Final Degree Dissertation
Degree in Mathematics

Distribution Theory and Fundamental Solutions of Differential Operators

Author:
Daniel Eceizabarrena Pérez
Supervisor:
Javier Duoandikoetxea Zuazo

Leioa, June 24th 2015

Contents

Introduction

The objective of this dissertation is to study the **theory of distributions** and some of its applications. Even if the vast majority of the theoretical results developed throughout the Degree correspond to the time before 20th century, we have to focus on the last 100 years to understand the matter of this document.

Certain concepts which we would include in the theory of distributions nowadays have been widely used in several fields of mathematics and physics. For example, many approaches to the very famous delta function were known to be used by mathematicians such as Cauchy and Poisson, and some more recent as Kirchoff and Hermite. In their works, they make use of some functions which they call *impulse functions*. Nevertheless, it was Dirac who first introduced the delta function as we know it, in an attempt to keep a convenient notation in his works in quantum mechanics. Their work contributed to open a new path in mathematics, as new objects, similar to functions but not of their same nature, were being used systematically.

Distributions are believed to have been first formally introduced by the Soviet mathematician **Sergei Sobolev**[1], in an attempt to find weak solutions to partial differential equations. Nevertheless, it is the name of **Laurent Schwartz**[2] which is most deeply connected to distribution theory. His book *Théorie des distributions*, published in 1950, is the source of the first systematic development of distributions, and it highlighted their utility.

The aim of this project is to show how distribution theory can be used to obtain what we call fundamental solutions of partial differential equations. The connection between these two areas is given by the Fourier transform, probably one of the most useful inventions of mathematics.

In short, distributions are linear and continuous functionals that assign a complex number to every function of a certain space. The definition of the space of **testing functions**, $\mathcal{D}$, and of the space of distributions, $\mathcal{D}'$, will be given first. Nevertheless, the Fourier transform does not fit with the first and more basic definition of distributions, so it will be necessary to introduce a more general space of functions, the **Schwartz space**, $\mathcal{S}$, in which the Fourier transform behaves properly. This fact will naturally imply the existence of a new space of distributions, $\mathcal{S}'$, which will be known as the space of **tempered distributions**.

[1]Sergei Lvovich Sobolev (1908, Saint Petersburg - 1989, Moscow) worked in analysis and partial differential equations. The very famous Sobolev spaces are named after him.

[2]See Appendix A.

Definition and properties of the **Fourier transform** in $\mathcal{S}$ will be described afterwards. This will allow to easily define the Fourier transform of tempered distributions, being the next and last step to work with differential operators and their **fundamental solutions**. To illustrate the usefulness of the developed theory, some examples concerning the heat equation, the Schrödinger equation, the Laplace equation and the Cauchy-Riemann equations will be covered.

Chapter 1

Distributions

1.1 The space $\mathcal{D}$ of testing functions

As it will be developed in this chapter, distributions are operators from a certain space of functions to the field of real or complex numbers. Several function spaces can be defined to do so; the space of testing functions will be our first approach.

Definition 1.1. Let $\Omega \subset \mathbb{R}^n$ be an open set and $\varphi : \Omega \to \mathbb{C}$ a function. We say that φ is a **testing function** if:

1. $\operatorname{supp} \varphi$ is compact, and
2. φ is C^∞.

The space of testing functions in Ω is denoted by $\mathcal{D}(\Omega)$.

Remark 1.2. Recall that the support of a continuous function is the closure of the set of points in which the value of the function is not zero. In other words, $\operatorname{supp} \varphi = Cl\{x \in \Omega \mid \varphi(x) \neq 0\}$. Because of this, it is enough to ask the support to be bounded, as it will always be closed.

The first immediate property that can be obtained from the definition is given in the following lemma. From now on, most of the times we will consider $t = (t_1, \cdots, t_n)$ to be the general variable in $\mathbb{R}^n$, even if sometimes we will use $x = (x_1, \cdots, x_n)$ too.

Lemma 1.3. *If* $\varphi \in \mathcal{D}(\Omega)$*, then* $\frac{\partial \varphi}{\partial t_i} \in \mathcal{D}(\Omega), \forall i = 1, \cdots, n.$

Proof. It is immediate, as $\frac{\partial \varphi}{\partial t_i}$ is C^∞ because of the definition of φ. Moreover, if $t \notin \operatorname{supp} \varphi$, as the complementary of the support is an open set, there exists a neighbourhood of t in which $\varphi = 0$. Thus, in that neighbourhood, $\frac{\partial \varphi}{\partial t_i} = 0$ and $\frac{\partial \varphi}{\partial t_i}(t) = 0$, which implies that $t \notin \operatorname{supp} \frac{\partial \varphi}{\partial t_i}$. In other words, $\operatorname{supp} \frac{\partial \varphi}{\partial t_i} \subseteq \operatorname{supp} \varphi$, which completes the proof. □

We will be able to give the space $\mathcal{D}$ an algebraic structure. The following proposition shows why:

Proposition 1.4. *Let $\varphi, \phi \in \mathcal{D}(\Omega)$ and $a, b \in \mathbb{C}$. Then, $a\varphi + b\phi \in \mathcal{D}(\Omega)$. Therefore, $\mathcal{D}(\Omega)$ is a linear space.*

Proof. Let $A = \operatorname{supp}\varphi$ and $B = \operatorname{supp}\phi$. Then, $a\varphi(x) + b\phi(x) = 0$ for $x \notin A \cup B$. This, together with the fact of $A \cup B$ being closed, implies that $\operatorname{supp} a\varphi + b\phi \subseteq A \cup B$. What is more, as A and B are bounded, $\operatorname{supp} a\varphi + b\phi$ is bounded too.

Also, the fact that both φ and ϕ are in $C^\infty(\Omega)$ implies $a\varphi + b\phi \in C^\infty(\Omega)$. Finally, being working with functions with values in a field, every property for a linear space holds trivially. □

Once we have defined the concept of testing function, let us show an example.

Example 1.5. Let $\zeta(t)$ be defined by parts:

$$\zeta(t) = \begin{cases} e^{\frac{1}{t^2-1}} & |t| < 1, \\ 0 & |t| \geq 1. \end{cases} \tag{1.1}$$

Clearly $\operatorname{supp}\zeta = [-1, 1]$, and ζ is infinitely smooth whenever $|t| \neq 1$, so we only have to check the situation at points $|t| = 1$. Anyway, as the function is even, it is enough to analyse the case $t = 1$. We need to work out the derivatives in $(-1, 1)$. First,

$$\zeta'(t) = -\frac{2t}{(t^2-1)^2} e^{\frac{1}{t^2-1}}, \quad |t| < 1,$$

and it is fairly easy to observe that every derivative at $|t| < 1$ will be of the form

$$\zeta^{(k)} = \frac{P_k(t)}{(t^2-1)^{2k}} e^{\frac{1}{t^2-1}}, \quad |t| < 1,$$

where every P_k is a polynomial. Now, after a change of variables,

$$\lim_{t \to 1^-} \frac{P_k(t)}{(t^2-1)^{2k}} e^{\frac{1}{t^2-1}} = \frac{P_k(1)}{4^k} \lim_{t \to 0^-} \frac{e^{1/2t}}{t^{2k}} = 0.$$

This limit shows that every derivative of ζ is continuous, as $\zeta^{(k)}(t) = 0$ in $|t| > 1$. As a consequence, $\zeta(t) \in \mathcal{D}(\mathbb{R})$.

Remark 1.6. Most of the times, we will consider $\Omega = \mathbb{R}^n$, so from now on, we will use the notation $\mathcal{D}(\mathbb{R}^n)$ (or simply $\mathcal{D}$ if it is clear the space we are working in) instead of the more general $\mathcal{D}(\Omega)$.

An interesting fact about testing functions is that one can produce new ones fairly easy. The following proposition advocates so.

Proposition 1.7. *Let $\varphi \in \mathcal{D}(\mathbb{R}^n)$, $\tau \in \mathbb{R}^n$, $a \in \mathbb{R} - \{0\}$ and $g \in C^\infty(\mathbb{R}^n)$. Then,*

1. $\varphi(t + \tau), \varphi(-t), \varphi(at) \in \mathcal{D}(\mathbb{R}^n)$.
2. $g(t)\varphi(t) \in \mathcal{D}(\mathbb{R}^n)$.

Proof. Every new function is infinitely smooth, as they are nothing but compositions and products of C^∞ functions. Also, their supports are compact for being translations, transpositions or dilations of that of φ, which is compact. □

We will now define the concept of **convergence in $\mathcal{D}$**, for it is key to understand distributions, as we will later see.

Definition 1.8. Let $\{\varphi_m(t)\}_{m\in\mathbb{N}}$ be a sequence of testing functions. We say it converges to zero in $\mathcal{D}(\mathbb{R}^n)$ if

1. there exists $K \subseteq \mathbb{R}^n$ a bounded domain such that $\operatorname{supp}\varphi_m \subseteq K, \forall m \in \mathbb{N}$, and
2. each sequence $\{D^k\varphi_m(t)\}_{m\in\mathbb{N}}$ converges uniformly to zero as $m \to \infty$.

More generally, we say that the sequence $\{\varphi_m(t)\}_{m\in\mathbb{N}}$ converges to $\varphi(t)$ in $\mathcal{D}(\mathbb{R}^n)$ if every $\operatorname{supp}\varphi_m$ is contained in a single bounded domain $K \subset \mathbb{R}^n$ and the sequence $\{\varphi_m(t) - \varphi(t)\}_{m\in\mathbb{N}}$ converges to zero in $\mathcal{D}$. In this case, we will write $\{\varphi_m\} \to \varphi$, or simply $\varphi_m \to \varphi$.

Notation. We will use the notation $D^k\varphi(t)$ with a multi-index $k = (k_1, \cdots, k_n)$ of non-negative integers to represent the partial derivatives of the function φ. In other words,

$$D^k\varphi(t) = \frac{\partial^{k_1+\cdots+k_n}}{\partial t_1^{k_1}\cdots\partial t_n^{k_n}}\varphi(t_1, \cdots, t_n).$$

We will also write $|k| = k_1 + \cdots + k_n$.

Once we have defined the basics of testing functions, we are ready to introduce the space of distributions.

1.2 The space $\mathcal{D}'$ of distributions

As we stated at the beginning of Section 1.1, distributions are a special kind of functionals, which assign a complex number to each function from a particular space of functions. Let us define formally what a distribution is.

1.2.1 Definition

Definition 1.9. A mapping $T : \mathcal{D}(\mathbb{R}^n) \to \mathbb{C}$ is called a **distribution** if:

1. it is linear, in the sense that if $\varphi, \phi \in \mathcal{D}(\mathbb{R}^n)$ and $a, b \in \mathbb{C}$, then
$$T(a\varphi + b\phi) = aT(\varphi) + bT(\phi).$$
2. it is continuous, in the sense that if $\{\varphi_m\} \to \varphi$ in $\mathcal{D}(\mathbb{R}^n)$, then
$$\{T(\varphi_m)\} \to T(\varphi)$$
in $\mathbb{C}$.

The space of distributions is denoted by $\mathcal{D}'(\mathbb{R}^n)$. Some authors also call distributions **generalised functions**.

Remark 1.10. In the same way we could define the space of testing functions over an open set $\Omega \subset \mathbb{R}^n$, we can also define distributions over Ω. Anyway, as stated in Remark 1.6, we will consider the whole space $\mathbb{R}^n$.

A distribution can be represented by capital letters such as T, but it is usual to use the same letters as for functions, such as f and g. Also, if $\varphi \in \mathcal{D}(\mathbb{R}^n)$, we will write its image with $f \in \mathcal{D}'(\mathbb{R}^n)$ as $\langle f, \varphi \rangle$.

As an immediate consequence of the definition, we can make the condition of continuity less restrictive if we know a functional is linear.

Proposition 1.11. *If a functional $f : \mathcal{D}(\mathbb{R}^n) \to \mathbb{C}$ is known to be linear, it is enough to see that, if $\varphi_m \to 0$ in $\mathcal{D}(\mathbb{R}^n)$, then $\langle f, \varphi_m \rangle \to 0$ in $\mathbb{C}$ in order for f to be continuous, and thus a distribution.*

Proof. Let $\varphi_m \to \varphi$ be a convergent sequence in $\mathcal{D}(\mathbb{R}^n)$. Then, by definition, $\{\varphi_m - \varphi\} \to 0$ in $\mathcal{D}$. As f holds the property of continuity for sequences which converge to zero, we can write $\langle f, \varphi_m - \varphi \rangle \to 0$. Because of linearity, $\langle f, \varphi_m - \varphi \rangle = \langle f, \varphi_m \rangle - \langle f, \varphi \rangle$, and we can assert that $\langle f, \varphi_m \rangle \to \langle f, \varphi \rangle$. □

1.2.2 Examples

After defining what a distribution is, it will be useful to present some examples of distributions that will be useful later on.

Example 1.12. We say a function on $\mathbb{R}^n$ is locally integrable if it is integrable on every compact subset of $\mathbb{R}^n$. In this situation, **every locally integrable function can be treated as a distribution**. In fact, let f be a locally integrable function on $\mathbb{R}^n$. Then, for any $\varphi \in \mathcal{D}(\mathbb{R}^n)$, we can define the integral

$$\langle f, \varphi \rangle = \int_{\mathbb{R}^n} f(x)\varphi(x)dx. \tag{1.2}$$

Observe that if $K = \operatorname{supp}\varphi$, then $\langle f, \varphi \rangle = \int_K f(x)\varphi(x)dx$, and as K is compact and φ is infinitely smooth, $f\varphi$ is integrable on K, and therefore on $\mathbb{R}^n$.

So once seen the functional is well-defined, let us prove it is a distribution. The linearity of the integral makes it to be linear. Thus, we must check it is continuous. By Proposition 1.11, we only have to consider the situation of sequences which converge to zero. So let $\varphi_n \to 0$ in $\mathcal{D}$. We have to see that

$$\lim_{n\to\infty} \langle f, \varphi_n \rangle = \lim_{n\to\infty} \int_{\mathbb{R}^n} f(x)\varphi_n(x)dx = 0.$$

The convergence in $\mathcal{D}$ implies that the sequence $\{\varphi_n(t)\}$ converges to zero uniformly, and also that there exists a compact $K \subset \mathbb{R}^n$ in which the support of every φ_n is contained. In other words,

$$\forall \epsilon > 0, \quad \exists N \in \mathbb{N} \quad | \quad \forall n > N, \quad |\varphi_n(t)| < \epsilon, \quad \forall t \in \mathbb{R}^n,$$

and

$$\int_{\mathbb{R}^n} f(x)\varphi_n(x)dx = \int_K f(x)\varphi_n(x)dx.$$

So $\forall n > N$ we get

$$\left| \int_K f(x)\varphi_n(x)dx \right| \leq \int_K |f(x)||\varphi_n(x)|dx \leq \epsilon \int_K |f(x)|dx$$

and as $\int_K |f(x)|dx$ is finite, $\lim_{n\to\infty} \int_{\mathbb{R}^n} f(x)\varphi_n(x)dx = 0$.

Therefore, making the identification (1.2), we can say that $f \in \mathcal{D}'(\mathbb{R}^n)$. Some authors also write T_f to denote this functional. This kind of distributions are called **regular distributions**.

The question that arises is whether all distributions are regular. Fortunately, there are many more. This fact is the reason to call distributions generalised functions, as they take in more objects than usual functions. The next two examples illustrate this fact.

Example 1.13. Let φ be a testing function in $\mathbb{R}^n$. The functional $\delta : \mathcal{D}(\mathbb{R}^n) \to \mathbb{C}$ given by

$$\langle \delta, \varphi \rangle = \varphi(0) \tag{1.3}$$

and called **Dirac delta function** is a distribution, and it plays a very important role in many fields not only in mathematics, but also in physics.

It is straightforward to see that the delta function is linear, as for every $\varphi, \phi \in \mathcal{D}$ and $a, b \in \mathbb{C}$,

$$\langle \delta, a\varphi + b\phi \rangle = a\varphi(0) + b\phi(0) = a\langle \delta, \varphi \rangle + b\langle \delta, \phi \rangle.$$

To see it is continuous, let $\varphi_n \to 0$ be a convergent sequence in $\mathcal{D}$. Observe that $\langle \delta, \varphi_n \rangle = \varphi_n(0)$, so we have to prove that $\varphi_n(0) \to 0$ in $\mathbb{C}$. But convergence in $\mathcal{D}$ implies uniform convergence, what means that $\lim_{n\to\infty} ||\varphi_n||_\infty = 0$, where $||\cdot||_\infty$ stands for the supremum norm. Therefore, $\lim_{n\to\infty} |\varphi_n(0)| = 0$.

Now that we have seen that the delta function (1.3) defines a distribution, we want to see that it is not regular. Let us check this fact for $\mathbb{R}$. By way of contradiction, let f be a locally integrable function such that

$$\int_{\mathbb{R}} f(x)\varphi(x)dx = \varphi(0), \quad \forall \varphi \in \mathcal{D}(\mathbb{R}).$$

Consider the testing function (1.1), and define

$$\zeta_n(t) = \zeta(nt), \quad \forall n \in \mathbb{N}.$$

It holds for every $n \in \mathbb{N}$ that $\zeta_n(0) = \zeta(0) = e^{-1}$, and $\operatorname{supp} \zeta_n = [-\frac{1}{n}, \frac{1}{n}]$. So we can write

$$\int_{\mathbb{R}} f(t)\zeta_n(t)dt = \int_{-1/n}^{1/n} f(t)\zeta_n(t)dt = e^{-1}.$$

If we take limits on both sides, we get

$$\lim_{n\to\infty} \int_{-1/n}^{1/n} f(t)\zeta_n(t)dt = e^{-1}. \tag{1.4}$$

By the definition of ζ, it is clear that $\operatorname{Im} \zeta = [0, e^{-1}]$, so every ζ_n is bounded by e^{-1}. Now observe that

$$\int_{-1/n}^{1/n} f(t)\zeta_n(t)dt = \int_{\mathbb{R}} \chi_{[-\frac{1}{n},\frac{1}{n}]}(t)f(t)\zeta_n(t)dt,$$

and $|\chi_{[-\frac{1}{n},\frac{1}{n}]}(t)f(t)\zeta_n(t)| \leq e^{-1}\chi_{[-1,1]}(t)|f(t)|$, so we can bound the integrand by an integrable function. Thus, we can take the limit in (1.4) inside by the dominated convergence theorem, and

$$\lim_{n\to\infty} \int_{-1/n}^{1/n} f(t)\zeta_n(t)dt = \int_{\mathbb{R}} \lim_{n\to\infty} \chi_{[-\frac{1}{n},\frac{1}{n}]}(t)f(t)\zeta_n(t)dt = 0.$$

This is a contradiction with (1.4), so the delta function is not a regular distribution.

Distributions which are not regular are called **singular distributions**.

Example 1.14. In general, we cannot define a distribution through the expression (1.2) if the function we choose is not locally integrable. This is the case of $f(t) = 1/t$, as it is not integrable around the origin in $\mathbb{R}$. We can think of a similar definition anyway. Let $T : \mathcal{D}(\mathbb{R}) \to \mathbb{R}$ be defined the following way:

$$\langle T, \varphi \rangle = \lim_{\epsilon\to 0} \int_{|x|\geq\epsilon} \frac{\varphi(x)}{x} dx. \tag{1.5}$$

This expression defines a distribution, and it is called **Cauchy principal value of $\boldsymbol{1/x}$**. We denote it by $T = \operatorname{Pv} \frac{1}{x}$. In order to prove it is a distribution, we must ensure the limit exists. For that, a change of variable allows us to write

$$\int_{|x|\geq\epsilon} \frac{\varphi(x)}{x} dx = \int_{-\infty}^{-\epsilon} \frac{\varphi(x)}{x} dx + \int_{\epsilon}^{\infty} \frac{\varphi(x)}{x} dx = \int_{\epsilon}^{\infty} \frac{\varphi(x) - \varphi(-x)}{x} dx$$

Now, we know that φ has compact support, so we can find a positive constant k such that $\operatorname{supp} \varphi \subseteq [-k, k]$, and thus the limit (1.5) turns into

$$\lim_{\epsilon\to 0} \int_{\epsilon}^{k} \frac{\varphi(x) - \varphi(-x)}{x} dx.$$

By setting χ for the characteristic function,

$$\langle \mathrm{Pv}\,\frac{1}{x}, \varphi\rangle = \lim_{n\to\infty} \int_0^\infty \chi_{[\frac{1}{n},k]}(x)\frac{\varphi(x)-\varphi(-x)}{x}dx.$$

We also know that $\varphi \in C^\infty(\mathbb{R})$, so by the mean value theorem we know that for any $x > 0$ we can find $\xi \in (-x, x)$ such that $\varphi(x) - \varphi(-x) = 2x\varphi'(\xi)$. Moreover, $\varphi \in \mathcal{D}$ implies that $\varphi' \in \mathcal{D}$, so it is bounded and thus

$$\frac{|\varphi(x)-\varphi(-x)|}{x} \le 2||\varphi'||_\infty.$$

This allows us to write

$$\left|\chi_{[\frac{1}{n},k]}(x)\frac{\varphi(x)-\varphi(-x)}{x}\right| \le 2||\varphi'||_\infty \chi_{[0,k]}(x)$$

and the integrand remains bounded by an integrable function on $\mathbb{R}$. By the dominated convergence theorem, limit and integral may be interchanged and

$$\langle \mathrm{Pv}\,\frac{1}{x}, \varphi\rangle = \int_0^\infty \lim_{n\to\infty} \chi_{[\frac{1}{n},k]}(x)\frac{\varphi(x)-\varphi(-x)}{x}dx = \int_0^k \frac{\varphi(x)-\varphi(-x)}{x}dx. \qquad (1.6)$$

Finally, the integral is finite because the singularity at $x = 0$ represents a removable discontinuity. Indeed, L'Hôpital's rule leads to

$$\lim_{x\to 0} \frac{\varphi(x)-\varphi(-x)}{x} = 2\varphi'(0).$$

After seeing the functional is well defined, linearity is obtained using expression (1.6) and the linearity of the integral. To check continuity, let $\{\varphi_n\}$ converge to zero in $\mathcal{D}$. We know that their supports are contained in a compact subset $K \subset \mathbb{R}$. Choose $k > 0$ big enough so that $K \subset [-k, k]$. Therefore,

$$\langle \mathrm{Pv}\,\frac{1}{x}, \varphi_n\rangle = \int_0^k \frac{\varphi_n(x)-\varphi_n(-x)}{x}dx, \quad \forall n \in \mathbb{N}.$$

Again, by the same reasoning as before,

$$\frac{\varphi_n(x)-\varphi_n(-x)}{x} \le 2||\varphi_n'||_\infty.$$

Observe also that convergence in $\mathcal{D}$ implies that the sequence $\{\varphi_n'\}$ converges uniformly to zero. As every φ_n' is bounded, then there exists $M > 0$ such that $||\varphi_n'||_\infty < M, \forall n \in \mathbb{N}$. This way, the dominated convergence theorem allows us to write

$$\lim_{n\to\infty} \left|\langle \mathrm{Pv}\,\frac{1}{x}, \varphi_n\rangle\right| \le \int_0^k \lim_{n\to\infty} \frac{|\varphi_n(x)-\varphi_n(-x)|}{x}dx = 0.$$

So eventually, $\lim_{n\to\infty} \left|\langle \mathrm{Pv}\,\frac{1}{x}, \varphi_n\rangle\right| = 0$ and the Cauchy principal value is continuous. As a consequence, the Cauchy principal value is a distribution.

1.2.3 Convergence

The space of distributions can be given a concept of convergence, which in some cases will be useful.

Definition 1.15. Let $\{f_n\}$ be a sequence of distributions. We say that the sequence converges to the distribution f if for every testing function φ, the sequence $\langle f_n, \varphi\rangle$ converges to $\langle f, \varphi\rangle$.

We also say that the sequence $\{f_n\}$ simply converges if every sequence $\langle f_n, \varphi\rangle$ converges.

It can be proven that if a sequence of distributions converges, then the functional f which assigns to every testing function the limit value of the sequence defines a distribution. A proof can be consulted in [10, p. 37-39].

1.2.4 Operations on distributions

Once we have analysed several examples of distributions, it is time to define some operations on them. First, we will define the most simple ones, which will allow to determine an algebraic structure over $\mathcal{D}'(\mathbb{R}^n)$. So let $f, g \in \mathcal{D}'(\mathbb{R}^n)$ and $\alpha \in \mathbb{C}$. We define:

1. The **sum of distributions**, $f + g$, as

$$\langle f + g, \varphi\rangle = \langle f, \varphi\rangle + \langle g, \varphi\rangle, \quad \forall\varphi \in \mathcal{D}.$$

2. **The multiplication by a constant**, αf, as

$$\langle \alpha f, \varphi\rangle = \alpha\langle f, \varphi\rangle, \quad \forall\varphi \in \mathcal{D}.$$

It is straightforward to check that these two functionals define distributions. With them, we have defined an internal operation in $\mathcal{D}'$ and an external product over $\mathbb{C}$. Again, for $\mathbb{C}$ being a field, **$\mathcal{D}'(\mathbb{R}^n)$ is a linear space**.

These first two operations are trivial to define and they stand naturally. There are some others that also arise naturally if we work with regular distributions, and we will be able to generalise them to even singular distributions. So let $\tau \in \mathbb{R}^n$ and $a \in \mathbb{R} - \{0\}$. We define:

3. The **shifting of a distribution**, $f(t - \tau)$, as

$$\langle f(t - \tau), \varphi\rangle = \langle f, \varphi(t + \tau)\rangle, \quad \forall\varphi \in \mathcal{D}.$$

This operator is well-defined, as we saw in Proposition 1.7 that a shifted testing function is a testing function. It is fairly simple to justify this definition. Indeed, if f is a regular distribution, then changing variables,

$$\begin{aligned}\langle f(t - \tau), \varphi(t)\rangle &= \int_{\mathbb{R}^n} f(t - \tau)\varphi(t)dt = \int_{\mathbb{R}^n} f(y)\varphi(y + \tau)dy \\ &= \langle f(t), \varphi(t + \tau)\rangle.\end{aligned}$$

4. The **transposition of a distribution**, $f(-t)$, as

$$\langle f(-t), \varphi \rangle = \langle f, \varphi(-t) \rangle, \quad \forall \varphi \in \mathcal{D}.$$

This is also a well-defined functional, because $\varphi(-t)$ is a testing function, as we checked in Proposition 1.7. The justification is again given in terms of regular distributions; if f is so, then

$$\langle f(-t), \varphi(t) \rangle = \int_{\mathbb{R}^n} f(-t)\varphi(t)dt = \int_{\mathbb{R}^n} f(y)\varphi(-y)dy = \langle f(t), \varphi(-t) \rangle.$$

5. The **dilation of a distribution**, $f(at)$, as

$$\langle f(at), \varphi \rangle = \langle f, \frac{1}{|a|^n}\varphi\left(\frac{t}{a}\right) \rangle, \quad \forall \varphi \in \mathcal{D}.$$

Proposition 1.7 ensures that this functional is well-defined. As in the previous definitions, if f is a regular distribution,

$$\begin{aligned}\langle f(at), \varphi(t) \rangle &= \int_{\mathbb{R}^n} f(at)\varphi(t)dt = \int_{\mathbb{R}^n} f(y)\varphi\left(\frac{y}{a}\right)\frac{1}{|a|^n}dy \\ &= \langle f(t), \frac{1}{|a|^n}\varphi\left(\frac{t}{a}\right) \rangle.\end{aligned}$$

Apart from seeing that these new functionals are well-defined, to check they define distributions is an easy exercise. Linearity is granted by that of f. To see they are continuous, let $\{\varphi_n\}$ be convergent to zero in $\mathcal{D}$. If we prove that the sequences $\{\varphi_n(t+x)\}, \{\varphi_n(-t)\}, \{\varphi_n\left(\frac{t}{a}\right)\}$ converge to zero in $\mathcal{D}$, we will get the desired result automatically by the continuity of f. Now, if $K \subset \mathbb{R}^n$ is the domain containing every $\operatorname{supp}\varphi_n$, then a translation, a transposition or a dilation of K contains every support of the elements of the new sequences. Furthermore, every derivative of the new functions is a translation, transposition or a dilation of the derivative of the original functions, without taking constant factors into account. This implies that the suprema of the new derivatives are the same as the ones of the original derivatives (avoiding amplification constants), making them converge uniformly to zero. Thus, the new sequences converge to zero in $\mathcal{D}$.

There is one more operation that we can define. Let $g \in C^\infty$. We define:

6. The **multiplication of a distribution by a smooth function**, gf, as

$$\langle gf, \varphi \rangle = \langle f, g\varphi \rangle, \quad \forall \varphi \in \mathcal{D}.$$

This is a completely trivial definition. Indeed, if we consider f to be a regular distribution, then as g is infinitely smooth, the product gf is also locally integrable, and it defines a regular distribution,

$$\langle gf, \varphi \rangle = \int_{\mathbb{R}^n} g(t)f(t)\varphi(t)dt = \int_{\mathbb{R}^n} f(t)(g(t)\varphi(t))dt = \langle f, g\varphi \rangle,$$

as Proposition 1.7 asserts that $g\varphi \in \mathcal{D}$. The general proof that gf is indeed a distribution is obtained by very similar arguments as in previous cases.

A particularly interesting case is the one concerning the delta function. In fact,

$$\langle g\delta, \varphi\rangle = \langle \delta, g\varphi\rangle = g(0)\varphi(0) = g(0)\langle\delta, \varphi\rangle,$$

so eventually,

$$g\delta = g(0)\delta, \quad \forall g \in C^{\infty}(\mathbb{R}^n). \tag{1.7}$$

After being able to define a product of a distribution with a certain kind of functions, we wonder whether a general product between distributions can be defined. Unfortunately, this is not the case. Think of the function $f(t) = 1/(\sqrt{|t|})$. This is a locally integrable function on $\mathbb{R}$, thus defines a distribution. But $f^2(t) = 1/|t|$, which is known not to be integrable on neighbourhoods of the origin. Hence, we cannot define a distribution through the expression

$$\langle \frac{1}{|t|}, \varphi\rangle = \int_{\mathbb{R}} \frac{\varphi(t)}{|t|} dt,$$

because the integral will not exist for every testing function φ. So in general, **we will not be able to define the product of distributions**. This is one of the major drawbacks, if not the principal, of the theory of distributions, although many attempts to fix it have been carried.

Up to now, we have defined what we can call basic operations in the space $\mathcal{D}'$. But it is possible to define much more powerful operations which will play an important role in reaching our ultimate objective. Differentiation will be discussed next, and some others as convolution and the direct product will be treated in Chapter 2.

1.2.5 Differentiation of distributions

Distributions have many convenient properties which do not hold for usual functions, making them very useful, for they can reach results we would not be able to explore considering only standard functions. Differentiation is one of them. In fact, even if a locally integrable function may not be differentiable at some points, it is possible to define derivatives of every order for distributions. So let us define the distribution differentiation process.

Definition 1.16. Let f be a distribution in $\mathcal{D}'(\mathbb{R}^n)$. The **partial derivatives** of f are defined as

$$\langle \frac{\partial f}{\partial t_i}, \varphi\rangle = \langle f, -\frac{\partial \varphi}{\partial t_i}\rangle, \quad \forall \varphi \in \mathcal{D}(\mathbb{R}^n),$$

for every $i \in \{1, \cdots, n\}$.

As most of the properties for distributions, this definition has its origins in the behaviour of regular distributions. To see this, let us consider a regular distribution

$f \in \mathcal{D}'(\mathbb{R})$, for which the function f is smooth. Then, its derivative f' is also smooth, so it defines a distribution and the integration by parts yields

$$\langle f', \varphi \rangle = \int_{\mathbb{R}} f'(t)\varphi(t)dt = \int_{\mathbb{R}} (f\varphi)'(t)dt - \int_{\mathbb{R}} f(t)\varphi'(t)dt.$$

But as φ is a testing function, its support is compact, and thus the fundamental theorem of calculus forces the first integral to be zero. Because of this, we can write

$$\langle f', \varphi \rangle = -\langle f, \varphi' \rangle, \quad \forall \varphi \in \mathcal{D}.$$

which is the one dimensional version of Definition 1.16.

We can expect the derivative of a distribution to be a distribution too. This is in fact true, as it will be shown in the next proposition.

Proposition 1.17. *Let f be a distribution. Then, the expression for $\frac{\partial f}{\partial t_i}$ in Definition 1.16 defines a distribution.*

Proof. The functional is well-defined, Lemma 1.3 showing that every partial derivative of a testing function is also a testing function. Its linearity is provided by that of the derivative. Now, take a sequence of testing functions $\{\varphi_n\}$ which converges to zero in $\mathcal{D}$. Then, as seen in the proof of Lemma 1.3, the support of the derivative of a testing function is contained in the support of the testing function itself, so as there exists a subset $K \subset \mathbb{R}^n$ containing every $\operatorname{supp}\varphi_n$, it also contains every $\operatorname{supp}\frac{\partial \varphi_n}{\partial t_i}$. Moreover, we know that every $\{D^\alpha \varphi_n\}_{n=1}^\infty$ converges to zero uniformly, α being a multi-index. So we are allowed to say that the sequence $\{\frac{\partial \varphi_n}{\partial t_i}\}$ converges to zero in $\mathcal{D}$. Finally, as f is a distribution,

$$\langle \frac{\partial f}{\partial t_i}, \varphi_n \rangle = -\langle f, \frac{\partial \varphi_n}{\partial t_i} \rangle$$

converges to zero in $\mathbb{C}$. This shows that the functional is continuous, and thus a distribution. □

Once we have fixed the definition of the derivatives of distributions, we are ready to analyse some properties, which in some cases will enhance the performance of usual functions. For example, recall Schwarz' theorem, which allows to commute partial derivatives of a twice differentiable function. Not only is this true for distributions, but it also holds with no restrictions at all.

Theorem 1.18. *Let f be a distribution. Then, for every $i, j \in \{1, \cdots, n\}$,*

$$\frac{\partial^2 f}{\partial t_i \partial t_j} = \frac{\partial^2 f}{\partial t_j \partial t_i}$$

Proof. It is an immediate consequence of the fact that testing functions are infinitely smooth, thus Schwarz' theorem holds for them. □

It is also interesting to note that the product rule also holds for distributions.

Proposition 1.19. *Let $f \in \mathcal{D}'(\mathbb{R}^n)$ and $g \in C^\infty(\mathbb{R}^n)$. Then, for any $i \in \{1, \cdots, n\}$,*

$$\frac{\partial}{\partial t_i}(gf) = \frac{\partial g}{\partial t_i} f + g \frac{\partial f}{\partial t_i}.$$

Proof. Let φ be a testing function in $\mathbb{R}^n$. Then,

$$\langle \frac{\partial (gf)}{\partial t_i}, \varphi \rangle = -\langle f, g\frac{\partial \varphi}{\partial t_i} \rangle = -\langle f, \frac{\partial}{\partial t_i}(g\varphi) \rangle + \langle f, \frac{\partial g}{\partial t_i}\varphi \rangle = \langle g\frac{\partial f}{\partial t_i}, \varphi \rangle + \langle \frac{\partial g}{\partial t_i} f, \varphi \rangle. \qquad \square$$

The derivative of distributions, by the property seen in Proposition 1.17, can be seen as an operator in the space of distributions. In some moments, it will be useful to know it is linear and continuous.

Proposition 1.20. *Differentiation is a linear operation in the space of distributions, and it is continuous in the sense that if a sequence of distributions f_n converges to f in $\mathcal{D}'$, then $D^k f_n$ converges to $D^k f$ in $\mathcal{D}'$.*

Proof. Linearity is trivial. For continuity, Proposition 1.17 and convergence of f_n implies that

$$\langle D^k f_n, \varphi \rangle = (-1)^{|k|} \langle f_n, D^k \varphi \rangle \to (-1)^{|k|} \langle f, D^k \varphi \rangle = \langle D^k f, \varphi \rangle. \qquad \square$$

An interesting question is what will the derivatives of some particular distributions look like. We have seen that if we are dealing with regular distributions which come from a differentiable function, the usual derivative coincides with the distributional derivative, but we know nothing about the remaining ones, such as the delta function.

Example 1.21 (Delta function). We know that for every $\varphi \in \mathcal{D}(\mathbb{R}^n)$,

$$\langle \delta, \varphi \rangle = \varphi(0),$$

so if we write the derivative with respect to variable $i \in \{1, \cdots, n\}$, we get

$$\langle \frac{\partial \delta}{\partial t_i}, \varphi \rangle = -\langle \delta, \frac{\partial \varphi}{\partial t_i} \rangle = -\frac{\partial \varphi}{\partial t_i}(0).$$

In general, from the above we immediately get that if $k = (k_1, \cdots, k_n) \in (\mathbb{Z}^+)^n$, then

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

Another interesting example is the Heaviside function, whose derivative will be the delta function.

Example 1.22 (Heaviside function). We define the **Heaviside function** as the characteristic function of the interval $[0, +\infty] \subset \mathbb{R}$. It is usually denoted by $H(t)$, but also as $1_+(t)$ or $\theta(t)$. If we compute its derivative, for $\varphi \in \mathcal{D}(\mathbb{R})$,

$$\langle H'(t), \varphi \rangle = -\langle H(t), \varphi' \rangle = -\int_{\mathbb{R}} H(t)\varphi'(t)dt = -\int_0^\infty \varphi'(t)dt.$$

Now, as φ has compact support, we can find $k > 0$ such that the support is contained in $[-k, k]$, and

$$-\int_0^\infty \varphi'(t)dt = -\int_0^k \varphi'(t)dt = -\varphi(k) + \varphi(0) = \varphi(0) = \langle \delta, \varphi \rangle.$$

Therefore, in terms of distributions, we get

$$H'(t) = \delta(t). \tag{1.8}$$

Notice that if we consider H as a function, $H'(t) = 0$ holds almost everywhere. This example shows that, in general, the distributional derivative of a function does not coincide with the distribution associated to its derivative if the latter exists only almost everywhere. Therefore, the derivative of a function must exist in every point of its domain in order to ensure the distributional derivative coincides with the usual derivative.

Chapter 2

Tempered Distributions

In this chapter we will introduce a new space of distributions, the space of tempered distributions. Its importance will be covered in the next chapter, as the main motivation to define it is to be able to define the Fourier transform for distributions.

The space of tempered distributions, which we call $\mathcal{S}'$, is a proper subspace of the space $\mathcal{D}'$ we already know. It can be defined, as we are to show, through a wider class of testing fuctions.

2.1 The Schwartz space $\mathcal{S}$

As we have proposed, we want to introduce a space of functions containing the space of testing functions described in Section 1.1.

Definition 2.1. Let $\phi : \mathbb{R}^n \to \mathbb{C}$ be a function. We say that ϕ is a **Schwartz function** if

1. ϕ is $\mathbb{C}^\infty$, and
2. for every $m \in \mathbb{Z}^+, k \in (\mathbb{Z}^+)^n$, there exists a constant $C_{m,k} > 0$ such that

$$|t|^m |D^k \phi(t)| \leq C_{m,k}, \quad \forall t \in \mathbb{R}^n.$$

The space of all Schwartz functions is called **Schwartz space**, and it is denoted by $\mathcal{S}(\mathbb{R}^n)$.

Remark 2.2. The Schwartz functions are also called **functions of rapid descent**. The reason for this is that Condition 2 in Definition 2.1 is equivalent to

$$|D^k \phi(t)| \leq \frac{C_{m,k}}{|t|^m}.$$

This way, every Schwartz function and its derivatives must decrease to zero as $|t|$ tends to infinity, not slower than any power of $1/|t|$.

Once we have defined the Schwartz space, it is time to present some very basic properties.

Proposition 2.3. *Let $\phi \in \mathcal{S}(\mathbb{R}^n)$ and $k \in (\mathbb{Z}^+)^n$ a multi-index. Then,*

1. *$\mathcal{S}(\mathbb{R}^n)$ is a linear space.*

2. *$D^k\phi \in \mathcal{S}(\mathbb{R}^n)$.*

Proof. By the triangle inequality, it is trivial to see that the sum of two Schwartz functions and the product of a Schwartz function with a complex constant are again Schwartz functions. This way, as every operation is performed in a field, $\mathcal{S}(\mathbb{R}^n)$ is a linear space. Finally, to see 2 holds it is enough to check that for every $i \in \{1, \cdots, n\}$, $\frac{\partial\phi}{\partial t_i} \in \mathcal{S}(\mathbb{R}^n)$. Again, the fact that ϕ is $\mathbb{C}^\infty$ implies that the partial derivative is so. For the second condition, we must see that for every $m \in \mathbb{Z}^+$ and $k' \in (\mathbb{Z}^+)^n$, there exists $C'_{m,k} > 0$ such that

$$|t|^m \left| D^{k'} \frac{\partial\phi}{\partial t_i} \right| \leq C'_{m,k'}, \quad \forall t \in \mathbb{R}^n.$$

Now, if $k' = (k_1, \cdots, k_n)$, and we call $k = (k_1, \cdots, k_i + 1, \cdots, k_n)$, we can write

$$|t|^m \left| D^{k'} \frac{\partial\phi}{\partial t_i} \right| = |t|^m |D^k\phi| \leq C_{m,k}, \quad \forall t \in \mathbb{R}^n,$$

so if $C_{m,k}$ is the constant corresponding to ϕ and m, k, then $C'_{m,k'} = C_{m,k}$, which completes the proof. □

In the beginning of the section, we have claimed our intention to define a space wider than the space of testing functions. Let us state this fact in the following proposition.

Theorem 2.4. *If $\mathcal{D}(\mathbb{R}^n)$ is the space of testing functions and $\mathcal{S}(\mathbb{R}^n)$ is the Schwartz space, then*

$$\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n),$$

the inclusion being strict.

Proof. Let $\phi \in \mathcal{D}$ be a usual testing function. Then, as it is C^∞, we only have to check the second condition in Definition 2.1. So let $m \in Z^+$ and $k \in (\mathbb{Z}^+)^n$. The objective is to see that $|t|^m |D^k\phi(t)|$ is bounded. We know that $\operatorname{supp}\phi$ is bounded, so there exists $r > 0$ such that $\operatorname{supp}\phi \subset \overline{B}(r)$, where $\overline{B}(r)$ denotes the closed ball centered in the origin and of radius r. We also proved in Lemma 1.3 that $\operatorname{supp} D^k\phi \subset \operatorname{supp}\phi$. So we analyse two different situations:

- If $t \notin \overline{B}(r)$, then $D^k\phi(t) = 0$, and thus $|t|^m |D^k\phi(t)| = 0$.

- If $t \in \overline{B}(r)$, then we are considering a C^∞ function in a compact domain. This implies that $D^k\phi(t)$ is bounded in $\overline{B}(r)$, so there exists $K > 0$ such that

$$|t|^m |D^k\phi(t)| \leq r^m K < \infty.$$

So we see that $|t|^m|D^k\phi(t)|$ is bounded, what shows that $\phi \in \mathcal{S}$, and thus $\mathcal{D} \subseteq \mathcal{S}$. We want to prove that indeed $\mathcal{D} \subset \mathcal{S}$. So let ϕ be

$$\phi(t) = e^{-|t|^2}. \tag{2.1}$$

We know that $\phi \in C^\infty$. Besides, it is clear that its support is not bounded, as $\phi(t) \neq 0, \forall t \in \mathbb{R}^n$. Thus, $\phi \notin \mathcal{D}$. The question is whether $\phi \in \mathcal{S}$. It is easy to see that for every $k \in (\mathbb{Z}^+)^n$, $D^k\phi(t) = P_k(t)\phi(t)$, where $P_k(t)$ represents a polynomial. So, we can write

$$|t|^m|D^k\phi(t)| = \frac{|t|^m|P_k(t)|}{e^{|t|^2}}, \tag{2.2}$$

expression that tends to zero as $|t|$ tends to infinity. This means that for every $\epsilon > 0$, we can find a constant $N > 0$ such that if $|t| > N$,

$$\frac{|t|^m|P_k(t)|}{e^{|t|^2}} < \epsilon.$$

Thus, if we consider $|t| \leq N$, the expression (2.2) is bounded for being a C^∞ function on a compact domain, say by a constant $M > 0$. Therefore, consider $C_{m,k} = \max\{M, \epsilon\}$, and so,

$$|t|^m|D^k\phi(t)| \leq C_{m,k},$$

which shows that $\phi \in \mathcal{S}$. $\square$

Knowing that continuity plays an important role in distributions, next step is to describe **convergence in the space $\mathcal{S}$**.

Definition 2.5. Let $\{\phi_n(t)\}_{n=1}^\infty$ be a sequence in the Schwartz space $\mathcal{S}$. We say that the sequence **converges** to zero in $\mathcal{S}$ if for every $m \in \mathbb{Z}^+, k \in (\mathbb{Z}^+)^n$, the sequence $\{|t|^m D^k\phi_n(t)\}_{n=1}^\infty$ converges to zero uniformly.

Following this definition, we say that the sequence $\{\phi_n(t)\}_{n=1}^\infty$ converges to ϕ if the sequence $\{\phi_n(t) - \phi(t)\}_{n=1}^\infty$ converges to zero.

Remark 2.6. As in the case of $\mathcal{D}$, for simplicity, the sequence $\{\phi_n(t)\}_{n=1}^\infty$ will also be denoted ϕ_n, and convergence to the function ϕ will be written as $\phi_n \to \phi$.

We have to take special care with the relationship between convergences in $\mathcal{D}$ and in $\mathcal{S}$. Suppose we are working with a sequence in $\mathcal{S}$. Then, this sequence may not be in $\mathcal{D}$, so it makes no sense to ask about convergence in $\mathcal{D}$. On the other hand, every sequence in $\mathcal{D}$ is a sequence in $\mathcal{S}$, so in order to keep a coherent structure, we will have to ensure that convergence in $\mathcal{D}$ implies convergence in $\mathcal{S}$. That is what we present in the following proposition.

Proposition 2.7. *Let $\{\phi_n\}$ be a convergent sequence in $\mathcal{D}$. Then, it is a convergent sequence in $\mathcal{S}$.*

Proof. Remember that as $\phi_n \in \mathcal{D} \subset \mathcal{S}$, we only have to prove that the sequences $\{|t|^m D^k \phi_n(t)\}$ converge uniformly, for every choice of $m \in \mathbb{Z}^+$ and $k \in (\mathbb{Z}^+)^n$. Also observe that as the general concept of convergence depends on convergence to zero, it is enough to check this case.

So let $\{\phi_n\}$ be a sequence convergent to zero in $\mathcal{D}$. We know that there exists a bounded set $K \subset \mathbb{R}^n$ which contains the support of every ϕ_n, and thus, we can find $T > 0$ such that $K \subset \overline{B}(T)$. Therefore, we can write

$$|t|^m |D^k \phi_n(t)| \leq T^m \sup_{|t| \leq T} |D^k \phi_n(t)|, \quad \forall t \in \mathbb{R}^n, \tag{2.3}$$

as the derivatives are zero outside K. Now, the convergence in $\mathcal{D}$ implies that

$$||D^k \phi_n||_\infty \to 0,$$

so the right hand side of (2.3) decays to zero as $n \to \infty$. As a consequence,

$$\sup_{\mathbb{R}^n} |t|^m |D^k \phi_n(t)| \to 0, \quad \text{as } n \to \infty.$$

This means that the sequence $\{|t|^m D^k \phi_n(t)\}$ converges uniformly to zero. The proof is now complete. □

Granted that the definition of the Schwartz space is consistent with everything we have defined before, we will introduce some operations on $\mathcal{S}$, in a similar way as in Proposition 1.7. But before that, we will present a little lemma, which will be useful both to prove the following proposition and to get some results in next sections.

Lemma 2.8. *Let $a, b > 0$ and $m \in \mathbb{N}$. Then,*

$$(a + b)^m \leq 2^m (a^m + b^m).$$

Proof. We can prove it directly if we observe that

$$(a + b)^m \leq (2 \max\{a, b\})^m = 2^m (\max\{a, b\})^m \leq 2^m (a^m + b^m). \qquad \square$$

Proposition 2.9. *Let $\phi \in \mathcal{S}(\mathbb{R}^n)$, $\tau \in \mathbb{R}^n$, $a \in \mathbb{R} - \{0\}$ and $p(t)$ a polynomial in $\mathbb{R}^n$. Then,*

1. *$\phi(t + \tau), \phi(-t), \phi(at) \in \mathcal{S}(\mathbb{R}^n)$.*
2. *$p(t)\phi(t) \in \mathcal{S}(\mathbb{R}^n)$.*

Proof. All of the newly defined functions are infinitely smooth for ϕ and p being so. Now, for each of the cases, and for every $m \in \mathbb{Z}^+, k \in (\mathbb{Z}^+)^n$, if we consider the constants $C_{m,k}$ for ϕ,

- Using the triangle inequality and Lemma 2.8, for $\phi(t + \tau)$,

$$\begin{aligned} |t|^m |D^k \phi(t + \tau)| &\leq (|t + \tau| + |\tau|)^m |D^k \phi(t + \tau)| \\ &\leq 2^m (|t + \tau|^m |D^k \phi(t + \tau)| + |\tau|^m |D^k \phi(t + \tau)|) \\ &\leq 2^m (C_{m,k} + |\tau|^m C_{0,k}) \end{aligned}$$

- For $\varphi(t) = \phi(-t)$,

$$\begin{aligned}|t|^m |D^k \varphi(t)| &= |t|^m |(-1)^{|k|} D^k \phi(-t)| = |-t|^m |D^k \phi(-t)| \\ &\leq C_{m,k}.\end{aligned}$$

- For $\varphi(t) = \phi(at)$,

$$\begin{aligned}|t|^m |D^k \varphi(t)| &= |t|^m |a^{|k|} D^k \phi(at)| = |a|^{|k|-m} |at|^m |D^k \phi(at)| \\ &\leq |a|^{|k|-m} C_{m,k}.\end{aligned}$$

- For $p(t)\phi(t)$, because of Proposition 2.3, it is enough to show $t_i \phi(t) \in \mathcal{S}, \forall i \in \{1, \cdots, n\}$. So, as the variable t_i only has influence in derivatives with respect to variable t_i, it is easy to see that

$$\frac{\partial^k}{\partial t_i^k}(t_i \phi(t)) = k \frac{\partial^{k-1}}{\partial t_i^{k-1}} \phi(t) + t_i \frac{\partial^k}{\partial t_i^k} \phi(t),$$

and therefore, if $k = (k_1, \cdots, k_i, \cdots, k_n)$, and $k' = (k_1, \cdots, k_i - 1, \cdots, k_n)$, then

$$D^k(t_i \phi(t)) = k_i D^{k'} \phi(t) + t_i D^k \phi(t),$$

and it is straightforward that the constant we need to bound $|t|^m |D^k(t_i \phi(t))|$ is $k_i C_{m,k'} + C_{m,k}$. □

There is one more property that plays a very important role when working with Schwartz functions.

Theorem 2.10. *Let $p \in [1, \infty]$. Then, $\mathcal{S}(\mathbb{R}^n) \subseteq L^p(\mathbb{R}^n)$.*

Proof. Let ϕ be a Schwartz function. We know that it is infinitely smooth, and for every $m \in \mathbb{Z}^+$ and $k \in (\mathbb{Z}^+)^n$, there exists a constant $C_{m,k} > 0$ such that

$$|t|^m |D^k \phi(t)| \leq C_{m,k}.$$

In particular, for $m = 0, k = (0, \cdots, 0)$, $|\phi(t)| \leq C_{0,0}$, so $\phi \in L^\infty(\mathbb{R}^n)$.

Now let $1 \leq p < \infty$. Write

$$\int_{\mathbb{R}^n} |\phi(t)|^p dt = \int_{\mathbb{R}^n} \frac{|\phi(t)|^p (1 + |t|^N)^p}{(1 + |t|^N)^p} dt,$$

for some $N \in \mathbb{N}$. Notice that the numerator can be bounded as follows:

$$|\phi(t)|^p (1 + |t|^N)^p = [|\phi(t)| + |t|^N |\phi(t)|]^p \leq (C_{0,0} + C_{N,0})^p < \infty.$$

Call $C = (C_{0,0} + C_{N,0})^p$. Then,

$$\int_{\mathbb{R}^n} |\phi(t)|^p dt \leq C \int_{\mathbb{R}^n} \frac{1}{(1 + |t|^N)^p} dt. \tag{2.4}$$

So we need to check that the integral in (2.4) is finite. By a change to polar coordinates, we can write

$$\begin{aligned}\int_{\mathbb{R}^n} \frac{1}{(1+|t|^N)^p} dt &= \sigma(\mathbb{S}^{n-1}) \int_0^\infty \frac{r^{n-1}}{(1+r^N)^p} dr \\ &\leq \sigma(\mathbb{S}^{n-1}) \left(\int_0^1 r^{n-1} dr + \int_1^\infty \frac{dr}{r^{Np-n+1}} \right),\end{aligned} \tag{2.5}$$

where σ is a measure over the sphere $\mathbb{S}^{n-1}$. The first integral in (2.5) is finite, and the second will be so if $Np - n + 1 > 1$, or equivalently, $N > n/p$. Thus, choosing N as asserted, $\phi \in L^p(\mathbb{R}^n)$. □

We have already seen enough about Schwartz functions to be able to introduce the space of tempered distributions.

2.2 The space $\mathcal{S}'$ of tempered distributions

Remember that we defined the space $\mathcal{D}'$ of distributions as a dual space of the space of testing functions $\mathcal{D}$. In Section 2.1, we have extended the latter into a larger one, $\mathcal{S}$, in which we encountered many more testing functions. So it is natural to try to define distributions in this new space.

2.2.1 Definition and some basic properties

Definition 2.11. A mapping $T : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ is called a **tempered distribution** if

1. it is linear, in the sense that if $\phi, \varphi \in \mathcal{S}(\mathbb{R}^n)$ and $a, b \in \mathbb{C}$, then

$$T(a\phi + b\varphi) = aT(\phi) + bT(\varphi).$$

2. it is continuous, in the sense that if $\{\phi_n\} \to \phi$ in $\mathcal{S}(\mathbb{R}^n)$, then

$$\{T(\phi_n)\} \to T(\phi)$$

in $\mathbb{C}$.

The space of tempered distributions is denoted by $\mathcal{S}'(\mathbb{R}^n)$. Some authors also call tempered distributions **distributions of slow growth**. In general, tempered distributions will be represented using letters such as f and g as usual functions, and the image $f(\phi)$ will be written as $\langle f, \phi \rangle$.

As we see, the concept of tempered distribution is completely analogous to that of usual distribution. In fact, corresponding to Proposition 1.11, we also have a similar result.

Proposition 2.12. *If a functional $f : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ is known to be linear, it is enough to see that, if $\{\phi_m\} \to 0$ in $\mathcal{S}$, then $\{\langle f, \phi_m\rangle\} \to 0$ in $\mathbb{C}$ in order for f to be continuous, and thus a tempered distribution.*

Proof. It is completely analogous to the one corresponding to Proposition 1.11. □

In Theorem 2.4 we asserted that every testing function is a Schwartz function. It is natural to wonder whether a similar relation holds for spaces $\mathcal{D}'$ and $\mathcal{S}'$. Observe that if we consider a tempered distribution $f \in \mathcal{S}'$, it assigns a complex number to every Schwartz function, and thus to every testing function. So we can expect that $f \in \mathcal{D}'$. Indeed, the following proposition shows so.

Theorem 2.13. *Every tempered distribution is a distribution. In short,*

$$\mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n),$$

where the inclusion is proper.

Proof. Let $f \in \mathcal{S}'$ be a tempered distribution. We know by Theorem 2.4 that for every $\varphi \in \mathcal{D}$, the image $\langle f, \varphi\rangle$ is well-defined, and because of the definition of being a tempered distribution, f is linear. So we only have to check whether it is continuous. Consider φ_n to be a sequence which converges to zero in $\mathcal{D}$. By Proposition 2.7, φ_n also converges in $\mathcal{S}$, so for the continuity of f in $\mathcal{S}'$, the sequence $\langle f, \varphi_n\rangle$ converges to zero in $\mathbb{C}$. This shows that $\mathcal{S}' \subseteq \mathcal{D}'$.

Now consider the function $f(t) = e^{|t|^2}$ in $\mathbb{R}^n$. This function is locally integrable for being C^∞, so according to (1.2), it defines a distribution in $\mathcal{D}'(\mathbb{R}^n)$, given by

$$\langle f, \varphi\rangle = \int_{\mathbb{R}^n} f(t)\varphi(t)dt.$$

But the function ϕ defined by (2.1) in Theorem 2.4 is a Schwartz function, for which

$$\langle f, \phi\rangle = \int_{\mathbb{R}^n} f(t)\phi(t)dt = \int_{\mathbb{R}^n} e^{|t|^2}e^{-|t|^2}dt = \int_{\mathbb{R}^n} dt = \infty.$$

This means that f cannot be a tempered distribution, thus $\mathcal{S}' \subsetneq \mathcal{D}'$. This completes the proof. □

Very briefly, we will define **convergence in the space of tempered distributions**, very similarly as we did in section 1.2.3.

Definition 2.14. Let f_n be a sequence of tempered distributions. We say that the sequence converges to the tempered distribution f if for every Schwartz function ϕ, the sequence $\langle f_n, \phi\rangle$ converges to $\langle f, \phi\rangle$.

2.2.2 Examples

In the same way we did in Section 1.2.2, it is interesting to analyse some examples of tempered distributions, for they will be important in the upcoming sections.

Example 2.15 (Delta function). Let us start with the most representative distribution, the delta function. We saw in Section 1.2.2 that it defines a distribution. We want to check it also defines a tempered distribution. We know that for every $\phi \in \mathcal{S}$, $\langle \delta, \phi \rangle = \phi(0)$, so clearly, it is well-defined. We also checked that it is linear. Now, consider a convergent sequence $\phi_n \to 0$ in $\mathcal{S}$. Then, we know that choosing constants $m = 0$ and $k = (0, \cdots, 0)$, the sequence $\{\phi_n(t)\}$ converges uniformly to zero. Because of this,

$$|\langle \delta, \varphi_n \rangle| = |\phi_n(0)| \leq \sup_{\mathbb{R}^n} \phi_n \to 0,$$

while $n \to \infty$. As a conclusion, **the delta function defines a tempered distribution**.

Example 2.16. The proof of Theorem 2.13 clearly shows that, in contrast to usual distributions, not every locally integrable function defines a tempered distribution. Anyway, we are interested in knowing for which kind of functions f an expression as

$$\int_{\mathbb{R}^n} f(t)\phi(t)dt \tag{2.6}$$

defines a tempered distribution. It seems clear that the problem resides in the behaviour of the function $f(t)\phi(t)$ when $|t|$ tends to infinity. We also know that $\phi \in \mathcal{S}$ represent a Schwartz function, or as we called it alternatively, a function of rapid descent. This means that the values $\phi(t)$ tend to zero as $|t|$ increases. So it seems that the only property we need to ask f is that it does not grow too much; it should grow not faster than the pace in which ϕ decreases.

With this idea in mind, we call f a **function of slow growth** if there exists a natural number N such that

$$\lim_{|t| \to \infty} |t|^{-N} |f(t)| = \lim_{|t| \to \infty} \frac{|f(t)|}{|t|^N} = 0. \tag{2.7}$$

We are interested in proving that **every locally integrable function of slow growth defines a tempered distribution** through the expression (2.6). So considering f to be so, the first we have to check is whether the functional is well-defined. In other words, we need to prove that the integral exists for every $\phi \in \mathcal{S}(\mathbb{R}^n)$. From (2.7) we can assert that for any $\epsilon > 0$ we can find $M > 0$ such that

$$\forall |t| > M, \quad |f(t)| < \epsilon |t|^N. \tag{2.8}$$

This property can be exploited if we split the integral (2.6) into two terms:

$$\int_{\mathbb{R}^n} f(t)\phi(t)dt = \int_{|t| \leq M} f(t)\phi(t)dt + \int_{|t| > M} f(t)\phi(t)dt. \tag{2.9}$$

Observe that if $|t| \le M$, then as $\phi \in \mathbb{C}^\infty(\overline{B}(M))$, it is bounded, and thus, the local integrability of f implies that the first term of the right-hand side of (2.9) is finite. It remains to deal with the second term. We can work with $1/|t|$, because we are far away from the origin. Thus, using (2.8),

$$\begin{aligned} \left| \int_{|t|>M} f(t)\phi(t)dt \right| &\le \int_{|t|>M} |f(t)||\phi(t)|dt \le \epsilon \int_{|t|>M} |t|^N |\phi(t)|dt \\ &= \epsilon \int_{|t|>M} \frac{|t|^{N+n+1}|\phi(t)|}{|t|^{n+1}} dt. \end{aligned} \tag{2.10}$$

Now, as $\phi \in \mathcal{S}(\mathbb{R}^n)$, there exists a constant $C_{N+n+1,0} > 0$ such that $|t|^{N+n+1}|\phi(t)| < C_{N+n+1,0}$, and thus, it follows from (2.10) that

$$\left| \int_{|t|>M} f(t)\phi(t)dt \right| \le \epsilon C_{N+n+1,0} \int_{|t|>M} \frac{dt}{|t|^{n+1}}.$$

Now, integrating in polar coordinates, we get that

$$\int_{|t|>M} \frac{dt}{|t|^{n+1}} = \sigma(\mathbb{S}^{n-1}) \int_M^\infty \frac{dr}{r^2} < \infty, \tag{2.11}$$

where σ represents a measure over the $(n-1)$-dimensional sphere $\mathbb{S}^{n-1}$. This shows that the operator is well defined.

Apart from that, it is clear that the functional (2.6) is linear. To see that it is continuous, we must prove that if ϕ_m converges to zero in $\mathcal{S}$, then $\langle f, \phi_m \rangle$ does so in $\mathbb{C}$. The idea is the same as in proving that the functional is well defined. Indeed, we split the integral into two parts as is (2.9) (considering ϕ_m instead of ϕ). The first part is bounded by

$$\sup_{|t|<M} |\phi_m(t)| \int_{|t|<M} |f(t)|dt,$$

in which the integral is finite and the supremum tends to zero, because the convergence of ϕ_m is uniform. On the other hand, for the second part we need (2.8) and a procedure similar to (2.10) to get a bound given by

$$\epsilon \sup_{\mathbb{R}^n} \{|t|^{N+n+1}|\phi_m(t)|\} \int_{|t|>M} \frac{dt}{|t|^{n+1}},$$

where the integral, the same as (2.11), is also finite and the supremum tends to zero by the convergence of ϕ_m in $\mathcal{S}$. So eventually we get the result and f defines a tempered distribution.

Example 2.17. As a remark, it turns out that every bounded function is of slow growth, as it can be easily checked from the definition given in (2.7). In particular, **every Schwartz function is of slow growth**, as they are bounded (Theorem

2.10). Moreover, as Schwartz functions are infinitely smooth, they are locally integrable, so each of them defines a tempered distribution. As a consequence, and considering Schwartz functions and their corresponding tempered distributions to be the same, this example and Theorems 2.4 and 2.13 show that we can display a chain of relationships among every space discussed so far,

$$\mathcal{D} \subset \mathcal{S} \subset \mathcal{S}' \subset \mathcal{D}'.$$

Example 2.18. Let $f \in L^p(\mathbb{R}^n)$ for some $p \geq 1$. Then, f defines a tempered distribution. Indeed, for every $\phi \in \mathcal{S}(\mathbb{R}^n)$, from the definition of being Schwartz we can find a constant K such that

$$\left| \int_{\mathbb{R}^n} f(t)\phi(t)dt \right| \leq \int_{R^n} |f(t)||\phi(t)| \frac{1+|t|^N}{1+|t|^N} dt \leq K \int_{R^n} \frac{|f(t)|}{1+|t|^N} dt. \tag{2.12}$$

Now, recall the computations in Theorem 2.10. We saw in the proof, in (2.4) and in (2.5) that the function $1/(1+|t|^N)$ is in L^p whenever $N > n/p$. Therefore, taking any $N > n$, it is in every L^p space, for $p \geq 1$. Therefore, considering q to be the Hölder conjugate of p, by Hölder's inequality we can write

$$\int_{R^n} \frac{|f(t)|}{1+|t|^N} dt \leq ||f||_p \left|\left| \frac{1}{1+|t|^N} \right|\right|_q < \infty.$$

Therefore, the functional is well-defined. Linearity is clear, and continuity is given by the same procedure as in (2.12) and the dominated convergence theorem. Thus, we can write

$$L^p(\mathbb{R}^n) \subseteq \mathcal{S}'(\mathbb{R}^n).$$

2.2.3 Basic operations on tempered distributions

We know that we are allowed to make certain operations with distributions in order to get new distributions. As $\mathcal{S}' \subset \mathcal{D}'$, we can, for instance, sum $f, g \in \mathcal{S}'$ and obtain $f + g \in \mathcal{D}'$. A well developed and consistent theory would make $f + g \in \mathcal{S}'$. Indeed, ours does so, and other operations defined in subsection 1.2.4 also perform well in $\mathcal{S}'$.

In this subsection, we will present some of the most important operations on $\mathcal{S}'$, which coincide with those on $\mathcal{D}'$. Notice that we only need to check continuity to ensure they produce tempered distributions, as the definitions are given in Subsection 1.2.4, they are well-defined because of the results presented in Propositions 2.3 and 2.9, and the property of linearity does not depend on the nature of the space of functions. So let $f, g \in \mathcal{S}'(\mathbb{R}^n)$, $\alpha \in \mathbb{C}$, $\tau \in \mathbb{R}^n$ and $a \in \mathbb{R}, a \neq 0$. We analyse the continuity of the sum, the multiplication by a constant, the shifting, the transposition and the dilation of tempered distributions. For that, consider ϕ_n to be a convergent sequence in $\mathcal{S}$ whose limit is zero.

1. The **sum of tempered distributions**, $f + g$.

$$|\langle f+g, \phi_n \rangle| \leq |\langle f, \phi_n \rangle| + |\langle g, \phi_n \rangle| \to 0.$$

2. The **multiplication by a constant**, αf.

$$|\langle \alpha f, \phi_n \rangle| \leq |\alpha||\langle f, \phi_n \rangle| \to 0.$$

3. The **shifting of a tempered distribution**, $f(t-\tau)$. In this case, as $\langle f(t-\tau), \phi \rangle = \langle f, \phi(t+\tau) \rangle$, it is enough to check that the sequence $\{\phi_n(t+\tau)\} \to 0$ in $\mathcal{S}$. So for every $m \in \mathbb{Z}^+$ and $k \in (\mathbb{Z}^+)^n$, we get, using Lemma 2.8,

$$|t|^m |D^k \phi_n(t+\tau)| \leq 2^m (|t+\tau|^m |D^k \phi_n(t+\tau)| + |\tau|^m |D^k \phi_n(t+\tau)|),$$

where both summands tend uniformly to zero because the convergence of ϕ_n in $\mathcal{S}$. This shows that the sequence $\{\phi_n(t+\tau)\}$ converges to zero in $\mathcal{S}$, and thus $|\langle f, \phi_n(t+\tau\rangle| \to 0$ in $\mathbb{C}$.

4. The **transposition of a tempered distribution**, $f(-t)$. As in the previous case, it is enough to see that the sequence $\{\phi_n(-t)\}$ converges to zero in $\mathcal{S}$. But this is trivial, because $|t| = |-t|$.

5. The **dilation of a tempered distribution**, $f(at)$. Again we only have to check that the sequence $\{\phi_n(t/a)\}$ converges to zero in $\mathbb{C}$. It is trivial too, because

$$|t|^m \left| D^k \phi_n \left(\frac{t}{a} \right) \right| = |a|^m \left| \frac{t}{a} \right|^m \left| D^k \phi_n \left(\frac{t}{a} \right) \right|,$$

and the right-hand side term converges uniformly to zero.

Remark 2.19. Notice that in this case, we cannot define the product of a tempered distribution with a smooth function. In fact, not every product of a smooth function and a Schwartz function is a Schwartz function. It is enough to consider $\phi(t) = e^{-|t|^2} \in \mathcal{S}$ as in 2.1 and $g(t) = e^{|t|^2}$.

However, it can be proved that a tempered distribution can be multiplied by any C^∞ function for which itself and all its derivatives are of slow growth. If ψ is of that kind, then $\phi\psi$ is Schwartz for every Schwartz function ϕ, and the sequence $\psi\phi_n$ will converge to zero if ϕ_n does so. Thus, for every $f \in \mathcal{S}'$, $\psi \cdot f$ will define a tempered distribution by the usual definition

$$\langle \psi \cdot f, \phi \rangle = \langle f, \psi\phi \rangle, \quad \forall \phi \in \mathcal{S}'.$$

In particular, as every Schwartz function is of slow growth and as all derivatives are Schwartz functions and thus of slow growth, **multiplication between tempered distributions and Schwartz functions is permitted**.

We cannot forget about one of the main feature of distributions, the derivative.

6. The **derivative of a tempered distribution**, $\frac{\partial f}{\partial t_i}$. It is immediate from the definition of convergence in $\mathcal{S}$ that the sequence $\{\frac{\partial \phi_n}{\partial t_i}\}$ converges to zero in $\mathcal{S}$, so the derivative of a tempered distribution is continuous, and thus a tempered distribution.

So as we have proved that derivatives remain in the space of tempered distributions, they will hold every property we analysed in section 1.2.5. It is also interesting to remark that differentiation is, in the same way as seen in Proposition 1.20, a linear and continuous operator in $\mathcal{S}$.

Proposition 2.20. *Differentiation is a linear operation in the space of tempered distributions, and it is continuous in the sense that if a sequence of tempered distributions f_n converges to f in $\mathcal{S}'$, then $D^k f_n$ converges to $D^k f$ in $\mathcal{S}'$.*

2.2.4 Convolution

The ones presented in 2.2.3 are not the only operations we can define in the space of tempered distributions. Indeed, we will need to work with not so basic ones when we deal with fundamental solutions of differential operators in Chapter 4. The first new operation we will work with is convolution. It is possible, and useful, to define the convolution between tempered distributions under certain conditions, and even on the space of usual distributions. However, **we will focus on the convolution between a tempered distribution and a Schwartz function**, as it will be enough to achieve our goal.

Recall the definition of the **convolution between two integrable functions** f and g,

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y)g(x-y)dy.$$

It is straightforward to see that convolution is a commutative operation, so $f * g = g * f$. What we want to do first is to show that the convolution between two Schwartz functions is still a Schwartz function. We know that the definition of a Schwartz function depends in a big way of differentiation. On the other hand, convolution implies dealing with an integral, so we can expect having troubles with the commutability of both. The dominated convergence theorem is of great help in this situation, and it will allow us to present a lemma determining conditions under which commutation of integration and derivation will be permitted.

Lemma 2.21. *Let g be a function in $\mathbb{R}^n \times \mathbb{R}^n$, and consider, being $x, t \in \mathbb{R}^n$,*

$$f(t) = \int_{\mathbb{R}^n} g(t,x)dx.$$

Let $k \in \{1, \cdots, n\}$, and suppose both g and $\frac{\partial g}{\partial t_k}$ are continuous. Moreover, suppose $\frac{\partial g}{\partial t_k}$ is bounded by an integrable function in the variable x. In this situation,

$$\frac{\partial f}{\partial t_k}(t) = \int_{\mathbb{R}^n} \frac{\partial g}{\partial t_k}(t,x)dx.$$

Proof. It is consequence of using the definition of the partial derivative as a limit, the mean value theorem and the dominated convergence theorem. □

Once we know how to manage this situation, we present the following auxiliar proposition.

Proposition 2.22. *Let $f, g \in \mathcal{S}(\mathbb{R}^n)$. Then, $f * g \in \mathbb{C}^\infty(\mathbb{R}^n)$.*

Proof. We know that both f and g are smooth. As we want to compute, for every $k \in \{1, \cdots, n\}$, the derivative

$$\frac{\partial}{\partial x_k} \int_{\mathbb{R}^n} f(x-y)g(y)dy,$$

we can use Lemma 2.21 to take the derivative inside, and thus

$$\begin{aligned} \frac{\partial}{\partial x_k} \int_{\mathbb{R}^n} f(x-y)g(y)dy &= \int_{\mathbb{R}^n} \frac{\partial}{\partial x_k}(f(x-y)g(y))dy \\ &= \int_{\mathbb{R}^n} \frac{\partial f}{\partial x_k}(x-y)g(y)dy = \left(\frac{\partial f}{\partial x_k} * g\right)(x), \end{aligned}$$

from where the result follows immediately. $\square$

Remark 2.23. If we read the previous proof carefully, we will realize that g need not be so regular. Indeed, it would be enough for it to be continuous and integrable. In any case, if g is also Schwartz, by commuting the definition of convolution, the same proof shows that

$$\left(\frac{\partial f}{\partial x_k} * g\right)(x) = \frac{\partial}{\partial x_k}(f * g)(x) = \left(f * \frac{\partial g}{\partial x_k}\right)(x).$$

Let us prove that convolution is an internal operation in $\mathcal{S}$.

Theorem 2.24. *Let $f, g \in \mathcal{S}(\mathbb{R}^n)$. Then $f * g \in \mathcal{S}(\mathbb{R}^n)$.*

Proof. Proposition 2.22 shows the convolution is smooth. Thus, take $m \in Z^+, k \in (\mathbb{Z}^+)^n$, and observe that

$$|x|^m |D^k(f * g)(x)| \leq \int_{\mathbb{R}^n} |x|^m |g(x-y)||D^k f(y)|dy.$$

Now, writing $|x| = |x - y + y|$, we can use Lemma 2.8 to assert that

$$\begin{aligned} |x|^m |D^k(f * g)(x)| \leq 2^m \Big(&\int_{\mathbb{R}^n} |x-y|^m |g(x-y)||D^k f(y)|dy \\ &+ \int_{\mathbb{R}^n} |y|^m |g(x-y)||D^k f(y)|dy \Big). \end{aligned} \tag{2.13}$$

Now, $|x-y|^m|g(x-y)|$ and $|y|^m|D^k f(y)|$ can be bounded for the definition of f and g being Schwartz, and as every Schwartz function is integrable (Theorem 2.10), both $D^k f(y)$ and $g(x-y)$ are so, thus the sum of integrals (2.13) is bounded by a constant depending on m and k, but not on x. $\square$

After seeing some properties of convolution in $\mathcal{S}$, we need to define the convolution between a tempered distribution and a Schwartz function. Again, it is a good way to check the situation with regular distributions, so let $\phi, \varphi \in \mathcal{S}(\mathbb{R}^n)$. By Theorem 2.24, $f * g \in \mathcal{S}(\mathbb{R}^n)$, so it can be seen as a regular distribution, and thus for all $\psi \in \mathcal{S}(\mathbb{R}^n)$,

$$\langle \phi * \varphi, \psi \rangle = \int_{\mathbb{R}^n} (\phi * \varphi)(x)\psi(x)dx = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \phi(y)\varphi(x-y)dy \; \psi(x)dx.$$

Observe that ϕ, φ, ψ are integrable and bounded, so Fubini's theorem can be applied to revert the order inside the integral, and

$$\langle \phi * \varphi, \psi \rangle = \int_{\mathbb{R}^n} \phi(y) \left(\int_{\mathbb{R}^n} \varphi(x-y)\psi(x)dx \right) dy.$$

Denote as $\tilde{\varphi}$ the transposition of the function φ. In other words, $\tilde{\varphi}(t) = \varphi(-t)$. Then, we can write

$$\int_{\mathbb{R}^n} \phi(y) \left(\int_{\mathbb{R}^n} \varphi(x-y)\psi(x)dx \right) dy = \int_{\mathbb{R}^n} \phi(y) \left(\int_{\mathbb{R}^n} \tilde{\varphi}(y-x)\psi(x)dx \right) dy,$$

which clearly shows a convolution between $\tilde{\varphi}$ and ψ, to finally get

$$\langle \phi * \varphi, \psi \rangle = \langle \phi, \tilde{\varphi} * \psi \rangle. \tag{2.14}$$

Notice that $\tilde{\varphi}$ is a Schwartz function, and so, we will be able to extend (2.14) to any kind of distributions.

Definition 2.25. Let $f \in \mathcal{S}'(\mathbb{R}^n)$ and $\psi \in \mathcal{S}(\mathbb{R}^n)$. We define the **convolution between a tempered distribution and a Schwartz function**, f and ψ, as follows:

$$\langle f * \psi, \phi \rangle = \langle f, \tilde{\psi} * \phi \rangle, \quad \forall \phi \in \mathcal{S}(\mathbb{R}^n),$$

where $\tilde{\psi}$ represents the transposition of the function ψ.

We can expect, by the reasoning done for Schwartz functions, this definition to lead to a tempered distribution. Fortunately, that property holds, as we can see in the following theorem.

Theorem 2.26. *For every $f \in \mathcal{S}'(\mathbb{R}^n)$ and $\psi \in \mathcal{S}(\mathbb{R}^n)$, the convolution $f * \psi$ defines a tempered distribution.*

Proof. The linearity of the functional defined in Definition 2.25 is a result of the linearity of both the convolution of functions and the distribution f. So let ϕ_l be a sequence convergent to zero in $\mathcal{S}(\mathbb{R}^n)$. We know that every $\tilde{\psi} * \phi_l$ is a Schwartz function, so we only have to check that every sequence $|x|^m D^k(\tilde{\psi} * \phi_l)(x)$ converges uniformly to zero, for usual indices m, k. We will reason as in equation (2.13). Indeed, substituting f for $\tilde{\psi}$ and g for ϕ_l, and taking suprema,

$$\begin{aligned} |x|^m |D^k(\tilde{\psi} * \phi_l)(x)| \leq 2^m \Big(&\sup_{x\in\mathbb{R}^n} |x|^m |\phi_l(x)| \int_{\mathbb{R}^n} |D^k \tilde{\psi}(y)| dy \\ &+ \sup_{x\in\mathbb{R}^n} |\phi_l(x)| \int_{\mathbb{R}^n} |y|^m |D^k \tilde{\psi}(y)| dy \Big). \end{aligned} \tag{2.15}$$

The two integrals on the right-hand side of (2.15) are finite. Indeed, the first one is concerning a Schwartz, thus integrable function. The second integrand can be bounded by a factor $C/|y|^{n+1}$, where C is a constant, for $\tilde{\psi}$ being Schwartz. Therefore, the convergence of ϕ_l in $\mathcal{S}$ makes both suprema tend to zero, what implies the convergence of $\tilde{\psi} * \phi_n$. Finally, the continuity of f implies that of $f * \psi$. □

We have seen that the derivatives are fairly easy to handle when working with Schwartz functions. The next result shows that the situation does not change when considering tempered distributions.

Proposition 2.27. *Let $f \in \mathcal{S}'(\mathbb{R}^n)$ and $\psi \in \mathcal{S}(\mathbb{R}^n)$. Then,*

$$\frac{\partial}{\partial x_k}(f * \psi) = \frac{\partial f}{\partial x_k} * \psi = f * \frac{\partial \psi}{\partial x_k}, \quad k = 1, \cdots, n.$$

Proof. It is an immediate consequence of the definition of the convolution and remark 2.23. □

The most interesting example of convoluting a distribution is that concerning the delta function. Indeed, let δ be the delta function, and consider any Schwartz function ψ. We want to work out the value of $\delta * \psi$. For that, if ϕ represents a Schwartz function,

$$\begin{aligned} \langle \delta * \psi, \phi \rangle = \langle \delta, \tilde{\psi} * \phi \rangle = \tilde{\psi} * \phi(0) &= \int_{\mathbb{R}^n} \tilde{\psi}(-y)\phi(y) dy \\ &= \int_{\mathbb{R}^n} \psi(y)\phi(y) dy = \langle \psi, \phi \rangle. \end{aligned}$$

This means that

$$\delta * \psi = \psi, \quad \forall \psi \in \mathcal{S}(\mathbb{R}^n). \tag{2.16}$$

In other words, **the delta function works as the identity element with respect to convolution**.

Remark 2.28. Even if we will not be able to define the convolution between two arbitrary distributions, it is possible to do so under certain restrictions, and in a more general way than we did in this section. We refer the reader to [10, p. 122-137] or to [8, p. 102-109].

2.2.5 Direct product

Another important operation on the space of tempered distributions is the **direct product**. Indeed, the direct product can be defined on the space $\mathcal{D}'$ of usual distributions. In this subsection, we will briefly discuss the direct product on $\mathcal{D}'$. The reason for this is that this new operation will be required for some minor, though necessary computations in Chapter 4. We will also mention the analogous results for $\mathcal{S}'$, and references for more details will be given.

First of all, let us present what intuitively the direct product could be. Let f and g be locally integrable functions on $\mathbb{R}^n$. Then, the function given by $f(x)g(y)$ on $\mathbb{R}^{m+n}$ is also locally integrable, thus defines a distribution,

$$\langle f(x)g(y), \varphi(x,y)\rangle = \int_{\mathbb{R}^{n+m}} f(x)g(y)\varphi(x,y)dxdy, \quad \forall \varphi \in \mathcal{D}(\mathbb{R}^{n+m}).$$

For φ being bounded, Fubini's theorem can be applied to consider the most convenient order. In this case, let us write

$$\int_{\mathbb{R}^{n+m}} f(x)g(y)\varphi(x,y)dxdy = \int_{\mathbb{R}^n} f(x)\left(\int_{\mathbb{R}^m} g(y)\varphi(x,y)\right)dx,$$

so considering g as a distribution, we could formally write

$$\langle f(x)g(y), \varphi(x,y)\rangle = \langle f(x), \langle g(y), \varphi(x,y)\rangle\rangle,$$

what suggests a definition for the direct product of distributions. In fact, we can assert that the definition is correct.

Definition 2.29. Let $f \in \mathcal{D}'(\mathbb{R}^n)$ and $g \in \mathcal{D}'(\mathbb{R}^m)$ be any two distributions. Their direct product is defined as

$$\langle f(x) \cdot g(y), \varphi(x,y)\rangle = \langle f(x), \langle g(y), \varphi(x,y)\rangle\rangle, \quad \forall \varphi \in \mathcal{D}(\mathbb{R}^{n+m}).$$

It is also usual to denote the direct product of f and g as $f(x) \times g(y)$.

We have seen that the definition works fine for locally integrable functions, but when it comes to distributions, we have to ensure that this definition makes sense. For instance, we should check that the expression $\langle g(y), \varphi(x,y)\rangle$ defines a testing function on $\mathbb{R}^n$ whenever g is a distribution on $\mathbb{R}^m$ and φ is a testing function on $\mathbb{R}^{n+m}$. The following lemma asserts so, although we will not prove it. A proof can be found in [8, p.96-98].

Lemma 2.30. *For every distribution $g \in \mathcal{D}'(\mathbb{R}^m)$ and every testing function $\varphi \in \mathcal{D}(\mathbb{R}^{n+m})$, the function*

$$\psi(x) = \langle g(y), \varphi(x,y)\rangle$$

is a testing function on $\mathbb{R}^n$. Moreover, its derivatives are given by

$$D^\alpha \psi(x) = \langle g(y), D_x^\alpha \psi(x,y)\rangle,$$

for every multi-index α, where the notation D_x^α stands for the partial derivatives with respect to x and with indices α. Also, if $\varphi_n \to \varphi$ in $\mathcal{D}(\mathbb{R}^{n+m})$, then

$$\psi_n = \langle g(y), \varphi_n(x, y)\rangle \to \psi$$

in $\mathcal{D}(\mathbb{R}^n)$.

This lemma allows us to prove that the direct product is indeed a distribution.

Proposition 2.31. *The direct product of two distributions is a distribution.*

Proof. Let f and g be distributions on $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. Lemma 2.30 ensures that the functional $f(x) \cdot g(y)$ is well-defined, and it is linear because both f and g are so. Now, according again to Lemma 2.30, if $\varphi_n \to \varphi$ in $\mathcal{D}(\mathbb{R}^{n+m})$, we know that

$$\langle g(y), \varphi_n(x, y)\rangle \to \langle g(y), \varphi(x, y)\rangle$$

in $\mathcal{D}(\mathbb{R}^n)$, and as f is continuous,

$$\langle f(x), \langle g(y), \varphi_n(x, y)\rangle\rangle \to \langle f(x), \langle g(y), \varphi(x, y)\rangle\rangle,$$

showing that $f(x) \cdot g(y)$ is continuous. □

It can also be proved that the direct product is a commutative operation. This means that for every distributions f and g, it holds that

$$f(x) \cdot g(y) = g(y) \cdot f(x). \tag{2.17}$$

A proof can be found in [8, p.99-100].

Results for the space $\mathcal{S}'$ of tempered distributions completely analogous to Lemma 2.30 and Proposition 2.31 can be proved. Proofs can be found in [8, 119-121].

To end with this section, we present a very simple example concerning the Delta function. Let $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$. Then,

$$\langle \delta(x) \cdot \delta(y), \varphi(x, y)\rangle = \langle \delta(x), \varphi(x, 0)\rangle = \varphi(0, 0) = \langle \delta(x, y), \varphi(x, y)\rangle,$$

for every $\varphi \in \mathcal{D}(\mathbb{R}^{n+m})$. Thus, it is clear that

$$\delta(x) \cdot \delta(y) = \delta(x, y). \tag{2.18}$$

Chapter 3

The Fourier Transform

In this chapter, we will study what the effect of applying the Fourier transform to functions in $\mathcal{S}$ is, in order to define the Fourier transform of distributions. Several problems appear when the Fourier transform is defined for the space $\mathcal{D}$, but it works properly when we extend the space to $\mathcal{S}$. We will discuss this along next pages.

3.1 The Fourier transform in L^1

First of all, we will introduce the Fourier transform for integrable functions.

Definition 3.1. Let $f : \mathbb{R}^n \to \mathbb{C}$ be an integrable function. We define the **Fourier transform** of f as the function $\hat{f} : \mathbb{R}^n \to \mathbb{C}$, given by

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-i(\xi \cdot x)} dx.$$

It is also common to denote the Fourier transform of f as $\mathcal{F}(f)$ or simply $\mathcal{F}f$.

We also define the **inverse Fourier transform** of f as the function $\check{f} : \mathbb{R}^n \to \mathbb{C}$, given by

$$\check{f}(\xi) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} f(x) e^{i(\xi \cdot x)} dx,$$

and it is usually denoted by $\mathcal{F}^{-1}(f)$ or simply $\mathcal{F}^{-1}f$.

It is important to notice that the Fourier transform of a function in L^1 is bounded and continuous, and thus well-defined. We present this fact in the following proposition.

Proposition 3.2. *Let $f \in L^1(\mathbb{R}^n)$. Then, $\mathcal{F}f \in L^\infty(\mathbb{R}^n)$, and it is continuous.*

Proof. The boundedness requires a really short proof. Indeed,

$$|\hat{f}(\xi)| \leq \int_{\mathbb{R}^n} |f(x)||e^{-i(\xi \cdot x)}| dx = \int_{\mathbb{R}^n} |f(x)| dx < \infty.$$

Continuity follows from the dominated convergence theorem, as we can write

$$|\hat{f}(\xi+h)-\hat{f}(\xi)| \leq \int_{\mathbb{R}^n} |f(x)||e^{-i(\xi\cdot x)}||e^{-i(h\cdot x)}-1|dx.$$

The integrand is bounded by $2|f(x)|$ and the limit can be taken inside. It is clear that it tends to zero. □

It is interesting to notice that, in general, the Fourier transform of a L^1 function need not be integrable, thus we will not be able to define its inverse transform. This will not be a problem in $\mathcal{S}$ nor in $\mathcal{S}'$. Moreover, based on this definition, the Fourier transform can also be defined for L^2 functions, through a process of density and convergence. It can be proven that $\mathcal{S}$ is dense in L^2, so it will be enough to work in the former. There is even another way to define it in L^2. Recall the result in Example 2.18. Thus, as $L^2 \subseteq \mathcal{S}'$, it will be enough to define it on $\mathcal{S}'$.

3.2 The Fourier transform in $\mathcal{S}$

Defining the Fourier transform in the space $\mathcal{S}$ is key to be able to define it for tempered distributions. We can think of using Definition 3.1 to do so. Remember that we saw in Theorem 2.10 that $\mathcal{S} \subseteq L^1$, so there is no trouble.

Definition 3.3. The Fourier transform on the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ is an operator $\mathcal{F}: \mathcal{S}(\mathbb{R}^n) \to L^\infty(\mathbb{R}^n)$ which asigns to every $\phi \in \mathcal{S}(\mathbb{R}^n)$ the function

$$\mathcal{F}\phi(\xi) = \hat{\phi}(\xi) = \int_{\mathbb{R}^n} \phi(x)e^{-i(\xi\cdot x)}dx.$$

Our next objective is to see that the Fourier transform is a bijective endomorphism (thus an automorphism). In other words, we want to prove:

1. that the Fourier transform of every Schwartz function is also a Schwartz function, and

2. that $\mathcal{F}\mathcal{F}^{-1} = 1_{\mathcal{S}} = \mathcal{F}^{-1}\mathcal{F}$, where $1_{\mathcal{S}}$ represents the identity operator.

To get these two results, it is vital to know how to manage the derivatives of the Fourier transform of a function, as the definition of $\mathcal{S}$ is heavily dependent on differentiation. For that reason, we will fist develop some properties concerning $\mathcal{F}$.

3.2.1 Some properties of the Fourier transform in $\mathcal{S}$

As we have recently advocated, we have to find a way to manage the derivatives of the Fourier transform of Schwartz functions. Its definition involves an integral, so we can expect an environment in which derivatives and integrals will appear at the same time. The very famous dominated convergence theorem will be extremely useful in this situation, by means of Lemma 2.21. This lemma will allow us to prove the properties which are probably the major advantage of working in the Fourier space. These properties are the ones which make a direct relation between differentiation and multiplication. Let us analyse them in detail.

Proposition 3.4. *Let $\phi \in \mathcal{S}(\mathbb{R}^n)$ be a Schwartz function. Then, for every $k \in (\mathbb{Z}^+)^n$, the following properties hold:*

1. $D^k \mathcal{F}(\phi)(\xi) = (-i)^{|k|} \mathcal{F}(x^k \phi)(\xi)$.
2. $\xi^k \mathcal{F}(\phi)(\xi) = (-i)^{|k|} \mathcal{F}(D^k \phi)(\xi)$.

where if $k = (k_1, \cdots, k_n)$, we denote $x^k = x_1^{k_1} \cdots x_n^{k_n}$.

Proof. 1. We want to make the derivative of the Fourier transform $\hat{\phi}(\xi)$. Let $i \in \{1, \cdots, n\}$, and

$$\frac{\partial \hat{\phi}}{\partial \xi_i}(\xi) = \frac{\partial}{\partial \xi_i} \int_{\mathbb{R}^n} \phi(x) e^{-i(\xi \cdot x)} dx.$$

Observe that

$$\left| \frac{\partial}{\partial \xi_i} \left(\phi(x) e^{-i(\xi \cdot x)} \right) \right| = \left| \phi(x)(-ix_i) e^{-i(\xi \cdot x)} \right| = |x_i \phi(x)|.$$

By Proposition 2.9, we know that $x_i \phi(x) \in \mathcal{S}(\mathbb{R}^n)$, and by Theorem 2.10, it is integrable. Therefore, Lemma 2.21 asserts that

$$\begin{aligned} \frac{\partial \hat{\phi}}{\partial \xi_i}(\xi) &= \int_{\mathbb{R}^n} \frac{\partial}{\partial \xi_i} \left(\phi(x) e^{-i(\xi \cdot x)} \right) dx \\ &= -i \int_{\mathbb{R}^n} x_i \phi(x) e^{-i(\xi \cdot x)} dx = -i \mathcal{F}(x_i \phi(x))(\xi). \end{aligned} \tag{3.1}$$

So clearly, applying (3.1) $|k|$ times, we get

$$D^k \hat{\phi}(\xi) = (-i)^{|k|} \mathcal{F}(x^k \phi(x))(\xi).$$

2. Choose again $i \in \{1, \cdots, n\}$. We will compute the Fourier transform of the derivative. By definition,

$$\mathcal{F}(\frac{\partial \phi}{\partial x_i})(\xi) = \int_{\mathbb{R}^n} \frac{\partial \phi}{\partial x_i}(x) e^{-i(\xi \cdot x)} dx.$$

Observe that the function we are integrating is in L^1, because $\left| \frac{\partial \phi}{\partial x_i}(x) e^{-i(\xi \cdot x)} \right| = \left| \frac{\partial \phi}{\partial x_i}(x) \right|$, and $\frac{\partial \phi}{\partial x_k} \in \mathcal{S} \subseteq L^1$, as we saw in Proposition 2.3 and in Theorem 2.10. So the order of the integral can be changed by Fubini's theorem, and

$$\int_{\mathbb{R}^n} \frac{\partial \phi}{\partial x_i}(x) e^{-i(\xi \cdot x)} dx = \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \frac{\partial \phi}{\partial x_i}(x) e^{-i(\xi \cdot x)} dx. \tag{3.2}$$

Now, integrating by parts, the boundary term vanishes for ϕ being Schwartz, and

$$\int_{\mathbb{R}} \frac{\partial \phi}{\partial x_i}(x) e^{-i(\xi \cdot x)} dx_i = i\xi_k \int_{\mathbb{R}} \phi(x) e^{-i(\xi \cdot x)} dx_i.$$

So rewriting (3.2),

$$\int_{\mathbb{R}^n} \frac{\partial \phi}{\partial x_i}(x) e^{-i(\xi \cdot x)} dx = i\xi_k \int_{\mathbb{R}^n} \phi(x) e^{-i(\xi \cdot x)} dx.$$

In other words, $\mathcal{F}(\frac{\partial \phi}{\partial x_i})(\xi) = i\xi_k \mathcal{F}(\phi)$. Applying this last expression $|k|$ times,

$$\mathcal{F}(D^k \phi)(\xi) = i^{|k|} \xi^k \mathcal{F}(\phi)(\xi). \qquad \square$$

Proposition 3.4 conveys how derivatives work in interaction with the Fourier transform. This is key to prove our first objective.

Theorem 3.5. *Let $\phi \in \mathcal{S}(\mathbb{R}^n)$. Then, $\mathcal{F}(\phi) \in \mathcal{S}(\mathbb{R}^n)$.*

Proof. We must check two facts:

1. $\hat{\phi} \in C^\infty$, and
2. for every $m \in \mathbb{Z}^+, k \in (\mathbb{Z}^+)^n, \quad \exists C_{m,k} > 0 \quad | \quad |\xi|^m |D^k \hat{\phi}(\xi)| < C_{m,k}$.

So let us prove each fact separately.

1. By Proposition 3.4,

$$D^k \hat{\phi}(\xi) = (-i)^{|k|} \mathcal{F}(x^k \phi(x))(\xi),$$

so any derivative can be computed in terms of Fourier transforms of $x^k \phi(x)$, which we know it is a Schwartz function by 2.9. The Fourier transform of any Schwartz function is continuous, so $\hat{\phi}$ is infinitely smooth.

2. We need to use both properties of Proposition 3.4. Consider $i \in \{1, \cdots, n\}$. Then,

$$\xi_i^m D^k \hat{\phi}(\xi) = (-i)^{|k|+m} \mathcal{F}(D_i^m (x^k \phi(x))).$$

Then, by the definition of the Fourier transform,

$$|\xi_i^m D^k \hat{\phi}(\xi)| \leq \int_{\mathbb{R}^n} |D_i^m (x^k \phi(x))| dx. \tag{3.3}$$

Observe now that $D_i^m (x^k \phi)$ is a sum of several products between polynomials and derivatives of ϕ. For that, $D_i^m (x^k \phi) \in \mathcal{S}$ holds. This means that it is integrable, and thus, $M_i = \int_{\mathbb{R}^n} |D_i^m (x^k \phi(x))| dx$ is finite.

Also observe that $|\xi|^m \leq \sqrt{n}^m \max_{i=1,\cdots,n} |\xi_i|^m$, so we can write

$$|\xi|^m |D^k \hat{\phi}(\xi)| \leq \sqrt{n}^m \max_{i=1,\cdots,n} |\xi_i^m D^k \hat{\phi}(\xi)| \leq \sqrt{n}^m \max_{i=1,\cdots,n} M_i < \infty. \tag{3.4}$$

$\square$

Some ideas of this proof yield to the fact that the Fourier transform $\mathcal{F}$ is a linear continuous operator from $\mathcal{S}$ to $\mathcal{S}$.

Theorem 3.6. *The operator $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ is a linear operator, and it is continuous in the sense that if $\phi_n \to \phi$ in $\mathcal{S}$, then $\mathcal{F}(\phi_n) \to \mathcal{F}(\phi)$ in $\mathcal{S}$.*

Proof. It is trivial to see that the operator $\mathcal{F}$ is linear. To see that it is continuous, for the same reason as in Proposition 1.11, it is enough to check the case $\{\phi_l\}_{l=1}^{\infty} \to 0$. Supposing so, we have to see that every sequence $|\xi|^m D^k \hat{\phi}_l(\xi) \to 0$ uniformly, being $m \in \mathbb{Z}^+, k \in (\mathbb{Z}^+)^n$. But if we consider (3.4) from the proof of Theorem 3.5, we know that

$$|\xi|^m |D^k \hat{\phi}_l(\xi)| \le \sqrt{n}^m \max_{i=1,\cdots,n} M_{i,l},$$

where

$$M_{i,l} = \int_{\mathbb{R}^n} |D_i^m (x^k \phi_l(x))| dx. \tag{3.5}$$

Therefore, it is enough to check that $\max_{i=1,\cdots,n} M_{i,l}$ tends to zero as $l \to \infty$.

As the result of expanding the derivative of (3.5) will be a finite sum of products of polynomials and derivatives of ϕ_l, the convergence of ϕ_l implies that the derivatives will converge uniformly to zero.

Moreover, if we choose $q \in \mathbb{N}$, the expression $|x|^q |D_i^m (x^k \phi_l)(x)|$ will be bounded by the finite sum mentioned above times $|x|^q$, which, by the convergence of ϕ_l in $\mathcal{S}$, will be bounded by a constant $C_{q,m}$ (it will be the maximum among the constants corresponding to each summand). This way,

$$|x|^q |D_i^m (x^k \phi_l)(x)| \le C_{q,m}.$$

Now, as we can have a bound with a integrable function, we can apply the dominated convergence theorem to assert that

$$\lim_{l \to \infty} M_{i,l} = 0, \quad \forall i = 1, \cdots, n.$$

The proof is now complete. $\square$

3.2.2 The Fourier transform of the Gaussian function

Before going on, it is convenient to present an example of the Fourier transform of a particular function. This example has great importance in several proofs and solutions of differential equations. The function we are talking about is the Gaussian function.

Definition 3.7. Let $x \in \mathbb{R}^n$. We call **Gaussian function** the following function:

$$g(x) = e^{-k|x|^2}, \quad \text{for some } k > 0. \tag{3.6}$$

Instead of using the definition, we will think in an alternative way. First of all, observe that, if we allow a little abuse of notation when considering g in one variable,

$$g(x) = e^{-k|x|^2} = e^{-k(x_1^2+\ldots+x_n^2)} = e^{-kx_1^2} \cdot \ldots \cdot e^{-kx_n^2} = g(x_1) \cdot \ldots \cdot g(x_n). \tag{3.7}$$

Also observe that as $g \in \mathcal{S}$ (recall Theorem 2.4, equation (2.1)), it is integrable, and Fubini's theorem allows us to split the integral using (3.7), and get

$$\hat{g}(\xi) = \hat{g}(\xi_1) \cdot \ldots \cdot \hat{g}(\xi_n). \tag{3.8}$$

Therefore, it is enough to work out the transform in one variable. So consider g in one variable. Then, it is easy to check that the following equation holds:

$$g'(x) + 2kxg(x) = 0. \tag{3.9}$$

Now, let us apply the Fourier transform to (3.9). We get

$$\mathcal{F}(g'(x)) + 2k\mathcal{F}(xg(x)) = 0. \tag{3.10}$$

Proposition 3.4 is offering a way to proceed:

$$\mathcal{F}(g)'(\xi) = -i\mathcal{F}(xg)(\xi) \quad \text{and} \quad \xi\mathcal{F}g(\xi) = -i\mathcal{F}(g')(\xi),$$

so (3.9) is equivalent to

$$i\xi\mathcal{F}g(\xi) + 2ki\mathcal{F}(g)'(\xi) = 0,$$

which in turn is

$$\mathcal{F}(g)'(\xi) + \frac{\xi}{2k}\mathcal{F}g(\xi) = 0.$$

A solution can be easily found; it must be of the form

$$\mathcal{F}g(\xi) = Ce^{-\frac{\xi^2}{4k}}, \quad C \in \mathbb{R}.$$

We need to determine that constant C. Indeed, by the definition,

$$\hat{g}(0) = \int_{\mathbb{R}} g(x)dx = \sqrt{\frac{\pi}{k}}.$$

The last integral value can be found by first squaring it and then using polar coordinates. Thus, considering (3.8),

$$\hat{g}(x) = \left(\sqrt{\frac{\pi}{k}}\right)^n e^{-|x|^2/4k}, \quad \text{in } \mathbb{R}^n, \tag{3.11}$$

the Fourier transform of the Gaussian.

It is also interesting to analyse the case in which the constant k is not real. Consider $k = a + bi = z$ with $a > 0$, and the Gaussian

$$g(x) = e^{-z|x|^2} = e^{-(a+bi)|x|^2}. \tag{3.12}$$

By repeating every calculation for the case where $k > 0$, we can assert that

$$\hat{g}(\xi) = Ke^{-\xi^2/4z},$$

where K is a constant to determine. In fact,

$$K = \hat{g}(0) = \int_{\mathbb{R}} g(x)dx = \int_{\mathbb{R}} e^{-zx^2}dx.$$

This integral cannot be computed as we did in the real case. It needs an special treatment. So, for $a > 0$, consider

$$I_a(t) = \int_{\mathbb{R}} e^{-(a+ti)x^2}dx.$$

Lemma 2.21 and integration by parts allow to write

$$I_a'(t) = \frac{-i}{2(a+ti)}I_a(t),$$

and therefore we get, for a function $k(a)$,

$$I_a(t) = \frac{k(a)}{\sqrt{a+ti}},$$

where we work with the principal branch of the square root in $\mathbb{C}$. Finally, the value $I_a(0)$ gives the value of $k(a)$. In fact, $k(a) = \sqrt{\pi}$, so

$$I_a(t) = \frac{\sqrt{\pi}}{\sqrt{a+ti}}, \quad \forall a > 0.$$

This little calculation conveys that $K = \sqrt{\pi}/\sqrt{z}$,

$$\hat{g}(\xi) = \sqrt{\frac{\pi}{z}}e^{-\xi^2/4z},$$

for every z with $\mathrm{Re}(z) > 0$, and in $\mathbb{R}^n$,

$$\hat{g}(\xi) = \left(\sqrt{\frac{\pi}{z}}\right)^n e^{-|\xi|^2/4z}. \tag{3.13}$$

3.2.3 The inverse Fourier transform in $\mathcal{S}$

Remember that we gave a definition for the inverse Fourier transform in Definition 3.1. In this subsection, our objective will be to prove that it is indeed the inverse operator corresponding to the Fourier transform $\mathcal{F}$. Our first approach could be to try to get the result directly. This way, if $\phi \in \mathcal{S}$,

$$\begin{aligned} \mathcal{F}^{-1}(\mathcal{F}(\phi))(y) &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}\phi(x)e^{-i(\xi\cdot x)}dx\right)e^{-i(z\cdot\xi)}d\xi \\ &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\phi(x)e^{-i\xi\cdot(x-z)}dxd\xi. \end{aligned} \tag{3.14}$$

Anyway, the integral in 3.14 is not absolutely convergent, and thus, Fubini's theorem cannot be applied. To solve this problem, we will need some properties of the Fourier transform, concerning its behaviour with respect to certain operations.

Proposition 3.8. *Let $\phi \in \mathcal{S}(\mathbb{R}^n)$. Then,*

1. *If $\tau \in \mathbb{R}^n$, then*
$$\mathcal{F}(\phi(x-\tau))(\xi) = e^{-i\xi\cdot\tau}\mathcal{F}\phi(\xi). \tag{3.15}$$

2. *If $\lambda \in \mathbb{R}$, then*
$$\mathcal{F}(\phi(\lambda x))(\xi) = \frac{1}{|\lambda|^n}\mathcal{F}\phi\left(\frac{\xi}{\lambda}\right). \tag{3.16}$$

3. *If $\phi, \varphi \in \mathcal{S}(\mathbb{R}^n)$, then*
$$\int_{\mathbb{R}^n} \mathcal{F}\phi(x)\varphi(x)dx = \int_{\mathbb{R}^n} \phi(x)\mathcal{F}\varphi(x)dx. \tag{3.17}$$

Proof. Properties 1 and 2 are a direct result of simply applying a change of variables in the integral of the definition of the Fourier transform. Finally, property 3 is a direct consequence of Fubini's theorem. Observe that Fubini's results can be applied as every Schwartz function is integrable (see Theorem 2.10). □

Having presented these properties, we are ready to face our main objective.

Theorem 3.9. *Let $\phi \in \mathcal{S}(\mathbb{R}^n)$. Then, the inversion formula given by*

$$\mathcal{F}^{-1}(\phi)(\xi) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \phi(x)e^{i(\xi\cdot x)}dx$$

is the inverse operator of $\mathcal{F}$, in the sense that

$$\mathcal{F}^{-1}(\mathcal{F}(\phi)) = \phi = \mathcal{F}(\mathcal{F}^{-1}(\phi)).$$

Proof. We already noticed that a direct proof cannot be achieved. So consider property (3.17) in Proposition 3.8, for $f, g \in \mathcal{S}(\mathbb{R}^n)$:

$$\int_{\mathbb{R}^n} \mathcal{F}f(x)g(x)dx = \int_{\mathbb{R}^n} f(x)\mathcal{F}g(x)dx.$$

Consider $\lambda \neq 0$ and $g(x/\lambda)$ instead of $g(x)$. Then, by property (3.16),

$$\mathcal{F}\left(g\left(\frac{x}{\lambda}\right)\right)(\xi) = |\lambda|^n\mathcal{F}g(\lambda\xi),$$

and coming back to (3.17),

$$\int_{\mathbb{R}^n} \hat{f}(x)g\left(\frac{x}{\lambda}\right)dx = |\lambda|^n\int_{\mathbb{R}^n} f(x)\hat{g}(\lambda x)dx. \tag{3.18}$$

Now, changing variables $y = \lambda x$ on the right hand side of (3.18), we get

$$\int_{\mathbb{R}^n} \hat{f}(x) g\left(\frac{x}{\lambda}\right) dx = \int_{\mathbb{R}^n} f\left(\frac{x}{\lambda}\right) \hat{g}(x) dx.$$

We want to take limits when λ tends to infinity in both sides. Observe that $\lambda \neq 0$ ensures we are working with Schwartz functions all the time, so we are allowed to write

$$|\hat{f}(x) g\left(\frac{x}{\lambda}\right)| \leq |\hat{f}(x)| ||g||_\infty,$$

where of course, by Theorem 2.10, $\hat{f} \in L^1$. With a similar procedure for the right hand side integral, we are allowed to apply the dominated convergence theorem and

$$\int_{\mathbb{R}^n} \hat{f}(x) g(0) dx = \int_{\mathbb{R}^n} f(0) \hat{g}(x) dx. \tag{3.19}$$

This is the time when the Gaussian comes into play. If we consider g to be the Gaussian function (3.6) with $k = 1/2$,

$$g(x) = e^{-|x|^2/2},$$

then (3.11) and (3.19) convey that

$$\int_{\mathbb{R}^n} \hat{f}(x) dx = f(0) (\sqrt{2\pi})^n \int_{\mathbb{R}^n} g(x) dx = (2\pi)^n f(0).$$

Reordering,

$$f(0) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{f}(x) dx, \tag{3.20}$$

which is the result we seek at $\xi = 0$, valid for every Schwartz function. The last step is to use (3.15). Let $\xi \in \mathbb{R}^n$. Considering $f(x + \xi) \in \mathcal{S}$, then by (3.20) and (3.15),

$$\begin{aligned} f(\xi) &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \mathcal{F}(f(x+\xi))(y) dy = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-i(y \cdot (-\xi))} \hat{f}(y) dy \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{f}(x) e^{i(\xi \cdot x)} dx, \end{aligned}$$

which is the definitive result. This shows that $\mathcal{F}^{-1}\mathcal{F}f = f$ for every Schwartz function f. The symmetrical result is achieved in a similar way. □

With this result, we are allowed to say that the inverse Fourier transform of every Schwartz function is again a Schwartz function, and that the operator $\mathcal{F}^{-1}$ is also linear and continuous, by the same arguments used for $\mathcal{F}$.

Theorem 3.10. *The inverse Fourier transform of every Schwartz function is a Schwartz function, and the operator $\mathcal{F}^{-1} : \mathcal{S} \to \mathcal{S}$ is a linear and continuous operator.*

Proof. That $\mathcal{F}^{-1}$ is an endomorphism in $\mathcal{S}$ is a direct consequence of Theorem 3.9. The linearity and continuity of the inverse operator can be obtained by a similar procedure as in Theorem 3.6. □

3.2.4 More properties

Once we have discovered the inverse operator, we can deduce some interesting properties, which may be useful in some calculations.

Proposition 3.11. *Let $\phi, \varphi \in \mathcal{S}(\mathbb{R}^n)$. Then,*

1. $\mathcal{F}^{-1}\phi(\xi) = \frac{1}{(2\pi)^n}\mathcal{F}\tilde{\phi}(\xi)$.
2. $\mathcal{F}\tilde{\phi}(\xi) = \mathcal{F}\phi(-\xi)$.
3. $\mathcal{F}^2\phi(\xi) = (2\pi)^n\tilde{\phi}(\xi)$.
4. $\mathcal{F}^4\phi(\xi) = (4\pi^2)^n\phi(\xi)$.

Proof. To prove properties 1 and 2, it is enough to transpose the variable in the integral. Property 3 uses the same change of variables, Theorem 3.9 and property 2. Using property 3 twice yields to property 4 directly. □

As we said in Section 2.2.4, convolution will play an important role. For this reason, we need to know how to manage its Fourier transform.

Proposition 3.12. *Let $\phi, \varphi \in \mathcal{S}(\mathbb{R}^n)$. Then,*

1. $$\mathcal{F}(\phi * \varphi)(\xi) = \mathcal{F}\phi(\xi)\mathcal{F}\varphi(\xi). \tag{3.21}$$

2. $$\mathcal{F}(\phi\varphi)(\xi) = \frac{1}{(2\pi)^n}(\mathcal{F}\phi * \mathcal{F}\varphi)(\xi). \tag{3.22}$$

Proof. Formula (3.21) has to be justified by Fubini's theorem, to eventually use a change of variables. For property 2, we know by (3.21) that

$$\mathcal{F}(\phi * \varphi) = \mathcal{F}\phi \cdot \mathcal{F}\varphi.$$

Now, by properties 1 and 2 in Proposition 3.11,

$$\frac{1}{(2\pi)^n}\mathcal{F}^{-1}(\phi * \varphi) = \mathcal{F}^{-1}\phi \cdot \mathcal{F}^{-1}\varphi,$$

and if we consider Fourier transforms $\hat{\phi}, \hat{\varphi}$ instead, we get

$$\phi \cdot \varphi = \mathcal{F}^{-1}\hat{\phi} \cdot \mathcal{F}^{-1}\hat{\varphi} = \frac{1}{(2\pi)^n}\mathcal{F}^{-1}(\hat{\phi} * \hat{\varphi}).$$

So finally, by Theorem 3.9,

$$\mathcal{F}(\phi \cdot \varphi) = \frac{1}{(2\pi)^n}(\hat{\phi} * \hat{\varphi}).$$ □

Another famous property concerning the Fourier transform is the Plancherel equality. It is an interesting feature, as it can be used to extend the transform from $\mathcal{S}$ to L^2.

Proposition 3.13 (Plancherel equality). *Let $\phi, \varphi \in \mathcal{S}$. Then,*

$$\int_{\mathbb{R}^n} \phi(t)\overline{\varphi}(t)dt = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \mathcal{F}\phi(t)\overline{\mathcal{F}\varphi}(t)dt.$$

In particular,

$$\int_{\mathbb{R}^n} |\phi(t)|^2 dt = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} |\mathcal{F}\phi(t)|^2 dt.$$

Proof. Substitute φ for $\overline{\mathcal{F}\varphi}$ in (3.17), and observe that

$$\overline{\mathcal{F}\varphi}(\xi) = \int_{\mathbb{R}^n} \overline{\varphi}(x)e^{i\xi\cdot x}dx = (2\pi)^n\mathcal{F}^{-1}\overline{\varphi}(\xi). \qquad \square$$

The equality $||\mathcal{F}\phi||_2 = (2\pi)^{n/2}||\phi||_2$ for $\phi \in \mathcal{S}$ and the fact that $\mathcal{S}$ is dense in L^2 can be used to extend the Fourier transform to L^2. Nevertheless, for our purposes, we need to extend it to the space of tempered distributions, where we know L^2 is contained.

3.3 The Fourier transform in $\mathcal{S}'$

Once we have defined the Fourier transform in the space $\mathcal{S}$ and studied some of its properties, it is time to capitalise on them to define the transform in the space $\mathcal{S}'$ of tempered distributions. In fact, Theorem 3.6 is a great boon to do so. Thus, to the operations described in Section 2.2.3 we will be able to add one more. Throughout this section, we will present many properties concerning the Fourier transform of tempered distributions. Because of the definition to be given, they will be a direct consequence of those concerning Schwartz functions, so for the proofs we will offer references to Section 3.2, without giving many details.

3.3.1 Definition and properties

Definition 3.14. Let $f \in \mathcal{S}'(\mathbb{R}^n)$ be a tempered distribution. We define the **Fourier transform** of f as a functional over $\mathcal{S}(\mathbb{R}^n)$, denoted by $\mathcal{F}f$ or $\hat{f}$, given by

$$\langle \mathcal{F}f, \phi\rangle = \langle f, \mathcal{F}\phi\rangle, \quad \forall \phi \in \mathcal{S}(\mathbb{R}^n).$$

In the same way every operation analysed before has a justification in terms of regular distributions, Definition 3.14 can also be justified by considering Schwartz functions as tempered distributions. In fact, if ϕ is a Schwartz function and we treat it as a distribution, then for every $\varphi \in \mathcal{S}$,

$$\langle \mathcal{F}\phi, \varphi\rangle = \int_{\mathbb{R}^n} \mathcal{F}\phi(x)\varphi(x)dx,$$

and by property 3 in Proposition 3.8, we can write

$$\int_{\mathbb{R}^n} \mathcal{F}\phi(x)\varphi(x)dx = \int_{\mathbb{R}^n} \phi(x)\mathcal{F}\varphi(x)dx = \langle \phi, \mathcal{F}\varphi \rangle,$$

thus the definition. It is important to check that we are defining a tempered distribution.

Theorem 3.15. *The Fourier transform of every tempered distribution is a tempered distribution.*

Proof. The hardest part of this proof is already done in previous sections. Definition 3.14 and Proposition 3.5 show the new functional $\mathcal{F}f$ is well-defined. The linearity of both the Fourier transform in $\mathcal{S}$ and the tempered distribution f show that $\mathcal{F}f$ is also linear. Last, we saw in Theorem 3.6 that the Fourier transform is a continuous operator in $\mathcal{S}$, showing that if a sequence ϕ_n converges to zero in $\mathcal{S}$, then so does the sequence $\hat{\phi}_n$. Thus, because of the continuity of f, the sequence $\langle \mathcal{F}f, \phi_n \rangle = \langle f, \hat{\phi}_n \rangle$ converges to zero in $\mathbb{C}$, showing that $\mathcal{F}f$ is continuous. □

We can also define the inverse Fourier transform in $\mathcal{S}'$ in the same way we did it for the direct one.

Definition 3.16. Let $f \in \mathcal{S}'(\mathbb{R}^n)$ be a tempered distribution. Its inverse Fourier transform, denoted by $\mathcal{F}^{-1}f$ or $\check{f}$, is defined as follows:

$$\langle \mathcal{F}^{-1}f, \phi \rangle = \langle f, \mathcal{F}^{-1}\phi \rangle, \quad \forall \phi \in \mathcal{S}(\mathbb{R}^n).$$

Theorem 3.17. *The inverse Fourier transform of a tempered distribution is a tempered distribution.*

Proof. It is a direct consequence of Theorem 3.10 and an identical reasoning as in Theorem 3.15. □

Theorem 3.18. *The Fourier inversion formula still works in $\mathcal{S}'$. More precisely, for every tempered distribution f,*

$$\mathcal{F}\mathcal{F}^{-1}f = f \qquad \textit{and} \qquad \mathcal{F}^{-1}\mathcal{F}f = f.$$

Proof. It is trivial from Definition 3.14 and Theorem 3.9. □

Some more properties are given in the following proposition.

Proposition 3.19. *Let f be a tempered distribution and denote by $\tilde{f}$ its transposition. Let also $k \in (\mathbb{Z}^+)^n$ be a multi-index. Then*

1. $D^k(\mathcal{F}f) = (-i)^{|k|}\mathcal{F}(x^k f)$.
2. $x^k \mathcal{F}f = (-i)^{|k|}\mathcal{F}(D^k f)$.
3. $\mathcal{F}^2 f = (2\pi)^n \tilde{f}$.

4. $\mathcal{F}^4 f = (4\pi^2)^n f$.

Proof. The properties are direct consequence of those which hold for Schwartz functions, given in Propositions 3.4 and 3.11. □

It is also easy to show that the Fourier transform is a linear and continuous operator in the space of tempered distributions. For this, we must recall the concept of convergence of tempered distributions, which is absolutely identical to that of usual distributions defined in Section 1.2.3.

Theorem 3.20. *The Fourier transform operator $\mathcal{F} : \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ and the inverse Fourier transform operator $\mathcal{F}^{-1} : \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ are linear and continuous operators.*

Proof. The linearity being clear, let f_n be a sequence of tempered distributions convergent to f. Then, for $\mathcal{F}$, the continuity is obtained via

$$\langle \mathcal{F} f_n, \phi \rangle = \langle f_n, \mathcal{F}\phi \rangle \to \langle f, \mathcal{F}\phi \rangle = \langle \mathcal{F} f, \phi \rangle.$$

The same reasoning is valid for $\mathcal{F}^{-1}$. □

The Fourier transform interacts with the convolution of distributions in a very similar way it does with that of Schwartz functions.

Proposition 3.21. *Let $f \in \mathcal{S}'(\mathbb{R}^n)$ and $\psi \in \mathcal{S}(\mathbb{R}^n)$. Then,*

$$\mathcal{F}(f * \psi) = \mathcal{F} f \cdot \mathcal{F}\psi.$$

Proof. Let $\phi \in \mathcal{S}(\mathbb{R}^n)$. Then,

$$\langle \mathcal{F}(f * \psi), \phi \rangle = \langle f, \tilde{\psi} * \mathcal{F}\phi \rangle.$$

Now, as $\mathcal{F}^{-1}\mathcal{F} f = f$, we can write $\langle f, \tilde{\psi} * \mathcal{F}\phi \rangle = \langle \mathcal{F} f, \mathcal{F}^{-1}(\tilde{\psi} * \mathcal{F}\phi) \rangle$. Recall equation (3.22) from Proposition 3.12. Applying $\mathcal{F}^{-1}$ to that equation, we can write

$$\langle \mathcal{F} f, \mathcal{F}^{-1}(\tilde{\psi} * \mathcal{F}\phi) \rangle = \langle \mathcal{F} f, (2\pi)^n \mathcal{F}^{-1}(\tilde{\psi}) \cdot \phi \rangle.$$

It enough to observe that, by property 1 of Proposition 3.11,

$$(2\pi)^n \mathcal{F}^{-1}(\tilde{\psi})(\xi) = \mathcal{F}\psi(\xi)$$

to assert that

$$\langle \mathcal{F}(f * \psi), \phi \rangle = \langle \mathcal{F} f, \mathcal{F}\psi \cdot \phi \rangle = \langle \mathcal{F} f \cdot \mathcal{F}\psi, \phi \rangle.$$ □

Finally, we want to see what is the effect of applying the Fourier transform to a direct product of tempered distributions. The result is surprisingly simple.

Proposition 3.22. *Let $f \in \mathcal{S}'(\mathbb{R}^n)$ and $g \in \mathcal{S}'(\mathbb{R}^m)$. Then,*

$$\mathcal{F}(f(x) \cdot g(x))(\xi, \eta) = \mathcal{F} f(\xi) \cdot \mathcal{F} g(\eta).$$

Proof. Consider $\phi \in \mathcal{S}(\mathbb{R}^{n+m})$. By definition, we have

$$\langle \mathcal{F}(f(x) \cdot g(y))(\xi, \eta), \phi(\xi, \eta) \rangle = \langle f(x), \langle g(y), \mathcal{F}\phi(x, y) \rangle \rangle.$$

In this point, we need to remark the fact that **the Fourier transform can be computed in different stages concerning different variables**. This can be easily checked using the definition. This means that, for example,

$$\mathcal{F}\phi(x, y) = \mathcal{F}_\eta(\mathcal{F}_\xi[\phi])(x, y), \tag{3.23}$$

where

$$\mathcal{F}_\xi \phi(x, \eta) = \int_{\mathbb{R}^n} \phi(\xi, \eta) e^{-i(x \cdot \xi)} d\xi.$$

This method will also be used several times in Chapter 4. This way, by (3.23),

$$\begin{aligned} \langle f(x), \langle g(y), \mathcal{F}\phi(x, y) \rangle \rangle &= \langle f(x), \langle g(y), \mathcal{F}_\eta(\mathcal{F}_\xi[\phi])(x, y) \rangle \rangle \\ &= \langle f(x), \langle \mathcal{F}g(\eta), \mathcal{F}_\xi(\phi)(x, \eta) \rangle \rangle \\ &= \langle f(x) \cdot \mathcal{F}g(\eta), \mathcal{F}_\xi(\phi)(x, \eta) \rangle. \end{aligned}$$

Remember that we saw in (2.17) that the direct product is commutative. Thus,

$$\begin{aligned} \langle f(x) \cdot \mathcal{F}g(\eta), \mathcal{F}_\xi(\phi)(x, \eta) \rangle &= \langle \mathcal{F}g(\eta) \cdot f(x), \mathcal{F}_\xi(\phi)(x, \eta) \rangle \\ &= \langle \mathcal{F}g(\eta), \langle f(x), \mathcal{F}_\xi(\phi)(x, \eta) \rangle \rangle \\ &= \langle \mathcal{F}g(\eta), \langle \mathcal{F}f(\xi), \phi(\xi, \eta) \rangle \rangle \\ &= \langle \mathcal{F}f(\xi) \cdot \mathcal{F}g(\eta), \phi(\xi, \eta) \rangle. \end{aligned}$$

□

3.3.2 Examples

The Fourier transform plays an important role in the pursuit of fundamental solutions of differential equation, as we will analyse in the following chapters. For this reason, it is interesting to work out the Fourier transform of some distributions. In this section, we will work with the delta function and polynomials.

Example 3.23. We want to check that for every multi-index $k \in (\mathbb{Z}^+)^n$ and $\tau \in \mathbb{R}^n$,

$$\mathcal{F}(D^k[\delta(x - \tau)])(\xi) = i^{|k|} \xi^k e^{-i\xi \cdot \tau}. \tag{3.24}$$

For that, we only need to use properties we presented in Section 3.3.1. Indeed, choosing any Schwartz function ϕ,

$$\begin{aligned} \langle \mathcal{F}(D^k[\delta(x - \tau)]), \phi \rangle &= (-1)^{|k|} \langle \delta(x - \tau), D^k \hat{\phi}(x) \rangle \\ &= (-1)^{|k|} (-i)^{|k|} \langle \delta(x - \tau), \mathcal{F}(\xi^k \phi) \rangle \\ &= i^{|k|} \mathcal{F}(\xi^k \phi)(\tau) = i^{|k|} \int_{\mathbb{R}^n} \xi^k \phi(\xi) e^{-i(\tau \cdot \xi)} d\xi \\ &= \langle i^{|k|} \xi^k e^{-i(\xi \cdot \tau)}, \phi \rangle. \end{aligned} \tag{3.25}$$

In particular, for $k = 0$ and $\tau = 0$,

$$\mathcal{F}\delta(\xi) = 1(\xi). \tag{3.26}$$

Example 3.24. **Polynomials** are of great importance in many areas of analysis, so it is interesting to be able to work with their Fourier transform. Let us consider a very general form of a polynomial, with $k \in (\mathbb{Z}^+)^n$ and $\tau \in \mathbb{R}^n$,

$$i^{|k|}x^k e^{-ix\cdot\tau}.$$

We can expect to obtain a distribution related to the Delta function, by a similarity to equation (3.24). In fact,

$$\mathcal{F}(i^{|k|}x^k e^{-ix\cdot\tau}) = (-1)^{|k|}(2\pi)^n D^k\delta(\xi+\tau), \tag{3.27}$$

and we will use the mentioned (3.24) to prove so. Indeed, applying $\mathcal{F}$ and recovering results from Proposition 3.19, we get

$$\mathcal{F}(i^{|k|}x^k e^{-ix\cdot\tau}) = \mathcal{F}^2(D^k[\delta(\xi-\tau)])(\xi) = (2\pi)^n(D^k[\delta(\xi-\tau)])(-\xi).$$

Now, for any Schwartz function ϕ,

$$\langle (D^k[\delta(\xi-\tau)])(-\xi), \phi(\xi)\rangle = (-1)^{2|k|}D^k\phi(-\tau) = (-1)^{|k|}\langle D^k\delta(\xi+\tau), \phi\rangle,$$

and we get the result. As a direct consequence, we obtain

$$\mathcal{F}(x^k) = (2\pi)^n i^{|k|} D^k\delta(\xi), \tag{3.28}$$

and more basically,

$$\mathcal{F}1 = (2\pi)^n\delta(\xi). \tag{3.29}$$

In the same way, we obtain the Fourier transform of the complex exponential,

$$\mathcal{F}(e^{-ix\cdot\tau})(\xi) = (2\pi)^n\delta(\xi+\tau). \tag{3.30}$$

Finally, if we have a polynomial in $\mathbb{R}$ given by $P(x) = \sum_{k=0}^n a_k x^k$, where $a_k \in \mathbb{C}$ for every $\{1, \cdots, n\}$, by linearity, its Fourier transform is given by the following formula:

$$\mathcal{F}(P(x)) = 2\pi\sum_{k=0}^n a_k i^k \delta^{(k)}(x). \tag{3.31}$$

Example 3.25. Recall the Gaussian function from Section 3.2.2. We analysed the case in which the exponent is a non-real number with positive real part. A particular case is when the exponent is purely imaginary, say $k = bi, b \in \mathbb{R}$. We cannot argue the same way, but we can use continuity features. Recall that the Fourier transform $\mathcal{F}$ is a continuous operator in $\mathcal{S}'$. It is immediate, by the dominated convergence theorem, that

$$g_\epsilon(x) = e^{-(\epsilon+bi)|x|^2} \to e^{-bi|x|^2} \quad \text{in } \mathcal{S}'(\mathbb{R}^n) \text{ as } \epsilon \to 0^+.$$

For that reason, the convergence does not change even if we apply the Fourier transform, so by (3.13),

$$\mathcal{F}(g_\epsilon)(\xi) = \left(\sqrt{\frac{\pi}{\epsilon+bi}}\right)^n e^{-|\xi|^2/4(\epsilon+bi)} \to \mathcal{F}(e^{-bi|x|^2})$$

as $\epsilon \to 0$. On the other hand, taking the limit of the sequence, we get

$$\left(\sqrt{\frac{\pi}{bi}}\right)^n e^{-|\xi|^2/4bi} = \mathcal{F}(e^{-bi|x|^2}), \tag{3.32}$$

what shows that the formula shown in (3.13) is also valid for an imaginary number.

Chapter 4

Fundamental Solutions of Differential Operators

In this last chapter, we will develop an efficient and general theory to reach what we will call fundamental solutions of some differential operators. Eventually, we will apply this theory to analyse some representative cases such as the heat equation, the Schrödinger equation and the Laplace equation.

4.1 Generalised solutions and fundamental solutions

The first we need is to fix some notation. Let the following be a linear differential equation of order m:

$$\sum_{|\alpha|=0}^{m} a_\alpha(x) D^\alpha u = f, \tag{4.1}$$

in which $f \in \mathcal{D}'(\mathbb{R}^n)$ is a distribution and the coefficients a_α are C^∞ functions. To shorten the expressions to be used, we will denote the differential operator

$$L(x, D) = \sum_{|\alpha|=0}^{m} a_\alpha(x) D^\alpha, \tag{4.2}$$

thus the equation (4.1) turning into

$$L(x, D)u = f. \tag{4.3}$$

In general, we are looking for a distribution $u \in \mathcal{D}'(\mathbb{R}^n)$ which satisfies (4.3). Notice that this solution may not be a usual solution we are used to, for it need not be a function, and even if it is a function, its derivatives need not be functions. This is where the concept of generalised solution comes from (also remember that another name for distributions is generalised functions).

Definition 4.1. Let $L(x, D)$ be a differential operator, $f \in \mathcal{D}'(\mathbb{R}^n)$ and an open subset $A \subseteq \mathbb{R}^n$. We say that a distribution $u \in \mathcal{D}'(\mathbb{R}^n)$ is a **generalised solution of the equation $\boldsymbol{L(x, D)u = f}$** in the region A if

$$\langle L(x, D)u, \varphi\rangle = \langle f, \varphi\rangle$$

for every testing function $\varphi \in \mathcal{D}(\mathbb{R}^n)$ whose support is contained in A.

Consider now a differential operator with constant coefficients. If we keep notation as in (4.2), in this case we will write

$$L(x, D) = \sum_{|\alpha|=0}^{m} a_\alpha(x) D^\alpha = \sum_{|\alpha|=0}^{m} a_\alpha D^\alpha = L(D). \tag{4.4}$$

In this situation, we will be able to obtain some particular solutions which will be of extreme importance.

Definition 4.2. Let $L(D)$ be a differential operator with constant coefficients. We say that a distribution $E \in \mathcal{D}'(\mathbb{R}^n)$ is a **fundamental solution of the differential operator $\boldsymbol{L(D)}$** if E satisfies

$$L(D)E = \delta$$

in $\mathcal{D}'(\mathbb{R}^n)$.

We will work with these fundamental solutions from now on, but of course, it is important to know the reason for this. Notice that the connection between recently defined generalised solutions and fundamental solutions remains unknown. The following theorem will help to enlighten the situation.

Theorem 4.3. *Let $L(D)$ be a differential operator with constant coefficients and E a tempered fundamental solution of it. Let also f be a Schwartz function on $\mathbb{R}^n$. Then, a solution to the equation $L(D)u = f$ is given by*

$$u = E * f.$$

Proof. We know that $L(D)E = \delta$, as E is a fundamental solution of the operator $L(D)$. Now, if we consider $E * f$ as a candidate solution, and considering notation (4.4),

$$L(D)(E * f) = \sum_{|\alpha|=0}^{m} a_\alpha D^\alpha (E * f).$$

Recall the rules of differentiation of the convolution given in (2.23); we can write

$$\sum_{|\alpha|=0}^{m} a_\alpha D^\alpha (E * f) = \sum_{|\alpha|=0}^{m} a_\alpha (D^\alpha E * f).$$

Once we know this, by linearity we get

$$L(D)(E * f) = L(D)E * f = \delta * f = f, \tag{4.5}$$

as we saw in (2.16) that the delta function is the identity element with respect to convolution. As a consequence, the distribution $E * f$ is a solution to the equation in the statement. □

The reason for we bring fundamental solutions into the limelight is clear now. If we manage to obtain one for an operator, much of the work will be done. In any case, it is important to remark that, in general, they are not unique. Suppose we have been able to get a solution to the homogeneous equation $L(D)u = 0$. If we call it E_0, then

$$L(D)(E + E_0) = L(D)E + L(D)E_0 = \delta + 0 = \delta,$$

showing that $E + E_0$ is also a fundamental solution.

It would be interesting to get, if possible, a procedure to work out fundamental solutions. Here is where the Fourier transform comes into play.

Theorem 4.4. *Let $L(D)$ be a differential operator with constant coefficients and $E \in \mathcal{S}'(\mathbb{R}^n)$. Then, E is a fundamental solution of $L(D)$ if and only if*

$$L(i\xi)\mathcal{F}(E)(\xi) = 1,$$

where, according to notation (4.4), $L(x) = \sum_{|\alpha|=0}^{m} a_\alpha x^\alpha$ and $(i\xi)^\alpha = i^{|\alpha|}\xi^\alpha$.

Proof. Remember that E is a fundamental solution of $L(D)$ if $L(D)E = \delta$. Knowing this, apply the Fourier transform to the left hand side, considering the formulas in Proposition 3.19:

$$\mathcal{F}(L(D)E)(\xi) = \sum_{|\alpha|=0}^{m} a_\alpha \mathcal{F}(D^\alpha E)(\xi) = \sum_{|\alpha|=0}^{m} a_\alpha i^{|\alpha|}\xi^\alpha \mathcal{F}E(\xi) = \mathcal{F}E(\xi)L(i\xi). \tag{4.6}$$

Once we know this, consider $L(D)E = \delta$. Then, applying the Fourier transform and by (4.6),

$$\mathcal{F}(E)(\xi)L(i\xi) = \mathcal{F}\delta = 1.$$

Conversely, suppose $\mathcal{F}(E)(\xi)L(i\xi) = 1$. Then, by the inverse Fourier transform and (4.6),

$$\mathcal{F}(L(D)E) = 1 \Rightarrow L(D)E = \mathcal{F}^{-1}\mathcal{F}\delta = \delta.$$ □

The question is how we could exploit this interesting result. It conveys that it is enough to solve

$$L(i\xi)\mathcal{F}(E)(\xi) = 1,$$

or what we could expect to be the same,

$$\mathcal{F}(E)(\xi) = \frac{1}{L(i\xi)}. \tag{4.7}$$

But we must be careful at this point. To consider (4.7), we need to be sure that the expression on the right hand side is a distribution. Indeed, it is a usual function. If it were locally integrable and of slow growth, we could treat it as a tempered distribution and work with expression (4.7), being our solution

$$E = \mathcal{F}^{-1}\left(\frac{1}{L(i\xi)}\right).$$

But matters are not so simple if the function turns out not to be locally integrable. In that case, we will need to manage to obtain a tempered distribution F solving the equation

$$L(i\xi)F = 1,$$

to finally obtain the result by the inverse Fourier transform, $E = \mathcal{F}^{-1}(F)$.

4.2 Three fundamental solutions

In this section, we will make use of the method developed in Section 4.1 to obtain the fundamental solutions to the heat operator, the Laplace operator and the Cauchy-Riemann operator.

4.2.1 Fundamental solution of the heat operator

The very well-known heat equation models the evolution of the temperature in a certain space, which could be a stick or a plane, and even objects in spaces of greater dimension. Consider variable $x \in \mathbb{R}^n$ to be representative of space and variable $t \in \mathbb{R}$ of time. Then, the heat equation is given as

$$\frac{\partial u}{\partial t}(x,t) = a^2 \Delta_x u(x,t), \tag{4.8}$$

where Δ_x is the Laplace operator concerning variable x and a is a positive constant representing thermal diffusivity.

Basing on the heat equation (4.8), the heat operator $L_H(D)$ is

$$L_H(D) = D_t - a^2\Delta_x = D_t - a^2 \sum_{i=1}^{n} D_{x_i}^2, \tag{4.9}$$

where D_{x_i} represents the partial derivative with respect to the variable x_i. We know that if we want to obtain a fundamental solution of $L_H(D)$, we need to solve the equation

$$L_H(D)E = D_t E(x,t) - a^2\Delta_x E(x,t) = \delta(x,t).$$

Instead of applying the general Fourier transform as we did in Section 4.1, we will only consider the Fourier transform on the variable x. For that,

$$\mathcal{F}_x\left(\frac{\partial E}{\partial t}\right) - a^2\mathcal{F}_x(\Delta_x E) = \mathcal{F}_x(\delta). \tag{4.10}$$

An application of Lemma 2.21 yields to say that the Fourier transform in x and the partial derivative in t commute, as they are concerning different variables, so

$$\mathcal{F}_x\left(\frac{\partial E}{\partial t}\right) = \frac{\partial}{\partial t}\mathcal{F}(E)(\xi, t). \tag{4.11}$$

Also observe that

$$\mathcal{F}_x(\Delta_x E) = \sum_{i=1}^{n} \mathcal{F}_x(D^2_{x_i} E),$$

and by properties seen in 3.19,

$$\mathcal{F}_x(D^2_{x_i} E) = i^2 \xi_i^2 \mathcal{F}_x(E) = -\xi_i^2 \mathcal{F}_x(E).$$

This means that

$$\mathcal{F}_x(\Delta_x E) = -\sum_{i=1}^{n} \xi_i^2 \mathcal{F}_x(E) = -|\xi|^2 \mathcal{F}_x(E). \tag{4.12}$$

It remains to handle the right hand side of (4.10). In this case, by properties (2.18) and (3.22),

$$\mathcal{F}_x(\delta(x,t)) = \mathcal{F}_x(\delta(x)\cdot\delta(t)) = \mathcal{F}_x(\delta(x))\cdot\delta(t),$$

as can be easily checked. We also know that $F_x(\delta(x)) = 1(\xi)$, so

$$\mathcal{F}_x(\delta(x,t)) = 1(\xi)\cdot\delta(t). \tag{4.13}$$

Eventually, combining (4.10), (4.11), (4.12) and (4.13), we obtain

$$\frac{\partial}{\partial t}\mathcal{F}(E)(\xi,t) + a^2|\xi|^2\mathcal{F}_x(E)(\xi,t) = 1(\xi)\cdot\delta(t). \tag{4.14}$$

Observe that we have a differential equation with respect to variable t,

$$\frac{\partial}{\partial t}F(\xi,t) + k(\xi)F(\xi,t) = 1(\xi)\cdot\delta(t), \tag{4.15}$$

which in turn, if we fix values for ξ, is similar to

$$\frac{d}{dt}G(t) + kG(t) = \delta(t).$$

If G were a function, it could be expressed in terms of an exponential. Nevertheless, this is not a big setback, as we know that the Delta function is the derivative of the Heaviside function, H, so we will be able to give a solution in terms of distributions as

$$G(t) = H(t)e^{-kt}. \tag{4.16}$$

Indeed, $G'(t) = \delta(t)e^{-kt} - kG(t)$, and observe that by (1.7), $\delta(t)e^{-kt} = \delta(t)$, from where we get the result. Following this idea, an identical calculation shows that a solution to (4.15) is given by

$$F(\xi,t) = (1(\xi)\cdot H(t))e^{-a^2|\xi|^2 t},$$

where the only variation is that $D_t(1(\xi)\cdot H(t)) = 1(\xi)\cdot D_t H(t)$. But we can simplify that expression. In fact,

$$\begin{aligned}\langle(1(\xi)\cdot H(t))e^{-a^2|\xi|^2 t}, \varphi(\xi,t)\rangle &= \langle H(t), \langle 1(\xi), e^{-a^2|\xi|^2 t}\varphi(\xi,t)\rangle\rangle \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}^n} H(t)e^{-a^2|\xi|^2 t}\varphi(\xi,t)d\xi dt \\ &= \langle H(t)e^{-a|\xi|^2 t}, \varphi(\xi,t)\rangle.\end{aligned}$$

For this,

$$F(\xi,t) = H(t)e^{-a^2|\xi|^2 t}.$$

In short, if the fundamental solution we are looking for is E, then by the inverse Fourier transform,

$$E = \mathcal{F}_\xi^{-1}(H(t)e^{-a^2|\xi|^2 t}).$$

Clearly, the transform has no effect on H, and as properties 1 and 2 from Proposition 3.11 can be generalised to distributions,

$$E = \frac{H(t)}{(2\pi)^n}\mathcal{F}_\xi(e^{-a^2|\xi|^2 t}).$$

But we know how to work out the Fourier transform of the Gaussian; we analysed it in section 3.2.2. Thus,

$$E(x,t) = \frac{H(t)}{(2\pi)^n}\left(\sqrt{\frac{\pi}{a^2 t}}\right)^n e^{-|x|^2/4a^2 t} = \frac{H(t)}{(2a\sqrt{\pi t})^n}e^{-|x|^2/4a^2 t}. \tag{4.17}$$

4.2.2 Fundamental solution of the Laplace operator

The Laplace equation can be used to describe the behaviour of several potentials such as the electric, the gravitational and the fluid potentials. This fact makes it also be called the equation of the potential. We are looking for distributions E such that

$$\Delta E(x) = \delta(x). \tag{4.18}$$

As suggested in Theorem 4.4, the Fourier transform will be of great help. Indeed, we do not need to make any new calculations, as in the case of the heat equation, in (4.12), we got the Fourier transform of the Laplace operator. Hence,

$$-|\xi|^2\mathcal{F}(E)(\xi) = 1(\xi).$$

As seen in (4.7), we expect to get

$$\mathcal{F}(E)(\xi) = -\frac{1}{|\xi|^2}. \tag{4.19}$$

It cannot be done always anyway. In fact, we have problems in the plane. This is because in the unit ball B in $\mathbb{R}^n$,

$$\int_B \frac{1}{|x|^2}dx = \sigma(\mathbb{S}^{n-1})\int_0^1 r^{n-3}dr,$$

which is finite if $n \geq 3$, but not if $n = 2$. Thus (4.19) is not locally integrable in $\mathbb{R}^2$, and therefore, it is not a distribution. On the other hand, for every $n \geq 3$, (4.19) is locally integrable and of slow growth in $\mathbb{R}^n$, so we can work with it. We will work with these two cases separately. When $n \geq 3$, we will be able to get the solutions in a direct way. On the other hand, we will check that a particular distribution fulfils (4.18) in $\mathbb{R}^2$.

- Case $n \geq 3$. In this case, according to (4.19), we are looking for

$$E = -\mathcal{F}^{-1}\left(\frac{1}{|\xi|^2}\right).$$

We will make use of the Fourier transform of the Gaussian to get the result. Indeed, recall from Section 3.2.2 that

$$\mathcal{F}(e^{-k|x|^2})(\xi) = \left(\sqrt{\frac{\pi}{k}}\right)^n e^{-|\xi|^2/4k}.$$

Also recall that $\langle \mathcal{F}f, \phi \rangle = \langle f, \mathcal{F}\phi \rangle$ for every Schwartz function ϕ, from the definition of the Fourier transform in $\mathcal{S}'$. So taking f to be the Gaussian, we can write the following equality.

$$\left(\sqrt{\frac{\pi}{k}}\right)^n \int_{\mathbb{R}^n} \phi(\xi) e^{-|\xi|^2/4k} d\xi = \int_{\mathbb{R}^n} \hat{\phi}(x) e^{-k|x|^2} dx. \tag{4.20}$$

Next step is to integrate (4.20) from 0 to infinity with respect to k, to obtain

$$\int_0^\infty \left(\sqrt{\frac{\pi}{k}}\right)^n \int_{\mathbb{R}^n} \phi(\xi) e^{-|\xi|^2/4k} d\xi dk = \int_0^\infty \int_{\mathbb{R}^n} \hat{\phi}(x) e^{-k|x|^2} dx dk. \tag{4.21}$$

We will analyse each side of (4.21) separately.

  - The right hand side is a trivial calculation, as by Fubini's theorem,

$$\int_0^\infty \int_{\mathbb{R}^n} \hat{\phi}(x) e^{-k|x|^2} dx dk = \int_{\mathbb{R}^n} \hat{\phi}(x) \frac{e^{-k|x|^2}}{-|x|^2}\bigg|_0^\infty dx = \int_{\mathbb{R}^n} \frac{\hat{\phi}(x)}{|x|^2} dx$$
$$= \langle \frac{1}{|x|^2}, \hat{\phi} \rangle = \langle \mathcal{F}\left(\frac{1}{|x|^2}\right), \phi \rangle,$$

  and the value is finite for $1/|x|^2$ being locally integrable and of slow growth.

  - The left hand side is a bit more tricky. Again by Fubini's theorem, we can write equivalently

$$\int_{\mathbb{R}^n} \phi(\xi) \left(\int_0^\infty \left(\frac{\pi}{k}\right)^{n/2} e^{-|\xi|^2/4k} dk\right) d\xi.$$

Let us compute the inner integral. With a change of variables given by $r = |\xi|^2/4k$, after several steps, we obtain

$$\frac{(2\sqrt{\pi})^n}{4|\xi|^{n-2}}\int_0^\infty r^{\frac{n}{2}-2}e^{-r}dr.$$

This expression should remind us of the Gamma function. Indeed, as

$$\Gamma(k) = \int_0^\infty x^{k-1}e^{-x}dx,$$

we can write

$$\frac{(2\sqrt{\pi})^n}{4|\xi|^{n-2}}\int_0^\infty r^{\frac{n}{2}-2}e^{-r}dr = \frac{(2\sqrt{\pi})^n}{4|\xi|^{n-2}}\Gamma\left(\frac{n}{2}-1\right).$$

Thus we can write the whole integral as

$$\frac{(2\sqrt{\pi})^n}{4}\Gamma\left(\frac{n}{2}-1\right)\int_{\mathbb{R}^n}\frac{\phi(\xi)}{|\xi|^{n-2}}d\xi = 2^{n-2}\pi^{\frac{n}{2}}\Gamma\left(\frac{n}{2}-1\right)\langle\frac{1}{|\xi|^{n-2}},\phi\rangle.$$

So after working out each side of equality (4.21),

$$2^{n-2}\pi^{\frac{n}{2}}\Gamma\left(\frac{n}{2}-1\right)\langle\frac{1}{|\xi|^{n-2}},\phi\rangle = \langle\mathcal{F}\left(\frac{1}{|x|^2}\right),\phi\rangle, \quad \forall\phi\in\mathcal{S}(\mathbb{R}^n).$$

Because of this, we have the following equality of distributions:

$$\mathcal{F}\left(\frac{1}{|x|^2}\right) = 2^{n-2}\pi^{\frac{n}{2}}\Gamma\left(\frac{n}{2}-1\right)\frac{1}{|\xi|^{n-2}}. \tag{4.22}$$

Remember that we seek $\mathcal{F}^{-1}(1/|\xi|^2)$. We can use equation (4.22) for that, because we know that

$$\mathcal{F}^{-1}\left(\frac{1}{|\xi|^2}\right) = \frac{1}{(2\pi)^n}\mathcal{F}\left(\frac{1}{|\xi|^2}\right) = \frac{\Gamma\left(\frac{n}{2}-1\right)}{4(\sqrt{\pi})^n}\frac{1}{|x|^{n-2}}.$$

So eventually, the solution to (4.18) in $\mathbb{R}^n$ for $n \geq 3$ is given by

$$E(x) = -\frac{\Gamma\left(\frac{n}{2}-1\right)}{4(\sqrt{\pi})^n}\frac{1}{|x|^{n-2}}.$$

- Case $n = 2$. As we stated before, we cannot make the same reasoning for $\mathbb{R}^2$. Instead, we will directly prove that the regular distribution $E(x) = \log|x|$ satisfies (4.18) excepting for a constant. We will prove it by a process of convergence. Consider

$$E_n(x) = \frac{1}{2}\log\left(|x|^2 + \frac{1}{n^2}\right), \quad n\in\mathbb{N},$$

and observe that every E_n defines a regular distribution for being continuous and thus locally integrable, being also of slow growth. We want to see that $E_n \to E$ in $\mathcal{S}'(\mathbb{R}^2)$ if $n \to \infty$. Let $\phi \in \mathcal{S}(\mathbb{R}^2)$. Then,

$$\langle E_n, \phi \rangle = \frac{1}{2} \int_{\mathbb{R}^2} \log \left(|x|^2 + \frac{1}{n^2} \right) \phi(x) dx.$$

Observe that $|\log(|x|^2 + 1/n^2)\phi(x)| \leq (|\log(|x|^2 + 1)| + |\log(|x|^2)|)|\phi(x)|$, and the right-hand side function is integrable, because logarithms are locally integrable functions of slow growth and ϕ is Schwartz, so the sum of the products is integrable. For this, we can use the dominated convergence theorem and say that

$$\lim_{n\to\infty} \langle E_n, \phi \rangle = \frac{1}{2} \int_{\mathbb{R}^2} \log |x|^2 \phi(x) dx = \int_{\mathbb{R}^2} \log |x| \phi(x) dx = \langle E, \phi \rangle.$$

This way, $E_n(x) \to E(x)$ in $\mathcal{S}'(\mathbb{R}^2)$. Now, observe that, as a function,

$$\frac{\partial^2}{\partial^2 x_i} E_n(x) = \frac{|x|^2 + \frac{1}{n^2} - 2x_i^2}{(|x|^2 + 1/n^2)^2}, \quad i = 1, 2,$$

so

$$\Delta E_n(x) = \frac{2/n^2}{(|x|^2 + 1/n^2)^2}.$$

Now, as the derivative is a continuous operator, as we saw in (2.20), we can write

$$\lim_{n\to\infty} \Delta E_n(x) = \Delta E(x).$$

Because of this, for every Schwartz function ϕ,

$$\langle \Delta E(x), \phi \rangle = \lim_{n\to\infty} \langle \Delta E_n(x), \phi \rangle = \lim_{n\to\infty} \int_{\mathbb{R}^2} \frac{2/n^2}{(|x|^2 + 1/n^2)^2} \phi(x) dx.$$

The integral of the right hand side can be transformed to the following by a change of variables $x \to x/n$:

$$\int_{\mathbb{R}^2} \frac{2/n^2}{(|x|^2 + 1/n^2)^2} \phi(x) dx = \int_{\mathbb{R}^2} \frac{2}{(|x|^2 + 1)^2} \phi\left(\frac{x}{n}\right) dx.$$

The question now is if we can commute the limit with the integral. The answer is positive, as for being Schwartz, ϕ is bounded, and by a change to polar coordinates,

$$\int_{\mathbb{R}^2} \frac{2}{(|x|^2 + 1)^2} dx = 2\pi \int_0^\infty \frac{r dr}{(1 + r^2)^2} = 2\pi.$$

So taking every step done together,

$$\langle \Delta E(x), \phi \rangle = \int_{\mathbb{R}^2} \frac{2}{(|x|^2 + 1)^2} \phi(0) dx = 2\pi\phi(0) = 2\pi\langle \delta, \phi \rangle.$$

In short, we have obtained $\Delta E(x) = 2\pi\delta(x)$. It is almost what we need. Eventually, linearity of distributions makes it possible to assert that a fundamental solution of the Laplace operator in $\mathbb{R}^2$ is

$$\frac{1}{2\pi}\log|x|.$$

4.2.3 Fundamental solution of the Cauchy-Riemann operator

The Cauchy-Riemann equations are a system of two partial differential equations which play an important role in the characterisation of holomorphic complex functions. Let $f(x,y) = u(x,y) + iv(x,y)$ be a complex function, where u and v are real-valued. In this case, the Cauchy-Riemann equations concerning f are given by

$$u_x = v_y \quad \text{and} \quad u_y = -v_x.$$

To obtain an expression for the operator, let us define the following operators in $\mathbb{R}^2$:

$$D_{\overline{z}} = \frac{1}{2}\left(D_x + iD_y\right) \quad \text{and} \quad D_z = \frac{1}{2}\left(D_x - iD_y\right).$$

One can prove that the condition $D_{\overline{z}}f = 0$ is the same as saying that u and v satisfy the Cauchy-Riemann conditions. It is also easy to check that

$$D_{\overline{z}}D_z f = D_z D_{\overline{z}} f = \frac{1}{4}\Delta f.$$

We have seen in Section 4.2.2 that the function $\frac{1}{4\pi}\log|x^2+y^2|$ is the fundamental solution of the Laplace operator in the plane. This can be used to obtain the fundamental solutions to both $D_{\overline{z}}$ and D_z. Indeed, calling $E = \frac{1}{4\pi}\log|x^2+y^2|$,

$$D_{\overline{z}}D_z E = D_z D_{\overline{z}} E = \frac{1}{4}\Delta E = \frac{1}{4}\delta(x,y).$$

Therefore, it is clear that the fundamental solutions of the operators $D_{\overline{z}}$ and D_z are, respectively, $4D_zE$ and $4D_{\overline{z}}E$, or explicitly,

$$\frac{1}{\pi}D_z\log|x^2+y^2| = \frac{1}{\pi z} \quad \text{and} \quad \frac{1}{\pi}D_{\overline{z}}\log|x^2+y^2| = \frac{1}{\pi\overline{z}}.$$

4.3 Fundamental solution of the Cauchy problem

It is also possible to apply the idea of fundamental solutions to Cauchy problems. Let $L(D)$ be the differential operator in (4.4). The Cauchy problem associated to it, with $f \in \mathcal{S}(\mathbb{R}^n)$, is to find $u(x,t)$ such that

$$\begin{cases} D_t u(x,t) - L(D)u(x,t) = 0, & \text{for } t > 0, x \in \mathbb{R}^n, \\ u(x,0) = f(x), & \text{for } x \in \mathbb{R}^n. \end{cases} \tag{4.23}$$

We say that the distribution $E(x,t)$ is a fundamental solution to (4.23) if

$$\begin{cases} D_t E(x,t) - L(D)E(x,t) = 0, & \text{for } t > 0, x \in \mathbb{R}^n, \\ E(x,0) = \delta(x), & \text{for } x \in \mathbb{R}^n. \end{cases}$$

Indeed, if E is a fundamental solution, then the convolution $u = E *_x f$ solves (4.23). By $*_x$ we are denoting the convolution with respect to the variable x. We will analyse a couple of examples to illustrate these facts.

4.3.1 Fundamental solution of the Cauchy problem for the heat operator

When considering the heat equation, it is very common to work in a domain with some physical meaning. In fact, we are trying to know how will the heat of a domain evolve with time, so it is natural to consider $t > 0$. But this extra condition requires that we fix an initial condition at time $t = 0$. This is what we call a Cauchy problem. In the case of the heat equation, it looks like this:

$$\begin{cases} u_t(x,t) - a^2 \Delta_x u(x,t) = 0, & x \in \mathbb{R}^n, \quad t \geq 0, \\ u(x,0) = f(x), & x \in \mathbb{R}^n. \end{cases} \tag{4.24}$$

As stated before, we will work with $f \in \mathcal{S}(\mathbb{R}^n)$. In the following lines, we will try to obtain a solution through a fundamental solution. In the case of the heat operator, we need to solve

$$\begin{cases} E_t(x,t) - a^2 \Delta_x E(x,t) = 0, & \text{for } t > 0, x \in \mathbb{R}^n, \\ E(x,0) = \delta(x), & \text{for } x \in \mathbb{R}^n. \end{cases} \tag{4.25}$$

The method for solving this problem is to use the Fourier transform in variable x in a similar way we did in 4.2.1. So applying $\mathcal{F}$ to the first equation in (4.25),

$$\mathcal{F}_x(E_t)(\xi,t) - a^2 \mathcal{F}_x(\Delta_x E)(\xi,t) = 0,$$

which by the same reasoning as in 4.2.1 leads to

$$\frac{\partial}{\partial t}\mathcal{F}_x(E)(\xi,t) + a^2|\xi|^2 \mathcal{F}_x(E)(\xi,t) = 0.$$

If we now consider E to be a standard function, we know that

$$\mathcal{F}_x(E)(\xi,t) = c(\xi)e^{-a^2|\xi|^2 t}, \tag{4.26}$$

for some function $c(\xi)$. This function can be determined by using the initial condition. In fact,

$$\mathcal{F}_x(E)(\xi,0) = \mathcal{F}_x(E(x,0))(\xi) = \mathcal{F}(\delta)(\xi) = 1(\xi),$$

and we deduce that,

$$\mathcal{F}_x(E)(\xi,t) = e^{-a^2|\xi|^2 t}. \tag{4.27}$$

So we have to compute its inverse Fourier transform in order to get the fundamental solution function E. Observe that the function of which inverse Fourier transform we have to compute is the same we computed to obtain (4.17), so

$$E(x,t) = \mathcal{F}_\xi^{-1}(e^{-a^2|\xi|^2 t}) = \frac{1}{(2a\sqrt{\pi t})^n} e^{-|x|^2/4a^2 t}$$

is the fundamental solution to the Cauchy problem of the heat operator. Now, as we said in the beginning of the section, the solution to the Cauchy problem is given by a convolution in the variable x,

$$u(x,t) = f *_x \frac{1}{(2a\sqrt{\pi t})^n} e^{-|x|^2/4a^2 t} = \frac{1}{(2a\sqrt{\pi t})^n} \int_{\mathbb{R}^n} f(y) e^{-|x-y|^2/4a^2 t} dy. \qquad (4.28)$$

4.3.2 Fundamental solution of the Cauchy problem for the Schrödinger operator

The Schrödinger equation describes the change in the quantum state of a physical system over time. It is of great importance in quantum mechanics and in the analysis of the wave function of several systems. One of the particular forms it takes is the following, involving a wave function u:

$$u_t(x,t) = ik\Delta_x u(x,t), \qquad (4.29)$$

where k is a positive number related to Planck's constant. Again, variable t represents time, and it is usual to have an initial condition,

$$u(x,0) = f(x).$$

In this section, we will try to obtain a solution for the Cauchy problem corresponding to the Schrödinger equation, so we will try to obtain a fundamental solution E which fulfils (4.29) and also

$$E(x,0) = \delta(x).$$

Once we get that, the solution will be of the form $E *_x f$.

In the same way we did for the heat equation, we will make use of the Fourier transform. The similarity of the Schrödinger operator to the heat operator is obvious; the only important change is the complex number i. Nevertheless, this little variation generates a need of caution in every computation, although the result will be analogous to (4.28). The procedure is the same, so as the Fourier transform does not affect constants, every step repeats until (4.27). Hence,

$$\mathcal{F}_x(E)(\xi,t) = e^{-ik|\xi|^2 t},$$

so we need to compute the inverse Fourier transform of a function which looks like a Gaussian. In this case, the exponent is a complex number whose real part is zero.

In this moment, we must recall Example 3.25, and more precisely, (3.32). With that result, and also with Property 1 in Proposition 3.11, we can write

$$E(x,t) = \frac{1}{(2\pi)^n} \left(\sqrt{\frac{\pi}{ikt}} \right)^n e^{-|x|^2/4kti},$$

which is the fundamental solution to the Cauchy problem for the Schrödinger operator. Finally, the solution to the original problem is given by

$$u(x,t) = E *_x f = \frac{1}{(2\sqrt{\pi kti})^n} \int_{\mathbb{R}^n} f(y) e^{i|x-y|^2/4kt} dy.$$

4.4 The Malgrange-Ehrenpreis theorem

The Malgrange-Ehrenpreis theorem was one of the first results in which the potential of distributions was clearly shown to the mathematical community. Dated in the mid 1950s, it was proved independently by B. Malgrange[1] and L. Ehrenpreis[2]. The same way we have been able to obtain fundamental solutions to the heat operator and the Laplace operator in Section 4.2, the theorem states that it is possible to do so for any partial differential operator with constant coefficients.

Theorem 4.5 (Malgrange-Ehrenpreis)**.** *Every non-zero partial differential operator with constant coefficients admits a fundamental solution in the space of distributions* $\mathcal{D}'$*.*

Several different proofs have been published so far, also using completely independent arguments. The former proofs were non-constructive, and they made use of the Hahn-Banach theorem. Other proofs were based on the L^2-theory. Some of them also assert that the fundamental solutions can be taken from the space of tempered distributions. In the last decades, constructive proofs have been developed, with explicit expressions for a solution which have been turning more and more compact. A recent short proof using Fourier transforms is in [9], where references to other proofs can be found.

[1]Bernard Malgrange (born in 1928, Paris) was a student of Laurent Schwartz. He has been professor at Strasbourg, Orsay and Grenoble, and he works on differential equations.

[2]Leon Ehrenpreis (1930-2010, Brooklyn, NY) was professor at Temple University, Philadelphia, PA.

Appendix A

Laurent Schwartz: Life and Work

Laurent Schwartz was a brilliant mathematician who not only focused on scientific activities during his life. He was also very committed to French and international social and political problems.

He was born in Paris in 1915 to a non-practicing Jewish family. His father was the first Jew surgeon in a hospital in Paris, in a time when anti-Semitism was on the rise in France. Mathematics were present in his childhood; his uncle was a teacher, and the famous Jacques Hadamard was the uncle of his mother. Even if he was good at them, he showed better abilities in literature and languages, but advised by some of his teachers, and after a hard work, he enrolled in the École Normale Supérieure (ENS), where he met the daughter of the also mathematician Paul Lévy, Marie-Hélène. They married some years later.

The environment at ENS made him get close to communist, indeed Trotskyist ideas, an inclination he would never leave behind, and which would cause him some problems. The young Laurent finished his studies in 1937, and after two years of military service, he moved, together with his wife, to Toulouse, where both of them started working in a research institute. Nevertheless, their condition of Jews made their situation extremely delicate after France fell under Nazi control in 1940 and the pro-axis Vichy government led by Marshal Philippe Pétain imposed restrictive laws in the area. They decided to move to the city of Clermont-Ferrand for academic reasons, where the University of Strasbourg had been transferred during the war. It was there where Schwartz made contact with the Bourbaki group, a collective which had a decisive influence in his academic development. After finding himself forced to move to the Italian border for political reasons, he stayed in Grenoble for some months. In this time, and during the following years, he started to investigate in several fields, such as Fréchet spaces and generalised derivatives, after reading articles about harmonic and polyharmonic functions. One night, he came up with the idea of generalised functions in *the most beautiful night of his life*.

Once the war was over, he became professor of the University of Grenoble for a little time before moving to Nancy, city where he worked in his ideas, and published several articles in which he presented many concepts and results concerning distributions. These innovative contributions made him well-known in the whole country and even internationally. In fact, he received a Fields medal for his work in 1950 in Harvard. In the following year he published his two-volume treatise *Théorie des distributions*, where he put all his discoveries together.

In 1952, he became professor of the University of Paris. Some years later, in 1958, his father-in-law Paul Lévy retired as professor in the prestigious, though lately old-fashioned École Polytechnique. Even if he was not intending to make a request for the position at first, he was persuaded by representatives of the École, in an attempt to modernise their institution. He accepted, and from that time on, until he died, he carried an important work in the modernisation not only of the École, but also of the whole French university system. He rapidly revitalised the mathematical activity in a centre which had fallen in a situation of academic paralysis.

He stayed at the École Polytechnique until 1980, but it was not his last work at all. In 1985, he became chairman of the National Evaluation Committee, an organisation created by the French government after his advise. He also was member of the French Academy of Science since 1975. He died in Paris in 2002 at the age of 87.

Distributions were not his unique area of interest, and he worked in several fields during his life. Before developing the theory which made him famous, he worked on polynomials, semi-periodic functions, harmonic analysis and exponential sums. His thesis in 1942 was indeed involving this last matter. Once he published his main work *Théorie des distributions*, a second edition with some corrections and more information was reprinted in 1966, and extended the field with vectorial distributions and the kernel theorem. Afterwards, he succeeded in discovering applications of his work to theoretical physics. He also made important contributions to the field of probability and integration, especially in his latter works.

As stated in the beginning, Laurent Schwartz was a politically active man. He enrolled in the Trotskyst party in France during his stay in ENS, and excepting the lapse of the war, he was an active militant. He even presented himself for candidate in the legislative elections in 1945 and in 1946, but he quitted in 1947. He no longer took part in political parties, but anyway he had serious trouble when entering the USA in 1950 to receive his Fields medal. His colleague Marshall Stone discovered he was classified as a *dangerous communist* for the federal government.

Schwartz also played an active part in the decolonisation processes of the French Indochina in 1954 and Morocco and Tunisia in 1956, and in 1960, together with many intellectuals as Jean Paul Sartre and Simone de Beauvoir, he made a public manifesto for the right of the French youth not to take part in the Algerian War. This last action had a professional cost for Schwartz; indeed, he was working at the Ministry of Defence-run École Polytechnique, from which he was fired in 1961. Nevertheless, he was readmitted in 1963.

In the next decade, he made a strong campaign against the Vietnam War. He even travelled to that country and met the communist leader Ho Chi Minh, being part of the Russell Tribunal to find evidence of war crimes. In the context of the Afghanistan War after the Soviet invasion, he was president of the *International*

Bureau for Afghanistan.

In a more academic context but in the same path of defence of the human rights, he tried to make public the repression some intellectuals were suffering in countries such as Bolivia, Chile, Czechoslovakia, Morocco, Uruguay or the USSR.

As an interesting anecdote, it is remarkable his deep interest in butterflies. In his journeys to tropical countries in which he was to give conferences, he would try to obtain new species for his large collection. There are even two species he discovered for the first time, and as the tradition goes, they were named after him (*Xylophanes schwartzi* and *Clanis schwartzi*). In his last years, his collection had almost 20.000 insects, which he donated to the Museum of Natural History.

References: [1], [4], [5].

Bibliography

[1] Fernando Bombal. Laurent Schwartz, el matemático que quería cambiar el mundo. *La Gaceta de la RSME, 6, 177-201*, 2003.

[2] Joan Cerdà. *Linear Functional Analysis.* American Mathematical Society, Providence, RI, 2010.

[3] Gerald B. Folland. *Introduction to Partial Differential Equations, second edition.* Princeton University Press, Princeton, NJ, 1955.

[4] Jesús Hernández. Matemático, entomólogo, persona decente. Laurent Schwartz, un mathématicien aux prises avec le siècle. *La Gaceta de la RSME, 2, 319-326*, 1999.

[5] Gilles Pisier. Vida y obra científica de Laurent Schwartz. *La Gaceta de la RSME, 7, 691-699*, 2004.

[6] Karen Saxe. *Beginning Functional Analysis.* Springer-Verlag New York, Inc., 2002.

[7] Robert Strichartz. *A Guide to Distribution Theory and Fourier Transforms.* CRC Press, Boca Ratón, FL, 1994.

[8] V. S. Vladimirov. *Equations of Mathematical Physics.* Marcel Dekker, Inc., New York, 1971.

[9] Peter Wagner. A New Constructive Proof of the Malgrange-Ehrenpreis Theorem. *Amer. Math. Monthly, 116, 457-462*, May 2009.

[10] A. H. Zemanian. *Distribution Theory and Transform Analysis: An Introduction to Generalized Functions, with Applications, reprint edition.* Dover Publications, New York, 1987.