. *For all couple of distributions $S, T \in \mathscr{D}'(\mathbb{R}^d)$ whose supports are convolutive, and for all sequence $\{\chi_n\}_{n \in \mathbb{N}} \subset \mathscr{D}(\mathbb{R}^d)$ such that $\chi_n = 1$ on $\overline{B}(0, n)$, define for all $\varphi \in \mathscr{D}(\mathbb{R}^d)$*

$$\langle S * T, \varphi \rangle = \lim_{n \to \infty} \langle (\chi_n S) * (\chi_n T), \varphi \rangle . \tag{2.6.1}$$

Then, for all $\varphi \in \mathscr{D}(\mathbb{R}^d)$, the sequence on the right-hand side of (2.6.1) *is constant for $n \in \mathbb{N}$ large enough. We call $S * T$ the convolution of S and T, and it does not depend on the choice of the sequence $\{\chi_n\}_{n \in \mathbb{N}}$ as above.*

Furthermore, all properties of Proposition 2.6.1 are satisfied for this generalised convolution.

Proof. Assume that $\text{supp}(\varphi) \subset B(0,R)$. We will show that the sequence $\langle (\chi_n S) * (\chi_n T), \varphi \rangle$ is constant for n large enough. If $n, m \geq \rho(R)$, we have

$$\begin{aligned}&\langle (\chi_n S) * (\chi_n T) - (\chi_m S) * (\chi_m T), \varphi \rangle \\ &= \langle ((\chi_n - \chi_m)S) * (\chi_n T), \varphi \rangle + \langle (\chi_n S) * ((\chi_n - \chi_m)T), \varphi \rangle \,.\end{aligned}$$

If the first term does *not* vanish, then $\text{supp}(\varphi) \cap (\text{supp}\,((\chi_n - \chi_m)S) + \text{supp}(\chi_n T)) \neq \varnothing$. However, if $x \in \text{supp}(\varphi) \cap (\text{supp}\,((\chi_n - \chi_m)S) + \text{supp}(\chi_n T))$, then $x = y + z$, where $|y| > n \geq \rho(R)$, while $|x| < R$ since $x \in \text{supp}(\varphi) \subset B(0,R)$. However, since $(y,z) \in \text{supp}(S) \times \text{supp}(T)$, those conditions contradict the hypothesis on convolutiveness of those those distributions.

Therefore, we see that the limit (2.6.1) defines a distribution.,Now, assume that $\text{supp}(\varphi)$ and $\text{supp}(S) + \text{supp}(T)$ are disjoint. Then, we have *a fortiori* $\text{supp}(\varphi) \cap (\text{supp}(\chi_n S) + \text{supp}(\chi_n T)) = \varnothing$ for all $n \in \mathbb{N}$, which shows by (2.6.1) that $\langle S * T, \varphi \rangle = 0$.

The proofs of the other properties are similar, and we omit them. □

Remark 2.6.6. The convolution product is not associative in general, as the following example shows:

$$\begin{cases}(1 * \delta_0') * H = 0 \\ 1 * (\delta_0' * H) = 1 * (\delta_0 * H') = 1 * (\delta_0 * \delta_0) = 1 * \delta_0 = \delta_0.\end{cases}$$

Finally, we introduce a last extension of convolution thanks to a new class of distributions that happen to be proto-Sobolev spaces (anticipating on the next chapter, they correspond to $W^{-\infty,p}$). We will do so in a new section.

2.6.3 A Second Extension: the $\mathscr{D}'_{\mathrm{L}^p}(\mathbb{R}^d)$ Spaces of Distributions

Definition 2.6.7. Let

$$\mathscr{D}_{\mathrm{L}^p}(\mathbb{R}^d) = C^\infty(\mathbb{R}^d) \cap \left\{ f : D^\alpha f \in L^p(\mathbb{R}^d) \text{ for all } \alpha \in \mathbb{N}^d \right\}$$

We equip it with the family of semi-norm $\|D^\alpha(\,\cdot\,)\|_{\mathrm{L}^p(\mathbb{R}^d)}$ where $\alpha \in \mathbb{N}^d$, which makes it a Fréchet space.

We also define $\mathscr{B}(\mathbb{R}^d) = \mathscr{D}_{\mathrm{L}^\infty}(\mathbb{R}^d)$, and

$$\dot{\mathscr{B}}(\mathbb{R}^d) = \mathscr{B}(\mathbb{R}^d) \cap \left\{ \varphi : D^\alpha \varphi(x) \underset{|x|\to\infty}{\longrightarrow} 0 \text{ for all } \alpha \in \mathbb{R}^d \right\}.$$

Remark 2.6.8. We only define this space on $\mathbb{R}^d$ since it will only be used for convolution and Fourier transform, but one could define it on an arbitrary open subset $\Omega \subset \mathbb{R}^d$.

Theorem 2.6.9. *$\mathscr{D}(\mathbb{R}^d)$ is dense in $\mathscr{D}_{\mathrm{L}^p}(\mathbb{R}^d)$ for $p < \infty$ and in $\dot{\mathscr{B}}(\mathbb{R}^d)$.*

Proof. Let $f \in \mathscr{D}_{\mathrm{L}^p}(\mathbb{R}^d)$, and $\{\chi_n\}_{n\in\mathbb{N}} \subset \mathscr{D}(\mathbb{R}^d)$ such that $\chi_n = 1$ on $B(0,n)$. Then, we have for all $\alpha \in \mathbb{N}^d$ by Leibniz's formula and Minkowski's inequality

$$\begin{aligned}\left(\int_{\mathbb{R}^d} |D^\alpha f - D^\alpha(\chi_n f)|^p \, dx \right)^{\frac{1}{p}} &\leq \left(\int_{\mathbb{R}^d \setminus \overline{B}(0,n)} (1-\chi_n)^p |D^\alpha f|^p dx \right)^{\frac{1}{p}} \\ + \sum_{\substack{\beta \leq \alpha \\ \beta \neq 0}} \left(\int_{\mathbb{R}^d \setminus \overline{B}(0,n)} |D^\beta \chi_n| |D^{\alpha-\beta} f|^p dx \right)^{\frac{1}{p}} &\leq \|D^\alpha f\|_{\mathrm{L}^p(\mathbb{R}^d \setminus \overline{B}(0,n))} \underset{n\to\infty}{\longrightarrow} 0\end{aligned}$$

by dominated convergence, provided that $\|\nabla \chi_n\|_{\mathrm{L}^\infty(\mathbb{R}^d)}$ is bounded. But this is very easy to find such a sequence, by taking a regularisation (by convolution, say) of

$$\psi_n(x) = \begin{cases} 1 & \text{for all } |x| \leq n \\ -\dfrac{1}{n}|x| + 2 & \text{for all } n < |x| < 2n \\ 0 & \text{for all } |x| \geq 2n. \end{cases}$$

Therefore, the proof of the theorem is complete. □

Definition 2.6.10. For $1 < p \leq \infty$, we define $\mathscr{D}'_{\mathrm{L}^p}(\mathbb{R}^d) \subset \mathscr{D}'(\mathbb{R}^d)$ as the dual of $\mathscr{D}_{\mathrm{L}^{p'}}(\mathbb{R}^d)$, where $\dfrac{1}{p} + \dfrac{1}{p'} = 1$, that we equip with the weak topology. We also define $\mathscr{B}'(\mathbb{R}^d) = \mathscr{D}'_{\mathrm{L}^\infty}(\mathbb{R}^d)$.

Theorem 2.6.11. *For $1 < p < \infty$, the space $\mathscr{D}_{\mathrm{L}^p}(\mathbb{R}^d)$ is reflexive, and in particular, the dual of $\mathscr{D}'_{\mathrm{L}^p}(\mathbb{R}^d)$ is $\mathscr{D}_{\mathrm{L}^{p'}}$. Furthermore the dual of $\mathscr{D}'_{\mathrm{L}^1}(\mathbb{R}^d)$ is $\mathscr{B}(\mathbb{R}^d)$.*

We do not prove those theorems here, since they will be proved in a different setting in the next chapter.

Theorem 2.6.12. *A distribution $T \in \mathscr{D}'_{\mathrm{L}^p}(\mathbb{R}^d)$ if and only if T is a finite sum of derivatives of L^p functions. Equivalently, $T \in \mathscr{D}'_{\mathrm{L}^p}(\mathbb{R}^d)$ if and only if for all $\varphi \in \mathscr{D}(\mathbb{R}^d)$, the regularised distribution $T * \varphi$ belongs to $L^p(\mathbb{R}^d)$.*

We also omit the proof.

Remark 2.6.13. Anticipating on the next chapter, we see that for $1 < p \leq \infty$, we have $T \in \mathscr{D}'_{\mathrm{L}^p}(\mathbb{R}^d)$ if and only if there exists $m \in \mathbb{N}$ such that $T \in W^{-m,p}(\mathbb{R}^d) = (W^{m,p'}(\mathbb{R}^d))$. In other words,

$$\mathscr{D}'_{\mathrm{L}^p}(\mathbb{R}^d) = \bigcup_{m \in \mathbb{N}} W^{-m,p}(\mathbb{R}^d).$$

This is why we need not prove those results here, and prove the decomposition in derivatives of L^p functions for each space $W^{-m,p}(\mathbb{R}^d)$, which is relatively easy. See [35] p. $199 - 205$ for more details on this general approach.

Theorem 2.6.14.

1. *If $T \in \mathscr{D}'_{\mathrm{L}^p}(\mathbb{R}^d)$ and $f \in \mathscr{D}_{\mathrm{L}^q}(\mathbb{R}^d)$, the product fT belongs to $\mathscr{D}'_{\mathrm{L}^r}(\mathbb{R}^d)$ for $r \geq 1$, provided that $\dfrac{1}{r} \leq \dfrac{1}{p} + \dfrac{1}{q}$.*

2. *If $S \in \mathscr{D}'_{\mathrm{L}^p}(\mathbb{R}^d)$, $T \in \mathscr{D}'_{\mathrm{L}^q}(\mathbb{R}^d)$, and $\dfrac{1}{p} + \dfrac{1}{q} - 1 \geq 0$, then we can define the convolution product $S * T$ and $S * T \in \mathscr{D}'_{\mathrm{L}^r}(\mathbb{R}^d)$.*

Proof. The existence of a convolution product in $\mathscr{D}'(\mathbb{R}^d)$ follows immediately from the decomposition of S and T into a sum of derivatives of L^p and L^q functions. However, since this decomposition is not-unique, we must needs prove that the convolution is independent of the chosen decompositions. Therefore, assume that

$$S = \sum_{|\alpha| \leq m} D^\alpha f_\alpha \qquad T = \sum_{|\beta| \leq n} D^\beta g_\beta.$$

Then, for all $|\alpha| \leq m$ and $|\beta| \leq n$, define

$$D^\alpha f_\alpha * D^\beta g_\beta = D^{\alpha+\beta}(f_\alpha * g_\beta).$$

Since $f_\alpha \in L^p(\mathbb{R}^d)$ and $g_\beta \in L^q(\mathbb{R}^d)$, we have $f_\alpha * g_\beta \in L^r(\mathbb{R}^d)$, and $D^\alpha f_\alpha * D^\beta g_\beta \in \mathscr{D}'_{\mathrm{L}^r}(\mathbb{R}^d)$. Therefore, we define the convolution of S and T as

$$S * T = \sum_{\substack{|\alpha| \leq m \\ |\beta| \leq n}} D^{\alpha+\beta}(f_\alpha * g_\beta).$$

Since the role of S and T is symmetric, assume that we have the alternative decomposition

$$S = \sum_{|\alpha| \leq m} D^\alpha f'_\alpha.$$

Now, let

$$U=\sum_{\substack{|\alpha|\leq m\\|\beta|\leq n}}D^{\alpha+\beta}(f'_\alpha*g_\beta).$$

For all $\varphi\in\mathscr{D}(\mathbb{R}^d)$, we have

$$\begin{aligned}U(\varphi)&=\sum_{\substack{|\alpha|\leq m\\|\beta|\leq n}}\left\langle f'_\alpha*g_\beta,(-1)^{|\alpha|+|\beta|}D^\alpha\varphi\right\rangle=\sum_{\substack{|\alpha|\leq m\\|\beta|\leq n}}\left\langle f'_\alpha,\widetilde{g}_\beta*\left((-1)^{|\alpha|+|\beta|}D^{\alpha+\beta}\varphi\right)\right\rangle\\&=\left\langle\sum_{|\alpha|\leq m}D^\alpha f'_\alpha,\sum_{|\beta|\leq n}\widetilde{g}_\beta*\left((-1)^{|\beta|}D^\beta\varphi\right)\right\rangle=\left\langle S,\sum_{|\beta|\leq n}\widetilde{g}_\beta*\left((-1)^{|\beta|}D^\beta\varphi\right)\right\rangle=\langle S*T,\varphi\rangle.\end{aligned}$$

Therefore, the definition is independent of the chosen decomposition by density of $\mathscr{D}(\mathbb{R}^d)$ in $\mathscr{D}_{L^{p'}}(\mathbb{R}^d)$. □

2.7 Tempered Distributions and Fourier Transform

In this section, all functions will be complex valued unless stated otherwise (which will never happen). As previously, we want to define the Fourier transform by duality. Explicitly, for all $T\in\mathscr{D}'(\mathbb{R}^d)$ and $\varphi\in\mathscr{D}(\mathbb{R}^d)$, we want to define

$$\left\langle\widehat{T},\varphi\right\rangle=\langle T,\widehat{\varphi}\rangle$$

where

$$\mathscr{F}(\varphi)(\xi)=\widehat{\varphi}(\xi)=\int_{\mathbb{R}^d}\varphi(x)e^{-ix\cdot\xi}dx.$$

Recall the following basic facts on the Fourier transform.

Theorem 2.7.1 (Riemann-Lebesgue lemma). *Let $f\in L^1(\mathbb{R}^d)$. Then $\widehat{f}\in C^0(\mathbb{R}^d)$, and*

$$\lim_{|\xi|\to\infty}\widehat{f}(\xi)=0. \tag{2.7.1}$$

The basic algebraic properties of the Fourier transform are listed below.

Proposition 2.7.2. *Let $f,g\in L^1(\mathbb{R}^d)$. Then, the following properties are verified.*

1. *For all $\lambda\in\mathbb{C}$, we have*

$$\mathscr{F}(\lambda f+g)=\lambda\mathscr{F}(f)+\mathscr{F}(g).$$

2. *If $m\in\mathbb{N}$ et $|x|^mf\in L^1(\mathbb{R}^d)$, then $\widehat{f}\in C^m(\mathbb{R}^d,\mathbb{C})$, and for all $|\alpha|\leq m$, we have*

$$D^\alpha\widehat{f}(\xi)=(-i)^{|\alpha|}\int_{\mathbb{R}}x^\alpha f(x)e^{-ix\cdot\xi}dx.$$

3. *If $f\in C^m(\mathbb{R}^d,\mathbb{C})$ for some $n\in\mathbb{N}$, and $D^\alpha f\in L^1(\mathbb{R}^d)$ for all $|\alpha|\leq m$, then for all $|\alpha|\leq m$, we have*

$$\mathscr{F}(D^\alpha f)(\xi)=i^{|\alpha|}\xi^\alpha\mathscr{F}(f)(\xi).$$

Theorem 2.7.3. *Let $f\in L^1(\mathbb{R}^d)$.*

1. **(Fourier Inversion Formula)** *If $\widehat{f} \in L^1(\mathbb{R}^d)$, then for all $x \in \mathbb{R}^d$, we have*

$$f(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \widehat{f}(\xi)e^{ix\cdot\xi}d\xi = \frac{1}{(2\pi)^d}\mathscr{F}(\mathscr{F}(f))(-x).$$

2. **(Plancherel Identity)** *For all $f \in L^2(\mathbb{R}^d)$, we have $\widehat{f} \in L^2(\mathbb{R}^d)$ and*

$$\int_{\mathbb{R}^d} |f(x)|^2 dx = \frac{1}{(2\pi)^d}\int_{\mathbb{R}} |\widehat{f}(\xi)|^2 d\xi. \tag{2.7.2}$$

3. **(Convolution Property)** *For all $f, g \in L^1(\mathbb{R}^d)$, we have*

$$\mathscr{F}(f * g) = \mathscr{F}(f)\cdot\mathscr{F}(g). \tag{2.7.3}$$

Remark 2.7.4. In other words, we have $\mathscr{F}^2 = (2\pi)^d, \mathrm{Id}\circ\iota$, où $\iota(x) = -x$.

However, the definition *cannot* make sense for all distributions. Indeed, for all $\varphi \in \mathscr{D}(\mathbb{R}^d)$, there exists $R > 0$ such that

$$\widehat{\varphi}(\xi) = \int_{B(0,R)} \varphi(x)e^{-ix\cdot\xi}dx.$$

In particular, the function $\widehat{\varphi}$ can be extended to a pluri-holomorphic function, and the maximum principle implies that $\widehat{\varphi}$ does not have compact support unless $\varphi = 0$. In particular, $\widehat{\varphi} \notin \mathscr{D}(\mathbb{R}^d)$ for all $\varphi \neq 0$, and the expression $\langle T, \widehat{\varphi}\rangle$ does not make sense in general. Therefore, we are confronted with the problem of finding a topological vector space $S \subset C^\infty(\mathbb{R}^d)\cap L^1(\mathbb{R}^d)$ such that $\mathscr{D}(\mathbb{R}^d) \subset S$, and with respect to the Fourier inversion formula, that possesses the following invariance property: $\mathscr{F}(S) = S$. Furthermore, we want to find a space on which the previous operation of differentiation is compatible. In other words, we require that for all $T \in S'$ (the dual of S), for all $\varphi \in S$, $\alpha \in \mathbb{N}^d$, the following quantities are well-defined

$$\langle \mathscr{F}(D^\alpha T), \varphi\rangle$$

and

$$\langle D^\alpha(\mathscr{F}(T)), \varphi\rangle.$$

Using both definitions of D^α and $\mathscr{F}$, we get

$$\langle \mathscr{F}(D^\alpha T), \varphi\rangle = \langle D^\alpha T, \mathscr{F}(\varphi)\rangle = (-1)^{|\alpha|}\langle T, D^\alpha \mathscr{F}(\varphi)\rangle = \langle T, \mathscr{F}(x^\alpha\varphi)\rangle = \langle \mathscr{F}(T), x^\alpha\varphi\rangle,$$

whilst

$$\langle D^\alpha(\mathscr{F}(T)), \varphi\rangle = (-1)^\alpha\langle \mathscr{F}(T), D^\alpha\varphi\rangle.$$

Combining both properties, we are led to the axioms

$$x^\beta D^\alpha\varphi \in S \quad \text{for all } \varphi \in S \text{ and for all } \alpha, \beta \in \mathbb{N}^d. \tag{2.7.4}$$

Furthermore, the stability condition $\mathscr{F}(S) = S$ for Fourier transform shows that for all $\alpha, \beta \in \mathbb{N}^d$, there exists $\psi \in S$ such that

$$x^\beta D^\alpha\varphi = \mathscr{F}(\psi).$$

In particular, the Riemann-Lebesgue lemma implies that

$$x^\beta D^\alpha\varphi(x) \underset{|x|\to\infty}{\longrightarrow} 0. \tag{2.7.5}$$

Therefore, both properties (2.7.4) and (2.7.5) lead to the definition of the following minimal space $\mathscr{S}(\mathbb{R}^d)$, that happens to be a Fréchet space, once equipped with a natural set of semi-norms.

Definition 2.7.5. The Schwartz space $\mathscr{S}(\mathbb{R}^d)$, or space of rapidly decreasing function, is defined as follows:

$$\mathscr{S}(\mathbb{R}^d) = C^\infty(\mathbb{R}^d) \cap \left\{\varphi : \sup_{x\in\mathbb{R}^d} |x^\beta||D^\alpha\varphi(x)| < \infty \text{ for all } (\alpha,\beta)\in\mathbb{N}^d\times\mathbb{N}^d\right\}.$$

The previous discussion shows that $\mathscr{S}(\mathbb{R}^d)$ is the minimal space satisfying the desirable axioms of Fourier transform—we want to take the smallest function space so that the dual space is the largest possible. It happens to be a solution to our problem, as we will easily check.

Theorem 2.7.6. *For all $\alpha,\beta\in\mathbb{N}^d$, define the semi-norm $\|\cdot\|_{\alpha,\beta}$ on $\mathscr{S}(\mathbb{R}^d)$ such that for all $\varphi\in\mathscr{S}(\mathbb{R}^d)$,*

$$\|\varphi\|_{\alpha,\beta} = \left\|x^\beta D^\alpha\varphi\right\|_{\mathrm{L}^\infty(\mathbb{R}^d)}. \tag{2.7.6}$$

Then, the topological vector space $(\mathscr{S}(\mathbb{R}^d), \{\|\cdot\|_{\alpha,\beta}\})$ is a Fréchet space, and the closure of $\mathscr{D}(\mathbb{R}^d)$ for the induced topology is $\mathscr{S}(\mathbb{R}^d)$.

Furthermore, the Schwartz space is stable under Fourier transform: $\mathscr{F}(\mathscr{S}(\mathbb{R}^d)) = \mathscr{S}(\mathbb{R}^d)$.

Proof. **Step 1:** Stability under Fourier transform.

We first show that $\mathscr{S}(\mathbb{R}^d)$ is stable under $\mathscr{F}$, since we trivially have $\mathscr{D}(\mathbb{R}^d)\subset\mathscr{S}(\mathbb{R}^d)$. The inverse Fourier formula will then show that $\mathscr{F}(\mathscr{S}(\mathbb{R}^d)) = \mathscr{S}(\mathbb{R}^d)$. Let $\varphi\in\mathscr{S}(\mathbb{R}^d)$. Then, for all $\alpha,\beta\in\mathbb{N}^d$, and we have $x^\alpha D^\beta f\in L^1(\mathbb{R}^d)$, which shows by the Riemann-Lebesgue lemma (Theorem 2.7.1) and Proposition 2.7.2 that

$$\|\widehat{\varphi}\|_{\alpha,\beta} = \sup_{\xi\in\mathbb{R}^d}|\xi^\beta||D^\alpha\widehat{\varphi}(\xi)| = \sup_{\xi\in\mathbb{R}^d}|\mathscr{F}(x^\alpha D^\beta\varphi)(\xi)|$$

Now, we have for all $\xi\in\mathbb{R}^d$

$$\begin{aligned}
\left|\mathscr{F}(x^\alpha D^\beta\varphi)(\xi)\right| &= \left|\int_{\mathbb{R}^d} x^\alpha D^\beta\varphi(x)e^{-ix\cdot\xi}dx\right| \\
&\leq \int_{B(0,1)}|x^\alpha||D^\beta\varphi(x)|dx + \left\|x^\alpha|x|^{2d}D^\beta\varphi\right\|_{\mathrm{L}^\infty(\mathbb{R}^d)}\int_{\mathbb{R}^d\setminus\overline{B}(0,1)}\frac{dx}{|x|^{2d}} \\
&\leq \alpha(d)\,\|\varphi\|_{\beta,\alpha} + \frac{\beta(d)}{d}\,\|\varphi\|_{\beta,\alpha+2d\,e_0} = \alpha(d)\left(\|\varphi\|_{\beta,\alpha} + \|\varphi\|_{\beta,\alpha+2d\,e_0}\right),
\end{aligned}$$

where $e_0 = (1,\cdots,1)$. Finally, we get the inequality

$$\|\widehat{\varphi}\|_{\alpha,\beta} \leq \alpha(d)\left(\|\varphi\|_{\beta,\alpha} + \|\varphi\|_{\beta,\alpha+2d\,e_0}\right). \tag{2.7.7}$$

Therefore, we have $\mathscr{F}(\mathscr{S}(\mathbb{R}^d))\subset\mathscr{S}(\mathbb{R}^d)$, which shows as we said above that $\mathscr{F}(\mathscr{S}(\mathbb{R}^d)) = \mathscr{S}(\mathbb{R}^d)$, since $\mathscr{F}^{-1} = (2\pi)^{-d}\mathscr{F}\circ\iota$.

Step 2: Density of test functions in the space of Schwartz functions.

Let $\varphi\in\mathscr{S}(\mathbb{R}^d)$, and $\{\eta_n\}_{n\in\mathbb{N}}\subset\mathscr{D}(\mathbb{R}^d)$ such that $\eta_n:\mathbb{R}^d\to[0,1]$, $\eta_n = 1$ on $B(0,n)$ and $\mathrm{supp}(\eta_n)\subset B(0,n+1)$ for all $n\in\mathbb{N}$, and consider $\varphi_n = \eta_n\varphi$. Then, we have for all $(p,\alpha)\in\mathbb{N}\times\mathbb{N}^d$

$$\begin{aligned}
&\sup_{x\in\mathbb{R}^d}(1+|x|)^p|D^\alpha(\varphi_n(x)-\varphi(x))| \leq \sup_{x\in B(0,n+1)\setminus\overline{B}(0,n)}(1-\eta_n)(1+|x|)^p|D^\alpha\varphi(x)| \\
&+\sum_{\substack{\beta\leq\alpha\\ \beta\neq 0}}\binom{\alpha}{\beta}\sup_{x\in B(0,n+1)\setminus\overline{B}(0,n)}|D^\beta\eta_n(x)D^{\alpha-\beta}\varphi(x)| \\
&\leq \sup_{x\in B(0,n+1)\setminus\overline{B}(0,n)}(1+|x|)^p|D^\alpha\varphi(x)| + C\sum_{\substack{\beta\leq\alpha\\ \beta\neq 0}}\sup_{x\in B(0,n+1)\setminus\overline{B}(0,n)}(1+|x|)^p|D^{\alpha-\beta}\varphi(x)| \underset{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

This concludes the proof of this lemma.

Step 3: Fréchet property.

It is very easy to show that $\mathscr{S}(\mathbb{R}^d)$ is a topological vector space, and since its topology is defined by a countable family of semi-norms, it is metrisable. Finally, the Cauchy property follows *mutadis mutandis* from the steps of the proof of Theorem 2.5.15 and we omit this *easier* proof. □

We can now move to the definition of tempered distributions.

Definition 2.7.7. The space of of *tempered distributions*, denoted by $\mathscr{S}'(\mathbb{R}^d)$ is the dual space of the Schwartz space $\mathscr{S}(\mathbb{R}^d)$.

Since $\mathscr{D}(\mathbb{R}^d)$ is a dense subset of $\mathscr{S}(\mathbb{R}^d)$, a tempered distribution $T \in \mathscr{S}'(\mathbb{R}^d)$ can also be seen as an element of $\mathscr{D}'(\mathbb{R}^d)$, and it is defined by its values on $\mathscr{D}(\mathbb{R}^d)$.

Definition 2.7.8. For all $T \in \mathscr{S}'(\mathbb{R}^d)$, its Fourier transform $\mathscr{F}(T) = \widehat{T}$ is the tempered distribution such that for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$, we have

$$\langle \mathscr{F}(T), \varphi\rangle = \langle T, \mathscr{F}(\varphi)\rangle.$$

That the Fourier transform maps $\mathscr{S}'(\mathbb{R}^d)$ into $\mathscr{S}'(\mathbb{R}^d)$ follows from the above discussion and the invariance property of $\mathscr{S}(\mathbb{R}^d)$ by $\mathscr{F}$. By the Fourier inversion formula, the Fourier transform is in fact an isometry.

Examples 2.7.9.

1. We have $\mathscr{F}(1) = (2\pi)^d\delta_0$. Indeed, by the Fourier inversion formula, we have

$$\langle \mathscr{F}(1), \varphi\rangle = \int_{\mathbb{R}^d} \widehat{\varphi}(\xi)d\xi = \int_{\mathbb{R}^d} \widehat{\varphi}(\xi)e^{i\xi\cdot 0}d\xi = (2\pi)^d\varphi(0) = (2\pi)^d\delta_0(\varphi).$$

Likewise, we have

$$\langle \mathscr{F}(\delta_0), \varphi\rangle = \widehat{\varphi}(0) = \int_{\mathbb{R}^d} \varphi(x)e^{-ix\cdot 0}dx = \langle 1, \varphi\rangle,$$

which shows that $\widehat{\delta_0} = 1$ and $\widehat{1} = (2\pi)^d\delta_0$, which is obviously consistent with the Fourier inversion formula.

2. Likewise, computing the Fourier transform of polynomials (that are trivially tempered distributions) is easy. Fix some $\alpha \in \mathbb{N}^d$. Then, we have

$$\langle \mathscr{F}(x^\alpha), \varphi\rangle = \int_{\mathbb{R}^d} \xi^\alpha \widehat{\varphi}(\xi)d\xi = (2\pi)^d i^{|\alpha|} D^\alpha\delta_0(\varphi)$$

Indeed, for all $x \in \mathbb{R}^d$, we have by integration by parts and the Fourier inversion formula

$$\int_{\mathbb{R}^d} \xi^\alpha\widehat{\varphi}(\xi)e^{ix\cdot\xi}d\xi = i^{|\alpha|}D_x^\alpha\int_{\mathbb{R}^d}\widehat{\varphi}(\xi)e^{ix\cdot\xi}d\xi = i^{|\alpha|}D_x^\alpha\left((2\pi)^d\varphi(x)\right) = (2\pi)^d i^{|\alpha|}D^\alpha\varphi(x).$$

Therefore, we have

$$\widehat{x^\alpha} = (2\pi)^d i^{|\alpha|}D^\alpha\delta_0,$$

while

$$\begin{aligned}\langle \mathscr{F}(D^\alpha\delta_0), \varphi\rangle &= (-1)^{|\alpha|}D^\alpha\widehat{\varphi}(0) = (-1)^{|\alpha|}D_\xi^\alpha\left(\int_{\mathbb{R}^d}\varphi(x)e^{-ix\cdot\xi}dx\right)_{|\xi=0}\\ &= (-1)^{|\alpha|}\int_{\mathbb{R}^d}(-i)^{|\alpha|}x^\alpha\varphi(x)dx = \langle i^{|\alpha|}x^\alpha, \varphi\rangle,\end{aligned}$$

which implies that

$$\widehat{D^\alpha\delta_0} = i^{|\alpha|}x^\alpha.$$

Once more, we see that those results are consistent with the Fourier inversion formula.

3. Let us now compute the Fourier transform of $\mathrm{v.p.}\dfrac{1}{x} \in \mathscr{S}'(\mathbb{R})$. We have by Example 2.5.12 the identity

$$x \cdot \mathrm{v.p.}\frac{1}{x} = 1.$$

Let $u \in \mathscr{S}'(\mathbb{R})$ such that $\widehat{u} = \mathrm{v.p.}\dfrac{1}{x}$. Then, recalling that $\widehat{-i\,\delta_0'} = x$, we have

$$\mathscr{F}(-i\,\delta_0')\mathscr{F}(u) = 1 = \mathscr{F}(\delta_0).$$

Furthermore, by the property of Fourier transform on convolution, we have

$$\mathscr{F}(-i\,\delta_0')\mathscr{F}(u) = \mathscr{F}(-i\,\delta_0' * u) = -i\,\mathscr{F}(\delta_0 * u') = -i\,\mathscr{F}(u'),$$

and the previous equation becomes

$$-i\,\mathscr{F}(u') = \mathscr{F}(\delta_0).$$

Since $\mathscr{F}$ is an automorphisme on $\mathscr{S}'(\mathbb{R}^d)$, we deduce that

$$u' = i\,\delta_0 = i\,H',$$

where H is the Heaviside function. Therefore, we have

$$(u - iH)' = 0.$$

We deduce that there exists $c \in \mathbb{C}$ such that

$$u = iH + c.$$

However, since $\mathrm{v.p.}\frac{1}{x}$ is an odd distribution, its (inverse) Fourier transform is an odd distribution (the proof is immediate by a change of variable), which implies that $c = -\frac{i}{2}$, and

$$\mathscr{F}^{-1}\left(\mathrm{v.p.}\frac{1}{x}\right)(\xi) = \frac{i}{2}\mathrm{sgn}(\xi).$$

The inverse Fourier transform (notice the change of sign!) shows that

$$\mathscr{F}\left(\mathrm{v.p.}\frac{1}{x}\right)(\xi) = -i\,\pi\,\mathrm{sgn}(\xi).$$

Conversely, we have (without using tricks this time)

$$\langle \mathscr{F}(\mathrm{sgn}(x)), \varphi\rangle = -\int_{-\infty}^{0} \widehat{\varphi}(\xi)d\xi + \int_{0}^{\infty} \widehat{\varphi}(\xi)d\xi = \lim_{\varepsilon \to 0}\left(-\int_{-\frac{1}{\varepsilon}}^{-\varepsilon} \widehat{\varphi}(\xi)d\xi + \int_{\varepsilon}^{\frac{1}{\varepsilon}} \widehat{\varphi}(\xi)d\xi\right).$$

The truncation that we have made allows us to use Fubini's theorem and get

$$\begin{aligned}
&-\int_{-\frac{1}{\varepsilon}}^{-\varepsilon} \widehat{\varphi}(\xi)d\xi + \int_{\varepsilon}^{\frac{1}{\varepsilon}} \widehat{\varphi}(\xi)d\xi = \int_{\mathbb{R}} \varphi(x)\left(-\int_{-\frac{1}{\varepsilon}}^{-\varepsilon} e^{-ix\xi} + \int_{\varepsilon}^{\frac{1}{\varepsilon}} e^{-ix\xi}d\xi\right) \\
&= -2i\int_{\mathbb{R}} \varphi(x)\left(\int_{\varepsilon}^{\frac{1}{\varepsilon}} \sin(x \cdot \xi)d\xi\right) = -2i\int_{\mathbb{R}} \varphi(x)\frac{\cos(\varepsilon x) - \cos\left(\frac{x}{\varepsilon}\right)}{x}dx \\
&= -i\int_{\mathbb{R}} \frac{\varphi(x) - \varphi(-x)}{x}\left(\cos(\varepsilon x) - \cos\left(\frac{x}{\varepsilon}\right)\right)dx.
\end{aligned}$$

Notice that for all $\varphi \in \mathscr{S}(\mathbb{R})$, we have (by dominated convergence for example)

$$\int_{\mathbb{R}} \frac{\varphi(x) - \varphi(-x)}{x}\cos(\varepsilon x)dx \underset{\varepsilon \to 0}{\longrightarrow} \int_{\mathbb{R}} \frac{\varphi(x) - \varphi(-x)}{x}dx = \left\langle \mathrm{p.v.}\frac{1}{x}, \varphi\right\rangle.$$

On the other hand, since $\psi(x) = \frac{\varphi(x)-\varphi(-x)}{x} \in \mathscr{S}(\mathbb{R})$, the Riemann-Lebesgue lemma shows that

$$\int_{\mathbb{R}} \psi(x) \cos\left(\frac{x}{\varepsilon}\right) dx = \mathrm{Re} \int_{\mathbb{R}} \psi(x) e^{i\varepsilon^{-1}x} dx \underset{\varepsilon\to 0}{\longrightarrow} 0,$$

which finally shows that

$$\mathscr{F}(\mathrm{sgn}(\xi)) = -2i\, \mathrm{p.v.}\frac{1}{x}. \tag{2.7.8}$$

Therefore, we have

$$-2\pi\, \mathrm{sgn}(\xi) = \mathscr{F}^2(\mathrm{sgn}(\xi)) = -2i\, \mathscr{F}\left(\mathrm{p.v.}\frac{1}{x}\right) = -2i\left(-i\pi\, \mathrm{sgn}(\xi)\right) = -2\pi\, \mathrm{sgn}(\xi)$$

as the Fourier inversion formula predicts.

For other examples, refer to [13] (p. 385).

The Hilbert transform $H : \mathscr{S}(\mathbb{R}) \to \mathscr{S}'(\mathbb{R})$ is defined by

$$H(\varphi) = \left(\frac{1}{\pi}\mathrm{p.v.}\frac{1}{x}\right) * \varphi,$$

that is, for all $x \in \mathbb{R}$,

$$H(\varphi)(x) = \lim_{\varepsilon\to 0} \int_{|y|>\varepsilon} \frac{\varphi(x-y)}{\pi y} dy.$$

Then, the previous result and Parceval formula show that H extends to an isometry $H : L^2(\mathbb{R}) \to L^2(\mathbb{R})$, since

$$\widehat{H(\varphi)} = i\, \mathrm{sgn}(\xi)\widehat{\varphi},$$

so that

$$\int_{\mathbb{R}} |H(\varphi)(x)|^2 dx = \frac{1}{2\pi}\int_{\mathbb{R}} |\widehat{H(\varphi)}(\xi)|^2 d\xi = \frac{1}{2\pi}\int_{\mathbb{R}} |\widehat{\varphi}(\xi)|^2 d\xi = \int_{\mathbb{R}} |\varphi(x)|^2 dx.$$

In fact, H belongs to a general class of bounded operators that well-behaved on L^p spaces, called *Calderón-Zygmund* operators. They are the basic objects studied in harmonic analysis, but their study goes beyond the scope of this course.

Application to the Laplace equation.

Let $d \geq 2$, $f \in \mathscr{S}(\mathbb{R}^d)$ and $u \in \mathscr{S}'(\mathbb{R}^d)$ such that

$$\Delta u = f.$$

Then, we have

$$-|\xi|^2\widehat{u}(\xi) = \widehat{f}(\xi),$$

which shows that

$$u = \mathscr{F}^{-1}\left(-\frac{1}{|\xi|^2}\widehat{f}(\xi)\right) = -\mathscr{F}^{-1}\left(\frac{1}{|\xi|^2}\right) * f \quad \text{for } d \geq 2,$$

whilst

$$u = -\mathscr{F}^{-1}\left(\mathrm{p.v.}\frac{1}{|\xi|^2}\right) * f \quad \text{for } d = 2.$$

We see that for $n \geq 2$, since the Fourier transform of a positive distribution is positive, the Fourier transform of radial distribution is radial, and the Fourier transform of a homogenous function of degree $\alpha > 0$ (which is a tempered distribution in particular) is a homogenous distribution of degree $-\alpha - d$, we deduce that

$$\mathscr{F}^{-1}\left(\frac{1}{|\xi|^2}\right)(x) = \frac{c_d}{|x|^{d-2}}$$

for some $c_d > 0$. Therefore, we have

$$u(x) = -c_d \int_{\mathbb{R}^d} \frac{f(y)}{|x-y|^{d-2}} dy.$$

In order to compute the constant $c_d > 0$, we integrate by parts for $r > 0$

$$\begin{aligned}
&\int_{\mathbb{R}^d \setminus \overline{B}(x,r)} \frac{\Delta u(y)}{|x-y|^{d-2}} dy \\
&= -\int_{\partial B(x,r)} \left(\frac{1}{|x-y|^{d-1}} \partial_\nu u(y) - \frac{-(d-2)(x-y)}{|x-y|^d} \cdot \frac{x-y}{|x-y|} u(y) \right) d\mathscr{H}^{d-1}(y) \\
&= -(d-2) \int_{\partial B(x,r)} \frac{u(y)}{r^{d-1}} d\mathscr{H}^{d-1} + \int_{\partial B(x,r)} \frac{\partial_\nu u(y)}{r^{d-2}} d\mathscr{H}^{d-1} \\
&\underset{r\to 0}{\longrightarrow} -(d-2)\mathscr{H}^{d-1}(S^{d-1}) u(x) = -(d-2)\beta(d) u(x),
\end{aligned}$$

where the sign is due to the negative orientation of $\partial B(x,r)$, and the limit follows from the continuity of u at x, and the bound

$$\left| \int_{\partial B(x,r)} \frac{\partial_\nu u(y)}{r^{d-2}} d\mathscr{H}^{d-1} \right| \leq \beta(d) r \left\| \nabla u \right\|_{\mathrm{L}^\infty(\mathbb{R}^d)} \underset{r\to 0}{\longrightarrow} 0.$$

Finally, we get for $d \geq 3$

$$u(x) = -\frac{1}{(d-2)\beta(d)} \int_{\mathbb{R}^d} \frac{f(y)}{|x-y|^{d-2}} dy.$$

For $d = 2$, we need to define the notion of (Hadamard) finite part in more generality.

Proposition 2.7.10. *Let $d \geq 1$, and $m \geq d$. Then, there exists a rational fraction Q in ε^{-1} and $\log(1/\varepsilon)$ such that for all $\varphi \in \mathscr{D}(\mathbb{R}^d)$, we have*

$$\int_{\mathbb{R}^d \setminus \overline{B}(0,\varepsilon)} |x|^{-m} \varphi(x) dx = P\left(\frac{1}{\varepsilon}, \log\left(\frac{1}{\varepsilon}\right), \{D^\alpha \varphi(0)\}_{\alpha \in \mathbb{N}^d} \right) + R_\varepsilon(\varphi),$$

where each monomial of P diverges for a suitable choice of φ as $\varepsilon \to 0$, and where $R_\varepsilon(\varphi)$ converges to a limit as $\varepsilon \to 0$. Then, this limit is a distribution of finite order $\max\{1, [m] - d\}$*, that we denote* $\mathrm{f.p.}|x|^{-m}$ *(where "*f.p.*" stands for* finite part*).*

Remark 2.7.11. This process of renormalisation has a fundamental importance in physics (see [41] and [2], [3] for mathematical applications of those ideas).

Proof. It follows immediately from the Taylor formula. By symmetry, all non-even polynomials in the Taylor expansion of φ have zero integral. Furthermore, since φ has compact support, there exists $R > 0$ such that

$$\begin{aligned}
\int_{\mathbb{R}^d \setminus \overline{B}(0,\varepsilon)} |x|^{-m} \varphi(x) dx &= \int_{B_R \setminus \overline{B}_\varepsilon(0)} |x|^{-m} \varphi(x) dx \\
&= \sum_{|\alpha| \leq [m]-d} \frac{1}{\alpha!} D^\alpha \varphi(0) \int_{B_R \setminus \overline{B}_\varepsilon(0)} \frac{x^\alpha}{|x|^m} dx + \int_{B_R \setminus \overline{B}_r(0)} O(|x|^{d-1+m-[m]}).
\end{aligned}$$

Then, we see that the last term in the right-hand side of the last equation is bounded as $\varepsilon \to 0$. Now, using polar coordinates, we have

$$\int_{B_R\setminus\overline{B}_\varepsilon(0)}\frac{x^\alpha}{|x|^m}dx=\int_\varepsilon^R\frac{1}{r^{m-d-|\alpha|+1}}\left(\int_{S^{d-1}}y^\alpha d\mathscr{H}^{d-1}\right)dr.$$

If some α_j is odd, then by symmetry, we have

$$\int_{S^{d-1}}y^\alpha d\mathscr{H}^{d-1}=0.$$

Otherwise, if $\alpha=2\beta\in(2\mathbb{N})^d$, then

$$\int_{S^{d-1}}y^\alpha d\mathscr{H}^{d-1}=\int_{S^{d-1}}\prod_{j=1}^d|y_j|^{2\beta_j}d\mathscr{H}^{d-1}=F(\beta).$$

To compute this integral, we use a trick of Federer ([15], **3.2.13**). Recall that for all $z\in\mathbb{C}\cap\{z:\mathrm{Re}\,(z)>1\}$, the Gamma function is defined by

$$\Gamma(z)=\int_0^\infty t^{z-1}e^{-t}dt.$$

Now, making the change of variable $t=y^2$, we deduce that

$$\Gamma(z)=2\int_0^\infty y^{2z-1}e^{-y^2}dy=\int_{\mathbb{R}}|y|^{2z-1}e^{-y^2}dy.$$

Therefore, we deduce by the Fubini theorem and polar coordinates that

$$\begin{aligned}\prod_{j=1}^d\Gamma(z_j)&=\int_{\mathbb{R}^d}\prod_{j=1}^d|y_j|^{2z_j-1}e^{-|y|^2}d\mathscr{L}^dy=\int_{S^{d-1}}\prod_{j=1}^d|y_j|^{2z_j-1}d\mathscr{H}^{d-1}(y)\int_0^\infty r^{2\sum_{j=1}^d z_j-1}e^{-r^2}dr\\&=\frac{1}{2}\Gamma\left(\sum_{j=1}^d z_j\right)\int_{S^{d-1}}\prod_{j=1}^d|y_j|^{2z_j-1}d\mathscr{H}^{d-1}(y).\end{aligned}$$

Therefore, we have

$$\int_{S^{d-1}}\prod_{j=1}^d|y_j|^{2z_j-1}d\mathscr{H}^{d-1}(y)=\frac{2\displaystyle\prod_{j=1}^d\Gamma(z_j)}{\Gamma\left(\displaystyle\sum_{j=1}^d z_j\right)}. \tag{2.7.9}$$

In particular, we deduce that

$$F(\beta)=\frac{2\displaystyle\prod_{j=1}^d\Gamma\left(\beta_j+\frac{1}{2}\right)}{\Gamma\left(|\beta|+\frac{1}{2}\right)}=\frac{2\displaystyle\prod_{j=1}^d\Gamma\left(\frac{\alpha_j+1}{2}\right)}{\Gamma\left(\frac{|\alpha|+d}{2}\right)}.$$

Finally, we have provided that $m\notin\mathbb{N}$

$$\begin{aligned}&\sum_{|\alpha|\leq[m]-d}\frac{1}{\alpha!}D^\alpha\varphi(0)\int_{B_R\setminus\overline{B}_\varepsilon(0)}\frac{x^\alpha}{|x|^m}dx\\&=\sum_{\substack{|\alpha|\leq[m]-d\\ \alpha\in(2\mathbb{N})^d}}\frac{1}{\alpha!}\frac{2\displaystyle\prod_{j=1}^d\Gamma\left(\frac{\alpha_j+1}{2}\right)}{\Gamma\left(\frac{|\alpha|+d}{2}\right)}\frac{1}{m-d-|\alpha|}\left(\frac{1}{\varepsilon^{m-d-|\alpha|}}-\frac{1}{R^{m-d-|\alpha|}}\right)D^\alpha\varphi(0).\end{aligned}$$

For $m\in\mathbb{N}$ and $|\alpha|=m-d$, we have

$$\frac{1}{\alpha!}D^\alpha\varphi(0)\int_{B_R\setminus\overline{B}_\varepsilon(0)}\frac{x^\alpha}{|x|^m}dx=\frac{1}{\alpha!}\frac{2\displaystyle\prod_{j=1}^{d}\Gamma\left(\frac{\alpha_j+1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)}\log\left(\frac{R}{\varepsilon}\right)D^\alpha\varphi(0).$$

Therefore, the polynomials are given by the formulae above, and the claim on the order follows immediately from those explicit formulae. □

In the case $d=2$ and $m=2$, we have

$$\begin{aligned}\int_{\mathbb{R}^2\setminus\overline{B}(0,\varepsilon)}\frac{\varphi(x)}{|x|^2}dx&=\int_{B_R\setminus\overline{B}_\varepsilon(0)}\frac{\varphi(0)+O(|x|)}{|x|^2}dx\\&=2\pi\log\left(\frac{1}{\varepsilon}\right)\varphi(0)-2\pi\left(\log R\right)\varphi(0)+\int_{B_R\setminus\overline{B}_\varepsilon(0)}\frac{\varphi(x)-\varphi(0)}{|x|}dx.\end{aligned}$$

Therefore, we deduce that for all $\varphi\in\mathscr{D}(\mathbb{R}^2)$, we have

$$\left\langle \mathrm{f.p.}\frac{1}{|x|^2},\varphi\right\rangle=\lim_{\varepsilon\to 0}\left(\int_{\mathbb{R}^2\setminus\overline{B}(0,\varepsilon)}\frac{\varphi(x)}{|x|^2}dx-2\pi\log\left(\frac{1}{\varepsilon}\right)\varphi(0)\right).$$

In particular, we have

$$\begin{aligned}\left\langle\mathscr{F}\left(\mathrm{f.p.}\frac{1}{|x|^2}\right),\varphi\right\rangle&=\lim_{\varepsilon\to 0}\left(\int_{\mathbb{R}^2\setminus\overline{B}(0,\varepsilon)}\frac{\widehat{\varphi}(\xi)}{|\xi|^2}d\xi-2\pi\log\left(\frac{1}{\varepsilon}\right)\widehat{\varphi}(0)\right)\\&=\lim_{\varepsilon\to 0}\left(\int_{B_{\frac{1}{\varepsilon}}\setminus\overline{B}(0,\varepsilon)}\frac{\widehat{\varphi}(\xi)}{|\xi|^2}d\xi-2\pi\log\left(\frac{1}{\varepsilon}\right)\int_{\mathbb{R}}\varphi(x)dx\right).\end{aligned}$$

To compute the Fourier transform of $\mathscr{F}$, the direct approach is impracticable. However, since this distribution is positive and radial, its Fourier transform is also radial and can easily be computed by an approximating method. Indeed, the Fourier transform is a continuous map $\mathscr{S}'(\mathbb{R}^d)\to\mathscr{S}'(\mathbb{R}^d)$. In particular, if a sequence $\{T_n\}_{n\in\mathbb{N}}\subset\mathscr{S}'(\mathbb{R}^d)$ converges towards a tempered distribution $T\in\mathscr{S}(\mathbb{R}^d)$, we have

$$\mathscr{F}(T_n)\underset{n\to\infty}{\longrightarrow}\mathscr{F}(T).$$

Notice that the sequence $\frac{1}{|x|^{2-\varepsilon}}$ diverges in $\mathscr{S}(\mathbb{R}^2)$ as $\varepsilon\to 0$, so we need to proceed in a different way. However, in dimension $d=2$, using the link of harmonic functions and holomorphic functions, we easily see that the Green's function is given by $G(x)=\frac{1}{2\pi}\log|x|$. In particular, we must show that the inverse Fourier transform of $-\mathrm{f.p.}\frac{1}{|x|^2}$ is (up to constant terms) $\log|x|$, or, equivalently, that the Fourier transform of $\log|x|$ is (up to Dirac masses at 0) $-2\pi\,\mathrm{f.p.}\frac{1}{|x|^2}$. We first compute the Fourier transform of functions $|x|^\alpha$ in $\mathbb{R}^d$ for $\alpha\in\mathbb{C}$ non-singular (the meaning of this word will be clarified later).

First assume that $-d<\alpha<0$. Then, since $|x|^\alpha$ is a radial positive function of degree α, which implies that its Fourier transform is a radial positive distribution of degree $-d-\alpha$. For $|\alpha|<\dfrac{d}{2}$, $|x|^\alpha\in\mathscr{D}'_{\mathrm{L}^2}(\mathbb{R}^d)$, which implies that its Fourier transform is a function. The above properties shows that there exists $c_\alpha\in\mathbb{R}_+^*$ such that

$$\mathscr{F}(|x|^\alpha)(\xi)=\frac{c_\alpha}{|\xi|^{d+\alpha}}.$$

To compute the constant $c_\alpha>0$, we use a trick due to Deny (see [35]). Recall that for all $\beta>0$, we have

$$\mathscr{F}(e^{-\beta|x|^2})=\left(\frac{\pi}{\beta}\right)^{\frac{d}{2}}e^{-\frac{|\xi|^2}{4\beta}}.$$

Therefore, using the Parceval formula with $\beta = \frac{1}{2}$, we find that

$$\int_{\mathbb{R}^d} |x|^\alpha e^{-\frac{|x|^2}{2}} dx = \frac{c_\alpha}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} |\xi|^{-d-\alpha} e^{-\frac{|\xi|^2}{2}} d\xi.$$

The integral on the left-hand side converges for $\alpha > -d$, and the one of the right-hand side for $-d-\alpha > -d$, *i.e.* for $\alpha < 0$. Therefore, both integral converge as expected. Now, using polar coordinates, we deduce that

$$\begin{aligned}\int_{\mathbb{R}^d} |x|^\alpha e^{-\frac{|x|^2}{2}} dx &= \beta(d) \int_0^\infty r^{\alpha+d-1} e^{-\frac{r^2}{2}} dr = \beta(d) \int_0^\infty (2t)^{\frac{\alpha+d-1}{2}} e^{-t} \frac{dt}{\sqrt{2t}} \\ &= 2^{\frac{\alpha+d}{2}-1} \beta(d) \int_0^\infty t^{\frac{\alpha+d}{2}-1} e^{-t} dt = 2^{\frac{\alpha+d}{2}-1} \Gamma\left(\frac{\alpha+d}{2}\right) \beta(d),\end{aligned}$$

by making the change of variable

$$r = \sqrt{2t} \implies dr = \frac{dt}{\sqrt{2t}}.$$

Therefore, replacing α by $-\alpha-d$, we get

$$\int_{\mathbb{R}^d} |\xi|^{-d-\alpha} e^{-\frac{|\xi|^2}{2}} d\xi = 2^{\frac{-\alpha}{2}-1} \Gamma\left(-\frac{\alpha}{2}\right) \beta(d),$$

which finally implies that

$$c_\alpha = (2\pi)^{\frac{d}{2}} 2^{\alpha+\frac{d}{2}} \frac{\Gamma\left(\frac{\alpha+d}{2}\right)}{\Gamma\left(-\frac{\alpha}{2}\right)} = 2^{\alpha+d} \pi^{\frac{d}{2}} \frac{\Gamma\left(\frac{\alpha+d}{2}\right)}{\Gamma\left(-\frac{\alpha}{2}\right)},$$

and

$$\mathscr{F}\left(|x|^\alpha\right)(\xi) = 2^{\alpha+d} \pi^{\frac{d}{2}} \frac{\Gamma\left(\frac{\alpha+d}{2}\right)}{\Gamma\left(-\frac{\alpha}{2}\right)} \frac{1}{|\xi|^{\alpha+d}} \qquad -\frac{d}{2} < \alpha < 0. \tag{2.7.10}$$

Exchanging α into $-\alpha-d$, we see that (2.7.10) is true for all $-d < \alpha < 0$. Furthermore, by analytic continuation, this formula is also true for $0 > \mathrm{Re}\,(\alpha) > -d$, and extends outside of poles of the two Gamma functions to all values $\alpha \in \mathbb{C}$, as long as one replaces the functions by their principal values. For poles of Γ, the regularisation is more complicated, and we will (refer to [35]) first treat the case of interest of $\alpha = d = 2$.

Now, we will first compute the Fourier transform of $\log|x|$, which will *en passant* give us an approximation of $\mathrm{f.p.}\frac{1}{|x|^2}$, that could also be found directly.

First, notice that for all $x \in \mathbb{R}^2 \setminus \{0\}$, we have

$$\log|x| = \lim_{\varepsilon \to 0} \varepsilon^{-1}\left(1 - |x|^{-\varepsilon}\right).$$

Therefore, we also get

$$\varepsilon^{-1}(1-|x|^{-\varepsilon}) \underset{\varepsilon\to 0}{\longrightarrow} \log|x| \quad \text{in } \mathscr{S}'(\mathbb{R}^2).$$

Now, recalling Examples 2.7.9 **1**), we get

$$\mathscr{F}\left(\varepsilon^{-1}\left(1-|x|^{-\varepsilon}\right)\right) = \varepsilon^{-1}\left((2\pi)^2\delta_0 - 2^{2-\varepsilon}\pi \frac{\Gamma\left(1-\frac{\varepsilon}{2}\right)}{\Gamma\left(\frac{\varepsilon}{2}\right)} |\xi|^{-2+\varepsilon}\right).$$

Recall the following Taylor and Laurent expansions:

$$\begin{aligned}\Gamma(z) &= 1 - \gamma(z-1) + O((z-1)^2) \\ \Gamma(z) &= \frac{1}{z} - \gamma + O(z),\end{aligned}$$

where γ is the Euler constant. Therefore, we have

$$\begin{aligned}2^{2-\varepsilon}\pi\frac{\Gamma(1-\frac{\varepsilon}{2})}{\Gamma\left(\frac{\varepsilon}{2}\right)} &= 4\pi\left(1-\varepsilon\log(2)+O(\varepsilon^2)\right)\frac{1+\frac{\gamma}{2}\varepsilon+O(\varepsilon^2)}{\frac{2}{\varepsilon}(1-\frac{\gamma}{2}\varepsilon+O(\varepsilon^2))}\\ &= 2\pi\varepsilon\left(1+(\gamma-\log(2))\varepsilon+O(\varepsilon^2)\right)\end{aligned}$$

Since $\Gamma(z)=\frac{1}{z}+O(1)$ as $z\to 0$, and $\Gamma(1)=1$, we get

$$2^{2-\varepsilon}\pi\frac{\Gamma\left(1-\frac{\varepsilon}{2}\right)}{\Gamma\left(\frac{\varepsilon}{2}\right)}|\xi|^{-2+\varepsilon}=2\pi\varepsilon\left(1+(\gamma-\log(2))\varepsilon+O(\varepsilon^2)\right)|\xi|^{-2+\varepsilon}.$$

Notice that for all $\varphi\in\mathscr{D}(\mathbb{R}^2,\mathbb{R})$, we have (provided that $\mathrm{supp}(\varphi)\subset B(0,R)$ for some $R>0$

$$\int_{\mathbb{R}^2}\frac{\varphi(x)}{|x|^{2-\varepsilon}}dx=\int_{B(0,R)}\frac{\varphi(0)}{|x|^{2-\varepsilon}}dx+\int_{\mathbb{R}^2}\frac{\varphi(x)-\varphi(0)}{|x|^{2-\varepsilon}}dx=2\pi\frac{R^\varepsilon}{\varepsilon}\varphi(0)+\int_{B(0,R)}\frac{\varphi(x)-\varphi(0)}{|x|^2}dx.$$

Therefore, we have

$$\begin{aligned}&\left\langle\varepsilon^{-1}\left(2\pi\delta_0-\varepsilon(1+O(\varepsilon))|\xi|^{-2+\varepsilon}\right),\varphi\right\rangle=2\pi\frac{1-R^\varepsilon}{\varepsilon}\varphi(0)-\int_{B(0,R)}\frac{\varphi(x)-\varphi(0)}{|x|^2}dx\\ &-2\pi(\gamma-\log(2))R^\varepsilon\varphi(0)+O(\varepsilon)\int_{B(0,R)}\frac{\varphi(x)-\varphi(0)}{|x|^2}dx+O(\varepsilon)\varphi(0)\\ &\underset{\varepsilon\to 0}{\longrightarrow}-2\pi\log(R)\varphi(0)-\int_{B(0,R)}\frac{\varphi(x)-\varphi(0)}{|x|^2}dx-2\pi(\gamma-\log(2))\varphi(0).\end{aligned}$$

Since

$$\begin{aligned}\left\langle\mathrm{f.p.}\frac{1}{|x|^2},\varphi\right\rangle&=\lim_{\varepsilon\to 0}\left(\int_{\mathbb{R}^2\setminus\overline{B}(0,\varepsilon)}\frac{\varphi(x)}{|x|^2}dx-2\pi\log\left(\frac{1}{\varepsilon}\right)\varphi(0)\right)\\ &=\lim_{\varepsilon\to 0}\left(\int_{B_R\setminus\overline{B}_\varepsilon(0)}\frac{\varphi(x)}{|x|^2}dx-2\pi\log\left(\frac{1}{\varepsilon}\right)\varphi(0)\right)\\ &=\lim_{\varepsilon\to 0}\left(\int_{B_R\setminus\overline{B}_\varepsilon(0)}\frac{\varphi(0)}{|x|^2}-2\pi\log\left(\frac{1}{\varepsilon}\right)\right)+\int_{B_R\setminus\overline{B}_\varepsilon(0)}\frac{\varphi(x)-\varphi(0)}{|x|^2}dx\\ &=2\pi\log(R)\varphi(0)+\int_{B_R\setminus\overline{B}_\varepsilon(0)}\frac{\varphi(x)-\varphi(0)}{|x|^2}dx,\end{aligned}$$

we deduce that

$$\mathscr{F}(\log|x|)(\xi)=-2\pi\,\mathrm{f.p.}\frac{1}{|\xi|^2}-(2\pi)^2(\gamma-\log(2))\delta_0.\tag{2.7.11}$$

Conversely, the previous discussion shows that

$$\frac{1}{|x|^{2-\varepsilon}}-\frac{2\pi}{\varepsilon}\delta_0\underset{\varepsilon\to 0}{\longrightarrow}\mathrm{f.p.}\frac{1}{|x|^2}\quad\text{in }\mathscr{S}'(\mathbb{R}^2).$$

Using formula (2.7.10), we get

$$\begin{aligned}\mathscr{F}\left(\frac{1}{|x|^{2-\varepsilon}}\right)(\xi)&=2^\varepsilon\pi\frac{\Gamma\left(\frac{\varepsilon}{2}\right)}{\Gamma\left(1-\frac{\varepsilon}{2}\right)}|\xi|^{-\varepsilon}=\pi(1+\varepsilon\log(2)+O(\varepsilon^2))\frac{\frac{2}{\varepsilon}-\gamma+O(\varepsilon)}{1+\frac{\gamma}{2}\varepsilon+O(\varepsilon^2)}|\xi|^{-\varepsilon}\\ &=\frac{2\pi}{\varepsilon}(1+(\log(2)-\gamma)\varepsilon+O(\varepsilon^2))|\xi|^{-\varepsilon}.\end{aligned}$$

First, we trivially have $\varepsilon|\xi|^{-\varepsilon}\underset{\varepsilon\to 0}{\longrightarrow}0$ in $\mathscr{S}'(\mathbb{R}^2)$, and since $\widehat{\delta_0}=1$, we have

$$\mathscr{F}\left(\frac{1}{|x|^{2-\varepsilon}}-\frac{2\pi}{\varepsilon}\delta_0\right)=\frac{2\pi}{\varepsilon}\left(|\xi|^{-\varepsilon}-1\right)+2\pi(\log(2)-\gamma)|\xi|^{-\varepsilon}+O(\varepsilon)|\xi|^{-\varepsilon}$$

$$\underset{\varepsilon\to 0}{\longrightarrow} -2\pi\log|\xi| - 2\pi(\gamma - \log(2)) \quad \text{in } \mathscr{S}'(\mathbb{R}^2). \tag{2.7.12}$$

Notice that since $\mathscr{F}^2(\log|x|) = (2\pi)^2\log|x|$, formula (2.7.11) implies that

$$(2\pi)^2\log|\xi| = -2\pi\mathscr{F}\left(\text{f.p.}\frac{1}{|x|^2}\right) - (2\pi)^2(\gamma - \log(2)),$$

or

$$\mathscr{F}\left(\text{f.p.}\frac{1}{|x|^2}\right) = -2\pi\log|\xi| - 2\pi(\gamma - \log(2)),$$

which is indeed the formula given by (2.7.12). Notice that the previous computations show that

$$\begin{aligned}\Delta\log|x| &= \Delta\left(\mathscr{F}\left(-\frac{1}{2\pi}\text{f.p.}\frac{1}{|x|^2}\right) + (\gamma - \log(2))\delta_0\right)\\ &= -\mathscr{F}\left(-|x|^2\left(-\frac{1}{2\pi}\text{f.p.}\frac{1}{|x|^2}\right) - (\gamma - \log(2))|x|^2\delta_0\right)\\ &= \mathscr{F}\left(\frac{1}{2\pi}\right) = 2\pi\,\delta_0.\end{aligned}$$

We recover the classical formula

$$u(x) = \frac{1}{2\pi}\int_{\mathbb{R}^2}\log|x-y|f(y)dy,$$

for the solution of the equation $\Delta u = f$ in $\mathbb{R}^2$ (for $f \in \mathscr{S}(\mathbb{R}^d)$).

We saw that computing Fourier transforms of distributions may be rather challenging. However, there is an easy way in the case of distributions with compact support. Let $\mathscr{O}_M(\mathbb{R}^d) \subset C^\infty(\mathbb{R}^d)$ be the space of polynomial growth functions, *i.e.* such that for all $\varphi \in \mathscr{O}_M(\mathbb{R}^d)$, and for all $\alpha \in \mathbb{N}$, there exists $m = m(\varphi, \alpha) \in \mathbb{N}$ such that $(1+|x|)^{-m}D^\alpha\varphi \in L^\infty(\mathbb{R}^d)$.

Proposition 2.7.12. *Let $T \in \mathscr{E}'(\mathbb{R}^d)$. Then, $\widehat{T} \in \mathscr{O}_M(\mathbb{R}^d)$, and for all $\xi \in \mathbb{R}^d$, we have*

$$\widehat{T}(\xi) = \langle T, x \mapsto e^{-ix\cdot\xi}\rangle.$$

Proof. Since T has compact support, and $x \mapsto e^{-ix\cdot\xi} \in \mathscr{E}(\mathbb{R}^d) = C^\infty(\mathbb{R}^d)$, we deduce that the function $\psi(\xi) = \langle T, x \mapsto e^{-ix\cdot\xi}\rangle$ is well-defined and continuous in ξ. Furthermore, by the derivation theorem for distributions depending on a parameter, $\psi \in C^\infty(\mathbb{R}^d)$, and for all $\alpha \in \mathbb{N}^d$, we have

$$D^\alpha\psi(\xi) = \langle T, (-i)^{|\alpha|}x^\alpha e^{-ix\cdot\xi}\rangle$$

Therefore, if $K \subset \mathbb{R}^d$ is a compact set containing the support of T, and T has order $m \in \mathbb{N}$, we deduce that

$$|D^\alpha\psi(\xi)| \leq C\sup_{x\in K, |\beta|\leq m}\left|D_x^\beta\left(x^\alpha e^{-ix\cdot\xi}\right)\right| \leq C_\alpha(1+|\xi|)^m.$$

Using the theorem of integration under the bracket, we deduce that

$$\langle\psi,\varphi\rangle = \int_{\mathbb{R}^d}\left\langle T_x, e^{-ix\cdot\xi}\varphi(\xi)\right\rangle d\xi = \left\langle T_x, \int_{\mathbb{R}^d}e^{-ix\cdot\xi}\varphi(\xi)d\xi\right\rangle = \langle T, \widehat{\varphi}\rangle = \left\langle\widehat{T},\varphi\right\rangle,$$

which shows that $\widehat{T} = \psi$. □

Finally, we prove in a rather general setting the exchange formula between convolution and Fourier transforms.

Theorem 2.7.13. *If $S \in \mathscr{E}'(\mathbb{R}^d)$ and $T \in \mathscr{S}'(\mathbb{R}^d)$, then the following formula holds:*

$$\mathscr{F}(S * T) = \mathscr{F}(S)\mathscr{F}(T). \tag{2.7.13}$$

Proof. Notice that the product makes sense in $\mathscr{S}'(\mathbb{R}^d)$ since $\mathscr{F}(S)$ has a polynomial growth at infinity thanks to Proposition 2.7.12. Using the structure Theorem 2.5.32, and the identity $\mathscr{F}(D^\alpha U) = (-1)^{|\alpha|}\xi^\alpha\mathscr{F}(U)$ for all $U \in \mathscr{S}'(\mathbb{R}^d)$, we need only prove this formula for $S = \mu$, where μ is a Radon measure. Let $\{\varphi_n\}_{n\in\mathbb{N}} \subset \mathscr{S}(\mathbb{R}^d)$ be a sequence such that $\varphi_n \underset{n\to\infty}{\longrightarrow} T$ in the weak topology. Then, for all $n \in \mathbb{N}$, we classically have

$$\mathscr{F}(\mu * \varphi_n) = \mathscr{F}(\mu)\mathscr{F}(\varphi_n).$$

Furthermore, since μ has compact support, the formula

$$\langle \mu * \varphi_n, \psi\rangle = \left\langle \varphi_n, \mu * \widetilde{\psi}\right\rangle$$

shows that $\mu * \varphi_n$ converges to $\mu * T = S * T$ in $\mathscr{S}'(\mathbb{R}^d)$. Therefore, we obtain formula (2.7.13) by taking $n \to \infty$. □

2.8 Appendix

Let us compute the Fourier transform of $\log|x|$ in $\mathbb{R}^d$ for any $d \geq 1$. As previously, we have

$$\log|x| = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left(1 - |x|^{-\varepsilon}\right).$$

Now, we have

$$\begin{aligned}
\mathscr{F}\left(\frac{1}{\varepsilon}\left(1-|x|^{-\varepsilon}\right)\right) &= \frac{1}{\varepsilon}\left((2\pi)^d\delta_0 - 2^{d-\varepsilon}\pi^{\frac{d}{2}}\frac{\Gamma\left(\frac{d}{2}-\frac{\varepsilon}{2}\right)}{\Gamma\left(\frac{\varepsilon}{2}\right)}|\xi|^{-d+\varepsilon}\right)\\
&= \frac{1}{\varepsilon}\left((2\pi)^d\delta_0 - 2^d\pi^{\frac{d}{2}}(1-\varepsilon\log(2)+O(\varepsilon^2))\frac{\Gamma\left(\frac{d}{2}\right)-\frac{1}{2}\Gamma'\left(\frac{d}{2}\right)\varepsilon + O(\varepsilon)^2}{\frac{2}{\varepsilon}\left(1-\frac{\gamma}{2}\varepsilon+O(\varepsilon^2)\right)}|\xi|^{-d+\varepsilon}\right)\\
&= \frac{(2\pi)^d}{\varepsilon}\delta_0 - 2^{d-1}\pi^{\frac{d}{2}}\Gamma\left(\frac{d}{2}\right)|\xi|^{-d+\varepsilon} + 2^d\pi^{\frac{d}{2}}\varepsilon\left(\frac{\gamma}{2} - \frac{1}{2}\frac{\Gamma'\left(\frac{d}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} - \log(2)\right)|\xi|^{-d+\varepsilon}.
\end{aligned}$$

Let $\varphi \in \mathscr{D}(\mathbb{R}^d)$ and $R > 0$ such that $\mathrm{supp}(\varphi) \subset B(0,R)$. Then, using polar coordinates ([15] **3.2.13**)

$$\begin{aligned}
\langle |x|^{-d+\varepsilon}, \varphi\rangle &= \int_{B_R\setminus\overline{B}_\varepsilon(0)}\varepsilon|x|^{-d+\varepsilon}\varphi(x)dx = \varphi(0)\int_{B_R\setminus\overline{B}_\varepsilon(0)}\frac{dx}{|x|^{d-\varepsilon}} + \int_{B_R\setminus\overline{B}_\varepsilon(0)}\frac{\varphi(x)-\varphi(0)}{|x|^{d-\varepsilon}}dx\\
&= \frac{\beta(d)}{\varepsilon}R^\varepsilon\varphi(0) + \int_{B_R\setminus\overline{B}_\varepsilon(0)}\frac{\varphi(x)-\varphi(0)}{|x|^{d-\varepsilon}}dx, \qquad (2.8.1)
\end{aligned}$$

where we wrote $\beta(d) = \mathscr{H}^{d-1}(S^{d-1})$. Since the right-hand side of (2.8.1) is bounded as $\varepsilon \to 0$, we deduce that

$$\left\langle \varepsilon|x|^{-d+\varepsilon}, \varphi\right\rangle \underset{\varepsilon\to 0}{\longrightarrow} \beta(d)\varphi(0) = \langle \beta(d)\delta_0, \varphi\rangle,$$

i.e.

$$\lim_{\varepsilon}\varepsilon|x|^{-d+\varepsilon} = \beta(d)\delta_0 \quad \text{in } \mathscr{S}'(\mathbb{R}^d).$$

Recall by the previous computation (take (2.7.9) with $z_j = \frac{1}{2}$ for all $1 \leq j \leq d$) that

$$\beta(d) = \mathscr{H}^{d-1}(S^{d-1}) = \frac{2\,\Gamma\left(\frac{1}{2}\right)^d}{\Gamma\left(\frac{d}{2}\right)} = \frac{2\,\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}.$$

Therefore, we have for all $\varphi \in \mathscr{D}(\mathbb{R}^d)$

$$\left\langle \frac{(2\pi)^d}{\varepsilon}\delta_0 - 2^{d-1}\pi^{\frac{d}{2}}\Gamma\left(\frac{d}{2}\right)|\xi|^{-d+\varepsilon}, \varphi\right\rangle = (2\pi)^d\frac{\varphi(0)}{\varepsilon} - 2^{d-1}\pi^{\frac{d}{2}} \times \frac{2\,\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}R^\varepsilon\frac{\varphi(0)}{\varepsilon}$$

$$
\begin{aligned}
&- 2^{d-1}\pi^{\frac{d}{2}}\int_{B_R\setminus\overline{B}_\varepsilon(0)}\frac{\varphi(x)-\varphi(0)}{|x|^{d-\varepsilon}}dx\\
&= (2\pi)^d\frac{1-R^\varepsilon}{\varepsilon}\varphi(0) - 2^{d-1}\pi^{\frac{d}{2}}\int_{B_R\setminus\overline{B}_\varepsilon(0)}\frac{\varphi(x)-\varphi(0)}{|x|^{d-\varepsilon}}dx\\
&\underset{\varepsilon\to 0}{\longrightarrow} -(2\pi)^d\log(R) - 2^{d-1}\pi^{\frac{d}{2}}\int_{B_R\setminus\overline{B}_\varepsilon(0)}\frac{\varphi(x)-\varphi(0)}{|x|^{d}}dx = -2^{d-1}\pi^{\frac{d}{2}}\,\mathrm{f.p.}\frac{1}{|x|^d}.
\end{aligned}
$$

Finally, we get

$$
\begin{aligned}
\mathscr{F}\left(\log|x|\right)(\xi) &= -2^{d-1}\pi^{\frac{d}{2}}\,\mathrm{f.p.}\frac{1}{|\xi|^d} + \frac{2^{d+1}\,\pi^d}{\Gamma\left(\frac{d}{2}\right)}\left(-\frac{\gamma}{2}+\frac{1}{2}\frac{\Gamma'\left(\frac{d}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}+\log(2)\right)\delta_0\\
&= -2^{d-2}\Gamma\left(\frac{d}{2}\right)\beta(d)\,\mathrm{f.p.}\frac{1}{|\xi|^d} + \frac{(2\pi)^d}{\Gamma\left(\frac{d}{2}\right)}\left(2\log(2)-\gamma+\frac{\Gamma'\left(\frac{d}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}\right)\delta_0, \qquad (2.8.2)
\end{aligned}
$$

and we check that this formula coincide with the previous one for $d=2$ since $\Gamma(1)=1$ and $\Gamma'(1)=-\gamma$. We also get

$$
(2\pi)^d\log|\xi| = -2^{d-1}\pi^{\frac{d}{2}}\mathscr{F}\left(\mathrm{f.p.}\frac{1}{|x|^d}\right)(\xi) + \frac{(2\pi)^d}{\Gamma\left(\frac{d}{2}\right)}\left(2\log(2)-\gamma+\frac{\Gamma'\left(\frac{d}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}\right),
$$

or

$$
\begin{aligned}
\mathscr{F}\left(\mathrm{f.p.}\frac{1}{|x|^d}\right)(\xi) &= -2\,\pi^{\frac{d}{2}}\log|\xi| + \frac{2\,\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}\left(2\log(2)-\gamma+\frac{\Gamma'\left(\frac{d}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}\right)\\
&= -\Gamma\left(\frac{d}{2}\right)\beta(d)\log|x| + \beta(d)\left(2\log(2)-\gamma+\frac{\Gamma'\left(\frac{d}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}\right). \qquad (2.8.3)
\end{aligned}
$$

We will also check that directly, using the formula

$$
\frac{1}{|x|^{d-\varepsilon}} - \frac{\beta(d)}{\varepsilon}\delta_0 \underset{\varepsilon\to 0}{\longrightarrow} \mathrm{f.p.}\frac{1}{|x|^d} \quad \text{in } \mathscr{S}'(\mathbb{R}^2).
$$

Indeed, we have for all $\varphi\in\mathscr{D}(\mathbb{R}^d)$ such that $\mathrm{supp}(\varphi)\subset B(0,R)$,

$$
\begin{aligned}
&\left\langle \frac{1}{|x|^{d-\varepsilon}} - \frac{\beta(d)}{\varepsilon}\delta_0, \varphi\right\rangle = \int_{B(0,R)}\frac{\varphi(x)}{|x|^{d-\varepsilon}}dx - \frac{\beta(d)}{\varepsilon}\varphi(0)\\
&= \int_{B(0,R)}\frac{\varphi(0)}{|x|^{d-\varepsilon}}dx + \int_{B(0,R)}\frac{\varphi(x)-\varphi(0)}{|x|^{d-\varepsilon}}dx - \frac{\beta(d)}{\varepsilon}\varphi(0)\\
&= \int_0^R\frac{1}{r^{d-\varepsilon}}\left(\int_{S^{d-1}}\varphi(0)d\mathscr{H}^{d-1}\right)r^{d-1}dr + \int_{B(0,R)}\frac{\varphi(x)-\varphi(0)}{|x|^{d-\varepsilon}}dx - \frac{\beta(d)}{\varepsilon}\varphi(0)\\
&= \beta(d)\varphi(0)\int_0^R\frac{dr}{r^{1-\varepsilon}}dr + \int_{B(0,R)}\frac{\varphi(x)-\varphi(0)}{|x|^{d-\varepsilon}}dx\\
&= \beta(d)\frac{R^\varepsilon-1}{\varepsilon}\varphi(0) + \int_{B(0,R)}\frac{\varphi(x)-\varphi(0)}{|x|^{d-\varepsilon}}dx - \frac{\beta(d)}{\varepsilon}\varphi(0)\\
&\underset{\varepsilon\to 0}{\longrightarrow} \beta(d)\log(R)\varphi(0) + \int_{B(0,R)}\frac{\varphi(x)-\varphi(0)}{|x|^d}dx = \left\langle \mathrm{p.f.}\frac{1}{|x|^d},\varphi\right\rangle,
\end{aligned}
$$

since

$$
\begin{aligned}
\int_{\mathbb{R}^d\setminus\overline{B}(0,\varepsilon)}\frac{\varphi(x)}{|x|^d}dx &= \int_{B_R\setminus\overline{B}_\varepsilon(0)}\frac{\varphi(0)}{|x|^d} + \int_{\mathbb{R}^d\setminus\overline{B}(0,\varepsilon)}\frac{\varphi(x)-\varphi(0)}{|x|^d}dx\\
&= \beta(d)\,\log\left(\frac{1}{\varepsilon}\right)\varphi(0) + \beta(d)\log(R)\varphi(0) + \int_{B(0,R)}\frac{\varphi(x)-\varphi(0)}{|x|^d}dx.
\end{aligned}
$$

Now, we have by (2.7.10)

$$\mathscr{F}\left(\frac{1}{|x|^{d-\varepsilon}}-\frac{\beta(d)}{\varepsilon}\delta_0\right)=2^{\varepsilon}\pi^{\frac{d}{2}}\frac{\Gamma\left(\frac{\varepsilon}{2}\right)}{\Gamma\left(\frac{d-\varepsilon}{2}\right)}\frac{1}{|\xi|^{\varepsilon}}-\frac{\beta(d)}{\varepsilon}.$$

Now, recalling that $\beta(d)=\dfrac{2\,\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}$, we deduce that

$$\begin{aligned}2^{\varepsilon}\pi^{\frac{d}{2}}\frac{\Gamma\left(\frac{\varepsilon}{2}\right)}{\Gamma\left(\frac{d-\varepsilon}{2}\right)}&=(1+\varepsilon\log(2)+O(\varepsilon^2))\pi^{\frac{d}{2}}\frac{\frac{2}{\varepsilon}-\gamma+O(\varepsilon)}{\Gamma\left(\frac{d}{2}\right)-\frac{1}{2}\Gamma'\left(\frac{d}{2}\right)\varepsilon+O(\varepsilon^2)}\\&=\frac{2\,\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}\left(1+\left(2\log(2)-\gamma+\frac{\Gamma'\left(\frac{d}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}\right)\varepsilon+O(\varepsilon^2)\right).\end{aligned}$$

Therefore, we have

$$\begin{aligned}2^{\varepsilon}\pi^{\frac{d}{2}}\frac{\Gamma\left(\frac{\varepsilon}{2}\right)}{\Gamma\left(\frac{d-\varepsilon}{2}\right)}\frac{1}{|\xi|^{\varepsilon}}-\frac{\beta(d)}{\varepsilon}&=\frac{\beta(d)}{\varepsilon}(|\xi|^{-\varepsilon}-1)+\beta(d)\left(2\log(2)-\gamma+\frac{\Gamma'\left(\frac{d}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}\right)|\xi|^{-\varepsilon}\\&\underset{\varepsilon\to 0}{\longrightarrow}-\beta(d)\log|\xi|+\beta(d)\left(2\log(2)-\gamma+\frac{\Gamma'\left(\frac{d}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}\right).\end{aligned}$$

Notice that for $d=1$, and for all $\varphi\in\mathscr{D}(\mathbb{R})$ such that $\mathrm{supp}(\varphi)\subset[-R,R]$, we have

$$\begin{aligned}\int_{\mathbb{R}\setminus[-\varepsilon,\varepsilon]}\frac{\varphi(x)}{|x|}dx&=2\int_{\varepsilon}^{R}\frac{\varphi(0)}{x}dx+\int_{[-R,-\varepsilon]\cup[\varepsilon,R]}\frac{\varphi(x)-\varphi(0)}{|x|}dx\\&=2\,\log\left(\frac{1}{\varepsilon}\right)\varphi(0)+2\,\log(R)\,\varphi(0)+\int_{[-R,-\varepsilon]\cup[\varepsilon,R]}\frac{\varphi(x)-\varphi(0)}{|x|}dx.\end{aligned}$$

Therefore, we have

$$\begin{aligned}\left\langle \mathrm{f.p.}\frac{1}{|x|},\varphi\right\rangle&=\lim_{\varepsilon\to 0}\left(\int_{\mathbb{R}\setminus[-\varepsilon,\varepsilon]}\frac{\varphi(x)}{|x|}dx-2\,\log\left(\frac{1}{\varepsilon}\right)\varphi(0)\right)\\&=2\,\log(R)\,\varphi(0)+\int_{-R}^{R}\frac{\varphi(x)-\varphi(0)}{|x|}dx.\end{aligned}$$

We have

$$\begin{aligned}\int_0^R\frac{\varphi(x)-\varphi(0)}{|x|}dx&=\int_0^R\frac{1}{|x|}\left(\int_0^x\varphi'(t)dt\right)dx=\int_0^R\varphi'(t)\left(\int_0^R\frac{\mathbf{1}_{\{0\leq t\leq x\}}}{|x|}dx\right)dt\\&=\int_0^R\varphi'(t)\log\left(\frac{R}{t}\right)dt\\&=-\int_0^R\varphi'(t)\log|t|dt-\log(R)\varphi(0),\end{aligned}$$

whilst

$$\int_{-R}^0\frac{\varphi(x)-\varphi(0)}{|x|}dx=\int_{-R}^0\varphi'(t)\log|t|dt-\log(R)\varphi(0),$$

Therefore, we deduce that

$$\begin{aligned}\left\langle \mathrm{f.p.}\frac{1}{|x|},\varphi\right\rangle&=-\int_0^{\infty}\varphi'(t)\log|t|dt+\int_{-\infty}^0\varphi'(t)\log|t|dt=-\left\langle\mathrm{sgn}(x)\log|x|,\varphi'\right\rangle\\&=\left\langle\frac{d}{dx}\left(\mathrm{sgn}(x)\log|x|\right),\varphi\right\rangle,\end{aligned}$$

which shows that

$$\text{f.p.}\frac{1}{|x|} = \frac{d}{dx}\left(\text{sgn}(x)\log|x|\right), \tag{2.8.4}$$

whilst

$$\text{p.v.}\frac{1}{x} = \frac{d}{dx}\left(\log|x|\right). \tag{2.8.5}$$

Remark 2.8.1. Notice that for any L^1_{loc} function u in $\mathscr{S}'(\mathbb{R}^d)$ whose Fourier transform is a L^1_{loc} function (or a finite part or principal value function), its Fourier transform is given for all non-singular values by the limit

$$\lim_{\varepsilon\to 0}\int_{B_{\frac{1}{\varepsilon}}(0)} e^{-ix\cdot\xi}u(x)dx.$$

However, this formula is of almost none practical use, even for $u(x) = \log|x|$. What are the needed algebraic transformations that would allow one to make appear the Euler constant γ? Therefore, we see that the idea of Taylor expansions and limits is the most efficient way to compute Fourier transforms of rather complicated functions.

Chapter 3

Sobolev Spaces

Let Ω be an open set of $\mathbb{R}^d$ for some $d \geq 1$, that we choose connected for convenience.

3.1 Definition and Basic Properties

Definition 3.1.1. Let $m \in \mathbb{N}$ and $1 \leq p \leq \infty$. A function $u : \Omega \to \mathbb{R}$ belongs to the Sobolev space $W^{m,p}(\Omega)$ if and only if for all $|\alpha| \leq m$, we have $D^\alpha u \in L^p(\Omega)$. If $p = 2$, we commonly write $W^{m,2}(\Omega) = H^m(\Omega)$. We equip $W^{m,p}(\Omega)$ with the following norm

$$\|u\|_{\mathrm{W}^{m,p}(\Omega)} = \sum_{|\alpha| \leq m} \|D^\alpha u\|_{\mathrm{L}^p(\Omega)} . \tag{3.1.1}$$

Theorem 3.1.2. *The space $W^{m,p}(\Omega)$ is a Banach space. The space $W^{m,p}(\Omega)$ is reflexive for $1 < p < \infty$ and separable for $1 \leq p < \infty$. The space $H^m(\Omega)$ is a separable Hilbert space.*

Proof. **Step 1.** $W^{m,p}$ is a Banach space.

Let $\{u_n\}_{n \in \mathbb{N}} \subset W^{m,p}(\Omega)$ be a Cauchy sequence. Since $L^p(\Omega)$ is a Banach space, there exists $u \in L^p(\Omega)$ and for all $0 < |\alpha| \leq m$, there exists $u_\alpha \in L^p(\Omega)$ such that $u_n \underset{n\to\infty}{\longrightarrow} u$ and $D^\alpha u_n \underset{n\to\infty}{\longrightarrow} u_\alpha$. Now, by Hölder's inequality, for all $\varphi \in \mathscr{D}(\Omega)$, we have

$$|\langle u_n, \varphi\rangle - \langle u, \varphi\rangle| \leq \|u_n - u\|_{\mathrm{L}^p(\Omega)} \|\varphi\|_{\mathrm{L}^{p'}(\Omega)} \underset{n\to\infty}{\longrightarrow} 0.$$

Therefore, $u_n \underset{n\to\infty}{\longrightarrow} u$ in the distributional sense, and since derivation is continuous under $\sigma(\mathscr{D}(\Omega), \mathscr{D}'(\Omega))$, we deduce that $u_\alpha = D^\alpha u$ for all $|\alpha| \leq m$, which concludes the proof.

Step 2. Other properties.

We have an isometry $W^{m,p}(\Omega) \to L^p(\Omega)^{N(d,m)}$ given by the natural map $u \mapsto \{D^\alpha u\}_{|\alpha| \leq m}$, where

$$N(d, m) = \mathrm{card}(\mathbb{N}^d \cap \{\alpha : |\alpha| \leq m\} .$$

In particular, $W^{m,p}(\Omega)$ is a closed space of $L^p(\Omega)^{N(d,m)}$, which implies the claims on reflexivity and separability. ☐

Remark 3.1.3. There are many generalisations of Sobolev spaces, using more complicated norms or weaker notions than functions. We will not list them all, but let us nevertheless mention the important class of function of *bounded variations*, commonly called BV functions, that are L^1 functions whose distributional derivative is a Radon measure. Those functions have applications to the study of minimal surfaces, and we send to Giusti's monograph for more details ([18]).

Theorem 3.1.4. *Let $u \in W^{m,p}(\Omega)$, with $1 \leq p < \infty$. Then, there exists a sequence $\{u_n\}_{n\in\mathbb{N}} \subset \mathscr{D}(\mathbb{R}^d)$ such that*

$$\begin{cases} \|u_n - u\|_{\mathrm{L}^p(\Omega)} \underset{n\to\infty}{\longrightarrow} 0 \\ \|D^\alpha(u_n - u)\|_{\mathrm{L}^p(\Omega')} \underset{n\to\infty}{\longrightarrow} 0 \qquad \text{for all } \Omega' \subset\subset \Omega. \end{cases} \tag{3.1.2}$$

Proof. Let $\{\rho_n\}_{n\in\mathbb{N}} \subset \mathscr{D}(\mathbb{R}^d)$ be an approximation of unity, *i.e.* a non-negative function with integral 1, support included in $B(0, \frac{1}{n})$, and such that $\rho_n \underset{n\to\infty}{\longrightarrow} \delta_0$ in $\mathscr{D}'(\mathbb{R}^d)$. Let $v_n = \rho_n * (u\,\mathbf{1}_\Omega)$. Then, the classical results of convolution show that

$$\|u_n - u\,\mathbf{1}_\Omega\|_{\mathrm{L}^p(\Omega)} \underset{n\to\infty}{\longrightarrow} 0,$$

which shows the first part of (3.1.2). Now, fix some relatively compact open subset Ω' of Ω, and let $\chi \in \mathscr{D}(\Omega)$ such that $\chi = 1$ on an open neighbourhood of Ω'. Then, for $n \in \mathbb{N}$ large enough, we have

$$\rho_n * (\chi\, u) = \rho_n * (u\,\mathbf{1}_\Omega).$$

Indeed, we have

$$\mathrm{supp}\,(\rho_n * (\chi\, u) - \rho_n * (u\,\mathbf{1}_\Omega)) = \mathrm{supp}(\rho_n * ((1_\Omega - \chi)u)) \subset \mathrm{supp}(\rho_n) + \mathrm{supp}(1_\Omega - \chi) \subset \Omega \setminus \overline{\Omega'}$$

for $n \in \mathbb{N}$ large enough. Indeed, $\mathrm{supp}\,(1_\Omega - \chi) \subset \Omega \setminus \overline{\Omega'}$ which is an open set, and since $\mathrm{supp}\,(\rho_n) \subset B(0, \frac{1}{n})$, for n large enough, we also have

$$B\left(0, \frac{1}{n}\right) + \mathrm{supp}(1_\Omega - \chi) \subset \Omega \setminus \overline{\Omega'}.$$

Now, we have by Proposition 2.6.1

$$D^\alpha(\rho_n * (\chi u)) = \rho_n * (u\, D^\alpha \chi + \chi\, D^\alpha u) \qquad \text{in } \mathscr{D}'(\mathbb{R}^d).$$

In particular, we have

$$\|D^\alpha(v_n - u)\|_{\mathrm{L}^p(\Omega')} \underset{n\to\infty}{\longrightarrow} 0.$$

Finally, if $\eta \in C^\infty(\Omega)$ is such that $\eta(t) = 1$ for $t \leq 1$ and $\eta(t) = 0$ for $t \geq 2$, defining $\eta_n(x) = \eta\left(\frac{|x|}{n}\right)$, the sequence $\{u_n = \eta_n\, v_n\}_{n\in\mathbb{N}}$ has the required properties. □

Remark 3.1.5. More generally, the Meyers-Serrin theorem shows that for all $u \in W^{m,p}(\Omega)$, there exists $\{u_n\}_{n\in\mathbb{N}} \subset W^{m,p}(\Omega) \cap C^\infty(\Omega)$ such that $u_n \underset{n\to\infty}{\longrightarrow} u$ in $W^{m,p}(\Omega)$.

3.2 Basic Properties of Sobolev functions

We prove two propositions on the composition and change of variable.

Proposition 3.2.1 (Composition of Sobolev functions)**.** *Let $1 \leq p \leq \infty$, and $G \in C^1(\mathbb{R}) \cap \mathrm{Lip}(\mathbb{R})$ such that $G(0) = 0$. Then, for all $u \in W^{1,p}(\Omega)$, we have $G \circ u \in W^{1,p}(\Omega)$, $\nabla(G \circ u) = G'(u)\nabla u$, and*

$$\|G \circ u\|_{\mathrm{W}^{1,p}(\Omega)} \leq \mathrm{Lip}(G)\, \|u\|_{\mathrm{W}^{1,p}(\Omega)}\,.$$

Proof. If $L = \mathrm{Lip}(G)$, we have by definition $|G(x)| = |G(x) - G(0)| \leq L|x|$ for all $x \in \Omega$. Therefore, we have $|G(u)| \leq L|u|$ and

$$\|G(u)\|_{\mathrm{L}^p(\Omega)} \leq L\, \|u\|_{\mathrm{L}^p(\Omega)}\,.$$

Now, let us show that $\nabla(G\circ u)=G'(u)\nabla u$. Let $\{u_n\}_{n\in\mathbb{N}}\subset\mathscr{D}(\mathbb{R}^d)$ such that

$$\begin{cases}u_n \underset{n\to\infty}{\longrightarrow} u & \text{in } L^p(\Omega)\\ \nabla u_n \underset{n\to\infty}{\longrightarrow} \nabla u & \text{in } L^p(\Omega') \text{ for all } \Omega'\subset\subset\Omega.\end{cases}$$

For all $\varphi\in\mathscr{D}(\Omega,\mathbb{R}^d)$, and for all $n\in\mathbb{N}$, we have by Stokes theorem

$$\int_\Omega (G\circ u_n)\,\mathrm{div}\,\varphi\,dx=-\int_\Omega G'(u_n)\nabla u_n\circ\varphi\,dx.$$

As $G(u_n)\underset{n\to\infty}{\longrightarrow} G(u)$ in $L^p(\Omega)$ and $G'(u_n)\nabla u_n\underset{n\to\infty}{\longrightarrow} G'(u)\nabla u$ in $L^p(\mathrm{supp}(\varphi)+B(0,\varepsilon))$ for $\varepsilon>0$ small enough, we deduce that in the limit

$$\int_\Omega (G\circ u)\,\mathrm{div}\,\varphi\,dx=-\int_\Omega G'(u)\nabla u\cdot\varphi\,dx.$$

For $p=\infty$, we apply the previous argument on $\Omega'=\mathrm{supp}(\varphi)+B(0,\varepsilon)$, and notice that $u\in W^{1,p}(\Omega')$ for all $p<\infty$. □

Remark 3.2.2. More generally, the result would hold for a Lipschitzian function.

Proposition 3.2.3 (Change of variable)**.** *Let $\Phi:U\to V$ be a C^1-diffeomorphism such that* $\mathrm{Jac}(\Phi)\in L^\infty(U)$ *and* $\mathrm{Jac}(\Phi^{-1})\in L^\infty(V)$. *Then, for all $u\in W^{1,p}(V)$, we have $u\circ\Phi\in W^{1,p}(U)$ and for all $1\leq i\leq d$, we have*

$$\partial_{x_i}(u\circ\Phi)(y)=\nabla u(\Phi(y))\cdot\partial_{x_i}\Phi(y).$$

while

$$\|u\circ\Phi\|_{\mathrm{W}^{1,p}(U)}\leq C(\Phi)\,\|u\|_{\mathrm{W}^{1,p}(V)}.$$

Proof. The proof is similar and we omit it. □

3.3 Extension Operator

A lot of properties of Sobolev spaces are easier to prove for the whole space. If $\Omega\neq\mathbb{R}^d$, an extension operator is a continuous linear map $E:W^{m,p}(\Omega)\to W^{m,p}(\mathbb{R}^d)$, such that $E(u)|_\Omega=u$ for all $u\in W^{m,p}(\Omega)$. Those extension operators do not exist for all domains, and we will treat the simple case of C^1 open sets (there are many generalisations of this result, see [1]).

The notion of manifolds with boundary will be relevant in this chapter, so we identify $\mathbb{R}^d$ with $\mathbb{R}^{d-1}\times\mathbb{R}$, and write $x=(x',x_d)\in\mathbb{R}^{d-1}\times\mathbb{R}$ those coordinates.

$$\begin{aligned}\mathbb{R}^d_+&=\mathbb{R}^d\cap\{(x',x_d):x_d>0\}\\ \mathbb{R}^{d-1}&=\partial\mathbb{R}^d_+=\mathbb{R}^{d-1}\times\{0\}.\end{aligned}$$

Definition 3.3.1. We say that an open subset $\Omega\subset\mathbb{R}^d$ is of class C^m—or is a C^m domain—if Ω is a C^1 manifold with boundary. Explicitly, for all $x\in\partial\Omega$, there exists $r>0$ and a C^m diffeomorphism $\varphi:\mathbb{R}^d\to B(x,r)$ such that $\varphi(\mathbb{R}^d_+)=\Omega\cap B(x,r)$, and $\varphi(\partial\mathbb{R}^d_+)=\partial\Omega\cap B(x,r)$.

Let us start by a simple lemma of extension by reflection. By induction, we need only prove the case of derivatives of first order, and we will therefore restrict to this case in the rest of this section.

Lemma 3.3.2 (Extension Lemma)**.** *Let $u\in W^{1,p}(\mathbb{R}^d_+)$, and $u_*:\mathbb{R}^d\to\mathbb{C}$ defined by*

$$u_*(x)=\begin{cases}u(x',x_d) & \textit{for all } x_d>0\\ u(x',-x_d) & \textit{for all } x_d\leq 0.\end{cases}$$

Then, $u_*\in W^{1,p}(\mathbb{R}^d)$, and

$$\|u_*\|_{\mathrm{L}^p(\Omega)}=2\,\|u\|_{\mathrm{L}^p(\Omega)},\qquad \|\nabla u_*\|_{\mathrm{L}^p(\Omega)}=2\,\|\nabla u\|_{\mathrm{L}^p(\Omega)}.$$

Proof. First assume that $u \in C^\infty(\mathbb{R}^d_+)$, and fix some $\varphi \in \mathscr{D}(\mathbb{R}^d)$. Let $\{\theta_n\}_{n\in\mathbb{N}} \subset C^\infty(\mathbb{R})$ be an even function such that $\theta_n = 1$ on $\mathbb{R}\setminus[-\frac{1}{n},\frac{1}{n}]$ and $\mathrm{supp}(\theta_n) = \mathbb{R}\setminus[-\frac{1}{2n},\frac{1}{2n}]$. Furthermore, by approximating a piece-wise linear function, we can assume that $|\theta_n'| \leq 4n$. Letting $\chi_n(x) = \theta_n(x_d)$, for all $\varphi \in \mathscr{D}(\mathbb{R}^d)$ and for all $1 \leq i \leq d-1$, we have

$$\int_{\mathbb{R}^d} u_*\,\partial_{x_i}(\chi_n\varphi)\,dx = \int_{\mathbb{R}^d_+} u\,\partial_{x_i}\left(\chi_n(\varphi+\varphi^*)\right)dx,$$

where $\varphi^*(x) = \varphi(x,-x_d)$, while

$$\int_{\mathbb{R}^d} \widetilde{u}\,\partial_{x_d}(\chi_n\varphi)dx = \int_{\mathbb{R}^d_+} u\,\partial_{x_d}\left(\chi_n(\varphi-\varphi^*)\right)dx.$$

Then, we have for all $1 \leq i \leq d-1$

$$\begin{aligned}\int_{\mathbb{R}^d_+} u\,\chi_n\,\partial_{x_i}(\varphi+\varphi^*)\,dx &= \int_{\mathbb{R}^d_+} u\,\partial_{x_i}(\chi_n(\varphi+\varphi^*))dx = -\int_{\mathbb{R}^d_+}(\varphi+\varphi_*)\,\chi_n\,\partial_{x_i}u\,dx\\ &= -\int_{\mathbb{R}^d_+}\varphi\,\chi_n\partial_{x_i}u_*dx.\end{aligned}$$

Letting $n\to\infty$, we deduce that

$$\int_{\mathbb{R}^d} u\,\partial_{x_i}\varphi\,dx = -\int_{\mathbb{R}^d}\varphi\,\partial_{x_i}u_*\,dx,$$

while

$$\begin{aligned}\int_{\mathbb{R}^d_+} u\,\chi_n\,\partial_{x_d}(\varphi-\varphi^*))dx &= \int_{\mathbb{R}^d_+} u\,\chi_n\,\partial_{x_d}(\chi_n(\varphi-\varphi^*))dx - \int_{\mathbb{R}^d_+} u\,(\varphi-\varphi^*)\partial_{x_d}\chi_n\,dx\\ &= -\int_{\mathbb{R}^d_+}\chi_n(\varphi-\varphi^*)\,\partial_{x_d}u\,dx - \int_{\mathbb{R}^d_+} u\,(\varphi-\varphi^*)\partial_{x_d}\chi_n\,dx.\end{aligned}$$

Now, we have $|\varphi(x,x_d)-\varphi(x',-x_d)| \leq 2\,\|\partial_{x_d}\varphi\|_{\mathrm{L}^\infty(\mathbb{R})}\,|x_d| = C|x_d|$, so that

$$\left|\int_{\mathbb{R}^d_+} u\,(\varphi-\varphi^*)\partial_{x_d}\chi_n\,dx\right| \leq \int_{\left\{0<x_d<\frac{1}{2n}\right\}} C|u||x_d||\theta_n'(x_d)|dx \leq 2C\int_{\left\{0<x_d<\frac{1}{2n}\right\}}|u|dx \underset{n\to\infty}{\longrightarrow} 0,$$

where we used $|\theta_n'| \leq 4n$. Therefore, we have

$$\partial_{x_d}u_* = \begin{cases}\partial_{x_d}u(x',x_d) & \text{for all } x_d > 0\\ -\,\partial_{x_d}u(x',-d) & \text{for all } x_d \leq 0,\end{cases}$$

which concludes the proof of the lemma. □

Using local charts, the obtention of a continuous extension operator is rather straightforward.

Theorem 3.3.3. *Let Ω be a C^m domain of $\mathbb{R}^d$ with a bounded boundary. Then, there exists a continuous* extension operator $T : W^{m,p}(\Omega) \to W^{m,p}(\mathbb{R}^d)$ *such that for all $u \in W^{m,p}(\Omega)$,*

1. $Tu|\Omega = u$,

2. $\|Tu\|_{\mathrm{L}^p(\mathbb{R}^d)} \leq C\,\|u\|_{\mathrm{L}^p(\Omega)}$

3. $\|Tu\|_{\mathrm{W}^p(\mathbb{R}^d)} \leq C\,\|u\|_{\mathrm{W}^p(\Omega)}$.

where the constant C depends only on Ω.

Proof. As previously, by induction, we need only treat the case $m = 1$. Since $\partial\Omega$ is compact, there exists finitely many points $x_1, \cdots, x_n \in \partial\Omega$ such that

$$\partial\Omega \subset \bigcup_{i=1}^{n} B(x_i, r_i),$$

and C^1 diffeomorphisms $\varphi_i : \mathbb{R}^d \to B(x_i, r_i)$ such that $\varphi(\mathbb{R}^d_+) = \Omega \cap B(x_i, r_i)$ and $\varphi_i(\partial\mathbb{R}^d_+) = \partial\Omega \cap B(x, r)$ $(1 \leq i \leq m)$. Then, if $\chi_0, \chi_1, \cdots, \chi_n$ is a partition of unity associated to $\mathbb{R}^d \setminus \partial\Omega, B(x_1, r_1), ..., B(x_n, r_n)$, we write

$$u = \sum_{i=0}^{n} \chi_i u = \sum_{i=0}^{n} u_i.$$

We extend u_0 by 0 and write $u_0^* = u_0\, \mathbf{1}_\Omega$. As $\mathrm{supp}(\chi_0) \subset \mathbb{R}^d \setminus \partial\Omega$, u_0 has compact support on Ω, and therefore extends to $\mathbb{R}^d$ as an element of $W^{1,p}(\mathbb{R}^d)$. Furthermore, we have

$$\partial_{x_i} u_0^* = \theta_0\, (\partial_{x_i} u_0)\, \mathbf{1}_\Omega + (\partial_{x_i}\theta_0) u\, \mathbf{1}_\Omega. \tag{3.3.1}$$

For all $1 \leq i \leq d$, we have $\nabla\theta_i \in L^\infty(\mathbb{R}^d)$ since $\theta_i \in \mathscr{D}(\mathbb{R}^d)$. Therefore, the identity

$$\sum_{i=0}^{n} \theta_i = 1$$

shows that

$$\nabla\theta_0 = -\sum_{i=1}^{n} \nabla\theta_i \in L^\infty(\mathbb{R}^d),$$

and by

$$\|\nabla u_0^*\|_{\mathrm{L}^p(\mathbb{R}^d)} \leq \|u_0\|_{\mathrm{L}^p(\Omega)} + \|\nabla\theta_0\|_{\mathrm{L}^\infty(\mathbb{R}^d)} \|\nabla u_0\|_{\mathrm{L}^p(\Omega)} \leq \left(1 + \|\nabla\theta_0\|_{\mathrm{L}^\infty(\mathbb{R}^d)}\right) \|u\|_{\mathrm{W}^{1,p}(\Omega)}. \tag{3.3.2}$$

Now, fix some $1 \leq i \leq n$, and let $v_i = u \circ \varphi|\mathbb{R}^d_+ : \mathbb{R}^d_+ \to \mathbb{C}$. An immediate argument by density shows that $v_i \in W^{1,p}(\mathbb{R}^d_+)$, and that

$$\nabla v_i = \nabla\varphi \cdot \nabla u_i \circ \varphi.$$

Extend v_i by reflection to a map $v_i^* : \mathbb{R}^d \to \mathbb{C}$ using Lemma 3.3.2, and let $w_i = v_i^* \circ \varphi^{-1} : B(x_i, r_i) \to \mathbb{C}$. Then, the previous comment on compositions shows that

$$\|w_i\|_{\mathrm{W}^{1,p}(B(x_i,r_i))} \leq C\, \|u\|_{\mathrm{W}^{1,p}(\Omega\cap B(x_i,r_i))}.$$

Finally, setting $u_i^* = \chi_i\, w_i\, 1_\Omega$ yields the following controlled extension:

$$Tu = \sum_{i=0}^{n} u_i^*.$$

This concludes the proof of the theorem. □

Remark 3.3.4. The theorem holds under milder hypotheses; a strong local Lipschitzian condition (see [1], Theorem 5.24) suffices by virtue of Stein Extension Theorem.

3.4 Sobolev Embedding Theorem

3.4.1 Super-Critical Case

As we mentioned previously, the Sobolev inequality shows that a distribution u such that $\nabla u \in L^p(\mathbb{R}^d)$ is in fact a locally L^q function for some exponent $q > 1$. Assuming that u belongs to some L^r space, we get a global estimate. In particular, the Sobolev inequality is particularly easy to state for $W^{1,p}$ functions. The argument generalises to $W^{m,p}$ spaces, and once more, we need only look at the case $m = 1$ to deduce more general Sobolev inequalities. The results depend on the relation between $1 \leq p \leq \infty$ and the ambient dimension d.

Theorem 3.4.1 (Sobolev)**.** *Assume that $d \geq 2$, and let $1 \leq p < d$. Then, we have a continuous embedding $W^{1,p}(\mathbb{R}^d) \hookrightarrow L^{p^*}(\mathbb{R}^d)$, where*

$$p^* = \frac{dp}{d-p}.$$

For $d = 1$, for all interval $I \subset \mathbb{R}$, we have a continuous embedding $W^{1,p}(I) \hookrightarrow C^0(I)$, and

$$\|u\|_{\mathrm{L}^\infty(I)} \leq C(I)\, \|u\|_{\mathrm{W}^{1,p}(I)}\,.$$

Proof. **Part 1.** Case $d \geq 2$. We first assume that $p = 1$. For all $u \in \mathscr{D}(\mathbb{R}^d)$, we have for all $1 \leq i \leq d$ by the triangle inequality

$$\begin{aligned} |u(x)| &= \left|\int_{-\infty}^{x_i} \partial_{x_i} u(x_1, \cdots, x_{i-1}, y, x_{i+1}, \cdots, x_d) dy\right| \\ &\leq \int_{\mathbb{R}} |\partial_{x_i} u(x_1, \cdots, x_{i-1}, y, x_{i+1}, \cdots, x_d)| dy = f(\widehat{x_i}), \end{aligned}$$

where $\widehat{x_i} = (x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_d) \in \mathbb{R}^{d-1}$. Now, we use the following straightforward lemma.

Lemma 3.4.2. *Let $f_1, \cdots, f_d \in L^{d-1}(\mathbb{R}^{d-1})$, and define*

$$f(x) = \prod_{i=1}^d f(\widehat{x_i}).$$

Then, $f \in L^1(\mathbb{R}^d)$, and

$$\|f\|_{\mathrm{L}^1(\mathbb{R}^d)} \leq \prod_{i=1}^d \|f_i\|_{\mathrm{L}^{d-1}(\mathbb{R}^{d-1})}\,. \tag{3.4.1}$$

Proof. The case $d = 2$ simply corresponds to Fubini's theorem:

$$\int_{\mathbb{R}^2} |f_1(x_1)||f_2(x_2)| dx_1 dx_2 = \|f_1\|_{\mathrm{L}^1(\mathbb{R})}\, \|f_2\|_{\mathrm{L}^1(\mathbb{R})}\,.$$

Now, assume that the results holds for $2 \leq i \leq d-1$. Fixing x_d, we have by Hölder's inequality

$$\int_{\mathbb{R}^{d-1}} |f(x', x_d)| dx' \leq \|f_d\|_{\mathrm{L}^{d-1}(\mathbb{R}^{d-1})} \left(\int_{\mathbb{R}^{d-1}} \prod_{i=1}^{d-1} |f_1(\widehat{x_i})|^{\frac{d-1}{d-2}} dx'\right)^{\frac{d-2}{d-1}}.$$

Applying the induction hypothesis to $|f_1|^{\frac{d-1}{d-2}}, \cdots, |f_{d-1}|^{\frac{d-1}{d-2}}$, we deduce that

$$\left(\int_{\mathbb{R}^d} \prod_{i=1}^{d-1} |f_1(\widehat{x_i})|^{\frac{d-1}{d-2}} dx'\right)^{\frac{d-2}{d-1}} \leq \prod_{i=1}^{d-1} \|f_i\|_{\mathrm{L}^{d-1}(\mathbb{R}^{d-2})} (x_d),$$

and

$$\int_{\mathbb{R}^{d-1}} |f(x', x_d)| dx' \leq \|f_d\|_{\mathrm{L}^{d-1}(\mathbb{R}^{d-1})} \prod_{i=1}^{d-1} \|f_i\|_{\mathrm{L}^{d-1}(\mathbb{R}^{d-2})} (x_d).$$

Now, integrating in x_d, the generalised Hölder's inequality shows that

$$\int_{\mathbb{R}} \prod_{i=1}^{d-1} \|f_i\|_{\mathrm{L}^{d-1}(\mathbb{R}^{d-2})} (x_d) dx_d \leq \prod_{i=1}^{d-1} \left\| \|f_i\|_{\mathrm{L}^{d-1}(\mathbb{R}^{d-2})} (x_d) \right\|_{\mathrm{L}^{d-1}(\mathbb{R})} = \prod_{i=1}^{d-1} \|f_i\|_{\mathrm{L}^{d-1}(\mathbb{R}^{d-1})}\,,$$

which concludes the proof of the lemma. ☐

Therefore, we have

$$|u(x)|^{\frac{d}{d-1}} \leq \prod_{i=1}^{d} f_i(\widehat{x_i})^{\frac{1}{d-1}},$$

which shows by the lemma that

$$\int_{\mathbb{R}^d} |u|^{\frac{d}{d-1}} dx \leq \prod_{i=1}^{d} \left\| f_i^{\frac{1}{d-1}} \right\|_{\mathrm{L}^{d-1}(\mathbb{R}^{d-1})} = \prod_{i=1}^{d} \|f_i\|_{\mathrm{L}^1(\mathbb{R}^{d-1})}^{\frac{1}{d-1}} \prod_{i=1}^{d} \|\partial_{x_i} u\|_{\mathrm{L}^1(\mathbb{R}^d)}^{\frac{1}{d-1}}. \tag{3.4.2}$$

We finally deduce that

$$\|u\|_{\mathrm{L}^{\frac{d}{d-1}}(\mathbb{R}^d)} \leq \prod_{i=1}^{d} \|\partial_{x_i} u\|_{\mathrm{L}^1(\mathbb{R}^d)}^{\frac{1}{N}} \leq \|\nabla u\|_{\mathrm{L}^1(\mathbb{R}^d)}.$$

For a general exponent $1 < p < d$, fix some $t > 1$ to be determined later, and apply (3.4.2) to $|u|^{t-1}u$. We get

$$\|u\|_{\mathrm{L}^{\frac{t\,d}{d-1}}(\mathbb{R}^d)}^{t} \leq t \prod_{i=1}^{d} \left\| |u|^{t-1} \partial_{x_i} u \right\|_{\mathrm{L}^1(\mathbb{R}^d)}^{\frac{1}{d}} \leq t\, \|u\|_{\mathrm{L}^{p'(t-1)}(\mathbb{R}^d)}^{t-1} \prod_{i=1}^{d} \|\partial_{x_i} u\|_{\mathrm{L}^p(\mathbb{R}^d)}^{\frac{1}{d}}. \tag{3.4.3}$$

Choosing $t > 1$ such that $\dfrac{t\,d}{d-1} = p'(t-1)$, we get

$$\|u\|_{\mathrm{L}^{p^*}(\mathbb{R}^d)} \leq \frac{p(d-1)}{d-p} \prod_{i=1}^{d} \|\partial_{x_i} u\|_{\mathrm{L}^p(\mathbb{R}^d)}^{\frac{1}{d}}.$$

Using the density of $\mathscr{D}(\mathbb{R}^d)$ in $W^{1,p}(\mathbb{R}^d)$ and Fatou lemma, we obtain the general proof.

Part 2. Now, we treat the case $d = 1$, and first establish an elementary lemma.

Lemma 3.4.3. *Let $g \in L^1_{\mathrm{loc}}(I)$, and fix some $x_0 \in I$. Define*

$$f(x) = \int_{x_0}^{x} g(y)dy.$$

Then, we have $f \in C^0(I)$, and $f' = g$ in $\mathscr{D}'(I)$.

Proof. The continuity follows from the classical theorems of continuous dependence of the Lebesgue integral (one can use the dominated convergence theorem for example). Now, for all $\varphi \in \mathscr{D}'(I)$, we have by Fubini's theorem

$$\begin{aligned}
&\int_I f(x)\,\varphi'(x)\,dz = \int_I \left(\int_{x_0}^{x} g(y)dy \right) \varphi'(x)dx \\
&= -\int_{\inf I}^{x_0} \int_I g(y)\varphi'(x)\mathbf{1}_{\{x \leq y \leq x_0\}} dxdy + \int_{x_0}^{\sup I} \int_I g(y)\varphi'(x)\mathbf{1}_{\{x_0 \leq y \leq x\}} dxdy \\
&= -\int_I \left(\int_{\inf I}^{x_0} \varphi'(x)\mathbf{1}_{\{x \leq y \leq x_0\}} dx \right) g(y)dy + \int_I \left(\int_{x_0}^{\sup I} \varphi'(x)\mathbf{1}_{\{x_0 \leq y \leq x\}} dx \right) g(y)dy \\
&= -\int_I \left(\int_{\inf I}^{y} \varphi'(x)dx \right) g(y)dy + \int_I \left(\int_{y}^{\sup I} \varphi'(x) \right) g(y)dy = -\int_I \varphi(y)g(y)dy,
\end{aligned}$$

where we used that φ has compact support in I. Therefore, we have in the distributional sense $f' = g$ in $\mathscr{D}'(I)$ as claimed. $\square$

Thanks to the lemma and Theorem 2.5.21, we deduce that for all $u \in W^{1,p}(I)$ and for all $x_0 \in I$, we have

$$u(x) - u(x_0) = \int_{x_0}^{x} u'(y)dy.$$

Provided that $I = \mathbb{R}$ and $u \in \mathscr{D}(\mathbb{R})$, we obtain similarly the formula

$$u(x)|u(x)|^{p-1} = \int_{-\infty}^{x} pu'(x)|u(x)|^{p-1}dx,$$

so that by Hölder's inequality

$$|u(x)|^p \leq p \left\|u'\right\|_{\mathrm{L}^p(\mathbb{R})} \left\|u\right\|_{\mathrm{L}^p(\mathbb{R})}^{p-1},$$

so that

$$\left\|u\right\|_{\mathrm{L}^\infty(\mathbb{R})} \leq p^{\frac{1}{p}} \left\|u\right\|_{\mathrm{W}^{1,p}(\mathbb{R})}.$$

The general result follows by density of $\mathscr{D}(\mathbb{R})$ in $W^{1,p}(\mathbb{R})$, and generalises to an arbitrary interval I thanks to the Extension Theorem 3.3.3. ☐

Recall the following elementary interpolation result.

Lemma 3.4.4. *Let (X, μ) be a measured space, $1 \leq p < q \leq \infty$, and $u \in L^p \cap L^q(X, \mu)$. Then, $u \in L^r(X, \mu)$ for all $p \leq r \leq q$, and we have*

$$\left\|u\right\|_{\mathrm{L}^r(X)} \leq \left\|u\right\|_{\mathrm{L}^p(X)}^{\alpha} \left\|u\right\|_{\mathrm{L}^q(X)}^{1-\alpha}, \tag{3.4.4}$$

where $\alpha \in [0, 1]$ is such that

$$\frac{1}{r} = \frac{\alpha}{p} + \frac{1-\alpha}{q} \tag{3.4.5}$$

Proof. Let $p < r < q$ and $0 < \alpha < 1$ such that

$$r = \alpha p + (1-\alpha)q.$$

By the Hölder's inequality, for all $1 < s < \infty$, we have

$$\int_X |u|^r d\mu = \int_X |u|^{\alpha p}|u|^{(1-\alpha)q} d\mu \leq \left(\int_X |u|^{\alpha ps} d\mu\right)^{\frac{1}{s}} \left(\int_X |u|^{(1-\alpha)qs'} d\mu\right)^{\frac{1}{s'}}.$$

We choose s such that

$$\begin{cases} \alpha ps = p \\ (1-\alpha)qs' = q \end{cases}$$

which leads to

$$\alpha = \frac{p(q-r)}{r(q-p)}$$

or

$$\alpha = \frac{\frac{1}{r} - \frac{1}{q}}{\frac{1}{p} - \frac{1}{q}}.$$

Since this last expression is equivalent to (3.4.5), we are done. ☐

Corollary 3.4.5. *Let $1 \leq p < d$, and $u \in W^{1,p}(\mathbb{R}^d)$. Then, for all $p \leq q \leq p^*$, we have a continuous injection $W^{1,p}(\mathbb{R}^d) \longhookrightarrow L^q(\mathbb{R}^d)$ and there exists a universal constant $C = C(p) < \infty$ such that*

$$\left\|u\right\|_{\mathrm{L}^q(\mathbb{R}^d)} \leq C \left\|\nabla u\right\|_{\mathrm{W}^{1,p}(\mathbb{R}^d)}. \tag{3.4.6}$$

3.4.2 Critical Case

Theorem 3.4.6. *We have a continuous embedding*

$$W^{1,d}(\mathbb{R}^d) \longrightarrow L^p(\mathbb{R}^d) \quad \textit{for all } d \le p < \infty.$$

Proof. Apply the inequality (3.4.3) with $p = d$ to get

$$\|u\|^t_{\mathrm{L}^{\frac{t\,d}{d-1}}(\mathbb{R}^d)} \le t\,\|u\|^{t-1}_{\mathrm{L}^{\frac{d(t-1)}{d-1}}(\mathbb{R}^d)}\,\|\nabla u\|_{\mathrm{L}^d(\mathbb{R}^d)}\,. \tag{3.4.7}$$

Choosing $t = d$, we get

$$\|u\|^d_{\mathrm{L}^{\frac{d^2}{d-1}}(\mathbb{R}^d)} \le d\,\|u\|^{d-1}_{\mathrm{L}^d(\mathbb{R}^d)}\,\|\nabla u\|_{\mathrm{L}^d(\mathbb{R}^d)}\,.$$

By interpolation, we deduce that $u \in L^q$ for all $d \le q \le \frac{d^2}{d-1}$. Applying (3.4.7) with $t = d+1, d+2$, etc, we deduce the statement of the theorem. □

Remark 3.4.7. By the Poincaré-Wirtinger inequality (see Theorem 3.5.4 below), we have a more precise result $W^{1,d}(\mathbb{R}^d) \hookrightarrow \mathrm{BMO}(\mathbb{R}^d)$, where BMO is the space of bounded-mean oscillation functions. We say that $f \in \mathrm{BMO}(\mathbb{R}^d)$ if

$$\|f\|_{\mathrm{BMO}} = \sup_{r>0} \fint_{B(x,r)} |f - \overline{f}_{B(x,r)}|d\mathscr{L}^d < \infty,$$

where

$$\overline{f}_{B(x,r)} = \fint_{B(x,r)} f\,d\mathscr{L}^d = \frac{1}{\alpha(d)r^d}\int_{B(x,r)} f\,d\mathscr{L}^d.$$

This Banach space, that contains L^∞ but is strictly greater that this space ($\log|x| \in \mathrm{BMO} \setminus L^\infty$, has deep applications to partial differential equations, and the celebrated Stein-Fefferman result identifies it as the pre-dual of the Hardy space. One studies such spaces in lectures about harmonic analysis (see [37], [7]).

The Poincaré Wirtinger inequality, scaling considerations and the convexity of $x \mapsfrom |x|^d$ imply that

$$\left(\fint_{B(x,r)} |u - \overline{u}_{B(x,r)}|d\mathscr{L}^d\right)^d \le \fint_{B(x,r)} |u - \overline{u}_{B(x,r)}|^d d\mathscr{L}^d \le C\int_{B(x,r)} |\nabla u|^d d\mathscr{L}^d,$$

where $C < \infty$ is independent of $x \in \mathbb{R}^d$ and $r > 0$.

3.4.3 Sub-Critical Case

Theorem 3.4.8. *Assume that $p > d$. Then, $W^{1,p}(\mathbb{R}^d) \hookrightarrow C^{0,\alpha} \cap L^\infty(\mathbb{R}^d)$, where $\alpha = 1 - \frac{d}{p} \in (0,1)$. Furthermore, there exists $C < \infty$ such that*

$$\|u\|_{\mathrm{L}^\infty(\mathbb{R}^d)} \le C\,\|u\|_{\mathrm{W}^{1,p}(\mathbb{R}^d)} \tag{3.4.8}$$

and

$$|u(x) - u(y)| \le C\,\|\nabla u\|_{\mathrm{L}^p(\mathbb{R}^d)}\,|x-y| \qquad \textit{for a.e. } x, y \in \mathbb{R}^d. \tag{3.4.9}$$

Remark 3.4.9. In other words, u admits a continuous and even Hölder-continuous representative. Beware that when one deals with Sobolev functions, he does not always have the latitude to replace u by its continuous representative. Indeed, generalised versions of the coarea formula ([15] **3.2.12**) that holds for $W^{1,p}(\mathbb{R}^d)$ functions ($1 \le p \le \infty$) are only valid for a specific choice of representative. One can find an example a $W^{1,p}$-homeomorphism (for all $1 \le p < d$) for which the corea formula is false. In other words, the theorem would only hold for a *discontinuous* representative. One can find similar counter-examples for $W^{1,d}$ maps that are *not* homeomorpism. However, it cannot happen for $p > d$ thanks to the Lusin property—a map has the Lusin property if it maps negligible sets to negligible sets. The restriction on the exponent $p < d$ in the first counter-example is optimal for $W^{1,d}$ homeomorphism satisfy the Lusin property (see [21] for more details).

Proof. Let $u \in \mathscr{D}(\mathbb{R}^d)$. Then, for all $x \in \mathbb{R}^d$, we have

$$u(x) - u(0) = \int_0^1 \frac{d}{dt} u(tx)dt = \sum_{i=1}^d \int_0^1 x_i\, \partial_{x_i} u(tx)dt.$$

If we fix $r > 0$, and assume that $x \in B(0,r)$, we get by the triangle inequality

$$|u(x) - u(0)| \leq r \sum_{i=1}^d \int_0^1 |\partial_{x_i} u(tx)|dt.$$

Integrating on $B(x,r)$, we get

$$|\overline{u}_{B(0,r)} - u(0)| \leq \frac{1}{\alpha(d) r^{d-1}} \sum_{i=1}^d \int_{B(0,r)} \int_0^1 |\partial_{x_i} u(tx)|dt. \tag{3.4.10}$$

Now, we have for all $0 < t < 1$

$$\begin{aligned}
\int_{B(0,r)} |\partial_{x_i} u(tx)|dx \underset{y=tx}{=} \frac{1}{t^d} \int_{B(0,tr)} |\partial_{x_i} u(y)|dy \leq \frac{1}{t^d} \left(\int_{B(x,tr)} |\nabla u|^p d\mathscr{L}^d \right)^{\frac{1}{p}} \left(\alpha(d)(tr)^d \right)^{\frac{1}{p'}} \\
= \frac{\alpha(d)^{\frac{1}{p'}} r^{\frac{d}{p'}}}{t^{\frac{d}{p}}} \left\| \nabla u \right\|_{\mathrm{L}^p(B(0,r))}, \qquad (3.4.11)
\end{aligned}$$

and since $\dfrac{d}{p} < 1$, we get by Fubini's theorem, (3.4.10) and (3.4.11)

$$\begin{aligned}
|\overline{u}_{B(0,r)} - u(0)| &\leq \frac{d}{\alpha(d) r^{d-1}} \times \alpha(d)^{\frac{1}{p'}} r^{\frac{d}{p'}} \left\| \nabla u \right\|_{\mathrm{L}^p(B(0,r))} \int_0^1 \frac{dt}{t^{\frac{d}{p}}} \\
&= \frac{d}{1 - \frac{d}{p}} \frac{1}{\alpha(d)^{\frac{1}{p}}} r^{1-\frac{d}{p}} \left\| \nabla u \right\|_{\mathrm{L}^p(B(0,r))}.
\end{aligned}$$

By translation, we deduce that for all $x \in \mathbb{R}^d$ and $y \in B(x,r)$

$$|u(y) - \overline{u}_{B(x,r)}| \leq \frac{d}{1 - \frac{d}{p}} \frac{1}{\alpha(d)^{\frac{1}{p}}} r^{1-\frac{d}{p}} \left\| \nabla u \right\|_{\mathrm{L}^p(B(x,r))}.$$

Using Lebesgue differentiation theorem (that is trivial of smooth functions), we deduce that

$$|u(x) - u(y)| \leq \frac{d}{1 - \frac{d}{p}} \frac{1}{\alpha(d)^{\frac{1}{p}}} r^{1-\frac{d}{p}} \left\| \nabla u \right\|_{\mathrm{L}^p(B(x,r))},$$

and by triangle inequality, we get

$$|u(y) - u(z)| \leq \frac{2d}{1 - \frac{d}{p}} \frac{1}{\alpha(d)^{\frac{1}{p}}} r^{1-\frac{d}{p}} \left\| \nabla u \right\|_{\mathrm{L}^p(B(x,r))} \quad \text{for all } y, z \in B(x,r),$$

which shows the second inequality. Choosing $r = 2|x-y|$, we get for all $y \in \mathbb{R}^d$

$$|u(y)| \leq |\overline{u}_{B(x,2|y-x|)}| + \frac{2^{2-\frac{d}{p}} d}{1 - \frac{d}{p}} \frac{1}{\alpha(d)^{\frac{1}{p}}} |x-y|^{1-\frac{d}{p}} \left\| \nabla u \right\|_{\mathrm{L}^p(B(x,r))} \leq C \left\| u \right\|_{\mathrm{W}^{1,p}(\mathbb{R}^d)},$$

which shows the first inequality. This proves the theorem in the case $u \in \mathscr{D}(\mathbb{R}^d)$, and the general result follows by density. □

3.4.4 General Result for $W^{m,p}(\Omega)$

Theorem 3.4.10. *Let $m \in \mathbb{N}$ and $1 \leq p < \infty$. We have the following results:*

1. *If $\dfrac{1}{p} - \dfrac{m}{d} > 0$, then $W^{m,p}(\mathbb{R}^d) \hookrightarrow L^q(\mathbb{R}^d)$ for $q = \dfrac{d\,p}{d-p}$.*
2. *If $\dfrac{1}{p} - \dfrac{m}{d} = 0$, then $W^{m,p}(\mathbb{R}^d) \hookrightarrow L^q(\mathbb{R}^d)$ for all $p \leq q < \infty$.*
3. *If $\dfrac{1}{p} - \dfrac{m}{d} < 0$, we have $W^{m,p}(\mathbb{R}^d) \hookrightarrow L^\infty(\mathbb{R}^d)$. Furthermore, if $\alpha = \left(m - \frac{d}{p}\right) - \left[m - \frac{d}{p}\right] > 0$, and $k = \left[m - \frac{d}{p}\right]$, we have $u \in C^{k,\alpha}(\mathbb{R}^d)$, and for all $|\beta| \leq k$, we have*
$$|D^\beta u(x) - D^\beta u(y)| \leq C\,\|u\|_{\mathrm{W}^{m,p}(\mathbb{R}^d)}\,.$$

Proof. The proof is done by induction thanks to the previous embedding theorems, and we leave it to the reader. ☐

Corollary 3.4.11. *Let Ω be a bounded open subset of class C^m and assume that $\partial\Omega$ is bounded. Then, the following results hold:*

1. *If $\dfrac{1}{p} - \dfrac{m}{d} > 0$, then $W^{m,p}(\Omega) \hookrightarrow L^q(\Omega)$ for $q = \dfrac{d\,p}{d-p}$.*
2. *If $\dfrac{1}{p} - \dfrac{m}{d} = 0$, then $W^{m,p}(\Omega) \hookrightarrow L^q(\Omega)$ for all $p \leq q < \infty$.*
3. *If $\dfrac{1}{p} - \dfrac{m}{d} < 0$, we have $W^{m,p}(\Omega) \hookrightarrow L^\infty(\mathbb{R}^d)$. Furthermore, if $\alpha = \left(m - \frac{d}{p}\right) - \left[m - \frac{d}{p}\right] > 0$, and $k = \left[m - \frac{d}{p}\right]$, we have $u \in C^{k,\alpha}(\mathbb{R}^d)$, and for all $|\beta| \leq k$, and for a.e. $x, y \in \Omega$ such that $B(x, 2|x-y|) \cup B(y, 2|x-y|) \subset \Omega$ we have*
$$|D^\beta u(x) - D^\beta u(y)| \leq C\,\|u\|_{\mathrm{W}^{m,p}(\Omega)}\,|x-y|.$$

Remark 3.4.12. We recall that the hypothesis of C^m open subset could be weakened to Lipschitzian subset by virtue of Stein Extension Theorem.

Theorem 3.4.13 (Rellich-Kondrachov)**.** *Assume that $d \geq 2$, and that Ω is a bounded open subset of class C^1 of $\mathbb{R}^d$. Then, we have*

1. *If $p < d$, then we have a* compact embedding *$W^{1,p}(\Omega) \hookrightarrow L^\infty(\Omega)$ for all $1 \leq q < p^*$, where $p^* = \frac{d\,p}{d-p}$.*
2. *If $p = d$, then we have a* compact embedding *$W^{1,p}(\Omega) \hookrightarrow L^p(\Omega)$ for all $1 \leq p < \infty$.*
3. *If $p > d$, we have a compact embedding $W^{1,p}(\Omega) \hookrightarrow C^0(\overline{\Omega})$.*

For all $-\infty < a < b < \infty$, we have a compact embedding $W^{1,p}(]a,b[) \hookrightarrow C^0([a,b])$ for $1 < p \leq \infty$ and a compact embedding $W^{1,1}(]a,b[) \hookrightarrow L^q(]a,b[)$ for all $1 \leq q < \infty$.

Proof. Thanks to Ascoli's theorem, we need only treat the case $p < d$.

We apply the following compactness criterion in L^p (*[11]*, IV.25).

Theorem 3.4.14 (Riesz-Fréchet-Kolmogorov)**.** *Let Ω be an open subset of $\mathbb{R}^d$, and $U \subset \Omega$ be a relatively compact open subset. Let $\mathscr{F}$ be a bounded domain of $L^p(\Omega)$ with $1 \leq p < \infty$. Assume that*
$$\forall \varepsilon > 0,\ \exists\,\delta > 0 \ \text{such that}\ \|\tau_h f - f\|_{\mathrm{L}^p(U)} < \varepsilon \quad \forall h \in B(0,\delta)\ \text{and}\ \forall f \in \mathscr{F},$$
where $\tau_h f(x) = f(x+h)$. Then, $\mathscr{F}_{|\mathrm{U}}$ is relatively compact in $L^p(U)$.

Fix some relatively compact open subset $U \subset \Omega$, to be determined later, and let $\varepsilon > 0$. Using the interpolation inequality from Lemma (3.4.4), for all $1 \leq q < p^*$, there exists $0 \leq \alpha < 1$ such that

$$\begin{aligned}\|\tau_h u - u\|_{\mathrm{L}^p(U)} &\leq \|\tau_h u - u\|_{\mathrm{L}^1(U)}^{\alpha} \|\tau_h u - u\|_{\mathrm{L}^{p^*}(U)}^{1-\alpha} \leq |h|^{\alpha} \|\nabla u\|_{\mathrm{L}^1(U)}^{\alpha} \|\tau_h u - u\|_{\mathrm{L}^{p^*}(U)}^{1-\alpha} \\ &\leq 2^{1-\alpha} |h|^{\alpha} \|\nabla u\|_{\mathrm{L}^1(U)}^{\alpha} \|u\|_{\mathrm{L}^{p^*}(U)}^{1-\alpha} = C|h|^{\alpha} < \varepsilon\end{aligned}$$

provided that h is small enough. On the other hand, we have by Hölder's inequality

$$\|u\|_{\mathrm{L}^q(\Omega \setminus \overline{U})} \leq \|u\|_{\mathrm{L}^{p^*}(\Omega)} \left(\mathscr{L}^n(\Omega \setminus \overline{U})\right)^{1-\frac{q}{p^*}} < \varepsilon,$$

provided that $\mathscr{L}^n(\Omega \setminus \overline{U})$ is small enough.

We omit the proof of the case $d = 1$ which is very similar. □

3.5 The Space $W_0^{m,p}(\Omega)$

3.5.1 Definition and first properties

Definition 3.5.1. Let $1 \leq p < \infty$. We define $W_0^{m,p}(\Omega) = \overline{\mathscr{D}(\Omega)}^{W^{m,p}}$ the closure of the space of compactly supported smooth functions in Ω for the $W^{m,p}$ topology. For $p = 2$, we write $H_0^m(\Omega) = W_0^{m,2}(\Omega)$. The space $W_0^{1,p}(\Omega)$ is a separable Banach space, and a reflexive space for $1 < p < \infty$. $H_0^m(\Omega)$ is a Hilbert space for the standard scalar product associated to $H^m(\Omega)$.

$W_0^{m,p}$ functions are functions whose traces up to the derivatives of order $m-1$ vanish on the boundary. However, in order to make the idea of trace precise, one needs to introduce fractional Sobolev spaces, that will be mentioned later in the course. We will therefore only prove two classical inequalities of fundamental importance.

Furthermore, in order to solve boundary problems for partial differential equations, the notion of trace is not formally needed in simple cases. Indeed, if $g \in W^{1,p}(\Omega)$, then we define the space

$$W_g^{1,p}(\Omega) = W^{1,p}(\Omega) \cap \left\{u : u - g \in W_0^{1,p}(\Omega)\right\}$$

of Sobolev functions whose trace on the boundary is g. This is not completely satisfactory for it requires to be able to extend g, but we will treat below the easier case of traces in $H^s(\Omega)$ (where $s \in \mathbb{R}$).

3.5.2 Poincaré Inequalities

Theorem 3.5.2 (Poincaré Inequality)**.** *Let $1 \leq p < \infty$, Ω be a bounded subset. Then, there exists a universal constant $C_P < \infty$ such that*

$$\|u\|_{\mathrm{L}^p(\Omega)} \leq C_P \|\nabla u\|_{\mathrm{L}^p(\Omega)} \qquad \textit{for all } u \in W_0^{1,p}(\Omega). \tag{3.5.1}$$

Proof. Let $\varphi \in \mathscr{D}(\Omega)$, and $R > 0$ be such that $\Omega \subset \mathbb{R}^d \cap \{x : |x_d| \leq R\}$. Then, we have

$$\varphi(x', x_d) = \int_{-R}^{x_d} \partial_{x_d} \varphi(x', t) dt.$$

By Hölder's inequality, we deduce that

$$|\varphi(x', x_d)|^p \leq (2R)^{p-1} \int_{-R}^{R} |\partial_{x_d} \varphi(x', t)|^p dt.$$

Therefore, Fubini's theorem implies that

$$\int_{\Omega} |\varphi(x)|^p dx \leq (2R)^p \int_{\Omega} |\partial_{x_d} \varphi|^p dx,$$

which yields the announced inequality by density of $\mathscr{D}(\Omega)$ in $W_0^{1,p}(\Omega)$. □

Remark 3.5.3. The proof shows that the statement is true for a set that is bounded in a single direction.

Theorem 3.5.4 (Poincaré-Wirtinger Inequality)**.** *Let $1 \leq p < \infty$, and Ω be a bounded domain of $\mathbb{R}^d$. Then, there exists a universal constant $C_{PW} < \infty$ such all $u \in W^{1,p}(\Omega)$, we have*

$$\int_\Omega |u - u_\Omega|^p \leq C_{PW} \int_\Omega |\nabla u|^p dx, \tag{3.5.2}$$

where

$$u_\Omega = \fint_\Omega u \, d\mathscr{L}^d = \frac{1}{\mathscr{L}^d(\Omega)} \int_\Omega u \, d\mathscr{L}^d$$

is the mean of u on Ω.

Proof. We argue by contradiction, and let $\{u_n\}_{n\in\mathbb{N}^*} \subset W^{1,p}(\Omega)$ such that

$$\|u_n - u_{n\Omega}\|_{\mathrm{L}^p(\Omega)} = 1$$
$$\|\nabla u_n\|_{\mathrm{L}^p(\Omega)} \leq \frac{1}{n}.$$

Let $v_n = u_n - u_{n\Omega}$. Then, $\{v_n\}_{n\in\mathbb{N}}$ is bounded in $W^{1,p}(\Omega)$, which implies by the Rellich-Kondrachov Theorem 3.4.13 that up to a subsequence, we have $v_n \underset{n\to\infty}{\longrightarrow} v \in L^p(\Omega)$ strongly, which implies in particular that $\|v\|_{\mathrm{L}^p(\Omega)} = 1$, and $v_\Omega = 0$. However, we also have by Fatou lemma

$$\|\nabla v\|_{\mathrm{L}^p(\Omega)} \leq \liminf_{n\to\infty} \|\nabla v_n\|_{\mathrm{L}^p(\Omega)} = 0.$$

Therefore, v is constant, but the condition $v_\Omega = 0$ implies that $v = 0$, contradiction. □

3.6 The Dual Spaces $W^{-m,p'}(\Omega)$

Definition 3.6.1. For all $1 \leq p < \infty$ and $m \in \mathbb{N}$, we denote by $W^{-m,p'}(\Omega)$ the dual space of $W_0^{m,p}(\Omega)$.

Theorem 3.6.2. *For all $F \in W^{-m,p'}(\Omega)$, there exists $f_\alpha \in L^{p'}(\Omega)$ $(\alpha \in \mathbb{N}^d)$ such that*

$$\langle F, u\rangle = \sum_{|\alpha|\leq m} \int_\Omega f_\alpha \, D^\alpha u \, d\mathscr{L}^d \quad \textit{for all } u \in W_0^{m,p}(\Omega). \tag{3.6.1}$$

Proof. The proof is, *mutadis mutandis*, the same as the one of Theorem 2.5.32, provided that one replaces compactly supported continuous functions by L^p functions. Indeed, let $N(d,m) = \mathrm{card}(\mathbb{N}^d \cap \{\alpha : |\alpha| \leq m\}) \in \mathbb{N}$, and let $T : W_0^{m,p}(\Omega) \to L^p(\Omega)^{N(d,m)}, u \mapsto (\{D^\alpha u\}_{|\alpha|\leq m})$. Then, T is an isometry of $W_0^{m,p}(\Omega)$ into $L^p(\Omega)^{N(d,m)}$. Let $T^{-1} : T(W_0^{m,p}(\Omega)) \to W_0^{m,p}(\Omega)$, and

$$L : T(W_0^{m,p}(\Omega)) \to \mathbb{R}$$
$$\{g_\alpha\}_{|\alpha|\leq m} \mapsto \left\langle F, T^{-1}(\{g_\alpha\}_{|\alpha|\leq m}\right\rangle.$$

By Hahn-Banach theorem, this continuous linear form admits a continuous linear extension to $L^p(\Omega)^{N(d,m)}$, that preserves the norm, that we denote by $\Phi : L^p(\Omega)^{N(d,m)+1} \to W_0^{m,p}(\Omega)$. According to the Riesz representation theorem, there exists $f_\alpha \in L^{p'}(\Omega)$ such that

$$\left\langle \Phi, (\{g_\alpha\}_{\alpha\in\mathbb{N}^d})\right\rangle = \sum_{|\alpha|\leq m} \int_\Omega f_\alpha \, g_\alpha \, d\mathscr{L}^d \quad \text{for all } g \in L^p(\Omega)^{N(d,m)}.$$

Therefore, we have

$$\|\Phi\|_{(L^p(\Omega)^{N(d,m)})'} = \max_{\substack{\alpha\in\mathbb{N}^d \\ |\alpha|\leq m}} \left\{\|f_\alpha\|_{\mathrm{L}^{p'}(\Omega)}\right\},$$

which concludes the proof of the theorem. □

Remark 3.6.3. In general, the functions f_α are not unique. Notice that our previous theorem on $\mathscr{D}_{\mathrm{L}^p}(\mathbb{R}^d)$ is proven.

3.7 The Hilbert Spaces $H^s(\mathbb{R}^d)$

3.7.1 Basic Properties

Those spaces will be the first examples of interpolation spaces, and they are easy to define.

Definition 3.7.1. For all $s \in \mathbb{R}$ define

$$H^s(\mathbb{R}^d) = \mathscr{S}'(\mathbb{R}^d) \cap \left\{u : (1+|\xi|^2)^{\frac{s}{2}}\mathscr{F}(u) \in L^2(\mathbb{R}^d)\right\},$$

and equip it with the following norm:

$$\|u\|_{\mathrm{H}^s(\mathbb{R}^d)} = \left(\int_{\mathbb{R}^d}(1+|\xi|^2)^s|\widehat{u}(\xi)|^2d\xi\right)^{\frac{1}{2}}. \tag{3.7.1}$$

Remark 3.7.2. In the case of S^1, the space $H^s(S^1)$ is defined as follows:

$$H^s(S^1) = \mathscr{D}'(S^1) \cap \left\{u : (1+|n|)^{\frac{s}{2}}\widehat{u} \in l^2(\mathbb{Z})\right\},$$

where for all $n \in \mathbb{Z}$, we have

$$\widehat{u}(n) = \langle u, e^{-in\theta}\rangle.$$

We equip $H^s(S^1)$ with the following norm:

$$\|u\|_{\mathrm{H}^s(\mathbb{Z})} = \left(\sum_{n\in\mathbb{Z}}(1+|n|^2)^s|\widehat{u}(n)|^2\right)^{\frac{1}{2}}. \tag{3.7.2}$$

Theorem 3.7.3. *For all $s \in \mathbb{R}$, $H^s(\mathbb{R}^d)$ is a separable Hilbert space, and for $m \in \mathbb{Z}$, $H^m(\mathbb{R}^d) = W^{m,2}(\mathbb{R}^d)$ with equivalent norms.*

Proof. The following quantity

$$\langle u, v\rangle_s = \int_{\mathbb{R}^d}(1+|\xi|^2)^s\widehat{u}(\xi)\overline{\widehat{v}(\xi)}d\xi \tag{3.7.3}$$

is a scalar product on H^s, and the map $u \mapsto (1+|\xi|^2)^{\frac{s}{2}}\widehat{u}$ is an isometric bijection between H^s and L^2. Since $L^2(\mathbb{R}^d)$ is complete, we deduce that H^s is complete for the norm above. We need only treat the second part in the case $m \geq 0$. By the properties of the Fourier transform, for all $u \in \mathscr{S}'(\mathbb{R}^d)$, we have $\mathscr{F}(D^\alpha u) = i^{|\alpha|}\xi^\alpha\widehat{u}$, which shows by Parseval identity that

$$\|D^\alpha u\|_{\mathrm{L}^2(\mathbb{R}^d)} = \frac{1}{(2\pi)^{\frac{d}{2}}}\left(\int_{\mathbb{R}^d}|\xi^\alpha|^2|\widehat{u}(\xi)|^2d\xi\right)^{\frac{1}{2}}. \tag{3.7.4}$$

Notice that here exists constants $0 < C_m < \infty$ such that

$$C_m^{-1}(1+|\xi|^2)^m \leq \sum_{|\alpha|\leq m}|\xi^\alpha|^2 \leq C_m(1+|\xi|^2)^2. \tag{3.7.5}$$

Indeed, for all $|\alpha| \leq m$, we have

$$|\xi^\alpha|^2 \leq |\xi|^{2|\alpha|} \leq (1+|\xi|^2)^m,$$

while

$$\sum_{|\alpha|\leq m}|\xi^\alpha|^2 \geq 1 + \sum_{j=1}^m|\xi_j^m|^2 \geq C_1(1+|\xi|^{2m}) \geq C_2(1+|\xi|^2)^m$$

thanks to the binomial formula. Finally, we deduce by (3.7.4) and (3.7.5) that both $H^m(\mathbb{R}^d)$ norms are equivalent. ☐

Theorem 3.7.4. *$\mathscr{D}(\mathbb{R}^d)$ is dense in $H^s(\mathbb{R}^d)$.*

Proof. We first show that $\mathscr{S}(\mathbb{R}^d)$ is dense in $H^s(\mathbb{R}^d)$. Indeed, the inverse isometry of $u \mapsto (1+|\xi|^2)^{\frac{s}{2}}\widehat{u}$ is the map $v \mapsto \mathscr{F}^{-1}((1+|\xi|^2)^{-s}) * \mathscr{F}^{-1}(v) = (2\pi)^{-d}\mathscr{F}((1+|\xi|^2)^{-s}) * \widehat{v}(-\xi)$, which sends $\mathscr{S}(\mathbb{R}^d)$ into $\mathscr{S}(\mathbb{R}^d)$ since $(1+|\xi|^2)^{-s} \in \mathscr{O}_M(\mathbb{R}^d)$ has polynomial growth. By density of $\mathscr{S}(\mathbb{R}^d)$ in $L^2(\mathbb{R}^d)$, we deduce that $\mathscr{S}(\mathbb{R}^d)$ is dense in $H^s(\mathbb{R}^d)$.

Now, notice that for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$, we have

$$\begin{aligned}
\|\varphi\|_{\mathrm{H}^s(\mathbb{R}^d)} &= \left(\int_{\mathbb{R}^d}(1+|\xi|^2)^s|\widehat{\varphi}(\xi)|^2 d\xi\right)^{\frac{1}{2}} \leq \left\|(1+|\xi|^2)^{s+d}\widehat{\varphi}\right\|_{\mathrm{L}^\infty(\mathbb{R}^d)}\int_{\mathbb{R}^d}\frac{d\xi}{(1+|\xi|^2)^d}\\
&= \sqrt{\frac{\beta(d)}{2d}}\left\|(1+|\xi|^2)^{s+d}\widehat{\varphi}\right\|_{\mathrm{L}^\infty(\mathbb{R}^d)} \leq \sqrt{\frac{\beta(d)}{2d}}2^{|s|+d-1}\left(\|\widehat{\varphi}\|_{\mathrm{L}^\infty(\mathbb{R}^d)} + \left\||\xi|^{2(|s|+d)}\widehat{\varphi}\right\|_{\mathrm{L}^\infty(\mathbb{R}^d)}\right)\\
&\leq \sqrt{\frac{\beta(d)}{2d}}2^{|s|+d-1}\left(\|\widehat{\varphi}\|_{\mathrm{L}^\infty(\mathbb{R}^d)} + (|s|+1+d)^{|s|+d}\sum_{i=1}^{d}\left\||\xi_j|^{2([s]+1+d)}\widehat{\varphi}\right\|_{\mathrm{L}^\infty(\mathbb{R}^d)}\right)\\
&= \sqrt{\frac{\beta(d)}{2d}}2^{|s|+d-1}\left(\|\widehat{\varphi}\|_{0,0} + (|s|+1+d)^{|s|+d}\sum_{j=1}^{d}\|\widehat{\varphi}\|_{2([s]+1+d)e_j,0}\right),
\end{aligned}$$

where we recall that the norms $\|\cdot\|_{\alpha,\beta}$ $(\alpha,\beta \in \mathbb{N}^d)$ are defined in (2.7.6), and $(e_1,\cdots,e_d)$ is the canonical basis of $\mathbb{R}^d$. Using the inequality (2.7.7), we deduce by density of $\mathscr{D}(\mathbb{R}^d)$ ib $\mathscr{S}(\mathbb{R}^d)$ that $\mathscr{D}(\mathbb{R}^d)$ is dense in $H^s(\mathbb{R}^d)$. □

3.7.2 Duality

Theorem 3.7.5 (Duality)**.** *For all $s \in \mathbb{R}$, for all $L \in (H^s(\mathbb{R}^d))'$, there exists a unique $v \in H^{-s}(\mathbb{R}^d)$ such that*

$$L(u) = \langle u, v\rangle = \int_{\mathbb{R}^d} u(x)\, v(x)\, dx \quad \text{for all } u \in H^s(\mathbb{R}^d).$$

Proof. First, for all $(u,v) \in H^s(\mathbb{R}^d) \times H^{-s}(\mathbb{R}^d)$, we have by Parseval identity:

$$\langle u, v\rangle = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^d}\widehat{u}(\xi)\widehat{v}(-\xi)d\xi.$$

In particular, we deduce that

$$|\langle u, v\rangle| = \frac{1}{(2\pi)^n}\left|\int_{\mathbb{R}^d}(1+|\xi|^2)^{\frac{s}{2}}\widehat{u}(\xi)\,(1+|\xi|^2)^{-\frac{s}{2}}\widehat{v}(\xi)\,d\xi\right| \leq \frac{1}{(2\pi)^n}\|u\|_{\mathrm{H}^s(\mathbb{R}^s)}\|v\|_{\mathrm{H}^{-s}(\mathbb{R}^d)},$$

which shows that the map $L_v : u \mapsto \langle u, v\rangle$ is a continuous linear form on $H^s(\mathbb{R}^d)$, *i.e.* an element of $(H^s(\mathbb{R}^d))'$, and that furthermore, we have

$$\|L_v\|_{(H^s(\mathbb{R}^d))'} \leq \frac{1}{(2\pi)^n}\|v\|_{\mathrm{H}^{-s}(\mathbb{R}^d)}. \tag{3.7.6}$$

In fact, we have the equality in (3.7.6), as we will see it below.

Conversely, let $L \in (H^s(\mathbb{R}^d))'$, and consider the isometric map $T : H^s(\mathbb{R}^d) \to L^2(\mathbb{R}^d), u \mapsto (1+|\xi|^2)^{\frac{s}{2}}\widehat{u}$. Then, T is an linear continuous isometry. Consider the map $T^{-1} : T(H^s(\mathbb{R}^d)) \to H^s(\mathbb{R}^d)$, and as previously, the map

$$\begin{aligned}
F : T(H^s(\mathbb{R}^d)) &\to \mathbb{R}\\
v &\mapsto \left\langle L, T^{-1}(v)\right\rangle.
\end{aligned}$$

By Hahn-Banach theorem, this linear map extends to a continuous linear map $\Phi : L^2(\mathbb{R}^d) \to \mathbb{R}$. Thanks to Riesz-Fréchet representation theorem (on the dual space of a Hilbert space), there exists $f \in L^2(\mathbb{R}^d)$ such that

$$\Phi(u) = \int_{\mathbb{R}^d} f(x)\, v(x)\, dx \quad \text{for all } v \in L^2(\mathbb{R}^d).$$

If $v = (1+|x|^2)^{\frac{s}{2}}\widehat{u}$ for some $u \in H^s(\mathbb{R}^d)$, the Parseval identity implies that

$$\begin{aligned} L(u) = \Phi(u) &= \int_{\mathbb{R}^d} (1+|x|^2)^{\frac{s}{2}} f(x)\widehat{u}(x)dx = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \mathscr{F}((1+|x|^2)^{\frac{s}{2}} f)(\xi)\mathscr{F}(\widehat{u})(-\xi)d\xi \\ &= \int_{\mathbb{R}^d} h(\xi)u(\xi)d\xi, \end{aligned}$$

where $h = \mathscr{F}((1+|x|^2)^{\frac{s}{2}} f) \in H^{-s}(\mathbb{R}^d)$ by definition, since

$$\widehat{h}(\xi) = \mathscr{F}^2((1+|x|^2)^{\frac{s}{2}} f) = (2\pi)^d(1+|\xi|^2)^{\frac{s}{2}} f(-\xi),$$

and $f \in L^2(\mathbb{R}^d)$. This concludes the proof of the theorem. □

3.7.3 Traces

Theorem 3.7.6. *For all $s > \frac{1}{2}$, the operator $\gamma : \mathscr{S}(\mathbb{R}^d) \to \mathscr{S}(\mathbb{R}^{d-1})$, such that*

$$\gamma(\varphi)(x') = \varphi(x', 0),$$

admits a unique continuous linear extension $H^s(\mathbb{R}^d) \to H^{s-\frac{1}{2}}(\mathbb{R}^d)$.

Proof. Fix some $\varphi \in \mathscr{S}(\mathbb{R}^d)$. Thanks to the Fourier inversion formula and Fubini's theorem, we have for all $x' \in \mathbb{R}^{d-1}$

$$\begin{aligned} \psi(x') = \varphi(x', 0) &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^{d-1}\times\mathbb{R}} \widehat{\varphi}(\xi', \xi_d) e^{i\, x'\cdot\xi'} d\xi' d\xi_d \\ &= \frac{1}{(2\pi)^{d-1}}\int_{\mathbb{R}^{d-1}} \left(\int_{\mathbb{R}} \widehat{\varphi}(\xi', \xi_d)d\xi_d\right) e^{i\, x'\cdot\xi'} d\xi'. \end{aligned}$$

Using once more the inverse Fourier formula, we deduce that

$$\widehat{\psi}(\xi') = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{\varphi}(\xi', t)dt = \frac{1}{2\pi}\int_{\mathbb{R}} \left(\widehat{\varphi}(\xi', t)(1+|\xi'|^2+t^2)^{\frac{s}{2}}\right)(1+|\xi'|^2+t^2)^{-\frac{s}{2}} dt.$$

Since $s > \frac{1}{2}$, we have

$$\int_{\mathbb{R}} \frac{dt}{(1+|\xi'|^2+t^2)^s} = \frac{1}{(1+|\xi'|^2)^{s-\frac{1}{2}}}\int_{\mathbb{R}} \frac{dt}{(1+t^2)^s} = \frac{c_s}{(1+|\xi'|^2)^{s-\frac{1}{2}}} < \infty,$$

which implies by Cauchy-Schwarz inequality that

$$|\widehat{\psi}(\xi')|^2 \leq \frac{c_s}{(2\pi)^2}\frac{1}{(1+|\xi'|^2)^{s-\frac{1}{2}}}\int_{\mathbb{R}} (1+|\xi'|^2+t^2)^s|\widehat{\varphi}(\xi', t)|^2 dt.$$

Another application of Fubini's theorem shows that

$$\begin{aligned} \|\gamma(\varphi)\|^2_{\mathrm{H}^{s-\frac{1}{2}}(\mathbb{R}^{d-1})} &= \int_{\mathbb{R}^{d-1}} (1+|\xi'|^2)^{s-\frac{1}{2}}|\widehat{\psi}(\xi')|^2 d\xi' \\ &\leq \frac{c_s}{(2\pi)^2}\int_{\mathbb{R}^d} (1+|\xi|^2)^s|\widehat{\varphi}(\xi)|^2 d\xi = \frac{c_s}{(2\pi)^2}\|u\|^2_{\mathrm{H}^s(\mathbb{R}^d)} \end{aligned}$$

By density of $\mathscr{S}(\mathbb{R}^d)$ in $H^s(\mathbb{R}^d)$, we deduce that $\gamma : H^s(\mathbb{R}^d) \to H^{s-\frac{1}{2}}(\mathbb{R}^{d-1})$ is a continuous linear map such that

$$\|\gamma\| \leq \frac{1}{2\pi}\left(\int_{\mathbb{R}} \frac{1}{(1+t^2)^s}\right)^{\frac{1}{2}} = \frac{1}{\pi}\sqrt{\frac{s}{2s-1}},$$

which concludes the proof of the theorem. □

Remark 3.7.7. In particular, if we have a continuous trace operator $H^1(B_{\mathbb{R}^2}(0,1)) \to H^{\frac{1}{2}}(S^1)$ (where $H^s(S^1)$ is defined in (3.7.2)), and more generally, the trace theorem is true for a C^1 domain, but requires to define the fractional Sobolev space, which will be seen later on.

Those results have applications to the solvability of the Dirichlet problem for $H^{\frac{1}{2}}$ data, which is crucial in many applications (see [9], and the exercises).

3.8 Link between fractional derivatives and fractional Sobolev spaces

3.8.1 Fractional derivatives

For all locally integrable function $f : \mathbb{R} \to \mathbb{R}$, and for all $n \in \mathbb{N}$, one shows by induction that the n-th primitive of f that vanishes at 0 is given by

$$\mathcal{I}^n f(x) = \frac{1}{\Gamma(n)} \int_0^x (x-t)^{n-1} f(t) dt, \tag{3.8.1}$$

where Γ is the Euler Γ function, such that for all $z \in \mathbb{C} \cap \{z : \mathrm{Re}\,(z) > 0\}$, we have

$$\Gamma(z) = \int_0^\infty t^z e^{-t} \frac{dt}{t}. \tag{3.8.2}$$

This is the Mellin transform of the function $t \mapsto e^{-t}$. By analogy, we define for all $\alpha > 0$ the fractional integral of order $\alpha > 0$ by

$$\mathcal{I}^\alpha f(x) = \frac{1}{\Gamma(\alpha)} \int_0^x (x-t)^{\alpha-1} f(t) dt, \tag{3.8.3}$$

well-defined for $x > 0$, and that one can eventually extend by analytic continuation to all values of x outside a half-line of the complex plane. In particular, to define the α-derivative—where $\alpha \in (0,1)$—we assert that

$$\mathcal{D}^\alpha f(x) = \frac{d}{dx} \mathcal{I}^{1-\alpha} f(x) = \frac{d}{dx} \left(\frac{1}{\Gamma(1-\alpha)} \int_0^x \frac{f(t)}{(x-t)^\alpha} dt \right) \tag{3.8.4}$$

is a good notion of fractional derivative. If $\alpha = [\alpha] + \alpha'$, we define

$$\mathcal{D}^\alpha f(x) = \mathcal{D}^{\alpha'} f^{([\alpha])}(x). \tag{3.8.5}$$

Indeed, the properties of the Γ function and the β integral

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \tag{3.8.6}$$

show by Fubini's theorem that

$$\begin{aligned} \mathcal{I}^\alpha \mathcal{I}^\beta f(x) &= \mathcal{I}^\alpha \left(t \mapsto \frac{1}{\Gamma(\beta)} \int_0^t (t-s)^{\beta-1} f(s) ds \right)(x) \\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^x (x-t)^{\alpha-1} \left(\int_0^t (t-s)^{\beta-1} f(s) ds \right) dt \\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^x (x-t)^{\alpha-1} \left(\int_0^x (t-s)^{\beta-1} \mathbf{1}_{\{0 \leq s \leq t\}} f(s) ds \right) dt \\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^x f(s) \left(\int_s^x (x-t)^{\alpha-1} (t-s)^{\beta-1} dt \right) ds. \end{aligned} \tag{3.8.7}$$

Making the linear change of variable $t - s = u$, we get

$$
\begin{aligned}
\int_s^x (x-t)^{\alpha-1}(t-s)^{\beta-1}dt &= \int_0^{x-s} (x-s-u)^{\alpha-1}u^{\beta-1}du \\
&= (x-s)^{\alpha-1}\int_0^{x-s}\left(1-\frac{u}{x-s}\right)^{\alpha-1}u^{\beta-1}du \\
&\underset{u=(x-s)v}{=} (x-s)^{\alpha-1}\int_0^1 (1-v)^{\alpha-1}((x-s)v)^{\beta-1}(x-s)dv \\
&= (x-s)^{\alpha+\beta-1}B(\alpha,\beta) \\
&= \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}(x-s)^{\alpha+\beta-1}. \qquad (3.8.8)
\end{aligned}
$$

Putting together (3.8.7) and (3.8.8), we deduce that

$$
\mathcal{I}^\alpha \mathcal{I}^\beta f(x) = \frac{1}{\Gamma(\alpha+\beta)}\int_0^x (x-s)^{\alpha+\beta-1}f(s)ds
$$

which shows that $\mathcal{I}^\alpha\mathcal{I}^\beta f = \mathcal{I}^{\alpha+\beta}f$. Therefore, we also have $\mathcal{D}^\alpha\mathcal{D}^\beta f = \mathcal{D}^{\alpha+\beta}f$ whenever both expressions make sense.

Example 3.8.1. Let us check that on an explicit example, with $f(x) = x$ and $\alpha = \beta = \frac{1}{2}$. Recalling that $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$, we have

$$
\begin{aligned}
\mathcal{I}^{\frac{1}{2}}x &= \frac{1}{\sqrt{\pi}}\int_0^x \frac{t}{\sqrt{x-t}}dt = \frac{1}{\sqrt{\pi}}\left\{\left[-2\,t\sqrt{x-t}\right]_0^x + 2\int_0^x \sqrt{x-t}\,dt\right\} \\
&= \frac{2}{\sqrt{\pi}}\left[-\frac{2}{3}(x-t)^{\frac{3}{2}}\right]_0^x = \frac{4}{3\sqrt{\pi}}x^{\frac{3}{2}},
\end{aligned}
$$

which implies that

$$
\mathcal{D}^{\frac{1}{2}}x = \frac{d}{dx}\left(\frac{4}{3\sqrt{\pi}}x^{\frac{3}{2}}\right) = \frac{2}{\sqrt{\pi}}\sqrt{x}.
$$

Now, we compute $\mathcal{D}^{\frac{1}{2}}\sqrt{x}$. We have

$$
\begin{aligned}
\mathcal{I}^{\frac{1}{2}}\sqrt{x} &= \frac{1}{\sqrt{\pi}}\int_0^x \sqrt{\frac{t}{x-t}}dt = \frac{1}{\sqrt{\pi}}\left\{\left[-2\sqrt{t(x-t)}\right]_0^x + \int_0^x \sqrt{\frac{x}{t}-1}\,dt\right\} \\
&\underset{\frac{x}{t}=y}{=} \frac{x}{\sqrt{\pi}}\int_1^\infty \sqrt{y-1}\,\frac{dy}{y^2}.
\end{aligned}
$$

Now, making the change of variable $y - 1 = u^2$, we get

$$
\begin{aligned}
\int_1^\infty \sqrt{y-1}\,\frac{dy}{y^2} &= \int_0^\infty \frac{2u^2}{(1+u^2)^2}du \\
&= \left[-\frac{1}{1+u^2}\cdot u\right]_0^\infty + \int_0^\infty \frac{du}{1+u^2} \\
&= [\arctan(u)]_0^\infty = \frac{\pi}{2},
\end{aligned}
$$

which shows that

$$
\mathcal{I}^{\frac{1}{2}}\sqrt{x} = \frac{\sqrt{\pi}}{2}x,
$$

and

$$
\mathcal{D}^{\frac{1}{2}}\sqrt{x} = \frac{d}{dx}\mathcal{I}^{\frac{1}{2}}\sqrt{x} = \frac{\sqrt{\pi}}{2},
$$

so that

$$\mathcal{D}^{\frac{1}{2}}\mathcal{D}^{\frac{1}{2}}x = \mathcal{D}^{\frac{1}{2}}\left(\frac{2}{\sqrt{\pi}}\sqrt{x}\right) = 1 = \frac{d}{dx}x$$

as expected.

Now, in order to allow for negative values of x, we will define the partial integral for Schwarz functions, and then generalise it to tempered distributions.

Definition 3.8.2. Let $f \in \mathscr{S}(\mathbb{R}^d)$. For all $1 \leq j \leq d$ and for all $\alpha \in (0,1)$, define

$$\mathcal{I}_j^\alpha f(x) = \frac{1}{\Gamma(\alpha)}\int_{-\infty}^{x_j}(x_j - t)^{\alpha-1} f(x_1, \cdots, x_{j-1}, t, x_{j+1}, \cdots, x_d)dt \tag{3.8.9}$$

and

$$\begin{aligned}\mathcal{D}_j^\alpha f(x) &= \frac{\partial}{\partial x_j}\mathcal{I}_j^{1-\alpha} f(x)\\ &= \frac{\partial}{\partial x_j}\left(\frac{1}{\Gamma(1-\alpha)}\int_{-\infty}^{x_j}\frac{f(\widehat{x_j}, t)}{(x_j - t)^\alpha}dt\right),\end{aligned} \tag{3.8.10}$$

where $\widehat{x_j} = (x_1, \cdots, x_{j-1}, 0, x_{j+1}, x_d)$ and we write by abuse of notation

$$f(\widehat{x_j}, t) = f(\widehat{x_j} + t\, e_j).$$

The α-gradient of f is defined by

$$\mathcal{D}^\alpha f = \left(\mathcal{D}_1^\alpha f, \cdots, \mathcal{D}_d^\alpha f\right). \tag{3.8.11}$$

For all $\alpha > 0$, if $\alpha = [\alpha] + \alpha'$, where $\alpha' \in (0,1)$, we define

$$\mathcal{D}_j^\alpha = \mathcal{D}_j^{\alpha'}\left(\frac{\partial^{[\alpha]}}{\partial x_j^{[\alpha]}} f\right)(x). \tag{3.8.12}$$

3.8.2 Computation in Fourier space

Now, taking the partial Fourier transform in x_j, if

$$P_\alpha(t) = \frac{1}{t^\alpha}\mathbf{1}_{\{t>0\}},$$

then we have

$$P_\alpha(t) + P_\alpha(-t) = \frac{1}{|t|^\alpha}.$$

Now, recall that by (2.7.10), we have in $\mathbb{R}^d$

$$\mathscr{F}\left(x \mapsto |x|^\alpha\right)(\xi) = 2^{\alpha+d}\pi^{\frac{d}{2}}\frac{\Gamma\left(\frac{\alpha+d}{2}\right)}{\Gamma\left(-\frac{\alpha}{2}\right)}\frac{1}{|\xi|^{\alpha+d}} \qquad -d < \alpha < 0.$$

Taking $d = 1$ and replacing α by $-\alpha$, we get

$$\mathscr{F}\left(x \mapsto \frac{1}{|x|^\alpha}\right)(\xi) = 2^{1-\alpha}\frac{\Gamma\left(\frac{1-\alpha}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)}\frac{1}{|\xi|^{1-\alpha}}.$$

Therefore, we first obtain

$$\mathscr{F}\left(P_\alpha + \widetilde{P_\alpha}\right)(\xi) = 2^{1-\alpha}\frac{\Gamma\left(\frac{1-\alpha}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)}\frac{1}{|\xi|^{1-\alpha}}. \tag{3.8.13}$$

if $\widetilde{\varphi}(x) = \varphi(-x)$. Since $P_\alpha \in \mathscr{D}'_{\mathrm{L}^2}(\mathbb{R})$ for $0 < \alpha < \frac{1}{2}$, its Fourier transform is a function, and we have in particular

$$\mathscr{F}(t \mapsto P_\alpha(-t))(\xi) = \widehat{P_\alpha}(-\xi),$$

which gives

$$\widehat{P_\alpha}(\xi) + \widehat{P_\alpha}(-\xi) = 2^{1-\alpha}\sqrt{\pi}\frac{\Gamma\left(\frac{1-\alpha}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)}\frac{1}{|\xi|^{1-\alpha}}.$$

Now, recalling (2.7.8), we have

$$\mathscr{F}(x \mapsto \mathrm{sgn}(x))(\xi) = -2i\,\mathrm{p.v.}\frac{1}{\xi},$$

and

$$P_\alpha(t) - P_\alpha(-t) = \frac{\mathrm{sgn}(t)}{|t|^\alpha}. \tag{3.8.14}$$

The Fourier transform of this function is also a function for $0 < \alpha < \dfrac{1}{2}$, and since it is an odd function, its Fourier transform is odd, purely imaginary and homogenous of degree $1 - \alpha$. Therefore, we have

$$\mathscr{F}\left(x \mapsto \frac{\mathrm{sgn}(x)}{|x|^\alpha}\right)(\xi) = -i\,c_\alpha\frac{\mathrm{sgn}(\xi)}{|\xi|^{1-\alpha}}$$

for some $c_\alpha \in \mathbb{R} \setminus \{0\}$. We will determine the constant thanks to the Parseval formula. Now, let $G(x) = e^{-\frac{x^2}{2}}$. Then, we have

$$\widehat{G}(\xi) = \sqrt{2\pi}G(\xi).$$

Furthermore, the properties of the Fourier transform imply that

$$\begin{aligned}\mathscr{F}(x \mapsto x\,G(x))(\xi) &= \int_{\mathbb{R}} x\,e^{-\frac{x^2}{2}}e^{-i\,x\,\xi}dx = \int_{\mathbb{R}} i\,\frac{\partial}{\partial\xi}\left(e^{-x^2}e^{-i\,x\,\xi}\right)dx\\ &= i\frac{\partial\widehat{f}}{\partial\xi}(\xi) = i\frac{\partial}{\partial\xi}\left(\sqrt{2\pi}e^{-\frac{\xi^2}{2}}\right) = -i\sqrt{2\pi}\,\xi\,e^{-\frac{\xi^2}{2}}.\end{aligned}$$

Using the Parseval identity, we deduce that for all function $f \in \mathscr{D}'_{\mathrm{L}^2}(\mathbb{R})$, we have

$$\begin{aligned}\int_{\mathbb{R}} f(x)\,x\,e^{-\frac{x^2}{2}}\,dx &= \frac{1}{2\pi}\int_{\mathbb{R}}\widehat{f}(\xi)\overline{-i\sqrt{2\pi}\,\xi\,e^{-\frac{\xi^2}{2}}}\,d\xi\\ &= \frac{i}{\sqrt{2\pi}}\int_{\mathbb{R}}\widehat{f}(\xi)\,\xi\,e^{-\frac{\xi^2}{2}}\,d\xi.\end{aligned} \tag{3.8.15}$$

This formula is well-suited for odd functions. Applying the formula to $f(x) = \frac{\mathrm{sgn}(x)}{|x|^\alpha}$, we deduce that

$$\begin{aligned}\int_{\mathbb{R}}\frac{\mathrm{sgn}(x)}{|x|^\alpha}x\,e^{-\frac{x^2}{2}}\,dx &= \frac{i}{\sqrt{2\pi}}\int_{\mathbb{R}}\left(-i\,c_\alpha\frac{\mathrm{sgn}(\xi)}{|\xi|^{1-\alpha}}\right)\xi\,e^{-\frac{\xi^2}{2}}\,d\xi\\ &= c_\alpha\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\frac{\mathrm{sgn}(\xi)}{|\xi|^{1-\alpha}}\,\xi\,e^{-\frac{\xi^2}{2}}\,d\xi\\ &= c_\alpha\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}|\xi|^\alpha\,e^{-\frac{\xi^2}{2}}\,d\xi.\end{aligned}$$

where we used $\mathrm{sgn}(x)x = |x|$ valid for all $x \in \mathbb{R} \setminus \{0\}$. Finally, we obtain

$$\int_{\mathbb{R}}|x|^{1-\alpha}\,e^{-\frac{x^2}{2}}\,dx = c_\alpha\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}|\xi|^\alpha e^{-\frac{\xi^2}{2}}\,d\xi. \tag{3.8.16}$$

Now, we compute by polar coordinates

$$\begin{aligned}\int_{\mathbb{R}} |x|^{1-\alpha}\, e^{-\frac{x^2}{2}}\, dx &= 2\pi \int_0^\infty r^{1-\alpha} e^{-\frac{r^2}{2}}\, r\, dr \\ &\underset{\frac{r^2}{2}=t}{=} 2\pi \int_0^\infty (2t)^{\frac{1-\alpha}{2}} e^{-t} dt \\ &= 2\pi\, 2^{\frac{1-\alpha}{2}}\, \Gamma\left(1+\frac{1-\alpha}{2}\right) \\ &= 2\pi\, 2^{\frac{1-\alpha}{2}}\, \frac{1-\alpha}{2} \Gamma\left(\frac{1-\alpha}{2}\right) \\ &= \pi\sqrt{2}\frac{1-\alpha}{2^{\frac{\alpha}{2}}}\Gamma\left(\frac{1-\alpha}{2}\right),\end{aligned}$$

where we used the formula $\Gamma(z+1) = z\,\Gamma(z)$. Therefore, we get

$$\pi\sqrt{2}\frac{1-\alpha}{2^{\frac{\alpha}{2}}}\Gamma\left(\frac{1-\alpha}{2}\right) = c_\alpha \frac{1}{\sqrt{2\pi}} \times \pi\sqrt{2}\frac{\alpha}{2^{\frac{1-\alpha}{2}}}\Gamma\left(\frac{\alpha}{2}\right),$$

which finally gives

$$c_\alpha = 2\sqrt{\pi}\frac{1-\alpha}{\alpha\, 2^\alpha}\frac{\Gamma\left(\frac{1-\alpha}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)}.$$

We deduce that

$$\mathscr{F}\left(x \mapsto \frac{\mathrm{sgn}(x)}{|x|^\alpha}\right)(\xi) = -i\, 2\sqrt{\pi}\frac{1-\alpha}{\alpha\, 2^\alpha}\frac{\Gamma\left(\frac{1-\alpha}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)}\frac{\mathrm{sgn}(\xi)}{|\xi|^{1-\alpha}} \tag{3.8.17}$$

which is valid for all $0 < \mathrm{Re}\,(\alpha) < 1$ by analytic continuation. Let us check this formula thanks to the Fourier inversion formula. We have

$$\mathscr{F}\left(x \mapsto \frac{\mathrm{sgn}(x)}{|x|^{1-\alpha}}\right)(\xi) = -i\, 2\sqrt{\pi}\frac{\alpha}{(1-\alpha)\, 2^{1-\alpha}}\frac{\Gamma\left(\frac{\alpha}{2}\right)}{\Gamma\left(\frac{1-\alpha}{2}\right)}\frac{\mathrm{sgn}(\xi)}{|\xi|^{\alpha}}.$$

Therefore, we have

$$\begin{aligned}\mathscr{F}^2\left(x \mapsto \frac{\mathrm{sgn}(x)}{|x|^\alpha}\right)(\xi) &= -i\, 2\sqrt{\pi}\frac{1-\alpha}{\alpha\, 2^\alpha}\frac{\Gamma\left(\frac{1-\alpha}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)} \times \left(-i\, 2\sqrt{\pi}\frac{\alpha}{(1-\alpha)\, 2^{1-\alpha}}\frac{\Gamma\left(\frac{\alpha}{2}\right)}{\Gamma\left(\frac{1-\alpha}{2}\right)}\frac{\mathrm{sgn}(\xi)}{|\xi|^{\alpha}}\right) \\ &= -(2\sqrt{\pi})^2\frac{1}{2}\frac{\mathrm{sgn}(\xi)}{|\xi|^\alpha} \\ &= -2\pi\frac{\mathrm{sgn}(\xi)}{|\xi|^\alpha} \\ &= 2\pi\frac{\mathrm{sgn}(-\xi)}{|\xi|^\alpha}\end{aligned}$$

which coincides with the formula given by the inverse Fourier transform.

By (3.8.14), we deduce that

$$\begin{aligned}\mathscr{F}\left(P_\alpha - \widetilde{P}_\alpha\right)(\xi) &= -i\, 2^{1-\alpha}\sqrt{\pi}\frac{1-\alpha}{\alpha}\frac{\Gamma\left(\frac{1-\alpha}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)}\frac{\mathrm{sgn}(\xi)}{|\xi|^{1-\alpha}} \\ &= -i\,\mathrm{sgn}(\xi)\frac{1-\alpha}{\alpha}\mathscr{F}\left(P_\alpha + \widetilde{P_\alpha}\right)(\xi).\end{aligned} \tag{3.8.18}$$

Recalling (3.8.13), we finally get

$$\mathscr{F}\left(x \mapsto \frac{1}{|x|^\alpha}1_{\{x>0\}}\right)(\xi) = \frac{1}{2}\left(2^{1-\alpha}\sqrt{\pi}\frac{\Gamma\left(\frac{1-\alpha}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)}\frac{1}{|\xi|^{1-\alpha}} - i\, 2^{1-\alpha}\sqrt{\pi}\frac{1-\alpha}{\alpha}\frac{\Gamma\left(\frac{1-\alpha}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)}\frac{\mathrm{sgn}(\xi)}{|\xi|^{1-\alpha}}\right)$$

$$= \frac{\sqrt{\pi}}{2^\alpha}\frac{\Gamma\left(\frac{1-\alpha}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)}\left(1-i\frac{1-\alpha}{\alpha}\mathrm{sgn}(\xi)\right)\frac{1}{|\xi|^{1-\alpha}}.$$

Now, thinking of the constant involved in the definition of the fractional derivative, we can simplify this formula thanks to Legendre's duplication formula and the Euler's reflection formula:

$$\begin{aligned}\Gamma(2z) &= \frac{2^{2z-1}}{\sqrt{\pi}}\Gamma(z)\Gamma\left(z+\frac{1}{2}\right)\\ \Gamma(z)\Gamma(1-z) &= \frac{\pi}{\sin(\pi z)}.\end{aligned}$$

Therefore, we have

$$\begin{aligned}\Gamma(1-z) &= \frac{2^{2\left(\frac{1-z}{2}\right)-1}}{\sqrt{\pi}}\Gamma\left(\frac{1-z}{2}\right)\Gamma\left(1-\frac{z}{2}\right)\\ &= \frac{1}{\sqrt{\pi}}\frac{1}{2^z}\Gamma\left(\frac{1-z}{2}\right)\times\frac{\pi}{\sin\left(\frac{\pi z}{2}\right)}\frac{1}{\Gamma\left(\frac{z}{2}\right)}\\ &= \frac{\sqrt{\pi}}{\sin\left(\frac{\pi z}{2}\right)}\frac{1}{2^z}\frac{\Gamma\left(\frac{1-z}{2}\right)}{\Gamma\left(\frac{z}{2}\right)},\end{aligned}$$

which finally gives

$$\frac{\sqrt{\pi}}{2^z}\frac{\Gamma\left(\frac{1-z}{2}\right)}{\Gamma\left(\frac{z}{2}\right)} = \sin\left(\frac{\pi z}{2}\right)\Gamma(1-z),$$

and

$$\mathscr{F}\left(x\mapsto\frac{1}{|x|^\alpha}\right)(\xi) = \sin\left(\frac{\pi\alpha}{2}\right)\Gamma(1-\alpha)\left(1-i\frac{1-\alpha}{\alpha}\mathrm{sgn}(\xi)\right)\frac{1}{|\xi|^{1-\alpha}}. \tag{3.8.19}$$

Now, recall that

$$\begin{aligned}\mathcal{D}_j^\alpha f(x) &= \frac{\partial}{\partial x_j}\left(\frac{1}{\Gamma(1-\alpha)}\int_{-\infty}^{x_j}\frac{f(\widehat{x_j},t)}{(x_j-t)^\alpha}dt\right)\\ &= \frac{1}{\Gamma(1-\alpha)}\frac{\partial}{\partial x_j}\left(P_\alpha * f(\widehat{x_j},\,\cdot\,)\right)(x_j).\end{aligned}$$

Using the product formula for the Fourier transform of a convolution and the elementary algebraic identity $\mathscr{F}(\partial_{x_j}S) = i\,\xi_j\mathscr{F}(S)$ valid for all tempered distribution $S\in\mathscr{S}'(\mathbb{R}^d)$, we deduce that

$$\mathscr{F}(P_\alpha * f(\widehat{x_j},\,\cdot\,))(\xi_j) = \widehat{P}_\alpha(\xi)\widehat{f}(\widehat{x_j},\xi_j) = \sin\left(\frac{\pi\alpha}{2}\right)\Gamma(1-\alpha)\left(1-i\frac{1-\alpha}{\alpha}\mathrm{sgn}(\xi_j)\right)\frac{1}{|\xi_j|^{1-\alpha}}\widehat{f}(\widehat{x_j},\xi_j),$$

and

$$\mathscr{F}\left(\mathcal{D}_j^\alpha f\right)(\widehat{x_j},\xi_j) = i\sin\left(\frac{\pi\alpha}{2}\right)\left(1-i\,\frac{1-\alpha}{\alpha}\mathrm{sgn}(\xi_j)\right)\frac{\xi_j}{|\xi_j|}|\xi_j|^\alpha\widehat{f}(\widehat{x_j},\xi_j). \tag{3.8.20}$$

Therefore, we obtain

$$\begin{aligned}\mathscr{F}(\mathcal{D}_j^\alpha f)(\xi) &= \mathscr{F}_{\widehat{x_j}}\left(\widehat{x_j}\mapsto\mathscr{F}\left(\mathcal{D}_j^\alpha f\right)(\widehat{x_j},\xi_j)\right)(\widehat{\xi_j})\\ &= i\sin\left(\frac{\pi\alpha}{2}\right)\left(1-i\,\frac{1-\alpha}{\alpha}\mathrm{sgn}(\xi_j)\right)\frac{\xi_j}{|\xi_j|}|\xi_j|^\alpha\widehat{f}(\xi)\\ &= \sin\left(\frac{\pi\alpha}{2}\right)\left(\frac{1-\alpha}{\alpha}+i\,\mathrm{sgn}(\xi_j)\right)|\xi_j|^\alpha\widehat{f}(\xi),\end{aligned} \tag{3.8.21}$$

and

$$\begin{aligned}|\mathscr{F}(\mathcal{D}_j^\alpha f)(\xi)| &= \sin\left(\frac{\pi\alpha}{2}\right)\sqrt{1+\frac{(1-\alpha)^2}{\alpha^2}}|\xi_j|^\alpha|\widehat{f}(\xi)|\\ &= d_\alpha|\xi_j|^\alpha|\widehat{f}(\xi)|.\end{aligned} \tag{3.8.22}$$

Therefore, we have proved the following theorem.

Theorem 3.8.3. *For all $s > 0$, the following norm*

$$\|u\|_{\mathcal{H}^s(\mathbb{R}^d)} = \sum_{\alpha\leq[s]} \|D^\alpha u\|_{\mathrm{L}^2(\mathbb{R}^d)} + \sum_{|\alpha|=[s]} \left\|\mathcal{D}^{s-[s]} D^\alpha u\right\|_{\mathrm{L}^2(\mathbb{R}^d)}$$

is equivalent to the norm $\|\cdot\|_{\mathrm{H}^s(\mathbb{R}^d)}$ defined in (3.7.1) *on $H^s(\mathbb{R}^d)$.*

Proof. By the Parseval identity and (3.8.22), we deduce that for all $\alpha \in \mathbb{N}^d$, we have

$$\int_{\mathbb{R}^d} |D^\alpha u(x)|^2 dx = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} |\xi^\alpha|^2 |\widehat{u}(\xi)|,$$

while

$$\int_{\mathbb{R}^d} |\mathcal{D}^{s-[s]} D^\alpha u(x)|^2 dx = \frac{d_{s-[s]}^2}{(2\pi)^d} \int_{\mathbb{R}^d} \sum_{j=1}^{d} |\xi_j|^{s-[s]} |\xi^\alpha|^2 |\widehat{u}(\xi)|^2 d\xi.$$

Noticing that there exists $0 < C_m < \infty$ such that

$$C_m^{-1}(1+|\xi|^2)^s \leq 1 + \sum_{|\alpha|\leq m} \sum_{j=1}^{d} |\xi_j|^{s-[s]} |\xi^\alpha|^2 \leq C_m(1+|\xi|^2)^s.$$

we conclude the proof as in □

Remark 3.8.4. An equivalent norm would be given by

$$\|u\|_{\mathbf{H}^s(\mathbb{R}^d)} = \|u\|_{\mathrm{L}^2(\mathbb{R}^d)} + \sum_{|\alpha|=[s]} \left\|\mathcal{D}^{s-[s]} D^\alpha u\right\|_{\mathrm{L}^2(\mathbb{R}^d)}.$$

3.9 Fractional Sobolev spaces in the non-Hilbertian case

3.9.1 Fractional Laplacian and fractional Sobolev spaces

Since the two characterisations of fractional Sobolev spaces coincide, we can find another equivalent norm thanks to the introduction of the fractional Laplacian. Recall that by the properties of the Fourier transformation, for all $u \in \mathscr{S}(\mathbb{R}^d)$, we have

$$\mathscr{F}(\Delta u)(\xi) = -|\xi|^2 \widehat{u}(\xi).$$

By analogy, for all $0 < s < 1$, we define the fractional Laplacian on $\mathscr{S}(\mathbb{R}^d)$ by

$$\mathscr{F}((-\Delta)^s u)(\xi) = |\xi|^{2s} \widehat{u}(\xi). \tag{3.9.1}$$

Therefore, we formally have

$$(-\Delta)^s u(x) = \mathscr{F}^{-1}\left(\xi \mapsto |\xi|^{2s}\widehat{u}(\xi)\right)(x) = (2\pi)^d \mathscr{F}^{-1}\left(\xi \mapsto |\xi|^{2s}\right) * u(x).$$

For $0 < s < 1$, $(-\Delta)^s$ is not a differential operator, but a pseudo-differential operator given formally by a convolution with a tempered distribution, and which is standardly used to define fractional Sobolev spaces. Indeed, one first shows that for all $u \in \mathscr{S}(\mathbb{R}^d)$, we have

$$(-\Delta)^s u(x) = c_s \,\mathrm{p.v.} \int_{\mathbb{R}^d} \frac{u(x)-u(y)}{|x-y|^{d+2s}} dy,$$

where

$$c_s = \frac{s\, 2^{2s}}{\pi^{\frac{d}{2}}} \frac{\Gamma\left(s+\frac{d}{2}\right)}{\Gamma(1-s)}. \tag{3.9.2}$$

Notice that the principal value is not indeed if $s < \dfrac{1}{2}$. Let us shows that this formula holds (we will however not directly prove it from the definition in (3.9.1)). First, a change of variable allows one to get rid of the principal value, and to get

$$(-\Delta)^s u(x) = -\frac{c_s}{2}\int_{\mathbb{R}^d} \frac{u(x+y)+u(x-y)-2u(x)}{|y|^{d+2s}}dy.$$

Let us compute the Fourier transform of this function. We have by Fubini's theorem since all integrals are absolutely convergent

$$\begin{aligned}
\mathscr{F}((-\Delta)^s u)(\xi) &= -\int_{\mathbb{R}^d}(-\Delta)^s u(x) e^{-i\,x\cdot\xi}dx \\
&= -\frac{c_s}{2}\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}\frac{u(x+y)+u(x-y)-2u(x)}{|y|^{d+2s}}dy\right)e^{-i\,x\cdot\xi}dx \\
&= -\frac{c_s}{2}\int_{\mathbb{R}^d}\frac{1}{|y|^{d+2s}}\left(\int_{\mathbb{R}^d}(u(x+y)+u(x-y)-2u(x))e^{-i\,x\cdot\xi}dx\right)dy \\
&= -\frac{c_s}{2}\int_{\mathbb{R}^d}\frac{1}{|y|^{d+2s}}\left(e^{i\,y\cdot\xi}+e^{-i\,y\cdot\xi}-2\right)\widehat{u}(\xi)dy \\
&= c_s\widehat{u}(\xi)\int_{\mathbb{R}^d}\frac{1-\cos(y\cdot\xi)}{|y|^{d+2s}}dy \\
&= c_s c_s'|\xi|^{2s}\widehat{u}(\xi),
\end{aligned}$$

where we used the identity

$$I(\xi) = \int_{\mathbb{R}^d}\frac{1-\cos(y\cdot\xi)}{|y|^{d+2s}}dz = c_s'|\xi|^{2s}.$$

Notice that I is a radial function, as one immediately checks by making a rotation and using the parity of cos. If $\xi \neq 0$, making a change of variable $y = \dfrac{z}{|\xi|}$, we get

$$\begin{aligned}
I(\xi) &= \int_{\mathbb{R}^d}\frac{1-\cos\left(z\cdot\frac{\xi}{|\xi|}\right)}{\left|\frac{z}{|\xi|}\right|^{d+2s}}|\xi|^d\,dz = |\xi|^{2s}I\left(\frac{\xi}{|\xi|}\right) = |\xi|^{2s}I(e_1) \\
&= |\xi|^s\int_{\mathbb{R}^d}\frac{1-\cos(z_1)}{|z|^{d+2s}}dz \\
&= c_s'|\xi|^s
\end{aligned}$$

by radiality. Therefore, we deduce that

$$c_s = I(e_1)^{-1}.$$

Let us compute c_s. Taking polar coordinates, we find by taking polar coordinates ([15], **3.2.13**)

$$\begin{aligned}
I(e_1) &= \int_{\mathbb{R}^d}\frac{1-\cos(x_1)}{|x|^{d+2s}}dx = 2\int_{\mathbb{R}^d}\frac{\sin^2\left(\frac{x_1}{2}\right)}{|x|^{d+2s}}dx = 2\int_{S^{d-1}}\left(\int_0^\infty\frac{\sin^2\left(\frac{ry_1}{2}\right)}{r^{1+2s}}dr\right)d\mathscr{H}^{d-1}(y) \\
&= 2\int_{S^{d-1}}\left(\int_0^\infty\frac{\sin^2\left(\frac{r|y_1|}{2}\right)}{r^{1+2s}}dr\right)d\mathscr{H}^{d-1}(y) \\
&\underset{\frac{r|y_1|}{2}=t}{=} 2^{1-2s}\int_{S^{d-1}}\left(|y_1|^{2s}\int_0^\infty\frac{\sin^2(t)}{t^{1+2s}}dt\right)d\mathscr{H}^{d-1}(y) \\
&= 2^{1-2s}\left(\int_0^\infty\frac{\sin^2(t)}{t^{1+2s}}dt\right)\left(\int_{S^{d-1}}|y_1|^{2s}d\mathscr{H}^{d-1}(y)\right).
\end{aligned}$$

Now, remember (2.7.9)

$$\int_{S^{d-1}}\prod_{j=1}^{d}|y_j|^{2z_j-1}d\mathscr{H}^{d-1}(y)=\frac{2\displaystyle\prod_{j=1}^{d}\Gamma(z_j)}{\Gamma\left(\displaystyle\sum_{j=1}^{d}z_j\right)}.$$

Taking $z_1=s+\dfrac{1}{2}$ and $z_j=\dfrac{1}{2}$ for all $2\leq j\leq d$, we get

$$\int_{S^{d-1}}|y_1|^{2s}d\mathscr{H}^{d-1}(y)=2\,\pi^{\frac{d-1}{2}}\frac{\Gamma\left(s+\frac{1}{2}\right)}{\Gamma\left(s+\frac{d}{2}\right)},$$

where we used $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$. For the second integral, see the Appendix 3.9.3, where we prove that

$$2^{1-2s}\left(\int_0^\infty\frac{\sin^2(t)}{t^{1+2s}}dt\right)=\frac{\sqrt{\pi}}{s2^{2s+1}}\frac{\Gamma(1-s)}{\Gamma\left(s+\frac{1}{2}\right)},$$

which finally gives

$$c_s^{-1}=I(e_1)=\frac{\pi^{\frac{d}{2}}}{s2^{2s}}\frac{\Gamma(1-s)}{\Gamma\left(s+\frac{d}{2}\right)},\tag{3.9.3}$$

which allows us to recover (3.9.2).

This prompts one to make the following definition.

Definition 3.9.1. For all $1\leq p\leq\infty$, and $s>0$ such that $s\notin\mathbb{N}$, define

$$W^{s,p}(\Omega)=W^{[s],p}(\Omega)\cap\left\{u:\text{for all }\;|\alpha|=[s],[D^\alpha u]^p_{\mathrm{W}^{(s-[s]),p}(\Omega)}=\int_\Omega\int_\Omega\frac{|D^\alpha u(x)-D^\alpha u(y)|^p}{|x-y|^{d+p(s-[s])}}dxdy<\infty\right\},$$

that we equip with the norm

$$\|u\|_{\mathrm{W}^{s,p}(\Omega)}=\|u\|_{\mathrm{W}^{[s],p}(\Omega)}+\sum_{|\alpha|=[s]}[D^\alpha u]_{\mathrm{W}^{(s-[s]),p}(\Omega)}.$$

Remarks 3.9.2. **1**. Thanks to what precedes, the definition coincides with the one of H^s for $p=2$ and $\Omega=\mathbb{R}^d$. If Ω is an extension domain, the definition can be replaced with the one involving fractional derivatives.

2. For $p=\infty$, we simply recover the Hölder space $C^{[s],\gamma}(\Omega)$, where

$$\gamma=\frac{d}{p}+s-[s].$$

The study of pseudo-differential operators is the topic of harmonic analysis, and goes beyond the scope of these lecture notes. However, let us remark that harmonic analysis gives a completely satisfying solution to the question relating the fractional derivatives and fractional Sobolev spaces.

Theorem 3.9.3. *For all $1\leq p\leq\infty$ and $s>0$, define*

$$\widetilde{W}^{s,p}(\mathbb{R}^d)=\mathscr{S}'(\mathbb{R}^d)\cap\left\{u:(1+|\xi|^2)^{\frac{s}{2}}\widehat{u}=\widehat{v},\;\; \textit{where } v\in\in L^p(\mathbb{R}^d)\right\},$$

and the associated norm

$$\|u\|_{\widetilde{\mathrm{W}}^{s,p}(\mathbb{R}^d)}=\left(\int_{\mathbb{R}^d}\left|\mathscr{F}^{-1}\left(\xi\mapsto(1+|\xi|^2)^{\frac{s}{2}}\widehat{u}(\xi)\right)(x)\right|^p dx\right)^{\frac{1}{p}}.$$

For all $1<p<\infty$, and for all $s>0$, we have $\widetilde{W}^{s,p}(\mathbb{R}^d)=W^{s,p}(\mathbb{R}^d)$ with equivalent norms.

The identity (3.8.22) implies in turns the following theorem.

Theorem 3.9.4. *For all $1 \leq p \leq \infty$ and $s > 0$, define on*

$$\mathcal{W}^{s,p}(\mathbb{R}^d) = \mathscr{S}'(\mathbb{R}^d) \cap \left\{u : \mathcal{D}^\alpha u \in L^p(\mathbb{R}^d) \text{ for all } \alpha \in \mathbb{R}^d_+ \text{ such as } |\alpha| \leq s\right\}$$

and the associated norm

$$\|u\|_{\mathcal{W}^{s,p}(\mathbb{R}^d)} = \sum_{\alpha \leq [s]} \|D^\alpha u\|_{\mathrm{L}^p(\mathbb{R}^d)} + \sum_{|\alpha|=[s]} \left\|\mathcal{D}^{s-[s]} D^\alpha u\right\|_{\mathrm{L}^p(\mathbb{R}^d)}.$$

For all $1 < p < \infty$, and for all $s > 0$, we have $\mathcal{W}^{s,p}(\mathbb{R}^d) = W^{s,p}(\mathbb{R}^d)$ with equivalent norms.

For $p = 1$, the theorem does not apply for Calderón-Zygmund operators do not generally map L^1 to L^1. Indeed, elliptic regularity shows that for all $1 < p < \infty$ and $f \in L^p(\mathbb{R}^d)$, if u solves in the distributional sense

$$\Delta u = f,$$

then $u \in W^{2,p}(\mathbb{R}^d)$. However, this result is false for $p = 1$. In Fourier, notice that

$$\begin{aligned}
\mathscr{F}(\partial^2_{x_i x_i} u)(\xi) = -\xi_i \xi_j \widehat{u}(\xi) &= \frac{\xi_i}{|\xi|}\frac{\xi_j}{|\xi|}(-|\xi|^2 \widehat{u}(\xi)) \\
&= \frac{\xi_i}{|\xi|}\frac{\xi_j}{|\xi|}\mathscr{F}(\Delta u)(\xi) \\
&= \frac{\xi_i}{|\xi|}\frac{\xi_j}{|\xi|}\widehat{f}(\xi).
\end{aligned}$$

We see in particular by the Parseval formula that

$$\left\|\nabla^2 u\right\|_{\mathrm{L}^2(\mathbb{R}^d)} = C_d \|f\|_{\mathrm{L}^2(\mathbb{R}^d)}.$$

By showing that the inverse Fourier transform of $\dfrac{i\xi_j}{|\xi|}$ is a Calderón-Zygmund operator*, we deduce that L^p boundedness for all $1 < p < \infty$. If one wants to obtain a L^1 estimate, a estimate of f in the Hardy space $\mathscr{H}^1(\mathbb{R}^d)$ (the pre-dual space of $\mathrm{BMO}(\mathbb{R}^d)$; it is a subspace of $L^1(\mathbb{R}^d)$) is necessary.

3.9.2 Return to the case $p = 2$

Let us prove that both notions coincide for $p = 2$.

Theorem 3.9.5. *For all $s > 0$, we have $W^{s,2}(\mathbb{R}^d) = H^s(\mathbb{R}^d)$ with equivalent norms.*

Proof. It suffices to treat the case $0 < s < 1$ thanks to the properties of the Fourier transform. The proof is an easy consequence of the Parseval formula. We have by Fubini's theorem and a linear change of variable

$$\begin{aligned}
\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{|u(x)-u(y)|^2}{|x-y|^{d+2s}} dx dy &= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} |u(z+y)-u(y)|^2 \, dy\right)\frac{dz}{|z|^{d+2s}} \\
&= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} |\mathscr{F}(\tau_z u - u)(\xi)|^2 \, d\xi\right)\frac{dz}{|z|^{d+2s}},
\end{aligned}$$

where we wrote as previously $\tau_z f(x) = f(x+z)$. Since

$$\mathscr{F}(\tau_z y)(\xi) = \int_{\mathbb{R}^d} u(x+z) e^{-i\, x\cdot\xi} dx \underset{x+z=y}{=} e^{i\, z\,\xi}\widehat{u}(\xi).$$

*Equal to p.v. $\dfrac{x}{|x|^{d+1}}$ up to a constant.

This identity implies that

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\frac{|u(x)-u(y)|^2}{|x-y|^{d+2s}}dxdy = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}|e^{i\,z\cdot\xi}-1|^2|\widehat{u}(\xi)|^2d\xi\right)\frac{dz}{|z|^{d+2s}}$$
$$= \frac{1}{2^{d-1}\pi^d}\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}\frac{1-\cos(z\cdot\xi)}{|z|^{d+2s}}dz\right)|\widehat{u}(\xi)|^2d\xi$$
$$= \frac{c_s'}{2^{d-1}\pi^d}\int_{\mathbb{R}^d}|\xi|^{2s}|\widehat{u}(\xi)|^2d\xi,$$

where we used the identity

$$I(\xi)=\int_{\mathbb{R}^d}\frac{1-\cos(z\cdot\xi)}{|z|^{d+2s}}dz = c_s'|\xi|^{2s}.$$

Notice that I is a radial function, as one immediately checks by making a rotation and using the parity of cos. If $\xi\neq 0$, making a change of variable $z=\frac{y}{|\xi|}$, we get

$$I(\xi)=\int_{\mathbb{R}^d}\frac{1-\cos\left(y\cdot\frac{\xi}{|\xi|}\right)}{\left|\frac{y}{|\xi|}\right|^{d+2s}}|\xi|^d\,dy = |\xi|^{2s}I\left(\frac{\xi}{|\xi|}\right)=|\xi|^{2s}I(e_1)$$
$$=|\xi|^{2s}\int_{\mathbb{R}^d}\frac{1-\cos(y_1)}{|y|^{d+2s}}dy$$
$$=c_s'|\xi|^{2s}$$

by radiality. This concludes the proof of the theorem. □

Using Theorem 3.7.5, we deduce the following result that permits to unify all notions of fractional derivatives.

Corollary 3.9.6. *For all $s>0$ such that $s\notin\mathbb{N}$, the semi-norms*

$$[u]_{\mathcal{H}^s(\mathbb{R}^d)}=\sum_{|\alpha|=[s]}\left\|\mathcal{D}^{s-[s]}D^\alpha u\right\|_{\mathrm{L}^2(\mathbb{R}^d)}$$

and

$$[u]_{\mathrm{W}^{s,2}(\mathbb{R}^d)}=\sum_{|\alpha|=[s]}\left(\int_\Omega\int_\Omega\frac{|D^\alpha u(x)-D^\alpha u(y)|^p}{|x-y|^{d+2(s-[s])}}dxdy\right)^{\frac{1}{2}}$$

are equivalent. In particular, $\mathcal{H}^s(\mathbb{R}^d)=W^{s,2}(\mathbb{R}^d)$ with equivalent norms.

3.9.3 Appendix

Let us compute the integral

$$J(s)=\int_0^\infty\frac{1-\cos(t)}{t^{1+2s}}dt = 2^{1-2s}\int_0^\infty\frac{\sin^2(t)}{t^{1+2s}}dt,$$

where $0<s<1$. Integrating by parts, we get

$$J(s)=\left[-\frac{1}{2s}\frac{1-\cos(t)}{t^{2s}}\right]_0^\infty+\frac{1}{2s}\int_0^\infty\frac{\sin(t)}{t^{2s}}dt$$
$$=\frac{1}{2s}\int_0^\infty\frac{\sin(t)}{t^{2s}}dt.$$

Since $J(s)<\infty$ for all $0<s<1$, we deduce that the integral above converges for all $0<s<1$ (which is trivial for $\frac{1}{2}<s<1$). For $0<s<\frac{1}{2}$, this integral is known (up to a change of variable $t=u^2$) as the

Fresnel integral, and can be easily computed with the help of a standard positive quarter-disk contour. We get after computations

$$J(s) = \frac{1}{2s}\sin\left((1-2s)\frac{\pi}{2}\right)\Gamma(1-2s).$$

Therefore, we have

$$\begin{aligned} c_s^{-1} &= 2\pi^{\frac{d-1}{2}}\frac{\Gamma\left(s+\frac{1}{2}\right)}{\Gamma\left(s+\frac{d}{2}\right)}\times\frac{1}{2s}\sin\left((1-2s)\frac{\pi}{2}\right)\Gamma(1-2s) \\ &= \frac{1}{s}\pi^{\frac{d-1}{2}}\frac{1}{\Gamma\left(s+\frac{d}{2}\right)}\sin\left((1-2s)\frac{\pi}{2}\right)\Gamma\left(s+\frac{1}{2}\right)\Gamma(1-2s). \end{aligned}$$

Now, recall the formulas

$$\begin{aligned} \Gamma(2z) &= \frac{2^{2z-1}}{\sqrt{\pi}}\Gamma(z)\Gamma\left(z+\frac{1}{2}\right) \\ \Gamma(z)\Gamma(1-z) &= \frac{\pi}{\sin(\pi z)}. \end{aligned}$$

We get

$$\sin\left((1-2s)\frac{\pi}{2}\right) = \sin\left(\pi\left(\frac{1}{2}-s\right)\right) = \frac{\pi}{\Gamma\left(\frac{1}{2}-s\right)\Gamma\left(\frac{1}{2}+s\right)}.$$

Therefore, we get

$$\begin{aligned} \sin\left((1-2s)\frac{\pi}{2}\right)\Gamma\left(s+\frac{1}{2}\right)\Gamma(1-2s) &= \pi\frac{\Gamma(1-2s)}{\Gamma\left(\frac{1}{2}-s\right)} = \pi\frac{\Gamma\left(2\left(\frac{1}{2}-s\right)\right)}{\Gamma\left(\frac{1}{2}-s\right)} \\ &= \pi\frac{2^{(1-2s)-1}}{\sqrt{\pi}}\Gamma\left(1-s\right) \\ &= \sqrt{\pi}\frac{1}{2^{2s}}\Gamma(1-s), \end{aligned}$$

which finally gives

$$\begin{aligned} c_s^{-1} &= \frac{1}{s}\pi^{\frac{d-1}{2}}\frac{1}{\Gamma\left(s+\frac{d}{2}\right)}\times\sqrt{\pi}\frac{1}{2^{2s}}\Gamma(1-s) \\ &= \frac{1}{s2^{2s}}\pi^{\frac{d}{2}}\frac{\Gamma(1-s)}{\Gamma\left(s+\frac{d}{2}\right)}. \end{aligned}$$

Notice that by analytic continuation, the formula proven for $0 < s < \dfrac{1}{2}$ is true for all $s \in \mathbb{C}$ such that $0 < \mathrm{Re}(s) < 1$.

3.10 Sobolev Spaces on Manifolds

We end this section with some comments on Sobolev spaces between Riemannian manifolds. A possible definition is the following. If M^m and N^n are Riemannian manifolds, by Nash's isometric embedding theorem ([27]), we can consider N^n as a subset of $\mathbb{R}^d$ with the induced metric from the canonical embedding $\iota : N^n \to \mathbb{R}^d$. Then, define

$$W^{s,p}(M^m, N^n) = W^{s,p}(M^n, \mathbb{R}^d) \cap \{u : u(x) \in N^n \text{ for almost every } x \in M^m\},$$

where M^m is equipped with its natural measure induced by its volume form (equivalently, we can consider the m-dimensional Hausdorff measure on M^m induced by its Riemannian distance).

In this general setting, very little is known in general, and questions of density of smooth functions are difficult. Furthermore, one *cannot* embed $W^{s,p}(M^m, N^n)$ with a structure of a Banach manifold in general (this is only possible provided that $sp > m$).

For further reading, one may look at Hélein's book ([22], and in the case of maps with values to the circle S^1, Brezis-Mironescu's new 500-page-long monograph! (Cf. [**?**])

Bibliography

[1] Robert A. Adams and John J. F. Fournier. *Sobolev spaces.* Academic press, Series of Pure and Applied Mathematics Series, Elsevier, 2003.

[2] Spyridon Alexakis and Rafe Mazzeo. Renormalized area and properly embedded minimal surfaces in hyperbolic 3-manifolds. *Comm. Math. Phys. 297, no. 3, 621–651.*, 2010.

[3] Spyros Alexakis and Rafe Mazzeo. Complete Willmore surfaces in $\mathbb{H}^3$ with bounded energy: boundary regularity and bubbling. *J. Differential Geom. 101, no. 3, 369–422.*, 2015.

[4] Serge Alinhac and Patrick Gérard. *Opérateurs pseudo-différentiels et théorème de Nash-Moser.* Savoirs Actuels. InterEditions, Paris; Éditions du Centre National de la Recherche Scientifique (CNRS), Meudon, 190 pp., 1991.

[5] Frederick J. Jr. Almgren. *Almgren's big regularity paper, Q-valued functions minimizing Dirichlet's integral and the regularity of area-minimizing rectifiable currents up to codimension 2. With a preface by Jean E. Taylor and Vladimir Scheffer.* World Scientific Monograph Series in Mathematics, 1. World Scientific Publishing Co., Inc., River Edge, NJ, xvi+955 pp. ISBN: 981-02-4108-9, 2000.

[6] Luigi Ambrosio. The regularity theory of area-minimizing integral currents [after almgren–de lellis–spadaro]. *Astérisque No. 380, Séminaire Bourbaki. Vol. 2014/2015, Exp. No. 1093, 139–169*, 2016.

[7] Pascal Auscher. Real Harmonic Analysis. 2012.

[8] John L. Bell and Francis Jellett. On the relationship between the boolean prime ideal theorem and two principles in functional analysis. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. 19, 191–194*, 1971.

[9] Fabrice Bethuel, Haïm Brezis, and Frédéric Hélein. *Ginzburg-Landau Vortices.* Birkhäuser, 1994.

[10] Jean-Michel Bony. *Cours d'analyse. Théorie des distributions et analyse de Fourier.* Les Éditions de l'École polytechnique, 276 pp. ISBN: 978-2-7302-0775-1, 2001.

[11] Haïm Brezis. *Analyse fonctionnelle. Théorie et applications.* Collection Mathématiques Appliquées pour la Maîtrise. Masson, Paris, xiv+234 pp., 1983.

[12] Haïm Brezis. *Functional analysis, Sobolev spaces and partial differential equations.* Universitext. Springer, New York, xiv+599 pp., 2011.

[13] Robert E. Edwards. *Functional analysis. Theory and applications.* Corrected reprint of the 1965 original. Dover Publications, Inc., New York, xvi+783 pp., 1995.

[14] Lawrence C. Evans. *Partial differential equations.* Second edition. Graduate Studies in Mathematics, 19. American Mathematical Society, Providence, RI, xxii+749 pp, 2010.

[15] Herbert Federer. *Geometric Measure Theory.* Springer-Verlag, 1969.

[16] Herbert Federer and Wendell H. Fleming. Normal and integral currents. *Ann. of Math. (2) 72, 458–520*, 1960.

[17] David Gilbarg and Neil S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer-Verlag Berlin Heidelberg GmbH, 1977.

[18] Enrico Giusti. *Minimal Surfaces and Functions of Bounded Variation*. Birkhäuser Boston, MA, Monographs in Mathematics, XII+240 pp., 1984.

[19] Alexandre Grothendieck. *Produits tensoriels topologiques et espaces nucléaires*. Mem. Amer. Math. Soc. 16, Chapter 1: 196 pp.; Chapter 2: 140 pp., 1955.

[20] Uffe Haagerup. A new upper bound for the complex Grothendieck constant. *Israel J. Math. 60 (1987), no. 2, 199–224*, 1987.

[21] Piotr Hajłasz. Sobolev mappings, co-area formula and related topics. *Proceedings on Analysis and Geometry (Russian) (Novosibirsk Akademgorodok, 1999), 227–254, Izdat. Ross. Akad. Nauk Sib. Otd. Inst. Mat., Novosibirsk*, 2000.

[22] Frédéric Hélein. *Applications harmoniques, lois de conservation, et repères mobiles*. Diderot éditeur, Sciences et Arts, 1996.

[23] Lars Hörmander. *The analysis of linear partial differential operators. I. Distribution theory and Fourier analysis*. Second edition. Grundlehren der mathematischen Wissenschaften, 256. Springer-Verlag, Berlin, 1990. xii+440 pp., 1983.

[24] Lars Hörmander. *The analysis of linear partial differential operators. II. Differential operators with constant coefficients*. Grundlehren der mathematischen Wissenschaften, 257. Springer-Verlag, Berlin, 1983. viii+391 pp., 1983.

[25] Lars Hörmander. *The analysis of linear partial differential operators. III. Pseudodifferential operators*. Grundlehren der mathematischen Wissenschaften, 274. Springer-Verlag, Berlin, 1985. viii+525 pp., 1985.

[26] Lars Hörmander. *The analysis of linear partial differential operators. IV. Fourier integral operators*. Grundlehren der mathematischen Wissenschaften, 275. Springer-Verlag, Berlin, 1985. vii+352 pp., 1985.

[27] John F. Nash Jr. The imbedding problem for riemannian manifolds. *Ann. of Math. (2) 63, 20–63*, 1956.

[28] Yitzhak Katznelson. *An introduction to harmonic analysis*. Second corrected edition. Dover Publications, Inc., New York, xiv+264 pp., 1976.

[29] Fernando Codá Marques and André Neves. Min-max theory, willmore conjecture and the energy of links. *Bull. Braz. Math. Soc. (N.S.) 44, no. 4, 681–707*, 2013.

[30] Fernando Codá Marques and André Neves. Min-Max theory and the Willmore conjecture. *Ann. of Math., 179, 683-782*, 2014.

[31] Fernando Codá Marques and André Neves. The Willmore Conjecture. *Jahresber Dtsch Math-Ver 116:201–222*, 2014.

[32] Gilles Pisier. Grothendieck's theorem, past and present. *Bull. Amer. Math. Soc. (N.S.) 49, no. 2, 237–323*, 2012.

[33] Jon T. Pitts. *Existence and regularity of minimal surfaces on Riemannian manifolds*. Mathematical Notes, 27. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo. iv+330 pp., 1981.

[34] Tristan Rivière. Analysis aspects of Willmore surfaces. *Invent. Math., 174, 1-45*, 2008.

[35] Laurent Schwartz. *Théorie des distributions*. Nouvelle édition, entiérement corrigée, refondue et augmentée. Publications de l'Institut de Mathématique de l'Université de Strasbourg, IX-X Hermann, Paris, xiii+420 pp., 1966.

[36] Laurent Schwartz. *Un mathématicien aux prises avec le siècle.* Éditions Odile Jacob, 528 pages, EAN13 : 9782738104625, 1997.

[37] Elias M. Stein. *Singular integrals and differentiability properties of functions.* Princeton Mathematical Series, No. 30 Princeton University Press, Princeton, N.J. xiv+290 pp., 1970.

[38] Luc Tartar. *An introduction to Sobolev spaces and interpolation spaces.* Lecture Notes of the Unione Matematica Italiana 3, Springer, Berlin, 2007.

[39] Hassler Whitney. The self-intersections of a smooth n-manifold in $2n$-space. *Ann. of Math., Vol. 45, No. 2, p. 220-246, April*, 1944.

[40] Hassler Whitney. *Geometric integration theory.* Princeton University Press, Princeton, N. J., xv+387 pp., 1957.

[41] C. Robin Graham Edward Witten. Conformal anomaly of submanifold observables in Ads/CFT correspondence. *Nuclear Phys. B 546, no. 1-2, 52–64*, 1999.