Lecture Notes[1] for the course

Distributions and Partial Differential Equations

Master Mathématiques Fondamentales et Applications

Université Paris-Saclay, Sep-Dec 2022

Stéphane Nonnenmacher

September 2022

1. following earlier versions by Thierry Ramond and Patrick Gérard

Contents

Introduction : Why Distributions ?

The notion of *distribution* emerged during the twentieth century, as a powerful tool in the study of partial differential equations (PDEs). The need to generalize the concept of a "function" also emerged from different parts of theoretical physics, in particular electromagnetism and quantum mechanics. On the mathematical side, major contributors to the formal definition of distributions are Jean Leray, Serguei Sobolev and Laurent Schwartz.

In this introduction, we present four simple motivations for the emergence of distributions.

Differentiating non differentiable functions

Taking the derivative of non-differentiable functions

In differential calculus, one immediately encounters the unpleasant fact that not every function is differentiable. The purpose of distribution theory is to remedy this flaw. Recall that, given a function f on an open interval $I \subset \mathbb{R}$, the derivative function f' is defined as

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h},$$

provided this limit exists at every point $x \in I$. Let us mention a classical elementary application of this notion. If f' exists, f is a nondecreasing function on I if and only if[1] its derivative $f'(x) \geq 0$ at every point $x \in I$. However, there are many examples of nondecreasing functions such that f' cannot be defined on the whole of I: a typical example is the Heaviside function,

$$H(x) = \begin{cases} 1 & \text{if} \quad x \geq 0 \\ 0 & \text{if} \quad x < 0 \end{cases}$$

Of course, H admits a derivative at every point $x \neq 0$, and this derivative is 0. But this is not sufficient to conclude that H is nondecreasing, since the derivative of $-H$ enjoys the same property ! It is therefore necessary to extend to $x = 0$ the derivative of H, in a way which takes into account the discontinuity of H at $x = 0$. In fact, to every locally integrable function f, we shall associate a mathematical object — a distribution — called the derivative of f, with the property that f is nondecreasing iff its derivative is nonnegative (the nonnegativity of a distribution will need to be properly defined).

1. in the sequel, the expression "if and only if" will be abbreviated by *iff*

Fourier series

In his famous memoir *Théorie analytique de la chaleur* (1822), Joseph Fourier introduced, for every "reasonable" 2π-periodic function f, the coefficients

$$c_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,\mathrm{e}^{-inx}\,dx\ ,\ n\in\mathbb{Z}$$

and he stated that

$$f(x) = \sum_{n=-\infty}^{+\infty} c_n(f)\,\mathrm{e}^{inx}\ .$$

Throughout the nineteenth and the early twentieth century, many mathematicians have tried to give sense to the above equality for the largest possible class of functions f. Dirichlet proved it for any C^1 function, yet Kolmogorov made the striking observation that there exist locally integrable functions f such that the above Fourier series is divergent for every $x\in\mathbb{R}$! We shall see below that, even in such an unfavorable situation, the above series is convergent, but in a different sense, namely *in the sense of distributions*. In fact, for any element u in the class of 2π-periodic distribution (this class includes, in particular, the locally integrable functions), we shall define a sequence $(c_n(u))_{n\in\mathbb{Z}}$ of Fourier coefficients such that the corresponding series is convergent in the sense of distributions, and its sum equals u. Furthermore, the derivative of the distribution u (which is also a 2π-periodic distribution) can be obtained by summing the series of derivatives, namely the series with coefficients $(inc_n(u))$. This interplay between Fourier series and differentiation was a major reason for their introduction by Fourier back in 1822; the theory of distributions manages to extend the applicability of this connection to a much larger class of objects.

Electrostatics

If f is a function on $\mathbb{R}^3$ which represents an electric charge distribution, the electric potential u generated by this charge distribution satisfies the Poisson equation

$$-\Delta u = f. \tag{0.0.1}$$

If f is smooth enough — say C^1 — and small enough at infinity (for instance it vanishes outside of some ball), one can prove that there exists a unique C^2 solution u of this equation which goes to 0 at infinity: this solution is given by the integral

$$u(x) = \frac{1}{4\pi}\int_{\mathbb{R}^3}\frac{f(y)}{|x-y|}\,dy\ . \tag{0.0.2}$$

This formula still has a meaning if f is bounded and has compact support, but is not necessarily continuous (we will note $f\in L^\infty_{comp}(\mathbb{R}^3)$). In that case u is no more C^2 in general, yet it is tempting to try to interpret the Poisson equation (0.0.1) in this more general situation as well.

Actually, in Physics charge distributions are often modelized as being supported on surfaces, curves, points, or forming microscopic dipoles... Such charge distributions cannot be described by L^∞_{comp}

functions. Yet, in all these cases, the formula (0.0.2) still makes sense, provided one integrates on the appropriate subset of $\mathbb{R}^3$; what is then the status of the Poisson equation (0.0.1)? We will show that, in all these cases, it is possible to interpret this equation *in the sense of distributions*.

Let us remark that the word *distribution* actually takes its origin in this application to electrostatics, a (mathematical) distribution primarily represented a distribution of electric charge.

Quantum mechanics

Quantum mechanics is a theory based on the Hilbert space $L^2(\mathbb{R}^3.\mathbb{C})$: the state of a quantum particle[2] is represented by a wavefunction, which is a square integrable function $u \in L^2(\mathbb{R}^3, \mathbb{C})$, with L^2 norm equal to unity. Nevertheless, the founding fathers of the theory (in particular Paul Dirac) felt the need to extend the notion of quantum state to objects which are not in L^2. In particular, Dirac introduced his famous "delta function", namely a "function" $\delta(x)$ which vanishes at each point $x \neq 0$, but such that, *formally*

$$\int_{\mathbb{R}^3} \delta(x)\, dx = 1.$$

If one had to give a value to $\delta(0)$, this value should be $+\infty$ for the above identity to hold. So this δ cannot be a function: it is not an element of L^2. Mathematically, this "delta function" will appear as one of the simplest nontrivial distributions, the delta distribution at the origin, often denoted by δ_0, .

Dirac also invented the "bracket" notation $\langle u|v\rangle$ for the scalar bracket on L^2, and extended the notation to his singular objects: his notation $\langle x_0|v\rangle$ will represent the distributional bracket between the delta distribution δ_{x_0} and the "test function" v; we will denote it instead by the brackets $\langle \delta_{x_0}, u\rangle_{\mathscr{D}',\mathscr{D}}$.

2. Here I only consider *scalar* particles, that is particles without spin.

Chapter 1

Distributions in one space dimension

We have chosen to start our presentation of the theory of distributions by the one dimensional setting. This will give simpler expressions, yet will already convey most of the ideas and methods of the theory. Note that this one dimensional theory is already useful in practice, for instance in signal analysis, or in 1-dimensional models in quantum or wave mechanics.

Distributions will appear below as “duals” of smooth functions (more precisely, linear forms acting on spaces of smooth functions), a distribution will be described by the way it acts on all smooth functions. For this reason, these smooth functions will be called *test functions*, since their rôle will be to test the distributions.

We will start our presentation by describing in some detail these classes of test functions.

1.1 Background on differential calculus on $\mathbb{R}$

In this section, we recall elementary facts about smooth functions of one real variable. The functions will be defined on some open (nonempty) interval $I \subset \mathbb{R}$. This interval may be bounded ($I =]a, b[$), semibounded ($I =]-\infty, b[$ or $I =]a, \infty[$), or simply $I = \mathbb{R}$. The theory will essentially be identical in all cases. For the reader’s ease, at first read it can be convenient to take $I = \mathbb{R}$.

1.1.1 Basic properties of smooth functions on I

A smooth function on an open interval $I \subset \mathbb{R}$ is a function $\varphi : I \to \mathbb{C}$ whose derivatives of any order $\varphi', \varphi'', \ldots, \varphi^{(k)}, \ldots$ exist and are continuous on I. A linear combination of smooth functions is a smooth function, and we denote $\mathcal{C}^\infty(I)$ the vector space formed by all smooth functions on I.

If $\varphi \in \mathcal{C}^\infty(I)$ and J is an open subset of I, the function defined on J by $x \mapsto \varphi(x)$ is a smooth function on J, that we denote by $\varphi_{|_J}$. This function is called the restriction of φ on J.

The product of two smooth functions is smooth, and the derivatives of the product are computed by the *Leibniz formula*[1]:

$$(\varphi_1\varphi_2)^{(k)} = \sum_{j=0}^{k} \binom{k}{j} \varphi_1^{(j)} \varphi_2^{(k-j)}.$$

A smooth function $\varphi : I \to \mathbb{C}$ satisfies the Taylor formula with integral remainder, for any base point $x_0 \in I$ and any order $m \in \mathbb{N}$:

$$\forall x \in I \; , \; \varphi(x) = \sum_{k=0}^{m} \frac{(x-x_0)^k}{k!} \varphi^{(k)}(x_0) + \frac{(x-x_0)^{m+1}}{m!} \int_0^1 (1-s)^m \varphi^{(m+1)}(x_0 + s(x-x_0))ds.$$

This formula can be proved by integrating by parts the last term on the right hand side. We will use it several times in those notes.

Exercise 1.1.1 Prove Hadamard's lemma: if $\varphi \in \mathcal{C}^\infty(I)$ satisfies $\varphi(x_0) = 0$, there exists a function $\psi \in \mathcal{C}^\infty(I)$ such that $\varphi(x) = (x - x_0)\psi(x)$ for any $x \in I$.
In other words, we can factorize from φ the monomial $(x - x_0)$, and the quotient is still a smooth function.

1.1.2 Support of a continuous function

If f is continuous on I and J is an open subset of I, we say that f vanishes on J if it vanishes at every point of J, or, equivalently, if $f_{|J}$ is the null function.

Definition 1.1.2 [Support of a function] Let $f : I \to \mathbb{C}$ be a continuous function. The support of f is the complement of the union of all the open sets in I where f vanishes. This set is denoted by supp f.

Note that the support of f is a closed set. It is also the closure of the set of $x \in I$ such that $f(x) \neq 0$. The following characterization is often useful:

$$x_0 \notin \text{supp}\, f \iff \exists V \text{ neighborhood of } x_0 \text{ such that } f_{|V} = 0.$$

Exercise 1.1.3 Show that

$$\text{supp}(f_1 f_2) \subset \text{supp}\, f_1 \cap \text{supp}\, f_2.$$

Are these two sets equal?

Of course, if $\varphi \in \mathcal{C}^\infty(I)$ vanishes on an open set $J \subset I$, all its derivatives vanish as well on J, and, therefore, for any integer k,

$$\text{supp}\, \varphi^{(k)} \subset \text{supp}\, \varphi.$$

1. Here $\binom{k}{j} = \frac{k!}{j!(k-j)!}$ are the binomial coefficiens.

1.2 Test functions

In this section we introduce and manipulate test functions on some open interval $I \subset \mathbb{R}$.

1.2.1 The space of test functions

Definition 1.2.1 We denote by $\mathscr{D}(I) = \mathcal{C}_0^\infty(I)$ (also $\mathcal{C}^\infty_{\text{comp}}(I)$) the vector space of functions which are $\mathcal{C}^\infty$ on I, and whose support is a compact subset of I. Equivalently, a smooth function belongs to $\mathscr{D}(I)$ if it vanishes outside some closed segment $[a,b] \subset I$.

If J is an open subinterval of I, one may identify a function $\varphi \in \mathscr{D}(J)$ with its extension $\tilde\varphi \in \mathscr{D}(I)$, that is the function defined as

$$\tilde\varphi(x) = \varphi(x) \text{ for } x \in J, \quad \tilde\varphi(x) = 0 \text{ for } x \in I \setminus J.$$

Indeed, $\tilde\varphi$ is automatically smooth and with compact support on I for $\varphi \in \mathscr{D}(J)$. On the other hand, a function $\varphi \in \mathcal{C}_0^\infty(\mathbb{R})$ can be identified with its restriction $\varphi_{|_I} \in \mathscr{D}(I)$ for any open interval I that contains supp f.

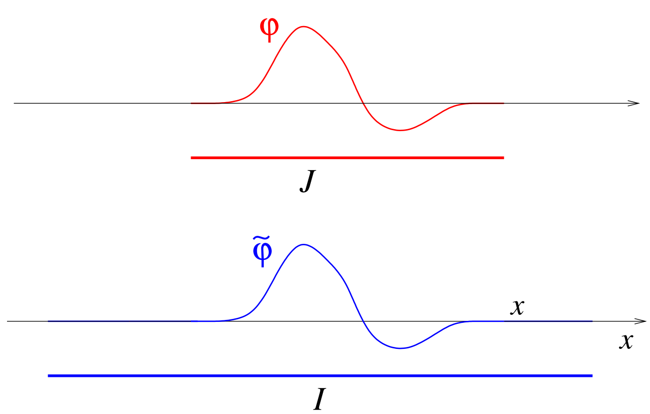


Figure 1.1: A test function on J and its extension to I.

Exercise 1.2.2 Let f and φ be two functions in $\in L^1(\mathbb{R})$. Show that the convolution $f * \varphi$ of f and φ, given by

$$f * \varphi(x) = \int f(x-y)\varphi(y)dy,$$

is defined almost everywhere, and is an L^1 function.

Suppose moreover that $\varphi \in \mathcal{C}_0^\infty(\mathbb{R})$, and show that $f * \varphi$ is then smooth. At last, if f is also continuous, show that

$$\text{supp } f * \varphi \subset \text{supp } f + \text{supp } \varphi.$$

Proofs: (see TD1, Ex. 5,6,7).

It is not immediately clear that $\mathscr{D}(I)$ is not reduced to the null function. One knows for example that the only compactly supported *real analytic function* on $\mathbb{R}$ is the null function, due to the principle of isolated zeroes. Yet, one has the following

Proposition 1.2.3 The space $\mathscr{D}(I)$ is nontrivial. More precisely, for every $x_0 \in I$ and $r > 0$ such that $[x_0-r, x_0+r] \subset I$, there exists a function $\varphi_{x_0,r} \in \mathcal{C}^\infty(I)$ such that $\operatorname{supp} \varphi_{x_0,r} = [x_0-r, x_0+r]$.

Proof.— Here we give a "standard" construction, which uses the function $\varphi : \mathbb{R} \to \mathbb{R}$ given by:

$$\varphi(t) = \begin{cases} e^{-1/t} & , t > 0, \\ 0 & , t \leq 0. \end{cases}$$

Let us check that φ is smooth on $\mathbb{R}$. Indeed, it is obviously smooth on $\mathbb{R}^*$, with $\varphi^{(k)}(t) = 0$ for all $t < 0$. On the other hand, for $t > 0$, one can prove by induction that

$$\forall k \in \mathbb{N},\ \varphi^{(k)}(t) = P_k\left(\frac{1}{t}\right) e^{-1/t}$$

where P_k is a polynomial of degree $2k$. Therefore, for any $k \geq 0$, $\varphi^{(k)}(t) \to 0$ as $t \to 0^+$. Hence $\varphi^{(k)}(t)$ admits the same limit when $t \to 0$ on both sides. This shows that $\varphi^{(k)}(0) = 0$, and that this function is continuous on $\mathbb{R}$.

Now let $x_0 \in I$, and $r > 0$ such that $[x_0 - r, x_0 + r] \subset I$. We then define the function $\varphi_{x_0,r} : I \to \mathbb{R}^+$ by

$$\varphi_{x_0,r}(x) = \varphi(r^2 - |x - x_0|^2), \quad \forall x \in \mathbb{R}.$$

This function is smooth, with $\operatorname{supp} \varphi_{x_0,r} = [x_0 - r, x_0 + r]$. In particular it belongs to $\mathscr{D}(I)$.

(A different construction of a function in $\mathscr{D}(\mathbb{R})$ is given in TD1, Ex. 11). □

We will often use a particular type of test function, equal to unity on some interval.

Proposition 1.2.4 (Cutoff functions) Let $I \subset \mathbb{R}$ an open set, and a segment $[a, b] \subset I$. There exists a function $\psi \in \mathscr{D}(I)$ such that

i) $\psi = 1$ on $[a, b]$,

ii) $\forall x \in I, \psi(x) \in [0, 1]$.

In French, such a function is called "fonction plateau", while in English it is rather referred to as a "cut-off function".

Proof.— First of all, we prove that, for every $\alpha < \beta$, there exists $\chi_{\alpha,\beta} \in \mathcal{C}^\infty(\mathbb{R})$ such that *i)* $\chi_{\alpha,\beta} = 0$ on $]-\infty, \alpha]$,

ii) $\forall x \in \mathbb{R}\,, \chi_{\alpha,\beta}(x) \in [0,1]$,

iii) $\chi_{\alpha,\beta} = 1$ on $[\beta, +\infty[$.
(such a $\chi_{\alpha,\beta}$ is called a "smooth step function").

Let $\varphi_{x_0,r} \in \mathcal{C}^\infty(\mathbb{R})$ be the function defined in Prop. 1.2.3, with the parameters

$$x_0 = \frac{\alpha+\beta}{2}\,,\; r = \frac{\beta-\alpha}{2}\,.$$

Then it is easy to check that

$$\chi_{\alpha,\beta}(x) \stackrel{\text{def}}{=} \frac{\int_{-\infty}^{x} \varphi_{x_0,r}(t)\,dt}{\int_{-\infty}^{+\infty} \varphi_{x_0,r}(t)\,dt}$$

satisfies the required properties.
Coming back to $[a,b] \subset I$, select $a', b' \in I$ such that $a' < a < b < b'$. Then we may just take the product

$$\psi(x) = \chi_{a',a}(x)(1 - \chi_{b,b'}(x))$$

□

1.2.2 Smooth partitions of unity

An important use of these cutoff functions lies in the construction of *smooth partitions of unity*. We start by defining the notion of *open cover* of a compact set.

Definition 1.2.5 (Open cover) Fix I an open interval, and let $K \subset I$ be a compact (bounded and closed) subset of I. A finite open cover of K inside I is a family of open intervals $I_1, \ldots, I_n \subset I$, such that

$$K \subset \bigcup_{j=1}^{n} I_j\,.$$

To any open cover we may associate a smooth partition of unity.

Proposition 1.2.6 (Smooth partitions of unity) Let $K \subset I$ be a compact subset of the open interval I, and let $I_1, \ldots, I_n \subset I$ be a finite open cover of K.

Then, there exist functions $\chi_1 \in \mathcal{C}_0^\infty(I_1, [0,1]), \ldots, \chi_n \in \mathcal{C}_0^\infty(I_n, [0,1])$, such that their sum

$$\chi = \sum_{i=1}^{n} \chi_i \quad \text{is equal to unity on a neighbouhood of } K,$$

and $\chi \in \mathcal{C}_0^\infty(I, [0,1])$.

As a consequence of this smooth partition of unity, we easily draw the following

Corollary 1.2.7 Let $\varphi \in \mathcal{C}_0^\infty(I)$, and assume $I_1, \ldots, I_n \subset I$ form a finite open cover of supp φ inside I:

$$\operatorname{supp} \varphi \subset I_1 \cup \cdots \cup I_n \ .$$

Then there exist $\varphi_1 \in \mathcal{C}_0^\infty(I_1), \ldots, \varphi_n \in \mathcal{C}_0^\infty(I_n)$, such that

$$\varphi = \varphi_1 + \cdots + \varphi_n \ .$$

Indeed, once we have at our disposal a partition of unity $\chi_i, \ldots, \chi_n$ associated with the cover of supp φ, it suffices to take $\varphi_j = \chi_j \varphi$ for $j = 1, \ldots, n$. □

Remark 1.2.8 In the whole chapter, our distributions are based on an open interval I. It is actually possible to generalize the construction of smooth partitions of unity to the case where the compact set K and its open cover $\bigcup_{i=1}^n I_i \supset K$ are contained in some general open set $\Omega \subset \mathbb{R}$, instead of an interval I. We leave to the reader the proofs of this generalization.

Proof.— Let us now prove Proposition 1.2.6. We start with an elementary

Lemma 1.2.9 ("Shrinking lemma") Take $I_1, \ldots, I_n$ a family of open subintervals of I. Then, for every compact subset K such that

$$K \subset I_1 \cup \cdots \cup I_n \ ,$$

there exist segments[2] $[a_1, b_1] \subset I_1, \ldots, [a_n, b_n] \subset I_n$ such that

$$K \subset]a_1, b_1[\cup \cdots \cup]a_n, b_n[\ .$$

To prove this Lemma we proceed by induction on n. For $n = 1$, K is a compact subset of the open interval I_1, hence there exists a segment $[\alpha_1, \beta_1] \subset I_1$ such that $K \subset [\alpha_1, \beta_1]$. For instance, one may choose $\alpha_1 = \min K$, $\beta_1 = \max K$. Then we just choose $a_1, b_1 \in I$ such that $a_1 < \alpha_1 < \beta_1 < b_1$. Assume now that for some $n \geq 2$, the result is true at the level $n-1$; let us and prove that it holds at the level n. The set

$$K' = K \setminus I_n$$

is closed and contained in K, hence it is a compact subset of I, and

$$K' \subset I_1 \cup \cdots \cup I_{n-1} \ .$$

Applying the induction assumption, there exist $[a_1, b_1] \subset I_1, \ldots, [a_{n-1}, b_{n-1}] \subset I_{n-1}$ such that

$$K' \subset]a_1, b_1[\cup \cdots \cup]a_{n-1}, b_{n-1}[\ .$$

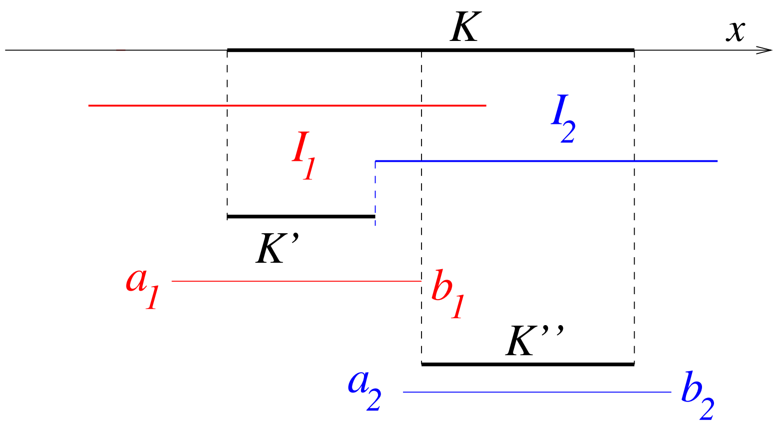


Figure 1.2: Intervals $]a_i, b_i[$ constructed in the "Shrinking Lemma" (here $n = 2$).

Now consider the compact subset

$$K'' = K \setminus \Big( ]a_1, b_1[\cup \cdots \cup]a_{n-1}, b_{n-1}[\Big) \subset I_n \ .$$

Applying the assumption at the level 1, there exists a segment $[a_n, b_n] \subset I_n$ such that

$$K'' \subset]a_n, b_n[\ .$$

As a consequence, we obtain

$$K \subset]a_1, b_1[\cup \cdots \cup]a_n, b_n[\ ,$$

as announced. □

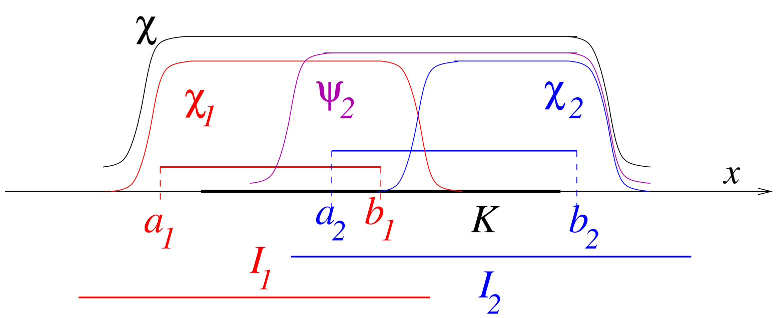


Figure 1.3: Partition of unity associated with the open cover $K \subset I_1 \cup I_2$, as constructed in Proposition 1.2.6.

Let us use this Lemma to prove Proposition 1.2.6. Let $[a_1, b_1], \ldots, [a_n, b_n]$ be the segments provided by the Lemma. From Proposition 1.2.4 we may construct cutoff functions $\psi_1 \in \mathcal{C}_0^\infty(I_1, [0,1])$,...,$\psi_n \in \mathcal{C}_0^\infty(I_n, [0,1])$ such that $\psi_1 = 1$ on $[a_1, b_1]$,...,$\psi_n = 1$ on $[a_n, b_n]$. We define

$$\begin{aligned}
\chi_1 &= \psi_1 \ , \\
\chi_2 &= \psi_2(1 - \psi_1) \ , \\
&\ \ldots \\
\chi_n &= \psi_n(1 - \psi_{n-1}) \ldots (1 - \psi_1) \ .
\end{aligned}$$

Then $\chi_1 \in \mathscr{D}(I_1,[0,1]),\dots,\chi_n \in \mathscr{D}(I_n,[0,1]),$, and the difference

$$\begin{aligned}1-(\chi_1+\dots+\chi_n) &= 1-\psi_1-\psi_2(1-\psi_1)-\dots-\psi_n(1-\psi_{n-1})\dots(1-\psi_1))\\ &= (1-\psi_1)(1-\psi_2-\dots-\psi_n(1-\psi_{n-1})\dots(1-\psi_2))\\ &= (1-\psi_1)\dots(1-\psi_n).\end{aligned}$$

By construction, the function $(1-\psi_1)\dots(1-\psi_n)$ vanishes identically on $[a_1,b_1]\cup\dots\cup[a_n,b_n]$, which is a neighbourhood of K, so $\chi=1$ on a neighbourhood of K as stated in the Proposition. □

1.2.3 Convergence in the space of test functions

Once we have defined the space of test functions $\mathscr{D}(I)$, it is important to define a proper notion of convergence on this space, or equivalently a topology on $\mathscr{D}(I)$. Test functions are smooth and compactly supported, it is natural that the topology on $\mathscr{D}(I)$ takes into account all these properties.

The natural notion of convergence for continuous functions is that of uniform convergence, since it is the simplest one for which the limit of a sequence of continuous functions is continuous. For test functions, the notion of convergence is provided by the next definition.

Definition 1.2.10 Let $(\varphi_j)_{j\in\mathbb{N}}$ be a sequence of functions in $\mathscr{D}(I)$, and $\varphi\in\mathscr{D}(I)$. We say that (φ_j) converges to φ in $\mathscr{D}(I)$ (or in the $\mathscr{D}(I)$-sense), if:

i) Uniform support There exists a segment $[a,b]\subset I$ such that $\operatorname{supp}\varphi_j\subset[a,b]$ for all $j\in\mathbb{N}$ (we will sometimes denote this property by $\varphi_j\in\mathscr{D}_{[a,b]}(I)$).

ii) Uniform convergence of all derivatives For any $k\in\mathbb{N}$,

$$\|\varphi_j^{(k)}-\varphi^{(k)}\|_\infty := \sup_{x\in I}|\varphi_j^{(k)}(x)-\varphi^{(k)}(x)| \xrightarrow{j\to\infty} 0.$$

We denote this convergence by

$$\varphi = \mathscr{D}-\lim_{j\to+\infty}\varphi_j\,.$$

Obviously, when it exists, the limit function φ of a sequence $(\varphi_j)_{j\in\mathbb{N}}$ is unique, and is also supported in $[a,b]$.

Remark 1.2.11 Under the conditions of the above definition, the limit function necessarily satisfies $\operatorname{supp}\varphi\subset[a,b]$.

Exercise 1.2.12 Assuming the uniform convergence $\|\varphi_j-\varphi\|_\infty\to 0$ alone does not imply, a priori, the uniform convergence of derivatives.
Construct a counterexample with "fast small oscillations".

Exercise 1.2.13 Let $\varphi \in \mathscr{D}(\mathbb{R})$, and, for $t \neq 1$, denote by $\psi_t \in \mathscr{D}(\mathbb{R})$ the function given by

$$\psi_t(x) = \frac{\varphi(tx) - \varphi(x)}{t-1}\,.$$

Show that the family (ψ_t) converges in $\mathscr{D}(\mathbb{R})$ as t tends to 1.

Remark 1.2.14 (Strict inductive topology on $\mathscr{D}(I)$) The definition 1.2.10 of the convergence of sequences in $\mathscr{D}(I)$ corresponds to a slighty subtle topology on $\mathscr{D}(I)$, called the *strict inductive topology*. Why "inductive"? Because one first defines a topology on the space of smooth functions supported on a given segment $K \subset I$ (which we denote by $\mathscr{D}_K(I)$). This topology is defined through the family of seminorms

$$N_{K,k}(\varphi) = \max_{0\leq j\leq k} \|\varphi^{(j)}\|_\infty, \qquad \varphi \in \mathscr{D}_K(I),$$

namely a local base of open neighbourhoods in $\mathscr{D}_K(I)$ of the zero function can be defined by

$$U_{K,k} = \{\varphi \in \mathscr{D}_K(I)\,,\, N_{K,k}(\varphi) < 1/k\}, \qquad k \geq 1\,.$$

With this base, one recovers the fact that a sequence $(\varphi_j) \subset \mathscr{D}_K(I)$ converges to zero in $\mathscr{D}_K(I)$ iff all seminorms $N_{K,k}(\varphi_j)$ converge to zero when $j \to \infty$.

To induce a topology on $\mathscr{D}(I)$, one considers a growing sequence of segments $K_1 \subset \cdots K_i \subset K_{i+1} \subset \cdots$, such that

$$I = \bigcup_{i\geq 1} K_i.$$

It is obvious that for any i and any function $\varphi \in \mathscr{D}_{K_i}(I)$, that function is also an element of $\mathscr{D}_{K_{i+1}}(I)$, and the seminorms $N_{K_{i+1},k}(\varphi) = N_{K_i,k}(\varphi)$. As a result, the intersection between the open set $U_{K_{i+1},k} \subset \mathscr{D}_{K_{i+1}}(I)$ and $\mathscr{D}_{K_i}(I)$ is just the set $U_{K_i,k}$, which is open in $\mathscr{D}_{K_i}(I)$; one says that the topology on $\mathscr{D}_{K_i}(I)$ is induced by the topology on $\mathscr{D}_{K_{i+1}}(I)$.

Now, there exists a single topology on $\mathscr{D}(I)$, called the strict inductive topology, such that for each open set $U \subset \mathscr{D}(I)$ of this topology and any index i, the set $U \cap \mathscr{D}_{K_i}(I)$ is open for the topology of $\mathscr{D}_{K_i}(I)$. A family of neighbourhood of the zero function in $\mathscr{D}(I)$ is given by the sets

$$U_k = \bigcup_{i\geq 1} U_{K_i,k}\,, \quad k \geq 1.$$

Let us quote a few properties of this inductive topology:

1. each $\mathscr{D}_{K_i}(I)$ is a closed subspace of $\mathscr{D}(I)$.
2. each space $\mathscr{D}_{K_i}(I)$ is separated and complete w.r.to its topology, as a consequence $\mathscr{D}(I)$ is separated and complete in the inductive topology.
3. although the topology on each $\mathscr{D}_{K_i}(I)$ is metrizable, the strict inductive topology on $\mathscr{D}(I)$ is *not* metrizable.
4. a sequence $(\varphi_j) \subset \mathscr{D}(I)$ converges to some $\varphi_0 \in \mathscr{D}(I)$ iff the two properties of Definition 1.2.10 are satisfied.

1.3 Definition of distributions on I

Equipped with our test functions, we are now able to formally define the notion of distribution on an open interval I. As mentioned in the Introduction, distributions correspond to a certain class of linear forms acting on test functions.

1.3.1 Definitions

Definition 1.3.1 Let $I \subset \mathbb{R}$ an open subset, and T a complex valued linear form on $\mathscr{D}(I)$. One says that T is a distribution on I if, for every segment $[a,b] \subset I$,

$$\exists C > 0,\ \exists m \in \mathbb{N},\ \forall \varphi \in \mathscr{D}_{[a,b]}(I)\,, \quad |T(\varphi)| \leq C \sum_{k=0}^{m} \|\varphi^{(k)}\|_\infty\,. \tag{1.3.1}$$

We denote by $\mathscr{D}'(I)$ the set of distributions on I, and for $T \in \mathscr{D}'(I)$, $\varphi \in \mathscr{D}(I)$, we use the bracket notation

$$\langle T, \varphi \rangle = \langle T, \varphi \rangle_{\mathscr{D}'(I),\mathscr{D}(I)} \stackrel{\text{def}}{=} T(\varphi)\,.$$

From the bound (1.3.1), we say that the distributional bracket $\langle T, \varphi \rangle$ is *controlled by* m derivatives of φ. As we will see later, this number of derivatives necessary to control T may depend on the interval $[a,b]$ on which the test function is controlled.

Definition 1.3.2 [Distribution of finite order] Assume T is a distribution on I. If the integer $m \in \mathbb{N}$ in (1.3.1) can be chosen independently of the interval $[a,b]$, then the distribution T is said to be of order $\leq m$; if m_0 is the smallest such integer m, then T is said to be of order m_0. Notice that the constant C in (1.3.1) may depend on $[a,b]$, though.

The above definition may appear a bit artificial. The following Proposition provides a more natural characterization, in terms of the topology of the space of test functions.

Proposition 1.3.3 (Continuity) A linear form T on $\mathscr{D}(I)$ is a distribution on I if and only if, for any sequence (φ_j) of functions in $\mathscr{D}(I)$ that converges to φ in the $\mathscr{D}(I)$-sense, one has $T(\varphi_j) \to T(\varphi)$.
In other words, the linear form $T : \mathscr{D}(I) \to \mathbb{C}$ is a distribution iff it is *continuous* with respect to the topology of $\mathscr{D}(I)$.

Strictly speaking, we prove the *sequential continuity* of T. On a general topological vector space, sequential continuity of a linear form T may not be equivalent with its continuity. Fortunately, because each of the subspaces $\mathscr{D}_K(I)$ is metrizable, sequential continuity is equivalent with continuity in this inductive topology.
Proof.— Let T be a distribution on I, and (φ_j) a sequence in $\mathscr{D}(I)$ which converges to φ in $\mathscr{D}(I)$.

There is a segment $[a, b] \subset I$ such that supp $\varphi_j \subset [a, b]$ for all $j \in \mathbb{N}$, and supp $\varphi \subset [a, b]$. There exist $C > 0$ and $m \in \mathbb{N}$ such that

$$\forall \psi \in \mathscr{D}(I) \text{ with supp } \psi \subset [a, b], \quad |T(\psi)| \leq C \sum_{k \leq m} \|\psi^{(k)}\|_\infty .$$

In particular, for any $j \in \mathbb{N}$,

$$|T(\varphi_j) - T(\varphi)| = |T(\varphi_j - \varphi)| \leq C \sum_{k \leq m} \|\varphi_j^{(k)} - \varphi^{(k)}\|_\infty .$$

Therefore $T(\varphi_j) \to T(\varphi)$ as $j \to +\infty$, and we have proved the *only if* part of the proposition.

We now want to prove the *if* part. Suppose that for any sequence (φ_j) of functions which converges in $\mathscr{D}(I)$, we have $T(\varphi_j) \to T(\varphi)$, where $\varphi = \mathscr{D} - \lim \varphi_j$. We will reason *ab absurdo*: let us assume that the linear form T is *not* a distribution: the converse of (1.3.1) gives the long statement:

(1.3.2) $$\exists [a, b] \subset I, \ \forall C > 0, \ \forall m \in \mathbb{N}, \ \exists \varphi \in \mathscr{D}_{[a,b]}(\Omega), \quad \text{such that} \quad |T(\varphi)| > C \sum_{k \leq m} \|\varphi^{(k)}\|_\infty .$$

(Notice that φ cannot be the null function). In particular, for any $j \in \mathbb{N}$, choosing $C = m = j$, there is a test function $\varphi_j \in \mathscr{D}_{[a,b]}(I)$ such that

$$|T(\varphi_j)| > j \sum_{k \leq j} \|\varphi_j^{(k)}\|_\infty .$$

Let $\psi_j \in \mathscr{D}_{[a,b]}(I)$ be defined by $\psi_j := \varphi_j / |T(\varphi_j)|$. One obviously has $|T(\psi_j)| = 1$, with supp $\psi_j \subset [a, b]$. On the other hand, for any given $k \in \mathbb{N}$ and any $j \geq k$, one has the bound

$$\|\psi_j^{(k)}\|_\infty \leq \sum_{\beta \leq j} \|\psi_j^{(\beta)}\|_\infty < \frac{1}{j} ,$$

therefore $\lim_{j \to \infty} \|\psi_j^{(k)}\|_\infty = 0$; this proves the limit

$$\mathscr{D} - \lim_{j \to \infty} \psi_j = 0.$$

The continuity of T then imposes $T(\psi_j) \to 0$, which contradicts the normalization $|T(\psi_j)| = 1$. We obtain a contradiction, hence the *if* part of the Proposition is proved. □

Remark 1.3.4 As a set of continuous linear forms, $\mathscr{D}'(I)$ forms a vector space on $\mathbb{C}$: if $T_1, T_2 \in \mathscr{D}'(I)$ and $\lambda_1, \lambda_2 \in \mathbb{C}$, then $\lambda_1 T_1 + \lambda_2 T_2$ is the linear form given by

$$\langle \lambda_1 T_1 + \lambda_2 T_2, \varphi \rangle := \lambda_1 \langle T_1, \varphi \rangle + \lambda_2 \langle T_2, \varphi \rangle ,$$

and this form is obviously continuous since T_1, T_2 are so.

For $T \in \mathscr{D}'(I)$, we may also denote by $\bar{T}$ or T^* the distribution given by

$$\langle \bar{T}, \varphi \rangle = \overline{\langle T, \overline{\varphi} \rangle}.$$

Then, any distribution T can be written as $T = T_1 + iT_2$ where T_1 and T_2 are *real distributions*, that is such that $\langle T, \varphi \rangle \in \mathbb{R}$ for any real valued function φ. Indeed, this relation holds with

$$T_1 = \frac{1}{2}(T + \bar{T}) = \mathrm{Re}(T) \text{ and } T_2 = \frac{1}{2i}(T - \bar{T}) = \mathrm{Im}(T).$$

Equipped with this general definition, we will now provide some examples of distributions.

1.3.2 Distributions defined by locally integrable functions

The first distributions we meet are well-known objects, namely *locally integrable functions*, that is measurable functions f on $\mathbb{R}$, such that for any bounded interval $[a,b]$, $\int_{[a,b]} |f(x)|\,dx < \infty$. We write $f \in L^1_{\text{loc}}(I)$.

For such a function f and any $\varphi \in \mathscr{D}(I)$, the function $f\varphi$ is integrable. More precisely,

$$T_f : \varphi \to \int_I f(x)\varphi(x)\,dx$$

is a linear form on $\mathscr{D}(I)$. Moreover, for a segment $[a,b] \subset I$, for any $\varphi \in \mathscr{D}_{[a,b]}(I)$, one easily checks that

$$|T_f(\varphi)| \leq \|f\|_{L^1([a,b])} \sup |\varphi| \,.$$

This bound shows that T_f is a distribution on I, and that $T_f(\varphi)$ is controlled by 0 derivative of φ: according to the Definition 1.3.2, it is a distribution of order 0.

As a matter of fact, one can identify $L^1_{\text{loc}}(I)$ with a subspace of $\mathscr{D}'(I)$, that is identify the function f with the distribution T_f. This is the content of the next

Proposition 1.3.5 The linear map

$$f \in L^1_{\text{loc}}(I) \longmapsto T_f \in \mathscr{D}'(I)$$

is injective.

Remark 1.3.6 The injectivity refers to elements of L^1_{loc}, which are not functions, but classes of functions, for the equivalence relation that $f \sim g$ if $f(x) = g(x)$ almost everywhere w.r.to the Lebesgue measure on I. In other words, if f is a locally integrable function, we want to show that $T_f = 0$ if and only if $f = 0$ almost everywhere. In the special case where we also assume that f is continuous, this is equivalent to $f = 0$ everywhere (indeed, $f^{-1}(\mathbb{R}^*)$ is then open and Lebesgue-negligible, it is thus the empty set).

Proof.— Assume $T_f = 0$. It is enough to prove that $f = 0$ a.e. on every segment of I, or equivalently that for any $\psi \in \mathscr{D}(I)$ we must have $\psi f = 0$ a.e.

Our first task is to approximate the function f by a smooth function. For this aim, we will use a family of *convolution kernels* [see TD1, Ex.8].

Let $\rho \in \mathscr{D}(\mathbb{R})$, non negative, supported in $[-1,1]$, and such that

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

We then rescale ρ by a factor[3] $\varepsilon > 0$, into the function

$$\rho_\varepsilon(x) = \frac{1}{\varepsilon}\rho\left(\frac{x}{\varepsilon}\right) \,. \tag{1.3.3}$$

3. In general, in analysis the letter ε is used for a "small" parameter.

We notice that the integral of ρ_ε remains equal to unity, while $\rho_\varepsilon \in \mathscr{D}(\mathbb{R})$ is supported in $[-\varepsilon, \varepsilon]$.

Lemma 1.3.7 (Smoothing by convolution) Let $g \in L^1(\mathbb{R})$. Then the convolution product

$$g_\varepsilon = g * \rho_\varepsilon$$

converges in $L^1(\mathbb{R})$ to g when $\varepsilon \to 0$.

The proof of this Lemma is the goal of Ex.8 in TD1.

Let us use this Lemma to prove the Proposition 1.3.5. We have the following expression for the regularization of ψf:

$$\forall x \in \mathbb{R}, \qquad (\psi f) * \rho_\varepsilon(x) = \int f(y)\psi(y)\rho_\varepsilon(x-y)\, dy = \langle T_f, \psi\rho_\varepsilon(x-.)\rangle = 0,$$

where we used the fact that the function $y \mapsto \psi(y)\rho_\varepsilon(x-y)$ belongs to $\mathscr{D}(I)$. Applying the lemma to $g = \psi f \in L^1$, we conclude $\psi f = 0$ in L^1, which completes the proof. □

1.3.3 The Dirac mass

The next example of distribution is a genuinely new object, in the sense that it is not a function (nor an equivalence class of functions). It admits various names: Dirac distribution, Dirac mass, Dirac measure, Dirac delta function,…, all associated with the physicist Paul Dirac, which introduced it in quantum mechanics, and called it "δ".

For a point $x_0 \in I$, we denote by $\delta_{x_0} : \mathscr{D}(I) \to \mathbb{C}$ the linear form defined by

$$\delta_{x_0}(\varphi) = \langle \delta_{x_0}, \varphi\rangle := \varphi(x_0)\,.$$

For any function $\varphi \in \mathscr{D}(I)$, one obviously has

$$|\delta_{x_0}(\varphi)| \leq \|\varphi\|_\infty,$$

so that δ_{x_0} is a distribution on I. Let us call it the *Dirac mass* at x_0, or the Diract distribution at x_0.

We notice that $\delta_{x_0}(\varphi)$ is controlled by 0 derivative of φ, like in the case of distributions T_f associated with an L^1_{loc} function f. It is thus also a distribution of order 0.

Yet, we claim that

Lemma 1.3.8 The Dirac mass at x_0 cannot be defined by a locally integrable function f.

Proof.—Let us assume the opposite, namely that $\delta_{x_0} = T_f$ for some $f \in L^1_{\text{loc}}$. Then, for any function $\varphi \in \mathscr{D}(I)$ such that $x_0 \notin \operatorname{supp}\varphi$,

$$\varphi(x_0) = 0 = \int f(x)\varphi(x)dx\,.$$

Let us split the interval I at the point x_0, to obtain the disjoint union

$$I = I_1 \sqcup \{x_0\} \sqcup I_2 .$$

Then for any $\varphi_1 \in \mathscr{D}(I_1)$ we will have

$$\int_I f(x)\varphi_1(x)\,dx = \int_{I_1} f_{|I_1}(x)\varphi_1(x)\,dx = 0.$$

The second integral is the distributional bracket associated with the function $f_{\restriction I_1} \in L^1_{\text{loc}}(I_1)$. The Proposition 1.3.5 applied to $f_{|I_1}1$, implies that $f_{|I_1} = 0$ a.e. on I_1. The same reasoning applies to the restriction $f_{|I_2}$, which hence vanishes a.e. on I_2. Finally, the full function $f = 0$ a.e. on I.

Now, if we select a plateau function ψ such that $\psi(x_0) = 1$, we would have

$$1 = \psi(x_0) = \langle T_f, \psi\rangle = \int_I f\psi = 0,$$

which is a contradiction. □

The Dirac mass is relevant in many contexts. Among the 4 motivations we presented in the introduction, 3 were involving the Dirac mass (on $\mathbb{R}$, or its generalization on $\mathbb{R}^3$):

i) the derivative of the Heaviside function is the Dirac distribution at the origin (see below section (1.4)).

ii) in electrostatics, a point charge at the origin can be modelized by a "density" given by the Dirac mass at $x_0 \in \mathbb{R}^3$.

iii) in quantum mechanics, we already explained that the quantum state localized at a point x_0 is modelized by the Dirac mass at x_0.

1.3.4 A distribution involving an infinite number of derivatives

In this section we want to explain why, in the bound (1.3.1) in the definition of a distribution, the number m of derivatives needed to control $T(\varphi)$ may depend on the support of φ (hence on the segment $[a, b]$ containing this support). Indeed, we construct below a distribution such that $m = m([a, b])$ converges to infinity when $[a, b]$ is made larger and larger, up to exhausting all of I.

We take $I =]0, 1[$, and consider a sequence $(x_n)_{n\in\mathbb{N}}$ of points in $]0, 1[$ converging to 0 as n tends to infinity. Let $(a_n)_{n\in\mathbb{N}}$ an arbitrary sequence of complex numbers. For every $\varphi \in \mathscr{D}(I)$, we claim that

$$\langle T, \varphi\rangle = \sum_{n=0}^{\infty} a_n \varphi^{(n)}(x_n) \tag{1.3.4}$$

defines a distribution on I. Indeed, let $a < b$ be points in $]0, 1[$ such that supp$(\varphi) \subset [a, b]$. Then there exists an integer N such that, for every $n > N$, $x_n < a$, so that

$$\sum_{n=0}^{\infty} a_n \varphi^{(n)}(x_n) = \sum_{n=0}^{N} a_n \varphi^{(n)}(x_n)$$

is well defined, and controlled by N derivatives of φ. For the same reason, if a sequence $(\varphi_k)_{k\in\mathbb{N}}$ converges to φ in $\mathscr{D}(I)$, then there exists a segment $[a,b]\subset I$ containing the supports of all the φ_k and of φ, so there will exist some N such that

$$\langle T,\varphi_k\rangle = \sum_{n=0}^{N} a_n\varphi_k^{(n)}(x_n) \xrightarrow{k\to\infty} \sum_{n=0}^{N} a_n\varphi^{(n)}(x_n) = \langle T,\varphi\rangle \ .$$

If an infinite set of parameters (a_n) are nonzero, we claim that the distribution T is not of finite order, meaning that one cannot control T with a fixed number of derivatives of φ, independent of the support of φ. This is due to the fact that derivatives $\varphi^{(n)}(x_n)$ cannot be controlled by norms $(\|\varphi^{(k)}\|_\infty)_{k<n}$ involving less derivatives. Following the Definition 1.3.2, we say that T is a distribution of infinite order.

Remark 1.3.9 In this example it is crucial that the sequence $(x_n)_{n\in\mathbb{N}}$ does not admit an accumulation point in I. Indeed, if there were such an accumulation point $x_a\in I$, then for a given choice of parameters (a_n) (assuming infinitely many of these parameters are nonzero), one can construct a test function $\varphi\in\mathscr{D}(I)$ such that the derivatives $\varphi^{(n)}(x_a)$ grow arbitrarily fast when $n\to\infty$, in particular such as to make the sum (1.3.4) divergent. Hence the linear form T would not be defined on this function φ, it would not be a distribution.

1.3.5 Cauchy's principal value of $1/x$

We now introduce a distribution which looks similar to the distributions T_f associated with L^1_{loc} functions. However, in the present case the function is $1/x$, which is not locally integrable at the origin. For this reason, one has to be a bit clever and proceed by a limiting argument, removing smaller and smaller neighbourhoods of the origin. We will see that the resulting distribution is more singular than distributions associated with L^1_{loc} functions.

Precisely, we consider the linear form $T:\mathscr{D}(\mathbb{R})\to\mathbb{C}$ given by

$$T(\varphi) = \lim_{\varepsilon\to0^+}\int_{|x|>\varepsilon}\frac{\varphi(x)}{x}dx.$$

Let us show that this limit exists for any $\varphi\in\mathscr{D}(\mathbb{R})$. Indeed, for such a function, we have

$$\begin{aligned}\int_{|x|>\varepsilon}\frac{\varphi(x)}{x}dx &= \int_{\varepsilon}^{+\infty}\frac{\varphi(x)}{x}\,dx + \int_{-\infty}^{-\varepsilon}\frac{\varphi(x)}{x}\,dx\\ &= \int_{\varepsilon}^{+\infty}\frac{\varphi(x)-\varphi(-x)}{x}\,dx \underset{\varepsilon\to0}{\longrightarrow} \int_{0}^{+\infty}\frac{\varphi(x)-\varphi(-x)}{x}\,dx \ .\end{aligned}$$

Here we have used the fact that the function

$$x\in\mathbb{R}\mapsto\frac{\varphi(x)-\varphi(-x)}{x}$$

is C^∞ and compactly supported (cf Ex. 1.1.1). In fact, both integrals on $\mathbb{R}_+$ and $\mathbb{R}_-$ diverge, but with opposite asymptotics, such that their sum remains bounded.

Now let us show that the well-defined linear form T is a distribution on $\mathbb{R}$. For any $\varphi \in \mathscr{D}(\mathbb{R})$, we have the bound

$$\left|\frac{\varphi(x)-\varphi(-x)}{x}\right| \leq 2\|\varphi'\|_\infty$$

hence, if $\operatorname{supp}\varphi \subset [-A, A]$,

$$\left|\int_0^{+\infty} \frac{\varphi(x)-\varphi(-x)}{x}\,dx\right| \leq 2A\|\varphi'\|_\infty\ .$$

This shows that T satisfies the estimate in Definition 1.3.1. Notice that the constant in the estimate is here $2A$, it does indeed depend on the support of φ. Another striking fact is that this distribution is controlled by the first derivative of φ. In this sense, it is more singular than the distributions T_f associated with locally integrable functions. One can show that this distribution is indeed of order 1.

This distribution is called the *Cauchy principal value*[4] of $1/x$, and we denote it by

$$\langle \mathrm{pv}\left(\frac{1}{x}\right), \varphi\rangle = \lim_{\varepsilon\to 0^+}\int_{|x|>\varepsilon}\frac{\varphi(x)}{x}dx = \int_0^{+\infty}\frac{\varphi(x)-\varphi(-x)}{x}\,dx\ .$$

There are other examples of extensions of functions near a singular point (Hadamard finite parts,…), which can often be seen as a *renormalization* procedure of a divergent expression: one substracts the "simple" divergent behaviour, to identify a bounded remainder, which is called the "renormalized" expression.

1.4 Differentiating distributions

This section introduces one of the most important operation on distributions, namely the derivative. In order to understand the definition below, let us notice that, if $f \in \mathcal{C}^1(I)$, a simple integration by parts yields

$$\forall\varphi \in \mathcal{C}_0^\infty(I), \quad \langle T_{f'}, \varphi\rangle = \int_I f'(x)\varphi(x)dx = -\int_I f(x)\varphi'(x)dx = -\langle T_f, \varphi'\rangle\ .$$

This suggest the following definition.

Proposition 1.4.1 Let $T \in \mathscr{D}'(I)$. The linear form on $\mathscr{D}(I)$ defined by

$$\varphi \mapsto -\langle T, \varphi'\rangle$$

is a distribution, that we call the derivative of T, and that we denote by T'.

Proof.— Let $[a, b] \subset I$ be a segment, and $C > 0$, $m \in \mathbb{N}$ the constants controlling the distribution $T \in \mathscr{D}'(I)$ for test functions supported on $[a, b]$ (see Def. 1.3.1). For $\varphi \in \mathscr{D}(I)$, supported in $[a, b]$,

4. In French, it is called "valeur principale de $1/x$", hence denoted as vp $\left(\frac{1}{x}\right)$.

we have

$$|\langle T',\varphi\rangle| = |\langle T,\varphi'\rangle| \leq C\sum_{k\leq m}\sup|\varphi^{(1+k)}| \leq C\sum_{k\leq m+1}\sup|\varphi^{(k)}|,$$

which shows that T' is a distribution, a little more singular that T: if T is of finite order m, then T' is of order $m+1$. □

Of course the above calculation is compatible with the case of functions:

Proposition 1.4.2 If $T = T_f$ with $f \in \mathcal{C}^1(I)$, then $T' = T_{f'}$.

The next examples are much more interesting.

Example 1.4.3 Let $H:\mathbb{R}\to\mathbb{C}$ denote the Heaviside function, namely $H(x) = \mathbb{1}_{\mathbb{R}_+}(x)$. This function belongs to $L^1_{\mathrm{loc}}(\mathbb{R})$, so we can consider $T = T_H$ the associated distribution. For any $\varphi \in \mathscr{D}(\mathbb{R})$,

$$\langle T',\varphi\rangle = -\langle T,\varphi'\rangle = -\int_0^{+\infty}\varphi'(x)dx = \varphi(0),$$

so that $T'_H = \delta_0$, as we stated in the introduction of these notes: in the sense of distributions, the derivative of the Heaviside function (which is not a differentiable function) is the delta mass at the origin. The non-differentiability (actually, the discontinuity) of H at the origin leads to a singular distribution at 0.

Example 1.4.4 (Primitives of L^1_{loc} functions) Let $f \in L^1_{\mathrm{loc}}(I)$ and $a \in I$. Consider the primitive

$$F(x) = \int_a^x f(t)\,dt\ .$$

A classical application of the dominated convergence theorem shows that F is continuous on I. In particular, F is locally integrable. But in general, F will not be differentiable everywhere. We claim that

Proposition 1.4.5

$$T'_F = T_f\ .$$

Proof.— Introduce, for every $t \in I$,

$$I_{<t} = \{x\in I : x<t\}\ ,\ I_{>t} = \{x\in I : x>t\}\ .$$

Let us compute, for every $\varphi \in \mathcal{C}^\infty_0(I)$,

$$\begin{aligned}\langle T'_F,\varphi\rangle &= -\langle T_F,\varphi'\rangle = -\int_I\left(\int_a^x f(t)\,dt\right)\varphi'(x)\,dx\\ &= \int_{I_{<a}}\left(\int_I \mathbb{1}_{]x,a[}(t)\,f(t)\,dt\right)\varphi'(x)\,dx - \int_{I_{>a}}\left(\int_I \mathbb{1}_{]a,x[}(t)\,f(t)\,dt\right)\varphi'(x)\,dx\ . \qquad (1.4.5)\end{aligned}$$

Let us apply the Fubini theorem to both integrals in the right hand side. Notice that, if the support of φ is included in $[\alpha,\beta]$ with $a\in[\alpha,\beta]$, the integrand of both integrals is supported by $[\alpha,\beta]\times[\alpha,\beta]$ and is bounded by

$$|f(t)||\varphi'(x)|$$

which is integrable on $[\alpha,\beta]\times[\alpha,\beta]$. Hence the Fubini theorem allows us to interchange the orders of integration. This yields

$$\begin{aligned}\langle T_F',\varphi\rangle &= \int_{I_{<a}} f(t)\left(\int_{I_{<t}}\varphi'(x)\,dx\right)dt-\int_{I_{>a}} f(t)\left(\int_{I_{>t}}\varphi'(x)\,dx\right)dt\\ &= \int_{I_{<a}} f(t)\varphi(t)\,dt+\int_{I_{>a}} f(t)\varphi(t)\,dt\\ &= \int_I f(t)\varphi(t)\,dt=\langle T_f,\varphi\rangle\ .\end{aligned}$$

□

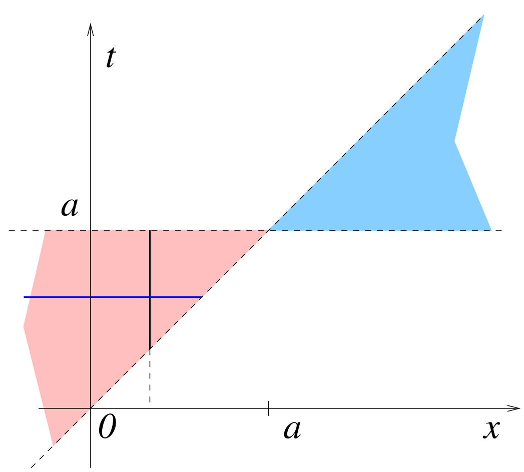


Figure 1.4: Applying Fubini's theorem on the two integrals of (1.4.5); the two domains of integration are drawn in pink, resp. light blue.

In this proof we notice a "meta-strategy" which will be used at many points in this course: transfer the computations (differentiation, product with a smooth function) from the distribution side to the test function side.

Remark 1.4.6 [Going further than Lebesgue !] The statement of Prop. 1.4.5 looks very similar with the one of Prop. 1.4.2. However the extension from $f\in C^1$ to $f\in L^1_{loc}$ contains some subtleties.

1. The derivative f is not defined everywhere, yet the corresponding distribution is unique, as explained in Prop. 1.3.5. On the other hand, the function F is continuous, hence unambiguously defined.

2.Lebesgue proved that the function $F(x)=\int^x f$ is almost everywhere differentiable, with a derivative a.e. equal to f; in particular F can be recovered from its derivative. This is not a general property of continuous functions[5]. Indeed, there exists continuous functions G, differentiable almost everywhere, but such that their (a.e. defined) derivative G' does not allow to reconstruct G. For instance, the "devil's staircase" function G is continuous on $[0,1]$, almost everywhere differentiable with $G'(x)=0$ a.e., yet G is not a constant function (see fig. 1.5). This uncomfortable phenomenon does not happen with distributional derivatives, as we shall see below in Proposition 1.4.8: for such a function G the distributional derivative T'_G is not associated with an L^1_{loc} function, but is more singular; yet it vanishes on a union of intervals of full measure on $]0,1[$.

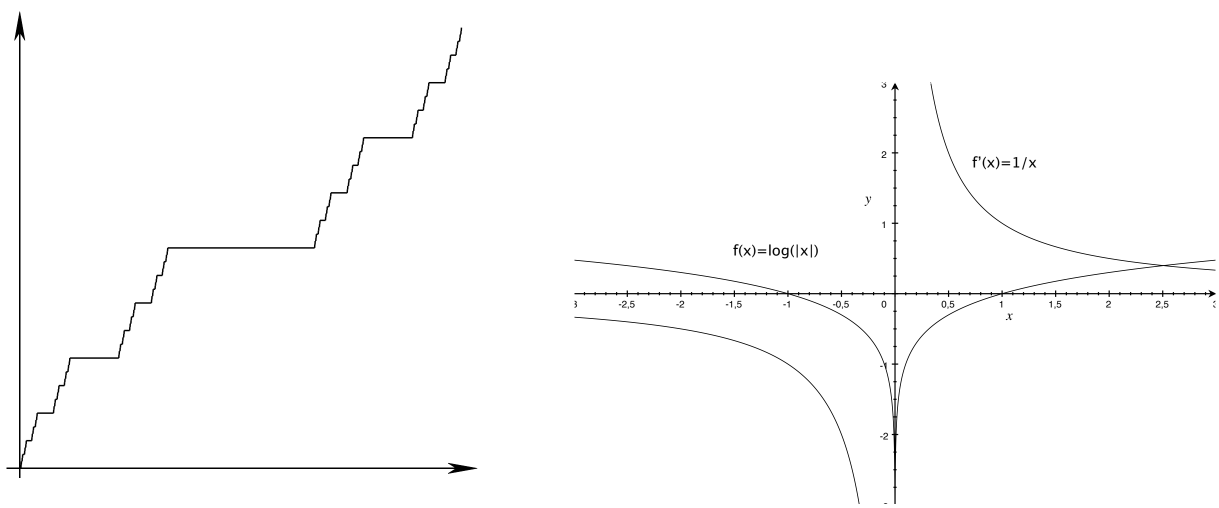


Figure 1.5: Left: the "devil staircase" function G, which satisfies $G'(x)=0$ a.e. Right: $f(x)=\mathrm{ln}(|x|)$ and its derivative.

Example 1.4.7 The function $f(x)=\mathrm{ln}(|x|)$ belongs to $L^1_{\mathrm{loc}}(\mathbb{R})$, thus defines a distribution $T=T_f\in\mathscr{D}'(\mathbb{R})$. We want to compute T'; it seems reasonable that T' is related to the function $x\mapsto 1/x$, but this function is not in $L^1_{\mathrm{loc}}(\mathbb{R})$. However, let $\varphi\in\mathcal{C}_0^\infty([-A,A])$. We compute

$$\begin{aligned}\langle T',\varphi\rangle &= -\int_{\mathbb{R}} \mathrm{ln}\,|x|\;\varphi'(x)\,dx = -\lim_{\varepsilon\to 0}\int_{|x|>\varepsilon}\varphi'(x)\,\mathrm{ln}(|x|)\,dx\\ &= \lim_{\varepsilon\to 0}\int_{|x|>\varepsilon}\frac{\varphi(x)}{x}\,dx+\lim_{\varepsilon\to 0}(\varphi(\varepsilon)-\varphi(-\varepsilon))\,\mathrm{ln}(\varepsilon)\end{aligned}$$

and the second limit in the right hand side vanishes, since φ is smooth. We conclude that

$$T'=\mathrm{pv}(1/x),$$

the distribution introduced in the previous paragraph.

The only functions with an identically null derivative on a full interval are constant functions. The next proposition shows that distributions satisfy the same property.

5. This property actually characterizes *absolute continuity* of the function F, which is a stronger property than continuity

Proposition 1.4.8 If $T \in \mathscr{D}'(I)$ satisfies $T' = 0$, then T is associated with a constant function.

The second point of Remark 1.4.6 shows that this statement is not obvious. Notice that this is the first differential equation we are solving in the space of distributions; it does not yield any "new" solution.

Proof.— We start the proof by a remark that may be useful in other contexts. A function $\varphi \in \mathscr{D}(I)$ is the derivative of a function $\psi \in \mathscr{D}(I)$ if and only if $\int_I \varphi = 0$ (we say that φ "has zero mass", or "is massless"). Indeed, if $\varphi = \psi'$ for ψ supported in $[a, b] \subset I$, then

$$\int_I \varphi(x)dx = \int_a^b \psi'(x)dx = [\psi(x)]_a^b = 0 \; .$$

Conversely, if $\int_I \varphi = 0$, the function $\psi : x \mapsto \int_{I\cap]-\infty,x[} \varphi(t)dt$ is compactly supported, with support included in any compact interval containing $\operatorname{supp}\varphi$, and satisfies $\psi' = \varphi$.

The constraint $T' = 0$ shows that for any function φ of zero mass, one has

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

Therefore, we already know the value of T on all massless functions. We now wish to compute the value of T on arbitrary functions $\varphi \in \mathscr{D}(I)$, including "massive" functions. For this aim, the idea is to split φ into a massless part (which will be killed by T) and a simple massive part.

To operate this splitting, we choose a "reference function" $\chi \in \mathscr{D}(I)$ such that $\int_I \chi = 1$. Using this function, any $\varphi \in \mathscr{D}(I)$ splits into the sum

$$\varphi = \varphi_0 + \varphi_m \;\text{ with }\; \varphi_0 = \varphi - \left(\int_I \varphi\right)\chi, \text{ and } \varphi_m = \left(\int_I \varphi\right)\chi,$$

that is a "massless" part and a "massive" one. Since $\int_I \varphi_0 = 0$, there exists a unique $\psi \in \mathscr{D}(I)$ such that $\varphi_0 = \psi'$. Hence, by linearity we may write

$$\begin{aligned}\langle T, \varphi\rangle &= \langle T, \psi'\rangle + \left(\int_I \varphi\right)\langle T, \chi\rangle \\ &= -\langle T', \psi\rangle + \left(\int_I \varphi\right)\langle T, \chi\rangle \,.\end{aligned}$$

The hypothesis $T' = 0$ yields

$$\langle T, \varphi\rangle = C\int_I \varphi = \langle T_C, \varphi\rangle,$$

where $C = \langle T, \chi\rangle$ is a constant independent of φ, and T_C is the distribution associated with the constant function equal to C. We notice that C is independent of the choice of reference function: indeed, had we chosen a different reference function $\tilde{\chi}$ of mass unity, since $\chi - \tilde{\chi}$ is massless, we would have $\langle T, \chi - \tilde{\chi}\rangle = 0$. □

Proposition 1.4.2 implies that the derivative of a distribution associated with a constant function is null. Therefore, the above statement is actually an equivalence.

This result can be easily extended to a differential equation with a nonhomogeneous term.

Corollary 1.4.9 Let $f \in L^1_{\mathrm{loc}}(I)$. The distributions $T \in \mathscr{D}'(I)$ solving the equation

$$T' = T_f$$

are all of the form $T = T_F$, where

$$F(x) = \int_a^x f(t)\, dt + c \ , \ a \in I \ , \ c \in \mathbb{C} \ .$$

Proof.— Let $a \in I$ and define

$$F_1(x) = \int_a^x f(t)\, dt \ .$$

By Proposition 1.4.5, $T'_{F_1} = T_f = T'$. Then just apply Proposition 1.4.9 to $T - T_{F_1}$. □

Using arguments from the proof of Proposition 1.4.8, we can actually establish the surjectivity of the mapping $T \mapsto T'$ on the whole of $\mathscr{D}'(I)$.

Proposition 1.4.10 (Primitive of a distribution) If $S \in \mathscr{D}'(I)$, there exists $T \in \mathscr{D}'(I)$ such that $T' = S$.

The previous proposition shows that, once we have found a particular solution T to the above equation, all the solutions are of the form $T + T_c$ for some constant $c \in \mathbb{C}\}$.

Proof.— As is becoming usual, the strategy will be to push the computations to the test functions. The identity $T' = S$ is equivalent to

$$\forall \varphi \in \mathscr{D}(I) \ , \quad \langle T, \varphi' \rangle = -\langle S, \varphi \rangle \ . \tag{1.4.6}$$

In view of the proof of Proposition 1.4.8, the above equation already imposes the value of T on all test functions of the form φ', namely on all massless test functions.

Following the same strategy of proof to compute the value of T on a general test function, we split the latter it into a massless and a massive part. So we choose a fixed "reference function" $\chi \in \mathscr{D}(I)$ such that $\int_I \chi\, dx = 1$, and for any test function $\varphi \in \mathscr{D}(I)$ we define its massive part $\varphi_m = \chi \int_I \varphi$, and is massless part $\varphi_0(x) = \varphi(x) - \chi(x) \int_I \varphi$. This massless part admits a primitive

$$P(\varphi)(x) := \int_{-\infty}^x \varphi_0(t)\, dt \ ,$$

which belongs to $\mathscr{D}(I)$. Note that the map $\varphi \mapsto P(\varphi)$ is linear. We may then define the value of T on φ by:

$$\langle T, \varphi \rangle := -\langle S, P(\varphi) \rangle \ . \tag{1.4.7}$$

Let us check that this definition satisfies the required property (1.4.6). If we take $\varphi = \psi'$ in the above formula, we note that this function is massless, so that $\varphi_0 = \varphi$, and $P(\varphi)$ is the unique primitive of φ which is a test function, namely $P(\varphi) = P(\psi') = \psi$; we then find (1.4.6) (with φ replaced by ψ).

There remains to check that the linear form (1.4.7) is a distribution, namely is continuous w.r.t. the topology on $\mathscr{D}(I)$. To show this, it suffices to show that the map $P : \mathscr{D}(I) \to \mathscr{D}(I)$ is continuous. The first property to check is the control on the support of $P(\varphi)$. Let us assume that the reference function χ is supported in an interval $[\alpha, \beta] \subset I$. If φ is supported in the interval $[a, b]$, then the function φ_0 is supported in $[\min(a, \alpha), \max(b, \beta)]$, and it will be the case for its primitive $P(\varphi)$ too. Hence, controlling supp(φ) allows to control supp $P(\varphi)$.

Then, the expression of φ_0 shows that for any $k \in \mathbb{N}$, the norm

$$\|\varphi_0^{(k)}\|_\infty \leq \|\varphi^{(k)}\|_\infty + \|\chi^{(k)}\|_\infty \|\varphi\|_\infty$$
$$C_{\chi,k} \sum_{j=0}^{k} \|\varphi^{(j)}\|_\infty \,,$$

where the constant $C_{\chi,k}$ takes into account the norms on the derivatives of the reference function χ. Taking the primitive of φ_0, we find the same type of estimates:

$$\|P(\varphi)\|_\infty \leq \tilde{C}_{\chi,0} \|\varphi\|_\infty,$$
$$\|P(\varphi)^{(k)}\|_\infty = \|\varphi_0^{(k-1)}\|_\infty \leq C_{\chi,k-1} \sum_{j=0}^{k-1} \|\varphi^{(j)}\|_\infty, \quad k \geq 1.$$

These estimates show that the map $\varphi \mapsto P(\varphi)$ is (sequentially) continuous: if a sequence (φ_n) converges to φ in $\mathscr{D}(I)$, then $P(\varphi_n)$ converges to $P(\varphi)$ in $\mathscr{D}(I)$. Therefore the linear form $T = -S \circ P : \mathscr{D}(I) \to \mathbb{C}$ is itself continuous, hence it is a distribution. □

1.4.1 Higher derivatives*

Since T' is itself a distribution, it can be differentiated. Iterating this process, one can define the successive derivatives $T^{(k)}$ of T for all $k \in \mathbb{N}$. This means that any distribution can be differentiated at any order, and we obtain the formula

$$\forall \varphi \in \mathscr{D}(I), \qquad \langle T^{(k)}, \varphi \rangle = (-1)^k \langle T, \varphi^{(k)} \rangle.$$

For instance, the k-th derivative of the Dirac mass at the origin is defined as

$$\langle \delta_{x_0}^{(k)}, \varphi \rangle = (-1)^k \varphi^{(k)}(x_0) \,.$$

For instance, the distribution described in subsection 1.3.4 can be expressed as the infinite sum

$$T = \sum_{n \geq 1} (-1)^n \, a_n \, \delta_{x_n}^{(n)} \,.$$

1.5 Product of a distribution with a smooth function

One important property of continuous (or C^k) functions is that they form an *algebra*: one can multiply functions with one another, staying the same class of continuous (or C^k) functions. This is already not the case for two L^1 functions, yet the product of an L^1 function with a bounded (L^∞) function remains in L^1.

Two distributions can in general NOT be multiplied with one another. However, we may multiply a distribution with smooth functions, as shown in the following

Proposition 1.5.1 Let $T \in \mathscr{D}'(I)$, and $f \in \mathcal{C}^\infty(I)$. The linear form

$$\varphi \mapsto \langle T, f\varphi \rangle$$

is a distribution in $\mathscr{D}'(I)$. We denote it by fT.

Proof.— Suppose (φ_j) is a sequence of functions in $\mathscr{D}(I)$, that converges to 0 in the $\mathscr{D}(I)$-sense. There is a segment $[a,b] \subset I$ such that supp $\varphi_j \subset [a,b]$ for all j, which implies supp $f\varphi_j \subset [a,b]$ for all j. Moreover, for any $\alpha \in \mathbb{N}$,

$$(f\varphi)^{(\alpha)} = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} f^{(\beta)}\, \varphi^{(\alpha-\beta)}.$$

so that if we denote

$$M = \max_{\beta \leq \alpha} \sup_{x \in [a,b]} |f^{(\beta)}(x)| := \|f\|_{C^\alpha([a,b])},$$

we see that

$$\|(f\varphi_j)^{(\alpha)}\|_\infty \leq M \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \|\varphi_j^{(\alpha-\beta)}\|_\infty.$$

Since each of the terms in the sum converges to 0 as $j \to +\infty$, the right hand side does so too, so that $(f\varphi_j) \to 0$ in $\mathscr{D}(I)$. Finally, since T is a distribution, $\langle T, f\varphi_j \rangle \to 0$. This proves that fT is continuous on $\mathscr{D}(I)$, hence is a distribution on I. □

Exercise 1.5.2 For $f \in \mathcal{C}^\infty(\mathbb{R})$, show that $f\delta_0 = f(0)\delta_0$.

Exercise 1.5.3 For $f \in \mathcal{C}^\infty(\mathbb{R})$ and $g \in L^1_{\mathrm{loc}}(\mathbb{R})$, show that $fT_g = T_{fg}$.

Exercise 1.5.4 Show that $x\,\mathrm{pv}(1/x) = T_1$, the distribution associated with the constant function on $\mathbb{R}$ (see TD 2, ex.3)

Exercise 1.5.5 For $f, g \in \mathcal{C}^\infty(\mathbb{R})$, show that $f(gT) = (fg)T$.

We have seen that distributions can be differentiated arbitrarily many times; so can smooth functions. In view of the above proposition, it makes sense to differentiate the product fT. Like for the product of two smooth functions, the derivatives of this product satisfy a Leibniz formula (we leave the proof of the following result to the reader).

Proposition 1.5.6 (Leibniz formula) For $f \in \mathcal{C}^\infty(I)$, and $T \in \mathscr{D}'(I)$, we have

$$\forall k \in \mathbb{N}, \quad (fT)^{(k)} = \sum_{j=0}^{k} \binom{k}{j} f^{(j)}\, T^{(k-j)}.$$

Proof.— Let us check the property for $k = 1$. By definition, for any $\varphi \in \mathscr{D}(I)$,

$$\langle (fT)', \varphi \rangle = -\langle fT, \varphi' \rangle = -\langle T, f\varphi' \rangle$$

On the other hand,

$$\langle f'T + fT', \varphi \rangle = \langle T, f'\varphi \rangle + \langle T', f\varphi \rangle = \langle T, f'\varphi \rangle - \langle T, (f\varphi)' \rangle = -\langle T, f\varphi' \rangle.$$

Since the two expressions coincide for any φ, we get $(fT)' = fT' + f'T$. Since this algebra of derivatives is identical to the one for the product of two functions, it can be iterated to any order k, leading to the same Leibniz rule as for the product of functions. □

Product with a polynomial

Let us now focus on the case where the function f is a polynomial, or even a monomial. The Exercise 1.5.2 implies the identity

$$(x - x_0)\delta_{x_0} = 0 \;,$$

Our next statement shows that the delta mass at x_0 is essentially the only distribution satisfying this identity.

Proposition 1.5.7 Let $x_0 \in I$, and let $T \in \mathscr{D}'(I)$ satisfy

$$(x - x_0)T = 0 \;,$$

Then there exists a constant $c \in \mathbb{C}$ such that $T = c\delta_{x_0}$.

Proof.— Hadamard's Lemma (see Ex. 1.1.1) shows that, for any test function φ vanishing at x_0, the function $\psi = \frac{\varphi}{(x-x_0)} \in \mathscr{D}(I)$. In that case, if T satisfies $(x - x_0)T = 0$, we have

$$\langle T, \varphi \rangle = \langle T, (x - x_0)\psi \rangle = \langle (x - x_0)T, \psi \rangle = 0\,.$$

Hence, we already know that such a distribution T kills all functions φ vanishing at x_0.

Like in Prop. 1.4.8, to compute T for a general test function φ, we split the latter in a part vanishing at x_0, and a "simple" remaining part. We will also use a reference function $\chi \in \mathscr{D}(I)$, which we now require to satisfy $\chi(x_0) = 1$.

The generic test function $\varphi \in \mathscr{D}(I)$ splits into a multiple of this reference function, and a term vanishing at x_0:

$$\varphi(x) = \varphi(x_0)\chi(x) + (x - x_0)\theta(x)\,,$$

where $\theta(x) = \frac{\varphi(x)-\varphi(x_0)\chi(x)}{x-x_0}$ is smooth (from Hadamard's Lemma), and is compactly supported since the numerator is so. We then infer

$$\langle T, \varphi\rangle = \langle T, \chi\rangle\varphi(x_0) + \langle (x - x_0)T, \theta\rangle = \langle T, \chi\rangle\varphi(x_0)\ ,$$

whence $T = c\delta_{x_0}$ with $c := \langle T, \chi\rangle$. Here as before, changing the reference function will not change the constant c. □

Example 1.5.8 We are now in measure to solve another first order differential equation on the space of distributions, more complicated than the equation $T' = S$ of the previous section. Let us look for solutions $T \in \mathscr{D}'(\mathbb{R})$ of the equation

$$xT' + T = 0\ . \tag{1.5.8}$$

By the Leibniz formula, this equation is equivalent to

$$(xT)' = 0\ .$$

Using Proposition 1.4.8, this is equivalent to

$$xT = T_{c_1} \quad \text{for some } c_1 \in \mathbb{C}\,.$$

Exercise 1.5.4 shows that $c_1\,\mathsf{pv}(1/x)$ is a solution of this equation, so we now have to solve

$$x\Big(T - c_1\,\mathsf{pv}\,(\frac{1}{x})\Big) = 0\ .$$

Finally, using Proposition 1.5.7, we obtain the general solution of the differential equation $xT'+T = 0$:

$$T = c_1\,\mathsf{pv}\,(\frac{1}{x}) + c_2\delta_0\ ,\quad (c_1, c_2) \in \mathbb{C}^2\ .$$

We have proved that the only global nontrivial solutions to (1.5.8) are genuine distributions which present singularies at the point $x = 0$; this singularity is due to the fact that the coefficient of the highest derivative (here T') vanishes at that point.

We also notice that the space of solutions has dimension 2, eventhough the ODE is of order 1.

Finally, we use elements of the proof of Prop. 1.4.10 to show the surjectivity of the mapping $T \mapsto (x - x_0)T$ on $\mathscr{D}'(I)$.

Proposition 1.5.9 Given $S \in \mathscr{D}'(I)$ and $x_0 \in I$, there exists $T \in \mathscr{D}'(I)$ such that

$$(x - x_0)T = S\ .$$

Of course, if $x_0 \notin I$, the statement trivially holds, since the function $(x-x_0)^{-1}$ is smooth on I, and $(x-x_0)^{-1}S$ is well-defined.

Proof.— We proceed similarly to the proof of Proposition 1.4.10. Indeed, $(x-x_0)T = S$ reads

$$\forall \varphi \in \mathscr{D}(I)\ , \quad \langle T, (x-x_0)\varphi\rangle = \langle S, \varphi\rangle\ , \tag{1.5.9}$$

which, in view of Hadamard's lemma, imposes the value of T on test functions vanishing at x_0. As in the proof of Proposition 1.5.7, we fix a reference function $\chi \in \mathscr{D}(I)$ such that $\chi(x_0) = 1$. For every $\psi \in \mathscr{D}(I)$, we split ψ into $\psi = \psi(x_0)\chi + (x-x_0)\theta(x)$, but we now want to remember that the function θ depends on φ, so we denote this function by $Q(\psi)$:

$$Q(\psi)(x) = \frac{\psi(x) - \psi(x_0)\chi(x)}{x - x_0} = \int_0^1 \Big[\psi'\big(tx + (1-t)x_0\big) - \psi(x_0)\chi'\big((tx + (1-t)x_0)\big)\Big]\, dt\ .$$

$Q : \mathscr{D}(I) \to \mathscr{D}(I)$ defines a linear mapping. Like in Prop. 1.4.10, we need to check that this mapping is continuous on $\mathscr{D}(I)$. Firstly, the inclusion

$$\text{supp}\, Q(\psi) \subset \text{supp}(\psi) \cup \text{supp}(\chi)$$

shows that controlling the support of φ allows to control the support of $Q(\varphi)$. Secondly, any derivative $Q(\varphi)^{(k)}$ is controlled by a certain number of derivatives of φ. Indeed, the integral representation of $Q(\varphi)$ shows that

$$\|Q(\psi)\|_\infty \leq \|\psi'\|_\infty + \|\chi'\|_\infty \|\varphi\|_\infty,$$

and similarly for higher derivatives of $Q(\psi)$. These properties show that Q is a continuous mapping.

Now, let us define the linear form T by

$$\forall \psi \in \mathscr{D}(I)\ , \quad \langle T, \psi\rangle = \langle S, Q(\psi)\rangle\ .$$

The continuity of Q and the fact that S is a distribution imply that $T = S \circ Q$ is a distribution as well. Finally, For every $\varphi \in \mathscr{D}(I)$, one checks that

$$Q((x-x_0)\varphi) = \varphi\ ,$$

hence the distribution T satisfies (1.5.9). □

Similarly, the differential equation of Example 1.5.8 can also be made nonhomogeneous.

Exercise 1.5.10 For a given $S \in \mathscr{D}'(\mathbb{R})$, describe the solutions $T \in \mathscr{D}'(\mathbb{R})$ of the equation

$$xT' + T = S\ .$$

To generalize Proposition 1.5.9, we may as well multiply T' by higher order monomials.

Corollary 1.5.11 If $x_0 \in I$ and m is a positive integer, the mapping $T \in \mathscr{D}'(I) \mapsto (x-x_0)^m T \in \mathscr{D}'(I)$ is surjective, and its kernel is the m-dimensional vector space generated by $\{\delta_{x_0}^{(j)},\ j = 0, 1, \dots, m-1\}$.

Proof.— We proceed by induction on m. The case $m = 1$ has been solved in Propositions 1.5.7 and 1.5.9. Assume $m \geq 2$ and that the statement is true for $m - 1$. The surjectivity of the mapping $T \in \mathscr{D}'(I) \mapsto (x - x_0)^m T \in \mathscr{D}'(I)$ follows immediately from the similar property at the level $m - 1$ and from Proposition 1.5.9.

Concerning the kernel, the identity

$$(x - x_0)^m T = (x - x_0)^{m-1}\big((x - x_0)T\big)$$

is solved by the induction hypothesis by

$$(x - x_0)T = \sum_{j=0}^{m-2} c_j \delta_{x_0}^{(j)}, \quad \text{with arbitrary coefficients } c_j \in \mathbb{C}.$$

Prop. 1.5.9 claims that this has a solution. We can actually find an explicit solution to this equation. For this, we will use the following lemma, which generalizes Prop. 1.5.7:

Lemma 1.5.12 For any point $x_0 \in I$ and any integer $k \geq 1$,

$$(x - x_0)\delta_{x_0}^{(k)} = -k\delta_{x_0}^{(k-1)}.$$

Indeed, for every test function φ, we have

$$\langle (x - x_0)\delta_{x_0}^{(k)}, \varphi\rangle = (-1)^k \big((x - x_0)\varphi\big)^{(k)}(x_0) = (-1)^k k\varphi^{(k-1)}(x_0) = -k\langle \delta_{x_0}^{(k-1)}, \varphi\rangle.$$

Coming back to the proof of Prop. 1.5.11, we may rewrite

$$(x - x_0)T = -(x - x_0)\sum_{j=0}^{m-2} \frac{c_j}{j+1}\delta_{x_0}^{(j+1)}],.$$

Finally, applying Prop. 1.5.7 to the distribution

$$T + \sum_{j=0}^{m-2} \frac{c_j}{j+1}\delta_{x_0}^{(j+1)}$$

adds a possible term $c\delta_{x_0}$, which completes the proof. □

After multiplying T by monomials, we may generalize to products with polynomials (we leave the proof to the reader).

Corollary 1.5.13 Let P be a non identically zero polynomial function. Then the mapping

$$T \in \mathscr{D}'(I) \mapsto PT \in \mathscr{D}'(I)$$

is surjective, and its kernel is the vector space generated by the distributions $\delta_a^{(j)}$, where a ranges over the zeroes of P inside I, and $j = 0, 1, \ldots, m(a) - 1$, where $m(a)$ denotes the multiplicity of a as a zero of P.

Finally, let us solve another (very simple) differential equation on $\mathscr{D}'(\mathbb{R})$.

Proposition 1.5.14 Let $I \subset \mathbb{R}$ be an open interval, and $a \in \mathcal{C}^\infty(I)$. The distributions in $\mathscr{D}'(I)$ that satisfy the differential equation

$$T' + aT = 0$$

are exactly the $\mathcal{C}^\infty$ solutions, that is the regular distributions T_f with $f : x \mapsto Ce^{-A(x)}$, for some constant $C \in \mathbb{C}$, where A is a primitive of a in I.

In other words, the distribution solutions of this differential equation coincide with the classical solutions. Notice that this is not always the case if the coefficient of the highest derivative vanishes at some poines — see example 1.5.8 above.

Proof.— Let A be a primitive of a on I. For $T \in \mathscr{D}'(I)$, we have, using Leibniz formula,

$$(e^A T)' = ae^A T + e^A T' = e^A(T' + aT).$$

Thus

$$T' + aT = 0 \iff (e^A T)' = 0 \iff e^A T = T_C \iff T = e^{-A} T_C = T_{Ce^{-A}}.$$

□

Exercise 1.5.15 Solve in $\mathscr{D}'(I)$ the inhomogeneous equation $T' + aT = f$, for $f \in L^1_{\text{loc}}(I)$. Furthermore, prove that, if $S \in \mathscr{D}'(I)$, there exists $T \in \mathscr{D}'(I)$ such that $T' + aT = S$.

1.6 Restriction and support

We have defined the support of a continuous function. For locally integrable functions, there exists a similar notion, called *essential support*, which needs a bit more care to define. Restrictions of functions are easy to define. Below we will define the support of an arbitrary distribution, as well as its restriction on an subinterval. These notions allow to "localize" the action of a distribution on a small interval; more importantly, they allow to construct a global distribution piece by piece, by gluing together locally defined pieces.

1.6.1 Restricting a distribution on a subinterval

Definition 1.6.1 Let $T \in \mathscr{D}'(I)$, and $J \subset I$ an open subinterval of I. The restriction of T to J is the distribution $T_{|J} \in \mathscr{D}'(J)$ defined as

$$\forall \varphi \in \mathscr{D}(J)\ ,\ \langle T_{|J}, \varphi \rangle = \langle T, \underline{\varphi} \rangle\ ,$$

where $\underline{\varphi}$ denotes the extension of φ by 0 on $I \setminus J$.

If $T_{|J} = 0$ we say that T vanishes in J.

Let us state without proof the following elementary properties of restriction with respect to derivation and multiplication by a smooth function.

$$(T_{|J})' = T'_{|J}\,, \qquad (fT)_{|J} = f_{|J}T_{|J}\ .$$

1.6.2 Nested construction of a distribution

As a first application of this important notion of support, let us state a very useful result, namely the possibility to construct a distribution piece by piece, on larger and larger intervals.

Proposition 1.6.2 [Nested distributions] Let $(I_n)_{n\in\mathbb{N}}$ be a growing sequence of bounded intervals in I, such that $I_n =]a_n, b_n[$, $[a_n, b_n] \subset I_{n+1}$ for every n, and

$$\bigcup_{n\in\mathbb{N}} I_n = I\ .$$

Suppose we are given, for every $n \in \mathbb{N}$, a distribution T_n on I_n, such that T_{n+1} extends T_n:

$$\forall n \in \mathbb{N}\ ,\ T_{n+1|I_n} = T_n\ .$$

Then there exists a unique distribution T on I which extends all the T_n, that is such that

$$\forall n \in \mathbb{N}\ ,\ T_{|I_n} = T_n\ .$$

Proof.— The key of the proof is the following elementary fact. For every segment $[a, b] \subset I$, there exists n such that $[a, b] \subset I_n$. Indeed, from the assumptions, the sequence (a_n) is decreasing, converging to the left boundary of I, while the sequence (b_n) is increasing, converging to the right boundary of I. This fact is also a consequence of the Borel–Lebesgue property for the compact set $[a, b]$. The bounded intervals (I_n) are said to form a *nested* (or *growing*) sequence, and their union cover all of I.

If such a distribution T exists, for every $\varphi \in \mathscr{D}(I)$, there exists n such that supp$(\varphi) \subset I_n$, and therefore we should have

$$\langle T, \varphi\rangle = \langle T_n, \varphi\rangle\ .$$

This proves the uniqueness of T, and suggests a way of constructing it. Indeed, given $\varphi \in \mathscr{D}(I)$, we may take n_0 to be the smallest n such that supp$(\varphi) \subset I_n$. The assumption $T_{n+1|I_n} = T_n$ then implies that

$$\langle T_n, \varphi\rangle = \langle T_{n_0}, \varphi\rangle \quad \text{for all } n \geq n_0\,.$$

For this function φ, we can therefore define

$$\langle T, \varphi\rangle \stackrel{\text{def}}{=} \langle T_{n_0}, \varphi\rangle\,.$$

This procedure clearly defines a linear form on $\mathscr{D}(I)$. Is it continuous on $\mathscr{D}(I)$? If φ_j converges to φ in $\mathscr{D}(I)$, then there exist n_0 such that,

$$\forall j\ ,\ \text{supp}(\varphi_j) \subset I_{n_0}\ .$$

Therefore, for all the functions φ_j we take

$$\langle T, \varphi_j \rangle = \langle T_{n_0}, \varphi_j \rangle,$$

and since T_{n_0} is a distribution on I_{n_0}, $\langle T, \varphi_j \rangle$ converges to $\langle T, \varphi \rangle$, showing that T is a distribution as well. □

Let us state a first consequence of this nested construction.

Corollary 1.6.3 Let $f \in \mathcal{C}^\infty(I)$ admitting only finite order zeroes in I, namely, if $f(a) = 0$, there exists $m \geq 1$ such that $f^{(m)}(a) \neq 0$. Then the mapping $T \in \mathscr{D}'(I) \mapsto fT \in \mathscr{D}'(I)$ is surjective.

This result generalizes the product with polynomials of Corollary 1.5.13, to the case of a smooth function f with possibly infinitely many zeros on I; yet, a crucial point is the fact that on any segment $[a, b] \subset I$, the number of zeros of f is at most finite.

Proof.— If $f(a) = 0$ and m is the smallest integer such that $f^{(m)}(a) \neq 0$, the Taylor formula yields

$$f(x) = \frac{(x-a)^m}{(m-1)!} \int_0^1 (1-t)^{m-1} f^{(m)}(a + t(x-a))\, dt\ .$$

From the continuity of $f^{(m)}$, we infer that there exists an open interval $J_a \subset I$ containing a, such that $f(x) \neq 0$ for every $x \in J_a \setminus \{a\}$. For any segment $[\alpha, \beta] \subset I$, the (compact) set of zeroes of f in $[\alpha, \beta]$ is covered by the family of open intervals J_a's, hence — from the Borel-Lebesque property — by a *finite* subfamily of J_a's, showing that this set of zeroes is finite. Consequently, in $]\alpha, \beta[$ we can write

$$f = Pg\ ,$$

where P is a polynomial vanishing on the zeroes of f (with the same multiplicities), and g is a smooth function which does not vanish in $]\alpha, \beta[$. Consequently, the Corollary 1.5.13 shows that the mapping

$$T \in \mathscr{D}'(]\alpha, \beta[) \mapsto fT = PgT \in \mathscr{D}'(]\alpha, \beta[)$$

is surjective, and its kernel consists of finite linear combinations of Dirac masses and their derivatives on the zeroes of f in $]\alpha, \beta[$. In particular, any distribution $U_{]\alpha,\beta[}$ in this kernel is the restriction to $]\alpha, \beta[$ of some distribution U on I, satisfying $fU = 0$ on I.

Let $S \in \mathscr{D}'(I)$. In order to solve the equation $fT = S$, we will now apply the above local construction to a nested cover $(I_n)_{n \geq 0}$ for the interval I, using Proposition 1.6.2.

We start with the two first intervals $I_0 \Subset I_1 \Subset I$. In view of the above local construction, there exists $T_0 \in \mathscr{D}'(I_0)$ such that $f_{|I_0} T_0 = S_{|I_0}$; similarly, there exists $\tilde{T}_1 \in \mathscr{D}'(I_1)$ such that $f_{|I_1} \tilde{T}_1 = S_{|I_1}$. A priori, the restriction $\tilde{T}_{1|I_0}$ is not necessarily equal to T_0, we must apply some correction to make them "match".

Denote by $\tilde{T}_{1,0} = \tilde{T}_{1|I_0}$ the restriction of $\tilde{T}_1$ to I_0. Of course,

$$f_{|I_0}(\tilde{T}_{1,0} - T_0) = 0\ ,$$

From the above observation on the kernel of $T \mapsto fT$, there exists $U_1 \in \mathscr{D}'(I_1)$ such that $U_{1|I_0} = \tilde{T}_{1,0} - T_0$ and $f_{|I_1} U_1 = 0$. If we modify $\tilde{T}_1$ into

$$T_1 \stackrel{\text{def}}{=} \tilde{T}_1 - U_1 \in \mathscr{D}'(I_1) ,$$

we observe that we have restored the "matching" between T_1 and T_0:

$$T_{1|I_0} = T_0 , \quad \text{and} \quad f_{|I_1} T_1 = S_{|I_1} .$$

By induction, we can therefore construct, for every $n \in \mathbb{N}$, a distribution T_n on I_n such that

$$T_{n+1|I_n} = T_n , \quad \text{and} \quad f_{|I_n} T_n = S_{|I_n} .$$

Applying Proposition 1.6.2, there exists $T \in \mathscr{D}'(I)$ such that, for every n, $T_{|I_n} = T_n$. Consequently,

$$\forall n \in \mathbb{N} , (fT)_{|I_n} = f_{|I_n} T_n = S_{|I_n} .$$

By the uniqueness part of Proposition 1.6.2, we have constructed this way a unique distribution T, which globally satisfies

$$fT = S .$$

□ Notice that, as opposed to the case of polynomials, the function f may have infinitely (but countably) many zeroes on I.

1.6.3 Support of a distribution

In Definition 1.1.2 we had introduced the support of a continuous function as the complement of the open set where the function vanishes. We can now adapt this notion to a distribution.

Definition 1.6.4 The support of a distribution $T \in \mathscr{D}'(I)$ is the complement of the union of all the open subintervals on which the restriction of T is 0. We denote this set by $\operatorname{supp} T$.

Like in the case of continuous functions, $\operatorname{supp} T$ is closed, and the following characterizations are convenient.

- $x_0 \notin \operatorname{supp} T$ if and only if there is an open neighborhood J of x_0 such that $T_{|J} = 0$.
- $x_0 \in \operatorname{supp} T$ if and only if, for any open neighborhood J_{x_0} of x_0, one can find $\varphi \in \mathcal{C}_0^\infty(J_{x_0})$ such that $\langle T, \varphi \rangle \neq 0$.

Example 1.6.5 *i)* Let $T = \delta_0$. If J is an open interval that does not contain $\{0\}$, then $\langle T, \varphi \rangle = \varphi(0) = 0$ for any $\varphi \in \mathscr{D}(\mathbb{R})$ such that $\operatorname{supp} \varphi \subset J$. Thus $\operatorname{supp} T \subset \{0\}$. On the other hand, if J is an open subinterval that contains 0, we can find a function $\psi \mathscr{D}(J)$ sucht that $\psi(0) = 1$; for this function we have $\langle T, \psi \rangle = \psi(0) = 1$. Therefore, $0 \in \operatorname{supp} T$, and finally $\operatorname{supp} T = \{0\}$.

ii) If $T = T_f$ for some $f \in \mathcal{C}^0(I)$, with I an open interval of $\mathbb{R}$, we have

$$\operatorname{supp} T = \operatorname{supp} f = \overline{\{x \in I, f(x) \neq 0\}} \ .$$

Indeed, suppose that $x_0 \notin \operatorname{supp} f$. There is an open neighborhood J of x_0 such that $f_{|J} = 0$. For any $\varphi \in \mathscr{D}(J)$, we have thus $\langle T_f, \varphi\rangle = 0$, so that T_f vanishes on J, and $x_0 \notin \operatorname{supp} T_f$. Conversely, if $x_0 \notin \operatorname{supp} T_f$, there is a neighborhood J of x_0 such that, for all $\varphi \in \mathcal{C}_0^\infty(J)$, we have $\int f\varphi dx = \langle T_f, \varphi\rangle = 0$. We have seen in Proposition 1.3.5 that this implies $f = 0$ in J, thus $x_0 \notin \operatorname{supp} f$.

iii) If $T = T_f$ for some $f \in L^1_{\mathrm{loc}}(I)$, the same argument leads to the characterization

$$\operatorname{supp}(T_f) = \operatorname{ess-supp}(f) = \{x_0 \in I, \forall r > 0, \operatorname{meas}\{x \in [x_0 - r, x_0 + r], f(x) \neq 0\} > 0\} \ .$$

1.6.4 Some properties of the support

Let us consider how the support behaves w.r.t. the two operations on distributions, namely differentiation and multiplication by a smooth function.

Lemma 1.6.6 For any $T \in \mathscr{D}'(I)$:

i) for any $k \in \mathbb{N}$, $\operatorname{supp} T^{(k)} \subset \operatorname{supp} T$.

ii) for any $f \in \mathcal{C}^\infty(I)$, $\operatorname{supp} fT \subset \operatorname{supp} f \cap \operatorname{supp} T$.

Proof.— Let $x_0 \notin \operatorname{supp} T$. There exists a neighborhood V of x_0 such that for all $\psi \in \mathcal{C}_0^\infty(V)$, $\langle T, \psi\rangle = 0$. But if $\varphi \in \mathcal{C}_0^\infty(V)$, $\psi := \varphi^{(k)}$ belongs to $\mathcal{C}_0^\infty(V)$ as well, thus

$$\langle T^{(k)}, \varphi\rangle = (-1)^k \langle T, \varphi^{(k)}\rangle = 0.$$

This proves that $T^{(k)}_{|V} = 0$, hence $x_0 \notin \operatorname{supp} T^{(k)}$, which proves *i)*.

Let us prove the second point. If $x_0 \notin \operatorname{supp} f \cap \operatorname{supp} T$, x_0 either belongs to $\complement \operatorname{supp} f$ or to $\complement \operatorname{supp} T$. In the first case, there exists a neighborhood V of x_0 such that $f_{|V} = 0$. For $\varphi \in \mathcal{C}_0^\infty(V)$, we have $\langle fT, \varphi\rangle = \langle T, f\varphi\rangle = 0$, thus $x_0 \notin \operatorname{supp}(fT)$. In the second case, there exists a neighborhood V of x_0 such that, for all $\psi \in \mathcal{C}_0^\infty(V)$, $\langle T, \psi\rangle = 0$. For $\varphi \in \mathcal{C}_0^\infty(V)$, $f\varphi$ also belongs to $\mathcal{C}_0^\infty(V)$, thus $\langle fT, \varphi\rangle = \langle T, f\varphi\rangle = 0$. In both cases, we have shown that $x_0 \notin \operatorname{supp} fT$, which proves the point $ii)$. □

The following result is now easy to show, yet it is fundamental.

Proposition 1.6.7 Let $\varphi \in \mathscr{D}(I)$ and $T \in \mathscr{D}'(I)$. If $\operatorname{supp} \varphi \cap \operatorname{supp} T = \emptyset$, then $\langle T, \varphi\rangle = 0$.

Proof.— Let $x \in \operatorname{supp}\varphi$. We have, by assumption, $x \notin \operatorname{supp} T$, thus there is an open subinterval J_x containing x, on which T vanishes: $T_{J_x} = 0$. Covering the compact subset supp(φ) with the open intervals J_x and applying the Borel-Lebesgue theorem, one can extract a finite covering

$$\operatorname{supp}\varphi \subset \bigcup_{j=1}^{n} J_{x_j}.$$

By Corollary 1.2.7, one can find a smooth partition of φ, namely functions $\varphi_j \in \mathscr{D}(J_{x_j})$ for $j = 1, \ldots, n$, such that

$$\varphi = \varphi_1 + \cdots + \varphi_n\,.$$

By linearity

$$\langle T, \varphi\rangle = \sum_{j=1}^{n} \langle T, \varphi_j\rangle = \sum_{j=1}^{n} \langle T_{|J_{x_j}}, \varphi_j\rangle = 0\ .$$

□

A word of caution: one may have $\varphi = 0$ on $\operatorname{supp} T$ and $\langle T, \varphi\rangle \neq 0$. For example, this is the case for $T = \delta_0'$ and $\varphi \in \mathscr{D}(\mathbb{R})$ such that $\varphi(0) = 0$, $\varphi'(0) = 1$.

An immediate consequence of Proposition 1.6.7 is the following important

Corollary 1.6.8 The only distribution with empty support is the null distribution $T = 0$.

We conclude this section by the important characterisation of distributions $T \in \mathscr{D}'(I)$ supported on a single point.

Proposition 1.6.9 Let $T \in \mathscr{D}'(I)$, with $T \neq 0$, and $x_0 \in I$. If $\operatorname{supp} T \subset \{x_0\}$, there exists $m \in \mathbb{N}$ and $m+1$ complex numbers a_k for $0 \leq k \leq m$ such that

$$T = \sum_{k=0}^{m} a_k \delta_{x_0}^{(k)}.$$

Proof.— We write the proof in the case $x_0 = 0$. Let $a < a' < 0 < b' < b$ such that the segment $[a, b] \subset I$, and $\chi \in \mathscr{D}(]a, b[)$ a cutoff function equal to unity on $[a', b']$. Since T is a distribution, there exist $C > 0$, $m \in \mathbb{N}$ such that

$$\forall \psi \in \mathscr{D}(I), \quad \operatorname{supp}\psi \subset [a, b] \Longrightarrow |\langle T, \psi\rangle| \leq\ C \sum_{k=0}^{m} \sup|\psi^{(k)}|,$$

and from now on we denote by m the smallest integer for which this property holds[6].

6. This value m is therefore the order of the distribution $T_{|]a,b[}$

For any function $\varphi \in \mathscr{D}(I)$, since supp$(\varphi - \chi\varphi) \cap$ supp $T = \emptyset$, we have

$$\langle T, \varphi \rangle = \langle T, \chi\varphi \rangle + \langle T, \varphi - \chi\varphi \rangle = \langle T, \chi\varphi \rangle.$$

Since supp$(\chi\varphi) \subset [a,b]$, we have

$$|\langle T, \varphi \rangle| = |\langle T, \chi\varphi \rangle| \leq C \sum_{k=0}^{m} \sup |(\chi\varphi)^{(k)}|.$$

The Leibniz formula gives

$$\sup |(\chi\varphi)^{(k)}| = \sup_{[a,b]} |(\chi\varphi)^{(k)}| \leq C \sum_{\ell=0}^{k} \sup_{[a,b]} |\varphi^{(\ell)}|,$$

We first consider a test function $\varphi \in \mathscr{D}(I)$ satisfying $\varphi^{(k)}(0) = 0$ for every $k \leq m$, then $\langle T, \varphi \rangle = 0$. Indeed, let χ be a cutoff function on $\left[-\frac{1}{2}, \frac{1}{2}\right]$, with supp$(\chi) \subset [-1,1]$. For $\varepsilon > 0$ small enough, we define a rescaled cutoff function

$$\chi_\varepsilon(x) = \chi\left(\frac{x}{\varepsilon}\right) \ .$$

The function $\varphi - \chi_\varepsilon\varphi$ vanishes in $[-\varepsilon/2, \varepsilon/2]$, a neighborhood of 0, so

$$\langle T, \varphi \rangle = \langle T, \chi_\varepsilon\varphi \rangle \ .$$

The assumption on φ and a generalization of the Hadamard Lemma shows that φ factorizes into

$$\varphi(x) = x^{m+1}\psi(x)$$

for some $\psi \in \mathscr{D}(I)$. Hence

$$\chi_\varepsilon\varphi(x) = \chi\big(\frac{x}{\varepsilon}\big)x^{m+1}\psi(x) = \varepsilon^{m+1}\rho\big(\frac{x}{\varepsilon}\big)\psi(x) \ ,$$

where we introduced the function $\rho(y) = y^{m+1}\chi(y)$. By the Leibniz formula, one easily shows the existence of some $B > 0$ such that, uniformly in ε:

$$\sup_{k \leq m} \sup_{x} \left| \Big(\rho\big(\frac{x}{\varepsilon}\big)\psi(x)\Big)^{(k)} \right| \leq \frac{B}{\varepsilon^m} \ .$$

As a consequence

$$|\langle T, \varphi \rangle| = |\langle T, \chi_\varepsilon\varphi \rangle| \leq C \sup_{k \leq m} \sup |(\chi_\varepsilon\varphi)^{(k)}| \leq CB\varepsilon \ .$$

Since ε can be chosen arbitrary small, while C, B do not depend on ε, we obtain the claimed result

$$\langle T, \varphi \rangle = 0 \ .$$

Let us now consider a general test function $\varphi \in \mathscr{D}(I)$. The Taylor formula at order m reads

$$\varphi(x) = \sum_{k \leq m} \frac{x^k}{k!}\varphi^{(k)}(0) + \tilde{r}(x),$$

where the remainder $\tilde{r} \in C^\infty(I)$ satisfies

$$\tilde{r}^{(k)}(0) = 0 \ , \text{ for all } k \leq m \ .$$

However, we would like the various terms on the right-hand side to be in $\mathscr{D}(I)$, so we choose $\varepsilon_0 > 0$ is small enough so that $[-\varepsilon_0, \varepsilon_0] \subset I$, and use the rescaled cutoff χ_{ε_0} to modify the above formula into:

$$\varphi(x) = \sum_{k \leq m} \frac{x^k}{k!} \varphi^{(k)}(0) \chi_{\varepsilon_0}(x) + \tilde{r}(x).$$

All terms, including the remainder r, are now in $\mathscr{D}(I)$. Besides, $r(x) = \tilde{r}(x)$ in $[-\varepsilon/2, \varepsilon/2]$, so in particular

$$r^{(k)}(0) = 0 \ , \text{ for all } k \leq m \ .$$

The previous statement hence applies to the function r: $\langle T, r \rangle = 0$, so by linearity we get

$$\langle T, \varphi \rangle = \sum_{k \leq m} \varphi^{(k)}(0) \langle T, \frac{x^k}{k!} \chi_{\varepsilon_0}(x) \rangle.$$

This is precisely what we have claimed, if we set $a_k = \langle T, \frac{x^k}{k!} \chi_{\varepsilon_0} \rangle$ (these numbers do not depend on the choice of cutoff function χ, nor on ε_0). □

1.7 Converging sequences of distributions

1.7.1 Convergence in $\mathscr{D}'$

We have defined the notion of convergence in $\mathscr{D}(I)$, and showed that a distribution is (sequentially) continuous w.r.t. this notion. We now define a dual notion of convergence in $\mathscr{D}'(I)$.

Definition 1.7.1 Let (T_j) be a sequence of distributions in $\mathscr{D}'(I)$. We say that (T_j) converges to $T \in \mathscr{D}'(I)$ when, for any function $\varphi \in \mathcal{C}_0^\infty(I)$, the sequence of complex numbers $(\langle T_j, \varphi \rangle)$ converges to $\langle T, \varphi \rangle$. In this case, we write $T_j \to T$ in $\mathscr{D}'(I)$.

If we view T_j and T as functions $\mathscr{D}(I) \to \mathcal{C}$, this definition amounts to the *simple* convergence of these functions.

Example 1.7.2 If (f_j) is a sequence of locally integrable functions on I such that, for some $f \in L^1_{\text{loc}}(I)$,

$$\forall [a,b] \subset I \ , \ \int_a^b |f_j(x) - f(x)| \, dx \to 0 \ ,$$

(we say that f_j converge to f in L^1_{loc}), then

$$T_{f_j} \to T_f \quad \text{in } \mathscr{D}'(I).$$

Indeed, for every test function φ on I with $\operatorname{supp}\varphi \subset [a,b]$,

$$\langle T_{f_j}, \varphi\rangle - \langle T_f, \varphi\rangle = \int_I f_j(x)\varphi(x)\,dx - \int_I f(x)\varphi(x)\,dx$$

so that

$$\left|\langle T_{f_j}, \varphi\rangle - \langle T_f, \varphi\rangle\right| \leq \sup|\varphi| \int_a^b |f_j(x) - f(x)|\,dx$$

which tends to 0 as j tends to the infinity.

Notice that a special case of this is the following situation. The sequence f_j converges to f almost everywhere on I, and there exists a locally integrable function h on I such that $\sup_j |f_j| \leq h$ almost everywhere on I. Indeed, the connection to the above condition of L^1 convergence is provided by the dominated convergence theorem in $L^1([a,b])$.

Example 1.7.3 Let $\rho \in L^1(\mathbb{R})$ such that $\int_{\mathbb{R}} \rho(x)\,dx = 1$, and (ρ_ε) the sequence of rescaled functions defined by

$$\rho_\varepsilon(x) = \varepsilon^{-1}\rho(\frac{x}{\varepsilon}).$$

We also denote $(T_\varepsilon) = (T_{\rho_\varepsilon})$ the associated family of distributions. For $\varphi \in \mathcal{C}_0^\infty(\mathbb{R})$, we have

$$\langle T_\varepsilon, \varphi\rangle = \int_{\mathbb{R}} \rho_\varepsilon(x)\varphi(x)dx = \int_{\mathbb{R}} \rho(y)\varphi(\varepsilon y)\,dy.$$

By the dominated convergence theorem, $\langle T_\varepsilon, \varphi\rangle \to \varphi(0)$ as $\varepsilon \to 0$. Therefore the sequence (T_{ρ_ε}) converges to δ_0 in $\mathscr{D}'(\mathbb{R})$.

The following Exercise generalizes the type of sequences (f_j) which converge towards a Dirac delta mass.

Exercise 1.7.4 Let (f_j) be a sequence of L^1 functions on I such that

$$\sup_j \int_I |f_j|\,dx < \infty\ ,\ \int_I f_j dx \to c\ ,\ \operatorname{supp}(T_{f_j}) \subset [x_0 - \varepsilon_j, x_0 + \varepsilon_j]\ ,\ \varepsilon_j \to 0\ .$$

Prove that

$$T_{f_j} \to c\delta_{x_0}\ .$$

One may also approximate distributions of strictly positive order, as shown in the following

Exercise 1.7.5 Let

$$f_\varepsilon(x) = \begin{cases} \frac{1}{\varepsilon^2} & \text{if } x \in]0,\varepsilon[\\ \frac{-1}{\varepsilon^2} & \text{if } x \in]-\varepsilon, 0[\\ 0 & \text{if } |x| > \varepsilon \end{cases}$$

Prove that

$$T_{f_\varepsilon} \underset{\varepsilon\to 0}{\longrightarrow} -\delta_0'\ .$$

Exercise 1.7.6 Let

$$f_\varepsilon(x) = \frac{x}{x^2+\varepsilon}\ ,\ \varepsilon > 0\ .$$

Prove that

$$T_{f_\varepsilon} \underset{\varepsilon\to 0}{\longrightarrow} \text{pv}\left(\frac{1}{x}\right) \quad \text{in } \mathscr{D}'(\mathbb{R}).$$

An important property is that the operations that we have defined on distributions are all continuous with respect to the notion of convergence in $\mathscr{D}'(I)$. More precisely,

Proposition 1.7.7 If (T_j) converges to T in $\mathscr{D}'(I)$, then

i) For any $k \in \mathbb{N}$, $(T_j^{(k)})$ converges to $T^{(k)}$.

ii) For any $f \in \mathcal{C}^\infty(I)$, (fT_j) converges to fT.

iii) For any open subinterval $J \subset I$, the sequence $(T_j)_{|J}$ converges to $T_{|J}$ in $\mathscr{D}'(J)$.

Proof.— Let $\varphi \in \mathcal{C}_0^\infty(I)$. We have clearly

$$\forall k \in \mathbb{N}, \qquad \langle (T_j)^{(k)}, \varphi\rangle = (-1)^k \langle T_j, \varphi^{(k)}\rangle \to (-1)^k\langle T, \varphi^{(k)}\rangle = \langle T^{(k)}, \varphi\rangle,$$

and

$$\langle fT_j, \varphi\rangle = \langle T_j, f\varphi\rangle \to \langle T, f\varphi\rangle = \langle fT, \varphi\rangle.$$

The statement $iii)$ is obvious. □

Exercise 1.7.8 Show that if (f_j) converges to f in $\mathcal{C}^\infty(I)$, in the sense of the uniform convergence of all derivatives on every compact subset, then (f_jT) converges to fT in $\mathscr{D}'(I)$.

1.7.2 The Uniform Boundedness Principle

Let us motivate the main statement of this paragraph by the following questions.

i) Assume that, for a sequence (T_j) of distributions, and for every $\varphi \in \mathscr{D}(I)$, the sequence $(\langle T_j, \varphi\rangle)$ has a limit $L(\varphi)$. This clearly defines a linear form L on $\mathscr{D}(I)$. Is it true that L is a distribution ?

ii) Assume $T_j \to T$ in $\mathscr{D}'(I)$ and $f_j \to f$ in $\mathcal{C}^\infty(I)$ in the sense of the uniform convergence of all derivatives on every compact subset, does it imply that $f_jT_j \to fT$? In order to to prove such a result, we write

$$\langle f_jT_j, \varphi\rangle - \langle fT, \varphi\rangle = \langle T_j, (f_j - f)\varphi\rangle + \big(\langle T_j, f\varphi\rangle - \langle T, f\varphi\rangle\big)\ .$$

The second term in the right hand side converges to 0 in view of the assumption. In the first term, we notice that $\psi_j = (f_j - f)\varphi$ converges to 0 in $\mathscr{D}(I)$, but it is seems difficult to conclude that $\langle T_j, \psi_j\rangle \to 0$ without a bound on the action of T_j which is uniform with respect to ψ_j.

We state below, without proof, an important theoretical result, which is a "distributional version" of the Banach-Steinhaus theorem[7], which provides such a uniform bound.

Proposition 1.7.9 [Uniform boundedness] Let (T_j) be a sequence in $\mathscr{D}'(I)$, and $[a,b] \subset I$ a segment. If, for any function $\varphi \in \mathcal{C}_0^\infty(I)$ with support in $[a,b]$,

$$\sup_j |\langle T_j, \varphi\rangle| < +\infty\,,$$

then there exists $C > 0$ and $m \in \mathbb{N}$, independent of j, such that

$$\forall \varphi \in \mathscr{D}(I),\ \forall j \in \mathbb{N}, \qquad \operatorname{supp}\varphi \subset [a,b] \Longrightarrow |\langle T_j, \varphi\rangle| \leq C \sum_{k=0}^{m} \sup |\varphi^{(k)}|.$$

In other words, the distributions $(T_j)_{j\in\mathbb{N}}$ are uniformly controlled on the interval $[a,b]$.

As a consequence, we obtain the following somewhat surprising result.

Corollary 1.7.10 Let (T_j) be a sequence of distributions on I. If, for all functions $\varphi \in \mathscr{D}(I)$, the sequence $(\langle T_j, \varphi\rangle)_{j\in\mathbb{N}}$ converges in $\mathbb{C}$, then the resulting linear form is actually a distribution $T \in \mathscr{D}'(I)$, and $T_j \to T$ in $\mathscr{D}'(I)$.

Proof.— Let $T : \mathscr{D}(I) \to \mathbb{C}$ be the linear form given by

$$T(\varphi) = \lim_{j\to+\infty} \langle T_j, \varphi\rangle.$$

We need to check that T is a distribution. So pick a segment $[a,b] \subset I$. Proposition 1.7.9 ensures that there is a constant $C > 0$ and an integer m such that, for any $\varphi \in \mathscr{D}(I)$ with $\operatorname{supp}\varphi \subset [a,b]$, we have

$$|\langle T_j, \varphi\rangle| \leq C \sum_{k=0}^{m} \sup |\varphi^{(k)}|.$$

Then we can pass to the limit $j \to +\infty$, and we get the same estimate for $|\langle T, \varphi\rangle|$. □

The second question above can also be addressed successfully (the proof is left as an exercise).

Corollary 1.7.11 Let $(T_j)_{j\in\mathbb{N}}$ be a sequence of distributions on I which converges to T in $\mathscr{D}'(I)$. Then, for every sequence (ψ_j) which converges to ψ in $\mathscr{D}(I)$, we have $\langle T_j, \psi_j\rangle \to \langle T, \psi\rangle$. In particular, if $f_j \to f$ in $\mathcal{C}^\infty(I)$ in the sense of the uniform convergence of all derivatives on every compact subset, then $f_j T_j \to fT$.

7. Let E, F be two Banach spaces, and $(A_j : E \to F)_j$ be a sequence of bounded operators, such that for any $u \in E$, $\sup_j \|A_j u\|_F < \infty$. Then the operators A_j are uniformly bounded: $\sup \|A_j\|_{\mathcal{L}(E,F)} < \infty$.

A proof of the uniform boundedness principle, as well as other applications, will be given in the next chapter.

1.8 An introduction to Sobolev spaces

In this section, we give a short overview of a family of functional spaces, introduced by the Russian mathematician Sergei Sobolev (1908-1989), which allowed to solve differential equations in a wide context. Here we concentrate on a very special case in one variable. A more general presentation will be provided in the next chapters.

1.8.1 Generalised derivatives

Definition 1.8.1 Let I be an open interval of $\mathbb{R}$, and let $u \in L^1_{\text{loc}}(I)$. We shall say that u admits a generalised derivative if there exists $f \in L^1_{\text{loc}}(I)$ such that, in the sense of distributions in I,

$$T_u' = T_f \ .$$

The element $f \in L^1_{\text{loc}}(I)$ is then unique, it is called the generalised derivative of u, and is denoted by $f = u'$.

Those elements u are said to belong to the space $W^{1,1}_{\text{loc}}(\mathbb{R})$, the (local) first Sobolev space based on L^1.

The uniqueness of f immediately follows from Proposition 1.3.5. The notation $f = u'$ is precisely fitted to satisfy a generalised integration by parts formula, which is nothing but a reformulation of $T_u' = T_f$,

$$\forall \varphi \in \mathcal{C}_0^\infty(I) \ , \quad \int_I u(x)\varphi'(x)\,dx = -\int_I u'(x)\varphi(x)\,dx \ .$$

From Corollary 1.4.9, we infer the following useful statement.

Proposition 1.8.2 Assume $u \in L^1_{\text{loc}}(I)$ admits a generalised derivative $u' \in L^1_{\text{loc}}(I)$. Let $x_0 \in I$. Then there exists $c \in \mathbb{C}$ such that

$$u(x) = \int_{x_0}^x u'(t)\,dt + c \qquad a.e.$$

In particular, u admits as reprensentative a continuous function.

Notice that, as a consequence of the Proposition 1.8.2, if the generalised derivative u' of u turns out to be a continuous function, then u is a $\mathcal{C}^1$ function, and u' is merely its derivative in the usual sense.

1.8.2 The Sobolev space $H^1(I)$.

We now introduce the more useful Sobolev space based on L^2, which one simply calls "Sobolev space".

Definition 1.8.3 Let I be an open interval of $\mathbb{R}$. We denote by $H^1(I)$ the subspace formed by the $u \in L^2(I)$ which admit a generalised derivative $u' \in L^2(I)$.

Notice that, since $L^2(I) \subset L^1_{\text{loc}}(I)$ by Cauchy-Schwarz, we also have $H^1(I) \subset W^{1,1}_{\text{loc}}(I)$.

One huge advantage of the Sobolev space $H^1(I)$ is that we can naturally equip it with an inner product:

$$(u,v)_{H^1} \overset{\text{def}}{=} \int_I u'(x)\overline{v'}(x)\,dx + \int_I u(x)\overline{v}(x)\,dx\ ,\ u \in H^1(I)\ ,\ v \in H^1(I)\ , \tag{1.8.10}$$

and the derived norm:

$$\|u\|_{H^1} = \Big(\|u'\|^2_{L^2} + \|u'\|^2_{L^2}\Big)^{1/2} . \tag{1.8.11}$$

The following proposition shows how the notion of convergence in $\mathcal{D}'(I)$ is useful in this context.

Proposition 1.8.4 Endowed with the inner product (1.8.10), the space $H^1(I)$ is a Hilbert space.

Proof.— Let (u_j) be a Cauchy sequence in $H^1(I)$ for the norm (1.8.11). This precisely means that (u_j) and (u'_j) are Cauchy sequences in $L^2(I)$. Since $L^2(I)$ is a Hilbert space, there exist u, v in $L^2(I)$ such that

$$u_j \to u\ ,\quad u'_j \to v \quad \text{in } L^2(I)\,.$$

Since the convergence in $L^2(I)$ implies the convergence on $L^1(]a,b[)$ for every segment $[a,b] \subset I$, we infer that

$$T_{u_j} \to T_u \quad \text{and} \quad T_{u'_j} \to T_v \quad \text{in } \mathscr{D}'(I)\,.$$

But, by definition of a generalised derivative, and using the continuity of derivation for the convergence in $\mathscr{D}'(I)$,

$$T_{u'_j} = T'_{u_j} \to T'_u\,.$$

This implies the identity $T'_u = T_v$, in other words u admits v as a generalised derivative, $u' = v \in L^2(I)$, which means that $u \in H^1(I)$ and

$$u_j \to u\ ,\quad u'_j \to u' \text{ in } L^2(I)\,.$$

This precisely means that u_j tends to u in $H^1(I)$. Hence $H^1(I)$ is a *complete* normed vector space.
□

1.8.3 First Sobolev space on a bounded interval

In this section, $I =]a, b[$, for $a, b \in \mathbb{R}$. In this case, we get more informations from Proposition 1.8.2 .

Proposition 1.8.5 [Sobolev inequality] If $u \in H^1(]a, b[)$, then u admits as representative a continuous function on $[a, b]$, and we have the Sobolev inequality,

$$\|u\|_{L^\infty(]a,b[)} \leq C\|u\|_{H^1(]a,b[)} ,$$

where C is a constant depending only on the length $(b - a)$ of the interval.

This Sobolev inequality achieves an unexpected goal: a pointwise control of the function u, in terms of "averaged" norms for u and u'. We will often replace the equivalence class $u \in L^2$ by the specific continuous representative $u \in C^0$. **Proof.—** We start from the identity in Proposition 1.8.2: fixing some $x_0 \in]a, b[$, there exists $c \in \mathbb{C}$ such that

$$\forall x \in [a, b], \qquad u(x) = \int_{x_0}^{x} u'(t)\, dt + c \quad (a.e.)$$

Since $L^2(]a, b[) \subset L^1(]a, b[)$, we infer that the right hand side has a limit when x tends to a and when x tends to b. Hence u is almost everywhere equal to a continuous function on $[a, b]$. Furthermore, by the Cauchy-Schwarz inequality, for any $x, y \in [a, b]$,

$$\forall x, y \in [a, b], \qquad \left| \int_y^x u'(t)\, dt \right| \leq \sqrt{b-a}\|u'\|_{L^2(]a,b[)} .$$

Consider the average of u,

$$m_u = \frac{1}{b-a} \int_a^b u(y)\, dy .$$

By the Cauchy–Schwarz inequality,

$$|m_u| \leq \frac{1}{\sqrt{b-a}} \|u\|_{L^2(]a,b[)} .$$

On the other hand,

$$u(x) - m_u = \frac{1}{b-a} \int_a^b (u(x) - u(y))\, dy = \frac{1}{b-a} \int_a^b \left(\int_y^x u'(t)\, dt \right) dy .$$

Therefore

$$|u(x) - m_u| \leq \frac{1}{b-a} \int_a^b \left| \int_y^x u'(t)\, dt \right| dy \leq \sqrt{b-a}\|u'\|_{L^2(]a,b[)} .$$

Summing up, we obtain

$$\|u\|_{L^\infty} \leq \|u - m_u\|_{L^\infty} + |m_u| \leq (b-a)^{-1/2}\|u\|_{L^2} + (b-a)^{1/2}\|u'\|_{L^2} \leq C\|u\|_{H^1} .$$

□

Corollary 1.8.6 For any point $x_0 \in [a,b]$, the linear form $u \in H^1(]a,b[) \mapsto u(x_0) \in \mathbb{C}$ is continuous (here u is implicitly the distinguished continuous representative).

Our next step is the introduction of an important closed subspace of $H^1(]a,b[)$.

Definition 1.8.7 We denote by $H_0^1(]a,b[)$ the closure of $\mathcal{C}_0^\infty(]a,b[)$ in $H^1(I)$.

Corollary 1.8.6 shows that $u \mapsto u(a)$ and $u \mapsto u(b)$ are continuous linear forms on $H^1(]a,b[)$. A remarkable fact is that $H_0^1(]a,b[)$ can be characterised by these linear forms.

Proposition 1.8.8 Given $u \in H^1(]a,b[)$, u belongs to $H_0^1(]a,b[)$ if and only if $u(a) = u(b) = 0$.

Proof.— Since the elements of $H^1(]a,b[)$ are continuous functions on $[a,b]$, and the H^1 norm controls the L^∞ norm on $[a,b]$ from Proposition 1.8.5, for any $u \in H_0^1(]a,b[)$ and any sequence (φ_n) in $\mathscr{D}(]a,b[)$ such that $\|\varphi_n - u\|_{H^1} \to 0$, we also have

$$\|\varphi_n - u\|_{L^\infty} \to 0\,,$$

namely the test functions converge *uniformly* towards u. As a result, since all the φ_n vanish on the boundaries of $]a,b[$, so does their uniform limit u: $u(a) = u(b) = 0$.

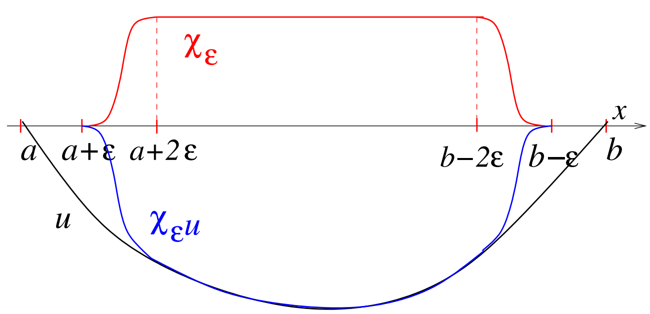


Figure 1.6: Truncation of $u \in H_0^1(]a,b[)$ near the boundaries of the interval.

Conversely, consider $u \in H^1(]a,b[)$ such that $u(a) = u(b) = 0$. Our first task will be to approximate it by functions $u_\varepsilon \in H^1$ which vanish near the boundaries of the interval. For this, take $\varepsilon > 0$ small enough, and consider a cutoff function function $\chi_\varepsilon \in C_0^\infty(I,[0,1])$, supported in $[a+\varepsilon, b-\varepsilon]$, and such that $\chi_\varepsilon = 1$ on $[a+2\varepsilon, b-2\varepsilon]$, with its derivative satisfying the estimate

$$\|\chi_\varepsilon'\|_{L^\infty} = \mathcal{O}\left(\frac{1}{\varepsilon}\right) \quad \text{when } \varepsilon \to 0\,.$$

An example of such a family of functions is

$$\chi_\varepsilon(x) = \chi\left(\frac{x-a}{\varepsilon}\right)\chi\left(\frac{b-x}{\varepsilon}\right),$$

where $\chi \in \mathbb{C}^\infty(\mathbb{R})$ is a "smooth step function", which satisfies

$$\chi(t) = \begin{cases} 0 \text{ if } & t \leq 1 \\ 1 \text{ if } & t \geq 2. \end{cases}$$

We then claim that $u_\varepsilon \stackrel{\text{def}}{=} \chi_\varepsilon u \to u$ in $H^1(]a,b[)$ as $\varepsilon \to 0$.

Indeed, by dominated convergence $\chi_\varepsilon u \to u$ in $L^2(]a,b[)$. Besides, by Leibniz we have

$$(\chi_\varepsilon u)' = \chi_\varepsilon u' + \chi_\varepsilon' u\ .$$

Since the first term converges to u' in L^2, we are therefore reduced to proving that $\chi_\varepsilon' u \to 0$ in $L^2(]a,b[)$. This is not completely obvious, since χ_ε' can take large values; however, it only takes nonzero values in small intervals near a and b. We have

$$\|\chi_\varepsilon' u\|_{L^2} \leq \mathcal{O}\left(\frac{1}{\varepsilon}\right)\left(\int_a^{a+2\varepsilon} |u(x)|^2\,dx + \int_{b-2\varepsilon}^b |u(x)|^2\,dx\right)^{\frac{1}{2}}\ .$$

Let us estimate the values of $u(x)$ in the small intervals near a and b. For each $a \leq x \leq a+2\varepsilon$, the explicit expression of $u(x)$ gives the bound

$$\begin{aligned} |u(x)| &= \left|u(a) + \int_a^x u'(t)\,dt\right| = \left|\int_a^x u'(t)\,dt\right| \\ &\stackrel{C-S}{\leq} \sqrt{2\varepsilon}\left(\int_a^{a+2\varepsilon} |u'(x)|^2\,dx\right)^{\frac{1}{2}} = \sqrt{\varepsilon}\,o(1)_{\varepsilon\to 0}\,, \end{aligned}$$

and the same estimate holds for $x \in [b-2\varepsilon, b]$. Here we used the fact that $\|u'\|_{L^2}$ is finite, so the integrals of $|u'|^2$ in smaller and smaller intervals decreases to zero. Integrating $|u(x)|^2$ over these intervals, we get:

$$\int_a^{a+2\varepsilon} |u(x)|^2\,dx = o(\varepsilon^2) \quad \text{when } \varepsilon \to 0,$$

and the same estimate for the integral on $[b-2\varepsilon, b]$. Putting all those estimates together, we find:

$$\|\chi_\varepsilon' u\|_{L^2} \leq \mathcal{O}(1/\varepsilon)\left(o(\varepsilon^2)\right)^{1/2} = o(1) \quad \text{when } \varepsilon \to 0\ .$$

We have proved that u can be approximated in H^1 by functions $u_\varepsilon = \chi_\varepsilon u \in H^1(]a,b[)$ supported in $[a+2\varepsilon, b-2\varepsilon]$. Applying the following Lemma, each of these u_ε can be approached, in the H^1 norm, by test functions $u_{\varepsilon,\eta}$. Playing with the two limits $\varepsilon, \eta \to 0$, we therefore approach u by test functions.

□

Lemma 1.8.9 Any compactly supported element of $H^1(]a,b[)$ can be approximated (w.r.t. the H^1 norm) by a sequence in $\mathscr{D}(]a,b[)$.

Proof.— The statement relies on an explicit regularisation argument. Let $v \in H^1(]a,b[)$ such that supp$(v) \subset [\alpha,\beta] \subset]a,b[$. We extend v by 0, obtaining an element of $H^1(\mathbb{R})$ supported in $[\alpha,\beta]$. Let $\rho \in \mathcal{C}_0^\infty(\mathbb{R})$ of integral 1, and rescale it on a small scale $\eta > 0$:

$$\rho_\eta(x) = \frac{1}{\eta}\rho\left(\frac{x}{\eta}\right) \ .$$

Then $\rho_\eta * v \in \mathcal{C}_0^\infty(\mathbb{R})$, is compactly supported in $]a,b[$ if η is small enough, and $\rho_\eta * v \to v$ in L^2 as $\eta \to 0$. Furthermore, by the Leibniz rule,

$$(\rho_\eta * v)'(x) = \int_{\mathbb{R}} \rho_\eta'(x-y)\, v(y)\, dy = \int_{\mathbb{R}} -\frac{d}{dy}[\rho_\eta(x-y)]\, v(y)\, dy = \int_{\mathbb{R}} \rho_\eta(x-y)\, v'(y)\, dy \ ,$$

by definition of the generalised derivative. In other words,

$$(\rho_\eta * v)' = \rho_\eta * v'$$

and therefore $(\rho_\eta * v)' \to v'$ in L^2 as $\eta \to 0$. Summing up, we have proved that $\rho_\eta * v \to v$ in H^1. □

Proposition 1.8.8 implies an important inequality for elements of $H_0^1(]a,b[)$.

Proposition 1.8.10 (The Poincaré inequality) For every $u \in H^1(]a,b[)$ such that $u(a) = 0$ or $u(b) = 0$, we have

(1.8.12) $$\|u\|_{L^2} \le (b-a)\|u'\|_{L^2} \ .$$

In particular,

$$(u,v)_{H_0^1} = \int_a^b u'(x)\, \overline{v'}(x)\, dx$$

defines an inner product on $H_0^1(]a,b[)$, which is equivalent to the H^1 inner product.

Proof.— Again we refer to Proposition 1.8.2,

$$u(x) = \int_{x_0}^x u'(t)\, dt + c \ .$$

Assume for instance $u(a) = 0$. Then making x tend to a, we infer $c = \int_a^{x_0} u'(t)\, dt$, and consequently

$$u(x) = \int_a^x u'(t)\, dt \ .$$

The inequality (1.8.12) then follows from the Cauchy-Schwarz inequality:

$$|u(x)|^2 \leq \int_a^x (x-a)\,|u'(t)|^2\,dt \Longrightarrow \int \|u\|_{L^2}^2 \leq \frac{(b-a)^2}{2}\|u'\|_{L^2}^2 \,.$$

In view of Proposition 1.8.8, the inequality (1.8.12) applies in particular to elements of $H_0^1(]a,b[)$, so that

$$\|u\|_{L^2}^2 \leq \frac{(b-a)^2}{2}\,(u,u)_{H_0^1} \,.$$

Since $(u,u)_{H^1} = (u,u)_{H_0^1} + \|u\|_{L^2}^2$, we conclude that the scalar products $(u,v)_{H^1}$ and $(u,v)_{H_0^1}$ on H_0^1 are equivalent. □

1.8.4 Application: solving the Dirichlet problem in dimension 1

The Hilbert structure on H_0^1 leads to a very efficient strategy for solving second order linear differential equations with homogeneous boundary conditions.

Theorem 1.8.11 Let $q \in L^1(]a,b[)$ such that $q \geq 0$, and let $f \in L^1(]a,b[)$. There exists a unique $u \in H^1(]a,b[)$ such that u' admits a generalised derivative u'' satisfying

$$-u'' + qu = f \,,\; u(a) = u(b) = 0 \,.$$

Proof.— Notice that qu is well defined in L^1 since $q \in L^1$ and $u \in L^\infty$. Decomposing f into its real and imaginary parts, we may assume f is real valued, so that u is to be real valued as well. So we shall work in the real Hilbert space made of real valued elements of H^1. This simple reduction allows to avoid the complex conjugation in the inner product, so that the connection with the distribution bracket is clearer.

In view of Proposition 1.8.8, the above problem is equivalent to

$$-T'_{u'} + T_{qu} = T_f \,,\; u \in H_0^1(]a,b[) \,,$$

or, for every $\varphi \in \mathcal{C}_0^\infty(]a,b[)$,

$$\int_a^b u'\varphi'\,dx + \int_a^b qu\varphi\,dx = \int_a^b f\varphi\,dx \,,\; u \in H_0^1(]a,b[) \,.$$

Since the left hand side and the right hand side of the above identity are continuous linear forms of φ for the H^1 norm, and since H_0^1 is the closure of $\mathcal{C}_0^\infty$ for this norm, those two linear forms on $\mathscr{D}(]a,b[)$ can be continuously extended to test functions $v \in H_0^1(]a,b[)$. Our problem is thus equivalent to finding $u \in H_0^1(]a,b[)$ such that

$$\forall v \in H_0^1(]a,b[) \,,\; \int_a^b u'v'\,dx + \int_a^b quv\,dx = \int_a^b fv\,dx \,.$$

Now observe that the left hand side is an inner product on the real space H_0^1, which is controlled by the H^1 inner product: Indeed,

$$(u,u)_{H_0^1} \leq \int_a^b (u')^2\,dx + \int_a^b qu^2\,dx \quad \leq \quad \int_a^b (u')^2\,dx + \left(\int_a^b q\,dx\right)\|u\|_{L^\infty}^2$$
$$\overset{Prop. 1.8.5}{\leq} \quad \max\left(1, C^2\left(\int_a^b q\,dx\right)\right)(u,u)_{H^1}$$

and Prop. 1.8.12 shows that the H^1 inner product is controlled by the H_0^1 product. Hence, the norm induced on H_0^1 by the inner product

$$(u,v)_q = \int_a^b u'v'\,dx + \int_a^b quv\,dx$$

is equivalent with the standard H_0^1 norm on that Hilbert space. Since, in view of Proposition 1.8.5, the linear form

$$L_f : v \mapsto \int_a^b fv\,dx$$

is continuous on this Hilbert space, the Riesz representation theorem allows to solve the equation $(u,\cdot)_q = L_f$ by a unique $u \in H_0^1$. □

Remark 1.8.12 Since u'' belongs to $L^1(]a,b[)$, we infer that u' extends as a continuous function on $[a,b]$. Furthermore, if q,f are continuous functions on $[a,b]$, then u'' is continuous on $[a,b]$, which means that u is $\mathcal{C}^2$ on $[a,b]$, so that the differential equation is satisfied in the usual sense.

1.9 Further properties of distributions and generalized derivatives

1.9.1 Characterisation of Lipschitz functions

Definition 1.9.1 Let k be a positive number and I be an open interval. A function $u : I \to \mathbb{C}$ is k–Lipschitz if, for every $x, y \in I$,

$$|u(x) - u(y)| \leq k|x-y| \ .$$

A typical example of a k–Lipschitz function is a C^1 function u such $\|u'\|_\infty \leq k$. Of course, there are Lipschitz functions which are not differentiable everywhere, like function $x \mapsto |x|$. However, a theorem by Rademacher states that a Lipschitz function on $\mathbb{R}$ is differentiable Lebesgue-almost everywhere.

The following result give a complete description of k–Lipschitz functions.

Theorem 1.9.2 A function $u : \mathbb{R} \to \mathbb{C}$ is k–Lipschitz if and only if there exists $f \in L^\infty(\mathbb{R})$, with $\|f\|_{L^\infty} \leq k$, such that

$$\forall x \in \mathbb{R}\ ,\ u(x) = u(0) + \int_0^x f(t)\, dt\ . \tag{1.9.13}$$

Remark 1.9.3 A similar result holds on any open interval I, from an adaptation of the proof below on $\mathbb{R}$.

Proof.— If $\forall x \in \mathbb{R}$, $u(x) = u(0) + \int_0^x f(t)\, dt$, we have

$$|u(x) - u(y)| = \left| \int_y^x f(t)\, dt \right| \leq \|f\|_\infty |x - y|\ .$$

Conversely, let u be a k–Lipschitz function. We are going to prove that u admits a generalised derivative f with $\|f\|_\infty \leq k$. According to Proposition 1.8.2, this will imply property (1.9.13).
As a first step, we are going to prove the following inequality,

$$\forall \varphi \in \mathscr{D}(I)\ ,\ |\langle T_u', \varphi\rangle| \leq k\|\varphi\|_{L^1}\ . \tag{1.9.14}$$

Indeed, we have,

$$\begin{aligned} \langle T_u', \varphi\rangle &= -\int_{\mathbb{R}} u(x)\varphi'(x)\, dx = -\int_{\mathbb{R}} u(x) \lim_{h\to 0} \frac{\varphi(x+h) - \varphi(x)}{h}\, dx \\ &= -\lim_{h\to 0} \int_{\mathbb{R}} u(x) \frac{\varphi(x+h) - \varphi(x)}{h}\, dx\ , \end{aligned}$$

by using either dominated convergence or uniform convergence. Next we decompose the integral as follows,

$$\begin{aligned} \int_{\mathbb{R}} u(x) \frac{\varphi(x+h) - \varphi(x)}{h}\, dx &= \frac{1}{h} \left(\int_{\mathbb{R}} u(x)\varphi(x+h)\, dx - \int_{\mathbb{R}} u(x)\varphi(x)\, dx \right) \\ &= \frac{1}{h} \left(\int_{\mathbb{R}} u(y-h)\varphi(y)\, dy - \int_{\mathbb{R}} u(x)\varphi(x)\, dx \right) \\ &= \int_{\mathbb{R}} \frac{u(y-h) - u(y)}{h} \varphi(y)\, dy\ , \end{aligned}$$

and, by the Lipschitz property of u, the modulus of the latter expression is bounded by $k\|\varphi\|_{L^1}$. This leads to inequality (1.9.14). At this stage we appeal to the following useful lemma of functional analysis.

Lemma 1.9.4 (Continuous extension of a linear form) Let $\mathcal{E}$ be a normed vector space, and let $\mathscr{D}$ be a dense vector subspace of $\mathcal{E}$. Let $L : \mathscr{D} \to \mathbb{C}$ be a linear form, which is continuous for the norm on $\mathcal{E}$. Then L admits a unique continuation $\tilde{L}$ as a continuous linear form on $\mathcal{E}$, and $\|\tilde{L}\| = \|L\|$.

Applying Lemma 1.9.4 to $\mathcal{E} = L^1(\mathbb{R})$, $\mathscr{D} = \mathscr{D}(\mathbb{R})$ and $L = T_u'$, we infer that T_u' extends to a continuous linear form on $L^1(\mathbb{R})$, of norm at most k. Since the dual of $L^1(\mathbb{R})$ is identified (isometrically) to $L^\infty(\mathbb{R})$, there exists $f \in L^\infty(\mathbb{R})$ such that $\|f\|_{L^\infty} \leq k$ and $L = T_f$. As a result,

$$\forall \varphi \in \mathscr{D}(I)\ ,\ \langle T_u', \varphi\rangle = \int_{\mathbb{R}} f(x)\varphi(x)\, dx\ .$$

This precisely means that f is the generalised derivative of u, whence (1.9.13). □

1.9.2 Differentiation and integration under the bracket

This paragraph provides two very useful rules of computation, namely the differentiation and the integration under the bracket, when our test functions smoothly depend on an auxiliary parameter z. These properties will be crucial when dealing with convolution of distributions — see the next paragraph.

Proposition 1.9.5 (Differentiation under the bracket) Let I, J be open intervals, $\psi : I \times J \to \mathbb{C}$ be a $\mathcal{C}^\infty$ function such that there exists a segment $[a, b] \subset I$ for which

$$\forall z \in J\ ,\ \operatorname{supp} \psi(\cdot, z) \subset [a, b]\ ,$$

where we have set $\psi(\cdot, z) : x \in I \to \psi(x, z) \in \mathbb{C}$. Consider a distribution $T \in \mathscr{D}'(I)$, and define

$$g(z) \stackrel{\text{def}}{=} \langle T, \psi(\cdot, z)\rangle\ ,\ z \in J\ .$$

Then $g \in \mathcal{C}^\infty(J)$, and

$$\forall z \in J\ ,\ g'(z) = \left\langle T, \frac{\partial \psi}{\partial z}(\cdot, z)\right\rangle\ .$$

It is important to realize that the distribution T only acts on the variable $x \in I$, not on the z variable; the latter just parametrizes a family of test functions $\{\psi(\cdot, z)\, ;\, z \in J\}$.

Proof.— Let us first prove that g is differentiable on J. We calculate, for $z \in J$ and $h \neq 0$ small enough,

$$\frac{g(z+h) - g(z)}{h} = \left\langle T, \frac{\psi(., z+h) - \psi(., z)}{h}\right\rangle\ .$$

Since $\psi \in \mathcal{C}^\infty(I \times J)$, we observe the pointwise convergence

$$\forall x \in I\ ,\ \frac{\psi(x, z+h) - \psi(x, z)}{h} \underset{h \to 0}{\longrightarrow} \frac{\partial \psi}{\partial z}(x, z)\ .$$

We claim that the convergence actually holds in the sense of $\mathscr{D}(I)$. Indeed, we already know that the left hand side is a smooth function of x which is supported in $[a, b]$. We claim that it converges *uniformly on* I to the right hand side. Indeed,

$$\frac{\psi(x, z+h) - \psi(x, z)}{h} = \int_0^1 \frac{\partial \psi}{\partial z}(x, z+th)\, dt\ .$$

Fix $\varepsilon_0 > 0$ such that $|h| \leq \varepsilon_0$. Then the continuous function $\frac{\partial \psi}{\partial z}$ is uniformly continuous on the compact subset $[a, b] \times [z - \varepsilon_0, z + \varepsilon_0]$. Consequently,

$$\sup_{x \in [a,b]} \left| \int_0^1 \frac{\partial \psi}{\partial z}(x, z+th)\, dt - \frac{\partial \psi}{\partial z}(x, z) \right| \underset{h \to 0}{\longrightarrow} 0\ .$$

The same argument holds for higher derivative in x:

$$\frac{\partial_x^k \psi(x, z+h) - \partial_x^k \psi(x, z)}{h} \underset{h \to 0}{\longrightarrow} \partial_z \partial_x^k \psi(x, z)\,,$$

uniformly for $x \in I$. We have thus proved that $\frac{\psi(\cdot, z+h) - \psi(\cdot, z)}{h}$ converges in $\mathscr{D}(I)$ towards $\partial_z \psi(\cdot, z)$.

Using the continuity of T on $\mathscr{D}(I)$, we infer

$$\frac{g(z+h) - g(z)}{h} = \left\langle T, \frac{\psi(\cdot, z+h) - \psi(\cdot, z)}{h} \right\rangle \underset{h \to 0}{\longrightarrow} \left\langle T, \frac{\partial \psi}{\partial z}(\cdot, z) \right\rangle\ .$$

This shows that g is differentiable, with the claimed formula for $g'(z)$. Finally, by applying this result iteratively, one easily proves by induction on n that g is n times derivable, with

$$g^{(n)}(z) = \left\langle T, \frac{\partial^n \psi}{\partial z^n}(\cdot, z) \right\rangle\ .$$

□ Let us now turn to the integration over an auxiliary parameter.

Proposition 1.9.6 (Integration under the bracket) Let I, J be open intervals, $\psi : I \times J \to \mathbb{C}$ be a $\mathcal{C}^\infty$ function such that there exists segments $[a, b] \subset I$ and $[c, d] \subset J$ such that

$$\operatorname{supp} \psi \subset [a, b] \times [c, d]\,.$$

Then

$$\int_J \langle T, \psi(\cdot, z) \rangle\ dz = \left\langle T, \int_J \psi(\cdot, z)\, dz \right\rangle.$$

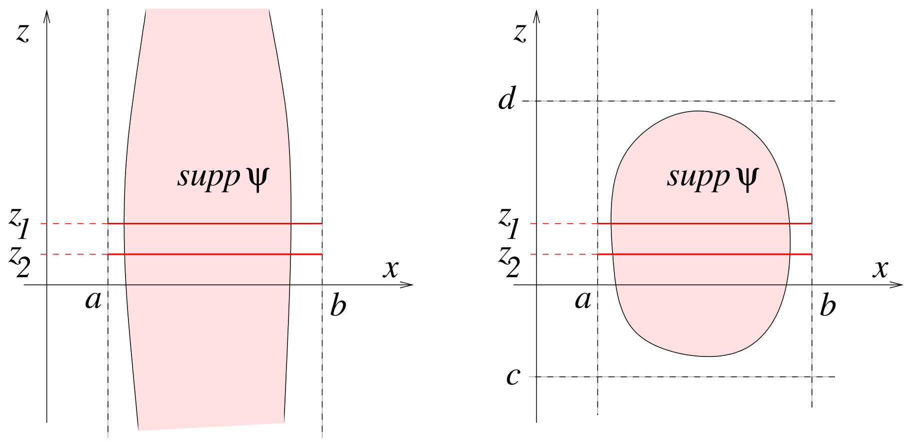


Figure 1.7: Supports of the parameter dependent test function $\psi(\cdot, x)$ for differentiation and integration under the bracket.

Proof.— Consider the function

$$\tilde{\psi}(x,z) = \int_{t\in J, t<z} \psi(x,t)\,dt\ .$$

Then $\tilde{\psi}$ satisfies the assumptions of Proposition 1.9.5, and

$$\frac{\partial\tilde{\psi}}{\partial z}(x,z) = \psi(x,z)\ .$$

Applying this proposition leads to

$$\frac{d}{dz}\langle T, \tilde{\psi}(\cdot,z)\rangle = \langle T, \psi(\cdot,z)\rangle\ .$$

We integrate both sides on J. This gives

$$\langle T, \tilde{\psi}(\cdot,d)\rangle - \langle T, \tilde{\psi}(\cdot,c)\rangle = \int_J \langle T, \psi(\cdot,z)\rangle\,dz\ .$$

In view of the assumptions on the support of ψ, we have $\tilde{\psi}(\cdot,d) = \int_J \psi(\cdot,z)\,dz$, $\tilde{\psi}(\cdot,c) = 0$, so that

$$\langle T, \tilde{\psi}(\cdot,d)\rangle = \left\langle T, \int_J \psi(\cdot,z)\,dz \right\rangle\ , \qquad \langle T, \tilde{\psi}(\cdot,c)\rangle = 0\ .$$

□

1.9.3 Convolution and regularisation

In this section, where we always take $I = \mathbb{R}$, we generalise the convolution of an L^1 function in with a test function, described for instance in Lemma 2.3.18, to the case of an arbitrary distribution. These convolutions will allow us to regularize arbitrary distributions.

We recall the case of a function $f \in L^1_{\text{loc}}(\mathbb{R})$ and a test function $\varphi \in \mathscr{D}(\mathbb{R})$. Their convolution is defined by the following smooth function:

$$\forall x \in I, \qquad f * \varphi(x) = \int_{\mathbb{R}} f(y)\, \varphi(x-y)\, dy\,.$$

Now, the right hand side can be written as the distributional bracket $f * \varphi(x) = \langle T_f, \varphi(x - \cdot)\rangle$, where $\varphi(x - \cdot)$ denotes the test function $y \mapsto \varphi(x-y)$.

This expression of $f * \varphi$ calls for a generalization to other types of distributions.

Definition 1.9.7 Let $T \in \mathscr{D}'(\mathbb{R})$ and $\varphi \in \mathscr{D}(\mathbb{R})$. For every $x \in \mathbb{R}$. Using the above notation, we define the convolution of T with φ by the following function:

$$\forall x \in \mathbb{R}, \qquad T * \varphi(x) = \langle T, \varphi(x - \cdot)\rangle\ .$$

Let us study some properties of this convolution. What is the regularity of $T * \varphi$?

Proposition 1.9.8 For every $(T, \varphi) \in \mathscr{D}'(\mathbb{R}) \times \mathscr{D}(\mathbb{R})$, $T * \varphi \in \mathcal{C}^\infty(\mathbb{R})$, and

$$(T * \varphi)' = T * \varphi' = T' * \varphi\ .$$

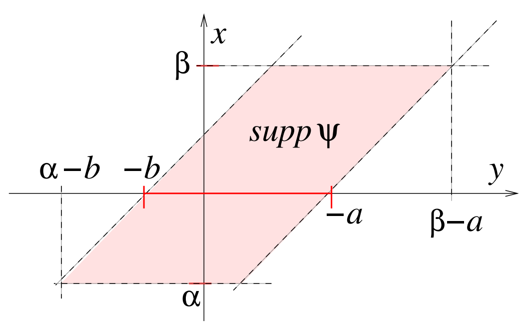


Figure 1.8: The function $\psi(y, x)$ used in the proof of Prop. 1.9.8.

Proof.— It is enough to prove that $T*\varphi$ is smooth on every finite open interval $]\alpha, \beta[$. If supp $\varphi \subset [a, b]$ and $x \in]\alpha, \beta[$, we have supp $\varphi(x - \cdot) \subset [\alpha - b, \beta - a]$, so that the assumptions of Proposition 1.9.5 are fulfilled with $I = \mathbb{R}$,$J =]\alpha, \beta[$ and $\psi(y, x) = \varphi(x - y)$ (notice that, here, x is the auxiliary parameter, while y is the variable on which T is acting).

Applying Proposition 1.9.5, we find that $T * \varphi \in \mathcal{C}^\infty(\mathbb{R})$, and

$$\frac{d}{dx}(T * \varphi)(x) = \langle T, \partial_x \varphi(x - \cdot)\rangle = \langle T, \varphi'(x - \cdot)\rangle = T * \varphi'(x)\ .$$

Furthermore, we have

$$\varphi'(x-y) = -\partial_y \varphi(x-y) ,$$

so that

$$(T*\varphi)'(x) = \langle T, -\partial_y \varphi(x-\cdot)\rangle = \langle T', \varphi(x-\cdot)\rangle = T' * \varphi(x) .$$

□

The convolution by a test function thus transforms T into a smooth function $T*\varphi \in C^\infty(\mathbb{R})$. By playing with the test function, it gives us a simple way to approximate the original distribution by a sequence of smooth functions, in other words, to regularise this distribution.

This regularisation works in the same way as in the case of L^1 functions (Lemma 2.3.18). Let $\rho \in \mathcal{C}_0^\infty(\mathbb{R})$ with

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

For every $\varepsilon > 0$, we set the rescaled function

$$\rho_\varepsilon(x) = \frac{1}{\varepsilon}\rho\left(\frac{x}{\varepsilon}\right) .$$

It is well known that, if $\varphi \in \mathcal{C}_0^\infty(\mathbb{R})$,

$$\rho_\varepsilon * \varphi \underset{\varepsilon\to 0}{\longrightarrow} \varphi \qquad \text{in } \mathscr{D}(\mathbb{R}).$$

Considering our distribution $T \in \mathscr{D}'(\mathbb{R})$, we set the function $f_\varepsilon = T * \rho_\varepsilon \in C^\infty(\mathbb{R})$.

Proposition 1.9.9 For every $T \in \mathscr{D}'(\mathbb{R})$,

$$T_{f_\varepsilon} \underset{\varepsilon\to 0}{\longrightarrow} T .$$

Proof.— We have to prove that

$$\forall \varphi \in \mathscr{D}(\mathbb{R}) ,\ \int_{\mathbb{R}} f_\varepsilon(x)\varphi(x)\, dx \underset{\varepsilon\to 0}{\longrightarrow} \langle T, \varphi\rangle . \tag{1.9.15}$$

As usual, the strategy is to "transfer the computations to the test function side". We notice that

$$\int_{\mathbb{R}} f_\varepsilon(x)\varphi(x)\, dx = \int_{\mathbb{R}} \langle T, \rho_\varepsilon(x-.)\rangle \varphi(x)\, dx = \int_{\mathbb{R}} \langle T, \rho_\varepsilon(x-.)\varphi(x)\rangle\, dx .$$

We write the test function on the right as a parameter-dependent function

$$\psi(y,x) \overset{\text{def}}{=} \rho_\varepsilon(x-y)\varphi(x) ,$$

where y is the variable on which T is acting, while x is the auxiliary parameter. This function ψ satisfies the assumption of Proposition 1.9.6 on $\mathbb{R}\times\mathbb{R}$. Indeed, if supp $\rho \subset [-C, C]$, supp $\varphi \subset [a,b]$, then

$$\text{supp}\, \psi \subset [a-\varepsilon C, b+\varepsilon C] \times [a,b] ,\ \psi(y,x) = 0 .$$

Applying Proposition 1.9.6, we obtain

$$\int_{\mathbb{R}} \langle T, \rho_\varepsilon(x - .)\varphi(x)\rangle \, dx = \left\langle T, \int_{\mathbb{R}} \rho_\varepsilon(x - .)\varphi(x) \, dx \right\rangle = \langle T, \tilde{\rho}_\varepsilon * \varphi\rangle \, ,$$

where $\tilde{\rho}(z) = \rho(-z)$. We have thus managed to transfer the convolution from the distribution side, to the test function side. Since $\tilde{\rho}$ satisfies the same assumptions as ρ, we have $\tilde{\rho}_\varepsilon * \varphi \underset{\varepsilon\to 0}{\longrightarrow} \varphi$ in $\mathscr{D}(\mathbb{R})$, and consequently

$$\int_{\mathbb{R}} f_\varepsilon(x)\varphi(x) \, dx \underset{\varepsilon\to 0}{\longrightarrow} \langle T, \varphi\rangle \ .$$

□

1.9.4 Positive distributions and increasing functions

In this section we define, and study the interesting case of positive distributions. Quite unexpectedly, the positivity assumption constrains quite much the structure and singularities of those distributions.

Definition 1.9.10 A distribution T on I is said positive (we note $T \geq 0$) if, for every $\varphi \in \mathscr{D}(I)$ valued in $\mathbb{R}_+$, we have $\langle T, \varphi\rangle \in \mathbb{R}_+$.

For instance, it is easy to check that, if $f \in L^1_{\mathrm{loc}}(I)$, $T_f \geq 0$ if and only $f \geq 0$ a.e. Another example is $T = c\delta_{x_0}$ with $x_0 \in I$ and $c \geq 0$. On the other hand, whatever the value of $c \neq 0$ is, it is clear that $c\delta'_{x_0}$ cannot be positive. In fact, we have the following general result.

Proposition 1.9.11 If T is a positive distribution, then T has order 0, namely

$$\forall [a, b] \subset I \ , \ \exists C > 0 \ , \ \forall \varphi \in \mathscr{D}(I), \mathrm{supp}\, \varphi \subset [a, b] \Rightarrow |\langle T, \varphi\rangle| \leq C\|\varphi\|_\infty \ .$$

Proof.— Let $[a, b] \subset I$ and $\chi \in \mathscr{D}(I)$ be a plateau function on $[a, b]$. Let $\varphi \in \mathscr{D}(I)$ be real valued and such that $\mathrm{supp}\, \varphi \subset [a, b]$. Then φ is bounded on both sides by the following functions:

$$-\chi\|\varphi\|_\infty \leq \varphi \leq \chi\|\varphi\|_\infty \ .$$

The positivity of T implies that its action on real valued functions gives a real value, and preserves the ordering between different test functions. In our case, $\langle T, \varphi\rangle \in \mathbb{R}$ and

$$-\langle T, \chi\rangle\|\varphi\|_\infty \leq \langle T, \varphi\rangle \leq \langle T, \chi\rangle\|\varphi\|_\infty \ ,$$

which implies

$$|\langle T, \varphi\rangle| \leq \langle T, \chi\rangle\|\varphi\|_\infty \ .$$

If φ is complex valued with $\operatorname{supp}\varphi \subset [a,b]$, we decompose

$$\varphi = \mathrm{Re}(\varphi) + i\mathrm{Im}(\varphi)\,,$$

and we conclude that

$$|\langle T,\varphi\rangle| \le 2\langle T,\chi\rangle \|\varphi\|_\infty\,.$$

The distribution T is thus of order 0. □

Remark 1.9.12 Since any compactly supported (positive) continuous function on I can be approximated uniformly by a sequence of (positive) test functions supported in a fixed segment of I, the above proposition implies that any positive distribution T can be continuously extended into a positive linear form on compactly supported continuous functions. Hence, according to the Riesz representation theorem, there exists a positive Borel measure μ on I, finite on segments (i.e., locally finite), such that

$$\forall\varphi \in \mathscr{D}(I)\ ,\ \langle T,\varphi\rangle = \int_I \varphi(x)\,d\mu(x)\ .$$

Let us come to the main result of this paragraph, which characterises increasing functions. For simplicity, we state and prove this result on $\mathbb{R}$, but a similar result holds on any open interval. Recall that an increasing function on $\mathbb{R}$ is a function $f:\mathbb{R}\to\mathbb{R}$ such that

$$\forall x \ge y\ ,\quad f(x) \ge f(y)\ .$$

(sometimes such functions are called nondecreasing functions).

Theorem 1.9.13 If $f:\mathbb{R}\to\mathbb{R}$ is an increasing function, then $T_f' \ge 0$. Conversely, if $T\in\mathscr{D}'(\mathbb{R})$ is such that $T'\ge 0$ and $\langle T,\varphi\rangle \in \mathbb{R}$ for every real valued test function φ, then there exists an increasing function f on $\mathbb{R}$ such that $T = T_f$.

Proof.— Let $f:\mathbb{R}\to\mathbb{R}$ be an increasing function. Note that f is bounded on every segment, hence it is locally integrable, and it makes sense to consider T_f. Let $\varphi\in\mathscr{D}(\mathbb{R})$, valued in $\mathbb{R}_+$. Let us calculate

$$\begin{aligned}
\langle T_f',\varphi\rangle &= -\int_{\mathbb{R}} f(x)\varphi'(x)\,dx = -\int_{\mathbb{R}} f(x)\,\lim_{h\to 0}\frac{\varphi(x+h)-\varphi(x)}{h}\,dx\\
&= -\lim_{h\to 0}\int_{\mathbb{R}} f(x)\,\frac{\varphi(x+h)-\varphi(x)}{h}\,dx\\
&= -\lim_{h\to 0}\int_{\mathbb{R}} \frac{f(y-h)-f(y)}{h}\varphi(y)\,dy\ ,
\end{aligned}$$

by the same arguments as in subsection 1.9.1. The monotonicity of f implies $\langle T_f',\varphi\rangle \ge 0$, hence T_f' is a positive distribution.

Conversely, let $T \in \mathscr{D}'(\mathbb{R})$ such that $T' \geq 0$. Choose $\rho \in \mathscr{D}(\mathbb{R})$, supported in $\mathbb{R}_+$, satisfying

$$\rho \geq 0 \ , \quad \int_{\mathbb{R}} \rho(x)\, dx = 1.$$

For $\varepsilon \in]0,1[$ we rescale ρ into ρ_ε like in (1.3.3) and recall that, from Proposition 1.9.9, the smooth function $f_\varepsilon = T * \rho_\varepsilon$ satisfies $T_{f_\varepsilon} \to T$ in $\mathscr{D}'(\mathbb{R})$ when $\varepsilon \to 0$. Notice moreover that, since T takes real values on real valued test functions, f_ε is real valued. Our strategy is to prove that, for an appropriate choice of ρ, f_ε is an increasing function and pointwise converges to a function f, with local dominated convergence. This will imply that $f_e \to f$ in L^1_{loc}, hence $T_{f_\varepsilon} \to T_f$, and finally $T = T_f$.

So let us first prove that f_ε is increasing. We know that

$$f'_\varepsilon = T' * \rho_\varepsilon$$

which is ≥ 0 because $T' \geq 0$ and $\rho_\varepsilon \geq 0$. Therefore f_ε is an increasing function on $\mathbb{R}$.
We then claim that for each x, $f_\varepsilon(x)$ is decreasing w.r.t. ε, equivalently, that $f_\varepsilon(x)$ increases as $\varepsilon \searrow 0$. In fact, applying Proposition 1.9.5 of derivation under the bracket, and considering f_ε as depending of the parameter $\varepsilon \in]0,1[$, we see that $f_\varepsilon(x)$ is a $\mathcal{C}^\infty$ function of ε, and that

$$\frac{d}{d\varepsilon} f_\varepsilon(x) = \langle T, \frac{\partial}{\partial \varepsilon} \rho_\varepsilon(x - \cdot) \rangle \ .$$

Observe that

$$\begin{aligned}
\frac{\partial}{\partial \varepsilon} \rho_\varepsilon(x-y) &= \frac{\partial}{\partial \varepsilon} \left[\frac{1}{\varepsilon} \rho\left(\frac{x-y}{\varepsilon} \right) \right] \\
&= -\frac{1}{\varepsilon^2} \left[\rho\left(\frac{x-y}{\varepsilon} \right) + \frac{x-y}{\varepsilon} \rho'\left(\frac{x-y}{\varepsilon} \right) \right] \\
&= \frac{1}{\varepsilon^2} \frac{\partial}{\partial y} \left[(x-y) \rho\left(\frac{x-y}{\varepsilon} \right) \right]
\end{aligned}$$

Consequently,

$$\frac{d}{d\varepsilon} f_\varepsilon(x) = -\frac{1}{\varepsilon} \left\langle T', \frac{(x-\cdot)}{\varepsilon} \rho\left(\frac{x-\cdot}{\varepsilon} \right) \right\rangle \leq 0 \ ,$$

because $T' \geq 0$ and $z\rho(z) \geq 0$ in view of the support of ρ.
Now we claim that, as $\varepsilon \searrow 0$, for every $x \in \mathbb{R}$, $f_\varepsilon(x)$ is bounded from above. Indeed, consider $\psi \in \mathscr{D}(I)$, valued in $\mathbb{R}_+$, supported in $[x, +\infty[$, and such that $\int_{\mathbb{R}} \psi(y)\, dy = 1$. Then, because $f_\varepsilon(y)$ is an increasing function of y,

$$f_\varepsilon(x) = f_\varepsilon(x) \int_{\mathbb{R}} \psi(y)\, dy \leq \int_{\mathbb{R}} f_\varepsilon(y)\, \psi(y)\, dy \underset{\varepsilon \to 0}{\longrightarrow} \langle T, \psi \rangle$$

because of Proposition 1.9.9. Since $f_\varepsilon(x)$ is increasing as ε decreases to 0, we conclude that

$$f_\varepsilon(x) \underset{\varepsilon \to 0}{\longrightarrow} f(x)$$

where f is a function on $\mathbb{R}$. Furthermore, if $x \in [a,b] \subset \mathbb{R}$, we have

$$f(x) \geq f_\varepsilon(x) \geq f_\varepsilon(a) \underset{\varepsilon \to 0}{\longrightarrow} f(a) \ ,$$

so the function f is increasing on $\mathbb{R}$, and thus automatically in L^1_{loc}. From this pointwise convergence, we can apply the dominated convergence theorem and conclude that, for every $\varphi \in \mathscr{D}(\mathbb{R})$,

$$\int_{\mathbb{R}} f_\varepsilon(x)\varphi(x)\,dx \underset{\varepsilon\to 0}{\longrightarrow} \int_{\mathbb{R}} f(x)\varphi(x)\,dx \ .$$

Summing up, applying again Proposition 1.9.9, we have proved that

$$\langle T, \varphi\rangle = \int_{\mathbb{R}} f(x)\varphi(x)\,dx \ ,$$

hence $T = T_f$. □

Remark 1.9.14 The first point of Theorem 1.9.13 says that, if f is increasing, T_f' is a positive distribution, hence is given by a positive Borel measure, locally finite. This measure is called the *Stieltjes measure* associated to f, and the formula for the integral of a compactly supported continuous function with respect to this measure is an extension of the definition of the usual integral of a continuous function by using Riemann sums. More precisely, if $\operatorname{supp}\varphi \subset [a,b]$,

$$\langle T_f', \varphi\rangle = \lim_{N\to\infty} \sum_{j=0}^{N-1} \varphi\left(a + j\frac{b-a}{N}\right)\left[f\left(a+(j+1)\frac{b-a}{N}\right) - f\left(a + j\frac{b-a}{N}\right)\right] \ .$$

For instance, if $f : \mathbb{R} \to \mathbb{R}$ is the cumulative distribution function of a random variable on $\mathbb{R}$, then T_f' is nothing but the law of this random variable.

1.9.5 The structure of distributions

In this section we show that distributions of finite order on $\mathbb{R}$ can be represented as the derivatives of functions.

Recall that, by Proposition 1.4.10, every distribution T admits a primitive distribution S. Furthermore, as can be observed from the proof of 1.4.10, if T is real, then one can choose S to be real as well. Now, if T is positive, then S satisfies the assumption of Theorem 1.9.13, so there exists an increasing function f such that $T_f = S$, hence $T = T_f'$. In other words, any positive distribution is the derivative of T_f, where f is an increasing (hence L^∞_{loc}) function. The next theorem shows that this property extends to any distribution of order 0.

Theorem 1.9.15 Any distribution T of order 0 on $\mathbb{R}$ is of the form $T = T_f'$, where $f \in L^\infty_{\text{loc}}(\mathbb{R})$.

Proof.— Let us recall the proof of Proposition 1.4.10, namely the construction of a primitive to a distribution S. Fix a function $\chi \in \mathscr{D}(\mathbb{R})$ supported in $]-1,1[$ and satisfying $\int_{\mathbb{R}} \chi(y)\,dy = 1$. We use this auxiliary function to "substract the mass" of test functions, as in the proof of Prop. 1.4.10. For every $\psi \in \mathscr{D}(\mathbb{R})$, we set

$$P(\psi)(x) = \int_{-\infty}^{x} \left[\psi(t) - \chi(t) \int_{\mathbb{R}} \psi(y)\, dy\right] dt \ .$$

Since $\psi - \chi \int \psi$ is massless, $P(\psi) \in \mathscr{D}(\mathbb{R})$ and, if supp$(\psi) \subset]-n, n[$ for some $n \geq 1$, then supp$(P(\psi)) \subset]-n, n[$ as well. Since $P(\varphi') = \varphi$, we infer that the distribution S defined by

$$\langle S, \psi \rangle = -\langle T, P(\psi) \rangle$$

satisfies $S' = T$.

Now, let us use the assumptioin that our distribution T is of order 0: for every $n \geq 1$, there exists $C_n > 0$ such that

$$\forall \varphi \in \mathscr{D}(]-n, n[) \ , \ |\langle T, \varphi \rangle| \leq C_n \|\varphi\|_\infty \ .$$

Consequently, if $\psi \in \mathscr{D}(]-n, n[)$,

$$|\langle S, \psi \rangle| = |\langle T, P(\psi) \rangle| \leq C_n \|P(\psi)\|_\infty \leq C_n B \|\psi\|_{L^1} \ .$$

Arguing as in proof of Theorem 1.9.2, we infer that the linear form $S_{|]-n,n[}$ acting on $\mathscr{D}(]-n, n[)$ can be continuously extended to functions $\varphi \in L^1(]-n, n[)$; since the dual of L^1 is L^∞, this linear form is represented by a function $f_n \in L^\infty(]-n, n[)$:

$$S_{|]-n,n[} = T_{f_n} \ .$$

Therefore we have constructed a sequence $(f_n)_{n \geq 1}$ of functions $f_n \in L^\infty(]-n, n[)$ such that

$$T_{f_{n+1}|]-n,n[} = (S_{|]-n-1,n+1[})_{|]-n,n[} = S_{|]-n,n[} = T_{f_n} \ .$$

This implies that $f_{n+1|]-n,n[} = f_n$. As a result, there exists $f \in L^\infty_{\text{loc}}(\mathbb{R})$ such that

$$\forall n \geq 1 \ , \ f_{|]-n,n[} = f_n \ .$$

Since

$$S_{|]-n,n[} = T_{f_{|]-n,n[}} \quad \text{for every } n \geq 1 \, ,$$

we conclude that $S = T_f$, and $T = S' = T_f'$. □

Let us now consider the case of distributions of finite order. We recall (see Def. 1.3.2) that for $m \in \mathbb{N}$, $T \in \mathscr{D}'(I)$ is of order $\leq m$ if

$$\forall [a, b] \subset I \ , \ \exists C > 0 \ , \ \forall \varphi \in \mathscr{D}(I) \ , \ \text{supp}(\varphi) \subset [a, b] \Longrightarrow |\langle T, \varphi \rangle| \leq C \sum_{k=0}^{m} \|\varphi^{(k)}\|_\infty \ .$$

Theorem 1.9.16 (Structure of distributions on $\mathbb{R}$) i) If $T \in \mathscr{D}'(\mathbb{R})$ is of order $\leq m$, there exists $f \in L^\infty_{\text{loc}}(\mathbb{R})$ such that $T = T_f^{(m+1)}$.

ii) If $T \in \mathscr{D}'(\mathbb{R})$ is arbitrary, there exists a sequence $(f_n)_{n\geq 1}$ in $L^\infty(\mathbb{R})$, such that on every segment of $\mathbb{R}$, f_n vanishes for n large enough, and

$$T = \sum_{n=1}^{\infty} T_{f_n}^{(n)} \ .$$

Proof.— The statement $i)$ follows from an induction argument on m, based on Theorem 1.9.15 and on the following

Lemma 1.9.17 If $T \in \mathscr{D}'(\mathbb{R})$ is of order $\leq m$ with $m \geq 1$, there exists $S \in \mathscr{D}'(\mathbb{R})$ of order $\leq m-1$ such that $S' = T$. In other words, taking the primitive of a distribution reduces its order by one.

The proof of this lemma is straightforward, taking into account the formula $\langle S, \psi\rangle = -\langle T, P(\psi)\rangle$ and the fact that, for every $m \geq 1$ and $\psi \in \mathscr{D}(]-n,n[)$,

$$\sup_{0\leq k\leq m} \|P(\psi)^{(k)}\|_\infty \leq B\|\psi\|_{L^1} + \sup_{0\leq k\leq m-1} \|\psi^{(k)}\|_\infty \leq B_n \sup_{0\leq k\leq m-1} \|\psi^{(k)}\|_\infty \ .$$

Let us prove the statement $ii)$. For every integer $j \geq 1$, there exists an integer m_j such that $T_{|]-j-1/2,j+1/2[}$ is of order $\leq m_j$. Furthermore, we may impose without loss of generality that $m_{j+1} > m_j$.

We are going to construct a sequence $(g_j)_{j\geq 1}$ of L^∞ functions on $\mathbb{R}$ such that, for every $j \geq 1$, supp$(T_{g_j}) \subset [-j,-j+1] \cup [j-1,j]$, and

$$\left(T - \sum_{\ell=1}^{j} T_{g_\ell}^{(m_\ell+1)}\right)_{|]-j,j[} = 0 \ .$$

Let us first construct g_1. Since T is of order $\leq m_1$ on $]-3/2,3/2[$, the statement $i)$ (adapted to an interval) implies that there exists $h_1 \in L^\infty_{\text{loc}}(]-3/2,3/2[)$ such that

$$T_{|]-3/2,3/2[} = T_{h_1}^{(m_1+1)} \ .$$

Then take $g_1 \stackrel{\text{def}}{=} \mathbb{1}_{]-1,1[}\, h_1 \in L^\infty(\mathbb{R})$; the corresponding distribution satisfies

$$\big(T - T_{g_1}^{(m_1+1)}\big)_{|]-1,1[} = 0\ , \quad \text{supp}(T_{g_1}) \subset [-1,1]\ , \quad (T - T_{g_1}^{(m_1+1)})_{|]-3/2,3/2[} \text{ is of order } \leq m_1\ .$$

Assuming $g_1, \ldots, g_j$ are constructed, let us construct g_{j+1}. Since the sequence m_j is increasing, the distribution

$$\tilde{T}_{j+1} \stackrel{\text{def}}{=} \big(T - \sum_{\ell=1}^{j} T_{g_\ell}^{(m_\ell+1)}\big)_{|]-j-3/2,j+3/2[}$$

is of order $\leq m_{j+1}+1$, like $T_{|]-j-3/2,j+3/2[}$, the terms in the sum being of smaller orders. Furthermore, the restriction of $\tilde{T}_{j+1}$ to $]-j,j[$ vanishes. Therefore, using again property $i)$, there exists $h_{j+1} \in L^\infty_{\text{loc}}(]-j-3/2,j+3/2[)$ such that

$$\tilde{T}_{j+1} = T_{h_{j+1}}^{(m_{j+1}+1)} .$$

Furthermore, the restriction of $T_{h_{j+1}}^{(m_{j+1}+1)}$ to $]-j,j[$ is 0: this implies from Corollary 1.4.9 that h_{j+1} coincides with a polynomial p_{j+1} on $]-j,j[$. We may "correct" h_{j+1} into

$$g_{j+1} \stackrel{\text{def}}{=} 1\!\!1_{|]-j-1,j+1[}\,(h_{j+1} - p_{j+1}) .$$

Then $g_{j+1} \in L^\infty(\mathbb{R})$, $T_{g_{j+1}}$ is supported in $[-j-1,-j] \cup [j,j+1]$, and

$$\big(\tilde{T}_{j+1} - T_{g_{j+1}}^{(m_{j+1}+1)}\big)_{|]-j-1,j+1[} = 0 .$$

Coming back to the expression of $\tilde{T}_{j+1}$, we have checked the induction assumption at rank $j+1$, so that the sequence $(g_j)_{j\geq 1}$ is constructed by induction on j. In view of the properties of the supports of T_{g_j}, for every j, only the terms of rank $\ell \leq j$ of the series

$$\sum_{\ell=1}^{\infty} T_{g_\ell}^{(m_\ell+1)}$$

have a nonzero restriction to $]-j,j[$. This proves that this series is convergent in $\mathscr{D}'(\mathbb{R})$. Furthermore, by the construction of the g_j, the sum of this series coincides with T on every interval $]-j,j[$, therefore

$$T = \sum_{\ell=1}^{\infty} T_{g_\ell}^{(m_\ell+1)} .$$

For the moment, the sequence of the order of derivatives $(m_\ell+1)_{\ell\geq 1}$ is a strictly increasing sequence of $\mathbb{N}$. To obtain the statement of the theorem, we reindex the sum to take all orders of derivation into account. Namely, if there exists $\ell \geq 1$ such that $n = m_\ell + 1$, we set $f_n \stackrel{\text{def}}{=} g_\ell$, while otherwise we set $f_n \stackrel{\text{def}}{=} 0$. We then end up with the sum

$$T = \sum_{n=1}^{\infty} T_{f_n}^{(n)} .$$

□

Chapter 2

Distributions in several variables

In this chapter we extend the notion of distribution to the Euclidean space $\mathbb{R}^d$, or more generally to an open subset $\Omega \subset \mathbb{R}^d$. While most definitions and properties will mimic those in one dimension, we will exhibit new examples of interesting distributions (like distributions supported on submanifolds, e.g. surface measures in $\mathbb{R}^3$), and our results will carry a certain geometric flavour. The applications to partial differential equations will be more "impressive" than in 1D.

We start this chapter by reviewing differential calculus on $\mathbb{R}^d$, which will set our notations.

2.1 A brief review differential calculus in several variables

2.1.1 Scalar product, norm, distance, topology

From now on we shall work on the vector space $\mathbb{R}^d$, made of points

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}$$

with $x_1, \ldots, x_d \in \mathbb{R}$. We denote by $(e_1, \ldots, e_d)$ the canonical basis, so that

$$x = \sum_{j=1}^{d} x_j e_j \ .$$

This canonical basis is an orthonormal basis for the canonical scalar product on $\mathbb{R}^d$,

$$x \cdot y = \sum_{j=1}^{d} x_j y_j \ ,$$

defining the Euclidean norm

$$|x| = \sqrt{x \cdot x} = \left(\sum_{j=1}^{d} x_j^2 \right)^{\frac{1}{2}} \ .$$

This norm classically defines the distance

$$d(x,y) = |x-y| ,$$

and induces a topology, for which a set Ω is open if and only if for every $a \in \Omega$, there exists $r>0$ such that the ball $B(a,r) \subset \Omega$. Here, $B(a,r)$ denotes the open ball $\{x \in \mathbb{R}^d ;\, d(x,a) < r\}$. As in every metric space, a compact subset $K \subset \mathbb{R}^d$ can be equivalently defined by the Borel–Lebesgue covering property, or by the Bolzano–Weierstrass extraction property for every sequence in K. Since we are on a finite dimensional vector space, compact subsets of $\mathbb{R}^d$ coincide with closed bounded subsets.

If F is a nonempty closed subset of $\mathbb{R}^d$, we shall often use the distance function to F,

$$d(x,F) \stackrel{\text{def}}{=} \inf_{z \in F} d(x,z) .$$

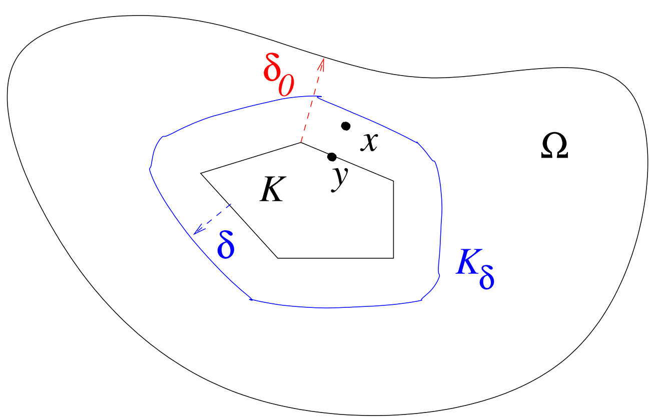
Notice that $d(x,F) = 0$ if and only if $x \in F$, and that, by the Bolzano–Weierstrass property for closed bounded subsets, this infimum is attained at some $z \in F$. Furthermore, the function $d(\cdot, F)$ is continuous on $\mathbb{R}^d$. In fact, by the triangle inequality, it is 1-Lipschitz continuous,

$$|d(x,F) - d(y,F)| \leq d(x,y) .$$

If K is a compact subset of $\mathbb{R}^d$ and $\delta > 0$, the set

$$K_\delta \stackrel{\text{def}}{=} \{x \in \mathbb{R}^d ; d(x,K) \leq \delta\}$$

is a compact subset containing K in its interior. The following lemma will be very useful.



Lemma 2.1.1 If K is a compact subset of an open subset $\Omega \subset \mathbb{R}^d$, then

$$\inf_{x \in K} d(x, \Omega^c) = \delta_0 > 0 .$$

For every $\delta \in]0, \delta_0[$, K_δ is contained in Ω.

Proof.— The first statement follow from the continuity of $d(\cdot,\Omega^c)$, which consequently attains its minimum on K, and from the assumption $K\cap\Omega^c=\emptyset$. The second statement is an elementary consequence of the triangle inequality, since, for every $z\in\Omega^c$, $y\in K$,

$$d(x,z)\geq |d(y,z)-d(x,y)|\,.$$

Hence, if $x\in K_\delta$, choosing $y\in K$ such that $d(x,y)=d(x,K)$, $d(x,\Omega^c)\geq\delta_0-\delta>0$. □ The boundary of Ω is denoted by $\partial\Omega=\overline{\Omega}\setminus\Omega$.

2.1.2 Partial derivatives, C^1 functions, differential, gradient

We denote by $\mathcal{C}^0(\Omega)$ the space of continuous functions $f:\Omega\to\mathbb{C}$ (note that $f(x)$ may explode when $x\to\partial\Omega$).

Definition 2.1.2 Given $f:\Omega\to\mathbb{C}$ and $a\in\Omega$, $j\in\{1,\ldots,d\}$, we say that f admits a j-th partial derivative at a if the function $t\mapsto f(a+te_j)$, locally defined for t in a neighbourhood of 0 in $\mathbb{R}$, has a derivative at $t=0$. We set

$$\frac{d}{dt}f(a+te_j)_{|t=0}=\frac{\partial f}{\partial x_j}(a)=\partial_j f(a)\ .$$

We say that f is a C^1 function on Ω if it admits a j–th partial derivative for every $j\in\{1,\ldots,d\}$ at every point $a\in\Omega$, and if the functions $\partial_j f$ are continuous on Ω. We denote by $\mathcal{C}^1(\Omega)$ the space of C^1 functions on Ω.

If $f\in\mathcal{C}^1(\Omega)$, one can prove that f is *differentiable* at every point a ; using the mean value theorem,

$$f(a+h)=f(a)+L_a(h)+o(|h|)\text{ as }h\to 0$$

where the increment in the direction h reads:

$$L_a(h)=\sum_{j=1}^d\partial_j f(a)h_j\ .$$

The linear map $L_a:\mathbb{R}^d\to\mathbb{C}$ is called the differential of f at the point a, and usually denoted by

$$L_a=d_af\ .$$

If f is real valued, d_af is a linear form on the Euclidean space $\mathbb{R}^d$, hence it can be represented by the scalar product with a vector, called the gradient of f at a, and denoted by $\nabla f(a)$:

$$d_af(h)=\nabla f(a)\cdot h\ ,\quad \nabla f(a)=\begin{pmatrix}\partial_1 f(a)\\ \vdots\\ \partial_d f(a)\end{pmatrix}\ .$$

Notice that the application $\nabla f:\Omega\to\mathbb{R}^d$ is continuous (since $f\in C^1$).

2.1.3 The chain rule

The chain rule allows to compute the differential of the composition of two C^1 functions. Its expression is a bit more complicated than in 1 dimension.

Proposition 2.1.3 Let $\Omega \subset \mathbb{R}^d$, $\Omega' \subset \mathbb{R}^p$ be open sets, and $\psi : \Omega \to \Omega'$, $\psi = (\psi_1, \psi_2, \ldots, \psi_p)$, a function of class $\mathcal{C}^1$.Let $f : \Omega' \to \mathbb{C}$ be a $\mathcal{C}^1$ function.

Then $f \circ \psi$ is $\mathcal{C}^1$ on Ω. Moreover, the differential of the composed function reads:

$$d_x(f \circ \psi) = \underbrace{d_{\psi(x)} f}_{\mathbb{C} \leftarrow \mathbb{R}^p} \circ \underbrace{d_x \psi}_{\mathbb{R}^p \leftarrow \mathbb{R}^d} = \underbrace{\nabla \psi(x)}_{\mathbb{R}^d \times \mathbb{R}^p} \cdot \underbrace{\nabla f(\psi(x))}_{\mathbb{R}^p \times \mathbb{C}},$$

or, for any component $j \in \{1, \ldots, n\}$,

$$\frac{\partial f \circ \psi}{\partial x_j}(x) = \sum_{k=1}^{p} \frac{\partial f}{\partial \psi_k}(\psi(x)) \frac{\partial \psi_k}{\partial x_j}(x) \, .$$

2.1.4 Higher order partial derivatives

More generally, for $m \geq 2$, we denote by $\mathcal{C}^m(\Omega)$ the vector space of functions $f \in C^1(\Omega)$ whose partial derivatives $\partial_1 f, \partial_2 f, \ldots, \partial_d f$ belong to $\mathcal{C}^{m-1}(\Omega)$. Moreover, $\mathcal{C}^\infty(\Omega)$ is the intersection of all $\mathcal{C}^m(\Omega)$.

For a general function, the order in which one computes repeated partial derivatives may matter; but this is not the case if we compute two derivatives of a $\mathcal{C}^2$ function:

Proposition 2.1.4 (Schwarz Lemma) If $f \in \mathcal{C}^2(\Omega)$, then, for any $j, k \in \{1, \ldots, d\}$,

$$\partial_j(\partial_k f) = \partial_k(\partial_j f)$$

In particular for $\mathcal{C}^\infty$ functions, one can compute partial derivatives of f in any order. It is therefore very convenient to use multi-indices, which only record how many times we differentiate in each direction, independently of the order in which we proceed.

2.1.5 Multiindices

Let $f \in \mathcal{C}^\infty(\Omega)$, and $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d) \in \mathbb{N}^d$ a multiindex. We denote $\partial^\alpha f$ the function

$$\partial^\alpha f = \partial_1^{\alpha_1} \partial_2^{\alpha_2} \cdots \partial_d^{\alpha_d} f = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_d^{\alpha_d}} f \, ,$$

where the number

$$|\alpha| \stackrel{\text{def}}{=} \alpha_1 + \alpha_2 + \cdots + \alpha_d,$$

is the order of the partial derivative, and it is called the length of α. The context usually avoids any confusion with the Euclidean norm ! We also set the factorial notation

$$\alpha! = \alpha_1!\alpha_2!\ldots\alpha_d!$$

and, for $\beta \in \mathbb{N}^d$ such that $\beta_j \leq \alpha_j$ for all j, which we will write $\beta \leq \alpha$, the multinomial coefficient

$$\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!\,(\alpha-\beta)!} = \binom{\alpha_1}{\beta_1}\binom{\alpha_2}{\beta_2}\cdots\binom{\alpha_d}{\beta_d}.$$

With these notations, the Leibniz formula for the derivatives of a product of two functions easily extends to the case of functions of several variables. Its proof is exactly the same.

Proposition 2.1.5 (Multidimensional Leibniz formula) Let f and g be functions in $\mathcal{C}^\infty(\Omega)$, and $\alpha \in \mathbb{N}^d$ a multiindex. We have

$$\partial^\alpha(fg) = \sum_{\beta\in\mathbb{N}^d,\ \beta\leq\alpha} \binom{\alpha}{\beta}\, \partial^\beta f\ \partial^{\alpha-\beta} g\,.$$

Proof.— We prove the result by induction over $|\alpha|$. If $|\alpha| = 1$, $\partial^\alpha = \partial_j$ for some $j \in \{1,\ldots,d\}$, and

$$\partial_j(fg) = (\partial_j f)g + f(\partial_j g),$$

which is the above formula. Suppose then that the formula is true for all multiindices of length $\leq m$. Let $\alpha \in \mathbb{N}^d$ such that $|\alpha| = m+1$. There exists $j \in \{1,\ldots,d\}$ and $\beta \in \mathbb{N}^d$ of length m such that

$$\alpha = \beta + 1_j,$$

where $1_j = (0,\ldots,0,1,0,\ldots,0)$ with a 1 as j-th coordinate. With these notations

$$\partial^\alpha(fg) = \partial^{\beta+1_j}(fg) = \partial^\beta(\partial_j(fg)) = \partial^\beta((\partial_j f)g) + \partial^\beta(f(\partial_j g)).$$

Since β is of length m, the induction assumption gives

$$\begin{aligned}
\partial^\alpha(fg) &= \sum_{\gamma\leq\beta}\binom{\beta}{\gamma}\partial^\gamma(\partial_j f)\ \partial^{\beta-\gamma} g + \sum_{\gamma\leq\beta}\binom{\beta}{\gamma}\partial^\gamma f\ \partial^{\beta-\gamma}(\partial_j g)\\
&= \sum_{\gamma\leq\beta}\binom{\beta}{\gamma}\partial^{\gamma+1_j} f\ \partial^{\beta-\gamma} g + \sum_{\gamma\leq\beta}\binom{\beta}{\gamma}\partial^\gamma f\ \partial^{\beta+1_j-\gamma} g\\
&= \sum_{\gamma\leq\beta}\binom{\beta}{\gamma}\partial^{\gamma+1_j} f\ \partial^{\alpha-(\gamma+1_j)} g + \sum_{\gamma\leq\beta}\binom{\beta}{\gamma}\partial^\gamma f\ \partial^{\alpha-\gamma} g
\end{aligned}$$

We change the multiindex in the first sum: $\gamma \leftarrow \gamma + 1_j$, and we get

$$\begin{aligned}\partial^\alpha(fg) &= \sum_{1_j\leq\gamma\leq\alpha}\binom{\beta}{\gamma-1_j}\partial^\gamma f\,\partial^{\alpha-\gamma}g + \sum_{\gamma\leq\beta}\binom{\beta}{\gamma}\partial^\gamma f\,\partial^{\alpha-\gamma}g\\ &= \sum_{\gamma\leq\beta,\ \gamma_j=0}\binom{\beta}{\gamma}\partial^\gamma f\,\partial^{\alpha-\gamma}g + \sum_{1_j\leq\gamma\leq\beta}\Big(\binom{\beta}{\gamma-1_j}+\binom{\beta}{\gamma}\Big)\partial^\gamma f\,\partial^{\alpha-\gamma}g + f\partial^\alpha g\\ &= \sum_{\gamma\leq\beta,\ \gamma_j=0}\binom{\beta}{\gamma}\partial^\gamma f\,\partial^{\alpha-\gamma}g + \sum_{1_j\leq\gamma\leq\beta}\binom{\beta+1_j}{\gamma}\partial^\gamma f\,\partial^{\alpha-\gamma}g + f\partial^\alpha g\,,\end{aligned}$$

where in the last line we used the generalization of Pascal's formula for multinomial coefficients. This last line is exactly the stated formula at the order α. □

We can continue the analogy with the 1 variable case: if $x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$, we denote by x^α the monomial

$$x^\alpha = x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_d^{\alpha_d}\,.$$

Then we can show, the same way as for the Leibniz formula, that for $x, y \in \mathbb{R}^d$ and $\alpha \in \mathbb{N}^d$, the following generalization of Newton's binomial formula holds:

$$(x+y)^\alpha = \sum_{\beta\leq\alpha}\binom{\alpha}{\beta}x^\beta y^{\alpha-\beta}.$$

With these notations, Taylor-Lagrange's formula can be written as concisely as in 1D.

Proposition 2.1.6 (Multidimensional Taylor formula) Let $f : \Omega \subset \mathbb{R}^d \to \mathbb{C}$ a function of class $\mathcal{C}^{m+1}$. Let $a, b \in \Omega$, such that the segment $[a, b]$ is included in Ω. We have

$$f(b) = \sum_{|\alpha|\leq m}\frac{(b-a)^\alpha}{\alpha!}\partial^\alpha f(a) + (m+1)\sum_{|\alpha|=m+1}\frac{(b-a)^\alpha}{\alpha!}\int_0^1(1-t)^m\partial^\alpha f(a+t(b-a))dt\,.$$

This is called the Taylor expansion of f at the point a, at the order m.

Proof.— We have already seen that if $\varphi : \mathbb{R} \to \mathbb{C}$ is smooth, the Taylor formula gives

$$\varphi(1) = \sum_{k=0}^m\frac{1}{k!}\varphi^{(k)}(0) + \frac{1}{m!}\int_0^1(1-s)^m\varphi^{(m+1)}(s)ds.$$

We shall use this result for the function $\varphi : t \mapsto f\big(a + t(b-a)\big)$. Notice that

$$\varphi'(t) = \sum_{j=1}^d(b-a)_j\,\partial_j f\big(a+t(b-a)\big)$$

and more generally that

$$\varphi^{(k)}(t) = \sum_{j_1,j_2,\dots,j_k=1}^{d} (b-a)_{j_1}\dots(b-a)_{j_k}(\partial_{j_1}\dots\partial_{j_k}f)\big(a+t(b-a)\big)\,.$$

This sum only contains terms of the form $(b-a)^\alpha \partial^\alpha f\big(a+t(b-a)\big)$ with $\alpha\in\mathbb{N}^d$ of length $|\alpha|=k$. Therefore we can write, for some coefficients $c_\alpha\in\mathbb{R}$,

$$\sum_{j_1,j_2,\dots,j_k=1}^{d} (b-a)_{j_1}\dots(b-a)_{j_k}\,\partial_{j_1}\dots\partial_{j_k}f\big(a+t(b-a)\big) = \sum_{|\alpha|=k} c_\alpha (b-a)^\alpha\,\partial^\alpha f\big(a+t(b-a)\big)\,.$$

Denoting $x=b-a$, this equality between two polynomials in x implies that the c_α are combinatorial factors:

$$c_\alpha = \#\{(j_1,j_2,\dots,j_k)\in\{1,\dots,d\}^k,\ x_1^{\alpha_1}x_2^{\alpha_2}\dots x_d^{\alpha_d} = x_{j_1}\dots x_{j_k}\}.$$

in words, c_α counts the ways to distribute α_1 indices $j=1$, α_2 indices $j=2$, etc, among $k=|\alpha|$ indices. It is a simple combinatorial problem, with solution:

$$c_\alpha = \binom{k}{\alpha_1}\binom{k-\alpha_1}{\alpha_2}\dots\binom{k-\alpha_1-\dots-\alpha_{n-1}}{\alpha_n} = \frac{k!}{\alpha!}.$$

Indeed, one has to first place α_1 indices $j=1$ among among k spots, then α_2 indices $j=2$ among the $k-\alpha_1$ remaining spots, etc.

Since $a+t(b-a)_{|t=0}=a$, and $a+t(b-a)_{|t=1}=b$, we obtain the stated formula. □

Exercise 2.1.7 Show that, for $k\in\mathbb{N}$ and $(x_1,x_2,\dots,x_d)\in\mathbb{R}^d$, we have

$$(x_1+x_2+\dots+x_d)^k = \sum_{|\alpha|=k}\frac{k!}{\alpha!}\,x^\alpha\,.$$

Applying the Taylor formula at order 0, we immediately obtain an extension of Hadamard's lemma in dimension d.

Corollary 2.1.8 (Hadamard's formula in dimension d) Let $f\in\mathcal{C}^{m+1}(\Omega)$, where Ω is a *convex* open subset of $\mathbb{R}^d$. If $a\in\Omega$ and $f(a)=0$, there exist d functions g_1, g_2,…g_d of class $\mathcal{C}^m(\Omega)$ such that

$$f(x) = \sum_{j=1}^{d}(x_j-a_j)\,g_j(x)\,.$$

Notice that the convexity assumption of Ω is necessary, because we need that the segment $[a,x]\subset\Omega$ for any $a,x\in\Omega$. In practice, we shall use this lemma when Ω is a ball.

2.2 Test functions

2.2.1 Definitions. Examples

If $f \in \mathcal{C}^0(\Omega)$, the support of f is the closure of $\{x \in \Omega, f(x) \neq 0\}$ for the topology induced on Ω. This is a closed subset of Ω, denoted by supp(f). If $m \in \mathbb{N} \cup \{\infty\}$, we denote by $\mathcal{C}_0^m(\Omega)$ the space of functions in $\mathcal{C}^m(\Omega)$ with a compact support. Notice that a compact set for the induced topology in Ω is also a compact subset of $\mathbb{R}^d$ contained in Ω. In particular, elements of

$$\mathscr{D}(\Omega) \stackrel{\text{def}}{=} \mathcal{C}_0^\infty(\Omega)$$

are called test functions on Ω. As in one space dimension, if $V \subset \Omega$ is an open subset, every element $\varphi \in \mathscr{D}(V)$ can be extended as an element $\underline{\varphi}$ of $\mathscr{D}(\Omega)$ by setting $\underline{\varphi} = 0$ in $\Omega \setminus V$. This allows to identify $\mathscr{D}(V)$ as the subspace of $\mathscr{D}(\Omega)$ defined by supp$(\varphi) \subset V$. We shall often make this implicit identification.

Proposition 2.2.1 (Cutoff functions) *i)* For every $a \in \mathbb{R}^d$, for every $r > 0$, there exists $\varphi \in \mathscr{D}(\mathbb{R}^d)$ such that $\varphi(x) \geq 0$ for every $x \in \mathbb{R}^d$ and supp$(\varphi) = \overline{B(a,r)}$.

ii) For every compact subset K of Ω, there exists $\chi \in \mathscr{D}(\Omega)$, valued in $[0,1]$, such that $\chi = 1$ on K. We call such χ a cutoff function ("fonction plateau") on K in Ω.

Proof.— The proof of the first statement is similar to its one dimensional analogue; it uses the radial symmetry of the ball. Recall that the function f defined by

$$f(t) = \begin{cases} \mathrm{e}^{-\frac{1}{t}} & \text{if } t > 0 \\ 0 & \text{if } t \leq 0 \end{cases}$$

is $\mathcal{C}^\infty$. Then the function φ defined by

$$\varphi(x) = f(r^2 - |x-a|^2)$$

satisfies the requirements.

The second statement requires a little more work. Let φ be as in the first statement with $a = 0$ and $r = 1$. Since $\varphi \geq 0$ and is not identically 0, its integral on $\mathbb{R}^d$ is > 0. Up to dividing φ by this number, we may assume that

$$\int_{\mathbb{R}^d} \varphi(x)\, dx = 1 ,$$

namely φ has a unit mass; it is smooth convolution kernel. For any $\delta \in]0,1]$, we rescale this kernel to

$$\varphi_\delta(x) = \frac{1}{\delta^d}\varphi\Big(\frac{x}{\delta}\Big)$$

Then, by Lemma 2.1.1, we may choose $\delta > 0$ such that $K_{2\delta} \subset \Omega$. Set

$$\chi(x) = \varphi_\delta * 1\!\!1_{K_\delta}(x) = \int_{K_\delta} \varphi\left(\frac{x-y}{\delta}\right) \frac{dy}{\delta^d} .$$

By the rule of derivation under the integral, $\chi \in \mathcal{C}^\infty(\mathbb{R}^d)$. Since $\varphi \geq 0$ and $1\!\!1_K \geq 0$, we have $\chi \geq 0$ and

$$\forall x \in \mathbb{R}^d, \qquad \chi(x) \leq \int_{\mathbb{R}^d} \varphi_\delta(x-y)\, dy = 1 \,, .$$

Hence χ is valued in $[0,1]$. If $x \in K$ and $y \in (K_\delta)^c$, then $|x-y| > \delta$, hence $(x-y) \notin$ supp(φ_δ), so the integrand vanishes at y. Consequently,

$$\forall x \in K\,, \quad \chi(x) = \int_{\mathbb{R}^d} \varphi_\delta(x-y)\, dy = 1\,.$$

Finally, if $x \notin K_{2\delta}$, $y \in K_\delta$, $|x-y| > \delta$, and the integrand identically vanishes, so that $\chi(x) = 0$. Hence the support of χ is included in $K_{2\delta}$, which is contained in the open set Ω, so that $\chi \in \mathcal{C}_0^\infty(\Omega)$. $\square$

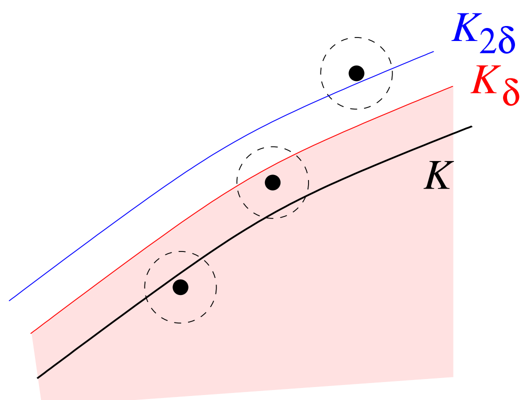


Figure 2.1: The cutoff function χ on a compact set K, obtained by convoluting $1\!\!1_{K_\delta}$ with a kernel φ_δ. Depending on the position x (black dot), the kernel is fully supported in K_δ, partially supported, or supported outside K_δ; accordingly, $\chi(x)$ smoothly drops from 1 to 0.

Remark 2.2.2 It may happen that we need cutoff functions on a compact neighborhood of K in Ω, for instance K_ε for $\varepsilon > 0$ small enough, according to Lemma 2.1.1. We shall call these functions *plateau functions on a neighbourhood of* K in Ω.

We now come to the contruction of partitions of unity, which turns out to be particularly useful in several variables, through the so–called *gluing principle*.

2.2.2 Partitions of unity

Proposition 2.2.3 (Partition of unity in Ω) Let $K \subset \Omega$ be a compact set covered by a finite collection $\Omega_1, \ldots, \Omega_n$ of open subsets of Ω. There exist $\chi_1 \in \mathcal{C}_0^\infty(\Omega_1), \ldots, \chi_n \in \mathcal{C}_0^\infty(\Omega_n)$, valued in $[0,1]$, such that

$$\chi_1 + \cdots + \chi_n = 1 \text{ on } K \ .$$

As a result, if $\varphi \in \mathcal{C}_0^\infty(\Omega)$ is supported in $\Omega_1 \cup \cdots \cup \Omega_n$, then one can find functions $\varphi_1 \in \mathcal{C}_0^\infty(\Omega_1), \ldots, \varphi_n \in \mathcal{C}_0^\infty(\Omega_n)$, such that

$$\varphi_1 + \ldots \varphi_n = \varphi \ .$$

Proof.— It is very similar to the one dimensional analogue. First, one proves the

Lemma 2.2.4 (Shrinking Lemma) If $K \subset \Omega_1 \cup \cdots \cup \Omega_n$, there exists open subsets $U_1, \ldots, U_n$ such that, for every $j \in \{1, \ldots, n\}$, $\overline{U}_j \subset \Omega_j$ is compact, and

$$K \subset U_1 \cup \cdots \cup U_n \ .$$

As in the one dimensional case, the proof proceeds by induction on $n \geq 1$. The case $n = 1$ follows from Lemma 2.1.1.

Once the shrinking lemma is proved, set

$$\chi_1 = \psi_1, \chi_2 = \psi_2(1 - \psi_1) \ , \ldots \ , \chi_n = \psi_n(1 - \psi_{n-1}) \ldots (1 - \psi_1) \ ,$$

where, for every $j \in \{1, \ldots, n\}$, ψ_j is a cutoff function on $\overline{U}_j$ in Ω_j.
Finally, the last assertion follows by writing $\varphi_j = \chi_j \varphi$, where the smooth partition $(\chi_j)_{j=1,\ldots,n}$ is associated wih the covering of supp(φ) by $\Omega_1, \ldots, \Omega_n$. □

2.3 Distributions on an open subset $\Omega \subset \mathbb{R}^d$

2.3.1 Definitions and examples

Definition 2.3.1 Let $(\varphi_j)_{j \geq 1}$ be a sequence of test functions in $\mathscr{D}(\Omega)$, and $\varphi \in \mathscr{D}(\Omega)$. We say that (φ_j) converges to φ in $\mathscr{D}(\Omega)$ (or in the $\mathscr{D}(\Omega)$-sense), when

i) There exits a compact subset $K \subset \Omega$ such that supp $\varphi_j \subset K$ for all j.

ii) For all $\alpha \in \mathbb{N}^d$, $\|\partial^\alpha \varphi_j - \partial^\alpha \varphi\|_\infty \to 0$ as $j \to \infty$.

In that case we may write

$$\varphi = \mathscr{D} - \lim_{j \to +\infty} \varphi_j .$$

Remark 2.3.2 Notice that, under the conditions of the above definition, we have $\mathrm{supp}(\varphi) \subset K$.

Definition 2.3.3 (Distributions on Ω) Let $\Omega \subset \mathbb{R}^d$ be an open subset, and T a complex valued linear form on $\mathscr{D}(\Omega)$. One says that T is a distribution on Ω if, for every compact set $K \subset \Omega$,

$$\exists C > 0, \exists m \in \mathbb{N}, \forall \varphi \in \mathcal{C}_0^\infty(\Omega) \text{ with } \mathrm{supp}\, \varphi \subset K\ ,\ |T(\varphi)| \leq C \sum_{|\alpha| \leq m} \sup |\partial^\alpha \varphi| = C\|\varphi\|_{C^m}.$$

We denote by $\mathscr{D}'(\Omega)$ the set of distributions on Ω, and for $T \in \mathscr{D}'(\Omega)$, $\varphi \in \mathscr{D}(\Omega)$, we write $\langle T, \varphi\rangle \stackrel{\text{def}}{=} T(\varphi)$.

We have the same characterization of distributions in 1 dimension, in terms of continuity w.r.t. converging sequences of test functions.

Proposition 2.3.4 A linear form T on $\mathscr{D}(\Omega)$ is a distribution on Ω if and only if $T(\varphi_j) \to T(\varphi)$ for any sequence (φ_j) of functions in $\mathscr{D}(\Omega)$ that converges to φ in the $\mathscr{D}(\Omega)$-sense.

The proof is similar to the one dimensional case, as well as the following examples.

Locally integrable functions.

Given $f \in L^1_{loc}(\Omega)$, the formula

$$\langle T_f, \varphi\rangle = \int_\Omega f(x)\,\varphi(x)\,dx$$

defines a distribution on Ω. Furthermore, the linear mapping

$$f \in L^1_{loc}(\Omega) \mapsto T_f \in \mathscr{D}'(\Omega)$$

is one to one. In the sequel, we shall identify f to T_f.

Dirac masses.

Given $a \in \Omega$, the formula

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

defines a distribution δ_a on Ω, called the Dirac mass at a. It is not defined by any $f \in L^1_{loc}(\Omega)$.

Like in the previous chapter, one can define the notion of finite order distributions, and the corresponding notion of order. In particular, distributions of order $\leq m$ can be extended to continuous linear forms on $\mathcal{C}_0^m(\Omega)$, non-negative distributions are of order 0 — hence are positive measures, and distributions of order 0 are finite linear combinations of nonnegative distributions.

We now come to the important notion of support, for which we shall be a little more specific than we were in Chapter 1.

2.3.2 Restriction and support

Definition 2.3.5 Let $T \in \mathscr{D}'(\Omega)$, and $V \subset \Omega$ an open subset. The restriction of T to V is the distribution $T_{|V} \in \mathscr{D}'(V)$ defined as

$$\forall \varphi \in \mathscr{D}(V)\ ,\ \langle T_{|V}, \varphi \rangle = \langle T, \underline{\varphi} \rangle\ ,$$

where $\underline{\varphi}$ denotes the extension of φ by 0 on $\Omega \setminus V$.

We say that T vanishes in V if $T_{|V} = 0$.

Definition 2.3.6 The support of a distribution $T \in \mathscr{D}'(\Omega)$ is the complement of the union of all the open subsets where T vanishes. We denote it by $\operatorname{supp} T$.

Notice that $\operatorname{supp} T$ is closed, and the following characterizations are convenient.

- $x_0 \notin \operatorname{supp} T$ if and only if there is an open neighborhood V of x_0 such that $T_{|V} = 0$.
- $x_0 \in \operatorname{supp} T$ if and only if for any open neighborhood V of x_0, one can find $\varphi \in \mathcal{C}_0^\infty(V)$ such that $\langle T, \varphi \rangle \neq 0$.

As in the one dimensional case, one can characterize distributions supported in one point a, as given by

$$\langle T, \varphi \rangle = \sum_{|\alpha| \leq m} c_\alpha \partial^\alpha \varphi(a)\ ,$$

where $(c_\alpha)_{|\alpha| \leq m}$ is a family of complex numbers, and $m \in \mathbb{N}$.

Proposition 2.3.7 Let $\varphi \in \mathcal{C}_0^\infty(\Omega)$ and $T \in \mathscr{D}'(\Omega)$. If $\operatorname{supp} \varphi \cap \operatorname{supp} T = \emptyset$, then $\langle T, \varphi \rangle = 0$. In particular, if $\operatorname{supp}(T) = \emptyset$, then $T = 0$.

The proof is similar to the one dimensional case. We now emphasize the following corollary, which is very useful.

Corollary 2.3.8 (The gluing principle in $\mathscr{D}'(\Omega)$) Assume

$$\Omega = \bigcup_{j\in J} \Omega_j$$

for some (possibly infinite, even uncountable) collection $(\Omega_j)_{j\in J}$ of open subsets. Suppose we are given, for every $j \in J$, a distribution T_j on Ω_j, so that the family $(T_j)_{j\in J}$ satisfies the following *compatibility property* : for every $j, k \in J$ such that $\Omega_j \cap \Omega_k \neq \emptyset$, then

$$(T_j)_{|\Omega_j\cap\Omega_k} = (T_k)_{|\Omega_j\cap\Omega_k} \ . \tag{2.3.1}$$

Then there exists a unique $T \in \mathscr{D}'(\Omega)$ such that, for every $j \in J$, $T_{|\Omega_j} = T_j$.

Remark 2.3.9 This construction of a distribution by gluing small pieces satisfying a compatibility condition (2.3.1) gives to the space of distributions $\mathscr{D}'(\Omega)$ the structure of a *sheaf* over Ω.

Proof.— The uniqueness of T follows from Proposition 2.3.7: if two distributions $T, \tilde{T}$ satisfy $T_{|\Omega_j} = \tilde{T}_{|\Omega_j}$ for any j, it means that $(T-\tilde{T})_{|\Omega_j} = 0$ for all j, hence supp$(T-\tilde{T}) = 0$, and thus $T - \tilde{T} = 0$.

Let us prove the existence of T. Given $\varphi \in \mathscr{D}(\Omega)$, we want to define $\langle T, \varphi\rangle$. From Borel-Lebesgue, supp φ can be covered by a finite subcollection supp $\varphi \subset \bigcup_{j\in J_\varphi} \Omega_j$. Proposition 2.2.3 shows that one can decompose

$$\varphi = \sum_{j\in J_\varphi} \varphi_j$$

where $\varphi_j \in \mathscr{D}(\Omega_j)$. In such a situation, if T exists, one must have

$$\langle T, \varphi\rangle = \sum_{j\in J_\varphi} \langle T, \varphi_j\rangle = \sum_{j\in J_\varphi} \langle T_j, \varphi_j\rangle \ .$$

This equality seems to defined $\langle T, \varphi\rangle$. However, we observe that the splitting $\varphi = \sum \varphi_j$ is not unique; the choice of subfamily J_φ itself may not be unique.

Therefore, we need to check that the above right hand side does not depend on these choices. This is the content of the following

Lemma 2.3.10 For every decomposition

$$\varphi = \sum_{j\in J} \varphi_j$$

where each $\varphi_j \in \mathscr{D}(\Omega_j)$ and $\varphi_j = 0$ except for a finite set of indices j, the value of the sum

$$\sum_{j\in J} \langle T_j, \varphi_j\rangle$$

only depends on φ, not on the decomposition.

Let us prove the lemma. Once we are given a decomposition, we consider the set

$$K = \bigcup_{j \in J} \mathrm{supp}(\varphi_j) \ .$$

Since $\mathrm{supp}(\varphi_j)$ is compact and is empty except for a finite set J_1 of indices j, K is a compact subset of Ω. Apply Proposition 2.2.3 to a finite covering

$$K \subset \bigcup_{r \in R} \Omega_r$$

extracted from the covering of K by the $(\Omega_j)_{j \in J}$'s. The family $(\chi_r)_{r \in R}$ satisfies

$$\mathrm{supp}(\chi_r) \subset \Omega_r \ , \quad \sum_{r \in R} \chi_r = 1 \text{ on } K \ .$$

In particular, for every $j \in J_1$, $\varphi_j = \sum_{r \in R} \chi_r \varphi_j$, therefore

$$\langle T_j, \varphi_j \rangle = \sum_{r \in R} \langle T_j, \chi_r \varphi_j \rangle \ .$$

For any $r \in R$, we claim that $\langle T_j, \chi_r \varphi_j \rangle = \langle T_r, \chi_r \varphi_j \rangle$. Indeed,
– either $\Omega_j \cap \Omega_r = \emptyset$, and $\chi_r \varphi_j = 0$, so both sides of this equality cancel;
– or $\Omega_j \cap \Omega_r \neq \emptyset$, and, since $\mathrm{supp}(\chi_r \varphi_j) \subset \Omega_j \cap \Omega_r$, the assumption merely says that

$$\langle T_j, \chi_r \varphi_j \rangle = \langle T_r, \chi_r \varphi_j \rangle.$$

Consequently,

$$\begin{aligned} \sum_{j \in J_1} \langle T_j, \varphi_j \rangle &= \sum_{j \in J_1} \sum_{r \in R} \langle T_r, \chi_r \varphi_j \rangle = \sum_{r \in R} \langle T_r, \chi_r (\sum_{j \in J_1} \varphi_j) \rangle \\ &= \sum_{r \in R} \langle T_r, \chi_r \varphi \rangle \ . \end{aligned}$$

This proves the lemma. Indeed, given any other decomposition $\varphi = \sum_{j \in J_2} \tilde{\varphi}_j$, just apply the above construction with

$$K = \bigcup_{j \in J_2} \mathrm{supp}(\varphi_j) \cup \bigcup_{j \in J} \mathrm{supp}(\tilde{\varphi}_j) \ ,$$

and we see that

$$\sum_{j \in J_1} \langle T_j, \varphi_j \rangle = \sum_{r \in R} \langle T_r, \chi_r \varphi \rangle = \sum_{j \in J_2} \langle T_j, \tilde{\varphi}_j \rangle. \qquad \square$$

Let us complete the proof of Corollary 2.3.8. Choosing any decomposition $\varphi = \sum_{j \in J} \varphi_j$ as in the Lemma, we can then consistently define

$$\langle T, \varphi \rangle \stackrel{\text{def}}{=} \sum_{j \in J} \langle T_j, \varphi_j \rangle \ .$$

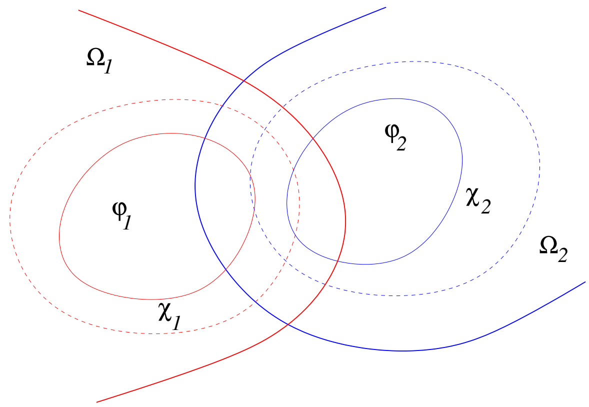


Figure 2.2: The supports of the functions φ_j and χ_j inside Ω_j, for $j = 1, 2$.

It is clear that T is a linear form on $\mathscr{D}(\Omega)$. To check it is a distribution, let K be a compact subset of Ω and $(\chi_r)_{r\in R}$ be a partition of unity associated to a finite covering $K \subset \cup_{r\in R}\Omega_r$. Then, for every test function φ supported in K, we have, from the lemma,

$$\langle T, \varphi\rangle = \sum_{r\in R} \langle T_r, \chi_r \varphi\rangle \ .$$

Since $T_r \in \mathscr{D}'(\Omega_r)$, we have

$$|\langle T_r, \chi_r\varphi\rangle| \leq C_r \|\chi_r\varphi\|_{C^{m_r}} \leq C_r' \|\varphi\|_{C^{m_r}} \ .$$

Then, with $m = \max_{r\in R} m_r$, we obtain

$$|\langle T, \varphi\rangle| \leq \left(\sum_{r\in R} C_r'\right) \|\varphi\|_{C^m} \ ,$$

so that $T \in \mathscr{D}'(\Omega)$. Finally, if $\varphi \in \mathscr{D}(\Omega_{j_0})$ for some $j_0 \in J$, we can write the decomposition $\varphi = \sum_{j\in J} \varphi_j$ with $\varphi_{j_0} = \varphi$ and $\varphi_j = 0$ if $j \neq j_0$. Using the lemma, we infer $\langle T, \varphi\rangle = \langle T_{j_0}, \varphi\rangle$. In other words, $T_{|\Omega_{j_0}} = T_{j_0}$. □

Let us close this paragraph by a few remarks concerning Corollary 2.3.8.

i) If every T_j is a nonnegative distribution, then so is T. Indeed, the elements χ_r of the partitions of unity can be chosen to be nonnegative.

ii) If every T_j is a $\mathcal{C}^m$ function, then so is T.

2.3.3 Multiplication by a smooth function

Definition 2.3.11 Given $q \in \mathcal{C}^\infty(\Omega)$ and $T \in \mathscr{D}'(\Omega)$, we define $qT \in \mathscr{D}'(\Omega)$ by

$$\forall \varphi \in \mathscr{D}(\Omega) \ , \quad \langle qT, \varphi \rangle = \langle T, q\varphi \rangle \ .$$

If $T \in L^1_{loc}(\Omega)$, the definition coincides with the usual product. Let us just mention the following generalization of Proposition 1.5.7 in Chapter 1.

Proposition 2.3.12 Let $T \in \mathscr{D}'(\Omega)$ and $a \in \Omega$ such that

$$\forall j \in \{1 \ldots, d\} \ , \ (x_j - a_j)T = 0 \ .$$

Then there exists $c \in \mathbb{C}$ such that $T = c\delta_a$.

Proof.— In view of what we did in Chapter 1, the proof reduces to the following

Lemma 2.3.13 (Hadamard lemma on an arbitrary open set Ω) If $\varphi \in \mathcal{C}_0^\infty(\Omega)$ satisfies $\varphi(a) = 0$, there exists $\psi_1, \ldots, \psi_d$ in $\mathcal{C}_0^\infty(\Omega)$ such that

$$\varphi(x) = \sum_{j=1}^{d} (x_j - a_j)\psi_j(x) \ .$$

Let us prove the lemma. Let $r > 0$ such that $B(a,r) \subset \Omega$, and let χ be a plateau function on $\overline{B}(a, r/2)$, supported in $B(a,r)$. From Corollary 2.1.8 — Hadamard's formula on a convex subset — we can write, for any $x \in B(a,r)$,

$$\varphi(x) = \sum_{j=1}^{d} (x_j - a_j) f_j(x) \ ,$$

with each $f_j \in \mathcal{C}^\infty(B(a,r)$. This implies that

$$\chi\varphi = \sum_{j=1}^{d} (x_j - a_j)\chi f_j \, .$$

This settles the part of φ supported near a. On the other hand, since $(1-\chi)$ vanishes near a, we can factorize $(1-\chi)\varphi$ as:

$$(1-\chi(x))\varphi(x) = \sum_{j=1}^{d} (x_j - a_j) \frac{(x_j - a_j)(1-\chi(x))}{|x-a|^2} \varphi(x) \ .$$

Summing both the above identities, the lemma follows with the functions

$$\psi_j(x) = \chi(x) f_j(x) + \frac{(x_j - a_j)(1 - \chi(x))}{|x - a|^2} \varphi(x) \ .$$

$\square$

2.3.4 Differentiation

As in one space dimension, we first consider the case of a C^1 function. In this case, we have the following elementary

Lemma 2.3.14 Let $f \in \mathcal{C}^1(\Omega)$, $\varphi \in \mathcal{C}^1_0(\Omega)$. Then, for any index $j \in \{1, \ldots, d\}$,

$$\int_\Omega \partial_j f(x) \, \varphi(x) \, dx = - \int_\Omega f(x) \, \partial_j \varphi(x) \, dx \ .$$

Proof.— Let χ be a plateau function on a neighborhood of the support of φ. It is easy to check that

$$\partial_j(\chi f)\varphi = \partial_j f \, \varphi \, , \qquad \chi f \partial_j \varphi = f \partial_j \varphi \ .$$

Furthermore, χf can be extended as a C^1 function on $\mathbb{R}^d$, vanishing outside of Ω. Therefore we are reduced to proving the lemma with $\Omega = \mathbb{R}^d$. This is an immediate consequence of the Fubini theorem and of integration by parts in the x_j variable. $\square$

On the basis of Lemma 2.3.14, we introduce the following definition for the partial derivatives of a distribution.

Definition 2.3.15 Let $T \in \mathscr{D}'(\Omega)$ and $j \in \{1, \ldots, d\}$. We define $\partial_j T \in \mathscr{D}'(\Omega)$ by

$$\langle \partial_j T, \varphi \rangle = -\langle T, \partial_j \varphi \rangle \ , \quad \varphi \in \mathscr{D}(\Omega) \ .$$

Similarly, for every $\alpha \in \mathbb{N}^d$, we define $\partial^\alpha T \in \mathscr{D}'(\Omega)$ by

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

At this stage, it is natural to ask for the multidimensional analogue of the identity $H' = \delta_0$ proved in Chapter 1, with $H = \mathbb{1}_{\mathbb{R}_+}$ the Heaviside function. A natural statement would be a formula for $\partial_j(\mathbb{1}_U)$, where U is an open subset of Ω. However, open subsets in $\mathbb{R}^d$ can be complicated enough, so that no such formula exists without additional assumptions on U. A relatively general formula of this kind will be the purpose of the next section, devoted to superficial measures and to the jump formula. At this stage, let us just consider a very simple examples in $\mathbb{R}^2$, that of the half-plane on $\mathbb{R}^2$.

For some $a \in \mathbb{R}$, consider the half-space $H_a = \{(x_1, x_2) \in \mathbb{R}^2 \,:\, x_1 > a\} \subset \mathbb{R}^2$. Then an elementary calculation gives, for any $\varphi \in \mathscr{D}(\mathbb{R}^2)$:

$$\langle \partial_1(\mathbb{1}_{H_a}), \varphi \rangle = \int_{\mathbb{R}} \varphi(a, x_2)\, dx_2 \,, \qquad \langle \partial_2(\mathbb{1}_{H_a}), \varphi \rangle = 0 \,,$$

so $\partial_1(\mathbb{1}_{H_a})$ is the superficial measure on the x_2-axis, while $\partial_2(\mathbb{1}_{H_a}) = 0$.

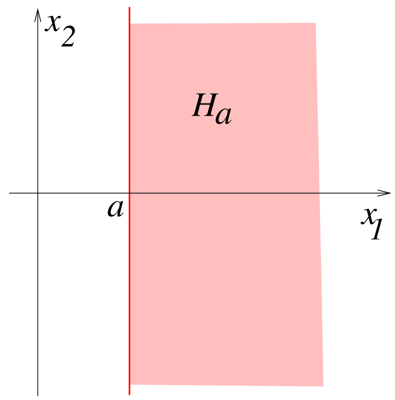


Figure 2.3: Derivatives of $\mathbb{1}_U$ for the half-plane.

Notice that $\partial_j(\mathbb{1}_{H_a})$ is a distribution of order 0, supported by the boundary of H_a. This fact will be generalized to more general open sets U in the next section.

2.3.5 Convergence

The convergence of a sequence of distributions is defined like in one dimension.

Definition 2.3.16 A sequence $(T_n)_{n\in\mathbb{N}}$ of distributions in $\mathscr{D}'(\Omega)$ converges to $T \in \mathscr{D}'(\Omega)$ if

$$\forall \varphi \in \mathscr{D}(\Omega) \,, \quad \langle T_n, \varphi \rangle \to \langle T, \varphi \rangle \,.$$

If $f \in L^1(\mathbb{R}^d)$ satisfies $\int_{\mathbb{R}^d} f(x)\, dx = 1$, then we can easily show (see Example 1.7.3 in 1 dimension) that

$$\frac{1}{\varepsilon^d} f\Big(\frac{x}{\varepsilon}\Big) \underset{\varepsilon\to 0}{\longrightarrow} \delta_0 \quad \text{in } \mathscr{D}'(\mathbb{R}^d) \,.$$

This observation leads to the following strategy of *regularization*. Fix $\rho \in \mathscr{D}(\mathbb{R}^d)$, supported in the unit ball, such that $\int_{\mathbb{R}^d} \rho(x)\, dx = 1$, and consider

$$\rho_\varepsilon(x) = \frac{1}{\varepsilon^d} \rho\Big(\frac{x}{\varepsilon}\Big) \,.$$

Given an open subset Ω in $\mathbb{R}^d$, set the ε-rectract of Ω,

$$\Omega^\varepsilon := \{x \in \Omega, d(x, \Omega^c) > \varepsilon\} .$$

If $x \in \Omega^\varepsilon$, the function

$$\rho_\varepsilon(x - .) : y \mapsto \rho_\varepsilon(x - y)$$

is supported in the closed ball of radius ε centered at x, which is included in Ω. Hence this function belongs to $\mathscr{D}(\Omega)$, and we may define the function

$$F^\varepsilon(x) = \langle T, \rho_\varepsilon(x - \cdot)\rangle \, , , \quad x \in \Omega^\varepsilon \, ,$$

which looks like the convolution of T by ρ_ε. In some sense, $F^\varepsilon(x)$ represents an "average" of T on the ball $B(x, \varepsilon)$.

Remark 2.3.17 The convolution we had described in Definition 1.9.7 in Chapter 1 was defined on $\mathbb{R}$, so there was no necessity of restricting it on some ε-retract.

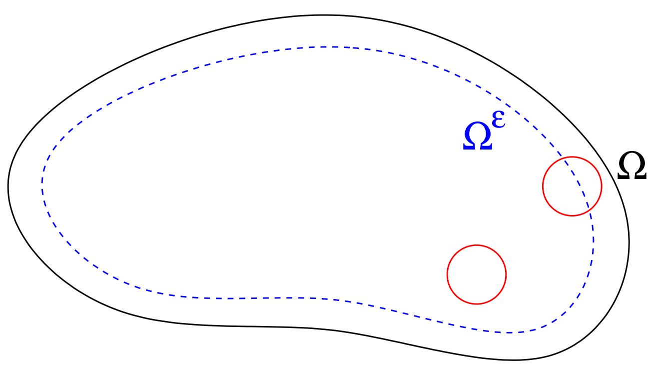


Figure 2.4: Regularization F^ε in the retract Ω^ε. Each red circle represents the support of the test function $\rho_\varepsilon(x - \cdot)$.

Proposition 2.3.18 The function F^ε is smooth on Ω^ε, with

$$\partial^\alpha F^\varepsilon(x) = \langle T, \partial^\alpha \rho_\varepsilon(x - \cdot)\rangle \, , \; \alpha \in \mathbb{N}^d \, ,$$

and, for every $\varphi \in \mathscr{D}(\Omega)$,

$$\int_{\Omega^\varepsilon} F^\varepsilon(x)\, \varphi(x)\, dx \xrightarrow[\varepsilon \to 0]{} \langle T, \varphi\rangle \, .$$

Notice that, given a compact subset K in Ω, the integral $\int_\Omega F^\varepsilon(x)\, \varphi(x)\, dx$ is well defined for ε small enough and every φ supported in K. The proof of Proposition 2.3.18 is a consequence of the following two results, which will be frequently used throughout the course, and are generalisations to several variables of Propositions 1.9.5 and 1.9.6 from Chapter 1.

Proposition 2.3.19 (Differentiation under the bracket) Let $\Omega \subset \mathbb{R}^d$, $Z \subset \mathbb{R}^p$ be open sets, and $T \in \mathscr{D}'(\Omega)$. For some compact set $K \subset \Omega$,let $\varphi \in \mathcal{C}^\infty(\Omega \times Z)$ be supported in $K \times Z$.

Then the function

$$G : z \in Z \mapsto \langle T, \varphi(\cdot, z)\rangle$$

is $\mathcal{C}^\infty$, and, for any $\alpha \in \mathbb{N}^p$,

$$\partial^\alpha G(z) = \langle T, \partial_z^\alpha \varphi(\cdot, z)\rangle \,.$$

Remark 2.3.20 If $T = f \in L^1_{loc}(\Omega)$, we have $G(z) = \int_\Omega f(x)\varphi(x,z)dx$, so that, under the above assumptions, we get $G \in \mathcal{C}^\infty(Z)$ and we recover the Leibniz rule of derivation under the integral sign,

$$\partial^\alpha G(z) = \int_\Omega f(x)\, \partial_z^\alpha \varphi(x,z)\, dx.$$

Proof.— Let $z_0 \in Z$ and $x \in \Omega$. For $h \in \mathbb{R}^q$, Taylor's formula at order 1 gives

$$\varphi(x, z_0 + h) = \varphi(x, z_0) + \sum_{j=1}^{p} \partial_{z_j}\varphi(x, z_0) h_j + r(x, z_0, h),$$

$$\text{with} \quad r(x, z_0, h) = 2 \sum_{|\alpha|=2} \frac{h^\alpha}{\alpha!} \int_0^1 (1-t)\partial_z^\alpha \varphi(x, z_0 + th)dt.$$

Since $x \mapsto r(x, z_0, h))$ is $\mathcal{C}^\infty$ with support in K, there exist a constant $C > 0$ and an integer $m \in \mathbb{N}$ (independent of z_0 or h) such that

$$|\langle T, r(\cdot, z_0, h)\rangle| \leq C \sum_{|\beta|\leq m} \sup |\partial_x^\beta r(x, z_0, h)|$$

But for $|h| \leq 1$,

$$|\partial_x^\beta r(x, z_0, h)| \leq 2 \sum_{|\alpha|=2} \frac{|h^\alpha|}{\alpha!} \int_0^1 (1-t)\, |\partial_x^\beta \partial_z^\alpha \varphi(x, z_0 + th)|\, dt \leq C|h|^2 \sum_{|\alpha|=2} \sup_{K\times \overline{B}(z_0,1)} |\partial_x^\beta \partial_z^\alpha \varphi(x,z)|,$$

Since $|h^\alpha| \leq |h|^2$ for all $|\alpha| = 2$, we get

$$|\langle T, r(\cdot, z_0, h)\rangle| = \mathcal{O}(|h|^2)\,,$$

so that

$$G(z_0 + h) = G(z_0) + \sum_{j=1}^{q} \langle T, \partial_{z_j}\varphi(\cdot, z_0)\rangle h_j + \mathcal{O}(|h|^2).$$

This equality shows that G is differentiable at z_0 (in particular G is continuous), with derivatives

$$\partial_j G(z) = \langle T, \partial_{z_j}\varphi(\cdot, z)\rangle,$$

which proves the formula for the first derivatives.

If we now replace $\varphi(x,z)$ by $\partial_{z_j}\varphi(x,z)$ in the above discussion, we see that for all j, $\partial_j\varphi$ is differentiable, thus in particular continuous. So G is $\mathcal{C}^1$, and the statement of the proposition is true for any $|\alpha|=1$. One easily shows the general case by induction. □

Let us now prove that the inverse operation, namely integration w.r.t. the auxiliary parameter z, can also be shifted inside the bracket.

Proposition 2.3.21 (Integration under the bracket) Let $\Omega\subset\mathbb{R}^d$ be an open set, and $T\in\mathscr{D}'(\Omega)$. Let also $\varphi\in\mathcal{C}_0^\infty(\Omega\times\mathbb{R}^p)$. Then

$$\int_{\mathbb{R}^p}\langle T,\varphi(\cdot,z)\rangle dz=\left\langle T,\int_{\mathbb{R}^p}\varphi(\cdot,z)dz\right\rangle$$

Proof.— We start with the case $p=1$, for which the proof is similar to the one in Prop. 1.9.6. Let $\varphi\in\mathcal{C}_0^\infty(\Omega\times\mathbb{R})$. We choose $A>0$ and a compact set $K\subset\Omega$ such that supp $\varphi\subset K\times[-A,A]$. We denote by $\psi:\Omega\times\mathbb{R}\to\mathbb{C}$ the function given by

$$\psi(x,z)=\int_{t<z}\varphi(x,t)\,dt\,,$$

namely a primitive of φ w.r.t. z. The function ψ belongs to $\mathcal{C}^\infty(\Omega\times\mathbb{R})$, and for any z, supp$(\psi(\cdot,z))$ is included in K. Therefore Proposition 2.3.19 applies. The function

$$G(z)=\langle T,\psi(\cdot,z)\rangle=\left\langle T,\int_{t<z}\varphi(\cdot,t)dt\right\rangle$$

is smooth and vanishes for $z\le -A$. Its derivative reads

$$G'(z)=\langle T,\partial_y\psi(\cdot,z)\rangle=\langle T,\varphi(\cdot,z)\rangle.$$

Integrating over the parameter z, we get

$$\left\langle T,\int_{t<z}\varphi(\cdot,t)dt\right\rangle=G(z)=\int_{t<z}G'(t)dt=\int_{t<z}\langle T,\varphi(\cdot,t)\rangle dt\ .$$

Taking $z=A$, we obtain the required statement in the case $p=1$.

For $p>1$, we proceed by induction on p. Let $\varphi\in\mathcal{C}_0^\infty(\Omega\times\mathbb{R}^p)$. We split the variable $z=(z',t)$ with $z'\in\mathbb{R}^{p-1}$. Using the result in the case $p=1$, we get, for every fixed $z'\in\mathbb{R}^{p-1}$:

$$\left\langle T,\int_{\mathbb{R}}\varphi(\cdot,z',t)\,dt\right\rangle=\int_{\mathbb{R}}\langle T,\varphi(\cdot,z',t)\rangle\,dt\,.$$

It remains to apply the induction assumption to $\tilde\varphi\in\mathcal{C}_0^\infty(\Omega\times\mathbb{R}^{p-1})$ defined by

$$\tilde\varphi(x,z')=\int_{\mathbb{R}}\varphi(x,z',t)\,dt\ .$$

Indeed, using Fubini's theorem we get

$$\int_{\mathbb{R}^p} \langle T, \varphi(\cdot, z', t)\rangle dz'\, dt = \int_{\mathbb{R}^{p-1}} \langle T, \tilde\varphi(\cdot, z')\rangle dz' = \left\langle T, \int_{\mathbb{R}^{p-1}} \tilde\varphi(\cdot, z')\, dz' \right\rangle = \left\langle T, \int_{\mathbb{R}^p} \varphi(\cdot, z', t)\, dz'\, dt \right\rangle .$$

□

Let us now finally come back to the proof of Proposition 2.3.18 on the regularization by convolution. Fix $\varphi \in \mathcal{C}_0^\infty(\Omega)$, and choose $\varepsilon > 0$ small enough such that $\operatorname{supp}\varphi \subset \Omega^\varepsilon$. Applying Proposition 2.3.19 — derivation under the bracket— to the function

$$\varphi(x)F^\varepsilon(x) = \langle T, \varphi(x)\rho_\varepsilon(x - \cdot)\rangle$$

(which is well-defined for the above values of ε), we infer $\varphi F^\varepsilon \in \mathcal{C}_0^\infty(\Omega)$ and

$$\forall \alpha \in \mathbb{N}^d\,, \quad \partial^\alpha(\varphi F^\varepsilon)(x) = \langle T, \partial_x^\alpha(\varphi(x)\rho_\varepsilon(x - \cdot))\rangle\ .$$

Take $x_0 \in \Omega^\varepsilon$, and choose φ a plateau function near $x_0 \in \Omega$. We then obtain the first statement of the proposition:

$$\forall \alpha \in \mathbb{N}^d\,, \quad \partial^\alpha F^\varepsilon(x_0) = \langle T, \partial^\alpha \rho_\varepsilon(x_0 - \cdot)\rangle\ .$$

As for the second statement, we apply Proposition 2.3.21 — integration under the bracket — to obtain

$$\int_\Omega \varphi(x)F^\varepsilon(x)\, dx = \langle T, \varphi_\varepsilon\rangle\ , \qquad \varphi_\varepsilon(y) := \int_{\mathbb{R}^d} \varphi(x)\rho_\varepsilon(x-y)\, dx = \phi * \tilde\rho_\varepsilon(x)\,,$$

where $\tilde\rho(x) := \rho(-x)$. Notice that a simple change of variables $x = y + \varepsilon z$ provides

$$\varphi_\varepsilon(y) = \int_{\mathbb{R}^d} \varphi(y + \varepsilon z)\rho(z)\, dz\ .$$

Since ρ is supported in the unit ball, we observe that φ_ε is supported in a fixed compact subset of Ω if ε is small enough. Furthermore, the standard derivation under the integral yields $\varphi_\varepsilon \in \mathcal{C}_0^\infty(\Omega)$ and

$$\forall \alpha \in \mathbb{N}^d\,,\ \partial^\alpha \varphi_\varepsilon(y) = \int_{\mathbb{R}^d} \partial^\alpha \varphi(y + \varepsilon z)\rho(z)\, dz\ .$$

Passing to the limit as ε tends, to 0, $\partial^\alpha\varphi_\varepsilon$ converges uniformly to $\partial^\alpha\varphi$. We conclude that φ_ε converges to φ in $\mathscr{D}(\Omega)$. Consequently,

$$\langle T, \varphi_\varepsilon\rangle \underset{\varepsilon\to 0}{\longrightarrow} \langle T, \varphi\rangle\ .$$

The proof of Proposition 2.3.18 is complete. □

Corollary 2.3.22 Every distribution on Ω is the limit (in $\mathscr{D}'(\Omega)$) of a sequence of test functions.

Proof.— For every $j \geq 1$, define

$$K_j = \left\{ x \in \Omega : d(x, \complement\Omega) \geq \frac{1}{j}\ ,\ |x| \leq j \right\} = \overline{\Omega^{1/j} \cap B(0,j)}\ .$$

Then K_j is a compact subset of Ω, $K_j \subset \mathring{K}_{j+1}$ and $\bigcup_{j\geq 1} K_j = \Omega$.

Choose $\chi_j \in \mathscr{D}(\mathring{K}_{j+1})$ a plateau function near K_j and consider a sequence (ε_j) converging to 0 such that

$$K_{j+1} \subset \Omega^{\varepsilon_j} \ .$$

Then the function

$$F_j(x) = \chi_j(x)\langle T, \rho_{\varepsilon_j}(x - .)\rangle$$

is well-defined since supp $\chi_j \subset \mathring{K}_{j+1}$, while the second factor (average of T in the ball $B(x, \varepsilon_j)$) is well-defined for $x \in K_{j+1} \subset \Omega^{\varepsilon_j}$. By Proposition 2.3.18 and by the above support information, we know that $F_j \in \mathcal{C}_0^\infty(\Omega)$. Furthermore, if $\varphi \in \mathscr{D}(\Omega)$, the support of φ is covered by a finite union of the open subsets $\mathring{K}_j$'s, hence is contained into some K_{j_0}. Then, for $j \geq j_0$, we have $\varphi = \chi_j \varphi$, hence

$$\int_\Omega \varphi(x)\, F_j(x)\, dx = \int_\Omega \varphi(x)\langle T, \rho_{\varepsilon_j}(x - .)\rangle\, dx \ ,$$

which converges to $\langle T, \varphi\rangle$ by Proposition 2.3.18. Therefore $F_j \to T$. □ We are now ready to

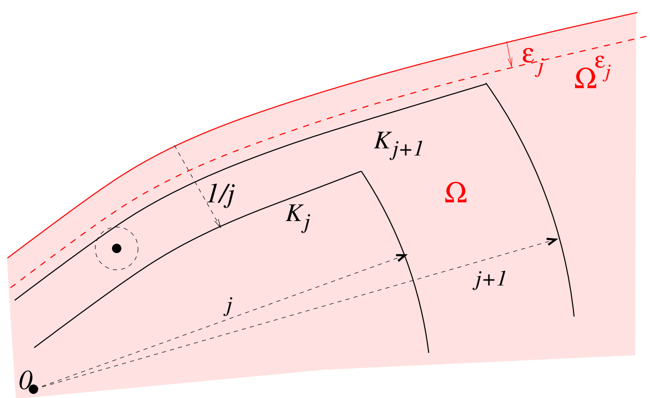


Figure 2.5: Construction of the function F_j approximating T.

characterize a C^1 distribution as if it were a function.

Corollary 2.3.23 Let $T \in \mathscr{D}'(\Omega)$ such that, for every $j \in \{1, \dots, d\}$, $\partial_j T \in \mathcal{C}^0(\Omega)$. Then $T \in \mathcal{C}^1(\Omega)$.

Remark 2.3.24 The proof below is significantly more intricate than its dimension 1 analogue. Indeed, the analogue of the integral formula used in Corollary 1.4.9 in Chapter 1 is much more complicated in several space dimensions, so we prefer to proceed differently. Moreover, let us mention that, as opposed to the one dimensional case, the assumptions $\partial_j T \in L^1_{loc}(\Omega), j = 1, \dots, d$, do not lead to $T \in \mathcal{C}^0(\Omega)$, but only to $T \in L^p_{loc}(\Omega)$ for $p = \frac{d}{d-1}$ (Sobolev estimates).

Let us now prove Corollary 2.3.23. By the gluing principle, it is enough to prove it locally, so we may assume that Ω is a ball B. For every j, denote by $f_j \in \mathcal{C}^0(B)$ the continuous function defined by

$\partial_j T$. As above, we define, on the smaller ball B^ε, the smooth function $F^\varepsilon(x) = \langle T, \rho_\varepsilon(x - .)\rangle$. Our goal is to show that F^ε converges to a C^1 function when $\varepsilon \to 0$.

Notice that, using the derivation under the bracket,

$$\partial_j F^\varepsilon(x) = \langle T, \partial_{x_j} \rho_\varepsilon(x - .)\rangle = -\langle T, \partial_{y_j}(\rho_\varepsilon(x - .))\rangle = \langle \partial_j T, \rho_\varepsilon(x - .)\rangle = \int_\Omega f_j(y) \rho_\varepsilon(x - y)\, dy\ .$$

To integrate this expression, we fix any ball $B' \subset\subset B$ and $\varepsilon_0 > 0$ such that $B' \subset B^{\varepsilon_0}$. Then the family $(\partial_j F^\varepsilon)_{\varepsilon<\varepsilon_0}$ converges uniformly to f_j on B' as ε tends to 0. Furthermore, picking $\chi \in \mathscr{D}(B')$ of integral 1, we have, using the Taylor formula:

$$F^\varepsilon(x) - \int_B \chi(y) F^\varepsilon(y)\, dy = \int_B \chi(y)(F^\varepsilon(x) - F^\varepsilon(y))\, dy = \sum_{j=1}^d \int_0^1 \int_B \chi(y)(x_j - y_j) \partial_j F^\varepsilon(y + t(x - y))\, dy\, dt.$$

The right hand side converges uniformly on B' to

$$\sum_{j=1}^d \int_0^1 \int_B \chi(y)\, (x_j - y_j)\, f_j(y + t(x - y))\, dy\, dt,$$

which is obviously a continuous function on B'. On the other hand, by Proposition 2.3.18, the constant term on the left hand side converges to

$$\int_B \chi(y) F^\varepsilon(y)\, dy \underset{\varepsilon\to 0}{\longrightarrow} \langle T, \chi\rangle\ ,$$

Hence, the remaining term F^ε converges uniformly on B' to some continuous function F. Since each $\partial_j F^\varepsilon$ is uniformly convergent on B', we conclude that each $\partial_j F = f_j$, and hence that $F \in \mathcal{C}^1(B')$. But we also know by Proposition 2.3.18 that $T_{|B'} = F$. Since B' is arbitrary, this completes the proof.
□

A very useful property of the convergence of distributions is the following Lemma, which is a consequence of the principle of uniform boundedness proved at the end of this chapter.

Lemma 2.3.25 (Bicontinuity of the bracket) Let (T_n) be sequence in $\mathscr{D}'(\Omega)$ and (φ_n) be a sequence in $\mathscr{D}(\Omega)$. We assume that $T_n \to T$ in $\mathscr{D}'(\Omega)$ and $\varphi_n \to \varphi$ in $\mathscr{D}(\Omega)$. Then

$$\langle T_n, \varphi_n\rangle \to \langle T, \varphi\rangle\ .$$

We close this subsection by a useful remark on the gluing principle for convergence of sequences of distributions.

Proposition 2.3.26 Let (T_n) be a sequence of distributions on Ω. Assume

$$\Omega = \bigcup_{j\in J} \Omega_j$$

for some collection $(\Omega_j)_{j\in J}$ of open subsets, and that, for every $j \in J$, $(T_n)_{|\Omega_j}$ converges to some $T^{(j)} \in \mathscr{D}'(\Omega_j)$. Then T_n is convergent in $\mathscr{D}'(\Omega)$.

Proof.— Let K be a compact subset of Ω. From Proposition 2.2.3 let $(\chi_j)_{j\in J}$ be a family of test functions, which are identically 0 except for a finite set of indices j, and such that

$$\mathrm{supp}(\chi_j) \subset \Omega_j \,, \qquad \sum_{j\in J} \chi_j = 1 \text{ on } K \,.$$

Then, for every test function φ supported in K, we have

$$\langle T_n, \varphi\rangle = \sum_{j\in J} \langle T_n, \chi_j \varphi\rangle \underset{n\to\infty}{\longrightarrow} \sum_{j\in J} \langle T^{(j)}, \chi_j \varphi\rangle \,.$$

Since each $T^{(j)}$ is a distribution, the right hand side is estimated by $C\|\varphi\|_{C^m}$ for some $C > 0$ and $m \in \mathbb{N}$ only depending on the compact K. Hence T_n is convergent in $\mathscr{D}'(\Omega)$. □

2.4 Superficial measures and the jump formula

2.4.1 Motivation

In this section, we come back to the problem of identifying the partial derivatives of the characteristic function of an open set $U \subset \Omega$. It is easy to check that these derivatives are distributions supported by the boundary ∂U of U. Under suitable assumptions on U, we shall show that these distributions are of order 0, and can be expressed in terms of a positive measure supported by ∂U, called the *superficial measure* on the hypersurface ∂U.

We first need to define the superficial measure on a hypersurface, which is the purpose of the next subsection. An intuitive way to define a superficial measure of a subset of an hypersurface is to thicken this subset into a *slab* of thickness ε, and to to take the limit, as ε tends to 0, of the ratio to ε of the Lebesgue measure of this slab. The next construction makes rigorous this intuitive definition.

2.4.2 Reminder on smooth hypersurfaces in $\mathbb{R}^d$

I thank Thomas Letendre for adding this subsection to the notes. The material in this subsection can be found for instance in the textbook *Introduction aux variétés différentielles* by Jacques Lafontaine, EDP Sciences, 2010 (Chap. 1, sections C and D).

Let $\Omega \subset \mathbb{R}^d$ be an open set, and consider a subset $\Sigma \subset \Omega$. Our goal is to characterize whether Σ is a C^k-smooth hypersurface of Ω.

Theorem 2.4.1 Consider $a \in \Sigma$, and $k \in \mathbb{N}^*$. Then the following statements are equivalent to one another:

$i)$ (local straightening) There exists U a neighbourhood of a in Ω, V a neighbourhood of $0 \in \mathbb{R}^d$, $\Phi : U \to V$ a C^k diffeomorphism, such that

$$\Phi(\Sigma \cap U) = (\{0\} \times \mathbb{R}^{d-1}) \cap V \,.$$

$ii)$ (zero set of a submersion) There exists U a neighbourhood of a in Ω, $f \in C^k(U, \mathbb{R})$ such that ∇f nowhere vanishes, and such that $\Sigma \cap U = f^{-1}(\{0\})$.

$iii)$ (local graph) Up to a permutation of the coordinates on $\mathbb{R}^d$, there exists U a neighbourhood of $a = (a_1, \ldots, a_d)$ in Ω, W a neighbourhood of $(a_2, \ldots, a_d)$ in $\mathbb{R}^{d-1}$, and $q \in C^k(W, \mathbb{R})$ such that

$$\Sigma \cap U = \mathsf{Gr}(q) = \{(x_1, y) \in \mathbb{R}^d\,;\, x_1 = q(y),\ y \in W\}.$$

Proof.— $i) \Rightarrow ii)$ Let $\Phi : U \to V$ be as indicated. Let us note $\pi : \mathbb{R}^d \to \mathbb{R}$ the first coordinate map given by $\pi(x) = x_1$, and define $f = \pi \circ \Phi : U \to \mathbb{R}$. Then the differential form $df(x) = d\pi \circ d\Phi(x)$ is surjective at each point $x \in U$, since the matrix $d\Phi(x)$ is invertible and $d\pi$ is surjective. As a result f is a submersion, and $\Sigma = f^{-1}(0)$.

$ii) \Rightarrow iii)$ We know that $\nabla f(x) \neq 0$ at each point $x \in U$. Up to a permutation of the coordinates, we may assume that near some $a \in \Sigma$, $\partial_{x_1} f(x) \neq 0$. By the implicit function theorem, there exists a neighbourhood V of a in U, and a C^k function q defined in a neighbourhood W of $(a_2, \ldots, a_d) \in \mathbb{R}^{d-1}$, such that, in the neighbourhood V,

$$f(x) = 0 \quad \text{iff} \quad x_1 = q(x_2, \ldots, x_d)\,.$$

$iii) \Rightarrow i)$ Assume that $\Sigma \cap U = \mathsf{Gr}(q)$. Up to translation of the coordinates, we may assume that the base point $a = 0$. Then, let us define

$$\Phi : (x_1, \ldots, x_d) \mapsto \big(x_1 - q(x_2, \ldots, x_d), x_2, \ldots, x_d\big).$$

It is easy to check that $\Phi : U \mapsto \Phi(U) \subset \mathbb{R}^d$ is a C^k diffeomorphism, which maps Σ to the local hyperplane $\{x_1 = 0\}$. □

Definition 2.4.2 If there exists $k \in \mathbb{N}^*$ such that the above equivalent statements hold for all $a \in \Sigma$, we say that Σ is a (boundaryless) hypersurface of Ω, of regularity C^k.

According to statement $i)$, a hypersurface (equivalently, a submanifold of dimension $d - 1$) locally resembles a piece of hyperplane (up to a C^k diffeomorphism).

The easiest characterization to check is usually $ii)$, eventhough it is the least intuitive one.

Definition 2.4.3 We say that a function $f \in C^k(\Omega, \mathbb{R})$ *vanishes transversally* if,

$$\forall x \in \Omega, \quad f(x) = 0 \Longrightarrow \nabla f(x) \neq 0\,.$$

Lemma 2.4.4 If $f \in C^k(\Omega, \mathbb{R})$ vanishes transversally, then its zero locus $\Sigma := f^{-1}(0)$ is a C^k hypersurface.

Proof.— When we restrict such a function f to the open set $\Omega_f := \{x \in \Omega, \nabla f(x) \neq 0\} \supset \Sigma$, it defines a submersion. The submersion theorem states that the zero locus of f, $\Sigma := f^{-1}(0) \subset \Omega_f$ is a C^k hypersurface. □

2.4.3 Examples of hypersurfaces

- if $U \subset \mathbb{R}^{d-1}$ is open, then $U \times \{0\}$ is a hypersurface in $\mathbb{R}^d$.
- for $U \subset \mathbb{R}^{d-1}$ open, the graph of any function $q \in C^k(U, \mathbb{R})$ is a hypersurface in $\mathbb{R}^d$.
- The affine hyperplanes in $\mathbb{R}^d$ are hypersurfaces.
- If one takes the function $f : x \in \mathbb{R}^d \mapsto |x|^2 - 1$, then its zero locus $f^{-1}(0) = \mathbb{S}^{d-1}$ is the $d-1$-dimensional unit sphere, which is a hypersurface in $\mathbb{R}^d$.
- the intersection of a hypersurface $\Sigma \subset \Omega$ with an open subset $U \subset \Omega$ is a hypersurface.
- in dimension $d = 1$, the hypersurfaces of $\Omega \subset \mathbb{R}$ are the discrete sets of points in Ω, that is the sets without accumulation points in Ω.
- in dimension $d = 2$, each connected component of a hypersurface Σ is a C^k curve without multiple points. The definition allows curves with boundaries.

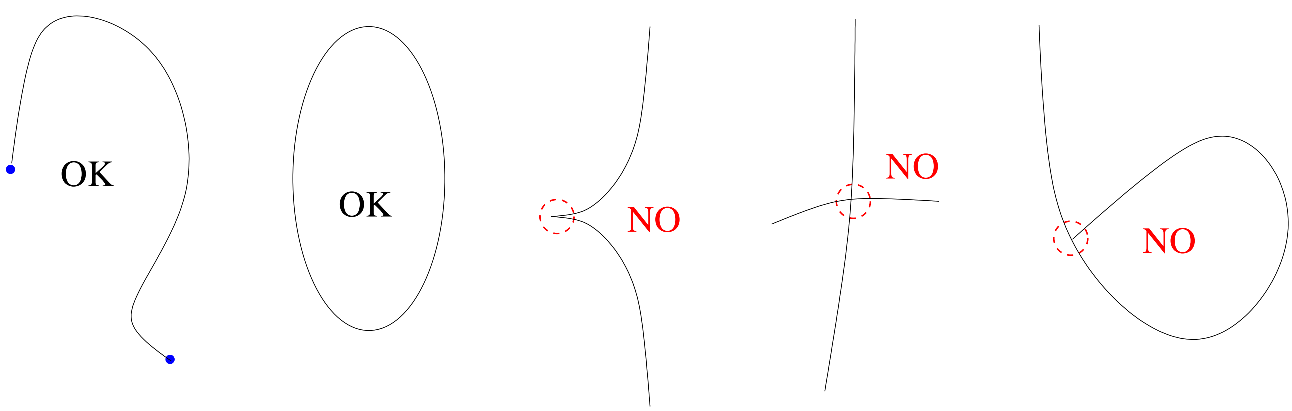


Figure 2.6: Smooth curves in $\mathbb{R}^2$ (the blue points are not part of the curves), and curves with singularities which are not hypersurfaces.

- in $d = 3$, hypersurfaces are the usual surfaces embedded in $\mathbb{R}^3$, with no singularity (edge, corner, conical point,...), but possible boundaries.

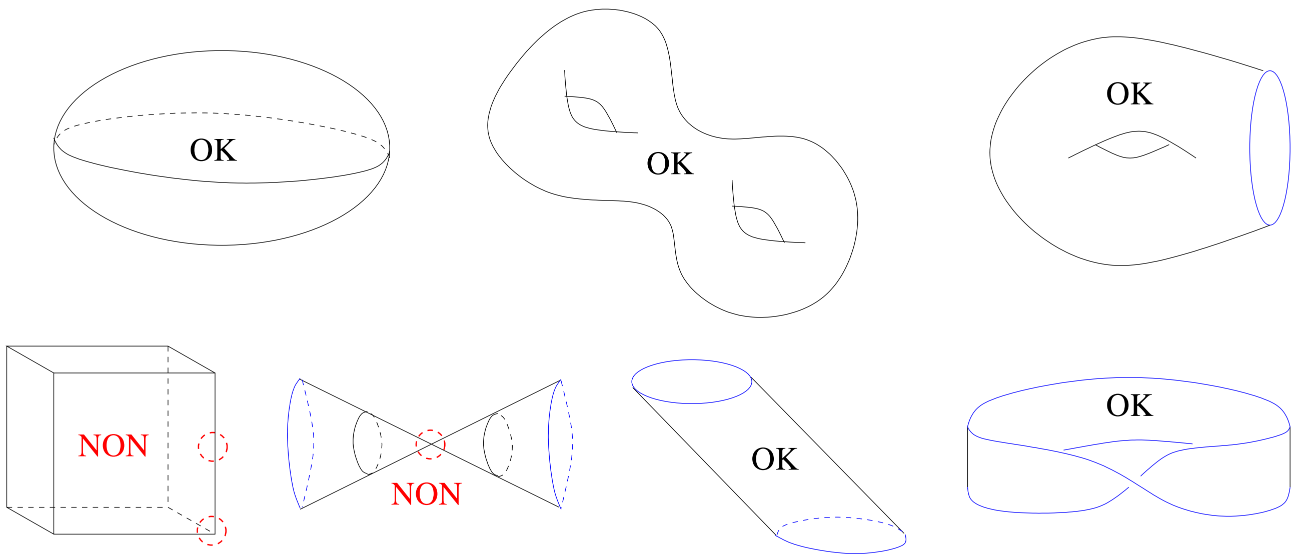


Figure 2.7: Surfaces in $\mathbb{R}^3$ (the blue boundaries are not part of the surface), and surfaces with singularities (edges, corners, conical point). The last one is a Möbius strip.

In the previous examples, we notice that some hypersurfaces Σ admit nonempty boundary $\partial\Sigma = \overline{\Sigma}\setminus\Sigma$, where the closure is taken in the ambient open set Ω. This is the case, for instance, of open curves in $d = 1$ (left on fig. 2.6), of open surfaces in $d = 2$ (see the "open handle" or the Möbius strip in fig. 2.7), if the ambient open set $\Omega = \mathbb{R}^d$. For the hypersurfaces to satisfy the conditions of Theorem 2.4.1, it is important that the boundary $\partial\Sigma$ is *not* contained in Σ: if we zoom onto a point on the boundary, we do not see a hyperplane, but rather a half-hyperplane.

In the following, we will be most often concerned with *closed* hypersurfaces, where we add the condition that Σ is closed in Ω. This condition will forbid the above situations with boundary.

2.4.4 Preliminaries on non–negative distributions

After this geometric interlude, let us come back to distributions. In this section we show that the nature of positive distributions in d dimensions is similar to the 1-dimensional case.

Definition 2.4.5 We say that $T \in \mathscr{D}'(\Omega)$ is a nonnegative distribution if $\langle T, \varphi\rangle \in \mathbb{R}^+$ for any function $\varphi \in \mathscr{D}(\Omega)$ with values in $\mathbb{R}_+$.

Proposition 2.4.6 If $T \in \mathscr{D}'(\Omega)$ is non-negative, then it is a distribution of order 0.

Proof.— Take $T \in \mathscr{D}'(\Omega)$ a nonnegative distribution. Let $K \subset \Omega$ be a compact subset, and $\chi \mathscr{D}(\Omega)$ a plateau function above K. For $\varphi \in \mathscr{D}_K(\Omega)$ real valued, one has

$$\forall x \in \Omega, \qquad -\chi \sup |\varphi| \leq \varphi(x) \leq \chi \sup |\varphi|,$$

hence, using the nonnegativity of T, we get:

$$\langle T, \varphi + \chi \sup |\varphi| \rangle \geq 0 \quad \text{and} \quad \langle T, \chi \sup |\varphi| - \varphi \rangle \geq 0.$$

These two inequalities can be summarized as:

$$|\langle T, \varphi \rangle| \leq \langle T, \chi \rangle \sup |\varphi|.$$

Now, if $\varphi \in \mathscr{D}_K(\Omega)$ is complex valued, we decompose it as $\varphi = \varphi_1 + i\varphi_2$ with φ_1, φ_2 real valued. Applying the above reasoning to φ_1, φ_2, we get:

$$|\langle T, \varphi \rangle| = |\langle T, \varphi_1 + i\varphi_2 \rangle| \leq |\langle T, \varphi_1 \rangle| + |\langle T, \varphi_2 \rangle| \leq C \sup |\varphi_1| + C \sup |\varphi_2| \leq C \sup |\varphi|,$$

which concludes the proof of the proposition. □

From Proposition 2.4.6 — and the fact that adapted regularisation procedures preserve the nonnegativity property —, nonnegative distributions therefore extend to nonnegative linear forms on $\mathcal{C}_0^0(\Omega)$. By the Riesz representation theorem, nonnegative linear forms on $\mathcal{C}_0^0(I)$ are in 1-to-1 correspondence with positive Borel measures on Ω which are finite on compact subsets, which are also called *Radon measures*.

Theorem 2.4.7 For every nonnegative distribution T on Ω, there exists a unique Radon measure μ such that

$$\forall \varphi \in \mathcal{C}_0^0(\Omega)\ ,\ \langle T, \varphi \rangle = \int_\Omega \varphi \, d\mu\ .$$

2.4.5 The measure $\delta(f)$ and the superficial measure associated to $\{f = 0\}$.

Let Ω be open subset of $\mathbb{R}^d$, and let $f : \Omega \to \mathbb{R}$ be a $\mathcal{C}^1$ function which vanishes transversally:

$$\forall x \in \Omega\ ,\ f(x) = 0 \Rightarrow \nabla f(x) \neq 0\ . \tag{2.4.2}$$

We denote by $\Sigma = \{x \in \Omega,\ f(x) = 0\}$. According to the Lemma 2.4.4, Σ is a $\mathcal{C}^2$ hypersurface in $\mathbb{R}^d$. The following theorem constructs positive measures supported on Σ.

Theorem 2.4.8 There exists a positive distribution $\delta(f) \in \mathscr{D}'(\Omega)$, supported by Σ, such that, for every function $\rho \in \mathcal{C}^0_0(\mathbb{R})$ such that

$$\int_{\mathbb{R}} \rho(t)\, dt = 1 \ ,$$

the family of distributions associated with the L^1_{loc} functions $\mu_\varepsilon : \Omega \to \mathbb{R}$ defined by

$$\mu_\varepsilon(x) := \rho \frac{1}{\varepsilon} \rho\Big(\frac{f(x)}{\varepsilon}\Big) = \rho_\varepsilon \circ f, \qquad \epsilon \in]0,1],$$

converges in $\mathscr{D}'(\Omega)$ to $\delta(f)$ when $\varepsilon \searrow 0$.

Furthermore, if $g : \Omega \to \mathbb{R}$ is another $\mathcal{C}^1$ function which satisfies (2.4.2) and $\Sigma = \{x \in \Omega, g(x) = 0\}$, then

$$|\nabla g|\, \delta(g) = |\nabla f|\, \delta(f) \ . \tag{2.4.3}$$

Proof.— 1) By the gluing principle for convergence of distributions — Proposition 2.3.26 —, it is enough to prove that every point a of Ω admits an open neighborhood V_a on which μ_ε converges in the sense of distributions. Moreover, the gluing principle shows that the positivity of this limit near each point a implies that the limit is a positive distribution on Ω — see the remark after Proposition 2.3.8.

2) Let $a \in \Omega$. If $f(a) \neq 0$, then the continuity of f implies that $|f| \geq c > 0$ on some neighborhood of a, so that, since ρ is compactly supported in $\mathbb{R}$, μ_ε is identically zero on this neighborhood if ε is small enough. Hence μ_ε tends to 0 in a neighborhood of a. This shows that the support of the limiting distribution must be contained in Σ.

3) Now assume that $a \in \Sigma$, in other words that $f(a) = 0$. By the assumption (2.4.2), $\nabla f(a) \neq 0$, so there exists $j \in \{1, \dots, d\}$ such that $\partial_j f(a) \neq 0$. Let us assume for instance that $j = 1$, and $\partial_1 f(a) > 0$.

Let us write points in $\mathbb{R}^d$ as $x = (x_1, y)$, with $x_1 \in \mathbb{R}$ and $y \in \mathbb{R}^{d-1}$. As used in Theorem 2.4.1, by the implicit function theorem there exists an open interval $I \subset \mathbb{R}$ and an open subset $W \subset \mathbb{R}^{d-1}$ such that $a = (a_1, b) \in I \times W \subset \Omega$, and a $\mathcal{C}^1$ function $q : W \to I$ such that

$$\forall (x_1, y) \in I \times W \ , \ f(x_1, y) = 0 \iff x_1 = q(y) \ .$$

We then use the Taylor-Lagrange expansion in the x_1 variable, to write

$$\begin{aligned} \forall (x_1, y) \in I \times W, \quad f(x_1, y) &= \underbrace{f(q(y), y)}_{0} + \big(x_1 - q(y)\big) \underbrace{\int_0^1 \partial_1 f\Big(q(y) + t\big(x_1 - q(y)\big), y\Big)\, dt}_{m(x_1,y)} \\ &= m(x_1, y)\big(x_1 - q(y)\big) \ , \end{aligned}$$

where we used the fact that $f(q(y), y) = 0$ for all $y \in W$. We observe that $m\big(q(y), y\big) = \partial_1 f\big(q(y), y\big) > 0$. By the continuity of $\partial_1 f$, if W, I are chosen small enough, $\partial_1 f(x_1, y) > 0$ for $(x_1, y) \in I \times W$; as a result, there exists $0 < c < C$ such that

$$\forall (x_1, y) \in I \times W, \qquad c \leq m(x_1, y) \leq C\,.$$

Let us now choose a test function $\varphi \in \mathscr{D}(I \times W)$, and compute

$$\langle \mu_\epsilon, \varphi \rangle = \int_{I \times W} \varphi(x)\, \mu_\varepsilon(x)\, dx = \int_W \int_I \varphi(x_1, y)\, \rho_\varepsilon\Big(m(x_1, y)\big(x_1 - q(y)\big)\Big)\, dx_1\, dy\ .$$

Extending φ by 0 to $(\mathbb{R} \setminus I) \times W$, we can write the inner integral as

$$\int_{\mathbb{R}} \varphi(x_1, y)\, \frac{1}{\varepsilon} \rho\left(\frac{m(x_1, y)\big(x_1 - q(y)\big)}{\varepsilon} \right)\, dx_1 = \int_{\mathbb{R}} \varphi\big(q(y) + \varepsilon z, y\big)\, \rho\Big(m\big(q(y) + \varepsilon z, y\big)\, z\Big)\, dz\,,$$

where we used the change of variables $x_1 = q(y) + \varepsilon z$ (here y is fixed). Since $m \geq c$ on $I \times W$ and $\operatorname{supp} \rho \subset [-A, A]$ for some $A > 0$, the argument in the integral is supported in $\{|z| \leq A/c\}$.

Let us restore the integration over $y \in W$. Since the integrand is dominated by $\|\rho\|_\infty \|\varphi\|_\infty$ on the bounded set $[-A/c, A/c] \times W$, we get by dominated convergence:

$$\int_W \int_{-A/c}^{A/c} \varphi\big(q(y) + \varepsilon z, y\big)\, \rho\Big(m\big(q(y) + \varepsilon z, y\big)\, z\Big)\, dz\, dy\ \epsilon \overset{\to}{\to} 0\ \int_W \int_{-A/c}^{A/c} \varphi\big(q(y), y\big)\, \rho\Big(m\big(q(y), y\big)\, z\Big)\, dz\, dy$$

$$= \int_W \frac{\varphi}{m}$$

$$= \int_W \frac{\varphi}{\partial_1}$$

In the second equality we have used the change of variables $t = m\big(q(y), y\big) z$, and the normalization $\int \rho(t)\, dt = 1$.

Relaxing the assumption $f\big(q(y), y\big) > 0$ to $f\big(q(y), y\big) \neq 0$, we prove more generally that:

$$\lim_{\epsilon \to 0} \int_{I \times W} \varphi(x)\, \mu_\varepsilon(x)\, dx = \int_W \frac{\varphi(q(y), y)}{|\partial_1 f\big(q(y), y\big)|}\, dy\,. \tag{2.4.4}$$

This shows that, on the neighbourhood $I \times W$, the distributions μ_ε have a limit when $\varepsilon \searrow 0$. The above formula shows that the limit distribution, which we denote by $\delta(f)_{|W \times I|}$ is supported on $\big\{(q(y), y);\ y \in W\big\} = \Sigma \cap (I \times W)$, and is nonnegative. This limit distribution does not depend on the choice of ρ, but it depends on the function f used to define Σ.

4) In order to let appear $|\nabla f|$ in the above expression, we rewrite the limiting distribution in a different way. The identity $f\big(q(y), y\big) = 0$ for all $y \in W$ can be differentiated w.r.t. $y \in W$: from the chain rule one gets

$$\forall j = 2, \ldots, d, \quad \partial_{y_j} q(y)\, \partial_{x_1} f\big(q(y), y\big) + \partial_{y_j} f\big(q(y), y\big) = 0\,.$$

Putting these equations together, we obtain the vector identity:

$$\nabla f\big(q(y), y\big) = \partial_{x_1} f\big(q(y), y\big) \begin{pmatrix} 1 \\ -\partial_{y_2} q(y) \\ \vdots \\ -\partial_{y_d} q(y) \end{pmatrix}.$$

Taking the norms of these vectors, we find

$$|\nabla f\big(q(y),y\big)| = |\partial_{x_1} f\big(q(y),y\big)|\,\big(1+|\nabla q(y)|^2\big)^{1/2}$$

Substituting this identity in (2.4.4) we find

$$\lim_{\epsilon\to 0}\int_{I\times W}\varphi(x)\,\mu_\varepsilon(x)\,dx = \int_W \varphi(q(y),y)\frac{\big(1+|\nabla q(y)|^2\big)^{1/2}}{||\nabla f\big(q(y),y\big)|}\,dy := \langle \delta(f)_{|I\times W},\varphi\rangle\,.$$

The distribution $\delta(f)_{|I\times W}$ is nonnegative, hence it defines a Radon measure supported on $\Sigma\cap(I\times W)$.

5) It makes sense to multiply this measure by the continuous function $|\nabla f|\in C^0(I\times W,\mathbb{R}_+^*)$: one obtains another Radon measure, which we denote $|\nabla f|\,\delta(f)$:

$$\forall\varphi\in\mathcal{C}_0^0(I\times W),\quad \langle|\nabla f|\,\delta(f),\varphi\rangle = \int_W \varphi(q(y),y)\big(1+|\nabla q(y)|^2\big)^{\frac{1}{2}}\,dy\,.$$

We notice that the integral on the right hand side does not depend on the function f any longer, but only on the function q defining Σ through its graph.

If g is another $\mathcal{C}^1$ transversally vanishing function such that and $\Sigma=\{x\in\Omega, g(x)=0\}$, the same computation as with f yields — possibly taking a different open neighbourhood of a, $I'\times W'$:

$$\forall\varphi\in\mathcal{C}_0^0(I'\times W'),\quad \langle|\nabla g|\,\delta(g),\varphi\rangle = \int_{W'} \varphi(q(y),y)\big(1+|\nabla q(y)|^2\big)^{\frac{1}{2}}\,dy\,.$$

As a result, the distributions $|\nabla g|\,\delta(g)$ and $|\nabla f|\,\delta(f)$ are equal in $(I\cap I')\times(W\cap W')$.

By the gluing principle, these distributions are equal in the whole of Ω:

$$|\nabla g|\,\delta(g) = |\nabla f|\,\delta(f)\,.$$

□

From this proposition, we infer that the positive distribution $|\nabla f|\,\delta(f)$ does not depend on the choice of the transversally vanishing function f used to define Σ. This positive distribution is intrinsically associated to Σ.

Definition 2.4.9 The superficial measure of Σ is the positive distribution

$$\sigma = |\nabla f|\,\delta(f)\;,$$

where f is any $\mathcal{C}^1$ transversally vanishing function such that $\Sigma=\{f=0\}$.

Let us retain from the above definition and from the proof of Theorem 2.4.8 the following two important facts about the superficial measure on Σ:

i) if $\Sigma = \{f = 0\}$ with f transversally vanishing, then, for every $\varphi \in \mathcal{C}_0^0(\Omega)$,

$$\int_\Omega \frac{1}{\varepsilon}\rho\left(\frac{f(x)}{\varepsilon}\right)\varphi(x)\,dx\ \varepsilon \overset{\to}{\to} 0 \int_\Omega \varphi(x)\,\frac{d\sigma(x)}{|\nabla f(x)|}\ .$$

ii) if $\Sigma \cap (I \times W) = \{(q(y), y), y \in W\}$, where $q : W \to I$ is a $\mathcal{C}^1$ function, then, for every continuous function φ supported into $I \times W$,

$$\int_{I\times W} \varphi(x)\,d\sigma(x) = \int_W \varphi(q(y), y)\left(1 + |\nabla q(y)|^2\right)^{\frac{1}{2}}\,dy\ . \tag{2.4.5}$$

This last expression admits a geometric interpretation: if we call $|dy|$ the area of an infinitesimal element of $\mathbb{R}^{d-1}$ near the point y, then $(1 + |\nabla q(y)|^2)^{\frac{1}{2}}|dy|$ is the area of the lift of this infinitesimal element to Σ; the extra factor is due to the slope of this lifted element.

Example 2.4.10 (The superficial measure on a sphere) Let $R > 0$. Then the function

$$\begin{array}{rcl} f_R : \mathbb{R}^d \setminus \{0\} & \to & \mathbb{R} \\ x & \mapsto & |x| - R \end{array}$$

is in $\mathcal{C}^1(\mathbb{R}^d \setminus \{0\})$, and $\{f_R = 0\}$ defines the sphere S_R of radius R centered at the origin. Notice that

$$\nabla f_R(x) = \frac{x}{|x|}\ ,$$

which nowhere vanishes. The intersection of S_R with the "upper half-space" $]0, +\infty[\times\mathbb{R}^{d-1}$ can be described by the equation $\{x_1 = q(y),\ |y| < R\}$, where

$$q(y) = \sqrt{R^2 - |y|^2} \quad |y| < R\ .$$

Notice that

$$\nabla q(y) = \frac{-y}{\sqrt{R^2 - |y|^2}}\ , \qquad 1 + |\nabla q(y)|^2 = \frac{r^2}{r^2 - |y|^2}\ ,$$

so the superficial measure σ_R on S_R is given by $\sigma_R = \delta(f_r)$. For any $\varphi \in C_0^0(\mathbb{R}_+^* \times \mathbb{R}^{d-1})$,

$$\int_{\mathbb{R}_+^*\times\mathbb{R}^{d-1}} \varphi(x)\,d\sigma_R(x) = \int_{|y|<R} \varphi\left(\sqrt{R^2 - |y|^2}, y\right)\,\frac{R\,dy}{\sqrt{R^2 - |y|^2}}\ .$$

Let us prove the following simple relation between the measures σ_R and σ_1:

$$\forall\varphi \in \mathcal{C}_0^0(\mathbb{R}^d) \qquad \int_{\mathbb{R}^d} \varphi(x)\,d\sigma_R(x) = R^{d-1}\int_{\mathbb{R}^d} \varphi(Ry)\,d\sigma_1(y)\ . \tag{2.4.6}$$

Indeed, using the original definition of σ_R from the function f_R, we find

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(x)\,d\sigma_R(x) &= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\int_{\mathbb{R}^d} \rho\left(\frac{|x| - R}{\varepsilon}\right)\varphi(x)\,dx \\ &\overset{x=Ry}{=} \lim_{\varepsilon\to 0} \frac{R}{\varepsilon}R^{d-1}\int_{\mathbb{R}^d} \rho\left(\frac{R(|y| - 1)}{\varepsilon}\right)\varphi(Ry)\,dy \\ &= R^{d-1}\int_{\mathbb{R}^d} \varphi(Ry)\,d\sigma_1(y)\ , \end{aligned}$$

using the rescaled small parameter $\varepsilon' = \frac{\varepsilon}{r}$.

Using this identity, in dimension $d = 2$ we may parametrize the measures σ_R by the angular coordinate $\theta \in [0, 2\pi[$, and obtain

$$\int_{\mathbb{R}^2} \varphi(x)\, d\sigma_R(x) = \int_0^{2\pi} \varphi(R\cos\theta, R\sin\theta)\, R\, d\theta \ .$$

2.4.6 Integration on level sets : the smooth coarea formula

The above definition of the superficial measure easily leads to an important formula of integral calculus, which can be viewed as a "nonlinear Fubini theorem".

Let $f = \Omega \to \mathbb{R}$ be a $\mathcal{C}^1$ function such that

$$\forall x \in \Omega\ ,\ \nabla f(x) \neq 0\ . \tag{2.4.7}$$

It is easy to prove that the set $f(\Omega) := \{f(x), x \in \Omega\}$ is an open subset of $\mathbb{R}$ — hence is an open interval if Ω is connected. Indeed, if $t_0 = f(a)$ with $a \in \Omega$, and if, say, $\partial_1 f(a) \neq 0$, then, on a small neighborhood $I \times W$ of $a = (a_1, b)$, we have $\partial_1 f \neq 0$, hence the function $x_1 \in I \mapsto f(x_1, b)$ is strictly monotone, and consequently its range is an open interval of $\mathbb{R}$ containing t_0.

For every $t \in f(\Omega)$, denote by Σ_t the level set $\{x \in \Omega\,;\ f(x) = t\}$, and by σ_t the superficial measure on Σ_t, constructed in the previous section. The hypersurfaces $(\Sigma_t)_{t\in f(\Omega)}$ form a *smooth foliation* of Ω.

The smooth coarea formula takes the following form. It can be viewed as a Fubini theorem, with respect to the foliation $\Omega = \bigsqcup_{t\in f(\Omega)} \Sigma_t$.

Proposition 2.4.11 (Smooth coarea formula) For every $\varphi \in \mathcal{C}_0^0(\Omega)$,

$$\int_\Omega \varphi(x)\, dx = \int_{f(\Omega)} \left(\int_{\Sigma_t} \varphi(x)\, \frac{d\sigma_t(x)}{|\nabla f(x)|} \right) dt\ .$$

Proof.— We will go back to the construction of σ_t as in Theorem 2.4.8. We thus use a convolution kernel $\rho \in \mathcal{C}_0^0(\mathbb{R})$ of integral 1. For every $x \in \Omega$, we observe that, for every $\varepsilon > 0$,

$$1 = \frac{1}{\varepsilon} \int_{\mathbb{R}} \rho\left(\frac{f(x) - t}{\varepsilon}\right) dt\ .$$

We plug this identity in the x integral and apply the Fubini theorem (the integrand is continuous and compactly supported in (x, t)):

$$\int_\Omega \varphi(x)\, dx = \int_{\mathbb{R}} \left(\int_\Omega \frac{1}{\varepsilon} \rho\left(\frac{f(x) - t}{\varepsilon}\right) \varphi(x)\, dx \right) dt\ .$$

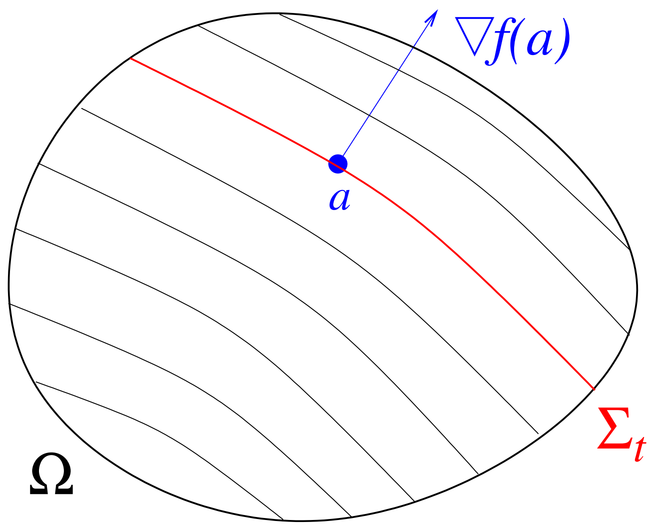


Figure 2.8: Foliation of Ω by the hypersurfaces $\Sigma_t = f^{-1}(t)$.

Let us call the inner integral

$$I_\varepsilon(t) := \int_\Omega \frac{1}{\varepsilon}\rho\left(\frac{f(x)-t}{\varepsilon}\right)\varphi(x)\,dx\,.$$

This is a continuous function of $t \in \mathbb{R}$. For $t \in \mathbb{R}$ fixed, the Theorem 2.4.8 shows that, when $\varepsilon \to 0$,

$$I_\varepsilon(t) \to I_0(t) := \begin{cases} 0 & \text{if } t \notin f(\Omega) \\ \int_{\Sigma_t} \varphi(x)\,\frac{d\sigma_t(x)}{|\nabla f(x)|} & \text{if } t \in f(\Omega)\,. \end{cases}$$

Since I_ε are continuous functions, their pointwise limit I_0 is necessarily measurable.

To show the requested formula, we need to show that $\int_\mathbb{R} I_\varepsilon(t)\,dt \to \int_\mathbb{R} I_0(t)\,dt$ as $\varepsilon \to 0$. This will be done by invoking the Dominated Convergence Theorem. First we notice that, because φ and ρ are compactly supported, $I_\varepsilon(t)$ is supported inside a compact subset $K \Subset \mathbb{R}$, which can be chosen uniform when $\varepsilon \in]0,1]$. So we just need to prove that the family $(I_\varepsilon)_{\varepsilon\in]0,1]}$ is dominated by an integrable function.

By the gluing principle, it is enough to assume that the test function φ is supported inside a small open subset V of Ω. We may then apply the methods of Theorem 2.4.8, but simultaneously for a family of hypersurfaces $(\Sigma_t)_{t\in J}$.

Assuming that $\partial_1 f \neq 0$ on a small open set $V = I \times W$. The equation $\{(x_1, y) \in V,\ f(x_1, y) = t\}$ defines the set $\Sigma_t \cap V$, which is nonempty only for t in some short interval $J \subset \mathbb{R}$. Then, the implicit function theorem shows that, if V is small enough, $\Sigma_t \cap V$ can be described by the graph of a $\mathcal{C}^1$ function $q(\cdot, t)$. Actually, q is $\mathcal{C}^1$ in both variables (y, t).

We can reproduce the proof of Theorem 2.4.8 with the additional parameter t. For any $\varphi \in \mathscr{D}(I \times W)$,

$$\begin{aligned} I_\varepsilon(t) &= \int_W \int_I \frac{1}{\varepsilon}\rho\left(\frac{m(x_1,y,t)\big(x_1 - q(y,t)\big)}{\varepsilon}\right)\varphi(x_1,y)\,dx_1\,dy \\ &= \int_W \int_{\mathbb{R}} \rho\Big(m\big(q(y,t)+\varepsilon z, y\big) z\Big)\,\varphi\big(q(y,t)+\varepsilon z, y\big)\,dz\,dy\,. \end{aligned}$$

The support integral is uniformly bounded, since φ is compactly supported and, as we showed Like in the proof of Theorem 2.4.8, we have a uniform lower bound $|m(x_1,y,t)| \geq c > 0$ for $(x_1,y,t) \in I \times W \times J$, which implies that the above integrand is supported in some bounded set $[-A,A] \times W$. The functions $(I_\varepsilon)_{\varepsilon \in]0,1]}$ are then dominated by $\|\varphi\|_\infty \|\rho\|_\infty \mathbb{1}_{[-A,A]\times W}$. This completes the proof of the dominated convergence theorem, hence of the coarea formula. □

Corollary 2.4.12 (Integration on spherical level sets) For every $\varphi \in \mathcal{C}^0_0(\mathbb{R}^d \setminus \{0\})$, we have

$$\int_{\mathbb{R}^d \setminus \{0\}} \varphi(x)\,dx = \int_0^\infty \int_{S_1} \varphi(r\omega)\,d\sigma_1(\omega)\,r^{d-1} dr\ .$$

Proof.— Apply Proposition 2.4.11 to the function $f : x \in \mathbb{R}^d \setminus \{0\} \mapsto |x|$, which defines the foliation into spheres $(S_t)_{t>0}$, and apply the connection (2.4.6) between measure on S_t and on S_1. □

2.4.7 Superficial measure on a closed hypersurface, and the jump formula for a regular open subset

In the previous section, we defined a hypersurface Σ as the zero locus of a single, transversally vanishing function $f \in C^1(\Omega,\mathbb{R})$, and used this definition to cook up the superficial measure σ on $\Sigma = \Sigma_f$.

In the present section, we will proceed differently, namely by starting from a hypersurface $\Sigma \subset \Omega$, as described in Section 2.4.2, without assuming that it can be defined as the zero locus of a global function f.

We start from a definition of a hypersurface which is more general than the one used in the previous section, but slightly more restrictive than the one used in Theorem 2.4.1.

Definition 2.4.13 Let Ω be an open subset of $\mathbb{R}^d$. A *closed* $\mathcal{C}^1$ hypersurface of Ω is a closed subset $\Sigma \subset \Omega$ with the following property: every $a \in \Sigma$ admits an open neighborhood V in Ω, such that there exists a transversally vanishing function $f \in \mathcal{C}^1(V,\mathbb{R})$ with

$$\Sigma \cap V = \{x \in V, f(x) = 0\}\,.$$

As opposed to the situation of the previous section, we do not impose the existence of a *global* function f whose zero locus equals Σ. We are then close to the definition of Theorem 2.4.1.

Compared with that theorem, we add the condition of *closedness* of Σ; this condition forbids the hypersurfaces with boundaries appearing in section 2.4.3 (see the discussion at the end of that section).

Proposition 2.4.14 Let Σ be a *closed* $\mathcal{C}^1$ hypersurface of Ω.

Then there exists a unique positive distribution σ supported in Σ, such that, for every open subset $V \subset \Omega$ and transversally vanishing function $f \in \mathcal{C}^1(V, \mathbb{R})$ defining $\Sigma \cap V$, we have as distributions in V:

$$|\nabla f| \, \delta(f) = \sigma_{|V} \ .$$

Furthermore, $\operatorname{supp} \sigma = \Sigma$.

Proof.— Let us show that the existence and uniqueness of σ follow from Theorem 2.4.8 and the gluing principle.

Indeed, given $a \in \Sigma$, consider $V_a \subset \Omega$ and $f_a : V_a \to \mathbb{R}$ as in the definition above. Applying Theorem 2.4.8, we define the measure $\sigma_a \in \mathscr{D}'(V_a)$ by

$$\sigma_a = |\nabla f_a| \, \delta(f_a) \ .$$

In the open set $\Omega \setminus \Sigma$, we consider the trivial distribution $\tilde{\sigma} = 0$ in $\mathscr{D}'(\Omega \setminus \Sigma)$. We have an open covering of Ω:

$$\Omega = (\Omega \setminus \Sigma) \cup \bigcup_{a \in \Sigma} V_a \, .$$

Let us try to apply the gluing principle to this situation. The last part of Theorem 2.4.8 shows that, if $V_a \cap V_b \neq \emptyset$, then the measures σ_a and σ_b, restricted on $V_a \cap V_b$, will coincide. Besides, since σ_a are all supported on Σ, σ_a and $\tilde{\sigma}$ coincide on $V_a \setminus \Sigma$. The local measures (σ_a) and $\tilde{\sigma}$ are thus compatible with one another, and define a single measure σ on Ω, supported on Σ. The uniqueness of each σ_a implies the uniqueness of σ.

Notice that the closedness of Σ is crucial in this proof: we use it when defining $\tilde{\sigma}$ on the open set $\Omega \setminus \Sigma$. □

Definition 2.4.15 The measure σ constructed in the previous proposition is called the superficial measure on the closed hypersurface Σ.

The above definition Σ allowed interesting topologies, for which there is no possible way to define an "interior" and an "exterior" to our hypersurface. The following definition will lead us to examples where the closed curve Σ automatically splits $\Omega \setminus \Sigma$ into an "interior region" and an "exterior region".

Definition 2.4.16 Let U be an open subset of Ω. We say that U is a *regular open subset of* Ω *of class* $\mathcal{C}^1$ if every point $a \in \partial U$ admits an open neighborhood V in Ω, such that there exists a transversally vanishing function $f \in \mathcal{C}^1(V, \mathbb{R})$ with

$$U \cap V = \{x \in V, f(x) > 0\}.$$

In such a situation, it easy to check that ∂U is a closed $\mathcal{C}^1$ hypersurface of Ω. Indeed, with the notation of the above definition, one verifies that

$$\partial U \cap V = \{x \in V, f(x) = 0\},$$

and the transversal vanishing of f implies that $\nabla f(x) \neq 0$ for $x \in \partial U \cap V$.

Yet, a regular open set is not only an open set whose boundary is a closed hypersurface. The above definition also imposes that the open subset U is locally *on one side* (say, the "interior") of this hypersurface ! As a counterexample, $U = \mathbb{R}^d \setminus \{x_1 = 0\}$ is *not* a regular subset of $\mathbb{R}^d$, though its boundary is the closed hypersurface $\{x_1 = 0\}$.

This distinction between regular hypersurfaces Σ admitting an "interior side", and regular hypersurfaces Σ not admitting one, is of topological nature. Locally, these hypersurfaces are not distinguishable.

The existence of an "interior side" allows to introduce the following objects.

Definition 2.4.17 Let Ω be an open subset of $\mathbb{R}^d$.

- If Σ is a closed $\mathcal{C}^1$ hypersurface of Ω and $a \in \Sigma$, the *normal line to* Σ *at* a is the affine line $a + \mathbb{R}\nabla f(a)$, where f is any function defining Σ locally near a as in Definition 2.4.13.
- If U is a regular open subset of Ω of class $\mathcal{C}^1$, the *inward unit normal vector* to U at $a \in \partial U$ is the unit vector
$$N^{\text{int}}(a) = \frac{\nabla f(a)}{|\nabla f(a)|},$$
where f is any function defining ∂U locally as in Definition 2.4.16.

  The outward unit normal vector to U at $a \in \partial U$ is
$$N^{\text{ext}}(a) = -N^{\text{int}}(a).$$

These definitions make sense because of the independence of the objects with respect to the specific function f. In the case of a closed hypersurface, we already checked, in the proof of Theorem 2.4.8, that the vectors $\nabla f(a)$ are all proportional to one another, so they define the same line. There exist only two unit vectors on the normal line to ∂U at $a \in \partial U$. In the general case, there is no canonical way to distinguish them. On the opposite, in the case of a regular open subset U, we can define an "inward", resp. an "outward" vector. These two vectors can be characterized by the following property: for $\varepsilon > 0$ small enough, $a + \varepsilon N^{\text{int}}(a) \in U$, while $a + \varepsilon N^{\text{ext}}(a) \in \Omega \setminus \overline{U}$.

Finally, notice that the mapping $N^{\text{int}} : \partial U \to \mathbb{R}^d$ is continuous. We call it the *inward unit normal vector field to U*. A similar definition holds for the *outward unit normal vector field to U*.

These stuctures attached to a regular open subset U will allow us to compute explicitly the first derivative of the characteristic function $\mathbb{1}_U$.

Theorem 2.4.18 (The jump formula) Let U be a regular open subset of Ω of class $\mathcal{C}^1$. Denote by σ the superficial measure on the closed hypersurface ∂U, and by $N^{\text{int}} : \partial U \to \mathbb{R}^d$ the inward unit normal vector field to U. Then, in $\mathscr{D}'(\Omega)$,

$$\forall j \in \{1, \dots, d\} \ , \ \partial_j(\mathbf{1}_U) = N_j^{\text{int}} \sigma \ ,$$
$$\text{or in condensed form:} \quad \nabla \mathbf{1}_U = N^{\text{int}} \sigma \, .$$

Proof.— By the gluing principle, it is enough to prove this identity near every point $a \in \Omega$.

1) If $a \notin \partial U$, then $\mathbf{1}_U$ is a constant function in a neighborhood of a, so that $\partial_j(\mathbf{1}_U) = 0$ in this neighborhood, which is also the case of σ.

2) Assume $a \in \partial U$, and let $f : V \to \mathbb{R}$ be a function near a as in Definition 2.4.16. Denote by $\chi : \mathbb{R} \to \mathbb{R}$ a smooth "step function", that is a $\mathcal{C}^1$ function such that

$$\chi(z) = \begin{cases} 0 & \text{if } z < 0 \\ 1 & \text{if } z \geq 1 \, . \end{cases}$$

Then it is easy to check that

$$\forall x \in V \ , \ \mathbf{1}_U(x) = \lim_{\varepsilon \to 0} \chi\left(\frac{f(x)}{\varepsilon}\right) \, .$$

Since the right hand side is uniformly bounded, the dominated convergence theorem implies that this convergence also holds in $\mathscr{D}'(V)$. By the continuity of the derivative map in $\mathscr{D}'(V)$, we have

$$\partial_j(\mathbf{1}_U) = \lim_{\varepsilon \to 0} \partial_j\left(\chi\left(\frac{f(x)}{\varepsilon}\right)\right) = \lim_{\varepsilon \to 0} \frac{\partial_j f}{\varepsilon} \chi'\left(\frac{f(x)}{\varepsilon}\right) \, .$$

The function $\rho = \chi'$ belongs to $\mathcal{C}_0^0(\mathbb{R})$ and satisfies

$$\int_{\mathbb{R}} \rho(z)\, dz = \int_0^\infty \chi'(z)\, dz = 1 \, .$$

Consequently, for every $\varphi \in \mathcal{C}_0^0(V)$,

$$\int_\Omega \varphi(x)\,\frac{1}{\varepsilon}\,\chi'\left(\frac{f(x)}{\varepsilon}\right)\,dx \to \langle \delta(f), \varphi\rangle\ ,$$

so that

$$\partial_j(\mathbf{1}_U) = \partial_j f\,\delta(f) = \frac{\partial_j f}{|\nabla f|}\,\sigma = N_j^{\text{int}}\,\sigma\ .$$

□ As a corollary of this jump formula, we obtain a well-known formula of integral calculus, dating back to the 19th century. It expresses the integral of a derivative over a regular open set U, in terms of boundary data.

Corollary 2.4.19 (The Gauss–Green formula, 1) Let U be a regular open subset of class $\mathcal{C}^1$ in Ω (U may be unbounded). For every $\varphi \in \mathcal{C}_0^1(\Omega)$ and any $j = 1, \ldots, d$:

$$\int_U \partial_j\varphi(x)\,dx = \int_{\partial U} N_j^{\text{ext}}(x)\varphi(x)\,d\sigma(x)\ .$$

Proof.— First assume $\varphi \in \mathscr{D}(\Omega)$, and write the integral on the left as a distributional bracket:

$$\int_U \partial_j\varphi(x)\,dx = \langle \mathbb{1}_U, \partial_j\varphi\rangle = -\langle \partial_j(\mathbf{1}_U), \varphi\rangle = -\langle N_j^{\text{int}}\sigma, \varphi\rangle = \int_{\partial U} N_j^{\text{ext}}(x)\varphi(x)\,d\sigma(x)\ ,$$

where we used the jump formula in the third equality.

The general case $\varphi \in \mathcal{C}_0^1(\Omega)$ follows by a density argument. Namely, one can construct a sequence $(\phi_n \in \mathscr{D}(\Omega))_n$ such that the φ_n are supported in a common compact set, and $\|\varphi_n - \phi\|_{\mathcal{C}^1} \to 0$ as $n \to \infty$. The Gauss-Green formula holds for all the φ_n, and the $\mathcal{C}^1$ convergence shows that each side of the equality converge to $\int_U \partial\phi\,dx$, resp. $\int_{\partial U} N_j^{\text{ext}}\varphi\,d\sigma$. This shows that these two limits are equal. □

Let us focus on the special case where U is a bounded regular open subset in $\mathbb{R}^d$.

Definition 2.4.20 Let U be a bounded regular open subset of class $\mathcal{C}^1$ in $\mathbb{R}^d$. We denote by $\mathcal{C}^1(\overline{U})$ the space of restrictions to U of $\mathcal{C}^1$ functions on $\mathbb{R}^d$.

This definition allows us to formulate a Gauss-Grenn formula, by specifying a function φ only in $\bar{U}$.

Corollary 2.4.21 (The Gauss–Green formula, compact set) Let U be a regular open subset of class $\mathcal{C}^1$ in $\mathbb{R}^d$, such that $\overline{U}$ is *compact*. For every $\varphi \in \mathcal{C}^1(\overline{U})$ and every $j = 1, \ldots, d$,

$$\int_U \partial_j\varphi(x)\,dx = \int_{\partial U} N_j^{\text{ext}}(x)\varphi(x)\,d\sigma(x)\ .$$

Proof.— If $\varphi \in \mathcal{C}^1(\overline{U})$, let $\tilde{\varphi} \in \mathcal{C}^1(\tilde{U})$ be an extension of φ to an open neighborhood $\tilde{U}$ of $\overline{U}$. Let χ be a cutoff function on $\overline{U}$, compactly supported in $\tilde{U}$ (here we use the fact that $\overline{U}$ is compact). Applying the Gauss–Green formula of Corollary 2.4.19 to $\chi\tilde{\varphi} \in \mathcal{C}_0^1(\mathbb{R}^d)$, we directly obtain the formula stated in the present corollary. □

Remark 2.4.22 Applying either Gauss-Green formula to a product $\varphi\psi$ of two functions, and using the Leibniz formula, we obtain a multidimensional integration by parts formula, for integrals over U. For instance, in the case $\overline{U} \Subset \mathbb{R}^d$ is compact and $\varphi, \psi \in \mathcal{C}^1(\overline{U})$, we get for each $j = 1, \ldots, d$:

$$\int_U \partial_j\varphi(x)\,\psi(x)\,dx = -\int_U \varphi(x)\partial_j\psi(x)\,dx + \int_{\partial U} N_j^{\text{ext}}(x)\varphi(x)\,\psi(x)\,d\sigma(x)\ .$$

The last integral over ∂U is the boundary term of this integration by parts.

2.5 Sobolev spaces in an open set

In this section we define, as we had done in Chapter 1, subspaces of L^2 functions admitting generalized derivatives, which are called Sobolev spaces. We then use these space to solve some elliptic PDEs on Ω.

We recall that, we identify a locally integrable function f with the associated distribution T_f.

2.5.1 Definition and general facts

Definition 2.5.1 Let $\Omega \subset \mathbb{R}^d$ be an open set, and $N \in \mathbb{N}$. A distribution $u \in \mathscr{D}'(\Omega)$ belongs to $H^N(\Omega)$ if, for all $\alpha \in \mathbb{N}^d$ such that $|\alpha| \leq N$, it holds that $\partial^\alpha u \in L^2(\Omega)$. We denote by $(\cdot|\cdot)_{H^N}$ the sesquilinear form defined on $H^N(\Omega) \times H^N(\Omega)$ by

$$(u|v)_{H^N} = \sum_{|\alpha|\leq N} (\partial^\alpha u|\partial^\alpha v)_{L^2}\ .$$

We also introduce the associated H^N norm

$$\|u\|_{H^N} = (u|u)_{H^N}^{\frac{1}{2}} = \left(\sum_{|\alpha|\leq N} \|\partial^\alpha u\|_{L^2}^2\right)^{\frac{1}{2}},$$

so that convergence of u_n to u in H^N is equivalent to the convergence of $\partial^\alpha u_n$ to $\partial^\alpha u$ in L^2 for every $|\alpha| \leq N$.

Notice that $H^0(\Omega) = L^2(\Omega)$. The structure of Hilbert space of $L^2(\Omega)$ is transferred to $H^N(\Omega)$, as shown by the next proposition.

Proposition 2.5.2 The sesquilinear form $(\cdot,\cdot)_{H^N}$ is a Hermitian scalar product, which makes $H^N(\Omega)$ a Hilbert space.

Proof.— The only nontrivial fact to show is the completeness of $H^N(\Omega)$ with respect to the H^N norm. Let (u_j) be a Cauchy sequence in $H^N(\Omega)$. For all $|\alpha| \leq N$, the sequence $(\partial^\alpha u_j)$ is Cauchy in L^2, thus converges in L^2 to some $v_\alpha \in L^2$. The convergence implies that $u_j \xrightarrow{L^2} v_0$ implies the convergence $u_j \xrightarrow{\mathscr{D}'} v_0$, so by continuity of the derivative on $\mathscr{D}'$, that $\partial^\alpha u_j \xrightarrow{\mathscr{D}'} \partial^\alpha v_0$. We thus identify the distributions $\partial^\alpha v_0 = v_\alpha$, and get the stronger convergence $\partial^\alpha u_j \xrightarrow{L^2} \partial^\alpha v_0$. This shows that $v_0 \in H^N$, and that $u_j \to v_0$ in $H^N(\Omega)$. □

Notice that, for $N \geq 1$, $u \in H^N(\Omega)$ if and only if $u \in H^1(\Omega)$ and, for every $j = 1,\ldots,d-1$, $\partial_j u \in H^{N-1}(\Omega)$. By induction on N, this reduces many properties of $H^N(\Omega)$ to the special case $N = 1$, on which we are going to focus in what follows.

2.5.2 Variational formulation of some elliptic problems

Definition 2.5.3 Given an arbitrary open subset Ω of $\mathbb{R}^d$, we denote by $H_0^1(\Omega)$ the closure of $\mathscr{D}(\Omega)$ in $H^1(\Omega)$.

Note that $H_0^1(\Omega)$ is a closed subspace of the Hilbert space $H^1(\Omega)$, hence it is a Hilbert space with the inner product $(\cdot|\cdot)_{H^1}$ defined in the previous section.

In what follows, we are going to use the Laplacian differential operator, acting on $T \in \mathscr{D}'(\Omega)$ by

$$\Delta T = \sum_{j=1}^{d} \partial_j^2 T\ .$$

Our goal will be to solve a well-known PDE involving the Laplacian. It is known in phyiscs as the *screened Poisson equation* arising in plasma physics, or the *(inhomogeneous) time-independent Klein-Gordon equation* in relativistic wave mechanics.

Theorem 2.5.4 Let Ω be an arbitrary open subset of $\mathbb{R}^d$. For each $f \in L^2(\Omega)$ (which is called the "source term", or "inhomogeneous term"), there exists a unique $u \in H_0^1(\Omega)$ such that

$$-\Delta u + u = f \quad \text{in } \mathscr{D}'(\Omega). \tag{2.5.8}$$

Proof.— The equation (2.5.8) in $\mathscr{D}'(\Omega)$ exactly means:

$$\forall\varphi \in \mathscr{D}(\Omega)\ , \quad \langle -\Delta u + u, \varphi\rangle = \int_\Omega f\varphi\, dx\ .$$

Since we are looking for $u \in H_0^1(\Omega)$, we know in particular that $u \in L^2(\Omega)$ and $\partial_j u \in L^2(\Omega)$ for $j = 1, \ldots, d$. Therefore, applying integration by parts in each of the variables x_j, the left hand side becomes

$$\langle -\Delta u + u, \varphi\rangle = \sum_{j=1}^d \int_\Omega \partial_j u \partial_j \varphi\, dx + \int_\Omega u\varphi\, dx = (\varphi|\overline{u})_{H^1}\ .$$

Notice the absence of boundary terms, due to the vanishing of φ near $\partial\Omega$.

Summing up, we are looking for some $u \in H_0^1(\Omega)$ such that,

$$\forall\varphi \in \mathscr{D}(\Omega)\ , \quad (\varphi|\overline{u})_{H^1} = \int_\Omega f\varphi\, dx\ .$$

Let us now extend the above identity to a larger class of test functions φ. Since both sides of the above equation are linear forms of φ which are continuous for the H^1 norm, this equality continuously extends to test functions in $H_0^1(\Omega)$, the closure of $\mathscr{D}(\Omega)$ in the H^1 topology.

Our problem is thus equivalent to finding $u \in H_0^1(\Omega)$ such that

$$\forall v \in H_0^1(\Omega)\ , \quad (v|\overline{u})_{H^1} = \int_\Omega f v\, dx\ .$$

We have transformed our PDE problem into a Hilbert space problem. Indeed, the linear form

$$v \in H_0^1(\Omega) \mapsto L(v) = \int_\Omega f v\, dx$$

is obviously continuous w.r.to the H^1 topology. As a result, by the Riesz representation theorem, this continuous linear form L can be represented by the inner product with a unique element $w_L \in H_0^1(\Omega)$: $L(v) = (v|w_L)_{H^1}$, for any $v \in H_0^1$. Hence this element w_L solves our problem by setting $u = \overline{w}$ — notice that the conjugate of a function in $H_0^1(\Omega)$ is still in $H_0^1(\Omega)$. □

Let us complete this subsection by some remarks about the possible extensions of the above Theorem.

Remark 2.5.5 *i)* The L^2 function f on the right hand side can be replaced more generally by any distribution which extends into a continuous linear form on $H_0^1(\Omega)$. The space of such distributions is denoted by $H^{-1}(\Omega)$. Hence for $T \in H^{-1}(\Omega)$, we can solve the equation $-\Delta u + u = T$ by a unique $u \in H_0^1(\Omega)$.

ii) If $q \in L^\infty(\Omega)$ and there exists $M \geq m > 0$ such that $M \geq q(x) \geq m$ almost everywhere on Ω, the sesquilinear form

(2.5.9) $$(u|v)_q := \sum_{j=1}^d (\partial_j u|\partial_j v)_{L^2} + \int_\Omega q u\overline{v}\, dx$$

is an inner product on $H^1(\Omega)$, inducing a norm which is equivalent to the H^1 norm. Indeed, for any $u \in H^1(\Omega)$,

$$m\|u\|_{L^2}^2 \leq \int q|u|^2\,dx \leq M\|u\|_{L^2}^2\,,$$

so that

$$\mathsf{min}(m,1)\|u\|_{H^1}^2 \leq (u|u)_q \leq \mathsf{max}(1,M)\|u\|_{H^1}^2$$

Therefore $H^1(\Omega)$ and $H_0^1(\Omega)$ are also Hilbert spaces for this new inner product. The equation

(2.5.10) $$-\Delta u + qu = T \quad \text{in } \mathscr{D}'(\Omega)\,,$$

for $T \in H^{-1}(\Omega)$, can thus be solved by the same method as above, producing a unique solution $u \in H_0^1(\Omega)$.

iii) The assumption $M \geq q \geq m > 0$ can be relaxed to $M \geq q \geq 0$ for special cases of open sets Ω, in particular if Ω is bounded. This comes from the following inequality, valid on bounded Ω.

Lemma 2.5.6 (Poincaré inequality) Let Ω be a *bounded* open subset of $\mathbb{R}^d$. For every $j \in \{1,\dots,d\}$, there exists $C > 0$ such that, for every $\varphi \in H_0^1(\Omega)$,

$$\|\varphi\|_{L^2} \leq C\|\partial_j\varphi\|_{L^2}\;.$$

By density, it is enough to prove this inequality for $\varphi \in \mathscr{D}(\Omega)$; by density of $\mathscr{D}$ in H_0^1, since both terms are controlled by the H^1 norm, the inequality will extend to all $\varphi \in H_0^1$.

Let us thus prove Lemma 2.5.6 for $\varphi \in \mathscr{D}(\Omega)$, say with $j = 1$. We split the variable $x = (x_1, y) \in \mathbb{R} \times \mathbb{R}^{d-1}$, and write for any $(x_1, y) \in \mathbb{R}^d$:

$$|\varphi(x_1,y)|^2 = -\int_{x_1}^{\infty} \partial_t(|\varphi(t,y)|^2)\,dt = -2\int_{x_1}^{\infty} \mathsf{Re}(\partial_t\varphi(t,y)\overline{\varphi}(t,y))\,dt\;.$$

Using the Cauchy-Schwarz inequality, we infer:

$$|\varphi(x_1,y)|^2 \leq 2\left(\int_{\mathbb{R}} |\partial_t\varphi(t,y)|^2\,dt\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}} |\varphi(t,y)|^2\,dt\right)^{\frac{1}{2}}\;.$$

We then integrate the above expressions over $y \in \mathbb{R}^{d-1}$, and apply the Cauchy-Schwarz inequality on the RHS:

$$\int_{\mathbb{R}^{d-1}} |\varphi(x_1,y)|^2\,dy \leq 2\left(\int_{\mathbb{R}\times\mathbb{R}^{d-1}} |\partial_t\varphi(t,y)|^2\,dt\,dy \int_{\mathbb{R}\times\mathbb{R}^{d-1}} |\varphi(t,y)|^2\,dt\,dy\right)^{\frac{1}{2}} = 2\,\|pa_t\varphi\|_{L^2}\,\|\varphi\|_{L^2}\;.$$

We see that the RHS is independent of x_1. Assuming $\Omega \subset]a, a+L[\times\mathbb{R}^{d-1}$, we may integrate over $x_1 \in]a, a+L[$, and get the bound

$$\|\,\varphi\|_{L^2}^2 = \int_{\mathbb{R}\times\mathbb{R}^{d-1}} |\varphi(x_1,y)|^2\,dx_1\,dy \leq 2L\,\|\partial_t\varphi\|_{L^2}\,\|\,\varphi\|_{L^2}\;.$$

This completes the proof of the lemma, with constant $C = 2L$. □

If Ω is bounded, in view of Lemma 2.5.6, the sesquilinear form $(.| .)_q$ defined in (2.5.9) is still an inner product on $H_0^1(\Omega)$, with a norm equivalent to the H^1 norm, provided $q \in L^\infty(\Omega)$ is just nonnegative — including q identically 0. Hence, under this more general assumption, it still possible to solve equation (2.5.10) on Ω bounded.

In particular, if Ω is bounded, for every $f \in H^{-1}(\Omega)$, there exists a unique solution $u \in H_0^1(\Omega)$ of the Poisson equation:

$$-\Delta u = f \quad \text{in } \mathscr{D}'(\Omega) .$$

We will come back to this equation at the end of this paragraph, in the particular case where Ω is regular. We will see that, in this case, the subspace $H_0^1(\Omega)$ of $H^1\Omega)$ can be described more explicitly.

2.5.3 Approximation by smooth functions

In this section, we will approximate functions $u \in H^1(\Omega)$ by smooth functions $\varphi \in C^infty(\overline{\Omega})$, provided Ω is a regular open subset of $\mathbb{R}^d$.

Let us start with the special case $\Omega = \mathbb{R}^d$.

Proposition 2.5.7 The space of test functions $\mathscr{D}(\mathbb{R}^d)$ is dense in $H^1(\mathbb{R}^d)$. In other words, $H_0^1(\mathbb{R}^d) = H^1(\mathbb{R}^d)$.

Proof.— Starting from some $u \in H^1(\mathbb{R}^d)$, we proceed in two steps: cutoff on a compact supports, then regularisation.

1) Let $\chi \in \mathscr{D}(\mathbb{R}^d)$ a cutoff function on the unit ball. For $n \geq 1$, set

$$\chi_n(x) = \chi\left(\frac{x}{n}\right) .$$

Given $v \in L^2(\mathbb{R}^d)$, we already know that $\chi_n v$ tends to v in L^2 as $n \to \infty$. Let us take our function $u \in H^1(\mathbb{R}^d)$, and prove that $\chi_n u \to u$ in H^1 as $n \to \infty$. That is, we want to establish that $\chi_n u \to u$ in L^2, which is already known, and that for every $j = 1, \ldots, d$, $\partial_j(\chi_n u) \to \partial_j u$. From the Leibniz formula for distributions,

$$\partial_j(\chi_n u) = \chi_n \partial_j u + (\partial_j \chi_n) u ,$$

and we know that $\chi_n \partial_j u \stackrel{L^2}{\to} \partial_j u$, so we just need to prove that $(\partial_j \chi_n) u \to 0$ in L^2, This follows from the fact that $u \in L^2$ and the bound $\|\partial_j \chi_n\|_\infty = \mathcal{O}(n^{-1})$. We have thus approximated u in H^1 by compactly supported functions in H^1.

2) We now want to prove that, if $u \in H^1(\mathbb{R}^d)$ is compactly supported, then u can be approximated in H^1 by a sequence of test functions. We will proceed by regularizing u using convolution by a smooth

kernel. Let $\rho \in \mathscr{D}(\mathbb{R}^d)$, supported in the unit ball B, and such that $\int_{\mathbb{R}^d} \rho(z)\,dz = 1$. Consider, for $\varepsilon \in]0,1]$, the convolution

$$u_\varepsilon(x) := \rho_\varepsilon * u(x) := \int_{\mathbb{R}^d} \frac{1}{\varepsilon^d} \rho\left(\frac{x-y}{\varepsilon}\right) u(y)\,dy\ .$$

Then $u_\varepsilon \in \mathscr{D}(\mathbb{R}^d)$, with supp$(u_\varepsilon) \subset$ supp$(u) + \varepsilon B$. We have shown in TD1 that $u_\varepsilon \stackrel{L^2}{\to} u$ as $\varepsilon \to 0$. Furthermore, for $j = 1, \ldots, d$,

$$\partial_j u_\varepsilon(x) = \int_{\mathbb{R}^d} \partial_j \rho_\varepsilon(x-y)\, u(y)\,dy = -\int_{\mathbb{R}^d} \partial_{y_j}[\rho_\varepsilon(x-y)]\, u(y)\,dy = \rho_\varepsilon * \partial_j u(x)\ .$$

Therefore $\partial_j u_\varepsilon \stackrel{L^2}{\to} \partial_j u$ as $\varepsilon \to 0$. Summing up, $u_\varepsilon \to u$ in H^1 as $\varepsilon \to 0$.

To summarize, starting from $u \in H^1(\mathbb{R}^d)$, we can approach u by $u\chi_n$, and approach the latter by $(u\chi_n) * \rho_\varepsilon$. By the triangle inequality, the latter function approaches u. □

Notice that the second step of the above proof implies the following useful result.

Proposition 2.5.8 Every compactly supported element u of $H^1(\Omega)$ is the limit of a sequence in $\mathscr{D}(\Omega)$ supported in an arbitrarily small neighbourhood of supp(u).

We now come to the main result of this subsection, which concerns the special case of a bounded regular open subset of $\mathbb{R}^d$.

Proposition 2.5.9 Let Ω be a bounded regular open subset of $\mathbb{R}^d$ with a $\mathcal{C}^\infty$ boundary. Then $\mathcal{C}^\infty(\overline{\Omega})$ is dense in $H^1(\Omega)$.

One difficulty in the above statement comes from the definition of the space $\mathcal{C}^\infty(\overline{\Omega})$ (see Definition 2.4.20): to construct an approximation u_ε of $u \in H^1(\Omega)$ lying in $\mathcal{C}^\infty(\overline{\Omega})$, it is not sufficient to construct u_ε inside Ω (e.g. by a smoothing of u), but also show that it can be extended beyond the boundary of Ω, into a smooth function on all $\mathbb{R}^d$. This capacity to extend u_ε outside Ω will rely on the regularity of the boundary $\partial\Omega$.

Proof.— Again, we will make our construction locally, using an open cover of Ω. Given $a \in \partial\Omega$, there exists an open neighbourhood V_a of a, such that $\partial\Omega \cap V_a$ can be represented as the graph of a smooth function, expressed in some variables. Since the V_a cover the compact subset $\partial\Omega$ of $\mathbb{R}^d$, we may extract a finite covering

$$\partial\Omega \subset \bigcup_{k=1}^N V_k$$

and we consider a partition of unity $\chi_0, \ldots, \chi_N$ associated to the covering

$$\overline{\Omega} \subset \Omega \cup \bigcup_{k=1}^N V_k\ ,$$

so that $\chi_0 \in \mathscr{D}(\Omega)$, $\chi_k \in \mathscr{D}(V_k$ for $k = 1, \ldots, N$, and

$$\chi_0 + \chi_1 + \cdots + \chi_N = 1$$

on $\overline{\Omega}$. In particular, on Ω, we may decompose

$$u = \chi_0 u + \sum_{k=1}^{N} \chi_k u \ .$$

We may apply Proposition 2.5.8 to $\chi_0 u$, which is compactly supported inside Ω.

Let us now turn to one of the functions $\chi_k u$ ($k = 1, \ldots, N$), and show that it can be approximated in $H^1(\Omega)$ by a sequence of $\mathcal{C}^\infty(\overline{\Omega})$.

For simplicity we drop the index k, and assume that $\partial\Omega \cap V$ can be represented by the graph of a smooth function $q : W \to I$, e.g. in the coordinates $y = (x_2, \ldots, x_N) \mapsto x_1 = q(y)$. The set $V = I \times W$ is open in $\mathbb{R}^d = \mathbb{R}_{x_1} \times \mathbb{R}^{d-1}_y$. We assume that $\Omega \cap V$ is "above" the hypersurface $\partial\Omega \cap V$

$$\Omega \cap V = \big\{(x_1, y) \in I \times W, x_1 > q(y) \ \big\} \ ,$$

and we recall tha $\chi \in \mathscr{D}(V)$. We use the change of variables $(x_1, y) \mapsto (z = x_1 - q(y), y)$, which represents a smooth vertical shear. This shear has the effect to map $\partial\Omega \cap V$ to the horizontal hyperplane $\{0\}_z \times W_y$; we have rectified the boundary to make it flat.

Applying this change of variables, our function u is mapped to the function

$$v(z, y) := (\chi u)(z + q(y), y) \ , \quad (z, y) \in]0, \infty[\times W \ .$$

Since $\chi \in \mathscr{D}(V)$, there is a compact subset $K \Subset W$ and some $a > 0$ such that

$$\text{supp}(v) \subset]0, a] \times K \ .$$

Once we have performed this "rectification" of $\partial\Omega \cap V$, it will prove easier to construct a smooh approximation to the function v. We first need to check the effect of this change of coordinates on the regularity of u.

Lemma 2.5.10 The function v belongs to $H^1(]0, \infty[\times W)$, and its generalized derivatives are given by:

$$\partial_z v(z, y) = \partial_{x_1}\big(\chi u)(z + q(y), y\big) \ , \quad \partial_{y_j} v(z, y) = \partial_{y_j}\big(\chi u)(z + q(y), y\big) + \partial_{y_j} q(y) \partial_{x_1}\big(\chi u)(z + q(y), y\big) \ .$$

Notice that those formulas are the "naive" ones, satisfied if u is differentiable.

Let us prove the Lemma. First of all, $v \in L^2(]0, \infty[\times W)$, since, by the change of variables formula,

$$\int_W \int_0^\infty |v(z, y)|^2 \, dz \, dy = \int_W \int_{q(y)}^\infty |(\chi u)(x_1, y)|^2 \, dx_1 \, dy \ < +\infty \ .$$

(in other words, the change of variables preserves the Lebesgue measure).

We now want to compute the generalized derivative $\partial_z v$, and show that it is L^2. Let $\varphi \in \mathscr{D}(]0,\infty[_z \times W_y)$. The bracket $\langle \partial_z v, \varphi\rangle$ gives the integral

$$\begin{aligned}
-\int_W \int_0^\infty v(z,y)\partial_z\varphi(z,y)\,dz\,dy \quad &\overset{z\to x_1}{=} \quad -\int_W \int_{q(y)}^\infty (\chi u)(x_1,y)\partial_z\varphi\big(x_1-q(y),y\big)\,dx_1\,dy \\
&= \quad -\int_\Omega (\chi u)(x_1,y)\frac{\partial}{\partial x_1}\Big[\varphi\big(x_1-q(y),y\big)\Big]\,dx_1\,dy \\
&\overset{IBP}{=} \quad \int_\Omega \frac{\partial}{\partial x_1}\big[(\chi u)(x_1,y)\big]\,\varphi\big(x_1-q(y),y\big)\,dx_1\,dy \\
&\overset{x_1\to z}{=} \quad \int_W \int_0^\infty \partial_{x_1}(\chi u)(z+q(y),y)\,\varphi(z,y)\,dz\,dy\ .
\end{aligned}$$

The integration by parts makes sense because $\chi u \in H^1(\Omega)$, so that $\partial_{x_1}(\chi u) \in L^2$. The last integral gives the expression of $\partial_z v$ stated in the Lemma.

Similarly, let us compute $\langle \partial_{y_j} v, \varphi\rangle = -\langle v, \partial_{y_j}\varphi\rangle$ for $\varphi \in \mathscr{D}(]0, infty[\times W)$:

$$\begin{aligned}
-\int_W \int_0^\infty v(z,y)\partial_{y_j}\varphi(z,y)\,dz\,dy \quad &\overset{z\to x_1}{=} \quad -\int_W \int_{q(y)}^\infty (\chi u)(x_1,y)\,\partial_{y_j}\varphi\big(x_1-q(y),y\big)\,dx_1\,dy \\
&= \quad -\int_\Omega (\chi u)(x_1,y)\,\frac{\partial}{\partial y_j}\Big[\varphi\big(x_1-q(y),y\big)\Big]\,dx_1\,dy \\
&\quad\quad -\int_\Omega (\chi u)(x_1,y)\,\partial_{y_j}q(y)\,\frac{\partial}{\partial x_1}\big[\varphi\big(x_1-q(y),y\big)\big]\,dx_1\,dy \\
&= \quad \int_\Omega \big[\partial_{y_j} + \partial_{y_j}q(y)\partial_{x_1}\big](\chi u)(x_1,y)\,\varphi\big(x_1-q(y),y\big)\,dx_1\,dy \\
&\overset{x_1\to z}{=} \quad \int_W \int_0^\infty [\partial_{y_j} + \partial_{y_j}q(y)\partial_{x_1}](\chi u)\big(z+q(y),y\big)\,\varphi(z,y)\,dz\,dy\ .
\end{aligned}$$

(in the third equality we performed integration by parts on x_1 and on y_j). The last integral provides the expression for $\partial_{y_j} v$, in terms of the derivatives of (χu), which are in $L^2(\Omega\cap V)$. Applying the shear to these derivatives preserves their L^2 character, hence the function

$$\begin{aligned}
&(z,y) \in]0,\infty[\times W \mapsto \partial_{x_1}(\chi u)(z+q(y),y) \\
&(z,y) \in]0,\infty[\times W \mapsto [\partial_{y_j} + \partial_{y_j}q(y)\partial_{x_1}](\chi u)(z+q(y),y)
\end{aligned}$$

belong to $L^2(]0,\infty[\times W)$. This completes the proof of Lemma 2.5.10. □

Once we understand the function $v \in H^1(\mathbb{R}_+^* \times W)$, we extend it to $\mathbb{R}\times W$ by symmetry w.r.t. the hyperplane $\{z=0\}$.

Lemma 2.5.11 Given $v \in H^1(\mathbb{R}_+^*\times W)$, we define $\tilde v \in L^2(\mathbb{R}\times W)$ by

$$\tilde v(z,y) = \begin{cases} v(z,y) & \text{if } z>0 \\ v(-z,y) & \text{if } z<0\,. \end{cases}$$

Then $\tilde v \in H^1(\mathbb{R}\times W)$.

Notice that, even if the function v were in $H^1 \cap C^1(\mathbb{R}_+^* \cap W)$, its derivative $\partial_z v(z,y)$ has no reason to go to zero when $z \searrow 0$, so the derivative $\partial_z \tilde{v}(z,y)$ generally has a jump at $z=0$.

Let us prove the Lemma. First we observe that

$$\|\tilde{v}\|^2_{L^2(\mathbb{R}\times W)} = 2\|\tilde{v}\|^2_{L^2(]0,\infty[\times W)} < \infty$$

so $\tilde{v} \in L^2(\mathbb{R}\times W)$. Like in the previous Lemma, let us compute the generalized derivatives of $\tilde{v}$. We now take test function $\varphi \in \mathscr{D}(\mathbb{R}\times W)$, which “see” the plane $\{z=0\}$. Let us calculate $\langle \partial_{y_j}\tilde{v}, \varphi\rangle$:

$$\begin{aligned}
-\int_{\mathbb{R}}\int_W \tilde{v}(z,y)\partial_{y_j}\varphi(z,y)\,dz\,dy &= -\int_0^{+\infty}\int_W v(z,y)\partial_{y_j}\varphi(z,y)\,dz\,dy - \int_{-\infty}^0\int_W v(-z,y)\partial_{y_j}\varphi(z,y)\,dz\,dy \\
&= -\int_0^{\infty}\int_W v(z,y)\big[\partial_{y_j}\varphi(z,y) + \partial_{y_j}\varphi(-z,y)\big]\,dz\,dy\,.
\end{aligned}$$

In order to use the information $v \in H^1(\mathbb{R}_+^*\times W)$, we need to reduce to test functions compactly supported in $\mathbb{R}_+^*\times W$. We introduce $\chi \in \mathcal{C}^\infty(\mathbb{R})$ such that $\chi(t)=0$ for $t\le \frac{1}{2}$ and $\chi(t)=1$ for $t\ge 1$, and we set, for $\varepsilon>0$,

$$\chi_\varepsilon(z) := \chi\left(\frac{z}{\varepsilon}\right)\ .$$

Then $\chi_\varepsilon(z) \to \mathbf{1}_{\mathbb{R}_+^*}(z)$ as $\varepsilon\to 0$, and by dominated convergence, we have

$$\begin{aligned}
-\int_{\mathbb{R}}\int_W \tilde{v}(z,y)\partial_{y_j}\varphi(z,y)\,dz\,dy &= -\lim_{\varepsilon\to 0}\int_0^{\infty}\int_W v(z,y)\,\chi_\varepsilon(z)\,\partial_{y_j}\big[\varphi(z,y)+\varphi(-z,y)\big]\,dz\,dy \\
&\overset{IBP}{=} \lim_{\varepsilon\to 0}\int_0^{\infty}\int_W \partial_{y_j}v(z,y)\,\chi_\varepsilon(z)\big[\varphi(z,y)+\varphi(-z,y)\big]\,dz\,dy \\
&\overset{\partial_{y_j}v\in L^2}{=} \int_0^{\infty}\int_W \partial_{y_j}v(z,y)\,\big[\varphi(z,y)+\varphi(-z,y)\big]\,dz\,dy \\
&:= \int_{\mathbb{R}}\int_W f_j(z,y)\,\varphi(z,y)\,dz\,dy\ ,
\end{aligned}$$

where

$$f_j(z,y) = \begin{cases}\partial_{y_j}v(z,y) & \text{if } z>0\\ \partial_{y_j}v(-z,y) & \text{if } z<0\end{cases}$$

defines an L^2 function on $\mathbb{R}\times W$.

We similarly compute $\langle \partial_z\tilde{v},\varphi\rangle$:

$$\begin{aligned}
-\int_{\mathbb{R}}\int_W \tilde{v}(z,y)\,\partial_z\varphi(z,y)\,dz\,dy &= -\int_0^{+\infty}\int_W v(z,y)\,\partial_z\varphi(z,y)\,dz\,dy - \int_{-\infty}^0\int_W v(-z,y)\,\partial_z\varphi(z,y)\,dz\,dy \\
&= -\int_0^{\infty}\int_W v(z,y)\partial_z\big[\varphi(z,y)-\varphi(-z,y)\big]\,dz\,dy \\
&= -\lim_{\varepsilon\to 0}\int_0^{\infty}\int_W v(z,y)\,\chi_\varepsilon(z)\,\partial_z\big[\varphi(z,y)-\varphi(-z,y)\big]\,dz\,dy \\
&= -\lim_{\varepsilon\to 0}\int_0^{\infty}\int_W v(z,y)\partial_z\big[\chi_\varepsilon(z)\big(\varphi(z,y)-\varphi(-z,y)\big)\big]\,dz\,dy \\
&\quad + \lim_{\varepsilon\to 0}\int_0^{\infty}\int_W v(z,y)\,\chi_\varepsilon'(z)\,\big[\varphi(z,y)-\varphi(-z,y)\big]\,dz\,dy
\end{aligned}$$

At this stage, we may integrate by parts the first integral. We will estimate the second one, with the following remarks:

i) the function χ'_ε is supported in the interval $[\varepsilon/2, \varepsilon]$, where it takes values $\mathcal{O}(\varepsilon^{-1})$.

ii) the function $\varphi(z,y) - \varphi(-z,y) = \mathcal{O}(z)$, so in the interval $[\varepsilon/2, \varepsilon]$ its values are $\mathcal{O}(\varepsilon)$, uniformly w.r.t. $y \in \pi\, \mathrm{supp}\, \varphi$.

iii) the function $v \in L^2$, so by Cauchy-Schwartz $\int_{\varepsilon/2}^{\varepsilon} |v| dz\, dy = \mathcal{O}(\varepsilon^{1/2})$.

We then deduce that the second integral vanishes when $\varepsilon \to 0$. There remains:

$$\begin{aligned}
\langle \partial_z \tilde{v}, \varphi \rangle &= \lim_{\varepsilon\to 0} \int_0^\infty \int_W \partial_z v(z,y)\, \chi_\varepsilon(z) \left[\varphi(z,y) - \varphi(-z,y)\right] dz\, dy \\
&= \int_0^\infty \int_W \partial_z v(z,y) \left[\varphi(z,y) - \varphi(-z,y)\right] dz\, dy \\
&= \int_{\mathbb{R}} \int_W g(z,y)\, \varphi(z,y)\, dz\, dy\ ,
\end{aligned}$$

where

$$g(z,y) = \begin{cases} \partial_z v(z,y) & \text{if } z > 0 \\ -\partial_z v(-z,y) & \text{if } z < 0 \end{cases}$$

defines an L^2 function on $\mathbb{R} \times W$. We notice that the formula for g is the "naive" formula for $\partial_z \tilde{v}$, without paying attention to the point $(z = 0, y)$ where

We have thus computed

$$\partial_{y_j} \tilde{v} = f_j\ , \quad \partial_z \tilde{v} = g,$$

given by "naive" expressions for $\partial_\bullet \tilde{v}$, without paying attention to the points $(z = 0, y)$ (notice that $\tilde{v}$ is generally not differentiable w.r.t. z at the points $z = 0$). These derivatives are L^2, which completes he proof that $\tilde{v} \in H^1(\mathbb{R} \times W)$. □

Let us now finish the proof of Proposition 2.5.9. Applying Lemma 2.5.11 to our "rectified function" $v(z,y) = (\chi u)(z + q(y), y)$, we observe that $\tilde{v} \in H^1(\mathbb{R} \times W)$ and $\mathrm{supp}(\tilde{v}) \subset [-a, a] \times K$, a compact subset of $\mathbb{R} \times W$. Now we appeal to Proposition 2.5.8 and obtain that there exist a family of test functions $(\tilde{v}_\varepsilon \in \mathscr{D}(\mathbb{R} \times W))_{\varepsilon \in]0,1]}$ such that $\mathrm{supp}(\tilde{v}_\varepsilon) \subset [-(a+\varepsilon, a+\varepsilon] \times K_\varepsilon$ and $\tilde{v}_\varepsilon\ \varepsilon \overrightarrow{\to}\ 0\ \tilde{v}$ in $H^1(\mathbb{R} \times W)$. Then the restriction v_ε of $\tilde{v}_\varepsilon$ to $\mathbb{R}_+^* \times W$ converges to v in $H^1(\mathbb{R}_+^* \times W)$, and coming back to the initial variables (x_1, y),

$$u_\varepsilon(x_1, y) := v_\varepsilon\big(x_1 - q(y), y\big)$$

is a family of functions in $\mathcal{C}^\infty(\overline{\Omega})$ which, in view of the formulae for derivatives established in Lemma 2.5.10, converges to χu in $H^1(\Omega)$ as $\varepsilon \to 0$. □

2.5.4 The trace theorem

In dimension 1, every function $u \in H^1(I)$ is automatically continuous and bounded, which allowed us to characterize the subspace $H_0^1(I)$ in terms of the limits of $u(x)$ when x approaches the boundary of I. In higher dimension, the Sobolev embedding theorem shows that $u \in H^1(\Omega)$ is not necessarily continuous, and not necessarily uniquely defined at each point. It is therefore not possible to characterize the property $u \in H_0^1(\Omega)$ in terms of the pointwise limits of $u(x)$ when x approaches $\partial\Omega$.

Yet, the following theorem will provide a way to characterize functions $u \in H_0^1(\Omega)$ in terms of a certain function $v = \gamma_0 u$ on $\partial\Omega$, called the *trace* of u on the boundary, which can be viewed as a weak form of "limit of u on $\partial\Omega$".

Theorem 2.5.12 Let Ω be a bounded regular open subset of $\mathbb{R}^d$, with a $\mathcal{C}^\infty$ boundary. Denote by σ the superficial measure on $\partial\Omega$ and by N^{ext} the exterior unit normal on $\partial\Omega$.

There exists a unique linear mapping

$$\gamma_0 : H^1(\Omega) \to L^2(\partial\Omega, \sigma)$$

such that, for every $\varphi \in \mathcal{C}^\infty(\overline{\Omega})$,

$$\gamma_0\varphi = \varphi_{|\partial\Omega} .$$

Furthermore, γ_0 satisfies the following identity, for all $j = 1, \dots, d$:

$$\forall u, v \in H^1(\Omega) \ , \ \int_\Omega u\, \partial_j v\, dx = \int_{\partial\Omega} \gamma_0 u\, \gamma_0 v\, N_j^{ext}\, d\sigma - \int_\Omega \partial_j u\, v\, dx \ . \tag{2.5.11}$$

This formula generalizes the Green-Gauss formula of Remark 2.4.22 to functions $u, v \in H^1(\Omega)$.

Finally, Ker γ_0 is equal to $H_0^1(\Omega)$, the closure of $\mathscr{D}(\Omega)$ in $H^1(\Omega)$.

Proof.— From Remark 2.4.22 we have, for every $\varphi, \psi \in \mathcal{C}^\infty(\overline{\Omega})$ and $j \in \{1, \dots, d\}$:

$$\int_\Omega \varphi\, \partial_j\psi\, dx = \int_{\partial\Omega} \varphi\, \psi\, N_j^{ext}\, d\sigma - \int_\Omega \partial_j\varphi\, \psi\, dx \ . \tag{2.5.12}$$

Consider the mapping

$$\gamma_0 : \varphi \in \mathcal{C}^\infty(\overline{\Omega}) \longmapsto \varphi_{|\partial\Omega} \in L^2(\partial\Omega, \sigma) \ .$$

We claim that this mapping is continuous if $\mathcal{C}^\infty(\overline{\Omega})$ is endowed with the H^1 norm, which amounts to proving the estimate

$$\|\varphi\|_{L^2(\partial\Omega,\sigma)} \le C\|\varphi\|_{H^1(\Omega)} \ . \tag{2.5.13}$$

Using (again) a partition of unity associated to a finite open cover of the compact set $\partial\Omega$, it is enough to prove the above inequality when φ is supported in the intersection of Ω with a small neighbourhood

V_a of some point $a \in \partial\Omega$. We may assume that, on such a neighbourhood V_a, a certain component $j = j_a$ of the normal vector satisfies

$$\forall x \in V_a \cap \partial\Omega, \qquad |N_j^{ext}(x)| \geq c,$$

for some $c = c_a > 0$. Then, considering the formula (2.5.12) with $\psi = \overline{\varphi}$ and the index j_a, and using the Cauchy–Schwarz inequality, we find that the cutoff function $\chi_a\varphi$ satisfies (2.5.13). Summing over the finitely many pieces $\chi_a\varphi$ composing φ, we obtain (2.5.13) for the full function φ.

At this stage, we appeal to Proposition 2.5.9, namely the density of $\mathcal{C}^\infty(\overline{\Omega})$ in $H^1(\Omega)$. Since γ_0 is a continuous linear mapping from a dense subspace of $\mathcal{C}^\infty(\overline{\Omega}) \subset H^1(\Omega)$ into the Hilbert space $L^2(\partial\Omega, \sigma)$, it admits a unique linear continuous extension from $H^1(\Omega)$ to $L^2(\partial\Omega, \sigma)$, which we will still denote by γ_0. This proves the existence an uniqueness of γ_0.

Let us come to the second point. Since both sides of Green's formula (2.5.11) are continuous bilinear maps on $H^1(\Omega) \times H^1(\Omega)$, and coincide on the dense subspace $\mathcal{C}^\infty(\overline{\Omega}) \times \mathcal{C}^\infty(\overline{\Omega})$ in view of (2.5.12), we infer that this equality holds as well on the closure $H^1(\Omega) \times H^1(\Omega)$.

Finally, let us characterise the kernel of γ_0. If $\varphi \in \mathscr{D}(\Omega)$, we have $\gamma_0\varphi = \varphi_{|\partial\Omega} = 0$, therefore $\operatorname{Ker}\gamma_0$ contains the closure $H_0^1(\Omega)$ of $\mathscr{D}(\Omega)$ in $H^1(\Omega)$.

The proof of the converse inclusion is more intricate. Let us start by a notation: for every $f \in L^2(\Omega)$, we denote by $\underline{f} \in L^2(\mathbb{R}^d)$ the extension of f by 0 to $\mathbb{R}^d \setminus \Omega$. Let us then take some $u \in \operatorname{Ker}\gamma_0$; our goal is to show that $u \in H_0^1(\Omega)$. We will consider the extension $\underline{u}$, and compute its ∂_j derivative in $\mathscr{D}'(\mathbb{R}^d)$. For this, consider some test function $\varphi \in \mathscr{D}(\mathbb{R}^d)$, and compute:

$$\langle \partial_j \underline{u}, \varphi\rangle = -\langle \underline{u}, \partial_j\varphi\rangle = -\int_\Omega u\,\partial_j\varphi\,dx\,.$$

We may the apply the Green identity (2.5.11), with u and $v := \varphi_{|\Omega}$:

$$\begin{aligned}
-\langle \partial_j \underline{u}, \varphi\rangle &= \int_\Omega \int_\Omega \partial_j u\,\varphi\,dx - \int_{\partial\Omega} \gamma_u\,\gamma_0\varphi\,d\sigma \\
&= \int_\Omega \int_\Omega \partial_j u\,\varphi\,dx\,,
\end{aligned}$$

where we have used the assumption $u \in \operatorname{Ker}\gamma_0$. This equality amounts to the identification of the following distributions in $\mathscr{D}'(\mathbb{R}^d)$:

$$\partial_j(\underline{u}) = \underline{\partial_j u}\,.$$

In particular, it shows that $\partial_j(\underline{u}) \in L^2(\mathbb{R}^d)$. Since this holds for for all $j = 1, \ldots, d$, we have proved that

$$u \in \operatorname{Ker}\gamma_0 \Longrightarrow \underline{u} \in H^1(\mathbb{R}^d).$$

In other words, the behaviour of u when approaching $\partial\Omega$ is sufficiently regular, so that taking the gradient of $\underline{u}$ does not produce any singularity on $\partial\Omega$.

Let us use the same partition of unity (χ_k) as in the proof of Proposition 2.5.9, and the same notations of that proof. We are reduced to proving that, for every $k = 1, \ldots, N$, $\chi_k u \in H_0^1(\Omega)$. Notice that

$\underline{\chi_k u} = \chi_k \underline{u} \in H^1(\mathbb{R}^d)$. Let us drop the index k and apply the change of variables $(x_1, y) \mapsto (z, y)$ to straighten $\partial\Omega \cap V$; we obtain the function

$$v(z,y) = (\chi u)\big(z + q(y), y\big)\ , \quad (z,y) \in \mathbb{R}_+^* \times W\ .$$

This function extends to $\mathbb{R}^d$:

$$\underline{v}(z,y) = \underline{\chi u}\big(z + q(y), y\big), \quad (z,y) \in \mathbb{R} \times \mathbb{R}^{d-1}\,.$$

Since $\underline{u} \in H^1(\mathbb{R}^d)$, we infer, by the same argument as in Lemma 2.5.10, that $\underline{v} \in H^1(\mathbb{R}^d)$.

The final arguments consists in shifting this function v upwards, to make it compactly supported in $\mathbb{R}_+^* \times W$. Namely, for $\varepsilon > 0$ small, we define the translated function

$$v_\varepsilon(z,y) := \underline{v}(z - \varepsilon, y) \quad (z,y) \in \mathbb{R}^d\,.$$

Of course $v_\varepsilon \in H^1(\mathbb{R}^d)$ and supp$(v_\varepsilon) \subset [\varepsilon, a+\varepsilon] \times \bar{W}$, a compact subset of $\mathbb{R}^d$. Furthermore, the continuous action of translations on L^2 implies that

$$\|v_\varepsilon - v\|_{H^1}^2 = \int_{\mathbb{R}_+} \int_W \Big(|v(z-\varepsilon, y) - v(z,y)|^2 + \sum_j |\partial_j v(z-\varepsilon, y) - \partial_j v(z,y)|^2 \Big)\, dz\, dy$$

converges to 0 when $\varepsilon \searrow 0$. We conclude that $v_\varepsilon \to v$ in $H^1(\mathbb{R}_+^* \times W)$, and, again by Lemma 2.5.10, that the family $(u_\varepsilon)_{\varepsilon \in]0,1]}$ defined by

$$u_\varepsilon(x_1, y) = v_\varepsilon\big(x_1 - q(y), y\big) = \chi \underline{u}(x_1 - \varepsilon, y)\,,$$

is supported in $\Omega \cap V$ for ε small enough, and converges to χu in $H^1(\Omega)$.

Finally, since, for every $\varepsilon > 0$, u_ε is supported in a compact subset of $\Omega \cap V$, Proposition 2.5.8 implies that $v_\varepsilon \in H_0^1(\Omega)$. Consequently, $\chi u \in H_0^1(\Omega)$. Summing up over all the indices $k = 1, \ldots, N$, we find that $u \in H_0^1(\Omega)$. □

Reformulating the Poisson equation

As an application of the trace theorem, we may express the following Dirichlet problem on a bounded regular subset Ω as follows.

Theorem 2.5.13 Let Ω be a bounded regular open subset of $\mathbb{R}^d$ with a $\mathcal{C}^\infty$ boundary. For every $f \in L^2(\Omega)$, there exists a unique $u \in H^1(\Omega)$ such that

$$\begin{cases} -\Delta u = f & \text{in } \Omega\ , \\ \gamma_0 u = 0 & \text{in } \partial\Omega\ . \end{cases}$$

Proof.— This is an immediate consequence of Remark 2.5.5 3. and of the trace Theorem 2.5.12. □

Remark 2.5.14 One can show that, under the assumptions of Theorem 2.5.13, the solution $u \in H^2(\Omega)$, so that the mapping
$$\Delta : H^2(\Omega) \cap H^1_0(\Omega) \to L^2(\Omega)$$
is an isomorphism if Ω is a regular bounded open subset. Furthermore, more regularity on f implies more regularity on u. For instance, if $f \in H^m(\Omega)$ for some $m \in \mathbb{N}$, then $u \in H^{m+2}(\Omega)$. As a result of Sobolev injection theorems on Ω, if $f \in \mathcal{C}^\infty(\overline{\Omega})$, then $u \in \mathcal{C}^\infty(\overline{\Omega})$. The latter statement is a consequence of the previous one and of

$$\bigcap_{m \in \mathbb{N}} H^m(\Omega) = \mathcal{C}^\infty(\overline{\Omega}) \ .$$

The proofs of these statements go beyond these lecture notes.

2.6 The uniform boundedness principle

In this section, we prove an important result about families of distributions, which implies Lemma 2.3.25. It is a generalization to the space of distributions of the uniform boundedness for Banach spaces, proved by Banach-Steinhaus.

Theorem 2.6.1 [Uniform boundedness for distributions] Let (T_n) be a family if distributions on Ω such that, for every $\varphi \in \mathscr{D}(\Omega)$,

$$\sup_n |\langle T_n, \varphi\rangle| < +\infty \ .$$

Then, for every compact subset K of Ω, there exist $C > 0$ and $m \in \mathbb{N}$ such that, for every $\varphi \in \mathscr{D}_K(\Omega)$,
$$\sup_n |\langle T_n, \varphi\rangle| \le C\|\varphi\|_{C^m} \ .$$

In other words, if for instance a sequence T_n is such that $\langle T_n, \varphi\rangle$ has a limit for every $\varphi \in \mathscr{D}(\Omega)$, one gets a uniform estimate on the action of the sequence T_n.

Before proving this theorem, let us show how it implies the Lemma 2.3.25. If $T_n \to T$, then the assumption of Theorem 2.6.1 is fulfilled. If $\varphi_n \to \varphi$ in $\mathscr{D}(\Omega)$, let K be a compact subset which contains the support of φ_n for every n, and hence the support of φ. By Theorem 2.6.1, we have, for some C, m independent of n,

$$|\langle T_n, \varphi_n - \varphi\rangle| \le C\|\varphi_n - \varphi\|_{C^m} \to 0 \ .$$

Then
$$\langle T_n, \varphi_n\rangle = \langle T_n, \varphi_n - \varphi\rangle + \langle T_n, \varphi\rangle \to \langle T, \varphi\rangle \ .$$

This completes the proof of Lemma 2.3.25. □

In the rest of this section, we give a proof of Theorem 2.6.1, in three steps. For every compact subset K of Ω, we denote by $\mathscr{D}_K = \mathscr{D}_K(\Omega)$ the subspace of $\mathscr{D}(\Omega)$ of test functions φ such supp$(\varphi) \subset K$.

2.6.1 Step1. Realizing $\mathscr{D}_K$ as a complete metric space

In this section we fix a compact subset $K \Subset \Omega$, and consider he space of test functions $\mathscr{D}_K$. We show that the topology on $\mathscr{D}_K$ can be metrized.

Lemma 2.6.2 [Metrization of $\mathscr{D}_K(\Omega)$] There exists a distance function d on $\mathscr{D}_K$ having the following properties:

i) if a sequence $(\varphi_n)_n$ and φ are in $\mathscr{D}_K$ then $\varphi_n \; n \overset{\rightarrow}{\rightarrow} \infty \; \varphi$ in $\mathscr{D}$ if and only if $d(\varphi_n, \varphi) \to 0$;

ii) the metric space $(\mathscr{D}_K, d)$ is complete;

iii) for every $\varphi, \psi, \theta \in \mathscr{D}_K$, $d(\varphi + \theta, \psi + \theta) = d(\varphi, \psi)$;

iv) for every $\varepsilon > 0$, there exist $m \in \mathbb{N}$ and $r > 0$ such that the "cylinder" $\{\varphi \in \mathscr{D}_K \,:\, \|\varphi\|_{C^m} \leq r\}$ is contained in the closed ball $\{\phi \in \mathscr{D}_K \,:\, d(\varphi, 0) \leq \varepsilon\}$.

Proof.— For $\varphi, \psi \in \mathscr{D}_K$, we define the function

$$d(\varphi, \psi) := \sum_{m=0}^{\infty} \min\left(2^{-m}, \|\varphi - \psi\|_{C^m}\right) \; .$$

It is clear that d takes values in $[0, \infty[$, and that $d(\varphi, \psi) = 0$ iff $\varphi = \psi$. The symmetry $d(\varphi, \psi) = d(\psi, \varphi)$ is trivial, as well as item $iii)$ of the Lemma. Finally, the triangle inequality is a consequence of the elementary inequality

$$\min(a, x + y) \leq \min(a, x) + \min(a, y) \; , \quad a \geq 0, x \geq 0, y \geq 0 \; .$$

Let us prove item $i)$. If a sequence φ_n and φ are in $\in \mathscr{D}_K$, the statement $\varphi_n \to \varphi$ in $\mathscr{D}$ is equivalent to

$$\forall m \in \mathbb{N}, \|\varphi_n - \varphi\|_{C^m} \to 0 \; . \tag{2.6.14}$$

Since $d(\varphi_n, \varphi) \geq \min\left(2^{-m}, \|\varphi_n - \varphi\|_{C^m}\right)$, it is clear that (2.6.14) is implied by $d(\varphi_n, \varphi) \to 0$. Conversely, (2.6.14) implies that every term of the series defining $d(\varphi_n, \varphi)$ tends to 0. Since this series is normally convergent, this implies $d(\varphi_n, \varphi) \to 0$.

Let us prove item $ii)$. Let (φ_n) be a Cauchy sequence in the metric space $(\mathscr{D}_K, d)$. This can be expressed as follows:

$$\forall \varepsilon > 0, \quad \exists N(\varepsilon) \in \mathbb{N}, \quad \forall n, p \geq N(\varepsilon), \qquad d(\varphi_n, \varphi_p) \leq \varepsilon \; .$$

Let $m \in \mathbb{N}$ and $\varepsilon > 0$ be such that $\varepsilon < 2^{-m}$. Since $d(\varphi_n, \varphi_p) \geq \min\left(2^{-m}, \|\varphi_n - \varphi_p\|_{C^m}\right)$, we conclude that, for $n, p \geq N(\varepsilon)$, $\|\varphi_n - \varphi_p\|_{C^m} \leq \varepsilon$. This means that (φ_n) is a Cauchy sequence in the Banach space $\mathcal{C}^m_K$ of $\mathcal{C}^m$ functions supported in K. Hence there exists $\varphi^{[m]} \in \mathcal{C}^m_K$ such that $\|\varphi_n - \varphi^{[m]}\|_{C^m} \to 0$. Since this is true for every $m \in \mathbb{N}$, and since the convergence in $\mathcal{C}^{m+1}_K$ implies the convergence in $\mathcal{C}^m_K$, we conclude that

$$\varphi^{[m]} = \varphi \in \bigcap_{m\in\mathbb{N}} \mathcal{C}^m_K = \mathscr{D}_K \ ,$$

and that $\varphi_n \to \varphi$ in $\mathscr{D}$, which, by item $i)$, means $d(\varphi_n, \varphi) \to 0$. Hence $(\mathscr{D}_K, d)$ is a complete metric space.

Finally, let us prove item $iv)$. Let $m \in \mathbb{N}$ be such that $2^{-m} \leq \frac{\varepsilon}{2}$, and let $r > 0$ be such that $(m+1)r \leq \frac{\varepsilon}{2}$. If $\|\varphi\|_{C^m} \leq r$, we have

$$\begin{aligned} d(\varphi, 0) &= \sum_{q=0}^{\infty} \min\left(2^{-q}, \|\varphi\|_{C^q}\right) \\ &\leq \sum_{q=0}^{m} \|\varphi\|_{C^q} + \sum_{q=m+1}^{\infty} 2^{-q} \\ &\leq (m+1)r + 2^{-m} \leq \varepsilon \ . \end{aligned}$$

This completes the proof of the Lemma. □

2.6.2 Step 2. The Baire lemma

The second step in the proof of the Uniform boundedness Theorem 2.6.1 is a standard result in general topology, namely Baire's Theorem on complete metric spaces.

Theorem 2.6.3 Let $(\mathcal{E}, d)$ be a complete metric space. Then every countable intersection of open dense subsets of $\mathcal{E}$ is dense in $\mathcal{E}$. By complementarity Every countable union of closed subsets of $\mathcal{E}$ with empty interior has empty interior.

Proof.— The second statement is equivalent to the first one, since the complement of an open set is a closed set, and the complement of a dense set has an empty interior.

Let us prove the first statement. Denote by $B(\varphi, r)$ and $B_f(\varphi, r)$ respectively the open and the closed balls of radius r centered at $\varphi \in \mathcal{E}$. Let $(\mathcal{O}_k)_{k\in\mathbb{N}}$ be a sequence of dense open subsets of $\mathcal{E}$. Fix $\varphi_0 \in \mathcal{E}$ and $\varepsilon > 0$. We want to prove that

$$B(\varphi_0, \varepsilon) \cap \bigcap_{k\in\mathbb{N}} \mathcal{O}_k \neq \emptyset \ .$$

Since $\mathcal{O}_0$ is dense, $B(\varphi_0, \varepsilon) \cap \mathcal{O}_0$ contains some element φ_1. Since $B(\varphi_0, \varepsilon) \cap \mathcal{O}_0$ is open, there exists $r_1 > 0$ such that

$$\overline{B(\varphi_1, r_1)} \subset B(\varphi_0, \varepsilon) \cap \mathcal{O}_0 \ ,$$

moreover we may assume that $r_1 \leq \frac{\varepsilon}{2}$.

Since $\mathcal{O}_1$ is dense, $B(\varphi_1, r_1) \cap \mathcal{O}_1$ contains some element φ_2. Since $B(\varphi_1, r_1) \cap \mathcal{O}_1$ is open, there exists $r_2 > 0$ such that

$$\overline{B(\varphi_2, r_2)} \subset B(\varphi_1, r_1) \cap \mathcal{O}_1 \ ,$$

moreover we may assume that $r_2 \leq \frac{\varepsilon}{2^2}$.

Continuing that way, we define a sequence $(\varphi_k)_{k\geq 1}$ of elements of $\mathcal{E}$ and a sequence $(r_k)_{k\geq 1}$ of positive numbers such that

$$\forall k \geq 1, \quad \overline{B(\varphi_{k+1}, r_{k+1})} \subset B(\varphi_k, r_k) \cap \mathcal{O}_k \ , \quad r_k \leq \frac{\varepsilon}{2^k} \ .$$

In particular, the sequence of closed balls $(\overline{B(\varphi_k, r_k)})_{k\geq 1}$ is decreasing, with a radius tending to 0. Moreover,

$$\overline{B(\varphi_k, r_k)} \subset \mathcal{O}_k \cap \cdots \cap \mathcal{O}_0 \cap B(\varphi_0, \varepsilon) \ .$$

Hence we are reduced to proving that

$$\bigcap_{k=1}^{\infty} \overline{B(\varphi_k, r_k)} \neq \emptyset \ .$$

We observe that $(\varphi_k)_{k\geq 1}$ is a Cauchy sequence in $\mathcal{E}$. Indeed, if $\ell \geq 0$, $\varphi_{k+\ell} \in \overline{B(\varphi_k, r_k)}$, hence

$$d(\varphi_{k+\ell}, \varphi_k) \leq r_k \ ,$$

which tends to 0 as k tends to ∞. Since $\mathcal{E}$ is a complete metric space, φ_k has a limit $\varphi \in \mathcal{E}$. Passing to the limit as ℓ tends to infinity in the above inequality, we get, for every $k \geq 1$,

$$d(\varphi, \varphi_k) \leq r_k \Longrightarrow \varphi \in \overline{B(\varphi_k, r_k)} \, ,$$

so that

$$\varphi \in \bigcap_{k=1}^{\infty} \overline{B(\varphi_k, r_k)} \subset \left(\bigcap_{k=1}^{\infty} \mathcal{O}_k \right) \cap B(\varphi_0, \varepsilon) \, .$$

This completes the proof of Baire's theorem. □

2.6.3 Step 3. A uniform boundedness theorem on a complete metric space

Let us use the notation of Theorem 2.6.1. We set, for every $k \in \mathbb{N}$,

$$\mathcal{F}_k = \{\varphi \in \mathscr{D}_K : \forall n, |\langle T_n, \varphi \rangle| \leq k\} \ .$$

Since each T_n is continuous on $\mathscr{D}_K$, $\mathcal{F}_k$ is a closed subset of $\mathscr{D}_K$. Furthermore, the assumption of Theorem 2.6.1 precisely means that

$$\bigcup_{k\in\mathbb{N}} \mathcal{F}_k = \mathscr{D}_K \ .$$

Let endow $\mathscr{D}_K$ with the distance function of Lemma 2.6.2. By Baire's theorem, we infer that there exists k_0 such that $\mathcal{F}_{k_0}$ has a nonempty interior, which means that there exists some $\varphi_0 \in \mathscr{D}_K$ and $\varepsilon > 0$ such that

$$\overline{B(\varphi_0, \varepsilon)} \subset \mathcal{F}_{k_0} \ .$$

By item $iii)$ of Lemma 2.6.2, we know that

$$\overline{B(0, \varepsilon)} = -\varphi_0 + \overline{B(\varphi_0, \varepsilon)},$$

so that, for every $\psi \in \overline{B(0, \varepsilon)}$, for every n,

$$|\langle T_n, \psi \rangle| \leq |\langle T_n, -\varphi_0 \rangle| + k_0 \leq A \ , \quad \text{with} \quad A = k_0 + \sup_n |\langle T_n, \varphi_0 \rangle| \, .$$

Using item $iv)$ of Lemma 2.6.2, there exists $m \in \mathbb{N}$ and $r > 0$ such that every $\psi \in \mathscr{D}_K$ such that $\|\psi\|_{C^m} \leq r$ belongs to $\overline{B(0, \varepsilon)}$, hence satisfies

$$\sup_n |\langle T_n, \psi \rangle| \leq A \ .$$

Given $\varphi \in \mathscr{D}_K \setminus \{0\}$, we may apply this fact to

$$\psi := r \frac{\varphi}{\|\varphi\|_{C^m}} \ ,$$

and we conclude that

$$\sup_n |\langle T_n, \varphi \rangle| \leq \frac{A}{r} \|\varphi\|_{C^m} \ .$$

The proof is complete. □

Remark 2.6.4 The above theorem is an adaptation of the Banach–Steinhaus theorem to the case of vector spaces admitting a distance enjoying the properties of Lemma 2.6.2. Such spaces are called Fréchet spaces.

Chapter 3

The Fourier Transformation

Classically, the Fourier transform of a function $f \in L^1(\mathbb{R}^d)$ is the function $\mathcal{F}(f) \in L^\infty(\mathbb{R}^d)$ given by

$$\forall \xi \in \mathbb{R}^d, \quad \mathcal{F}(f)(\xi) = \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x)dx, \quad \text{so that} \quad \|\mathcal{F}(f)\|_{L^\infty} \leq \|f\|_{L^1}.$$

Here the variable $\xi \in \mathbb{R}^d$ lives in the Fourier space, or reciprocal space, or (in quantum mechanics) momentum space; the corresponding space $\mathbb{R}^d$ is dual to the "direct space" $\mathbb{R}^d \ni x$.

The Fourier transform was invented in order to help solving certain linear PDEs from physics (originally the heat equation). This role is mainly due to the following algebraic property of the Fourier transform: when these objects are well-defined,

$$\boxed{\forall \xi \in \mathbb{R}^d, \qquad \mathcal{F}(\partial_j f)(\xi) = i\xi_j \mathcal{F}(f)(\xi)}$$

In other words, $\mathcal{F}$ transforms a differential operator with constant coefficients into the product by a polynomial. This would be worthless without an inversion formula giving back the function f in terms of $\mathcal{F}(f)$:

$$\forall x \in \mathbb{R}^d, \quad f(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{ix\cdot\xi} \mathcal{F}(f)(\xi)d\xi,$$

Unfortunately, this formula only makes sense when $\mathcal{F}(f) \in L^1(\mathbb{R}^d)$, and this is not the case in general for $f \in L^1$. In this chapter, we are going to introduce a subspace of distributions, which contains $L^1(\mathbb{R}^d)$, and a Fourier transformation on this space, which sastifies the two above identities, in an appropriate sense. This space will thus help us solve certain linear PDEs.

In most of the chapter, we will consider functions and distributions defined in the full space $\mathbb{R}^d$. Yet, some applications will also deal with functions or distributions on some open set $\Omega \subset \mathbb{R}^d$.

3.1 The Schwartz space

Before introducing distributions on $\mathbb{R}^d$ in the previous chapter, we needed to properly define the space of test functions $\mathscr{D}(\mathbb{R}^d)$. Test functions are in $L^1(\mathbb{R}^d)$, so we may define their Fourier transforms.

However, we will see that the Fourier transform of a nonzero test function φ is never compactly supported, showing that $\mathcal{F}$ does not map the space $\mathscr{D}(\mathbb{R}^d)$ to itself.

This is the reason why, in the next section, we will introduce a new space of smooth functions, which are not necessarily compactly supported, but which will decay fast at infinity, called the space of Schwartz functions. We will then show that this the Fourier transform maps this space of functions to itself. Besides, the Schwartz space is left invariant by differentiation and multiplication by polynomials. We will view this space as an "enlargement" of the space $\mathscr{D}(\mathbb{R}^d)$ of test functions.

In a second step, we will introduce a subspace of distributions which can be extended into linear forms on the Schwartz space. By duality, the Fourier transform will act on this space of distributions (called tempered distributions, or Schwartz distributions); this will allow us to solve certain PDEs on this space of distributions.

3.1.1 Definitions and examples

Definition 3.1.1 A function $f : \mathbb{R}^d \to \mathbb{R}$ is said to be rapidly decreasing if, as $|x| \to \infty$,

$$\forall \alpha \in \mathbb{N}^d, \qquad x^\alpha f(x) \to 0 \ .$$

Notice that, if f is rapidly decreasing and continuous, all the functions $x^\alpha f$ are bounded on $\mathbb{R}^d$. Conversely, if f is continuous on $\mathbb{R}^d$ and all functions $x^\alpha f$ are in $L^\infty(\mathbb{R}^d)$, then f is rapidly decreasing.

Definition 3.1.2 We denote by $\mathscr{S}(\mathbb{R}^d)$ the set of functions $\varphi \in \mathcal{C}^\infty(\mathbb{R}^d)$ which are rapidly decreasing, as well as all their derivatives. In other words, they satisfy:

$$\forall (\alpha, \beta) \in \mathbb{N}^d, \qquad x^\alpha \partial^\beta \varphi(x) \to 0 \quad \text{as } |x| \to \infty \,.$$

The set $\mathscr{S}(\mathbb{R}^d)$ is a vector space called the Schwartz space on $\mathbb{R}^d$, or the space of Schwartz functions on $\mathbb{R}^d$.

Examples 3.1.3 *i)* $\mathcal{C}_0^\infty(\mathbb{R}^d) \subset \mathscr{S}(\mathbb{R}^d)$.

ii) For $z \in \mathbb{C}$ such that $\operatorname{Re} z > 0$, the Gaussian function $\varphi(x) = e^{-z|x|^2}$ belongs to $\mathscr{S}(\mathbb{R}^d)$.

iii) If φ_1, $\varphi_2 \in \mathscr{S}(\mathbb{R}^d)$, then $\varphi_1 \varphi_2 \in \mathscr{S}(\mathbb{R}^d)$.

iv) No nontrivial rational function belongs to $\mathscr{S}(\mathbb{R}^d)$.

The topology on $\mathscr{S}(\mathbb{R}^d)$ we will work with is generated by the family of norms $(N_p)_{p\in\mathbb{N}}$ given by

$$\boxed{N_p(\varphi) = \sup_{|\alpha|,|\beta| \le p} \sup |x^\alpha \partial^\beta \varphi(x)|}.$$

It is clear that for $\varphi \in \mathcal{C}^\infty(\mathbb{R}^d)$, we have the equivalence

$$\varphi \in \mathscr{S}(\mathbb{R}^d) \iff \forall p \in \mathbb{N}, \quad N_p(\varphi) < +\infty \iff \forall \alpha, \beta \in \mathbb{N}^d, \quad x^\alpha \partial^\beta \varphi \in L^\infty .$$

The next Proposition shows a stability of $\mathscr{S}$ through differentiation or multiplication by polynomials.

Proposition 3.1.4 If $\varphi \in \mathscr{S}(\mathbb{R}^d)$ then $x^\alpha \partial^\beta \varphi \in \mathscr{S}(\mathbb{R}^d)$ for every $\alpha, \beta \in \mathbb{N}^d$.

Proof.— This follows immediately from the fact that for multi-indices $|\alpha|, |\beta| \le q$,

(3.1.1) $$N_p(x^\alpha \partial^\beta \varphi) = \sup_{|\lambda|,|\mu| \le p} \sup_x |x^\lambda \partial^\mu (x^\alpha \partial^\beta \varphi(x))| \le C_{p,\alpha,\beta}\, N_{p+q}(\varphi) .$$

Here we applied the Leibniz formula to transform $\partial^\mu(x^\alpha \partial^\beta \varphi)$ into a sum of terms $x^\bullet \partial^\bullet \varphi$. The constant $C_{p,\alpha,\beta}$ results from the corresponding combinatorial factors, and of the sum over all the terms. □

3.1.2 Convergence in $\mathscr{S}(\mathbb{R}^d)$ and density results

Definition 3.1.5 Let (φ_n) be a sequence of functions in $\mathscr{S}(\mathbb{R}^d)$. One says that (φ_n) converges to φ in $\mathscr{S}(\mathbb{R}^d)$ when, for every $\alpha, \beta \in \mathbb{N}^d$, $x^\alpha \partial^\beta \varphi_n(x) \to x^\alpha \partial^\beta \varphi(x)$ uniformly in $x \in \mathbb{R}^d$. Equivalently, for all $p \in \mathbb{N}$,

$$N_p(\varphi_n - \varphi) \to 0 \text{ as } n \to +\infty.$$

Remark 3.1.6 $i)$ Comparing this definition to the convergence in $\mathscr{D}(\mathbb{R}^d)$ (see Definition 1.2.10), we replaced the condition of uniform support for the φ_n, into a condition of "uniform fast decay" when x goes to infinity.

$ii)$ Let $\alpha \in \mathbb{N}^d$. It follows from (3.1.1) that, if (φ_n) converges to φ in $\mathscr{S}(\mathbb{R}^d)$, then $(x^\alpha \varphi_n)$ converges to $x^\alpha \varphi$ and $(\partial^\alpha \varphi_n)$ converges to $(\partial^\alpha \varphi)$ in $\mathscr{S}(\mathbb{R}^d)$. Otherwise stated, multiplication by a polynomial and differentiation act continuously in $\mathscr{S}(\mathbb{R}^d)$.

We already know that $\mathcal{S} \subset L^\infty$, from the fact that $\|\varphi\|_{L^\infty} = N_0(\varphi)$. The next proposition states a similar property for L^1.

Proposition 3.1.7 $\mathscr{S}(\mathbb{R}^d) \subset L^1(\mathbb{R}^d)$. More precisely, there exists a constant $C_d > 0$ such that

$$\forall \varphi \in \mathscr{S}(\mathbb{R}^d) \ , \quad \|\varphi\|_{L^1} \leq C_d \, N_{d+1}(\varphi) \ .$$

This estimate shows that the injection $\mathscr{S} \hookrightarrow L^1$ is continuous.

Proof.— We have the polynomial expansion

$$(1+|x|^2)^{d+1} = (1 + x_1^2 + \cdots + x_d^2)^{d+1} = \sum_{|\alpha| \leq d+1} c_{\alpha,d} \, x^{2\alpha}$$

for some combinatorial coefficients $c_{\alpha,d} \in \mathbb{N}$, hence for some constant $B_d > 0$ and any $\varphi \in \mathscr{D}(\mathbb{R}^d)$ we find:

$$\forall x \in \mathbb{R}^d, \qquad (1+|x|^2)^{d+1} |\varphi(x)|^2 \leq B_d \, N_{d+1}(\varphi)^2 \ .$$

Consequently,

$$\int_{\mathbb{R}^d} |\varphi(x)| \, dx \leq \sqrt{B_d} N_{d+1}(\varphi) \int_{\mathbb{R}^d} \frac{dx}{(1+|x|^2)^{(d+1)/2}} \leq C_d \, N_{d+1}(\varphi) \ ,$$

since $x \mapsto (1+|x|^2)^{-(d+1)/2}$ is integrable on $\mathbb{R}^d$. $\square$

Remark 3.1.8 For every $q \in \mathbb{N}$, one can establish similarly

$$\forall \varphi \in \mathscr{S}(\mathbb{R}^d) \ , \quad \forall x \in \mathbb{R}^d \ , \qquad (1+|x|^q)|\varphi(x)| \leq C_{q,d} N_q(\varphi) \ .$$

Since $\mathcal{C}_0^\infty(\mathbb{R}^d)$ is dense in each space $L^p(\mathbb{R}^d)$ space for $p \in [1, +\infty[$, we get the

Corollary 3.1.9 For each $p \in [1, +\infty[$, the space $\mathscr{S}(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d)$.

We also have the important density of test functions:

Proposition 3.1.10 The space $\mathscr{D}(\mathbb{R}^d) = \mathcal{C}_0^\infty(\mathbb{R}^d)$ is dense in $\mathscr{S}(\mathbb{R}^d)$: for any $\varphi \in \mathscr{S}(\mathbb{R}^d)$, there is a sequence $(\varphi_n \in \mathscr{D}(\mathbb{R}^d))_{n \geq 1}$ which converges to φ in the topology of $\mathscr{S}(\mathbb{R}^d)$.

Proof.— Take a function $\varphi \in \mathscr{S}(\mathbb{R}^d)$. We want to construct approximations $\chi_n \in \mathscr{D}(\mathbb{R}^d)$, such that for each $p \in \mathbb{N}$, $N_p(\varphi_n - \varphi) \to 0$ when $n \to \infty$.

We will proceed by smoothly truncating φ in larger and larger balls. Let $\chi \in \mathscr{D}(\mathbb{R}^d)$ be a plateau function over $B(0,1)$. For each $n \geq 1$ we set $\varphi_n(x) = \varphi(x)\chi(x/n)$. Each function φ_n is in $\mathscr{D}(\mathbb{R}^d)$, and equals φ inside $B(0,n)$.

In order to control the norms $N_p(\varphi_n - \varphi)$, we use the Leibniz formula to compute:

$$\partial^\beta(\varphi - \varphi_n)(x) = \partial^\beta\varphi(x)\,(1 - \chi(x/n)) + \sum_{|\gamma|\geq 1, \gamma\leq\beta} \binom{\beta}{\gamma} \frac{1}{n^{|\gamma|}}\, \partial^{\beta-\gamma}\varphi(x)\,(\partial^\gamma\chi)(x/n)\,.$$

If we multiply this expression by x^α, we get for a certain constant $C > 0$:

$$\forall n \geq 1, \qquad \|x^\alpha\partial^\beta(\varphi - \varphi_n)(x)\|_\infty \leq \sup_{|x|\geq n} |x^\alpha\partial^\beta\varphi(x)| + \frac{C}{n}\sum_{\gamma\leq\beta} \|x^\alpha\partial^{\beta-\gamma}\varphi\|_\infty\,.$$

The first term on the right hand-side goes to zero when $n \to \infty$, due to the fast decay of $\partial^\beta\varphi$. The sum on the RHS is controlled by the norm $N_p(\varphi)$ such that $p = \max(|\alpha|, |\beta|)$. Due to the prefactor $1/n$, the second term also vanishes when $n \to \infty$. This achieves to show that $N_p(\varphi - \varphi_n) \to 0$. □

3.2 The Fourier transformation in $\mathscr{S}(\mathbb{R}^d)$

In this section we define, and study the properties of the Fourier transformation acting on functions $\varphi \in \mathscr{S}(\mathbb{R}^d)$.

3.2.1 Definition and first properties

Any $\varphi \in \mathscr{S}(\mathbb{R}^d)$ satisfies $\varphi \in L^1(\mathbb{R}^d)$, so that $\mathcal{F}(\varphi)$ is well defined ans belongs to $L^\infty(\mathbb{R}^d)$.

Definition 3.2.1 For $\varphi \in \mathscr{S}(\mathbb{R}^d)$, we denote $\hat\varphi$, $\mathcal{F}(\varphi)$ or even $\mathcal{F}_{x\to\xi}(\varphi(x))$, the function in $L^\infty(\mathbb{R}^d)$ given by

$$\forall \xi \in \mathbb{R}^d, \qquad \hat\varphi(\xi) = \mathcal{F}(\varphi)(\xi) = \mathcal{F}_{x\to\xi}(\varphi(x)) = \int_{\mathbb{R}^d} e^{-ix\cdot\xi}\varphi(x)dx.$$

The linear map $\mathcal{F} : \varphi \mapsto \hat\varphi$ is called the Fourier transformation.

Here follow the elementary, yet important properties of the Fourier transform acting on the space $\mathscr{S}(\mathbb{R}^d)$.

Proposition 3.2.2 Let $\varphi \in \mathscr{S}(\mathbb{R}^d)$. Then $\hat\varphi \in \mathscr{S}(\mathbb{R}^d)$. More precisely,

i) For all $j \in \{1, \ldots, n\}$ $\mathcal{F}_{x\to\xi}(x_j\varphi(x)) = i\partial_{\xi_j}\mathcal{F}(\varphi)(\xi)$.

ii) For all $j \in \{1, \ldots, n\}$, we have $\mathcal{F}(\partial_{x_j}\varphi)(\xi) = i\xi_j\mathcal{F}(\varphi)(\xi)$.

iii) For $a \in \mathbb{R}^d$, $\mathcal{F}_{x\to\xi}(\varphi(x-a)) = e^{-ia\cdot\xi}\mathcal{F}(\varphi)(\xi)$.

iv) For $a \in \mathbb{R}^d$, $\mathcal{F}_{x\to\xi}(e^{ia\cdot x}\varphi(x)) = \mathcal{F}(\varphi)(\xi - a)$.

v) For all integrer $p \geq 0$, $N_p(\hat{\varphi}) \leq C_{p,d}\, N_{p+d+1}(\varphi)$ for some absolute constant $C_{p,d} > 0$.

Proof.— $i)$ The function $(x,\xi) \mapsto e^{-ix\cdot\xi}\varphi(x)$ is $\mathcal{C}^1$ on $\mathbb{R}^d \times \mathbb{R}^d$, and

$$\forall \xi \in \mathbb{R}^d, \qquad |\partial_{\xi_j}(e^{-ix\cdot\xi}\varphi(x))| = |-ix_j e^{-ix\cdot\xi}\varphi(x)| = |x_j\varphi(x)| \in L^1(\mathbb{R}^d).$$

By the theorem of derivation under the integral, we get that $\mathcal{F}(\varphi)$ is differentiable at each point ξ, and

$$\partial_{\xi_j}\mathcal{F}(\varphi)(\xi) = \int_{\mathbb{R}^d} -ix_j e^{-ix\cdot\xi}\varphi(x) = \mathcal{F}(-ix_j\varphi(x)).$$

The theorem of continuity under the integral shows that this function of ξ is continuous. Altogether $\mathcal{F}(\varphi)$ is $\mathcal{C}^1$.

ii) Let us first write the proof for $j = 1$. Integrating by parts, we get

$$\int_{\mathbb{R}} \partial_1\varphi(x)e^{-ix\cdot\xi}dx_1 = i\xi_1 \int_{\mathbb{R}} \varphi(x)e^{-ix\cdot\xi}dx_1$$

Now we integrate with respect to the variable x'. By Fubini, since $\varphi, \partial_1\varphi \in \mathscr{S}(\mathbb{R}^d) \subset L^1(\mathbb{R}^d)$, we have

$$\int_{\mathbb{R}^d} \partial_1\varphi(x)e^{-ix\cdot\xi}dx = i\xi_1 \int_{\mathbb{R}^d} \varphi(x)e^{-ix\cdot\xi}dx.$$

To get $iv)$, we only have to write

$$\mathcal{F}_{x\to\xi}(e^{ia\cdot x}\varphi(x)) = \int_{\mathbb{R}^d} e^{-ix\cdot\xi}e^{ia\cdot x}\varphi(x)dx = \int_{\mathbb{R}^d} e^{-i(\xi-a)\cdot x}\varphi(x)dx = \mathcal{F}(\varphi)(\xi - a).$$

Eventually, performing a change of variable, we have

$$\mathcal{F}_{x\to\xi}(\varphi(x-a)) = \int_{\mathbb{R}^d} e^{-ix\cdot\xi}\varphi(x-a)dx = \int_{\mathbb{R}^d} e^{-i(x+a)\cdot\xi}\varphi(x)dx = e^{-ia\cdot\xi}\mathcal{F}(\varphi)(\xi),$$

and this is property $iii)$.

Since $x_j\varphi \in L^1$, the property $i)$ implies that $\partial_j\hat{\varphi}$ is in L^∞. Similarly, $ii)$ implies hat $\xi_j\hat{\varphi} \in L^\infty$. Since $\varphi \in \mathscr{S}$, $\partial^\alpha(x^\beta\varphi) \in L^1$ for any multiindices α, β, so iterating the properties $i)$, $ii)$ we get that $\xi^\alpha\partial_\xi^\beta\hat{\varphi} \in L^\infty$; this shows that $\hat{\varphi} \in \mathscr{S}(\mathbb{R}^d)$.

Furthermore, using Proposition 3.1.7 we find explicit bounds for the norms $\|\xi^\alpha\partial_\xi^\beta\hat{\varphi}\|_\infty$, which lead to the following bounds for the norms $N_p(\hat{\varphi})$:

$$\begin{aligned} N_p(\hat{\varphi}) &= \max_{|\alpha|,|\beta|\leq p} \|\xi^\alpha\partial^\beta\hat{\varphi}\|_{L^\infty} = \max_{|\alpha|,|\beta|\leq p} \|\mathcal{F}(\partial^\alpha(x^\beta\varphi))\|_{L^\infty} \\ &\leq \max_{|\alpha|,|\beta|\leq p} \|\partial^\alpha(x^\beta\varphi)\|_{L^1} \leq B_d \max_{|\alpha|,|\beta|\leq p} N_{d+1}(\partial^\alpha(x^\beta\varphi)) \\ &\leq C_{p,d}\, N_{p+d+1}(\varphi)\ . \end{aligned}$$

These controls show that the the Fourier transform $\mathcal{F}$ acts continuously $\mathscr{S} \to \mathscr{S}$. □

Notation

Because of the presence of a factor $i = \sqrt{-1}$ in $i)$ and $ii)$, it is sometimes convenient to use the notation

$$D_j = \frac{1}{i}\partial_j.$$

With this notation, $ii)$ becomes $\mathcal{F}(D_{x_j}\varphi) = \xi_j \mathcal{F}(\varphi)$, and $i)$ reads $D_{\xi_j}\mathcal{F}(\varphi) = -\mathcal{F}(x_j\varphi)$. Summing up, we have

$$\begin{cases} \widehat{D_{x_j}\varphi} = \xi_j \widehat{\varphi}, \\ \widehat{x_j\varphi} = -D_{\xi_j}\widehat{\varphi}. \end{cases}$$

3.2.2 Example of the Gaussian function

Let us compute the Fourier transform of the simplest "nontrivial" Schwartz function, namely the Gaussian. For a parameter $\lambda > 0$, we set

$$G_\lambda(x) = \exp\big(-\lambda\frac{|x|^2}{2}\big)\ , \qquad x \in \mathbb{R}^d\ .$$

Proposition 3.2.3 The Fourier transform of the Gaussian G_λ is given by

$$\hat{G}_\lambda(\xi) = \Big(\frac{2\pi}{\lambda}\Big)^{\frac{d}{2}} G_{\frac{1}{\lambda}}(\xi), \qquad \xi \in \mathbb{R}^d\ .$$

That is, the Fourier tranform maps a Gaussian to another Gaussian (up to a change of global prefactor and parameter).

Proof.— If we denote by g_λ the 1-dimensional Gaussian with parameter λ, we notice that $G_\lambda(x_1,\ldots,x_d) = \prod_{j=1}^d g_\lambda(x_j)$. Since the exponential $e^{-ix\cdot\xi}$ also factorizes, we find $\mathcal{F}(G_\lambda)(\xi) = \prod_{j=1}^d \mathcal{F}(g_\lambda)(\xi_j)$. We are thus reduced to computing the Fourier transform of the 1-dimensional Gaussian g_λ.

For that aim, we observe that g_λ satisfies a simple ODE:

$$\forall x \in \mathbb{R}, \qquad g_\lambda'(x) = -\lambda x\, g_\lambda(x)\ .$$

The properties $i)$ and $ii)$ of Proposition 3.2.2 lead to

$$i\xi\hat{g}_\lambda(\xi) = -i\lambda\frac{d}{d\xi}\hat{g}_\lambda(\xi)\ .$$

Solving this ODE leads to the solution

$$\hat{g}_\lambda(\xi) = \hat{g}_\lambda(0)\,\mathrm{e}^{-\frac{\xi^2}{2\lambda}} = \hat{g}_\lambda(0)\, g_{1/\lambda}(\xi)\ .$$

To finish the proof we just need to compute the value at $\xi = 0$, which is given by the Gauss integral:

$$\hat{g}_\lambda(0) = \int_{\mathbb{R}} \mathrm{e}^{-\lambda \frac{x^2}{2}}\, dx = \sqrt{\frac{2\pi}{\lambda}}\ .$$

$\square$

Exercise 3.2.4 Alternatively, one can compute $\hat{g}_\lambda(\xi)$ by completing the square in the the exponent in the Fourier integral:

$$-\lambda x^2/2 - ix\xi = -\frac{\lambda}{2}\big(x + i\xi/\lambda\big)^2 - \frac{\xi^2}{2\lambda}\,,$$

which results in the complex contour integral $e^{-\frac{\xi^2}{2\lambda}} \int_{\mathrm{Im}\,\zeta=\xi/\lambda} e^{-\lambda\frac{\zeta^2}{2}}\, d\zeta$. The integral can be brought to the above Gauss integral by a contour deformation (which needs to be carefully justified).

3.2.3 The Inversion formula

A very important property of the Fourier transform on $\mathscr{S}(\mathbb{R}^d)$ is the fact that it acts bijectively $\mathscr{S} \to \mathscr{S}$, and that the inverse map is as simple as $\mathcal{F}$ itself.

Theorem 3.2.5 Let $\varphi \in \mathscr{S}(\mathbb{R}^d)$. Then

$$\forall x \in \mathbb{R}^d, \qquad \varphi(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \mathrm{e}^{ix\cdot\xi} \hat{\varphi}(\xi)\, d\xi\ .$$

Compared with the expression of $\mathcal{F}$ itself, we just have a change of sign in the exponential, and a global prefactor $(2\pi)^{-d}$.

Proof.— We first observe that it is enough to prove this identity at the point $x = 0$. Indeed, in view of property $iii)$ in Proposition 3.2.2, the general statement follows from applying the formula

$$\psi(0) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{\psi}(\xi)\, d\xi \tag{3.2.2}$$

to the function ψ_x defined by $\psi_x(\bullet) = \varphi(x + \bullet)$.

To prove (3.2.2), we consider two linear form on $\mathscr{S}(\mathbb{R}^d)$:

$$L_0 : \psi \mapsto \psi(0), \qquad {}_1 : \psi \mapsto \int_{\mathbb{R}^d} \hat{\psi}(\xi)\, d\xi\,.$$

We claim that, if $L_0(\psi) = 0$, then $L_1(\psi) = 0$. Indeed, applying the Hadamard lemma 2.3.13 and its proof, we may decompose $\psi \in \mathrm{Ker}(L_0)$ as

$$\psi(x) = \sum_{j=1}^{d} x_j \psi_j(x)\ ,$$

where the explicit expressions for ψ_j show that these functions are in $\mathscr{S}(\mathbb{R}^d)$. Taking the Fourier transform of both sides and applying property $i)$ of Proposition 3.2.2, we infer

$$\hat{\psi}(\xi) = \sum_{j=1}^{d} i\partial_j \hat{\psi}_j(\xi) \ .$$

Integrating both sides on $\mathbb{R}^d$, we conclude

$$\int_{\mathbb{R}^d} \hat{\psi}(\xi)\, d\xi = 0 \ , \quad \text{that is,} \quad L_1(\psi) = 0 \, .$$

We have proved that the kernel of L_0 is contained in the kernel of L_1. Since these linear forms act continuously on $\mathscr{S}$, their kernels are closed hyperplanes in $\mathscr{S}$. The inclusion implies that those kernels are the same hyperplane, thus that the two forms are proportional each other: there exists $c \in \mathbb{C}^*$ such that $L_1 = c\, L_0$.

To determine the constant c, it is sufficient to apply this identity to a nontrivial particular case. Taking $\psi = G_1$ and using Proposition 3.2.3 shows that $c = (2\pi)^d$, hence the formula (3.2.2). $\square$

In order to reformulate this important theorem, we introduce the operator of symmetry w.r.to the origin

$$\sigma : \mathscr{S} \to \mathscr{S} \quad \text{defined by} \quad \sigma\varphi(x) := \varphi(-x) \ , \quad \forall x \in \mathbb{R}^d \, ,$$

and observe that

$$\sigma \circ \mathcal{F} = \mathcal{F} \circ \sigma =: \check{\mathcal{F}} \ .$$

The transformation $\check{\mathcal{F}}$ is given by the integral in Definition 3.2.1, up to replacing $e^{-ix\cdot\xi}$ by $e^{ix\cdot\xi}$.

Corollary 3.2.6 The Fourier transformation is an isomorphism on the vector space $\mathscr{S}(\mathbb{R}^d)$. Its inverse $\mathcal{F}^{-1}$ is given by

$$\mathcal{F}^{-1} = (2\pi)^{-d} \check{\mathcal{F}} \ .$$

Equivalently, we have

$$\mathcal{F} \circ \mathcal{F} = (2\pi)^d \sigma \ . \tag{3.2.3}$$

Remark 3.2.7 As an immediate consequence of (3.2.3), $\mathcal{F}^4 = (\mathcal{F} \circ \mathcal{F})^2 = (2\pi)^{2d}\, Id$.

3.2.4 The "change of head" lemma

The following proposition is elementary but crucial for the whole chapter.

Proposition 3.2.8 Let $\varphi, \psi \in \mathscr{S}(\mathbb{R}^d)$. Then

$$\int_{\mathbb{R}^d} \hat{\varphi}(\xi)\, \psi(\xi)\, d\xi = \int_{\mathbb{R}^d} \varphi(x)\, \hat{\psi}(x)\, dx \ .$$

Proof.— Since $\varphi, \psi \in \mathscr{S}$, these functions are smooth and fast decaying, so it is possible to use Fubini's theorem:

$$\begin{aligned}\int_{\mathbb{R}^d} \hat{\varphi}(\xi)\psi(\xi)\,d\xi &= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} \mathrm{e}^{-ix.\xi}\varphi(x)\,dx\right)\psi(\xi)\,d\xi \\ &\overset{Fubini}{=} \int_{\mathbb{R}^d}\varphi(x)\left(\int_{\mathbb{R}^d}\mathrm{e}^{-ix.\xi}\psi(\xi)\,d\xi\right)\,dx = \int_{\mathbb{R}^d}\varphi(x)\hat{\psi}(x)\,dx\,.\end{aligned}$$

□

This change-of-head property will be used in the next section, which will prove an important identity.

3.2.5 The Plancherel identity

Proposition 3.2.9 (Plancherel formula) Let φ and ψ be two functions in $\mathscr{S}(\mathbb{R}^d)$. Then

$$\int_{\mathbb{R}^d}\varphi(x)\,\overline{\psi}(x)\,dx = (2\pi)^{-d}\int_{\mathbb{R}^d}\hat{\varphi}(\xi)\,\overline{\hat{\psi}}(\xi)\,d\xi\,.$$

Proof.— In the proof we will pile up several successive operations: Fourier transform, complex conjugation, inversion. The symbols above the functions should be read from bottom to top.

We first observe the following identity:

$$\overline{\mathcal{F}(\varphi)} = \overline{\hat{\varphi}} = \sigma\big(\hat{\overline{\varphi}}\big)\,, \tag{3.2.4}$$

which is merely a reformulation of

$$\overline{\hat{\varphi}}(\xi) = \overline{\int_{\mathbb{R}^d}\mathrm{e}^{-ix\cdot\xi}\,\varphi(x)\,dx} = \int_{\mathbb{R}^d}\mathrm{e}^{ix.\xi}\,\overline{\varphi}(x)\,dx\,.$$

Applying the "change-of-head" Proposition 3.2.8, we obtain

$$\int_{\mathbb{R}^d}\hat{\varphi}\,\overline{\hat{\psi}} = \int_{\mathbb{R}^d}\varphi\hat{\overline{\hat{\psi}}}.$$

Applying the identity (3.2.4) to $\hat{\psi}$, and then identity (3.2.3), we get

$$\hat{\overline{\hat{\psi}}} = \sigma\overline{\hat{\hat{\psi}}} = (2\pi)^d\sigma\overline{\sigma\psi} = (2\pi)^d\overline{\psi}\,.$$

This completes the proof. □

The above identity can be written in terms of the Hermitian scalar product on $L^2(\mathbb{R}^d,\mathbb{C})$:

$$(\varphi,\psi)_{L^2} = \frac{1}{(2\pi)^d}(\hat{\varphi},\hat{\psi})_{L^2}.$$

In particular, for $\psi=\varphi$, we obtain the famous Plancherel formula, here for functions in $\mathscr{S}(\mathbb{R}^d)$:

Corollary 3.2.10 (Plancherel) For any $\varphi \in \mathscr{S}(\mathbb{R}^d)$, it holds that

$$\|\varphi\|_{L^2(\mathbb{R}^d)} = (2\pi)^{-d/2}\|\hat{\varphi}\|_{L^2(\mathbb{R}^d)}\,.$$

Now that we have explored the space of Schwartz functions and the action of the Fourier transform on it, we are ready to define the tempered (or Schwartz) distributions.

3.3 The space $\mathscr{S}'(\mathbb{R}^d)$ of tempered distributions

3.3.1 Definition, examples

Initially, tempered distributions are just a subclass of distributions on $\mathbb{R}^d$.

Definition 3.3.1 A distribution $T \in \mathscr{D}'(\mathbb{R}^d)$ is said to be *tempered* when there exists $C > 0$ and $p \in \mathbb{N}$ such that, for all test function $\varphi \in \mathscr{D}(\mathbb{R}^d)$,

$$|\langle T, \varphi\rangle \leq C N_p(\varphi)\,.$$

Tempered distributions are also called Schwartz distributions. We denote by $\mathscr{S}'(\mathbb{R}^d)$ the space of tempered distributions on $\mathbb{R}^d$.

Example 3.3.2 If $T \in \mathcal{E}'(\mathbb{R}^d)$, the space of distributions of compact supports, then there is $C > 0$, $m \in \mathbb{N}$, such that for any $\varphi \in \mathscr{D}(\mathbb{R}^d)$ (independently of the support of φ),

$$|\langle T, \varphi\rangle| \leq C \sum_{|\alpha|\leq m} \text{sup}\,|\partial^\alpha \varphi| \leq C N_m(\varphi).$$

Thus $\mathcal{E}'(\mathbb{R}^d) \subset \mathscr{S}'(\mathbb{R}^d)$: compactly supported distributions are tempered.

Example 3.3.3 For $p \in [1, +\infty]$, we have $L^p(\mathbb{R}^d) \subset \mathscr{S}'(\mathbb{R}^d)$. Indeed, if $f \in L^p(\mathbb{R}^d)$ and $\varphi \in \mathscr{D}(\mathbb{R}^d)$, and for $q \in [1, +\infty]$ such that $1/p + 1/q = 1$, we have

$$|\langle T_f, \varphi\rangle| \leq |\int f(x)\varphi(x)dx| \overset{Hlder}{\leq} \|f\|_{L^p}\|\varphi\|_{L^q}\,.$$

A simple adaptation of Proposition 3.2.2 shows that if an integer n satisfies $n > \frac{d}{q}$, then $\|\varphi\|_{L^q} \leq C\, N_n(\varphi)$. We thus obtain, for some constant $C_{p,n,d} > 0$:

$$|\langle T_f, \varphi\rangle| \leq C_{p,n,d}\,\|f\|_{L^p}\, N_n(\varphi)\,.$$

In particular, in view of Proposition 3.1.7, we see that $\mathscr{S}(\mathbb{R}^d) \subset \mathscr{S}'(\mathbb{R}^d)$.

Example 3.3.4 Let $f \in L^1_{loc}(\mathbb{R}^d)$ be growing not too fast when $|x| \to \infty$ in the following averaged sense: for some $C > 0$ and $p \in \mathbb{N}$,

$$\forall R > 0, \quad \int_{|x|\leq R} |f(x)|\, dx \leq C_f (1+R)^p \,. \tag{3.3.5}$$

Then we claim that f is a tempered distribution. Indeed, if $\varphi \in \mathscr{D}(\mathbb{R}^d)$, we have

$$|\langle T_f, \varphi\rangle| \leq \int_{\mathbb{R}^d} |f(x)||\varphi(x)|\, dx \,.$$

The integral on $\mathbb{R}^d$ will be split by using a dyadic partition of $\mathbb{R}^d$ into the unit ball and the annuli $A_k := \{2^k \leq |x| < 2^{k+1}\}$, for $k \in \mathbb{N}$.

$$|\langle T_f, \varphi\rangle| \leq \int_{|x|<1} |f(x)|\, |\varphi(x)|\, dx + \sum_{k=0}^{\infty} \int_{A_k} |f(x)|\, |\varphi(x)\, dx.$$

The hypothesis (3.3.5) easily leads to the following bounds:

$$\int_{|x|<1} |f(x)|\, |\varphi(x)|\, dx \leq C\,(1+1)^p \|\varphi\|_\infty = C_f\, 2^p\, N_0(\varphi)\,,$$
$$\int_{A_k} |f(x)|\, |\varphi(x)|\, dx \leq C_f\,(1+2^{k+1})^p \sup_{x\in A_k} |\varphi(x)|$$

Notice that we did not just use $\|phe\|_\infty$ in the second line, because we want to use some decay property of φ in order to sum over $k \in \mathbb{N}$.

Since $|x| \geq 2^k$ on A_k, we have obviously

$$2^{kp+k} \sup_{x\in A_k} |\varphi(x)| \leq \sup_{x\in A_k} |x|^{p+1}|\varphi(x)|\,, \quad \text{hence}$$
$$(1+2^{k+1})^p \sup_{x\in A_k} |\varphi(x)| \leq C_p\, 2^{-k} \sup_{x\in A_k} |x|^{p+1}|\varphi(x)| \leq C'_p\, 2^{-k}\, N_{p+1}(\varphi)\,.$$

This rewriting allows us to sum over $k \in \mathbb{N}$, and obtain finally

$$|\langle T_f, \varphi\rangle| \leq C''_p\, C_f\, N_{p+1}(\varphi)\,.$$

Example 3.3.5 The previous example admits the following partial converse. If $f \in L^1_{loc}(\mathbb{R}^d)$ has nonnegative values and is a tempered distribution, then, there exists $C > 0$ and $p \in \mathbb{N}$ such that:

$$\forall R > 0, \qquad \int_{|x|\leq R} f(x)\, dx \leq C\,(1+R)^p.$$

Indeed, if f is tempered, we have, for some $A > 0$ et $p \in \mathbb{N}$,

$$\forall \varphi \in \mathscr{D}(\mathbb{R}^d), \qquad \left| \int_{\mathbb{R}^d} f(x)\, \varphi(x)\, dx \right| \leq A\, N_p(\varphi)\ .$$

Let $\chi \geq 0$ be a plateau function over the ball $B(0,1)$, supported in the ball $B(0,2)$. Then

$$0 \leq \int_{|x|\leq R} f(x)\,dx \leq \int_{\mathbb{R}^d} f(x)\chi\big(\frac{x}{R}\big)\,dx \leq A\,N_p\Big(\chi\big(\frac{\cdot}{R}\big)\Big)\,,$$

and an explicit computation (using the fact that supp $\chi\big(\frac{\cdot}{R}\big) \subset B(0,2R)$) shows that

$$N_p\Big(\chi\big(\frac{\cdot}{R}\big)\Big) \leq C\,(1+R)^p\,.$$

In particular, on $\mathbb{R}$ the function e^x is not tempered, since, for every $p \in \mathbb{N}$,

$$R^{-p}\int_0^R e^x\,dx \underset{R\to\infty}{\longrightarrow} +\infty\ .$$

We conclude this first paragraph with a simple observation.

Proposition 3.3.6 If $T \in \mathscr{S}'(\mathbb{R}^d)$, then $x^\alpha \partial^\beta T \in \mathscr{S}'(\mathbb{R}^d)$ for all $\alpha, \beta \in \mathbb{N}^d$. In other words, the space of tempered distributions is left invariant by differentiation and multiplication by polynomials.

Proof.— Assume that $T \in \mathscr{S}'$ is controlled by the norm N_p. It is sufficient to show that $x_j T$ and $\partial_j T$ are tempered distributions. For any test function $\varphi \in \mathscr{D}(\mathbb{R}^d) \subset \mathscr{S}(\mathbb{R}^d)$, we have

$$|\langle x_j T, \varphi\rangle| = |\langle T, x_j\varphi\rangle| \leq C\,N_p(x_j\varphi) \leq C\,C'_p\,N_{p+1}(\varphi),$$

where we made use of (3.1.1). A similar computation shows that $\partial_j T$ is controlled by N_{p+1} as well, hence that $\partial_j T \in \mathscr{S}'$. □

Exercise 3.3.7 Show that the function $x \mapsto e^x e^{ie^x}$ is not bounded by a polynomial, but nevertheless belongs to $\mathscr{S}'(\mathbb{R})$. Hint: it is the derivative of a tempered distribution.

The multiplication of a tempered distribution by a smooth function does not always yield a tempered distribution. However, it is the case when the function has *moderate growth*, a notion we now define.

Definition 3.3.8 A function $f \in \mathcal{C}^\infty(\mathbb{R}^d)$ has *moderate growth* if, for any $\beta \in \mathbb{N}^d$, there is $C_\beta > 0$ and $m_\beta \in \mathbb{N}$ such that

$$|\partial^\beta f(x)| \leq C_\beta(1+|x|)^{m_\beta}\,,\quad \forall x \in \mathbb{R}^d\,.$$

We denote by $\mathcal{O}_M(\mathbb{R}^d)$ the space of such functions.

We notice that polynomials are of moderate growth. It is also the case of continuous rational functions.

This notion allows us to extend the stability of $\mathscr{S}'$ through multiplication by polynomials.

Proposition 3.3.9 If $T \in \mathscr{S}'(\mathbb{R}^d)$ and $f \in \mathcal{O}_M(\mathbb{R}^d)$, then $fT \in \mathscr{S}'(\mathbb{R}^d)$.

Proof.— Let $\varphi \in \mathscr{S}(\mathbb{R}^d)$, and $\alpha, \beta \in \mathbb{N}^d$. The Leibniz formula gives

$$|x^\alpha \partial^\beta (f\varphi)| \leq \sum_{\gamma \leq \beta} \binom{\beta}{\gamma} |x^\alpha \partial^\gamma f| \, |\partial^{\beta-\gamma}\varphi| \leq C_\beta \sum_{\gamma \leq \beta} \binom{\beta}{\gamma} |x^\alpha| \, (1+|x|)^{m_\gamma} |\partial^{\beta-\gamma}\varphi|.$$

Then, as in (3.1.1), we obtain

$$N_p(f\varphi) \leq C N_{p+M}(\varphi),$$

where $M = |\alpha| + \max_{|\gamma| \leq p} m_\gamma$. Thus, for $\varphi \in \mathscr{D}(\mathbb{R}^d)$, we obtain

$$|\langle fT, \varphi \rangle| = |\langle T, f\varphi \rangle| \leq C N_p(f\varphi) \leq C' N_{p+M}(\varphi),$$

showing that fT is a tempered distribution. □

Exercise 3.3.10 Show that $\mathrm{vp}(1/x) \in \mathscr{S}'(\mathbb{R})$.

3.3.2 The fundamental characterization of $\mathscr{S}'$

So far our tempered distributions were acting on $\mathscr{D}(\mathbb{R}^d)$, like any distribution. The following Proposition shows that they can be naturally extended as linear forms on the space of Schwartz functions, which explains the notation $\mathscr{S}'$ for those distributions.

Proposition 3.3.11 If T is a tempered distribution, then T extends in a unique way as a continuous linear form $\widetilde{T}$ on $\mathscr{S}(\mathbb{R}^d)$, in the following sense: if $\varphi_n \to \varphi$ in $\mathscr{S}(\mathbb{R}^d)$, then $\langle \widetilde{T}, \varphi_n \rangle \to \langle \widetilde{T}, \varphi \rangle$.

Proof.— Let $\varphi \in \mathscr{S}(\mathbb{R})$. From Corollary 3.1.10, we may construct a sequence (φ_j) in $\mathscr{D}(\mathbb{R}^d)$ such that $\varphi_j \to \varphi$ in $\mathscr{S}(\mathbb{R}^d)$ when $j \to \infty$. The sequence $(\langle T, \varphi_j \rangle)_j$ is a Cauchy sequence in $\mathbb{C}$ since T is tempered:

$$(3.3.6) \qquad |\langle T, \varphi_j - \varphi_k \rangle| \leq C N_p(\varphi_j - \varphi_k) \quad \text{when } j, k \to \infty\,.$$

Its limit does not depend on the choice of the sequence (φ_j), since, for any other sequence $\psi_j \to \varphi$, we have

$$|\langle T, \varphi_j - \psi_j \rangle| \leq C N_p(\varphi_j - \psi_j) \; j \overset{\to}{\to} \infty \; 0\,.$$

Thus we can define $\tilde{T}$ the linear form on $\mathscr{S}(\mathbb{R}^d)$ by

$$\langle \tilde{T}, \varphi \rangle = \lim_{j \to +\infty} \langle T, \varphi_j \rangle,$$

where (φ_j) is *any* sequence in $\mathscr{D}(\mathbb{R}^d)$ which converges to φ. Since N_p is a continuous function on $\mathscr{S}$, taking the limit of the inequalities

$$|\langle T, \varphi_j \rangle| \leq C N_p(\varphi_j),$$

we get

$$|\langle \tilde{T}, \varphi\rangle| \leq C N_p(\varphi)\,.$$

This control holds for any $\varphi \in \mathscr{S}(\mathbb{R}^d)$. This shows that $\tilde{T}$ is a continuous linear form on $\mathscr{S}$.

Finally, if T_1 were another continuous extension of T to $\mathscr{S}(\mathbb{R}^d)$, we should have

$$\langle T_1, \varphi\rangle = \langle T_1, \varphi - \varphi_j\rangle + \langle T_1, \varphi_j\rangle \; j \overset{\to}{\to} \infty \; 0 + \langle \tilde{T}, \varphi\rangle,$$

so finally $T_1 = \tilde{T}$. □

Remark 3.3.12 For simplicity, we shall drop the $tilde$ above T in the sequel, and write T in place of $\tilde{T}$; that is, we identify the tempered distribution T with its extension to $\mathscr{S}$.

Notice that, conversely, every continuous linear form on $\mathscr{S}(\mathbb{R}^d)$ must satisfy a bound of the form

$$\forall \varphi \in \mathscr{S}, |\langle T, \varphi\rangle| \leq C\, N_p(\varphi)\,,$$

hence its restriction to $\mathscr{D}(\mathbb{R}^d)$ is a tempered distribution. Summing up, tempered distributions identify to continuous linear forms on the Schwartz space $\mathscr{S}(\mathbb{R}^d)$. This explains the notation $\mathscr{S}'$. We shall use this identification systematically.

3.3.3 Convergence in $\mathscr{S}'(\mathbb{R}^d)$

The notion of convergence for distributions in $\mathscr{S}'(\mathbb{R}^d)$ is similar to the one in $\mathscr{D}'(\mathbb{R}^d)$, but is slightly more restrictive.

Definition 3.3.13 Let (T_n) be a sequence of tempered distributions. One says that (T_n) tends to T in $\mathscr{S}'(\mathbb{R}^d)$ if, for any function $\varphi \in \mathscr{S}(\mathbb{R}^d)$, it holds that $\langle T_n, \varphi\rangle \; n \overset{\to}{\to} \infty \; \langle T, \varphi\rangle$.

As it is the case in $\mathscr{D}'(\mathbb{R}^d)$, this notion of convergence, a weak one, implies a stronger one. We admit the following *uniform boundedness* result.

Proposition 3.3.14 (Uniform boundedness in $\mathscr{S}'$) If $T_n \to T$ in $\mathscr{S}'$, there exists $C > 0$ and $p \in \mathbb{N}$ such that

$$\forall \varphi \in \mathscr{S}\ ,\ \forall n\ , \qquad |\langle T_n, \varphi\rangle| \leq C\, N_p(\varphi)\ .$$

This uniform boundedness principle can be proven by identifying the convergence on $\mathscr{S}$ as the one on a complete metric space.

Remark 3.3.15 When $T_n \to T$ in $\mathscr{S}'(\mathbb{R})$, it is true that $T_n \to T$ in $\mathscr{D}'(\mathbb{R})$, since $\mathscr{D}(\mathbb{R}^d) \subset \mathscr{S}(\mathbb{R}^d)$. The converse is not true in general, as shown by the following example: for any sequence (a_n) of complex numbers, the sequence $(a_n \delta_n)$ converges to 0 in $\mathscr{D}'(\mathbb{R})$. However it only converges (necessarily to 0) in $\mathscr{S}'(\mathbb{R}^d)$ if (a_n) has moderate growth, i.e. there is $C > 0$, $p \in \mathbb{N}$ such that for all n, $|a_n| \leq C(1+n)^p$.

Remark 3.3.16 One can compare the convergence in $\mathscr{S}'$ to more standard ones.

If $f_n \to f$ in $L^p(\mathbb{R}^d)$, then $f_n \to f$ in $\mathscr{S}'(\mathbb{R}^d)$.

If $T_n \to T$ in $\mathscr{S}'(\mathbb{R}^d)$ and $f \in \mathcal{O}_M(\mathbb{R}^d)$, then $fT_n \to fT$ in $\mathscr{S}'(\mathbb{R}^d)$.

3.4 The Fourier Transformation in $\mathscr{S}'(\mathbb{R}^d)$

3.4.1 Definition

Let $T \in \mathscr{S}'(\mathbb{R}^d)$ be a tempered distribution. We know that for any $\varphi \in \mathscr{S}$, its Fourier transform $\hat{\varphi}$ is also in $\mathscr{S}$, so the expression $\langle T, \hat{\varphi}$ is well-defined. The map $\varphi \mapsto \langle T, \hat{\varphi}\rangle$ is obviously a linear form on $\mathscr{S}(\mathbb{R}^d)$, let us check that it is continuous, hence a tempered distribution.

There exist $C > 0$ and $p \in \mathbb{N}$ such that

$$|\langle T, \hat{\varphi}\rangle| \leq C\, N_p(\hat{\varphi}) \leq C'\, N_{p+d+1}(\varphi)\,,$$

thanks to Proposition 3.2.6.

We also remember the change-of-head formula for $\varphi, \psi \in \mathscr{S}$, which can be rephrased as follows, remembering that T_ψ is a tempered distribution:

$$\langle T_\psi, \hat{\varphi}\rangle = \langle T_{\hat{\psi}}, \varphi\rangle\,.$$

This expression suggests to define the Fourier transform of a tempered distribution as follows

Definition 3.4.1 For $T \in \mathscr{S}'(\mathbb{R}^d)$, we denote by $\hat{T} = \mathcal{F}(T)$ the tempered distribution given by

$$\langle \hat{T}, \varphi\rangle := \langle T, \hat{\varphi}\rangle, \qquad \forall \varphi \in \mathscr{S}(\mathbb{R}^d).$$

Examples 3.4.2 *i)* For $f \in L^1$, we have by Fubini's theorem,

$$\begin{aligned}\langle \hat{f}, \varphi\rangle &= \int_{\mathbb{R}^d} f(x)\hat{\varphi}(x)\, dx = \int_{\mathbb{R}^d} f(x) \left(\int_{\mathbb{R}^d} \varphi(\xi)\, e^{-ix.\xi}\, d\xi\right) dx \\ &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(x)\, e^{-ix.\xi}\, dx\right) \varphi(\xi)\, d\xi = \int_{\mathbb{R}^d} \hat{f}(\xi)\varphi(\xi) d\xi,\end{aligned}$$

so that $\widehat{T_f} = T_{\hat{f}}$. Hence in this case we recover the classical Fourier transformation on L^1 (see also Proposition 3.4.6 below).

ii) For $\varphi \in \mathscr{S}(\mathbb{R}^d)$, $\langle \widehat{\delta_0}, \varphi \rangle = \int \varphi(x)dx$, which shows that $\widehat{\delta_0} = 1$ (constant function).

Proposition 3.4.3 The Fourier Transform $\mathcal{F}$ is an isomorphism on $\mathscr{S}'(\mathbb{R}^d)$. Its inverse is $\mathcal{F}^{-1} = (2\pi)^{-d}\check{\mathcal{F}}$. Moreover $\mathcal{F}$ and $\mathcal{F}^{-1}$ are continuous on $\mathscr{S}'(\mathbb{R}^d)$, in the following sense: if $T_n \to T \in \mathscr{S}'(\mathbb{R}^d)$, then $\mathcal{F}(T_n) \to \mathcal{F}(T)$ in $\mathscr{S}'(\mathbb{R}^d)$.

These results follow immediately form the above definition and Proposition 3.2.6. We may also easily transfer to $\mathscr{S}'(\mathbb{R}^d)$ the properties of the Fourier transform on $\mathscr{S}(\mathbb{R}^d)$ of Proposition 3.2.2, thereby obtaining the following identities.

Corollary 3.4.4 For any distribution $T \in \mathscr{S}'(\mathbb{R}^d)$, its Fourier transform intertwines as follows with differentiation, multiplication, translation:

$$\forall j = 1, \ldots, d, \qquad \mathcal{F}(D_j T) = \xi_j \hat{T}, \quad \mathcal{F}(x_j T) = -D_j \hat{T},$$
$$\forall x_0,\, \xi_0 \in \mathbb{R}^d, \quad \mathcal{F}(\tau_{x_0} T) = e^{-ix_0\cdot\xi}\, \mathcal{F}(T), \quad \mathcal{F}\big(e^{i\xi_0\cdot x} T\big) = \tau_{\xi_0}\mathcal{F}(T)\ .$$

Example 3.4.5 $\mathcal{F}(1) = \mathcal{F} \circ \mathcal{F}(\delta_0) = (2\pi)^d\, \check{\delta_0} = (2\pi)^d\, \delta_0$.

3.4.2 The Fourier Transformation on $L^1(\mathbb{R}^d)$ and $L^2(\mathbb{R}^d)$

Here we briefly sum up the main properties of the Fourier transform of a tempered distribution associated with an L^1 or an L^2 function.

Proposition 3.4.6 If $T = f \in L^1(\mathbb{R}^d)$, then $\mathcal{F}(T) = \hat{f}$. More precisely

i) $\mathcal{F}(T)$ is the continuous function given by $\mathcal{F}(T)(\xi) = \int e^{-ix\cdot\xi} f(x)dx$, and $\mathcal{F}(T)(\xi) \to 0$ when $|\xi| \to +\infty$.

ii) If moreover $\mathcal{F}(T)$ belongs to $L^1(\mathbb{R}^d)$, then $\mathcal{F}^{-1}(\mathcal{F}(T)) = T$ almost everywhere.

Proof.— The fact that $\hat{f}$ is a continuous function follows easily from the Lebesgue theorem of continuity of an integral w.r.to parameters, and the fact that $\hat{f}$ decays at infinity is called the Riemann-Lebesgue lemma. For $\varphi \in \mathscr{S}(\mathbb{R}^d)$, we have

$$\langle \hat{T}, \varphi \rangle = \langle T, \hat{\varphi} \rangle = \int f(\xi)\hat{\varphi}(\xi)d\xi = \int f(\xi)\Big(\int e^{-ix\cdot\xi}\varphi(x)dx\Big)d\xi.$$

Since the function $(x,\xi) \mapsto f(\xi)\, e^{-ix\cdot\xi}\varphi(x)$ belongs to $L^1(\mathbb{R}^d_x \times \mathbb{R}^d_\xi)$, we have by Fubini

$$\langle \hat{T}, \varphi \rangle = \int \varphi(x)\Big(\int e^{-ix\cdot\xi} f(\xi)d\xi\Big)dx = \langle \hat{f}, \varphi \rangle,$$

and this ends the proof of i). We also know that $\mathcal{F}^{-1}(\hat{f}) = f$ in $\mathscr{S}'(\mathbb{R}^d)$. If $\hat{f} \in L^1(\mathbb{R}^d)$, we thus have $\mathcal{F}^{-1}(\hat{f}) = f$ in $\mathscr{D}'(\mathbb{R}^d)$, and this equality of two distributions which are both in L^1 implies that $\mathcal{F}^{-1}(\hat{f}) = f$ almost everywhere. □

Let us now focus on distributions associated with functions $f \in L^2(\mathbb{R}^d)$.

Proposition 3.4.7 The map $T \in \mathscr{S}'(\mathbb{R}^d) \mapsto (2\pi)^{-d/2}\mathcal{F}(T) \in \mathscr{S}'(\mathbb{R}^d)$ induces a bijective isometry on the subspace $L^2(\mathbb{R}^d) \subset \mathscr{S}'(\mathbb{R}^d)$.

Proof.— Let $f \in L^2(\mathbb{R}^d)$. For every $\varphi \in \mathscr{S}(\mathbb{R}^d)$, we have by definition

$$\langle \hat{T}_f, \varphi \rangle = \int_{\mathbb{R}^d} f(x)\hat{\varphi}(x)\, dx\,.$$

By the Cauchy–Schwarz inequality and the Plancherel identity (Corollary 3.2.10),

$$|\langle \hat{T}_f, \varphi \rangle| \leq \|f\|_{L^2}\, \|\hat{\varphi}\|_{L^2} = (2\pi)^{\frac{d}{2}} \|f\|_{L^2}\, \|\varphi\|_{L^2}\,.$$

This shows that the distribution $\hat{T}_f$ can be extended into a continuous linear form on $L^2(\mathbb{R}^d)$. The Riesz representation theorem then shows that this linear form is represented by a single element of $L^2(\mathbb{R}^d)$, which we denote by $\hat{f} \in L^2(\mathbb{R}^d)$. We then have the identification $\hat{T}_f = T_{\hat{f}}$. Besides, the above inequality shows that

$$\|\hat{f}\|_{L^2} \leq (2\pi)^{\frac{d}{2}} \|f\|_{L^2}\,.$$

Applying this inequality to $\hat{f}$ in place of f and using the inversion formula from Proposition 3.4.3, we infer

$$(2\pi)^d \|f\|_{L^2} = \|\hat{\hat{f}}\|_{L^2} \leq (2\pi)^{\frac{d}{2}} \|\hat{f}\|_{L^2}$$

so finally,

$$\|f\|_{L^2} = (2\pi)^{-\frac{d}{2}} \|\hat{f}\|_{L^2}\,.$$

This shows that $(2\pi)^{-d/2}$ acts as an isometry $L^2 \to L^2$. Since it is clear from Proposition 3.4.3 that $\mathcal{F}$ is bijective on $\mathscr{S}'$, this isometry is actually a unitary bijection. □

Remark 3.4.8 There are functions f in $L^2(\mathbb{R}^d)$ such that $x \mapsto e^{-ix\cdot\xi} f(x)$ is not integrable whatever the value of ξ is (for example, for $d = 1$, $f(x) = (1 + |x|)^{-3/4}$), so one cannot compute $\hat{f}(\xi)$ by a direct computation of the integral $\int f(x)\, e^{-ix\cdot\xi}\, dx$.

However, we know that the truncated functions $f\mathbb{1}_{B(0,R)}$ converge to f in the $L^2(\mathbb{R}^d)$ sense, when $R \to \infty$. The functions $f\mathbb{1}_{B(0,R)}$ are in $L^1(\mathbb{R}^d)$, so their Fourier transforms are well-defined classically, given by

$$g_R(\xi) = \int_{|x|<R} e^{-ix\cdot\xi} f(x)\, dx, \quad \xi \in \mathbb{R}^d\,,$$

Thanks to Proposition 3.4.7, the L^2 convergence of $f\mathbb{1}_{B(0,R)}$ implies the convergence of the Fourier transforms:

$$g_R \overset{L^2}{\to} \hat{f} \quad \text{when } R \to \infty\,.$$

3.4.3 The Fourier Transform of compactly supported distributions

The Fourier transformation exchanges the speed of decay at infinity of a function with the regularity of its image, as shown for example by the following inequality

$$\|D^\alpha \hat{\varphi}\|_{L^\infty} = \|\mathcal{F}(x^\alpha \varphi)\|_{L^\infty} \leq \|x^\alpha \varphi\|_{L^1}, \quad \varphi \in \mathscr{S}(\mathbb{R}^d)\,.$$

The strongest speed of decay at infinity is achieved for compactly supported function ; the result of this subsection explores this phenomenon more precisely.

Proposition 3.4.9 Let $T \in \mathscr{D}'(\mathbb{R}^d)$ be compactly supported, and take $\chi \in \mathscr{D}(\mathbb{R}^d)$ a cutoff function above $\operatorname{supp} T$. Then the Fourier transform $\mathcal{F}(T)$ is the smooth function on $\mathbb{R}^d$ given by

$$\forall \xi \in \mathbb{R}^d, \quad \mathcal{F}(T)(\xi) = \langle T, \chi\, e^{-ix\cdot\xi}\rangle.$$

Moreover, there is an integer $m \in \mathbb{N}$ such that, for any $\alpha \in \mathbb{N}^d$, there is a $C_\alpha > 0$ satisfying

$$|\partial^\alpha \mathcal{F}(T)(\xi)| \leq C_\alpha (1 + |\xi|)^m.$$

In particular, $\mathcal{F}(T)$ is a function of moderate growth.

Proof.— Differentiating w.r.to ξ under the bracket, we see that the function $v(\xi)$ given by

$$v(\xi) := \langle T, \chi e^{-ix\cdot\xi}\rangle$$

is a $\mathcal{C}^\infty$ function. We also have

$$\partial^\alpha v(\xi) = \langle T, \chi\, (-ix)^\alpha\, e^{-ix\cdot\xi}\rangle,$$

All those test function are supported on $K = \operatorname{supp} \chi$, so there are $C > 0$, $m \in \mathbb{N}$ such that

$$\begin{aligned} |\partial^\alpha v(\xi)| &\leq C \sum_{|\beta|\leq m} \|\partial_x^\beta \big(\chi\, (-ix)^\alpha\, e^{-ix\cdot\xi}\big)\|_\infty \\ &\leq C \sum_{|\beta|\leq m} C_{\chi,\beta} \sup_{x\in K} \big|\partial_x^\beta \big((-ix)^\alpha\, e^{-ix\cdot\xi}\big)\big| \leq C_{\alpha,K} (1 + |\xi|)^m\,. \end{aligned}$$

Let us finally check that this function v represents $\hat{T}$. Indeed, for $\varphi \in \mathscr{D}(\mathbb{R}^d)$, thanks to the theorem of integration under the bracket, we find

$$\langle \hat{T}, \varphi\rangle = \Big\langle T, \chi \int e^{-ix\cdot\xi} \varphi(\xi) d\xi \Big\rangle = \int \langle T, \chi\, e^{-ix\cdot\xi}\rangle\, \varphi(\xi) d\xi,$$

hence $\hat{T}$ is given by the function $\hat{T}(\xi) = \langle T, \chi\, e^{-ix\cdot\xi}\rangle = v(\xi)$. □

3.4.4 Fourier characterization of Sobolev spaces on $\mathbb{R}^d$

We recall that Sobolev spaces $H^n(\Omega)$ were introduced in Chapter 2 for $n \in \mathbb{N}$ and Ω an open subset of $\mathbb{R}^d$. In this subsection, we use the Fourier transform to provide more information about these spaces when $\Omega = \mathbb{R}^d$.

Notation

For $\xi \in \mathbb{R}^d$, we define the "Japanese bracket notation"

$$\langle \xi \rangle := \sqrt{1 + |\xi|^2}\,, \qquad , \xi \in \mathbb{R}^d\,.$$

The function $\xi \mapsto \langle \xi \rangle$ is smooth on $\mathbb{R}^d$, and there is a constant $C > 0$ such that, for $|\xi| \geq 1$,

$$\frac{1}{C}|\xi| \leq \langle \xi \rangle \leq C|\xi|.$$

Thus, $\langle \xi \rangle$ is a regularized version of $|\xi|$, in the sense that it has the same behavior at infinity. Furthermore, since this function is smooth and since all its derivatives have growth at most polynomial at infinity, it is of moderate growth, so it acts on $\mathscr{S}'(\mathbb{R}^d)$ by multiplication.

The following proposition is a characterization of the Sobolev spaces $H^n(\mathbb{R}^d)$ using the Fourier transform.

Proposition 3.4.10 Let $n \in \mathbb{N}$ and $u \in \mathscr{S}'(\mathbb{R}^d)$. Then $u \in H^n(\mathbb{R}^d)$ if and only if the distribution $\langle \xi \rangle^n \hat{u} \in$ belongs to $L^2(\mathbb{R}^d)$. Furthermore, there exists $C_n > 0$ such that

$$\forall u \in H^n(\mathbb{R}^d)\ , \qquad C_s^{-1} \|\langle \xi \rangle^n \hat{u}\|_{L^2} \leq \|u\|_{H^n} \leq C_n \|\langle \xi \rangle^n \hat{u}\|_{L^2}\ .$$

In other words, $\|\langle \xi \rangle^n \hat{u}\|_{L^2}$ defines an norm on $H^n(\mathbb{R}^d)$, equivalent with the usual norm.

The characterization can be stated as the fact that the distribution $\hat{u}$ belongs to the weighted L^2 space $L^2(\mathbb{R}^d, \langle \xi \rangle^{2n}\, d\xi)$.

Proof.— If $u \in H^n(\mathbb{R}^d)$, then for all $|\alpha| \leq n$ we have $\partial^\alpha u \in L^2(\mathbb{R}^d)$. By the Plancherel theorem, its Fourier transform $\xi^\alpha \hat{u}$ is also in $L^2(\mathbb{R}^d)$, with the same norm (up to a constant $(2\pi)^{d/2}$).

Now observe the polynomial expansion

$$\langle \xi \rangle^{2n} = (1 + \xi_1^2 + \cdots + \xi_d^2)^n = \sum_{|\alpha| \leq s} c_{\alpha,n} \xi^{2\alpha}$$

for some positive constants $c_{\alpha,n}$. Consequently,

$$\langle \xi \rangle^{2n} \, |\hat{u}(\xi)|^2 = \sum_{|\alpha| \leq n} c_{\alpha,n} \xi^{2\alpha} \, |\hat{u}(\xi)|^2 \quad \text{is integrable,}$$

whence $\langle\xi\rangle^n \hat{u} \in L^2(\mathbb{R}^d)$, with

$$\|\langle\xi\rangle^n \hat{u}\|_{L^2} \leq C \sum_{|\alpha|\leq n} \|\xi^\alpha\, \hat{u}\|_{L^2}$$

Conversely, if $\langle\xi\rangle^n \hat{u} \in L^2(\mathbb{R}^d)$, then, for any multiindex such that $|\alpha| \leq n$, the function $\xi^\alpha \langle\xi\rangle^{-n}$ is bounded on $\mathbb{R}^d$, and we have

$$\xi^\alpha \hat{u} = \xi^\alpha \langle\xi\rangle^{-n}\, \langle\xi\rangle^n\, \hat{u} \in L^2(\mathbb{R}^d), \quad \text{with} \quad \|\xi^\alpha \hat{u}\|_{L^2} \leq C_\alpha\, \|\langle\xi\rangle^n\, \hat{u}\|_{L^2}\,.$$

which precisely means that $\partial^\alpha u \in L^2(\mathbb{R}^d)$. □

Proposition 3.4.10 suggests to generalize the definition of $H^n(\mathbb{R}^d)$ to every *real* index $s \in \mathbb{R}$ as follows.

Definition 3.4.11 Let $s \in \mathbb{R}$. A tempered distribution $u \in \mathscr{S}'(\mathbb{R}^d)$ belongs to the Sobolev space $H^s(\mathbb{R}^d)$ iff $\langle\xi\rangle^s \hat{u} \in L^2(\mathbb{R}^d)$. Equivalently, iff $\hat{u} \in L^2(\mathbb{R}^d, \langle\xi\rangle^{2s}\, d\xi)$.

The natural scalar product on $L^2(\mathbb{R}^d, \langle\xi\rangle^{2s}\, d\xi)$ defines a Hermitian scalar product on $H^s(\mathbb{R}^d)$:

$$(u, v)_{H^s} := \int_{\mathbb{R}^d} \hat{u}(\xi)\, \overline{\hat{v}(\xi)}\, \langle\xi\rangle^{2s}\, d\xi\,,$$

which makes H^s a Hilbert space, with norm denoted by $\|\cdot\|_{H^s}$.

We notice that the H^s inject in one another as Hilbert spaces: if $s' > s$ then $H^{s'} \subset H^s$ and the definition of the norms directly shows that the injection $H^{s'} \to H^s$ is continuous:

$$\forall u \in H^{s'}, \quad \|u\|_{H^s} \leq \|u\|_{H^{s'}}\,.$$

Examples 3.4.12 *i)* $\delta_0 \in H^s(\mathbb{R}^d)$ if and only if $s < \frac{-d}{2}$. Indeed $\hat{\delta}_0 = 1$, so that $\langle\xi\rangle^s \hat{\delta}_0 \in L^2(\mathbb{R}^d)$ if and only if $2s < -d$.

ii) The constant function 1 does not belong to any $H^s(\mathbb{R}^d)$, since $\hat{1} = (2\pi)^d \delta_0$ is not in L^1_{loc}, hence not in any $L^2(\langle\xi\rangle^{2s}\, d\xi)$ (any of these weighted L^2 spaces is contained in L^1_{loc}).

The above example shows that

$$\bigcup_{s\in\mathbb{R}} H^s(\mathbb{R}^d) \subsetneq \mathscr{S}'(\mathbb{R}^d)\,.$$

Below follows another illustration of the fact that H^s contains elements that are more and more singular as s decreases.

Proposition 3.4.13 Let $T \in \mathcal{E}'(\mathbb{R}^d)$ be a compactly supported distribution, with order $m \geq 0$. Then $T \in H^s(\mathbb{R}^d)$for any $s < -m - \frac{d}{2}$.

In particular, we have the inclusion

$$\mathcal{E}'(\mathbb{R}^d) \subsetneq \bigcup_{s\in\mathbb{R}} H^s(\mathbb{R}^d)\,.$$

Proof.— For $T \in \mathcal{E}'(\mathbb{R}^d)$, we already know that $\hat{T} \in \mathcal{C}^\infty$. Moreover,

$$|\langle\xi\rangle^s \hat{T}(\xi)| = |\langle\xi\rangle^s \langle T, \chi\, e^{-ix\cdot\xi}\rangle| \leq C\langle\xi\rangle^s \sum_{|\alpha|\leq m} \sup|\partial_x^\alpha(e^{-ix\cdot\xi})| \leq C\langle\xi\rangle^{s+m}$$

Thus $T \in H^s(\mathbb{R}^d)$ when $\langle\xi\rangle^{s+m} \in L^2$, thus when $s+m < -d/2$. □

We close this subsection by proving an important result in functional analysis, namely a connection between the Sobolev regularity (which is, in some sense, a regularity "on average") and the pointwise ($\mathcal{C}^k$) regularity.

We denote by $\mathcal{C}^k_{\to 0}(\mathbb{R}^d)$ the space of $\mathcal{C}^k$ functions which decay to 0 at infinity, as well as all their derivatives of order $\leq k$.

> **Proposition 3.4.14 (Sobolev embedding 1)** If $s > \frac{d}{2}+k$, then every element of $H^s(\mathbb{R}^d)$ belongs to $\mathcal{C}^k_{\to 0}(\mathbb{R}^d)$, and the embedding $H^s(\mathbb{R}^d) \to \mathcal{C}^k_{\to 0}(\mathbb{R}^d)$ is continuous.

Proof.— Let $u \in H^s(\mathbb{R}^d)$. For $\alpha \in \mathbb{N}^d$ such that $|\alpha| \leq k$, we have $\xi^\alpha \hat{u} \in L^1$. Indeed,

$$|\xi^\alpha \hat{u}(\xi)| \leq \frac{|\xi|^{|\alpha|}}{\langle\xi\rangle^s}\langle\xi\rangle^s|\hat{u}(\xi)| \leq \langle\xi\rangle^{k-s}\langle\xi\rangle^s|\hat{u}(\xi)|,$$

and $\langle\xi\rangle^{k-s} \in L^2(\mathbb{R}^d)$ since $k-s < -d/2$. By the Cauchy-Schwarz inequality, we thus get

$$\|\xi^\alpha \hat{u}\|_{L^1} \leq C_{s,d}\|u\|_{H^s}. \tag{3.4.7}$$

Therefore $\partial^\alpha u = \mathcal{F}^{-1}((i\xi)^\alpha \hat{u}) \in \mathcal{C}^0_{\to 0}$ by Proposition 3.4.6. The fact that the injection from $H^s(\mathbb{R}^d)$ to $\mathcal{C}^k_{\to 0}(\mathbb{R}^d)$ is continuous is just a way to read the inequalities

$$\forall|\alpha| \leq k, \quad \|\partial^\alpha u\|_{L^\infty} \leq \|\xi^\alpha \hat{u}\|_{L^1} \leq C_{s,d}\|u\|_{H^s}.$$

□

Another interesting consequence of Proposition 3.4.10 is that we recover solutions to the damped Poisson equation $-\Delta u + u = f$ on $\mathbb{R}^d$ (see Theorem 2.5.4) in a very general setup, with optimal regularity of the solution.

> **Proposition 3.4.15** For every $s \in \mathbb{R}^d$, the mapping $u \in H^s(\mathbb{R}^d) \mapsto u - \Delta u \in H^{s-2}(\mathbb{R}^d)$ is an isomorphism.

Proof.— This mapping is conjugated, through Fourier transform, to the mapping

$$\hat{u} \in L^2(\langle\xi\rangle^{2s} d\xi) \mapsto \langle\xi\rangle^2 \hat{u} \in L^2(\langle\xi\rangle^{2s-4} d\xi)$$

which is trivially an isomorphism (actually, a unitary one). □

Combining the above three propositions, we obtain he following extension of Theorem 1.9.16 to several dimensions.

Theorem 3.4.16 For every distribution T on $\mathbb{R}^d$ with compact support, there exist $p \in \mathbb{N}$ and a family $(f_\alpha)_{|\alpha| \leq p}$ of continuous functions on $\mathbb{R}^d$ such that

$$T = \sum_{|\alpha| \leq p} \partial^\alpha (f_\alpha) \ .$$

Proof.— From Proposition 3.4.13, we know that, for some $s \in \mathbb{R}$, $T \in H^s(\mathbb{R}^d)$ Let q be a positive integer such that

$$s + 2q > \frac{d}{2} \ .$$

By the isomorphism of Proposition 3.4.15 iterated q times, there exists $g \in H^{s+2q}$ unique such that

$$(I - \Delta)^q g = T \ .$$

From the Sobolev embedding of Proposition 3.4.14, the distribution g is a continuous function. It remains to write

$$T = (I - \Delta)^q g = \big(1 - \sum_{j=1}^{d} \partial_j^2\big)^q g = \sum_{|\beta| \leq q} c_\beta \partial^{2\beta} g \,,$$

and the proof is complete by identifying the even multiindices $\alpha = 2\beta$, $|\alpha| \leq 2q$, and functions $f_\alpha = c_\beta g$. □

3.4.5 Local Sobolev spaces on Ω

In this subsection we come back to the general setup of distributions on an arbitrary open set $\Omega \subset \mathbb{R}^d$ of Chapter 2. Using the Sobolev spaces $H^s(\mathbb{R}^d)$, we define subspaces of $\mathscr{D}'(\Omega)$, called *local* Sobolev spaces on Ω. Those spaces will be very useful when studying the local regularity of PDEs on Ω.

Definition 3.4.17 Let Ω be an open subset of $\mathbb{R}^d$, and let $s \in \mathbb{R}$. We denote by $H^s_{\mathrm{loc}}(\Omega)$ the space of distributions $T \in \mathscr{D}'(\Omega)$ such that, for any $\chi \in \mathscr{D}(\Omega)$, the extension to $\mathbb{R}^d$ by 0 of the compactly supported distribution χu belongs to $H^s(\mathbb{R}^d)$.

We first show the following important regularity theorem.

Proposition 3.4.18

$$\bigcap_{s\in\mathbb{R}} H^s_{\mathrm{loc}}(\Omega) = \mathcal{C}^\infty(\Omega)\ .$$

This “local” result may be compared with the strict inclusions

$$\mathscr{S}(\mathbb{R}^d) \subsetneq \bigcap_{s\in\mathbb{R}} H^s(\mathbb{R}^d) \overset{\mathcal{F}}{\Longleftrightarrow} \mathscr{S}(\mathbb{R}^d) \subsetneq \bigcap_{s\in\mathbb{R}} L^2(\mathbb{R}^d, \langle\xi\rangle^{2s}\, d\xi)\,.$$

Proof.— If $u \in \mathcal{C}^\infty(\Omega)$, then, for any test function$\chi \in \mathscr{D}(\Omega)$, the product $\chi u \in \mathscr{D}(\Omega)$, so that its extension to $\mathbb{R}^d$ by 0 belongs to $\mathscr{S}(\mathbb{R}^d)$, and admits a Fourier transform $\widehat{\chi u} \in \mathscr{S}(\mathbb{R}^d)$; consequently $\chi u \in H^s$ for every $s \in \mathbb{R}$.

Conversely, if $u \in H^s_{\mathrm{loc}}(\Omega)$ for every $s \in \mathbb{R}$, then for any $\chi \in \mathscr{D}(\Omega)$ we have by definition $\chi u \in H^s(\mathbb{R}^d)$, for every s. Proposition 3.4.14 implies that $\chi u \in \mathcal{C}^k_{\to 0}(\mathbb{R}^d)$ for every k, in particular $\chi u \in C^\infty(\mathbb{R}^d)$. Since this holds for any $\chi \in \mathscr{D}(\Omega)$, we obtain $u \in \mathcal{C}^\infty$. □

Like all “local” results in $\mathscr{D}'(\Omega)$, the inclusion $u \in H^s_{\mathrm{loc}}(\Omega)$ does control how fast the “local” H^s norms $\|\chi u\|_{H^s}$ may grow when $\operatorname{supp}\chi$ approaches the boundary of Ω.

3.5 Some applications to PDEs

In this section we use the Fourier transform as a very convenient tool to study the solutions of Partial Differential equations on $\mathbb{R}^d$, or on $\Omega \subset \mathbb{R}^d$, especially those with constant coefficients.

3.5.1 Partial differential equations with constant coefficients

Let $p \in \mathbb{C}[X_1, \ldots, X_d]$ be a polynomial of d variables with complex coefficients,

$$p(X) = \sum_{|\alpha|\leq m} a_\alpha X^\alpha, \quad X \in \mathbb{R}^d, \quad \text{with } a_\alpha \in \mathbb{C}\,.$$

We assume that, for at least one multiindex $|\alpha| = m$, the coefficient $a_\alpha \neq 0$: p is then a polynomial of degree m.

Given an open subset $\Omega \subset \mathbb{R}^d$, we denote by $p(\partial)$ the operator on $\mathscr{D}'(\Omega)$ given by

$$\mathscr{D}'(\Omega) \ni T \mapsto p(\partial)T = \sum_{|\alpha|\leq m} a_\alpha \partial^\alpha T \in \mathscr{D}'(\Omega).$$

Those operators are called linear partial differential operators with constant coefficients. The equation

$$p(\partial)u = f,$$

where $f \in \mathscr{D}'(\Omega)$ is given, and $u \in \mathscr{D}'(\Omega)$ is the unknown, is called a Partial Differential Equation (PDE) of order m with constant coefficients. When $f \neq 0$, it is said to be *inhomogeneous*, or with *source term* f.

Remark 3.5.1 For $d = 1$, the equation $p(\partial)u = f$ is a linear differential equation of order m with constant coefficients, which can be explicitly solved by transforming it into a linear system of differential equations of order 1, of the form $U' = MU + F$ with U a vector valued distribution, M a matrix of constant coefficients, F a vector valued inhomogeneous term.

On the opposite, in $d > 1$ the situation is drastically different: in general we have no explicit, or even approximate expression for the solutions u, and it may even be difficult to show the existence of solutions. The theory of PDEs tries to give as precise informations as possible of the solutions, without giving explicit formulae: their existence in some space of distributions, their regularity, sometimes a more precise structure.

Example 3.5.2 If $p(X_1, \dots, X_d) = X_1^2 + \cdots + X_d^2$, then

$$p(\partial)T = \partial_1^2 T + \cdots + \partial_d^2 T = \Delta T \ .$$

We already encountered operator Δ which is called the Laplacian (or Laplacean), and equation the equation $\Delta u = f$ is called the Laplace (or Poisson) equation.

A particularity of Δ is that this operator is covariant under the rotations of $\mathbb{R}^d$ centered at the origin: for any rotation matrix $R \in O(d, \mathbb{R})$ and $f \in C^2(\mathbb{R}^d)$, one has

$$\forall x \in \mathbb{R}^d, \qquad \Delta_x \big[f(Rx)\big] = [\Delta f](Rx) \, .$$

Moreover, one can show that every linear partial differential operator with constant coefficients on $\mathbb{R}^d$ which is invariant by rotations, is of the form $P = q(\Delta)$, where q is a polynomial of one variable. This invariance property explains why the Laplacian appears so frequently in many areas of Mathematical Physics: to name a few, in the heat equation, wave equation, Schrödinger equation...

The following proposition, which is at the heart of the use of $\mathcal{F}$ in PDEs, is a direct consequence of the properties of the Fourier transform on $\mathscr{S}'(\mathbb{R}^d)$ given in Corollary 3.4.4.

Proposition 3.5.3 Let $p \in \mathbb{C}[X_1, \dots, X_d]$ and $P = p(\partial)$ be the corresponding differential operator. For every $u \in \mathscr{S}'(\mathbb{R}^d)$, we have

$$\widehat{Pu} = p(i\xi)\hat{u} \ .$$

Therefore, the action of $\mathcal{F}$ transforms a linear differential operator with constant coefficients, into the operator of multiplication by a polynomial, a priori a much simpler object to study. The importance of the above proposition appears through the multiplicity of its consequences. We shall draw some of them in the next subsections.

3.5.2 Local results on Ω

Let $P = p(\partial)$ be a linear partial differential operators with constant coefficients. In this section, we study some properties of P on $\mathscr{D}'(\Omega)$, where Ω is an open subset of $\mathbb{R}^d$. We start with a very natural statement.

> **Proposition 3.5.4** If P is a linear partial differential operators with constant coefficients of order $\leq m$. Let $s \in \mathbb{R}$. Then $P : H^s_{\mathsf{loc}}(\Omega) \to H^{s-m}_{\mathsf{loc}}(\Omega)$. That is, for any $u \in H^s_{\mathsf{loc}}(\Omega)$, Pu belongs to $H^{s-m}_{\mathsf{loc}}(\Omega)$.

Proof.— The result $Pu \in H^{s-m}_{\mathsf{loc}}(\Omega)$ means that, or any $\varphi \in \mathscr{D}(\Omega)$, φPu belongs to $H^{s-m}(\mathbb{R}^d)$. On the other hand, the assumption involves the functions of the form ψu, $\psi \in \mathscr{D}(\Omega)$. It is thus necessary to connect these two types of function. The following Lemma is thus relevant.

> **Lemma 3.5.5** Let P be a linear partial differential operators with constant coefficients of order $\leq m$, and $\varphi \in \mathscr{D}(\Omega)$. Then there exist functions $\big(\varphi_\beta \in \mathscr{D}(\Omega)\big)_{|\beta|\leq m-1}$, such that, for every $u \in \mathscr{D}'(\Omega)$,
> $$P(\varphi u) = \varphi P u + \sum_{|\beta|\leq m-1} \partial^\beta(\varphi_\beta u)\ .$$

Let us use this Lemma. Take $u \in H^s_{\mathsf{loc}}(\Omega)$ and choose some $\varphi \in \mathscr{D}(\Omega)$. Apply the Lemma: the functions φu and $\varphi_\beta u$ belong to $H^s(\mathbb{R}^d)$ by the assumption on u. From the orders of the differential operators P and ∂^β, we get

$$P(\varphi u) \in H^{s-m}(\mathbb{R}^d)\ , \quad \partial^\beta(\varphi_\beta u) \in H^{s-m+1}(\mathbb{R}^d)\ .$$

This leads to $\varphi P u \in H^{s-m}(\mathbb{R}^d)$, which means that $Pu \in H^{s-m}_{\mathsf{loc}}(\mathbb{R}^d)$. □

Let us now prove the Lemma 3.5.5. By linearity, it is enough to prove it for $P = \partial^\alpha$. We proceed by induction on $|\alpha|$. If $|\alpha| = 0$, the formula trivially holds.

Assume that the statement holds for multiindices $|\alpha| \leq m-1$, and let us prove it at the rank $|\alpha| = m$. By the Leibniz formula,

$$\partial^\alpha(\varphi u) = \varphi \partial^\alpha u + \sum_{\beta\leq\alpha, |\beta|\leq m-1} \binom{\alpha}{\beta} \partial^{\alpha-\beta}\varphi\, \partial^\beta u\ . \tag{3.5.8}$$

Each term in the last sum is of the form $\varphi_{\alpha-\beta}\, \partial^\beta u$, with $|\beta| \leq m-1$ and $\varphi_{\alpha-\beta} \in \mathscr{D}(\Omega)$. We may then apply the induction hypothesis, to rewrite this term as

$$\varphi_{\alpha-\beta}\, \partial^\beta u = \partial^\beta(\varphi_{\alpha-\beta}\, u) - \sum_{\gamma\leq\beta, |\gamma|<|\beta|} \binom{\beta}{\gamma} \partial^{\beta-\gamma}\varphi_{\alpha-\beta}\partial^\gamma u\,,$$

so altogether $\varphi_{\alpha-\beta}\, \partial^\beta u$ is a sum of terms of the type $\partial^\gamma(\psi_{\alpha-\beta,\gamma} u)$ with $|\gamma| \leq m-1$ and $\psi_{\gamma\alpha-\beta,\gamma} \in \mathscr{D}(\Omega)$. Summing all these terms in (3.5.8), we get the statement of the Lemma. □

Elliptic regularity

We showed above that if $P = p(\partial)$ is a differential operator of order $\leq m$, then it maps $H^{s+m}_{\text{loc}}(\Omega)$ to $H^s_{\text{loc}}(\Omega)$. A classical problem in PDEs asks the reverse question: knowing the local regularity of f, and assuming that $u \in \mathscr{D}'(\Omega)$ solves $Pu = f$, what can we say about the local regularity of solutions u?

It is possible to give a general answer to this question under an assumption on the polynomial p, which we now define.

Definition 3.5.6 An operator $p(\partial) = \sum_{|\alpha|\leq m} a_\alpha \partial^\alpha$ is called *elliptic of order* m if the highest homogeneous part $p_m = \sum_{|\alpha|=m} a_\alpha X^\alpha$ of the polynomial p satisfies

$$\forall \xi \in \mathbb{R}^d \setminus \{0\}\,, \quad p_m(\xi) \neq 0\,,$$
$$\text{equivalently,} \quad \forall \xi \in S^{d-1}, \quad p_m(\xi) \neq 0\,.$$

For instance, in any dimension $d \geq 1$ the Laplacian Δ is elliptic, while the wave operator operator $\partial_1^2 - \partial_2^2$ on $\mathbb{R}^2$ is not.

Theorem 3.5.7 (Elliptic regularity) Let $\Omega \subset \mathbb{R}^d$ be an open subset, and let P be a linear partial differential operator with constant coefficients, *elliptic* of order m, acting on Ω. If $u \in \mathscr{D}'(\Omega)$ satisfies $Pu \in H^s_{\text{loc}}(\Omega)$ for some $s \in \mathbb{R}$, then $u \in H^{s+m}_{\text{loc}}(\Omega)$. In particular, if $Pu \in C^\infty(\Omega)$, then $u \in C^\infty(\Omega)$.

Proof.— 1. Let us first prove the result under the additional assumption that $u \in H^\sigma_{\text{loc}}(\Omega)$ for some $\sigma \in \mathbb{R}$. Since P is elliptic and since the unit sphere S^{d-1} is compact, there exists $c > 0$ such that,

$$\forall \xi \in S^{d-1}, \quad |p_m(i\xi)| \geq c\,.$$

By homogeneity of p_m, this implies

$$\forall \xi \in \mathbb{R}^d\,,\ |p_m(\xi)| \geq c\,|\xi|^m\,.$$

We are actually intested in $p(i\xi)$, which is the Fourier multiplier corresponding to the operator $p(\partial)$. Comparing $p(i\xi)$ with $p_m(i\xi)$, we find

$$p(i\xi) - p_m(i\xi) = \mathcal{O}(|\xi|^{m-1}) \quad \text{when } |\xi| \to \infty\,.$$

As a result, there exits $R > 0$ such that, for every $\xi \in \mathbb{R}^d$ with $|\xi| \geq R$,

$$|p(i\xi)| \geq \frac{c}{2}\,|\xi|^m\,. \qquad (3.5.9)$$

We see here one important aspect of the ellipticity assumption: it gives nontrivial information in the regime of high frequencies $|\xi| \gg 1$. Here, we find that the Fourier multiplier $p(i\xi)$ does not vanish for large frequencies.

2. Using this nonvanishing property, we will define a quasi-inverse of this multiplier. Let $\chi \in \mathscr{D}(\mathbb{R}^d)$ be a plateau function on the ball $B(0, R)$. Then the function

$$q(\xi) := \begin{cases} \frac{1-\chi(\xi)}{p(i\xi)}, & |x| \geq R, \\ 0, & |x| < R, \end{cases}$$

is well-defined and C^∞ on $\mathbb{R}^d$ and satisfies, due to estimate (3.5.9),

$$|q(\xi)| \leq C\langle\xi\rangle^{-m}) \,. \tag{3.5.10}$$

An easy computation shows that the same estimate holds for any derivative $\partial^\alpha q$, so that q is a function of moderate growth. The multiplication by function $q(\xi)$ thus acts on $\mathscr{S}'(\mathbb{R}^d)$. Viewing $q(\xi)$ as a Fourier multiplier amounts to defining the operator Q on $\mathscr{S}'$, such that $\widehat{Q(u)} = q(\xi)\hat{u}$. The decay (3.5.10) shows that for every $\tau \in \mathbb{R}$, this operator Q maps $H^\tau(\mathbb{R}^d)$ to $H^{\tau+m}(\mathbb{R}^d)$.

This operator Q is a quasi-inverse of P in the following sense:

$$QP = Id - R, \quad \text{where the operator } R \text{ is defined by } \widehat{R(u)} = \chi\,\hat{u}\,.$$

The operator R kills the high frequencies, it is called a regularizing operator. Indeed, for any $v \in H^\sigma(\mathbb{R}^d)$, $R(u)$ will belong to all the Sobolev spaces H^τ, $\tau \in \mathbb{R}$, in particular $R(u)$ will be a smooth function.

3. Now let us take $u \in H^\sigma_{\mathrm{loc}}(\Omega)$ such that $Pu \in H^s_{\mathrm{loc}}(\Omega)$. For every $\varphi \in \mathscr{D}(\Omega)$, Lemma 3.5.5 yields

$$P(\varphi u) = \varphi Pu + \sum_{|\alpha|\leq m-1} \partial^\alpha(\varphi_\alpha u)\,,$$

where $\varphi_\alpha \in \mathscr{D}(\Omega)$. Applying the operator Q on both sides, we find

$$QP(\varphi u) = Q(\varphi Pu) + \sum_{|\alpha|\leq m-1} Q(\partial^\alpha(\varphi_\alpha u))\,.$$

On the right hand side, the first term belongs to $H^{s+m}(\mathbb{R}^d)$ while the other terms belong to $H^{\sigma+1}(\mathbb{R}^d)$, so the full RHS belongs to $H^{\min(s+m,\sigma+1)}(\mathbb{R}^d)$. Turning to the left hand side, the decomposition

$$QP(\varphi u) = \varphi u - R(\varphi u)\,,$$

and the regularizing properties of R, show that φu belongs to the same space $H^{\min(s+m,\sigma+1)}(\mathbb{R}^d)$ as $QP(\varphi u)$.

We have thus proved that

$$u \in H^{\min(s+m,\sigma+1)}_{\mathrm{loc}}(\Omega)\,.$$

4. If $\sigma + 1 < s + m$, we iterate this result, taking into account the new information $u \in H^{\sigma+1}_{\mathrm{loc}}$: we may thus start from this assumption, and prove that $H^{\min(s+m,\sigma+2)}(\mathbb{R}^d)$. After a finite number of steps, we finally obtain the required result $u \in H^{s+m}_{\mathrm{loc}}(\Omega)$.

4. Now let assume $u \in \mathscr{D}'(\Omega)$ satisfying $Pu \in H^s_{\mathrm{loc}}(\Omega)$, without a priori knowledge that $u \in H^\sigma_{\mathrm{loc}}(\Omega)$. We will obtain such a property by restricting u in a bounded set of Ω. Indeed, let V be an open subset of Ω, so that $\overline{V} \subset \Omega$ is compact. If $\psi \in \mathscr{D}(\Omega)$ is a plateau function on $\overline{V}$, Proposition 3.4.13 implies that there exists $\sigma \in \mathbb{R}$ such that $\psi u \in H^\sigma(\mathbb{R}^d)$. The following Lemma 3.5.8 implies that, for any $\chi \in \mathscr{D}(V)$, the function $\chi u = \chi\psi u$ belongs to $H^\sigma(\mathbb{R}^d)$ as well. This exactly shows that the restriction $u_{|V} \in H^\sigma_{\mathrm{loc}}(V)$. Similarly, from $f \in H^s_{\mathrm{loc}}(\Omega)$ we draw $f_{|V} \in H^s_{\mathrm{loc}}(V)$. On the other hand, $u_{|V}$ is a solution of

$$P(u_{|V}) = (Pu)_{|V} = f_{|V} \in H^s_{\mathrm{loc}}(V)\,.$$

We may thus apply the above proof to this equation of V, and get that $u_{|V} \in H^{\sigma+m}_{\mathrm{loc}}(V)$.

Finally, since this conclusion holds for any bounded open subset $V \Subset \Omega$, it holds on all Ω. □

Lemma 3.5.8 If $\sigma \in \mathbb{R}$, $v \in H^\sigma(\mathbb{R}^d)$ and $\varphi \in \mathscr{D}(\mathbb{R}^d)$, then $\varphi v \in H^\sigma(\mathbb{R}^d)$.

Proof.— Proposition 3.4.9 shows that,

$$\widehat{\varphi v}(\xi) = \langle v, \varphi\, e^{-ix\cdot\xi}\rangle\,,$$

and this function of ξ is smooth. Call $\psi_\xi(x) = \varphi(x)\, e^{-ix\cdot\xi}$. Using the inversion formula, we compute

$$\langle v, \psi_\xi\rangle = \langle v, \mathcal{F}\circ\mathcal{F}^{-1}(\psi_\xi)\rangle = \langle \mathcal{F}(v), (2\pi)^{-d}\sigma\circ\mathcal{F}(\psi_\xi)\rangle\,.$$

The properties of the Fourier transform (see Corollary 3.4.4) show that $\mathcal{F}(\psi_\xi)(\eta) = \hat\varphi(\eta+\xi)$, and finally

$$\widehat{\varphi v}(\xi) = (2\pi)^{-d}\langle \hat v, \hat\varphi(\xi-\cdot)\rangle = (2\pi)^{-d}\int_{\mathbb{R}^d} \hat v(\eta)\,\hat\varphi(\xi-\eta)\,d\eta\,.$$

We want to show that $\langle\xi\rangle^\tau\widehat{\varphi v}(\xi)$ is square integrable on $\mathbb{R}^d$. Since this function is smooth, it is L^2_{loc}, so the convergence question only comes from $|\xi|\to\infty$. After a change of variables in the above integral, we have

$$\langle\xi\rangle^\sigma\widehat{\varphi v}(\xi) = (2\pi)^{-d}\int_{\mathbb{R}^d}\langle\xi\rangle^\sigma\,\hat v(\xi-\eta)\hat\varphi(\eta)\,d\eta\,.$$

The information we have is the square-integrability of $\langle\xi\rangle^\sigma\hat v(\xi)$. We decompose the integral into two parts, corresponding to relative size of η w.r.t. $|\xi|/2$, and define:

$$I_<(\xi) = \int_{|\eta|<|\xi|/2}\langle\xi\rangle^\sigma\,|\hat v(\xi-\eta)\hat\varphi(\eta)|\,d\eta, \qquad I_>(\xi) = \int_{|\eta|>|\xi|/2}\langle\xi\rangle^\sigma\,|\hat v(\xi-\eta)\hat\varphi(\eta)|\,d\eta\,.$$

Let us start by estimating the integral $I_>(\xi)$. The function $\hat\varphi \in \mathscr{S}$ decays rapidly, namely $\hat\varphi(\eta) = \mathcal{O}\big(\langle\eta\rangle^{-N}\big)$ for any $N>0$, so

$$I_>(\xi) \le C_N\langle\xi\rangle^\sigma\int_{|\eta|>|\xi|/2}|\hat v(\xi-\eta)|\,\frac{\langle\xi-\eta\rangle^\sigma}{\langle\xi-\eta\rangle^\sigma\,\langle\eta\rangle^N}d\eta\,.$$

A Cauchy-Schwartz argument gives the following bound for the integral

$$\Big(\int_{|\eta|>|\xi|/2}|\hat v(\xi-\eta)\,\langle\xi-\eta\rangle^\sigma|^2\,d\eta\int_{|\eta|>|\xi|/2}\langle\xi-\eta\rangle^{-2\sigma}\,\langle\eta\rangle^{-2N}\,d\eta\Big)^{1/2}$$

The first integral above converges from the assumption on $v \in H^\sigma$. Whatever the sign of σ, the second integral is bounded above by

$$\begin{aligned}\int_{|\eta|>|\xi|/2} \langle \xi-\eta\rangle^{2|\sigma|}\,\langle \eta\rangle^{-2N}\,d\eta &\leq C\int_{|\eta|>|\xi|/2}\frac{\langle\xi\rangle^{2|\sigma|}+\langle\eta\rangle^{2|\sigma|}}{\langle\eta\rangle^{2N}}\,d\eta\\ &\leq C'\int_{|\eta|>|\xi|/2}\frac{1}{\langle\eta\rangle^{2N-2|\sigma|}}\,d\eta\\ &\leq C''\langle\xi\rangle^{-2N+2|\sigma|+d}.\end{aligned}$$

Since N can be chosen arbitrary large, we see that $I_>(\xi)$ is rapidly decreasing when $|\xi|\to\infty$, in particular it is square-integrable.

Let us now deal with the integral $I_<(\xi)$.Applyint the Cauchy–Schwarz inequality, we get

$$I_<(\xi) \leq \Big(\int_{|\eta|\leq|\xi|/2}|\hat\varphi(\eta)|\,d\eta\Big)^{1/2}\Big(\int_{|\eta|\leq|\xi|/2}\langle\xi\rangle^{2\sigma}|\hat v(\xi-\eta)|^2|\hat\varphi(\eta)|\,d\eta\Big)^{1/2}.$$

The first integral on the RHS obviously converges. Since $|\eta|\leq|\xi|/2$, we have $\langle\xi\rangle\leq A\langle\xi-\eta\rangle$ for some constant $A>0$; this allows to replace $\langle\xi\rangle^{2\sigma}|\hat v(\xi-\eta)|^2$ by $A\langle\xi-\eta\rangle^{2\sigma}|\hat v(\xi-\eta)|^2$, which we know to be square integrable. We thus get:

$$I_<(\xi)^2 \leq B\int_{\mathbb{R}^d}\langle\xi-\eta\rangle^{2\sigma}|\hat v(\xi-\eta)|^2|\hat\varphi(\eta)|\,d\eta\,.$$

If we integrate this expression over $\xi\in\mathbb{R}^d$, we may apply Fubini's theorem and first integrate over ξ, which gives a finite result since $v\in H^\sigma$. The remaining η-integral also converges. This proves that $I_<\in L^2$ as well, so finally $I_<+I_<\in L^2$. □

3.5.3 Global results: equations over $\mathbb{R}^d$

Homogeneous equations on $\mathbb{R}^d$

We start by studying homogeneous equations of the form $Pu=0$ for $u\in\mathscr{S}'(\mathbb{R}^d)$.

Corollary 3.5.9 With the notation of Proposition 3.5.3, we have

i) If $p(i\xi)\neq 0$ for every $\xi\in\mathbb{R}^d$, then the only tempered distribution solution of $Pu=0$ is $u=0$.

ii) If $p(i\xi)\neq 0$ for every $\xi\in\mathbb{R}^d\setminus\{0\}$, then every tempered solution of $Pu=0$ is a polynomial function.

Proof.— If $u\in\mathscr{S}'$ satisfies $Pu=0$, then $p(i\xi)\hat u=0$, which implies that the distribution $\hat u$ is supported in the set $\{\xi\in\mathbb{R}^d, p(i\xi)=0\}$. If this set is empty, this implies $\hat u=0$, hence $u=0$, whence $i)$.

If this set reduces to $\{0\}$, this implies, from the structure of the distributions supported at a single point, that $\hat{u}$ is a finite linear combination of derivatives of δ_0. Applying $\mathcal{F}^{-1}$, we conclude that u is a polynomial function, which is $ii)$. $\square$

Remark 3.5.10 A consequence of Corollary 3.5.9 is that, under the assumption $ii)$, bounded solutions of $Pu = 0$ on $\mathbb{R}^d$ are constant functions. Indeed, since $L^\infty(\mathbb{R}^d) \subset \mathscr{S}'(\mathbb{R}^d)$, this follows from the elementary fact that a polynomial function which is bounded on $\mathbb{R}^d$ is constant. This is well-known if $d = 1$, but less obvious if $d \geq 2$. Let us sketch a proof of this fact. Let $p \in \mathbb{C}[X_1, \dots, X_d]$ be a polynomial such that the function $p(x)$ is bounded on $x \in \mathbb{R}^d$. For every $y \in \mathbb{R}^d$, consider

$$p_y(t) = p(ty) \ , \ t \in \mathbb{R} \ .$$

The function p_y is a polynomial function of one variable, which is bounded, hence it is a constant of t. Decomposing

$$p(X) = \sum_{|\alpha| \leq m} a_\alpha X^\alpha,$$

this implies that, for every $r = 1, \dots, m$, and every $y \in \mathbb{R}^d$, the coefficient of t^r in p_y vanishes:

$$\sum_{|\alpha| = r} a_\alpha y^\alpha = 0 \ .$$

Taking the ∂^α derivative of the left hand side, we infer $a_\alpha = 0$, so all these r-homogeneous parts vanish, and only the constant component $r = 0$ remains.

Examples 3.5.11
- Case $i)$ is fulfilled by $P = \Delta + \lambda$ if $\lambda \in \mathbb{C} \setminus \mathbb{R}_+$.
- Case $ii)$ is fulfilled by $P = \Delta$ on $\mathbb{R}^d$. Consequently, tempered harmonic functions on $\mathbb{R}^d$ are polynomial functions. In particular, we infer the strong Liouville theorem : any bounded harmonic function on $\mathbb{R}^d$ is a constant.

  Case $ii)$ is also fulfilled by $P = \partial_1 + i\partial_2$ on $\mathbb{R}^2$. In other words, the only tempered entire functions on $\mathbb{C}$ are polynomial, and again (Liouville), we recover that the only bounded entire functions are constant.

Fundamental solutions

The second application concerns inhomogeneous equations, $Pu = f$. It starts with the special case where $f = \delta_0$.

Definition 3.5.12 Let $P = p(\partial)$ be a linear partial differential operator with constant coefficients. One says that $E \in \mathscr{D}'(\mathbb{R}^d)$ is a fundamental solution of P if it satisfies $PE = \delta_0$.

In Physics, fundamental solutions are often called Green's functions. The importance of fundamental solutions is provided by the following

Proposition 3.5.13 If P has a fundamental solution $E \in \mathscr{D}'(\mathbb{R}^d)$, then for all $\varphi \in \mathscr{D}(\mathbb{R}^d)$, the equation $Pu = \varphi$ has a solution given by the convolution $u = E * \varphi$.

Proof.— In chapter 2, we saw that for any $T \in \mathscr{D}'(\mathbb{R}^d)$ and $\varphi \in \mathscr{D}(\mathbb{R}^d)$, $\partial_j(T*\varphi) = T*\partial_j\varphi = \partial_j T*\varphi$. Hence

$$P(E * \varphi) = (PE) * \varphi = \delta_0 * \varphi = \varphi,$$

so that $u = E * \varphi$ is a solution of $Pu = \varphi$. □

Using a convenient definition of the convolution between a distribution and a compactly supported distribution, it is possible to extend the above proposition to any right hand side which is a compactly supported distribution. We shall see an example of this in Theorem 3.5.17 below.

B. Malgrange and L. Ehrenpreis have proved, independently in 1954/1955, that any non trivial linear partial differential operator with constant coefficients has a fundamental solution. It is fact possible to prove that one can choose this distribution to be tempered. The proofs of these results are beyond the scope of these lectures. In what follows, we rather study the important case of the Laplace operator.

Proposition 3.5.14 Let $d \geq 3$, and let $E_d \in \mathscr{D}'(\mathbb{R}^d)$ be the $L^1_{\mathrm{loc}}(\mathbb{R}^d)$ function defined by

$$E_d(x) = \frac{1}{(d-2)\sigma(S^{d-1})} \frac{1}{|x|^{d-2}}.$$

Then

$$-\Delta E_d = \delta_0 \ .$$

Proof.— 1. We are going to use Fourier transformation. Indeed, by Proposition 3.5.3, in $\mathscr{S}'(\mathbb{R}^d)$, the equation

$$-\Delta E = \delta_0$$

is equivalent to

$$|\xi|^2 \hat{E} = 1 \ .$$

Since $d \geq 3$, the function $1/|\xi|^2$ is locally integrable in $\mathbb{R}^d$. More precisely, by decomposing it as the sum of the contributions for $|\xi| \leq 1$ and for $|\xi| > 1$, one notes that this function is the sum of an L^1 function and of an L^∞ function. As a consequence, $1/|\xi|^2$ belongs to $\mathscr{S}'(\mathbb{R}^d)$, therefore the distribution

(3.5.11) $$E_d := \mathcal{F}^{-1}\left(\frac{1}{|\xi|^2}\right)$$

is a tempered fundamental solution of $-\Delta$ (a priori, there can be many different solutions in $\mathscr{S}'$, differing by harmonic polynomials.

2.We now want to evaluate the particular solution E_d. Since any fundamental solution satisfies $-\Delta E = 0$ on $\mathbb{R}^d \setminus \{0\}$, we know from Theorem 3.5.7 that E_d is a C^∞ function on $\mathbb{R}^d \setminus \{0\}$. We want to show that the particular fundamental solution E_d is of the form

$$F_d(x) = \frac{c_d}{|x|^{d-2}}, \quad \forall x \in \mathbb{R}^d \setminus 0\,,$$

for some constant c_d. This functional form is characterized by the following symmetries: for every rotation $R \in O(d, \mathbb{R})$ and every $\lambda > 0$,

$$\forall x \in \mathbb{R}^d \setminus \{0\}, \quad F_d(\lambda R x) = \lambda^{2-d} F_d(x)\ ,$$

namely F_d is invariant by the rotations centered at the origin, and $2-d$-homogeneous w.r.to dilations. Equivalently, for every $\varphi \in \mathscr{D}(\mathbb{R}^d \setminus \{0\})$,

$$\int_{\mathbb{R}^d} F_d(\lambda R x)\varphi(x)\,dx = \lambda^{2-d} \int_{\mathbb{R}^d} F_d(x)\varphi(x)\,dx\ . \tag{3.5.12}$$

We will group together rotation and dilation in a single matrix $A := \lambda R$, and defined its action of φ as: $\varphi_A(x) := |\det A|^{-1}\varphi(A^{-1}x)$. Le us now study the action of this matrix on the solution E_d defined in (3.5.11). Using the same $\varphi \in \mathscr{D}(\mathbb{R}^d \setminus \{0\})$, we get

$$\int_{\mathbb{R}^d} E_d(Ax)\,\varphi(x)\,dx = \int_{\mathbb{R}^d} E_d(x)\,\varphi_A(x)\,dx, \quad \text{with} \quad \varphi_A(x) := |\det A|^{-1}\varphi(A^{-1}x)\,.$$

The Plancherel formula leads to

$$\begin{aligned}\int_{\mathbb{R}^d} E_d(x)\,\varphi_A(x)\,dx &= (2\pi)^{-d}\int_{\mathbb{R}^d} \sigma(\hat{E}_d)(\xi)\,\widehat{\varphi_A}(\xi)\,d\xi \\ &= (2\pi)^{-d}\int_{\mathbb{R}^d} \frac{\widehat{\varphi_A}(\xi)}{|\xi|^2}\,d\xi\,.\end{aligned}$$

The action of A on φ reads as follows on the Fourier transform side:

$$\widehat{\varphi_A}(\xi) = \int_{\mathbb{R}^d} |\det A|^{-1}\varphi(A^{-1}x)\mathrm{e}^{-ix\cdot\xi}\,dx = \int_{\mathbb{R}^d} \varphi(x)\mathrm{e}^{-iAx\cdot\xi}\,dx = \hat{\varphi}({}^tA\xi)\ .$$

Consequently,

$$\int_{\mathbb{R}^d} \frac{\widehat{\varphi_A}(\xi)}{|\xi|^2}\,d\xi = \int_{\mathbb{R}^d} \frac{\hat{\varphi}({}^tA\xi)}{|\xi|^2}\,d\xi = |\det A|^{-1}\int_{\mathbb{R}^d} \frac{\hat{\varphi}(\xi)}{|{}^tA^{-1}\xi|^2}\,d\xi\ .$$

Since ${}^tA^{-1} = \lambda^{-1}R$ and $|\det A| = \lambda^d$, we conclude

$$\int_{\mathbb{R}^d} \frac{\widehat{\varphi_A}(\xi)}{|\xi|^2}\,d\xi = \lambda^{2-d}\int_{\mathbb{R}^d} \frac{\hat{\varphi}(\xi)}{|\xi|^2}\,d\xi, \quad \text{hence} \quad \int_{\mathbb{R}^d} E_d(Ax)\,\varphi(x)\,dx = \lambda^{2-d}\int_{\mathbb{R}^d} E_d(x)\,\varphi(x)\,dx\,.$$

so that E_d satisfies the identity (3.5.12), and is therefore of the form F_d on $\mathbb{R}^d \setminus 0$.

3. Before computing the constant c_d, we need to verify that the solution E_d is just given, near 0, by the L^1_{loc} function F_d, without any extra singular piece supported at the origin. Such extra piece would be a linear combination of $\partial^\alpha\delta_0$. On what grounds can we exclude them? Above we tested E_d only with functions $\varphi \in \mathscr{D}(\mathbb{R}^d \setminus 0)$, which cannot detect such singularities at the origin. But the above computations, which show that for $A = \lambda R$,

$$\langle E_d, \varphi_A\rangle = \lambda^{2-d}\,\langle E_d, \varphi\rangle,$$

applies as well to $\varphi \in \mathscr{D}(\mathbb{R}^d)$. Let us show that singular components at the origin $\partial^\alpha \delta_0$ cannot satisfy this identity. The action of the dilation $A = \lambda\, Id$ will suffice: for any $\alpha \in \mathbb{N}^d$, the expression

$$\langle \partial^\alpha \delta_0, \varphi_A \rangle = (-1)^{|\alpha|} \langle \delta_0, \partial^\alpha [\lambda^{-d} \varphi(\lambda \cdot)] \rangle = \lambda^{-d-|\alpha|} \langle \partial^\alpha \delta_0, \varphi \rangle$$

satisfies a wrong homogeneity, so it cannot be part of E_d. Finally, the distribution E_d is equal to the L^1_{loc} function F_d.

4. Let us finally calculate the factor c_d. For this aim, it is sufficient to check E_d against the Gaussian function G_1, and apply the Plancherel formula :

$$c_d \int_{\mathbb{R}^d} \frac{G_1(x)}{|x|^{d-2}}\, dx = (2\pi)^{-d} \int_{\mathbb{R}^d} \frac{\hat{G}_1(\xi)}{|\xi|^2}\, d\xi \ .$$

Since $\hat{G}_1(\xi) = (2\pi)^{d/2} G_1(\xi)$, we infer

$$c_d \int_{\mathbb{R}^d} \frac{\mathsf{e}^{-|x|^2/2}}{|x|^{d-2}}\, dx = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \frac{\mathsf{e}^{-|\xi|^2/2}}{|\xi|^2}\, d\xi \ .$$

Passing in spherical coordinates, this reads

$$c_d \int_0^\infty r \mathsf{e}^{-r^2/2}\, dr = (2\pi)^{-d/2} \int_0^\infty r^{d-3} \mathsf{e}^{-r^2/2}\, dr \ ,$$

or

$$c_d = (2\pi)^{-d/2} 2^{(d-4)/2} \int_0^\infty t^{(d-4)/2} \mathsf{e}^{-t}\, dt = \frac{1}{4} \pi^{-d/2} \Gamma\left(\frac{d}{2} - 1\right)$$

We recall that the superficial measure of the unit sphere S^{d-1} is given by

$$\sigma(S^{d-1}) = \frac{2\pi^{d/2}}{\Gamma(d/2)} = \frac{4\pi^{d/2}}{(d-2)\Gamma(d/2-1)} \ ,$$

so that

$$c_d = \frac{1}{(d-2)\sigma(S^{d-1})} \ .$$

□

Remark 3.5.15 The value of c_d can also be determined by applying the Gauss–Green formula to the integral of $|x|^{2-d} \Delta\varphi(x)$ outside a ball of radius ε.

Remark 3.5.16 It is possible to extend Proposition 3.5.14 to $d = 1, 2$. For $d = 1$, it is easy to check that

$$E_1(x) = -\frac{1}{2}|x|$$

is a fundamental solution of $-\partial^2$.

The case $d = 2$ is more delicate. The above approach by Fourier analysis has to be slightly modified, because $1/|\xi|^2$ is not locally integrable in $\mathbb{R}^2$. A possible way out of this problem is to consider the operator

$$P = \partial_1 + i\partial_2 = 2\partial_{\bar{z}} \ , \quad \text{if we take } z = x_1 + ix_2 \, .$$

Since $1/(\xi_1 + i\xi_2) \in L^1_{\mathrm{loc}}(\mathbb{R}^2)$, the distribution $E \in \mathscr{S}'(\mathbb{R}^2)$ defined by its Fourier transform

$$\hat{E} = \frac{1}{i(\xi_1 + i\xi_2)}, \quad \text{which is in } L^1_{loc}(\mathbb{R}^d) \cap \mathscr{S}'(\mathbb{R}^d),$$

satisfies $PE = \delta_0$. Furthermore, from Theorem 3.5.7, E is a holomorphic function outside $\{0\}$, and the above proof shows that it is homogeneous of degree -1. These properties lead to the expression $E(x) = c\,(x_1 + ix_2)^{-1}$ on $\mathbb{R}^2 \setminus 0$, and the constant c can be computed again by testing PE again G_1, leading to:

$$E(x) = \frac{1}{2\pi(x_1 + ix_2)}.$$

Now, observing that $\Delta = (\partial^1 + i\partial_2)(\partial_1 - i\partial_2)$ and that

$$E = (\partial_1 - i\partial_2)\big[(4\pi)^{-1} \log(x_1^2 + x_2^2)\big] \quad \text{in } \mathscr{D}'(\mathbb{R}^2),$$

we conclude that

$$E_2 = -\frac{1}{2\pi} \log|x|$$

is a fundamental solution of $-\Delta$ on $\mathbb{R}^2$.

Poisson equation on $\mathbb{R}^3$

We finally study the Poisson equation $-\Delta u = f$ on $\mathbb{R}^3$, where f is an arbitrary compactly supported distribution. Notice that, from Theorem 3.5.7, every solution u of this equation is a C^∞ function outside of supp(f). In physics, this is the equation satisfied by the electric potential u generated by a charge distribution f.

Theorem 3.5.17 For every compactly supported distribution f on $\mathbb{R}^3$, there exists a unique distribution u on $\mathbb{R}^3$ satisfying $-\Delta u = f$ and such that, outside supp f, $u(x) \to 0$ when $x \to \infty$. Furthermore, for any $x \in \mathbb{R}^3 \setminus \text{supp}(f)$, the function $u(x)$ is given by

(3.5.13) $$u(x) = \left\langle f, \frac{\chi}{4\pi|x - .|} \right\rangle,$$

for every $\chi \in \mathscr{D}(\mathbb{R}^3)$ equal to 1 in a neighbourhood of supp(f) and equal to 0 near x (so that the function in the bracket is indeed in $\mathscr{D}(\mathbb{R}^3)$).

Proof.— The uniqueness of u is immediate in view of the Liouville theorem for harmonic functions — see Example 3.5.11. For the existence, we look for u in $\mathscr{S}'(\mathbb{R}^3)$, and Proposition 3.5.3 leads to the equation

$$|\xi|^2\, \hat{u} = \hat{f}.$$

Since $\hat{f}$ is smooth on $\mathbb{R}^3$ with moderate growth and $1/|\xi|^2$ belongs to $\mathscr{S}'(\mathbb{R}^d)$, the function

$$\hat{u}(\xi) = \frac{\hat{f}(\xi)}{|\xi|^2} \quad \text{belongs to } \mathscr{S}'(\mathbb{R}^d)\,.$$

We define $u = \mathcal{F}^{-1}\hat{u} \in \mathscr{S}'(\mathbb{R}^d)$. Our task is now to prove the identity (3.5.13). Notice that this expression automatically satisfies $u(x) \to 0$ as $x \to \infty$.

As a first step, we are going to prove this identity when $f = \varphi \in \mathscr{D}(\mathbb{R}^3)$. In this case, let us denote by g the inverse Fourier transform of

$$\hat{g} = \frac{\hat{\varphi}}{|\xi|^2}\ .$$

Since $\hat{g}$ is integrable on $\mathbb{R}^3$ and decays rapidly at infinity, g is smooth and converges to 0 at infinity. Consider the following smooth function, obtained by convolution of $E_3 \in \mathscr{S}'$ with $\varphi \in \mathscr{D}$:

$$E_3 * \varphi(x) = \int_{\mathbb{R}^3} \frac{\varphi(y)}{4\pi|x-y|}\, dy\ .$$

Then we've seen before that

$$-\Delta(E_3 * \varphi) = \varphi = -\Delta g\,,$$

and it is clear on the above expression that $E_3 * \varphi(x) \to 0$ as $x \to \infty$. Therefore, by the Liouville uniqueness theorem, $g = E_3 * \varphi$.

Let us come to the general case of a compactly supported distribution f. The strategy is to approximate f by smooth regularizations $f_\varepsilon \in \mathscr{D}$. Let $\rho \in \mathscr{D}(\mathbb{R}^3)$ be a convolution kernel and, for every $\varepsilon > 0$, defined the renormalized kernel

$$\rho_\varepsilon(x) = \frac{1}{\varepsilon^3}\rho\left(\frac{x}{\varepsilon}\right)\ .$$

Then $f_\varepsilon := \rho_\varepsilon * f$ is in $\mathscr{D}(\mathbb{R}^3)$ and converges to f in $\mathscr{D}'(\mathbb{R}^3)$ when $\varepsilon \to 0$. As we already observed in similar cases, the distributions $(f_\varepsilon)_{\varepsilon\in]0,1]}$ are all supported inside some compact neighbourhood of supp(f), so that

$$\hat{f}_\varepsilon(\xi) = \langle f_\varepsilon, \chi\, e^{-ix\cdot\xi}\rangle$$

converge to $\hat{f}(\xi)$, and these brackets are controlled by a uniform norm $C\|\chi\, e^{-ix\cdot\xi}$
$|_{C^m} \leq C\,\langle\xi\rangle^m$. Let $u_\varepsilon = E_3 * f_\varepsilon$. We know from the above identity that

$$\hat{u}_\varepsilon = \frac{\hat{f}_\varepsilon}{|\xi|^2}\,,$$

and the right hand side converges to $\hat{u}$ in $\mathscr{S}'(\mathbb{R}^3)$. Hence u_ε converges to u in $\mathscr{S}'(\mathbb{R}^3)$, and for every test function $\varphi \in \mathscr{D}(\mathbb{R}^3)$,

$$\begin{aligned}
\langle u, \varphi\rangle &= \lim_{\varepsilon\to 0}\langle u_\varepsilon, \varphi\rangle = \lim_{\varepsilon\to 0}\int_{\mathbb{R}^3}(E_3 * f_\varepsilon)(x)\,\varphi(x)\,dx \\
&= \lim_{\varepsilon\to 0}\int\int_{\mathbb{R}^3\times\mathbb{R}^3} E_3(x-y)f_\varepsilon(y)\varphi(x)\,dy\,dx \\
&\overset{Fubini}{=} \lim_{\varepsilon\to 0}\int_{\mathbb{R}^3} f_\varepsilon(y)E_3 * \varphi(y)\,dy \\
&= \langle f, \chi(E_3 * \varphi)\rangle
\end{aligned}$$

for every $\chi \in \mathscr{D}$ equal to 1 on the supports the f_ε for ε small enough. Notice that

$$\chi(y)\,(E_3 * \varphi)(y) = \int_{\mathbb{R}^3} \frac{\chi(y)}{4\pi|x-y|}\,\varphi(x)\,dx\ .$$

If supp$(\varphi) \cap$ supp$(\chi) = \emptyset$, we can apply Proposition 2.3.21 of integration under the bracket to obtain

$$\langle u, \varphi\rangle = \langle f, \chi(E_3 * \varphi)\rangle = \int_{\mathbb{R}^3} \varphi(x) \left\langle f, \frac{\chi}{4\pi|x-.|} \right\rangle dx$$

and this yields formula (3.5.13). □

Remark 3.5.18 Notice that formula (3.5.13) leads to an expansion of $u(x)$ as $x \to \infty$. In particular,

$$u(x) = \frac{q}{4\pi|x|} + O\left(\frac{1}{|x|^2}\right)\,,$$

where $q = \langle f, \chi\rangle$ for every $\chi \in \mathscr{D}(\mathbb{R}^3)$ a cutoff function in a neighbourhood of supp(f). This expansion supports the well known fact in Physics that any charge distribution f with nonzero total charge q can be seen at infinity as a point distribution at the origin with charge q.

PDEs with variable coefficients

We conclude this section by a remark about the generalization of the above approach to partial differential equations with nonconstant coefficients. The Fourier transformation can be used to study such equations, along several directions:

i) The first and most direct topic concerns equations with affine coefficients. Indeed, the Fourier transform converts such equations into first order linear equations with polynomial coefficients, which can be solved in general. A famous example in dimension $d = 1$ is the Airy differential equation on $\mathbb{R}$,

$$u''(x) = xu(x)\,.$$

It is well known from the classical theory of differential equations that the solutions of this ODE make up a 2-dimensional space of smooth functions. Let us look for solutions u which are moreover tempered. Then the equation satisfied by $\hat{u}$ is

$$(i\xi)^2\hat{u} = i\frac{d}{d\xi}\hat{u}, \quad \text{that is} \quad \frac{d\hat{u}}{d\xi} = i\xi^2\hat{u}\ .$$

The solutions in $\mathscr{D}'(\mathbb{R})$ are given by

$$\hat{u}(\xi) = c\,\mathrm{e}^{i\frac{\xi^3}{3}}\ ,\ c \in \mathbb{C}\ ,$$

which are indeed tempered distributions, since in $L^\infty(\mathbb{R})$. Consequently, the space of *tempered* solutions of the Airy differential equation is one dimensional, generated by the Airy function,

$$Ai(x) = \mathcal{F}^{-1}\big(\mathrm{e}^{i\frac{\xi^3}{3}}\big)\ .$$

ii) The second and much more general topic concerns the regularity theory of distribution, solutions to partial differential equations with arbitrary smooth coefficients in arbitrary open subsets of $\mathbb{R}^d$, using cutoff functions in a smart way. This is the starting point of the theory of pseudodifferential operators, which is beyond the scope of this course.