Surveys in Mathematics and its Applications

ISSN 1842-6298 (electronic), 1843-7265 (print)
Volume 15 (2020), 1 – 137

A MATHEMATICAL PRESENTATION OF LAURENT SCHWARTZ'S DISTRIBUTIONS

Josefina Alvarez

Abstract. This is a mathematical presentation of Laurent Schwartz's distributions.

1 Introduction

The theory of distributions, as developed by Laurent Schwartz (1915-2002), is considered a great mathematical achievement of the twentieth century. Its development provided a rigorous setting in which formal objects, such as the so-called Dirac function, were fully understood and justified. It also augmented the bank of initial data and solutions for partial differential equations, and further incorporated the methods of functional analysis into the study of differential operators. Moreover, it further developed the theory of the Fourier transform, aided by the Lebesgue integral, another great development of the twentieth century. With the theory of distributions, a whole array of new spaces appeared, which, in turn, became some of the foremost examples in the theory of topological linear spaces.

This blend of fundamental subjects is one of the reasons why the theory of distributions has become an integral part of mainstream mathematics. To be sure, this has not always been the case. Lars Hörmander relates in [14] how his doctoral advisor reproached him for using distributions in his dissertation. John Synowiec (b. 1934) reproduces in [28] the following exchange with several of his professors, during his years as a doctoral student: "Distributions? You mean probability distributions? -No, Laurent Schwartz distributions.- Oh ... if you are interested in that sort of thing, you will have to talk to someone else. I don't have much use for them in my work."

These reticencies are long in the past. As an evidence of the position enjoyed by the theory of distributions, we recall the words of Jeffrey Rauch (b. 1945) in [20]: "Distribution theory has become so ubiquitous that one of the good things that a course in partial differential equations can do is to familiarize students with it."

As it is the case with most subjects, the theory of distributions did not spring

2020 Mathematics Subject Classification: 46F05; 46F10; 46F12.
Keywords: Distributions, test functions, derivatives, tensor product, convolution product, multiplicative product, Fourier transform.

full fleshed from Schwartz's forehead. In his book [25], Schwartz gives a thorough account of the many insights and open problems that inspired his work. The article by János Horváth (1927-2015) [15] is another good source for such matters.

We must mention also that Schwartz's work is not the only attempt to extend the notion of function. The article by Rauch [20], already mentioned, ends with an extensive bibliography that includes presentations *à la Schwartz*, as well as presentations that follow various other approaches. Thus, anyone wanting to become familiarized with the theory of distributions, can choose from an array of possibilities. Still, it might be fair to say that Schwartz's approach has been the most successful. In fact, for his work on distributions, Schwartz was awarded the Fields medal at the International Congress of Mathematicians held in Cambridge, Massachusetts, in 1950.

Understanding why and how a piece of mathematics has become what is now, requires learning about its historical development. It is surprising, and even amusing, to see the origins of a topic, firmly rooted in common sense and human need, regardless of the level of abstraction it might have achieved today. The theory of distributions is no exception. An account of its very interesting historical development is given in the book by Jesper Lützen (b. 1951) [18].

As for the purpose of our article, it is to provide a mathematical presentation of the theory of distributions on $\mathbb{R}^n$, *à la Schwartz*. The emphasis is on getting our hands "dirty", developing many calculations in detail. Except for a few remarks, rigorously worded, we do not dwell much on the topological structures that make distributions what they are. To put it bluntly, our presentation chooses, whenever possible, to do a calculation, over citing a theorem. To be sure, this is not always possible or practical, thus showing that, for instance, the theory of topological linear spaces has a sure footing in the foundations of Schwartz's distribution theory.

None of the material discussed in this article is new. It can be found, in one form or another, in Schwartz's book [25], or in other references that will be mentioned at the appropriate time. If any novelty can be adjudicated, it is, to the way in which some of the topics are developed.

This article has been a long time in the making. It began as a set of lecture notes, with some exercises, that I wrote in the seventies, as an introduction to a course on partial differential equations, taught by Professor Alberto P. Calderón, at the University of Buenos Aires. It was my great honor to be his doctoral student and his teaching assistant. At one point, Professor Calderón read the notes and judged them to be "very good" for its purpose. This is the best assurance I can offer, because he was a master at tailoring the tools he used, to the scope of the results he wanted.

The original notes, as well as later versions from which this article stems, have been used in doctoral level courses, at several universities in Argentina and at New Mexico State University in the USA.

For the purpose of preparing this article, the notes have gone through a final rewrite, expanding on the examples as well as on the treatment of some of the topics, and incorporating, fully explained, many of the exercises.

As for the necessary background knowledge, we make frequent use of results from measure, integration, and function space theory, mostly at the level of, for instance, [27]. Occasionally, we invoke results from other sources, which we cite. We also bring in a fair amount of functional analysis, for which we give references at the appropriate time.

The organization of the article is as follows: Before concluding this introductory section, we go over the notation to be used. In Section 2 we collect, with proofs, a few auxiliary results. Section 3 and Section 4 are dedicated to define and study several spaces central to the subject. Section 5 and Section 6 deal with operations such as differentiation, tensor product, convolution product, and multiplicative product. The common theme of the last three sections is the Fourier transform, which we present in the classical setting, as well as in the sense of distributions. The article ends with a list of references.

1.1 Notation

With $\mathbb{C}$ we denote the space of complex numbers, while $\mathbb{R}^n$ denotes the n-dimensional euclidean space consisting of n-tuples $x = (x_1, ..., x_n)$ of real numbers, with the euclidean norm $|x| = \left(x_1^2 + ... + x_n^2\right)^{1/2}$.

We denote $\mathbb{N}$ and $\mathbb{N}^n$ the space of non-negative integers and the n-tuples of non-negative integers, respectively. Given $\alpha = (\alpha_1, ..., \alpha_n)$ in $\mathbb{N}^n$, we write $|\alpha| = \alpha_1 + ... + \alpha_n$ and $\alpha! = \alpha_1!...\alpha_n!$.

If α and β belong to $\mathbb{N}^n$, $\alpha \leq \beta$ means $\alpha_j \leq \beta_j$ for all j. The notation $\alpha < \beta$, $\alpha \geq \beta$, etc. should then be clear.

Given α and β in $\mathbb{N}^n$, if $\alpha \geq \beta$ we write

$$\binom{\alpha}{\beta} \stackrel{def}{=} \binom{\alpha_1}{\beta_1}...\binom{\alpha_n}{\beta_n},$$

which in turn can be written as

$$\frac{\alpha!}{\beta!\,(\alpha-\beta)!}.$$

If $x \in \mathbb{R}^n$ and $\alpha \in \mathbb{N}^n$,

$$x^\alpha \stackrel{def}{=} x_1^{\alpha_1}...x_n^{\alpha_n}.$$

Given $x, \xi \in \mathbb{R}^n$, $x\xi$ denotes the scalar product $x_1\xi_1 + ... + x_n\xi_n$.

When $\alpha \in \mathbb{N}^n$, ∂^α is the partial derivative

$$\partial_{x_1}^{\alpha_1}...\partial_{x_n}^{\alpha_n}.$$

If necessary, we indicate the variable on which the partial derivative is calculated as ∂_x^α. If it is important to recognize the order in which the partial derivatives are performed, we write, for instance,

$$\partial_{x_{j_1}} ... \partial_{x_{j_m}}.$$

In the case of derivatives in one variable, we might write $\frac{d}{dx}$ or f', $\frac{d^k}{dx^k}$ or $f^{(k)}$, etc..

If a space is generically considered on $\mathbb{R}^n$, we just write, L^1, $\mathcal{D}$, $\mathcal{S}$, etc. If the space is considered in a different context, such as $\mathbb{R}$ or a subset U of $\mathbb{R}^n$, etc., we indicate it as $L^1(\mathbb{R})$, $\mathcal{D}(U)$, etc..

Depending on the particular situation, a letter might denote a number, or a function, or a derivative. The specific meaning will always be clear.

The domain of a sequence or the summation range of a series will be, for instance, $l \geq 1$. The meaning of $r \geq 1$ or $r > 1$, etc., should be clear. If we want to indicate $m \geq 1$ and $m \in \mathbb{N}$, we will write $m = 0, 1, ...$ or $m = 1, 2, ...$

That a parameter, say C, or j, or δ, depends on other parameters, say n, m, ε, will be denoted $C_{n,m}$, j_m, δ_ε, etc.

As usual, the letter C, with or without subindexes attached, will indicate a positive constant that might be different at different occurrences.

Other notation will be explained at the appropriate time.

2 Preliminaries

Lemma 1. *(multinomial expansion)Given $x \in \mathbb{R}^n$ for $n \geq 2$, and $m \geq 1$,*

$$|x|^{2m} = \sum_{|\alpha|=m} \binom{m}{\alpha_1 + ... + \alpha_{n-1}} \binom{\alpha_1 + ... + \alpha_{n-1}}{\alpha_1 + ... + \alpha_{n-2}} ... \binom{\alpha_1 + \alpha_2}{\alpha_1} x^{2\alpha}. \tag{2.1}$$

The expansion (2.1) has $\binom{m+n-1}{m}$ terms.

Proof. To prove (2.1) we will use induction on n. For $n = 2$, (2.1) becomes the binomial expansion

$$\left(x_1^2 + x_2^2\right)^m = \sum_{j=0}^{m} \binom{m}{j} x_1^{2j} x_2^{2(m-j)}. \tag{2.2}$$

So, if we assume (2.1) to be true for some $n \geq 2$, we can write, with $\beta' =$

$(\beta_1, ..., \beta_n)$, $\beta = (\beta_1, ..., \beta_{n+1})$, $x' = (x_1, ...x_n)$ and $x = (x_1, ...x_{n+1})$,

$$\begin{aligned}\left(\left|x'\right|^2 + x_{n+1}^2\right)^m &= \sum_{j=0}^{m} \binom{m}{j} \left|x'\right|^{2j} x_{n+1}^{2(m-j)} \quad (2.3)\\ &= \sum_{j=0}^{m} \binom{m}{j} \sum_{|\beta'|=j} \binom{j}{\beta_1 + ... + \beta_{n-1}} \binom{\beta_1 + ... + \beta_{n-1}}{\beta_1 + ...\beta_{n-2}} ... \\ &\quad \times \binom{\beta_1 + \beta_2}{\beta_1} x'^{2\beta'} x_{n+1}^{2(m-j)} \\ &= \sum_{|\beta|=m} \binom{m}{\beta_1 + ... + \beta_n} \binom{\beta_1 + ... + \beta_n}{\beta_1 + ... + \beta_{n-1}} ... \\ &\quad \times \binom{\beta_1 + \beta_2}{\beta_1} x^{2\beta}.\end{aligned}$$

As for the number of terms in the expansion, for $n = 2$, (2.2) shows that we have $m + 1 = \binom{m+1}{m}$ terms. If we assume that, for n, there are $\binom{m+n-1}{m}$ terms in the multinomial expansion (2.1), according to (2.3) there will be

$$\sum_{j=0}^{m} \binom{j + n - 1}{j}$$

terms for $n+1$. That is to say, if indeed we have $\binom{n+m}{m}$ terms for $n+1$, the equality

$$\sum_{j=0}^{m} \binom{j + n - 1}{j} = \binom{m + n}{m} \quad (2.4)$$

must be true for each $m \geq 1$.

We prove (2.4) by induction on m.

If $m = 1$, the left-hand side of (2.4) is $\binom{n-1}{0} + \binom{n}{1} = 1 + n$, while the right-hand side is $\binom{n+1}{1} = n + 1$.

If we assume that (2.4) holds for some $m \geq 1$,

$$\begin{aligned}\sum_{j=0}^{m+1} \binom{j + n - 1}{j} &= \sum_{j=0}^{m} \binom{j + n - 1}{j} + \binom{m + n}{m + 1} \\ &= \binom{m + n}{m} + \binom{m + n}{m + 1} = \binom{m + 1 + n}{m + 1}, \quad (2.5)\end{aligned}$$

which is the right-hand side of (2.4), with $m + 1$ instead of m.

Let us observe that in (2.5) we have used the identity

$$\binom{p}{q - 1} + \binom{p}{q} = \binom{p + 1}{q}, \quad (2.6)$$

which is fairly simple to verify.

This completes the proof of the lemma. □

Corollary 2. *Given $x \in \mathbb{R}^n$ for $n \geq 2$, and $m \geq 1$,*

$$\left(1+|x|^2\right)^m = \sum_{|\alpha|=0}^{m} \binom{m}{\alpha_1+...+\alpha_n} \binom{\alpha_1+...+\alpha_n}{\alpha_1+...+\alpha_{n-1}} ... \binom{\alpha_1+\alpha_2}{\alpha_1} x^{2\alpha}. \tag{2.7}$$

The expansion (2.7) has $\binom{m+n}{m}$ terms .

Proof. We begin by writing (2.1) with $y \in \mathbb{R}^{n+1}$ and $\beta \in \mathbb{N}^{n+1}$, for $n \geq 2$ and $m \geq 1$.

$$|y|^{2m} = \sum_{|\beta|=m} \binom{m}{\beta_1+...+\beta_n} \binom{\beta_1+...+\beta_n}{\beta_1+...+\beta_{n-1}} \binom{\beta_1+\beta_2}{\beta_1} y'^{2\beta'} x_{n+1}^{\beta_{n+1}}, \tag{2.8}$$

where $y' = (y_1, ..., y_n)$ and, as in Lemma 1, $\beta' = (\beta_1, ..., \beta_n)$.

If we take $y_{n+1} = 1$ in (2.8), we have

$$\left(1+|y'|^2\right)^m = \sum_{|\beta|=0}^{m} \binom{m}{\beta_1+...+\beta_n} \binom{\beta_1+...+\beta_n}{\beta_1+...+\beta_{n-1}} \binom{\beta_1+\beta_2}{\beta_1} y'^{2\beta'}.$$

To obtain (2.7) we only have to write x instead of y' and α instead of β'.

The number of terms in (2.7) equals the number of terms in the sum $\sum_{0\leq j\leq m} \sum_{|\alpha|=j}$. The work done in Lemma 1 shows that this sum is $\sum_{0\leq j\leq m} \binom{j+n-1}{j}$, which, according to (2.4), equals $\binom{m+n}{m}$.

This completes the proof of the corollary. □

Remark 3. The coefficients in (2.1), called multinomial coefficients, can be written in various ways. For instance,

$$\begin{aligned} & \binom{m}{\alpha_1+...+\alpha_{n-1}} \binom{\alpha_1+...+\alpha_{n-1}}{\alpha_1+...+\alpha_{n-2}} ... \binom{\alpha_1+\alpha_2}{\alpha_1} \\ = & \frac{m!}{\alpha_n!\,(\alpha_1+...+\alpha_{n-1})!} \frac{(\alpha_1+...+\alpha_{n-1})!}{\alpha_{n-1}!\,(\alpha_1+...+\alpha_{n-2})!} ... \frac{(\alpha_1+\alpha_2)!}{\alpha_2!\alpha_1!} \\ = & \frac{m!}{\alpha_1!\alpha_2!...\alpha_{n-1}!\alpha_n!} = \frac{m!}{\alpha!}. \end{aligned} \tag{2.9}$$

Likewise, here is a shorter formula for the coefficients in (2.7):

$$\begin{aligned} & \binom{m}{\alpha_1+...+\alpha_n} \binom{\alpha_1+...+\alpha_n}{\alpha_1+...+\alpha_{n-1}} ... \binom{\alpha_1+\alpha_2}{\alpha_1} \\ = & \frac{m!}{(m-\alpha_1-...-\alpha_n)!\,(\alpha_1+...+\alpha_n)!} \frac{(\alpha_1+...+\alpha_n)!}{\alpha_n!\,(\alpha_1+...+\alpha_{n-1})!} ... \frac{(\alpha_1+\alpha_2)!}{\alpha_1!\alpha_2!} \\ = & \frac{m!}{(m-\alpha_1-...-\alpha_n)!\alpha_1!\alpha_2!...\alpha_n!}. \end{aligned} \tag{2.10}$$

If we denote $C_{m,\alpha}$ the expression in (2.9) or (2.10), it is clear that, in each case, there is a constant $C_{m,n} > 0$ so that

$$\sup C_{m,\alpha} = C_{m,n},$$

where the supremum is taken over the appropriate values of α. In most cases, this bound and the number of terms in the expansion is all that matters.

Lemma 4. *(Leibniz's rule) Given functions φ and ψ continuous with continuous derivatives of order $\leq m$ for some $m \geq 1$, we have, for $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq m$,*

$$\partial^{\alpha}(\varphi\psi) = \sum_{\beta=0}^{\alpha} \binom{\alpha}{\beta} \left(\partial^{\beta}\varphi\right)\left(\partial^{\alpha-\beta}\psi\right). \tag{2.11}$$

There are $(\alpha_1+1)\ldots(\alpha_n+1)$ terms in (2.11).

Proof. When $|\alpha| = 1$, we can write $\partial^{\alpha} = \partial_{x_j}$, for some $1 \leq j \leq n$. So, (2.11) becomes

$$\partial^{\alpha}(\varphi\psi) = \partial_{x_j}(\varphi\psi) = \left(\partial_{x_j}\varphi\right)\psi + \varphi\left(\partial_{x_j}\psi\right),$$

which is the rule for taking one derivative of a product.

If $m \geq 2$, we assume (2.11) to be true for $|\alpha| = k$ with $k < m$. We can write, for some $1 \leq j \leq n$,

$$\begin{aligned} \partial_{x_j}\partial^{\alpha}(\varphi\psi) &= \partial_{x_j}\sum_{\beta=0}^{\alpha}\binom{\alpha}{\beta}\left(\partial^{\beta}\varphi\right)\left(\partial^{\alpha-\beta}\psi\right) \\ &= \sum_{\beta=0}^{\alpha}\binom{\alpha}{\beta}\left(\partial_{x_j}\partial^{\beta}\varphi\right)\left(\partial^{\alpha-\beta}\psi\right) + \sum_{\beta=0}^{\alpha}\binom{\alpha}{\beta}\left(\partial^{\beta}\varphi\right)\left(\partial_{x_j}\partial^{\alpha-\beta}\psi\right). \end{aligned} \tag{2.12}$$

Let us denote α' and β' the $(n-1)$-tuples that result from removing α_j and β_j from α and β, respectively. Then, we can write (2.12) as

$$\begin{aligned} &\sum_{\beta'=0}^{\alpha'}\binom{\alpha'}{\beta'}\sum_{\beta_j=0}^{\alpha_j}\binom{\alpha_j}{\beta_j}\left(\partial_{x_j}^{\beta_j+1}\partial^{\beta'}\varphi\right)\left(\partial_{x_j}^{\alpha_j-\beta_j}\partial^{\alpha'-\beta'}\psi\right) \\ &+\sum_{\beta'=0}^{\alpha'}\binom{\alpha'}{\beta'}\sum_{\beta_j=0}^{\alpha_j}\binom{\alpha_j}{\beta_j}\left(\partial_{x_j}^{\beta_j}\partial^{\beta'}\varphi\right)\left(\partial_{x_j}^{\alpha_j-\beta_j+1}\partial^{\alpha'-\beta'}\psi\right) \\ =\; &(1)+(2). \end{aligned}$$

If we write $\beta_j + 1 = \gamma_j$

$$(1) = \sum_{\beta'=0}^{\alpha'}\binom{\alpha'}{\beta'}\sum_{\gamma_j=1}^{\alpha_j+1}\binom{\alpha_j}{\gamma_j-1}\left(\partial_{x_j}^{\gamma_j}\partial^{\beta'}\varphi\right)\left(\partial_{x_j}^{\alpha_j+1-\gamma_j}\partial^{\alpha'-\beta'}\psi\right). \tag{2.13}$$

So, substituting β_j for γ_j in (2.13), we have

$$\begin{aligned}(1)+(2) &= \sum_{\beta'=0}^{\alpha'}\binom{\alpha'}{\beta'}\binom{\alpha_j}{\alpha_j}\left(\partial_{x_j}^{\alpha_j+1}\partial^{\beta'}\varphi\right)\left(\partial^{\alpha'-\beta'}\psi\right)\\ &+\sum_{\beta'=0}^{\alpha'}\binom{\alpha'}{\beta'}\binom{\alpha_j}{0}\left(\partial^{\beta'}\varphi\right)\left(\partial_{x_j}^{\alpha_j+1}\partial^{\alpha'-\beta'}\psi\right)\\ &+\sum_{\beta'=0}^{\alpha'}\binom{\alpha'}{\beta'}\sum_{\beta_j=1}^{\alpha_j}\left[\binom{\alpha_j}{\beta_j-1}+\binom{\alpha_j}{\beta_j}\right]\left(\partial_{x_j}^{\beta_j}\partial^{\beta'}\varphi\right)\left(\partial_{x_j}^{\alpha_j+1-\beta_j}\partial^{\alpha'-\beta'}\psi\right).\end{aligned}$$

According to (2.6),

$$\binom{\alpha_j}{\beta_j-1}+\binom{\alpha_j}{\beta_j}=\binom{\alpha_j+1}{\beta_j}.$$

So

$$\partial_{x_j}\partial^{\alpha}\left(\varphi\psi\right)=\sum_{\beta'=0}^{\alpha'}\binom{\alpha'}{\beta'}\sum_{\beta_j=0}^{\alpha_j+1}\binom{\alpha_j+1}{\beta_j}\left(\partial_{x_j}^{\beta_j}\partial^{\beta'}\varphi\right)\left(\partial_{x_j}^{\alpha_j+1-\beta_j}\partial^{\alpha'-\beta'}\psi\right).$$

As for the number of terms in (2.11), we only need to remember that the sum $\sum_{0\leq\beta\leq\alpha}$ means $\sum_{0\leq\beta_1\leq\alpha_1}\ldots\sum_{0\leq\beta_n\leq\alpha_n}$. Thus, the number of terms in the expansion is $(\alpha_1+1)\ldots(\alpha_n+1)$.

This completes the proof of the lemma. □

Remark 5. Let us observe that if $|\alpha|\leq m$ for some $m\geq 1$,

$$\sup_{0\leq\beta\leq\alpha}\binom{\alpha}{\beta}=C_{m,n},$$

for some constant $C_{m,n}>0$, independent of α and β. As in the case of the multinomial expansion, an estimate of the coefficients and the number of terms in the expansion is, often, all that is needed.

The next two results not only are well known, but also appear in several references (see, for instance, [27], p. 100, Theorem 1 and Theorem 3). Nevertheless, we are going to invoke them often, so it will be convenient to include them in this section, in a suitable form.

Theorem 6. *(continuity of an integral depending on a parameter) Let (S,Σ,μ) be a measure space and let $f:\mathbb{R}^n\times S\to\mathbb{C}$ be a function satisfying the following conditions:*

1. *$f(x,\cdot): S \to \mathbb{C}$ is μ-integrable for each $x \in \mathbb{R}^n$.*

2. *$f(\cdot,s): \mathbb{R}^n \to \mathbb{C}$ is continuous at $x_0 \in \mathbb{R}^n$, for each $s \in S$.*

3. *There is a μ-integrable function $g: S \to [0,\infty)$ so that $|f(x,s)| \le g(s)$ for all $x \in \mathbb{R}^n, s \in S$.*

Then, the function $F(x) = \int_S f(x,s)\,d\mu(s)$ is continuous at x_0.

Proof. Let $\{x_j\}_{j\ge 1}$ be a sequence in $\mathbb{R}^n$ converging to x_0 as $j \to \infty$. According to 2), $f(x_j,s)$ must converge to $f(x_0,s)$ in $\mathbb{C}$ as $j \to \infty$ for every $s \in S$. Moreover, according to 3), $|f(x_j,s)| \le g(s)$ for every $j \ge 1$. Then, Lebesgue's dominated convergence theorem implies that $F(x_j)$ converges to $F(x_0)$ in $\mathbb{C}$ as $j \to \infty$. Thus, the function F is continuous at x_0.

This completes the proof of the theorem. □

Remark 7. The continuity of a function at a point is a local property. Therefore, it would suffice to assume in 1) and 3) of Theorem 6 that the assumption holds for all x in a particular open neighborhood of x_0.

Theorem 8. *(derivative of an integral depending on a parameter) Let (S,Σ,μ) be a measure space and let $f: \mathbb{R}^n \times S \to \mathbb{C}$ be a function satisfying the following conditions:*

1. *$f(x,\cdot): S \to \mathbb{C}$ is μ-integrable for each $x \in \mathbb{R}^n$.*

2. *$\left(\partial_{x_j} f\right)(x,s)$ exists for some $1 \le j \le n$ and for each $s \in S$.*

3. *There is a μ-integrable function $g: S \to [0,\infty)$ so that $\left|\left(\partial_{x_j} f\right)(x,s)\right| \le g(s)$ for all $x \in \mathbb{R}^n, s \in S$.*

Then, $\left(\partial_{x_j} F\right)(x)$ exists and

$$\left(\partial_{x_j} F\right)(x) = \int_S \left(\partial_{x_j} f\right)(x,s)\,d\mu(s),$$

for $x \in \mathbb{R}^n$.

Proof. First of all, the function f can be written as $f_1 + if_2$, where the functions f_1 and f_2 are real-valued and satisfy conditions 1), 2) and 3). Furthermore, to simplify the notation, we assume that $j=1$ and we write, if $n>1$, $x' = (x_2,\dots,x_n)$.

Now, for $x = (x_1,\dots,x_n) \in \mathbb{R}^n$ fixed, we consider every sequence $\{a_k\}_{k\ge 1}$ converging to x_1 in $\mathbb{R}$ as $k \to \infty$, for which $a_k \ne x_1$ for all $k \ge 1$. Then, there is

$$\lim_{k\to\infty} \frac{f(a_k,x',s) - f(x,s)}{a_k - x_1} = \left(\partial_{x_1} f\right)(x,s),$$

for each $s \in S$. Moreover, the function

$$\frac{f(a_k, x', s) - f(x, s)}{a_k - x_1}$$

is a μ-integrable function of $s \in S$, for a_k and x fixed. Thus, for each $s \in S$ and $x' \in \mathbb{R}^{n-1}$ fixed, the mean value theorem tells us that there are real numbers y_{k,x_1} and z_{k,x_1} between a_k and x_1, so that

$$\frac{f(a_k, x', s) - f(x, s)}{a_k - x_1} = (\partial_{x_1} f_1)\left(y_{k,x_1}, x', s\right) + i\left(\partial_{x_1} f_2\right)\left(z_{k,x_1}, x', s\right).$$

Therefore, according to 3),

$$\left|\frac{f(a_k, x', s) - f(x, s)}{a_k - x_1}\right| \leq 2g(s),$$

for each $s \in S$ and $k \geq 1$. Then, Lebesgue's dominated convergence theorem implies that there is

$$\lim_{k \to \infty} \int_S \frac{f(a_k, x', s) - f(x, s)}{a_k - x_1} d\mu(s) = \int_S (\partial_{x_1} f)(x, s)\, d\mu(s).$$

On the other hand,

$$\int_S \frac{f(a_k, x', s) - f(x, s)}{a_k - x_1} d\mu(s) = \frac{F(a_k, x') - F(x)}{a_k - x_1}.$$

Thus, $(\partial_{x_1} F)(x)$ exists and

$$(\partial_{x_1} F)(x) = \int_S (\partial_{x_1} f)(x, s)\, d\mu(s).$$

This completes the proof of the theorem. □

Remark 9. As with continuity, the existence of a partial derivative at a point is a local property. So, it would be enough to assume in 1) and 3) of Theorem 8 that the statement holds for all x in a particular open subset of $\mathbb{R}^n$.

Remark 10. In both, Theorem 6 and Theorem 8, it can be assumed that the function f is defined on $U \times S$, for some $U \subseteq \mathbb{R}^n$ open. Still, in most cases, we will be working with $U = \mathbb{R}^n$.

When the measure space in Theorems 6 and 8 is the Lebesgue measure space on $\mathbb{R}^n$, we will write dx or dy, etc., as appropriate, instead of $d\mu(s)$.

We are now quite ready to go onto our presentation on distributions.

3 The spaces $\mathcal{D}$ and $\mathcal{D}'$

We begin with a definition.

Definition 11. *Let $\varphi : \mathbb{R}^n \to \mathbb{C}$ be a function. The support of φ, denoted $supp\,(\varphi)$, is defined as the closure, in $\mathbb{R}^n$, of the set*

$$\{x \in \mathbb{R}^n : \varphi(x) \neq 0\}.$$

That is, $x \notin supp\,(\varphi)$ exactly when φ is zero on an open neighborhood of x.

Example 12. 1. The support of the sine function is $\mathbb{R}$, since the sine function is zero only on isolated points.

2. The support of the characteristic function of $\mathbb{Q}$, the set of rational numbers, is also $\mathbb{R}$.

3. The support of the function

$$H_1(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases},$$

is the set $\{x \in \mathbb{R} : x \geq 0\}$. This function is the one-dimensional Heaviside function, named after the electrical engineer and theoretical physicist Oliver Heaviside (1850-1925).

4. Given functions $f, g : \mathbb{R}^n \to \mathbb{C}$,

$$supp\,(fg) \subseteq supp\,(f) \bigcap supp\,(g).$$

In fact, let us assume that x_0 does not belong to $supp\,(f)$. Then, there is an open neighborhood U of x_0 where f is identically zero. Therefore, $f(x)\,g(x) = 0$ for all $x \in U$, which means, by definition, that $x_0 \notin supp\,(fg)$.

Let us observe that the inclusion can be strict. For instance if $\chi_{\mathbb{Q}}, \chi_{\mathbb{I}}$ denote, respectively, the characteristic function of the rational numbers and the characteristic function of the irrational numbers, $supp\,(\chi_{\mathbb{Q}}\chi_{\mathbb{I}}) = \varnothing$, while

$$supp\,(\chi_{\mathbb{Q}}) \bigcap supp\,(\chi_{\mathbb{I}}) = \mathbb{R}^n \bigcap \mathbb{R}^n = \mathbb{R}^n.$$

Definition 13. *A function $\varphi : \mathbb{R}^n \to \mathbb{C}$ is called smooth if it is continuous and it has continuous partial derivatives of all orders.*

Definition 14. *Given a compact set $K \subseteq \mathbb{R}^n$, we denote*

$$\mathcal{D}_K = \{\varphi : \mathbb{R}^n \to \mathbb{C} : \varphi \text{ is smooth and } supp\,(\varphi) \subseteq K\}.$$

We also denote

$$\mathcal{D} = \bigcup \{\mathcal{D}_K : K \subseteq \mathbb{R}^n \text{ compact}\}.$$

Remark 15. It follows from the definitions that $\mathcal{D}_K$ and $\mathcal{D}$ are complex linear spaces.

The functions in $\mathcal{D}$ are called test functions. It should be clear that $\mathcal{D} \subseteq L^p$ for $1 \leq p \leq \infty$.

Example 16. The function $\rho : \mathbb{R}^n \to \mathbb{R}$ defined as

$$\rho(x) = \begin{cases} e^{-\frac{1}{1-|x|^2}} & \text{if } |x| < 1 \\ 0 & \text{if } |x| \geq 1 \end{cases} \tag{3.1}$$

belongs to $\mathcal{D}$ and $supp(\rho) = \{x \in \mathbb{R}^n : |x| \leq 1\}$.

The function ρ is clearly continuous when $|x| < 1$ and when $|x| > 1$. If we fix $x_0 \in \mathbb{R}^n$ with $|x_0| = 1$ and $|x| < 1$, there is

$$\lim_{x \to x_0} e^{-\frac{1}{1-|x|^2}} = 0.$$

Thus, ρ is continuous everywhere. It should be clear that ρ has continuous derivatives of all orders when $|x| < 1$ and when $|x| > 1$. It only remains to show that every derivative exists and is continuous at x_0 for $|x_0| = 1$.

We claim that, for $x \in \mathbb{R}^n$ with $|x| < 1$ and for each $\alpha \in \mathbb{N}^n$, $|\alpha| \geq 1$,

$$(\partial^\alpha \rho)(x) = e^{\left(|x|^2-1\right)^{-1}} \left(|x|^2 - 1\right)^{-|\alpha|} P\left(x; \left(|x|^2 - 1\right)^{-1}\right), \tag{3.2}$$

where P is a polynomial function, of degree $|\alpha|$, in the variable $\left(|x|^2 - 1\right)^{-1}$, with coefficients that are polynomial functions in the variable x, of degree $\leq |\alpha|$. Let us observe that the notation ∂^α assumes a particular order in which the partial derivatives are taken. In all fairness, (3.2) should be stated in terms of a completely arbitrary derivative, $\partial_{x_{r_1},\dots,x_{r_k}}$, for $1 \leq r_1, \dots, r_k, \leq n$. Nevertheless, to keep the formulas as neat as possible, we will use the notation ∂^α. So, let us prove (3.2) by induction on $|\alpha|$. If $|\alpha| = 1$,

$$\left(\partial_{x_j} \rho\right)(x) = e^{\left(|x|^2-1\right)^{-1}} \left(|x|^2 - 1\right)^{-2} 2x_j = e^{\left(|x|^2-1\right)^{-1}} \left(|x|^2 - 1\right)^{-1} (-2)\, x_j \left(|x|^2 - 1\right)^{-1}.$$

Thus, (3.2) holds for $|\alpha| = 1$. If we assume (3.2) to be true for $|\alpha| = m$ for some

$m \geq 1$,

$$\begin{aligned}
\left(\partial_{x_j}\left(\partial^\alpha \rho\right)\right)(x) &= e^{\left(|x|^2-1\right)^{-1}}(-1)\left(|x|^2-1\right)^{-2} 2x_j \left(\left(|x|^2-1\right)^{-1}\right)^{|\alpha|} P\left(x;\left(|x|^2-1\right)^{-1}\right) \\
&+e^{\left(|x|^2-1\right)^{-1}} |\alpha| \left(\left(|x|^2-1\right)\right)^{-|\alpha|-1} 2x_j P\left(x;\left(|x|^2-1\right)^{-1}\right) \\
&+e^{\left(|x|^2-1\right)^{-1}} \left(|x|^2-1\right)^{-|\alpha|} \\
&\times\left[\left(\partial_{x_j} P\right)\left(x;\left(|x|^2-1\right)^{-1}\right)+P'\left(x;\left(|x|^2-1\right)^{-1}\right)(-1)\left(|x|^2-1\right)^{-2} 2x_j\right] \\
&= e^{\left(|x|^2-1\right)^{-1}}\left(|x|^2-1\right)^{-|\alpha|-1} \\
&\times\left[\begin{array}{c}
-2x_j\left(|x|^2-1\right)^{-1} P\left(x;\left(|x|^2-1\right)^{-1}\right) \\
+2|\alpha| x_j P\left(x;\left(|x|^2-1\right)^{-1}\right) \\
+\left[\left(\partial_{x_j} P\right)\left(x;\left(|x|^2-1\right)^{-1}\right)\left(|x|^2-1\right)\right. \\
\left.-2x_j\left(|x|^2-1\right)^{-1} P'\left(x;\left(|x|^2-1\right)^{-1}\right)\right]
\end{array}\right],
\end{aligned}$$

which is (3.2) for $\partial_{x_j}\partial^\alpha$.

Next, if we fix $x_0 \in \mathbb{R}^n$ with $|x_0| = 1$ and if $|x| < 1$, it should be clear that there is

$$\lim_{x \to x_0} \left(\partial^\alpha \rho\right)(x) = 0,$$

for all $\alpha \in \mathbb{N}^n$.

Finally, we show, by induction on $|\alpha|$, that $\left(\partial^\alpha \rho\right)(y)$ exists and it is equal to zero, for all $|y| = 1$ and for all $\alpha \in \mathbb{N}^n$, $|\alpha| \geq 1$. The same warning about the use of the notation ∂^α applies here.

For $y \in \mathbb{R}^n$ with $|y| = 1$ fixed and for $x_1 < 1$, there is

$$\lim_{x_1 \to y_1} \frac{\rho\left(x_1, y'\right)}{x_1 - y_1} = \lim_{x_1 \to y_1} \frac{e^{\left(x_1^2+|y'|^2-1\right)^{-1}}}{x_1 - y_1} = 0,$$

so $\left(\partial_{x_1}\rho\right)(0) = 0$. The same holds true for $\left(\partial_{x_j}\rho\right)(0)$.

With the same notation, we assume now that $\left(\partial^\alpha \rho\right)(y)$ exists and it is equal to zero, for all $|y| = 1$ and for all $\alpha \in \mathbb{N}^n$, $|\alpha| \leq m$ for some $m \geq 1$ fixed. Then, according to (3.2),

$$\left(\partial^\alpha \rho\right)\left(x_1, y'\right) = e^{\left(x_1^2+|y'|^2-1\right)^{-1}}\left(x_1^2+\left|y'\right|^2-1\right)^{-|\alpha|} P\left(x;\left(x_1^2+\left|y'\right|^2-1\right)^{-1}\right),$$

so, there is

$$\lim_{x_1 \to y_1} \frac{(\partial^\alpha \rho)(x_1, y')}{x_1 - y_1} = 0.$$

Thus, the space $\mathcal{D}$ contains non-identically zero functions, which is a surprising fact, given the constrains placed by Definition 14.

Remark 17. For each $K \subseteq \mathbb{R}^n$ compact, the space $\mathcal{D}_K$ is a Fréchet space (see, for instance, [26], p. 24, Example 27; p. 35). That is to say, $\mathcal{D}_K$ is a complete linear metric space, where the metric is derived from the countable family of norms

$$\|\varphi\|_m = \sup_{|\alpha| \leq m} \sup_{x \in \mathbb{R}^n} |\partial^\alpha \varphi(x)|,$$

using the definition

$$d(\varphi, \psi) = \sum_{m \geq 0} \frac{1}{2^m} \frac{\|\varphi - \psi\|_m}{1 + \|\varphi - \psi\|_m}.$$

Thus, a sequence $\{\varphi_j\}_{j \geq 1}$ converges to φ in $\mathcal{D}_K$ as $j \to \infty$ exactly when, for each $\alpha \in \mathbb{N}^n$, the sequence $\{\partial^\alpha \varphi_j\}_{j \geq 1}$ converges to $\partial^\alpha \varphi$ uniformly on $\mathbb{R}^n$ as $j \to \infty$.

We define in $\mathcal{D}$ a strong topology as the inductive limit topology of the spaces $\mathcal{D}_K$ for $K \subseteq \mathbb{R}^n$ compact. With this topology, the linear space $\mathcal{D}$ becomes a complete, locally convex and Hausdorff topological linear space (see, for instance, [26], Chapter III, Section 22).

Under the inductive limit topology, a sequence $\{\varphi_j\}_{j \geq 1}$ converges to φ in $\mathcal{D}$ as $j \to \infty$ exactly when there is a compact set $K \subseteq \mathbb{R}^n$ so that $\varphi_j, \varphi \in \mathcal{D}_K$ for all $j \geq 1$ and $\{\varphi_j\}_{j \geq 1}$ converges to φ in $\mathcal{D}_K$ as $j \to \infty$ (see, for instance,[26], p. 274, Example 14). Moreover, a linear functional $T : \mathcal{D} \to \mathbb{C}$ is continuous if, and only if, $T : \mathcal{D}_K \to \mathbb{C}$ is continuous for each $K \subseteq \mathbb{R}^n$ compact (see, for instance, [26], p. 268, Corollary 2).

As we pointed out in Remark 15, the space $\mathcal{D}$ is included in L^p for $1 \leq p \leq \infty$. Furthermore,

Lemma 18. *The inclusion of $\mathcal{D}$ in L^p is continuous for $1 \leq p \leq \infty$.*

Proof. If $\{\varphi_j\}_{j \geq 1}$ converges to φ in $\mathcal{D}$ as $j \to \infty$, we have the estimates:

$$\|\varphi_j - \varphi\|_{L^p} \leq \left(\sup_{x \in \mathbb{R}^n} |\varphi_j(x) - \varphi(x)| \right) (meas(K))^{1/p},$$

for some $K \subseteq \mathbb{R}^n$ compact and all $1 \leq p < \infty$, and

$$\|\varphi_j - \varphi\|_{L^\infty} = \sup_{x \in \mathbb{R}^n} |\varphi_j(x) - \varphi(x)|.$$

This completes the proof of the lemma. □

Definition 19. *We denote $\mathcal{D}'$ the topological dual of $\mathcal{D}$. That is to say, $\mathcal{D}'$ is the complex linear space consisting of all the linear and continuous functionals $T : \mathcal{D} \to \mathbb{C}$. Such functionals are called distributions.*

The action of a distribution T on a test function φ is denoted $T(\varphi)$ or (T, φ). When it is important to indicate the duality, we will write $(T, \varphi)_{\mathcal{D}', \mathcal{D}}$.

All we have said in Remark 17, allows us to rephrase Definition 19 in the following equivalent way:

Definition 20. *A map $T : \mathcal{D} \to \mathbb{C}$ is called a distribution if it is linear and $(T, \varphi_j) \to (T, \varphi)$ in $\mathbb{C}$ as $j \to \infty$, whenever $\varphi_j \to \varphi$ in $\mathcal{D}$ as $j \to \infty$.*

Let us repeat that a sequence $\{\varphi_j\}_{j \geq 1}$ converges to φ in $\mathcal{D}$ as $j \to \infty$ exactly when there is a compact set $K \subseteq \mathbb{R}^n$ so that $\varphi_j, \varphi \in \mathcal{D}_K$ for all $j \geq 1$ and $\{\varphi_j\}_{j \geq 1}$ converges to φ in $\mathcal{D}_K$ as $j \to \infty$.

That is to say, in spite of the topological complexity of $\mathcal{D}$, which is a non-metrizable space ([25], p. 65; [26], Chapter I, Section 3), it behaves in many ways as it were a metric space, where continuity can be checked using sequences. This is a very convenient characteristic of the space $\mathcal{D}$, which we will use quite often. For an exposition of the topological structure of $\mathcal{D}$, which fills in the details left out in Remark 17, see ([25], Chapter III, Section 1).

Remark 21. Given $K_1 \subseteq K_2$ compact subsets of $\mathbb{R}^n$, it should be clear that the identity map from $\mathcal{D}_{K_1}$ into $\mathcal{D}_{K_2}$ is continuous. Furthermore, $\mathcal{D}_{K_1}$ is a closed linear subspace of $\mathcal{D}_{K_2}$.

We now discuss the strong and the weak topologies in $\mathcal{D}'$, as well as the weak topology in $\mathcal{D}$.

Remark 22. In the inductive limit topology of $\mathcal{D}$, a subset $\mathcal{B}$ is bounded when it satisfies the following conditions ([25], Chapter III, Section 2):

1. $\mathcal{B} \subseteq \mathcal{D}_K$ for some $K \subseteq \mathbb{R}^n$ compact.

2. For each $\alpha \in \mathbb{N}^n$,
$$\sup_{\varphi \in \mathcal{B}} \sup_{x \in \mathbb{R}^n} |\partial^\alpha \varphi(x)| < \infty.$$

Thus, $\mathcal{B}$ is a bounded subset of $\mathcal{D}$ exactly when $\mathcal{B}$ is a bounded subset of the Fréchet space $\mathcal{D}_K$, for some $K \subseteq \mathbb{R}^n$ compact.

Furthermore ([25]. p. 69, Theorem VI), a subset $\mathcal{B}$ of $\mathcal{D}$ is bounded if, and only if, for each $T \in \mathcal{D}'$,
$$\sup_{\varphi \in \mathcal{B}} |T(\varphi)| < \infty.$$

The strong topology in $\mathcal{D}'$ is defined by the uncountable family of semi-norms ([25], p. 71)

$$\|T\|_{\mathcal{D}',\mathcal{B}} = \sup_{\varphi\in\mathcal{B}} |T(\varphi)|,$$

where $\mathcal{B}$ is every bounded subset of $\mathcal{D}$. With this topology, $\mathcal{D}'$ is a complete, locally convex and non-metrizable topological linear space ([25], p. 71). Moreover, the space $\mathcal{D}$ with the inductive limit topology is the topological dual of the space $\mathcal{D}'$ with the strong topology ([25], p. 75, Theorem XIV).

By definition, a set $\mathcal{B}' \subseteq \mathcal{D}'$ is bounded in the strong topology of $\mathcal{D}'$ exactly when

$$\sup_{T\in\mathcal{B}'} \|T\|_{\mathcal{D}',\mathcal{B}} = \sup_{T\in\mathcal{B}'} \sup_{\varphi\in\mathcal{B}} |T(\varphi)| < \infty,$$

for each $\mathcal{B} \subseteq \mathcal{D}$ bounded. Theorem IX in ([25], p. 72) provides necessary conditions and sufficient conditions for $\mathcal{B}' \subseteq \mathcal{D}'$ to be bounded. For example, if

$$\sup_{T\in\mathcal{B}'} |T(\varphi)| < \infty$$

for each $\varphi \in \mathcal{D}$, then $\mathcal{B}'$ is bounded.

From the definition of the semi-norms $\|\cdot\|_{\mathcal{D}',\mathcal{B}}$, it should be clear that a sequence $\{T_j\}_{j\geq 1}$converges to zero in the strong topology of $\mathcal{D}'$ as $j\to\infty$ if, and only if, $T_j(\varphi) \to 0$ as $j \to \infty$, uniformly on $\mathcal{B}$ for each bounded subset $\mathcal{B}$ of $\mathcal{D}$. This form of convergence is called the strong convergence.

As for the weak topology in $\mathcal{D}$, we can say that it is generated by the family

$$\left\{T^{-1}(U) : T \in \mathcal{D}', U \subseteq \mathbb{C} \text{ open}\right\}. \tag{3.3}$$

In other words, a set $\mathcal{U} \subseteq \mathcal{D}$ is open in the weak topology if, and only if, it can be written as an arbitrary union of finite intersections of sets of the form (3.3). That is to say, it is the smallest topology that makes each T continuous. By definition, every $T \in \mathcal{D}'$ is continuous in the strong topology of $\mathcal{D}$, which is finer, or has more sets, than the weak topology.

A sequence $\{\varphi_j\}_{j\geq 1}$ in $\mathcal{D}$ converges to zero weakly as $j \to \infty$ exactly when $T(\varphi_j) \to 0$ as $j \to \infty$, for each $T \in \mathcal{D}'$. Furthermore, ([25], p. 70), in the bounded subsets of $\mathcal{D}$, the strong topology of $\mathcal{D}$ coincides with the weak topology of $\mathcal{D}$. Since a bounded subset of $\mathcal{D}$ is a subset of $\mathcal{D}_K$ for some $K \subseteq \mathbb{R}^n$ compact, we conclude that convergent sequences in the weak topology of $\mathcal{D}$ are exactly the convergent sequences in the inductive limit, or strong, topology of $\mathcal{D}$. Likewise, a Cauchy sequence in the weak topology of $\mathcal{D}$ is also a Cauchy sequence in the strong topology of $\mathcal{D}$. In other words ([25], p. 70), we can say the following:

Given a sequence $\{\varphi_j\}_{j\geq 1} \subseteq \mathcal{D}$, suppose that the sequence $\{T(\varphi_j)\}_{j\geq 1}$ converges in $\mathbb{C}$, for each $T \in \mathcal{D}'$. Then there exists $\varphi \in \mathcal{D}$ such that $T(\varphi_j) \to T(\varphi)$ in $\mathbb{C}$ as $j \to \infty$ and $\varphi_j \to \varphi$ as $j \to \infty$ in the strong topology of $\mathcal{D}$.

Let us recall that a functional defined on a topological linear space $\mathcal{T}$ is bounded, by definition, if it maps bounded subsets of $\mathcal{T}$ into bounded subsets of $\mathbb{C}$. As on a normed linear space, linear functionals defined on a Fréchet space are continuous if, and only if, they are bounded. A linear functional $T : \mathcal{D} \to \mathbb{C}$ is continuous, that is to say it is a distribution, exactly when is continuous from $\mathcal{D}_K$ to $\mathbb{C}$ for each $K \subseteq \mathbb{R}^n$ compact. Thus, we conclude that a linear functional $T : \mathcal{D} \to \mathbb{C}$ is a distribution if, and only if, it is bounded.

Since the spaces $\mathcal{D}$ and $\mathcal{D}'$ with their strong topologies are reflexive, we can use the functions in $\mathcal{D}$, to construct the weak topology in $\mathcal{D}'$. This topology will be generated by the subsets of $\mathcal{D}'$ defined as

$$\left\{\varphi^{-1}(U) : \varphi \in \mathcal{D}, U \subseteq \mathbb{C} \text{ open}\right\}.$$

So, a sequence $\{T_j\}_{j\geq 1}$ converges to zero weakly in $\mathcal{D}'$ as $j \to \infty$ exactly when $T_j(\varphi) \to 0$ as $j \to \infty$, for each $\varphi \in \mathcal{D}$.

The strong topology and the weak topology of $\mathcal{D}'$ coincide in the subsets of $\mathcal{D}'$ that are bounded in the strong topology of $\mathcal{D}'$ ([25], p. 74). Moreover, we have the following result ([25], p. 74):

Given a sequence $\{T_j\}_{j\geq 1} \subseteq \mathcal{D}'$, suppose that the sequence $\{T_j(\varphi)\}_{j\geq 1}$ converges in $\mathbb{C}$, for each $\varphi \in \mathcal{D}$. Then, the linear functional $T : \mathcal{D} \to \mathbb{C}$ defined as $T(\varphi) = \lim_{j\to\infty} T_j(\varphi)$ is a distribution and $T_j \to_{j\to\infty} T$ in the strong topology of $\mathcal{D}'$.

Finally, if we view (T, φ) as a bilinear functional on $\mathcal{D}' \times \mathcal{D}$,

$$\sup_{T\in\mathcal{B}',\varphi\in\mathcal{B}} |(T, \varphi)| < \infty,$$

for each bounded subset $\mathcal{B}$ of $\mathcal{D}$ and for each bounded subset $\mathcal{B}'$ of $\mathcal{D}'$ ([25], p. 73). Furthermore, if $\{T_j\}_{j\geq 1}$ converges to zero in the strong topology of $\mathcal{D}'$ and $\{\varphi_j\}_{j\geq 1}$converges to zero in the strong topology of $\mathcal{D}$, then $(T_j, \varphi_j) \to_{j\to\infty} 0$ in $\mathbb{C}$.

Although some of these results are true for filters and nets ([25], pp. 66, 70, 71, 73, 74), in our presentation we will only use sequences.

This concludes the overview of topological matters concerning the spaces $\mathcal{D}$ and $\mathcal{D}'$. It is amply sufficient for our purposes. Still, much has been left unsaid, for which we refer to [25].

It is time now to discuss our first examples of distributions.

Example 23. 1. Let us fix p with $1 \leq p \leq \infty$ and let

$$L^p_{loc} = \{f : \mathbb{R}^n \to \mathbb{C} : \chi_K f \in L^p \text{ for each } K \subseteq \mathbb{R}^n \text{ compact}\},$$

where χ_K denotes the characteristic function of K. It should be clear that L^p_{loc} is a complex linear space.

We consider in L^p_{loc} the topology defined by the family of semi-norms

$$\|f\|_{p,K} = \|\chi_K f\|_{L^p},$$

for $K \subseteq \mathbb{R}^n$ compact.

Since it suffices to use a countable covering of $\mathbb{R}^n$ by compact subsets, L^p_{loc} is a Fréchet space. In this space, a sequence $\{f_j\}_{j\geq 1}$ converges to f as $j \to \infty$ exactly when $\{\chi_K f_j\}_{j\geq 1}$ converges to $\chi_K f$ in L^p as $j \to \infty$, for each $K \subseteq \mathbb{R}^n$ compact. Moreover, if $1 \leq p_1 < p_2 \leq \infty$, we have the continuous inclusion $L^{p_2}_{loc} \subseteq L^{p_1}_{loc}$.

Indeed, given $f \in L^{p_2}_{loc}$ and $K \subseteq \mathbb{R}^n$ compact, if $p_2 = \infty$,

$$\|\chi_K f\|_{L^{p_1}} = \left(\int_{\mathbb{R}^n} |f|^{p_1} \chi_K dx\right)^{\frac{1}{p_1}} \leq (meas\,(K))^{\frac{1}{p_1}} \|\chi_K f\|_{L^\infty},$$

while if p_2 is finite,

$$\begin{aligned} \|\chi_K f\|_{L^{p_1}} &= \left(\int_{\mathbb{R}^n} |f|^{p_1} \chi_K \chi_K dx\right)^{\frac{1}{p_1}} \underset{(1)}{\leq} \left(\int_{\mathbb{R}^n} |f|^{p_2} \chi_K dx\right)^{\frac{1}{p_2}} \left(\int_{\mathbb{R}^n} \chi_K dx\right)^{\frac{1}{p_1}\left(1-\frac{p_1}{p_2}\right)} \\ &= meas\,(K)^{\frac{1}{p_1}-\frac{1}{p_2}} \|\chi_K f\|_{L^{p_2}}. \end{aligned}$$

We have used in (1) Hölder's inequality with the conjugate exponents p_2/p_1 and $p_2/\left(p_2 - p_1\right)$.

Given $f \in L^p_{loc}$ and $\varphi \in \mathcal{D}$, Hölder's inequality tells us that $f\varphi = \left(f\chi_{supp(\varphi)}\right)\varphi$ is an integrable function. Thus, we can consider the linear map $T_f : \mathcal{D} \to \mathbb{C}$ defined as

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

We claim that T_f is a distribution. Since T_f is linear, to prove continuity, it suffices to use sequences converging to zero. Indeed, if we fix $K \subseteq \mathbb{R}^n$ compact and if $\{\varphi_j\}_{j\geq 1}$ converges to zero in $\mathcal{D}_K$ as $j \to \infty$, we can write,

$$\begin{aligned} |T_f(\varphi_j)| &\leq \int_K |f(x)|\,|\varphi_j(x)|\,dx \\ &\underset{(1)}{\leq} \|f\|_{p,K} \|\varphi_j\|_{L^q} \\ &\leq (meas\,(K))^{1/q} \|f\|_{p,K} \sup_{x\in\mathbb{R}^n} |\varphi_j(x)| \underset{j\to\infty}{\to} 0. \end{aligned}$$

We have used in (1) Hölder's inequality.

Conversely, we say that a distribution T is a function, if there exists $f \in L^p_{loc}$, for some $1 \leq p \leq \infty$, such that $T = T_f$. That is,

$$(T, \varphi) = (T_f, \varphi) = \int_{\mathbb{R}^n} f(x)\,\varphi(x)\,dx,$$

for all $\varphi \in \mathcal{D}$.

Due to the continuous inclusion $L^{p_2}_{loc} \subseteq L^{p_1}_{loc}$ for $1 \leq p_1 < p_2 \leq \infty$, we will generally assume that $f \in L^1_{loc.}$

Thus, when we say that distributions extend the concept of function, we always mean locally integrable functions.

Strictly speaking, the elements of L^p, L^p_{loc} and other spaces to be considered, are classes under the equivalence relation $f \sim g$ if f and g are equal except on a set of Lebesgue measure zero. Nevertheless, for the purpose of our exposition, we will always work with functions, that is with class representatives. In every instance in which we consider distributions of the form T_f, it will be straightforward to verify that the matter at hand remains unchanged if we pick a different representative.

2. Given $a \in \mathbb{R}^n$, we consider the linear map $\delta_a : \mathcal{D} \to \mathbb{C}$ defined as

$$\delta_a(\varphi) = \varphi(a).$$

The estimate

$$|\delta_a(\varphi)| \leq \sup_{x \in \mathbb{R}^n} |\varphi(x)|,$$

for every $\varphi \in \mathcal{D}$, tells us that δ_a is a distribution. It is named the Dirac distribution after the theoretical physicist Paul Dirac (1902-1984), co-winner with Erwin Schröedinger (1887-1961) of the 1933 Nobel prize in physics.

The Dirac distribution is not a function, that is to say, there is no $f \in L^1_{loc}$ such that $\delta_a = T_f$. To prove this claim, we assume that there is $f \in L^1_{loc}$ so that

$$\int_{\mathbb{R}^n} f(x)\varphi(x)\,dx = \varphi(a),$$

for all $\varphi \in \mathcal{D}$. In particular, let us pick the test function $j^n\rho(j(x-a))$ for $j \geq 1$, where ρ is defined as in (3.1). Then,

$$\int_{\mathbb{R}^n} f(x)\, j^n\rho(j(x-a))\,dx = j^n\rho(j(0)) = j^n,$$

or

$$\int_{\mathbb{R}^n} f(x)\,\rho(j(x-a))\,dx = 1,$$

for all $j \geq 1$.

However,

$$\begin{aligned}\left|\int_{\mathbb{R}^n} f(x)\,\rho(j(x-a))\,dx\right| &= \left|\int_{|j(x-a)|\leq 1} f(x)\,\rho(j(x-a))\,dx\right| \\ &\leq \sup_{x\in\mathbb{R}^n}(\rho(x))\int_{|(x-a)|\leq 1/j}|f(x)\,\chi_K(x)| \underset{j\to\infty}{\to} 0,\end{aligned}$$

where K indicates the compact set $\{x \in \mathbb{R}^n : |x - a| \leq 1\}$ and χ_K is the characteristic function of K.

Since we arrive at a contradiction, such an f cannot exist.

Lützen has this to say about the early days of the so-called Dirac function ([18], p. 110): "The δ-function must have had a very sad childhood since neither mathematicians nor physicists recognized it as belonging to their domain. If mathematicians used it, it was an intuitive physical notion with no mathematical reality ... On the other hand, physicists usually considered the δ-function, or the point mass, as a pure mathematical idealization which did not exist in nature."

3. If we fix $a \in \mathbb{R}^n$ and $m = 1, 2, \ldots$, the linear map $T_a^m : \mathcal{D} \to \mathbb{C}$ defined as

$$T_a^m(\varphi) = \left(\partial_{x_1}^m \varphi\right)(a) \tag{3.4}$$

is a distribution, since we have the estimate

$$|T_a^m(\varphi)| \leq \sup_{x \in \mathbb{R}^n} \left|\left(\partial_{x_1}^m \varphi\right)(x)\right|, \tag{3.5}$$

for every $\varphi \in \mathcal{D}_K$, where $K \subseteq \mathbb{R}^n$ is compact.

4. For $f \in L^1_{loc}$ and $k \in \mathbb{R}$, $k \neq 0$, we define the operator d_k as

$$d_k(f)(x) = f(kx).$$

If $\varphi \in \mathcal{D}$,

$$\begin{aligned} \left(T_{d_k(f)}, \varphi\right) &= \int_{\mathbb{R}^n} f(kx)\, \varphi(x)\, dx = \frac{1}{|k|^n} \int_{\mathbb{R}^n} f(x)\, \varphi\left(\frac{x}{k}\right) dx \\ &= \left(T_{f(x)}, \frac{\varphi\left(\frac{x}{k}\right)}{|k|^n}\right) = \left(T_f, \frac{d_{1/k}(\varphi)}{|k|^n}\right). \end{aligned}$$

Therefore, given $T \in \mathcal{D}'$, we define the linear map d_k as

$$(d_k(T), \varphi) = \left(T, \frac{d_{1/k}(\varphi)}{|k|^n}\right).$$

In particular, if $k = -1$,

$$(d_{-1}(T), \varphi) = (T, \varphi(-\cdot)).$$

When $d_{-1}(T) = -T$, we say that the distribution is odd. If $d_{-1}(T) = T$, we say that the distribution is even. In particular, if $f \in L^1_{loc}$, T_f is odd if, and only if, the function f is odd. Likewise, T_f is even if, and only if, the function f is even.

It should be clear that δ_0 is even, while T_0^1, as defined by (3.4), is odd.

Given a bounded subset $\mathcal{B}$ of $\mathcal{D}$, the set $\mathcal{B}_k = d_{1/k}(\mathcal{B})$ is bounded in $\mathcal{D}$. Therefore,

$$\|d_k(T)\|_{\mathcal{D}',\mathcal{B}} = \frac{1}{|k|^n} \sup_{\varphi\in\mathcal{B}} \left|\left(T, d_{1/k}(\varphi)\right)\right| = \frac{1}{|k|^n} \sup_{\varphi\in\mathcal{B}_k} |(T,\varphi)| = \frac{1}{|k|^n} \|T\|_{\mathcal{D}',\mathcal{B}_k}.$$

So, the linear operator d_k is continuous from $\mathcal{D}'$ into itself.

5. The series $\sum_{j\geq 1}\delta_j$ defines a distribution in $\mathcal{D}'(\mathbb{R})$. Indeed, if we fix $K\subseteq\mathbb{R}$ compact and $\varphi\in\mathcal{D}_K$, there is $j_K\geq 1$ so that $\varphi(j)=0$ for $j>j_K$. Therefore,

$$\left|\left(\sum_{j\geq 1}\delta_j,\varphi\right)\right| = \left|\sum_{j=1}^{j_K}\varphi(j)\right| \leq C_K \sup_{x\in\mathbb{R}}|\varphi(x)|.$$

Moreover, the sequence $\left\{\sum_{1\leq j\leq k}\delta_j\right\}_{k\geq 1}$ converges to $\sum_{j\geq 1}\delta_j$ in $\mathcal{D}'(\mathbb{R})$ as $k\to\infty$. In fact, if we fix a bounded subset $\mathcal{B}$ of $\mathcal{D}(\mathbb{R})$,

$$\sup_{\varphi\in\mathcal{B}}\left|\sum_{j>k}\varphi(j)\right| = 0$$

for $k>k_{\mathcal{B}}$, according to Remark 22.

6. The work done in 5) shows that the series $\sum_{j\geq 1}\delta_{1/j}$ defines a distribution in $\mathcal{D}'(0,\infty)$. The formula does not define a distribution on $\mathbb{R}$, since the pairing

$$\left(\sum_{j\geq 1}\delta_{1/j},\varphi\right)$$

is not a complex number if $\varphi(x)=1$ on a neighborhood of zero.

Lemma 24. *Given $\varphi\in\mathcal{D}_K$ for some $K\subseteq\mathbb{R}^n$ compact and given $f\in L^1_{loc}$, the convolution*

$$\int_{\mathbb{R}^n}\varphi(x-y)\,f(y)\,dy \tag{3.6}$$

defines a function $F:\mathbb{R}^n\to\mathbb{C}$ that is smooth. Moreover,

$$(\partial^\alpha F)(x) = \int_{\mathbb{R}^n}(\partial^\alpha\varphi)(x-y)\,f(y)\,dy, \tag{3.7}$$

for all $\alpha\in\mathbb{N}^n$, $|\alpha|\geq 1$.

Proof. First of all, for $x \in \mathbb{R}^n$ fixed, $supp_y(\varphi(x-y))$ is a compact subset of $\mathbb{R}^n$ which is a translation of K. Thus, the integral (3.6) exists.

Next, we verify the hypotheses of Theorem 6, in the form suggested in Remark 7.

Let us fix an open subset U of $\mathbb{R}^n$ with compact closure. For each $x \in U$ fixed, the integral in (3.6) is carried out on a compact subset of $\mathbb{R}^n$ which is a translation of K. Thus, 1) in Theorem 6 holds. If we fix $x_0 \in U$, it should be clear, once again, that $\varphi(\cdot - y) f(y)$ is continuous at x_0 for each $y \in \mathbb{R}^n$. So, 2) in Theorem 6 is true. Lastly, there is a compact subset K_1 of $\mathbb{R}^n$ so that $supp_y(\varphi(x-y)) \subseteq K_1$, for all $x \in U$. Thus,

$$|\varphi(x-y) f(y)| \leq \left(\sup_{z\in\mathbb{R}^n} |\varphi(z)|\right) |\chi_{K_1}(y) f(y)|,$$

for $x \in U$ and $y \in \mathbb{R}^n$. This shows that 3) in Theorem 6 is also true. So, the function F is continuous on U. Since U is arbitrary, we conclude that F is continuous on $\mathbb{R}^n$.

Let us now fix $\alpha \in \mathbb{N}^n$ with $|\alpha| = 1$. So, $\partial^\alpha = \partial_{x_j}$ for some $1 \leq j \leq n$. We want to verify the hypotheses of Theorem 8, in the form suggested in Remark 9.

As before, 1) holds for each $x \in U$ and all $y \in \mathbb{R}^n$. It should be clear that

$$\left(\partial_{x_j}(\varphi(\cdot - y) f(y))\right)(x) = \left(\partial_{x_j}\varphi\right)(x-y) f(y)$$

exists for each $x \in U$ and all $y \in \mathbb{R}^n$. Thus, 2) holds as well.

Finally,

$$\left|\left(\partial_{x_j}\varphi\right)(x-y) f(y)\right| \leq \left(\sup_{z\in\mathbb{R}^n} \left|\left(\partial_{z_j}\varphi\right)(z)\right|\right) |\chi_{K_1}(y) f(y)|,$$

for $x \in U$ and $y \in \mathbb{R}^n$. That is to say, 3) is also true.

So, $\left(\partial_{x_j}F\right)(x)$ exists for $x \in U$. Since U is arbitrary, we conclude that $\partial_{x_j}F$ exists on $\mathbb{R}^n$. Moreover, $\partial_{x_j}F(x)$ is given by (3.7) with $\partial^\alpha = \partial_{x_j}$.

The calculations we just performed show that $\left(\partial_{x_j}\varphi\right)(x-y) f(y)$ satisfies, for $x \in U$, the hypotheses of Theorem 6, as outlined in Remark 7. Thus, $\partial_{x_j}F$ is continuous on $\mathbb{R}^n$.

Likewise, we can verify that $\partial_{x_k}\left(\left(\partial_{x_j}\varphi\right)(x-y) f(y)\right)$ satisfies the hypotheses of Theorem 8, in the form suggested in Remark 9, for $x \in U$ $1 \leq k \leq n$. Therefore, $\partial^2_{x_k,x_j}F$ exists on $\mathbb{R}^n$ and

$$\left(\partial^2_{x_k,x_j}F\right)(x) = \int_{\mathbb{R}^n} \left(\partial^2_{x_k,x_j}\varphi\right)(x-y) f(y)\, dy.$$

Continuing in this fashion, we can prove that F has continuous derivatives of all orders and that (3.7) holds.

This completes the proof of the lemma. □

Theorem 25. *Given $K \subseteq \mathbb{R}^n$ compact and $\varepsilon > 0$, there is a function $\varphi \in \mathcal{D}$ satisfying the following properties:*

1. $0 \leq \varphi(x) \leq 1$ *for all* $x \in \mathbb{R}^n$,
2. $\varphi(x) = 1$ *for all* $x \in K$,
3. $supp(\varphi) \subseteq \varepsilon$*-neighborhood*$(K) = \{x \in \mathbb{R}^n : d(x, K) < \varepsilon\}$, *where* $d(x, K) = \inf_{y \in K} |x - y|$.

Proof. Given the test function ρ defined in Example 3.1, let

$$\rho_j(x) = \frac{j^n}{c} \rho(jx)$$

for each $j \geq 1$, where $c = \int_{\mathbb{R}^n} \rho(x)\, dx$. It should be clear that ρ_j is a test function with integral one and support in the ball $\{x \in \mathbb{R}^n : |x| \leq 1/j\}$.

If χ_{K_ε} denotes the characteristic function of the set $K_\varepsilon = \varepsilon/3$-neighborhood$(K)$, we write

$$\begin{aligned} \varphi_j(x) &= (\chi_{K_\varepsilon} * \rho_j)(x) = \int_{K_\varepsilon} \rho_j(x - y)\, dy \\ &= \int_{|y| \leq 1/j} \chi_{K_\varepsilon}(x - y)\, \rho_j(y)\, dy. \end{aligned}$$

Lemma 24 tell us that φ_j is a smooth function. Moreover,

$$0 \leq \varphi_j(x) \leq \int_{\mathbb{R}^n} \rho_j(x - y)\, dy = 1.$$

Now, given $x \in K$ and $y \in \mathbb{R}^n$,

$$d(x - y, K) \leq |y|.$$

Thus, for $j_0 > 3/\varepsilon$, $x - y \in K_\varepsilon$ when $x \in K$ and $|y| \leq 1/j_0$. So, if $x \in K$,

$$\varphi_{j_0}(x) = \int_{|y| \leq 1/j_0} \chi_{K_\varepsilon}(x - y)\, \rho_{j_0}(y)\, dy = \int_{\mathbb{R}^n} \rho_{j_0}(y)\, dy = 1.$$

As for the support of φ_{j_0}, let us recall that given $z_1, z_2 \in \mathbb{R}^n$,

$$|d(z_1, K) - d(z_2, K)| \leq |z_1 - z_2|.$$

In particular,

$$d(z_1, K) \geq d(z_2, K) - |z_1 - z_2|.$$

Now, if $z_1 = x - y$ and $z_2 = x$ for $x \notin 2\varepsilon/3$-neighborhood(K), and $|y| \leq 1/j_0$,

$$d(x - y, K) \geq d(x, K) - |y| > \frac{2\varepsilon}{3} - \frac{1}{j_0} > \frac{\varepsilon}{3}.$$

So, $\chi_{K_\varepsilon}(x - y) = 0$, and then $\varphi_{j_0}(x) = 0$.

This completes the proof of the theorem. □

Remark 26. Theorem 25 illustrates the effect of convolving with the function ρ_j. It smooths out the discontinuous function χ_K, yielding a smooth function with compact support, which is "almost" the characteristic function, of a slightly bigger set. For this reason, ρ is called a smoothing, or regularizing, function.

There are versions of Theorem 25 that do not involve compact sets. Here is an example:

Theorem 27. *There is a smooth function ψ so that $0 \leq \psi(x) \leq 1$ for all $x \in \mathbb{R}$ and*

$$\psi(x) = \begin{cases} 0 & \text{if } x \leq 2 \\ 1 & \text{if } x \geq 3 \end{cases} .$$

Proof. Given the characteristic function χ of the half-line $(2+\varepsilon, \infty)$ for some $0 < \varepsilon < 1$, we define

$$\begin{aligned} \psi_j(x) &= (\chi * \rho_j)(x) = \int_{\mathbb{R}} \chi(y) \rho_j(x-y) \, dy \\ &= \int_{\mathbb{R}} \chi(x-y) \rho_j(y) \, dy, \end{aligned}$$

where ρ_j is the one-dimensional version of the smoothing function used in Theorem 25.

According to Lemma 24, the function ψ_j is smooth for $j \geq 1$. Moreover,

$$0 \leq \psi_j(x) \leq \int_{\mathbb{R}} \rho_j(x-y) \, dy = 1,$$

for $x \in \mathbb{R}$.

If $x \geq 3$ and $-1/j \leq y \leq 1/j$,

$$2 \leq 3 - 1/j \leq x - y$$

so, $\chi(x-y) = 1$. Therefore,

$$\psi_j(x) = \int_{|y| \leq 1/j} \rho_j(y) \, dy = 1.$$

If $x \leq 2$ and $-1/j \leq y \leq 1/j$,

$$x - y \leq x + 1/j \leq 2 + \varepsilon/2$$

for some $j = j_0$. So, $\chi(x-y) = 0$ and then, $\psi_{j_0}(x) = 0$.

This completes the proof of the theorem. □

Theorem 28. *The space $\mathcal{D}$ is densely included in L^p for $1 \leq p < \infty$.*

Proof. Let us recall that continuous functions with compact support are dense in L^p (see, for instance, [21], p. 69, Theorem 3.14). Thus, it suffices to approximate continuous function with compact support, by functions in $\mathcal{D}$.

If f is such a function, let

$$f_j(x) = (f * \rho_j)(x) = \int_{\mathbb{R}^n} f(y)\, \rho_j(x-y)\, dy,$$

where ρ_j is the function used in the proof of Theorem 25. According to Lemma 24, f_j is a smooth function. Moreover, not only f_j has compact support, but $supp\,(f_j)$ is contained in a fix compact set, independently of j.

Let K be the support of f. We claim that $supp\,(f_j) \subseteq$ 2-neighborhood(K). In fact, if $x \notin$ 2-neighborhood(K) and $y \in K$,

$$2 < d(x, K) \leq |x-y|,$$

so $\rho_j(x-y) = 0$.

Thus, $f_j \in \mathcal{D}$ for all $j \geq 1$. It remains to show that $f_j \to f$ in L^p as $j \to \infty$.

First of all, the function f is uniformly continuous, because it is continuous and it has compact support. Thus, given $\varepsilon > 0$, there is $j_0 = j_0(\varepsilon) \geq 1$ so that

$$|f(x-y) - f(x)| \leq \varepsilon,$$

when $|y| < 1/j$ for $j \geq j_0$, for all $x \in \mathbb{R}^n$. So,

$$\begin{aligned} |f_j(x) - f(x)| &= \left| \int_{\mathbb{R}^n} f(x-y)\, \rho_j(y)\, dy - f(x) \int_{\mathbb{R}^n} \rho_j(y)\, dy \right| \\ &\leq \int_{\mathbb{R}^n} |f(x-y) - f(x)|\, \rho_j(y)\, dy \leq \varepsilon, \end{aligned}$$

for all $x \in \mathbb{R}^n$. Thus, $\|f_j - f\|_{L^p} \leq \varepsilon$ for $j \geq j_0$.

$$\|f_j - f\|_{L^p}.$$

This completes the proof of the theorem. □

Remark 29. 1. As a consequence of Theorem 28, given $f \in L^p$ for $1 \leq p < \infty$, the sequence $\{f * \rho_j\}_{j \geq 1}$ converges to f in L^p as $j \to \infty$.

First of all, the function $f * \rho_j$ is smooth, according to Lemma 24, although will not, in general, have compact support. Furthermore, Young's convolution theorem (see, for instance, [29], p. 146, Theorem 9.2; p. 145, Theorem 9.1) implies that

$$\|f * \rho_j\|_{L^p} \leq \|f\|_{L^p} \|\rho_j\|_{L^1} = \|f\|_{L^p}.$$

Thus, $f * \rho_j \in L^p$ for all $j \geq 1$.

Next, let $f_k = \chi_k f$, where χ_k is the characteristic function of $\{x \in \mathbb{R}^n : |x| \leq k\}$ for $k \geq 1$. Then,

$$\begin{aligned}
\|f - f * \rho_j\|_{L^p} &\leq \|f - f_k\|_{L^p} + \|f_k - f_k * \rho_j\|_{L^p} + \|(f_k - f) * \rho_j\|_{L^p} \\
&\leq 2\underbrace{\|f - f_k\|_{L^p}}_{(1)} + \underbrace{\|f_k - f_k * \rho_j\|_{L^p}}_{(2)}.
\end{aligned} \tag{3.8}$$

$$(1) = \int_{|x|\geq k} |f(x)|^p \, dx \underset{j\to\infty}{\to} 0,$$

according to Lebesgue's dominated convergence theorem. Thus, given $\varepsilon > 0$ there is $k_0 = k_0(\varepsilon) \geq 1$ so that $\|f - f_{k_0}\|_{L^p} \leq \varepsilon$. We fix $k = k_0$ throughout (3.8). As we said in the proof of Theorem 28, there is a continuous function g with compact support such that $\|f_{k_0} - g\|_{L^p} \leq \varepsilon$. Therefore,

$$\begin{aligned}
(2) &\leq \|f_{k_0} - g\|_{L^p} + \|g - g * \rho_j\|_{L^p} + \|(g - f_{k_0}) * \rho_j\|_{L^p} \\
&\leq 2\,\|f_{k_0} - g\|_{L^p} + \|g - g * \rho_j\|_{L^p} \leq 2\varepsilon + \|g - g * \rho_j\|_{L^p}.
\end{aligned}$$

Finally, it was shown in the proof of Theorem 28 that $\|g - g * \rho_j\|_{L^p} \to 0$ as $j \to \infty$.

2. With the notation used in 1), given $\varepsilon > 0$ there are $k_0 = k_0(\varepsilon) \geq 1$ and $j_0 = j_0(\varepsilon) \geq 1$ so that
$$\|f_{k_0} * \rho_{j_0} - f\|_{L^p} \leq \varepsilon.$$
Indeed,
$$\|f_k * \rho_j - f\|_{L^p} \leq \|f_k * \rho_j - f_k\|_{L^p} + \|f_k - f\|_{L^p}$$
That is to say, we can approximate functions in L^p with test functions, just by using a truncation followed by smoothing.

Theorem 30. *For each $1 \leq p \leq \infty$, the map*

$$\begin{array}{ccc} L^p_{loc} & \to & \mathcal{D}' \\ f & \to & T_f \end{array}$$

is linear, one-to-one, and continuous in the sense that $f_j \to f$ in L^p_{loc} as $j \to \infty$, implies strong convergence in $\mathcal{D}'$.

Proof. According to the inclusion result proved in 1) of Example 23, we can assume that $p = 1$.

It should be clear that the map is linear from L^1_{loc} into $\mathcal{D}'$. So, to prove that it is one-to-one, given $f \in L^1_{loc}$, we need to show that $\int_{\mathbb{R}^n} f\varphi dx = 0$ for all $\varphi \in \mathcal{D}$ implies that f is zero a.e., that is, $f(x) = 0$ except for x in a set of Lebesgue measure zero.

We consider

$$(f\varphi * \rho_j)(x) = \int_{\mathbb{R}^n} f(y)\,\varphi(y)\,\rho_j(x-y)\,dy,$$

where ρ_j is the function used in the proof of Theorem 25. For each $x \in \mathbb{R}^n$, the function $\varphi(\cdot)\,\rho_j(x-\cdot)$ belongs to $\mathcal{D}$. Therefore, $(f\varphi * \rho_j)(x) = 0$. Furthermore, as we have shown in 1) of Remark 29, $f\varphi * \rho_j \to f\varphi$ in L^1 as $j \to \infty$. Thus, $f\varphi$ is zero a.e.. Let us write

$$\mathbb{R}^n = \bigcup_{k\geq 1} B_k,$$

where $B_k = \{x \in \mathbb{R}^n : |x| \leq k\}$. Theorem 25 tells us that there is $\varphi_k \in \mathcal{D}$ so that $\varphi_k(x) = 1$ for $x \in B_k$. So, there is a null set $N_k \subseteq B_k$ such that $f(x) = 0$ for $x \in B_k \backslash N_k$. Finally, $f(x) = 0$ except for x in the null set $N = \bigcup_{k\geq 1} N_k$.

Let $\{f_j\}_{j\geq 1}$ be a sequence converging to zero in L^1_{loc} as $j \to \infty$ and let $\mathcal{B}$ be a bounded subset of $\mathcal{D}$. That is, $\mathcal{B}$ is a bounded subset of $\mathcal{D}_K$ for some $K \subseteq \mathbb{R}^n$ compact. Then,

$$\sup_{\varphi\in\mathcal{B}} \left| \int_{\mathbb{R}^n} f_j \varphi dx \right| \leq \|f_j\|_{1,K} \sup_{x\in\mathbb{R}^n} |\varphi(x)| \to_{j\to\infty} 0,$$

for every $\varphi \in \mathcal{B}$.

This completes the proof of the theorem. $\square$

Remark 31. Theorem 30 gives us an opportunity to test some of the properties stated in Remark 17 and Remark 22.

If $\{f_j\}_{j\geq 1}$ converges to zero in L^1_{loc} as $j \to \infty$, the family $\{T_{f_j}\}_{j\geq 1}$ is a bounded subset in the strong topology of $\mathcal{D}'$. Indeed, if we fix a bounded subset $\mathcal{B}$ in the strong topology of $\mathcal{D}$, there is $K \subseteq \mathbb{R}^n$ so that $\mathcal{B} \subseteq \mathcal{D}_K$. Thus,

$$\left|T_{f_j}(\varphi)\right| \leq \sup_{j\geq 1} \|f_j\|_{1,K} \sup_{\varphi\in\mathcal{B}} \sup_{x\in\mathbb{R}^n} |\varphi(x)| < \infty.$$

So,

$$\sup_{j\geq 1} \left\|T_{f_j}\right\|_{\mathcal{D}',\mathcal{B}} = \sup_{j\geq 1} \sup_{\varphi\in\mathcal{B}} \left|T_{f_j}(\varphi)\right| \leq \sup_{j\geq 1} \|f_j\|_{1,K} \sup_{\varphi\in\mathcal{B}} \sup_{x\in\mathbb{R}^n} |\varphi(x)| < \infty.$$

Since the strong convergence and the weak convergence, of sequences, coincide on the bounded set $\{T_{f_j}\}_{j\geq 1}$, the sequence $\{T_{f_j}\}_{j\geq 1}$ actually converges to zero in the strong topology of $\mathcal{D}'$, a fact that is fairly easy to verify directly. Indeed, as before, for each bounded subset $\mathcal{B}$ in the strong topology of $\mathcal{D}$, there is $K \subseteq \mathbb{R}^n$ so that $\mathcal{B} \subseteq \mathcal{D}_K$. Thus,

$$\left\|T_{f_j}\right\|_{\mathcal{B}} = \sup_{\varphi\in\mathcal{B}} \left|T_{f_j}(\varphi)\right| \leq \left(\sup_{\varphi\in\mathcal{B}} \sup_{x\in\mathbb{R}^n} |\varphi(x)| \right) \|f_j\|_{1,K} \to_{j\to\infty} 0.$$

At the end of Remark 22, we mentioned several properties of the pairing (T, φ) seen as a bilinear map from $\mathcal{D}' \times \mathcal{D}$ into $\mathbb{C}$. Let us illustrate them in the particular case of (T_f, φ) for $f \in L^1_{loc}$.

That B is a bounded subset of L^1_{loc} means that for each $K \subseteq \mathbb{R}^n$ compact we have

$$\sup_{f \in B} \|f\|_{1,K} < \infty.$$

Then, for each bounded subset $\mathcal{B}$ of $\mathcal{D}_K$ for each $K \subseteq \mathbb{R}^n$ compact,

$$|(T_f, \varphi)| \leq \|f\|_{1,K} \sup_{x \in \mathbb{R}^n} |\varphi(x)|,$$

for $f \in B$ and $\varphi \in \mathcal{B}$. So,

$$\sup_{f \in B, \varphi \in \mathcal{B}} |(T_f, \varphi)| < \infty.$$

If $\{f_j\}_{j \geq 1}$ converges to zero in L^1_{loc} and $\{\varphi_j\}_{j \geq 1}$converges to zero in the strong topology of $\mathcal{D}$, then

$$\left|\left(T_{f_j}, \varphi_j\right)\right| \leq \|f_j\|_{1,K} \sup_{x \in \mathbb{R}^n} |\varphi_j(x)| \to_{j \to \infty} 0, \tag{3.9}$$

where $\varphi_j \in \mathcal{D}_K$ for some $K \subseteq \mathbb{R}^n$ compact and for all $j \geq 1$.

For an arbitrary pairing (T, φ), we cannot say, generally, anything else. However, in our particular case, it should be clear from (3.9) that

1. $(T_f, \varphi_j) \to 0$ in $\mathbb{C}$ as $j \to \infty$ when f belongs to a bounded subset of L^1_{loc} and $\{\varphi_j\}_{j \geq 1}$ converges to zero in the strong topology of $\mathcal{D}$.

2. $\left(T_{f_j} \varphi\right) \to 0$ in $\mathbb{C}$ as $j \to \infty$ when $\{f_j\}_{j \geq 1}$ converges to zero in L^1_{loc} and φ belongs to a bounded subset of $\mathcal{D}$.

Definition 32. *Given $T \in \mathcal{D}'$ and an open subset Ω of $\mathbb{R}^n$, the restriction of T to Ω, denoted $T\,|\Omega$, is defined as*

$$(T\,|\Omega)(\varphi) = T(\varphi),$$

for all $\varphi \in \mathcal{D}$ with $supp(\varphi) \subseteq \Omega$.

Remark 33. The word restriction in Definition 32, should not be taken literally, since $\mathbb{R}^n$ is not the domain of T. It really means that the distribution T is being restricted to the linear subspace of $\mathcal{D}$ consisting of those functions φ with support contained in Ω. Later on, we will explicitly restrict a distribution to a linear subspace of its domain, thus using the word restriction with its proper meaning.

If $\mathcal{D}(\Omega)$ is the linear space of those functions $\varphi \in \mathcal{D}$ with $supp(\varphi) \subseteq \Omega$, all we have said in Remarks 17 and 22 applies to $\mathcal{D}(\Omega)$. Thus, $T\,|\Omega$ is a linear and continuous functional on $\mathcal{D}(\Omega)$, or in other words, belongs to $\mathcal{D}'(\Omega)$, the topological dual of $\mathcal{D}(\Omega)$.

For instance, the function $f(x) = 1/x$ does not define a distribution of the form T_f on $\mathbb{R}$, because f is not locally integrable on $\mathbb{R}$. However, T_f belongs to $\mathcal{D}'(\mathbb{R}\setminus\{0\})$ since $1/x$ is a locally integrable function on $\mathbb{R}\setminus\{0\}$.

Definition 34. *A distribution $T \in \mathcal{D}'$ is zero on an open subset Ω of $\mathbb{R}^n$, if $T|\Omega = 0$, that is, if $T(\varphi) = 0$ for every $\varphi \in \mathcal{D}(\Omega)$.*

For instance, the Dirac distribution δ_a is zero on $\mathbb{R}^n\setminus\{a\}$.

Definition 35. *Given $T_1, T_2 \in \mathcal{D}'$ we say that they coincide on an open subset Ω of $\mathbb{R}^n$ if $(T_1 - T_2)|\Omega = 0$.*

We consider the following question: Suppose that the distribution $T \in \mathcal{D}'$ is zero on an open subset Ω_i of $\mathbb{R}^n$ for $i \in I$, where I is an arbitrary set of indexes. If $\Omega = \bigcup_{i\in I}\Omega_i$, is it true that T is zero on Ω?. We will answer this question on the affirmative, but to do that, we will need to "cut up" a test function into pieces that are still test functions, but with specific properties. The following theorem is the tool we will use.

Theorem 36. *Given $K \subseteq \mathbb{R}^n$ compact and given a finite covering $\{\Omega_j\}_{1\le j\le k}$ of K by open sets, there are functions $\{\alpha_j\}_{1\le j\le k}$ so that*

1. *$\alpha_j \in \mathcal{D}(\Omega_j)$ for each $1 \le j \le k$,*
2. *$0 \le \alpha_j(x) \le 1$ for all $x \in \mathbb{R}^n$ and for each $1 \le j \le k$,*
3. *$\sum_{1\le j\le k}\alpha_j = 1$ for each x in a certain open neighborhood of K.*

Proof. It should be clear that there is $\varepsilon > 0$ and there are compact sets K_j so that the 2ε-neighborhood$(K_j) \subseteq \Omega_j$ for $1 \le j \le k$ and the ε-neighborhood$(K) \subseteq \bigcup\limits_{1\le j\le k} K_j$. For each $1 \le j \le k$, we select a function φ_j satisfying the conditions stated in Theorem 25, with respect to the compact set that is the closure of the ε-neighborhood(K_j). Then, we define

$$\begin{aligned}
\alpha_1 &= \varphi_1,\\
\alpha_2 &= \varphi_2(1-\varphi_1),\\
\alpha_3 &= \varphi_3(1-\varphi_1)(1-\varphi_2), ...,\\
\alpha_k &= \varphi_k(1-\varphi_1)(1-\varphi_2)...(1-\varphi_{k-1}).
\end{aligned}$$

Thus, the function α_j satisfies 1) and 2). As for 3), it can be proved by induction on k that

$$\sum\nolimits_{1\le j\le k}\alpha_j = 1 - (1-\varphi_1)...(1-\varphi_k).$$

So, if $x \in \varepsilon$-neighborhood(K), it should be clear that $\sum_{1\le j\le k}\alpha_j(x) = 1$.
This completes the proof of the theorem. □

Definition 37. *Given $K \subseteq \mathbb{R}^n$ compact and given a finite covering $\{\Omega_j\}_{1\leq j\leq k}$ of K by open sets, a family $\{\alpha_j\}_{1\leq j\leq k}$ of functions satisfying the conditions in Theorem 36 is called a partition of unity associated with the covering $\{\Omega_j\}_{1\leq j\leq k}$.*

Theorem 38. *Let us assume that the distribution $T \in \mathcal{D}'$ is zero on an open subset Ω_i of $\mathbb{R}^n$ for $i \in I$, where I is an arbitrary set of indexes. If $\Omega = \bigcup_{i\in I} \Omega_i$, then, T is zero on Ω.*

Proof. Given $\varphi \in \mathcal{D}(\Omega)$, let $K = supp(\varphi)$. Since $\{\Omega_i\}_{i\in I}$ is a covering of the compact set K, we can select a finite covering, $\left\{\Omega_{i_j}\right\}_{1\leq j\leq k}$. According to Theorem 36, there is a partition of unity $\{\alpha_j\}_{1\leq j\leq k}$ associated with the covering $\left\{\Omega_{i_j}\right\}_{1\leq j\leq k}$. Then,

$$\varphi = \sum_{j=1}^{k} \alpha_j \varphi,$$

where $\alpha_j\varphi \in \mathcal{D}\left(\Omega_{i_j}\right)$. Since, by definition, $\mathcal{D}\left(\Omega_{i_j}\right) \subseteq \mathcal{D}(\Omega) \subseteq \mathcal{D}$ for all $1 \leq j \leq k$, we can write

$$T(\varphi) = \sum_{j=1}^{k} T(\alpha_j\varphi),$$

which is zero by hypothesis. Thus, $T|\Omega = 0$.

This completes the proof of the theorem. □

Definition 39. *The support of a distribution $T \in \mathcal{D}'$, denoted $supp(T)$, is defined as the complement of the largest open subset of $\mathbb{R}^n$ where T is zero.*

Theorem 38 allows us to say that

$$supp(T) = \mathbb{R}^n \setminus \bigcup \{\Omega \subseteq \mathbb{R}^n \text{ open} : T|\Omega = 0\}.$$

Example 40. 1. The support of the Dirac distribution δ_a is $\{a\}$. The support of the distribution T_a^m defined in (3.4) is also $\{a\}$.

When the support of a distribution is a single point $\{a\}$, it is customary to say that the distribution is concentrated on $\{a\}$.

2. If T_f is the distribution defined by a locally integrable function f,

$$supp(T_f) \subseteq supp(f).$$

Indeed, if $x_0 \notin supp(f)$, there is an open set $U \subseteq \mathbb{R}^n$ so that $f(x) = 0$ for every $x \in U$. Then, $T_f(\varphi) = 0$ for every $\varphi \in \mathcal{D}(U)$. Thus, $x_0 \notin supp(T_f)$.

The inclusion $supp(T_f) \subseteq supp(f)$ can be strict. Indeed, if $f = \chi_{\mathbb{Q}}$, the characteristic function of the rationals, given $\varphi \in \mathcal{D}(\mathbb{R})$, $\int_{\mathbb{R}} \chi_{\mathbb{Q}}\varphi dx = 0$. Thus, $supp\left(T_{\chi_{\mathbb{Q}}}\right) = \varnothing$. However, as we saw in Example 11, $supp(\chi_{\mathbb{Q}}) = \mathbb{R}$.

Let us observe that the distribution T_f defined by a locally integrable function f, remains unchanged when f is replaced by a function g equal to f except on a null set. Thus, $supp\,(T_f)$ does not change either. Of course, $supp\,(g)$ may differ from $supp\,(f)$.

3. If $f : \mathbb{R}^n \to \mathbb{C}$ is continuous,

$$supp\,(T_f) = supp\,(f)\,.$$

We only need to verify that $supp\,(f) \subseteq supp\,(T_f)$. If $x_0 \notin supp\,(T_f)$, there is an open ball B_ε centered at x_0 so that

$$\int_{\mathbb{R}^n} f\,(y)\,\varphi\,(y)\,dy = 0$$

for all $\varphi \in \mathcal{D}\,(B_\varepsilon)$. If we assume that $x_0 \in supp\,(f)$, there is $x_1 \in B_\varepsilon$ so that $f\,(x_1) \neq 0$. Since f is continuous at x_1, there is $\varepsilon' < \varepsilon$ such that $f\,(x) \neq 0$ for $x \in B_{\varepsilon'}$, an open ball contained in B_ε and centered at x_1. The function $\varphi\,(\cdot)\,\rho_j\,(x - \cdot)$ belongs to $\mathcal{D}\,(B_\varepsilon)$ for $x \in \mathbb{R}^n$. Thus,

$$\int_{\mathbb{R}^n} f\,(y)\,\varphi\,(y)\,\rho_j\,(x-y)\,dy = 0.$$

Furthermore, as we have shown in 1) of Remark 29, $f\varphi * \rho_j \to f\varphi$ in L^1 as $j \to \infty$. Thus, $f\varphi$ is zero a.e. on $\mathbb{R}^n$. In particular, if we pick a test function φ that is equal to one on $B_{\varepsilon'}$, we conclude that f is zero a.e. on B_ε, which is not possible. Therefore, $x_0 \notin supp\,(f)$.

Lemma 41. *Let $T \in \mathcal{D}'$ and $\varphi \in \mathcal{D}$. If $supp\,(T) \bigcap supp\,(\varphi) = \varnothing$, then $T\,(\varphi) = 0$.*

Proof. That $supp\,(T) \bigcap supp\,(\varphi) = \varnothing$ implies that $supp\,(\varphi) \subseteq \mathbb{R}^n \backslash supp\,(T)$. Since T is zero on the open set $U = \mathbb{R}^n \backslash supp\,(T)$ and $\varphi \in \mathcal{D}\,(U)$, we can conclude, by definition, that $T\,(\varphi) = 0$.

This completes the proof of the lemma. □

Remark 42. 1. There is a more general version of Theorem 36 (see [25], p. 22, Theorem II), which would give us the following extension result ([25], p. 27, Theorem IV):

For each $i \in I$, an arbitrary set of indexes, we fix a distribution $T_i \in \mathcal{D}'\,(\Omega_i)$, where Ω_i is an open subset of $\mathbb{R}^n$. Furthermore, we stipulate the following compatibility condition: If $\Omega_i \bigcap \Omega_j$ is not empty, $T_i\,|\Omega_i \bigcap \Omega_j = T_j\,|\Omega_i \bigcap \Omega_j$. Then, if $\Omega = \bigcup_{i \in I} \Omega_i$, there is a distribution, in fact unique, $T \in \mathcal{D}'\,(\Omega)$, so that $T\,|\Omega_i = T_i$.

Roughly speaking, we can "stitch" together an arbitrary bunch of compatible distributions, by means of a certain partition of unity, obtaining a distribution that is a unique extension of all the pieces.

What we said in Remark 33 about the use of the word restriction applies to the word extension as well.

2. Given $\Omega_1 \subseteq \Omega_2$ open subsets of $\mathbb{R}^n$ and given a distribution T_1 in $\mathcal{D}'(\Omega_1)$, the Hahn-Banach theorem for locally convex and Hausdorff topological linear spaces (see, for instance, [16], p. 180, Proposition 1) implies that there is a distribution T_2 in $\mathcal{D}'(\Omega_2)$ such that $T_2\,|\Omega_1 = T_1$.

Example 43. Sometimes, there is an explicit *ad hoc* method to extend a particular distribution. Let us look at three examples.

1. The first one involves the distribution T_f in $\mathcal{D}'(\mathbb{R}\backslash\{0\})$ for $f(x) = 1/x$. Let us observe that the function $1/x$ is measurable, but not locally integrable, on $\mathbb{R}$. Given $\varphi \in \mathcal{D}(\mathbb{R})$ and given $j \geq 1$, let us consider

$$T_j(\varphi) = \int_{|x|>1/j} \frac{\varphi(x)}{x} dx. \tag{3.10}$$

It should be clear that $T_j \in \mathcal{D}'(\mathbb{R})$ for each $j \geq 1$. We claim that (3.10) has a limit in $\mathbb{C}$ as $j \to \infty$. In fact, we can assume that $supp(\varphi) \subseteq B_r = \{x{\in}\mathbb{R} : |x| \leq r\}$ for some $r \geq 1$. Then, if $1 \leq j' < j$,

$$\begin{aligned}&\left|\int_{r>|x|>1/j'} \frac{\varphi(x)}{x} dx - \int_{r>|x|>1/j} \frac{\varphi(x)}{x} dx\right| \\ &= \left|\int_{1/j'\geq|x|>1/j} \frac{\varphi(x)}{x} dx\right| \leq \left|\int_{1/j'\geq|x|>1/j} \frac{\varphi(x)-\varphi(0)}{x} dx\right| \\ &+ |\varphi(0)| \left|\int_{1/j'\geq|x|>1/j} \frac{dx}{x}\right| = (1) + (2).\end{aligned}$$

Since the function $1/x$ is odd, the integral in (2) is zero. As for (1),

$$\left|\int_{1/j'\geq|x|>1/j} \frac{\varphi(x)-\varphi(0)}{x} dx\right| \leq 2\sup_{x\in\mathbb{R}} |\varphi'(x)| \left(\frac{1}{j'} - \frac{1}{j}\right) \underset{j,j'\to\infty}{\to} 0.$$

So, there is $\lim T_j(\varphi)$ in $\mathbb{C}$ as $j \to \infty$, for each $\varphi \in \mathcal{D}(\mathbb{R})$. Let us call the limit $T(\varphi)$. Since

$$|T_j(\varphi)| = \left|\int_{r>|x|>1/j} \frac{\varphi(x)}{x} dx\right| \leq 2\left(r - \frac{1}{j}\right) \sup_{x\in\mathbb{R}} |\varphi'(x)|,$$

we conclude that

$$|T(\varphi)| \leq 2r \sup_{x\in\mathbb{R}} |\varphi'(x)|, \tag{3.11}$$

for $\varphi \in \mathcal{D}_{B_r}(\mathbb{R})$.

Therefore, according to Remark 17 and Remark 22, T belongs to $\mathcal{D}'(\mathbb{R})$. The distribution T is called the principal value of $1/x$ and it is denoted $pv\,1/x$. That is,

$$\left(pv\frac{1}{x},\varphi\right)=\lim_{j\to\infty}\int_{|x|>1/j}\frac{\varphi(x)}{x}dx,$$

for $\varphi \in \mathcal{D}$.

It should be clear that $pv\,1/x$ extends T_f to $\mathbb{R}$, when $f(x)=1/x$.

2. In the second example, we consider the distribution T_f in $\mathcal{D}'(\mathbb{R}\setminus\{0\})$ for $f(x)=1/x^2$. Once again, the function is measurable, but not locally integrable, on $\mathbb{R}$.

   We begin in the same way as in the first example. Given $\varphi\in\mathcal{D}(\mathbb{R})$ and given $j\geq 1$, let us write

$$T_j(\varphi)=\int_{|x|>1/j}\frac{\varphi(x)}{x^2}dx. \tag{3.12}$$

   As before, $T_j\in\mathcal{D}'(\mathbb{R})$ for each $j\geq 1$. If $\mathrm{supp}(\varphi)\subseteq B_r=\{x\in\mathbb{R}:|x|\leq r\}$ for some $r\geq 1$ and if $1\leq j'<j$,

$$\begin{aligned}
&\left|\int_{r>|x|>1/j'}\frac{\varphi(x)}{x^2}dx-\int_{r>|x|>1/j}\frac{\varphi(x)}{x^2}dx\right|\\
=&\left|\int_{1/j'\geq|x|>1/j}\frac{\varphi(x)}{x^2}dx\right|\leq\left|\int_{1/j'\geq|x|>1/j}\frac{\varphi(x)-\varphi(0)-x\varphi'(0)}{x^2}dx\right|\\
&+|\varphi(0)|\left|\int_{1/j'\geq|x|>1/j}\frac{dx}{x^2}\right|+|\varphi'(0)|\left|\int_{1/j'\geq|x|>1/j}\frac{dx}{x}\right|\\
=&\ (1)+(2)+(3).
\end{aligned}$$

   As in the first example, the integral in (3) is zero. Since $|\varphi(x)-\varphi(0)-x\varphi'(0)|\leq\frac{x^2}{2}\sup_{x\in\mathbb{R}}|\varphi''(x)|$, we can write

$$(1)\leq\sup_{x\in\mathbb{R}}|\varphi''(x)|\left(\frac{1}{j'}-\frac{1}{j}\right)\underset{j,j'\to\infty}{\to}0.$$

   Therefore, there is

$$\lim_{j\to\infty}\int_{|x|>1/j}\frac{\varphi(x)-\varphi(0)-x\varphi'(0)}{x^2}dx \tag{3.13}$$

for each $\varphi \in \mathcal{D}_{B_r}(\mathbb{R})$. As for (2),

$$\left| \int_{1/j' \geq |x| > 1/j} \frac{dx}{x^2} \right| = 4\left(j - j'\right).$$

This divergence is unavoidable, so we simply put it aside.

From the estimate

$$\left| \int_{r \geq |x| > 1/j} \frac{\varphi(x) - \varphi(0) - x\varphi'(0)}{x^2} dx \right| \leq 2\left(r - \frac{1}{j}\right) \frac{1}{2} \sup_{x \in \mathbb{R}} \left|\varphi''(x)\right|$$

for $\varphi \in \mathcal{D}_{B_r}(\mathbb{R})$, we can conclude that (3.13) defines a distribution in $\mathcal{D}'(\mathbb{R})$, which is called the finite part of $1/x^2$ and it is denoted $fp\,1/x^2$. That is,

$$\left(fp\frac{1}{x^2}, \varphi \right) = \lim_{j \to \infty} \int_{|x| > 1/j} \frac{\varphi(x) - \varphi(0) - x\varphi'(0)}{x^2} dx$$

and

$$\left| \left(fp\frac{1}{x^2}, \varphi \right) \right| \leq \frac{r}{2} \sup_{x \in \mathbb{R}} \left|\varphi''(x)\right|, \tag{3.14}$$

where $\varphi \in \mathcal{D}(B_r)$.

It should be clear that $fp\,1/x^2$ extends T_f to $\mathbb{R}$ when $f(x) = 1/x^2$.

3. Finally, we consider the distribution T_f in $\mathcal{D}'(\mathbb{R} \backslash \{0\})$ for $f(x) = 1/|x|$. Once again, the function is measurable, but not locally integrable, on $\mathbb{R}$. We begin in the same way as in the first example.

   Given $\varphi \in \mathcal{D}(\mathbb{R})$ and given $j \geq 1$, let us write

$$T_j(\varphi) = \int_{|x| > 1/j} \frac{\varphi(x)}{|x|} dx.$$

   As before, $T_j \in \mathcal{D}'(\mathbb{R})$ for each $j \geq 1$. If supp$(\varphi) \subseteq B_r = \{x \in \mathbb{R} : |x| \leq r\}$ for some $r \geq 1$ and if $1 \leq j' < j$,

$$\begin{aligned} &\left| \int_{r > |x| > 1/j'} \frac{\varphi(x)}{|x|} dx - \int_{r > |x| > 1/j} \frac{\varphi(x)}{|x|} dx \right| \\ &= \left| \int_{1/j' \geq |x| > 1/j} \frac{\varphi(x)}{|x|} dx \right| \leq \left| \int_{1/j' \geq |x| > 1/j} \frac{\varphi(x) - \varphi(0)}{|x|} dx \right| \\ + |\varphi(0)| \left| \int_{1/j' \geq |x| > 1/j} \frac{dx}{|x|} \right| &= (1) + (2). \end{aligned}$$

Since the function $1/|x|$ is even,

$$\left|\int_{1/j'\geq|x|>1/j}\frac{dx}{|x|}\right| = 2\int_{1/j}^{1/j'}\frac{dx}{x} = 2\ln\frac{j}{j'}.$$

This divergence is unavoidable, so we simply put it aside, as we did in 2).

As in 1),

$$(1) = \left|\int_{1/j'\geq|x|>1/j}\frac{\varphi(x)-\varphi(0)}{|x|}dx\right| \leq 2\sup_{x\in\mathbb{R}}\left|\varphi'(x)\right|\left(\frac{1}{j'}-\frac{1}{j}\right)\underset{j,j'\to\infty}{\to} 0.$$

So, there is

$$\lim_{j\to\infty}\int_{r>|x|>1/j}\frac{\varphi(x)-\varphi(0)}{|x|}dx$$

in $\mathbb{C}$, for each $\varphi\in\mathcal{D}(\mathbb{R})$. Let us call the limit $T(\varphi)$. Since

$$\left|\int_{r>|x|>1/j}\frac{\varphi(x)-\varphi(0)}{|x|}dx\right| \leq 2\left(r-\frac{1}{j}\right)\sup_{x\in\mathbb{R}}\left|\varphi'(x)\right|,$$

we conclude that

$$|T(\varphi)| \leq 2r\sup_{x\in\mathbb{R}}\left|\varphi'(x)\right|,$$

for $\varphi\in\mathcal{D}_{B_r}(\mathbb{R})$.

Therefore, according to Remark 17 and Remark 22, T belongs to $\mathcal{D}'(\mathbb{R})$. The distribution T is called the finite part of $1/|x|$ and it is denoted $fp\,1/|x|$. That is,

$$\left(fp\frac{1}{|x|},\varphi\right) = \lim_{j\to\infty}\int_{|x|>1/j}\frac{\varphi(x)-\varphi(0)}{|x|}dx,$$

for $\varphi\in\mathcal{D}$.

It should be clear that $fp\,1/|x|$ extends T_f to $\mathbb{R}$ when $f(x)=1/|x|$.

Remark 44. 1. In Example 43, we could have used $\varepsilon>0$ instead of $1/j$ for $j\geq 1$, taking the limit as $\varepsilon\to 0$. But, strictly speaking, we would had been using a net, while we promised to work only with sequences. Of course, this is just a minor detail.

2. The distribution $pv\,1/x$ is odd, while the distributions $fp\,1/x^2$ and $fp\,1/|x|$ are even. Indeed, according to 4) in Example 23, if $\varphi \in \mathcal{D}$,

$$\begin{aligned}
\left(d_{-1}\left(pv\frac{1}{x}\right),\varphi\right) &= \lim_{j\to\infty}\int_{|x|>1/j}\frac{\varphi(-x)}{x}dx = \lim_{j\to\infty}\left(\int_{1/j}^{\infty}\frac{\varphi(-x)}{x}dx + \int_{-\infty}^{-1/j}\frac{\varphi(-x)}{x}dx\right) \\
&= \lim_{j\to\infty}\left(\int_{-1/j}^{-\infty}\frac{\varphi(x)}{x}dx + \int_{\infty}^{1/j}\frac{\varphi(x)}{x}dx\right) = -\lim_{j\to\infty}\int_{|x|>1/j}\frac{\varphi(x)}{x}dx \\
&= -\left(pv\frac{1}{x},\varphi\right),
\end{aligned}$$

while

$$\begin{aligned}
\left(d_{-1}\left(fp\frac{1}{x^2}\right),\varphi\right) &= \lim_{j\to\infty}\int_{|x|>1/j}\frac{\varphi(-x)-\varphi(0)-(-x)\varphi'(0)}{x^2}dx \\
&= \left(fp\frac{1}{x^2},\varphi\right)
\end{aligned}$$

and

$$\begin{aligned}
\left(d_{-1}\left(fp\frac{1}{|x|}\right),\varphi\right) &= \lim_{j\to\infty}\int_{|x|>1/j}\frac{\varphi(-x)-\varphi(0)}{|x|}dx \\
&= \left(fp\frac{1}{|x|},\varphi\right).
\end{aligned}$$

4 The spaces $\mathcal{S}$, $\mathcal{S}'$, $\mathcal{E}$, $\mathcal{E}'$, $\mathcal{D}^{(m)}$ and $\mathcal{D}^{(m)\prime}$

We begin with a definition.

Definition 45. *Let*

$$\mathcal{E} = \{\varphi : \mathbb{R}^n \to \mathbb{C} : \varphi \textit{ is smooth}\}.$$

The space $\mathcal{E}$ is a complex linear space. We consider in $\mathcal{E}$ the topology defined by the countable family of semi-norms

$$\|\varphi\|_{m,B_k} = \sup_{|\alpha|\le m}\sup_{x\in B_k}|(\partial^\alpha\varphi)(x)|, \tag{4.1}$$

for each $m = 0, 1, \ldots$ and each $B_k = \{x \in \mathbb{R}^n : |x| \le k\}$ for $k \ge 1$.

With this topology, $\mathcal{E}$ becomes a Fréchet space. A sequence $\{\varphi_j\}_{j\ge1}$ converges to φ in $\mathcal{E}$ exactly when, for each $\alpha \in \mathbb{N}^n$, the sequence $\{\partial^\alpha\varphi_j\}_{j\ge1}$ converges to $\partial^\alpha\varphi$ as $j \to \infty$, uniformly on each compact subset of $\mathbb{R}^n$.

Example 46. Let $\psi \in \mathcal{D}$ be so that $0 \leq \psi(x) \leq 1$ for all $x \in \mathbb{R}^n$. Moreover let

$$\psi(x) = \begin{cases} 1 & \text{if } |x| \leq 1 \\ 0 & \text{if } |x| \geq 2 \end{cases}.$$

Theorem 25 shows how to construct such a function.

Let $\psi_j(x) = \psi(x/j)$ for $j \geq 1$. Then, the sequence $\{\psi_j\}_{j\geq 1}$ converges in $\mathcal{E}$ as $j \to \infty$ to the function identically equal to one.

Indeed, if we fix $K \subseteq \mathbb{R}^n$ compact, there is $j_0 = j_0(K) \geq 1$ so that $|\psi_j(x) - 1| = 0$ for $x \in K$ and $j \geq j_0$. Thus, trivially, $\{\psi_j\}_{j\geq 1}$ converges to one uniformly on K, as $j \to \infty$.

If we fix $\alpha \in \mathbb{N}^n$, $|\alpha| \geq 1$ and $x \in K$,

$$\begin{aligned} |\partial^\alpha (\psi_j - 1)(x)| &= |\partial^\alpha (\psi\varphi_j(x))| = \frac{1}{j^{|\alpha|}} \left| (\partial^\alpha \psi)\left(\frac{x}{j}\right) \right| \\ &\leq \frac{1}{j^{|\alpha|}} \sup_{x\in K} |(\partial^\alpha \psi)(x)| \underset{j\to\infty}{\to} 0. \end{aligned}$$

Remark 47. A test function like ψ in Example 46, is called a cut-off function. When preserving smoothness is not an issue, the characteristic function of an appropriate set can be used as a cut-off function. This was done, for instance, when we truncated an L^p-function in Remark 29.

Definition 48. *A smooth function $\varphi : \mathbb{R}^n \to \mathbb{C}$ is rapidly decreasing at infinity with all its derivatives if for each $\alpha, \beta \in \mathbb{N}^n$,*

$$\sup_{x\in\mathbb{R}^n} \left| x^\alpha \left(\partial^\beta \varphi\right)(x) \right| < \infty. \tag{4.2}$$

We denote $\mathcal{S}$ the complex linear space consisting of those smooth functions that are rapidly decreasing at infinity with all its derivatives. The space $\mathcal{S}$ becomes a Fréchet space with the topology induced by the countable family of semi-norms

$$\|\varphi\|_{a,b} = \sup_{|\alpha|\leq a, |\beta|\leq b} \sup_{x\in\mathbb{R}^n} \left| x^\alpha \left(\partial^\beta \varphi\right)(x) \right|, \tag{4.3}$$

for $a, b = 0, 1, \ldots$. A sequence $\{\varphi_j\}_{j\geq 1}$ converges to φ in $\mathcal{S}$ exactly when $x^\alpha \partial^\beta \varphi_j$ converges to $x^\alpha \partial^\beta \varphi$ as $j \to \infty$, uniformly on $\mathbb{R}^n$, for each $\alpha, \beta \in \mathbb{N}^n$.

We will refer to functions in $\mathcal{S}$, simply, as smooth functions that are rapidly decreasing with all its derivatives.

Example 49. 1. The function

$$\varphi(x) = e^{-|x|^2}$$

belongs to $\mathcal{S}$. Indeed, since φ is the composition of two smooth functions, $t \to e^{-t}$ and $x \to |x|^2$, it is smooth. Moreover, it should be clear that φ is rapidly decreasing.

As for the derivatives, we claim that

$$\partial_x^\alpha \left(e^{-|x|^2}\right)(x) = e^{-|x|^2} P(x), \tag{4.4}$$

where $P(x)$ is a polynomial function of degree $|\alpha|$.

Indeed, when $|\alpha| = 1$,

$$\partial_{x_j} \left(e^{-|x|^2}\right)(x) = -2x_j e^{-|x|^2}.$$

If we assume that (4.4) is true when $|\alpha| \leq k$ for some $k \geq 1$,

$$\partial_{x_j} \partial^\alpha \left(e^{-|x|^2}\right)(x) = \partial_{x_j} \left(e^{-|x|^2} P(x)\right) = -2x_j P(x) e^{-|x|^2} + \left(\partial_{x_j} P\right)(x).$$

Therefore, (4.4) is true for all α. Then, it should be clear that all the derivatives of $e^{-|x|^2}$ are rapidly decreasing.

2. The function

$$\psi(x) = e^{-|x|^2} e^{ie^{|x|^2}},$$

which is smooth and rapidly decreasing, does not belong to $\mathcal{S}$, since

$$\partial_{x_j} \psi(x) = -2x_j e^{-|x|^2} e^{ie^{|x|^2}} + ie^{ie^{|x|^2}} e^{|x|^2} 2x_j$$

is not rapidly decreasing.

Lemma 50. *As a consequence of (4.2), there is*

$$\lim_{j \to \infty} \sup_{|x| > j} \left| x^\alpha \left(\partial^\beta \varphi\right)(x) \right| = 0,$$

for each $\alpha, \beta \in \mathbb{N}^n$, or equivalently, there is

$$\lim_{|x| \to \infty} \left| x^\alpha \left(\partial^\beta \varphi\right)(x) \right| = 0,$$

for each $\alpha, \beta \in \mathbb{N}^n$.

Proof. Let us fix $\alpha, \beta \in \mathbb{N}^n$ and $m = 1, 2, \ldots$. According to Corollary 2 and Remark 3,

$$\begin{aligned} \left(1 + |x|^2\right)^m \left| x^\alpha \left(\partial^\beta \varphi\right)(x) \right| &\leq \sum_{|\gamma|=0}^{m} \frac{m!}{(m - \gamma_1 - \ldots - \gamma_n)! \gamma_1! \gamma_2! \ldots \gamma_n!} \left| x^{2\gamma+\alpha} \left(\partial^\beta \varphi\right)(x) \right| \\ &\leq C_{m,n} \binom{m+n}{m} \sup_{|\gamma| \leq m} \sup_{x \in \mathbb{R}^n} \left| x^{2\gamma+\alpha} \left(\partial^\beta \varphi\right)(x) \right|. \end{aligned} \tag{4.5}$$

Thus,

$$\begin{aligned}
&\sup_{|x|>j}\left|x^{\alpha}\left(\partial^{\beta}\varphi\right)(x)\right| \\
\leq\ &\left[C_{m,n}\binom{m+n}{m}\sup_{|\gamma|\leq m}\sup_{x\in\mathbb{R}^{n}}\left|x^{2\gamma+\alpha}\left(\partial^{\beta}\varphi\right)(x)\right|\right]\sup_{|x|>j}\left(1+|x|^{2}\right)^{-m} \\
\leq\ &\left[C_{m,n}\binom{m+n}{m}\sup_{|\gamma|\leq m}\sup_{x\in\mathbb{R}^{n}}\left|x^{2\gamma+\alpha}\left(\partial^{\beta}\varphi\right)(x)\right|\right]\left(1+\frac{1}{j^{2}}\right)^{-m}\underset{j\to\infty}{\to}0.
\end{aligned}$$

This completes the proof of the lemma. □

Lemma 51. *Pointwise multiplication by a fixed function $\eta\in\mathcal{E}$ is a linear and continuous operator from $\mathcal{E}$ into $\mathcal{E}$.*

Proof. It should be clear that the operator is well defined and it is linear. Given $\varphi\in\mathcal{E}$ and $\alpha\in\mathbb{N}^{n}$, if $\alpha\neq 0$ we can write, according to Lemma 4,

$$\left(\partial^{\alpha}(\eta\varphi)\right)(x)=\eta(x)\,\partial^{\alpha}(\varphi)(x)+\sum_{0<\beta\leq\alpha}\binom{\alpha}{\beta}\left(\partial^{\beta}\eta\right)(x)\left(\partial^{\alpha-\beta}\varphi\right)(x). \tag{4.6}$$

If $\alpha=0$, we only have the first term in (4.6). Therefore, it suffices to work with $\alpha\neq 0$. From the estimates

$$\sup_{x\in B_k}\left|\eta(x)\,\partial^{\alpha}(\varphi)(x)\right|\leq\left(\sup_{x\in B_k}|\eta(x)|\right)\sup_{x\in B_k}\left|\partial^{\alpha}(\varphi)(x)\right|$$

and

$$\sup_{x\in B_k}\left|\sum_{0<\beta\leq\alpha}\binom{\alpha}{\beta}\left(\partial^{\beta}\eta\right)(x)\left(\partial^{\alpha-\beta}\varphi\right)(x)\right|\leq C_{\alpha}\left(\sup_{\beta\leq\alpha}\sup_{x\in B_k}\left|\left(\partial^{\beta}\eta\right)(x)\right|\right)\sup_{\beta\leq\alpha}\sup_{x\in B_k}\left(\partial^{\alpha-\beta}\varphi\right)(x),$$

follows the continuity of the multiplication.

This completes the proof of the lemma. □

Remark 52. Let us observe that the pointwise multiplication by a fixed function $\eta\in\mathcal{E}$ remains well defined and continuous, if considered between other appropriate spaces, for instance from $\mathcal{D}$ into $\mathcal{D}$. It is not well defined, in general, from $\mathcal{S}$ into $\mathcal{S}$, as can be seen by taking, for instance, $\eta(x)=e^{|x|^{2}}$.

Theorem 53. *We have the following continuous, dense, and strict inclusions:*

$$\mathcal{D}\hookrightarrow\mathcal{S}\hookrightarrow\mathcal{E}$$

Proof. It should be clear that, as sets, $\mathcal{D}$ is included in $\mathcal{S}$ and $\mathcal{S}$ is included in $\mathcal{E}$. The inclusions are strict, since, for instance, $e^{-|x|^2}$ belongs to $\mathcal{S}\backslash\mathcal{D}$ and $e^{|x|^2}$ belongs to $\mathcal{E}\backslash\mathcal{S}$.

If we fix $K \subseteq \mathbb{R}^n$ compact, and $\alpha, \beta \in \mathbb{N}^n$, the estimate

$$\sup_{x\in\mathbb{R}^n} \left| x^\alpha \left(\partial^\beta \varphi\right)(x)\right| = \sup_{x\in K} |x^\alpha| \sup_{x\in\mathbb{R}^n} \left|\left(\partial^\beta \varphi\right)(x)\right|$$

for $\varphi \in \mathcal{D}_K$, implies that the inclusion of $\mathcal{D}$ into $\mathcal{S}$ is continuous.

If we fix $K \subseteq \mathbb{R}^n$ compact, $\alpha \in \mathbb{N}^n$, the estimate

$$\sup_{x\in K} \left|\left(\partial^\beta \varphi\right)(x)\right| \leq \sup_{x\in\mathbb{R}^n} \left|\left(\partial^\beta \varphi\right)(x)\right|$$

for $\varphi \in \mathcal{S}$, implies that the inclusion of $\mathcal{S}$ into $\mathcal{E}$ is continuous.

So, we are left to prove that $\mathcal{D}$ is dense in $\mathcal{S}$ and that $\mathcal{S}$ is dense in $\mathcal{E}$. In both cases, we will use the sequence $\{\psi_j\}_{j\geq 1}$ defined in Example 46.

If we fix $\varphi \in \mathcal{S}$, we claim that the sequence $\{\varphi\psi_j\}_{j\geq 1}$, contained in $\mathcal{D}$, converges to φ in $\mathcal{S}$. To prove this claim, we need to show that given $a, b = 0, 1, ...,$ the sequence $x^\alpha \partial^\beta (\varphi(\psi_j - 1))(x)$ converges to zero as $j \to \infty$, uniformly with respect to $x \in \mathbb{R}^n$, for each $|\alpha| \leq a, |\beta| \leq b$.

If $\beta \neq 0$, using Leibniz's rule we can write

$$\begin{aligned} x^\alpha \partial^\beta (\varphi(\psi_j - 1))(x) &= (\psi_j - 1)(x)\, x^\alpha \left(\partial^\beta \varphi\right)(x) && (4.7)\\ &+ \sum_{0<\gamma\leq\beta} \binom{\beta}{\gamma} x^\alpha \left(\partial^{\beta-\gamma}\varphi\right)(x) \frac{1}{j^{|\gamma|}} \left(\partial^\gamma \psi\right)\left(\frac{x}{j}\right). && (4.8) \end{aligned}$$

If $\beta = 0$, we only have (4.7) with $\beta = 0$. Thus, it suffices to consider the case $\beta \neq 0$. Each term in (4.8) can be estimated as

$$\frac{1}{j^{|\gamma|}} \binom{\beta}{\gamma} \sup_{x\in\mathbb{R}^n} \left| x^\alpha \left(\partial^{\beta-\gamma}\varphi\right)(x)\right| \sup_{x\in\mathbb{R}^n} \left|\left(\partial^\gamma \psi\right)(x)\right| \underset{j\to\infty}{\to} 0.$$

As for (4.7), since $\varphi \in \mathcal{S}$ and $\psi_j(x) = 1$ for $|x| \leq j$,

$$\sup_{x\in\mathbb{R}^n} \left|(\psi_j - 1)(x)\, x^\alpha \left(\partial^\beta \varphi\right)(x)\right| \leq \sup_{|x|>j} \left| x^\alpha \left(\partial^\beta \varphi\right)(x)\right| \underset{j\to\infty}{\to} 0.$$

Hence, we have proved that $\mathcal{D}$ is dense in $\mathcal{S}$. To show that $\mathcal{S}$ is dense in $\mathcal{E}$, it is enough to show that $\mathcal{D}$ is dense in $\mathcal{E}$. So, let us fix $\eta \in \mathcal{E}$ and let us consider the sequence $\{\eta\psi_j\}_{j\geq 1}$, contained in $\mathcal{D}$. According to Lemma 51 and Example 46, the sequence $\{\eta\psi_j\}_{j\geq 1}$ converges to η in $\mathcal{E}$ as $j \to \infty$.

This completes the proof of the theorem. $\square$

Lemma 54. *If $\mathcal{A}$ denotes one of the spaces $\mathcal{D}, \mathcal{E}$ or $\mathcal{S}$, the operator ∂^{α} is linear and continuous from $\mathcal{A}$ into itself. Moreover,*

$$supp\left(\partial^{\alpha}\varphi\right) \subseteq supp\left(\varphi\right),$$

for $\varphi \in \mathcal{A}$.

Proof. Let us consider the case $\mathcal{A} = \mathcal{S}$, the other cases being similar. It should be clear that the operator ∂^{α} is well defined and it is linear. If we fix $a, b = 0, 1, ...,$

$$\begin{aligned} \left\|\partial^{\alpha}\varphi\right\|_{a,b} &= \sup_{|\gamma|\leq a, |\beta|\leq b} \sup_{x\in\mathbb{R}^n} \left|x^{\gamma}\left(\partial^{\beta+\alpha}\varphi\right)(x)\right| \\ &\leq \sup_{|\gamma|\leq a, |\beta|\leq b+|\alpha|} \sup_{x\in\mathbb{R}^n} \left|x^{\gamma}\left(\partial^{\beta}\varphi\right)(x)\right|, \end{aligned}$$

which shows that the operator ∂^{α} is continuous from $\mathcal{S}$ into itself.

Finally, if $x_0 \notin supp\left(\varphi\right)$, it means that φ is zero on an open neighborhood of x_0. Consequently, each partial derivative $\partial_{x_j}\varphi$ must be zero on the neighborhood. Iterating this process, we conclude that $\partial^{\alpha}\varphi$ is zero on that neighborhood.

This completes the proof of the lemma. □

Definition 55. *A continuous linear functional $T : \mathcal{S} \to \mathbb{C}$ is called a tempered distribution.*

In other words, a linear functional $T : \mathcal{S} \to \mathbb{C}$ is a tempered distribution exactly when $T\left(\varphi_j\right) \to 0$ in $\mathbb{C}$ as $j \to \infty$, whenever $\varphi_j \to 0$ in the Fréchet space $\mathcal{S}$ as $j \to \infty$.

The complex linear space consisting of all the tempered distributions is denoted $\mathcal{S}'$. By definition, $\mathcal{S}'$ is the topological dual of $\mathcal{S}$.

Remark 56.

1. According to Theorem 53, if $T \in \mathcal{S}'$, then $T\,|\mathcal{D} : \mathcal{D} \to \mathbb{C}$ belongs to $\mathcal{D}'$. That is, we can say that a tempered distribution is a distribution in the sense of Definition 19 or Definition 20.

   Since $\mathcal{D}$ is dense in $\mathcal{S}$, a tempered distribution is uniquely determined by its restriction to $\mathcal{D}$.

2. Given a distribution T in $\mathcal{D}'$, it is not always possible to extend it to a distribution in $\mathcal{S}'$.

   For example, let us consider the distribution in $\mathcal{D}'$ defined by the function $e^{|x|^2}$. It should be clear that $T_{e^{|x|^2}}$ is not a tempered distribution, since, for instance, the pairing $\left(T_{e^{|x|^2}}, e^{|x|^2}\right)$ does not equal a complex number.

   Furthermore, there is no tempered distribution T such that

$$(T, \varphi) = \int_{\mathbb{R}^n} e^{|x|^2}\varphi(x)\,dx$$

for all $\varphi \in \mathcal{D}$. To prove it, we consider the sequence $\left\{e^{-|x|^2}\psi_j\right\}_{j\geq 1}$, where $\psi_j(x) = \psi(x/j)$ is the function in Example 46.

As shown in the proof of Theorem 53, there is $\lim_{j\to\infty} \varphi\psi_j = \varphi$ in $\mathcal{S}$. Therefore, if such a tempered distribution T would exist,

$$\begin{aligned}\left(T, e^{-|x|^2}\right) &= \lim_{j\to\infty}\left(T, e^{-|x|^2}\psi_j\right) = \lim_{j\to\infty}\int_{\mathbb{R}^n} e^{|x|^2}e^{-|x|^2}\psi\left(\frac{x}{j}\right)dx \\ &\geq \lim_{j\to\infty} meas\left(B_j\right)\end{aligned}$$

where $B_j = \{x \in \mathbb{R}^n : |x| \leq j\}$, which is not possible.

Theorem 57. *The space $\mathcal{S}$ is continuously included in L^p for all $1 \leq p \leq \infty$. Moreover, the inclusion is dense for $1 \leq p < \infty$.*

Proof. Given $\varphi \in \mathcal{S}$,

$$\|\varphi\|_{L^\infty} = \sup_{x\in\mathbb{R}^n}|\varphi(x)|.$$

Thus, $\mathcal{S}$ is continuously included in L^∞.

If $\varphi \in \mathcal{S}$, $1 \leq p < \infty$ and $m = 1, 2, \ldots$,

$$\|\varphi\|_{L^p}^p \leq \underbrace{\left(\sup_{x\in\mathbb{R}^n}\left|\left(1+|x|^2\right)^m\varphi(x)\right|^p\right)}_{(1)}\int_{\mathbb{R}^n}\left(1+|x|^2\right)^{-mp}dx.$$

Estimate (4.5) takes care of (1). As for the integral,

$$\begin{aligned}\int_{\mathbb{R}^n}\left(1+|x|^2\right)^{-mp}dx &= C_n\int_0^\infty\left(1+r^2\right)^{-mp}r^{n-1}dr \\ &\leq C_n\left(\int_0^1 r^{n-1}dr + \int_1^\infty\left(1+r^2\right)^{-mp}dr\right) < \infty, \quad (4.9)\end{aligned}$$

if $m > n/2p$.

That $\mathcal{S}$ is dense in L^p for $1 \leq p < \infty$, follows from Theorem 28 and Theorem 53. This completes the proof of the theorem. □

The following result is a fairly immediate consequence of Theorem 57.

Corollary 58. *Given $\varphi \in \mathcal{S}$, the functions $\partial^\alpha\left(x^\beta\varphi\right)$ and $x^\beta\partial^\alpha\varphi$ belong to L^p for all $1 \leq p \leq \infty$ and for all $\alpha, \beta \in \mathbb{N}^n$.*

Example 59. 1. According to Theorem 57, if $f \in L^p$ for any $1 \leq p \leq \infty$, the distribution T_f belongs to $\mathcal{S}'$.

2. If P is a complex polynomial function, $P(x) = \sum_{0\leq|\alpha|\leq d} c_\alpha x^\alpha$ with $c_\alpha \in \mathbb{C}$, the distribution T_P is tempered. Indeed, given $\varphi \in \mathcal{S}$,

$$\begin{aligned}\int_{\mathbb{R}^n} |P(x)\varphi(x)|\,dx &\leq \sum_{|\alpha|=0}^{d} |c_\alpha| \int_{\mathbb{R}^n} |x^\alpha \varphi(x)|\,dx \\ &\leq C_P \underbrace{\left(\sup_{x\in\mathbb{R}^n} \left|\left(1+|x|^2\right)^m x^\alpha \varphi(x)\right|\right)}_{(1)} \underbrace{\int_{\mathbb{R}^n} \left(1+|x|^2\right)^{-m} dx}_{(2)}.\end{aligned}$$

   Now, (4.5) takes care of (1), while (4.9) takes care of (2).

3. A function $f : \mathbb{R}^n \to \mathbb{C}$ is slowly increasing at infinity if there is a number $k \in \mathbb{N}$ such that

$$\sup_{x\in\mathbb{R}^n} \left|\frac{f(x)}{\left(1+|x|^2\right)^k}\right| < \infty.$$

   For instance, a polynomial function P as in 1), is slowly increasing at infinity.

   We will refer to these functions, simply, as slowly increasing.

4. In general, if f is a slowly increasing and Lebesgue measurable function, T_f is a tempered distribution. The verification requires calculations that are very similar to those used in 2).

5. The distributions $pv\,1/x$, $fp\,1/x^2$, and $fp\,1/|x|$, defined in Example 43, are tempered distributions.

   Indeed, given $\varphi \in \mathcal{S}$,

$$\int_{|x|>1/j} \frac{\varphi(x)}{x}dx = \underbrace{\int_{|x|\geq 2} \frac{\varphi(x)}{x}dx}_{(1)} + \underbrace{\int_{2>|x|>1/j} \frac{\varphi(x)}{x}dx}_{(2)}$$

   As in 1) of Example 43, there is $\lim_{j\to\infty}(2)$ and

$$|(2)| = \left|\int_{2>|x|>1/j} \frac{\varphi(x)-\varphi(0)}{x}dx\right| \leq 2\left(2-\frac{1}{j}\right)\sup_{x\in\mathbb{R}} \left|\varphi'(x)\right|,$$

   while

$$|(1)| \leq \sup_{x\in\mathbb{R}} |x\varphi(x)| \int_{|x|\geq 2} \frac{dx}{x^2}.$$

As in 2) of Example 43, there is

$$\lim_{j\to\infty}\int_{2>|x|>1/j}\frac{\varphi(x)-\varphi(0)-x\varphi'(0)}{x^2}dx.$$

Therefore,

$$\begin{aligned} &\left|fp\int_{|x|>1/j}\frac{\varphi(x)}{x^2}dx\right| \qquad (4.10)\\ \leq\ &\left|\int_{|x|\geq 2}\frac{\varphi(x)}{x^2}dx\right|+\left|\int_{2>|x|>1/j}\frac{\varphi(x)-\varphi(0)-x\varphi'(0)}{x^2}dx\right|\\ \leq\ &\sup_{x\in\mathbb{R}}|\varphi(x)|\int_{|x|\geq 2}\frac{dx}{x^2}+2\sup_{x\in\mathbb{R}}\left|\varphi''(x)\right|\left(2-\frac{1}{j}\right), \end{aligned}$$

where fp in (4.10) indicates the finite terms in the integral.

Finally, there is

$$\lim_{j\to\infty}\int_{2>|x|>1/j}\frac{\varphi(x)-\varphi(0)}{|x|}dx,$$

so,

$$\begin{aligned} &\left|fp\int_{|x|>1/j}\frac{\varphi(x)}{|x|}dx\right|\\ \leq\ &\left|\int_{|x|\geq 2}\frac{\varphi(x)}{|x|}dx\right|+\left|\int_{2>|x|>1/j}\frac{\varphi(x)-\varphi(0)}{|x|}dx\right|\\ \leq\ &\sup_{x\in\mathbb{R}}|x\varphi(x)|\int_{|x|\geq 2}\frac{dx}{x^2}+2\sup_{x\in\mathbb{R}}\left|\varphi'(x)\right|\left(2-\frac{1}{j}\right). \end{aligned}$$

6. With the same work done in 5) of Example 23, we can prove that the series $\sum_{j\geq 1}e^{j^2}\delta_j$ defines a distribution T in $\mathcal{D}'(\mathbb{R})$. However, T is not a tempered distribution, since the pairing

$$\left(\sum_{j=1}^{k}e^{j^2}\delta_j, e^{-x^2}\right)=k,$$

does not have a limit in $\mathbb{C}$ as $k\to\infty$. Furthermore, there is no tempered distribution T such that

$$T\,|\mathcal{D}=\sum_{j\geq 1}e^{j^2}\delta_j.$$

To see it, we only need to follow the proof of 2) in Remark 56.

The proof of the next result should be pretty straightforward, after the numerous estimates we have done on functions in the space $\mathcal{S}$.

Theorem 60. *1. Given $\varphi \in \mathcal{E}$, the following statements are equivalent:*

(a) $\varphi \in \mathcal{S}$.

(b) For each polynomial function $P(x)$ and each $\alpha \in \mathbb{N}^n$, there is $C_{P,\alpha} > 0$ so that

$$\sup_{x\in\mathbb{R}^n} |\partial^\alpha (P(x)\varphi)(x)| \leq C_{P,\alpha}.$$

(c) Given $\alpha \in \mathbb{N}^n$ and $r \in \mathbb{R}$, there is $C_{\alpha,r} > 0$ so that

$$\sup_{x\in\mathbb{R}^n} \left| \left(1 + |x|^2\right)^r (\partial^\alpha \varphi)(x) \right| \leq C_{\alpha,r}.$$

2. The formulas

$$\sup_{|\alpha|\leq m, |\beta|\leq l} \sup_{x\in\mathbb{R}^n} \left| \partial^\alpha \left(x^\beta \varphi \right)(x) \right|$$

and

$$\sup_{|\alpha|\leq m} \sup_{x\in\mathbb{R}^n} \left| \left(1 + |x|^2\right)^r (\partial^\alpha \varphi)(x) \right|$$

define families of semi-norms that are equivalent to the family of semi-norms given by (4.3).

Definition 61. *The linear space consisting of all the linear and continuous functionals $T : \mathcal{E} \to \mathbb{C}$ is denoted $\mathcal{E}'$. By definition, $\mathcal{E}'$ is the topological dual of $\mathcal{E}$.*

According to Theorem 53, given $T \in \mathcal{E}'$, the restriction $T : \mathcal{S} \to \mathbb{C}$ is a tempered distribution. Moreover, T is uniquely determined by its values on $\mathcal{D}$. Furthermore, we have the following result:

Theorem 62. *The space $\mathcal{E}'$ can be identified, as a linear space, with the linear subspace of $\mathcal{D}'$ consisting of distributions with compact support.*

Proof. Let

$$\mathcal{K} = \left\{ T \in \mathcal{D}' : supp\,(T) \text{ is compact} \right\}.$$

We will prove that the restriction operator $r : \mathcal{E}' \to \mathcal{D}'$, defined as $r(T) = T\,|\mathcal{D}$, is a bijection between $\mathcal{E}'$ and $\mathcal{K}$.

Let us begin by showing that $r(\mathcal{E}') \subseteq \mathcal{K}$.

Given $T \in \mathcal{E}'$, let $T_1 = T\,|\mathcal{D}$. If T_1 does not have compact support, for each $k \geq 1$ there is $\varphi_k \in \mathcal{D}$ so that $supp\,(\varphi_k) \subseteq \{x \in \mathbb{R}^n : |x| > k\}$ and $T_1(\varphi_k) \neq 0$. We can assume that $T_1(\varphi_k) = 1$ for all $k \geq 1$.

The sequence $\{\varphi_k\}_{k\geq 1}$ converges trivially to zero in $\mathcal{E}$ as $k \to \infty$. Hence, $T_1(\varphi_k) = T(\varphi_k)$ must converge to zero in $\mathbb{C}$ as $k \to \infty$, which is a contradiction. So $T_1 \in \mathcal{K}$. It should be clear that r is linear.

Next we show that the restriction map r is injective from $\mathcal{E}'$ into $\mathcal{K}$.

Given $\varphi \in \mathcal{E}$, according to Theorem 53 there is a sequence $\{\varphi_j\}_{j\geq 1}$ in $\mathcal{D}$ so that $\varphi_j \to \varphi$ in $\mathcal{E}$ as $j \to \infty$. Now, given $T \in \mathcal{E}'$, if $r(T) = 0$, we have

$$(T, \varphi)_{\mathcal{E}',\mathcal{E}} = \lim_{j\to\infty} (T, \varphi_j)_{\mathcal{E}',\mathcal{E}} = \lim_{j\to\infty} (r(T), \varphi_j)_{\mathcal{D}',\mathcal{D}} = 0,$$

since $(r(T), \varphi_j)_{\mathcal{D}',\mathcal{D}} = 0$ for every $j \geq 1$. Thus, $T = 0$.

Let us observe that the subscripts indicate, as mentioned in Definition 19, the duality used.

To prove that r is surjective, we will define a linear map $e : \mathcal{K} \to \mathcal{E}'$ such that $r \circ e$ is the identity map on $\mathcal{K}$.

Given $T \in \mathcal{K}$, we pick a function $\alpha \in \mathcal{D}$ so that $\alpha = 1$ on an open neighborhood of $supp\,(T)$. Then, we define the linear functional T_2 from $\mathcal{E}$ into $\mathbb{C}$ as

$$T_2(\varphi) = T(\alpha\varphi).$$

First, we claim that $T_2 \in \mathcal{E}'$. In fact, if $\varphi_j \to 0$ in $\mathcal{E}$ as $j \to \infty$, it should be clear that $\alpha\varphi_j \to 0$ in $\mathcal{D}$ as $j \to \infty$ and therefore, $T_2(\varphi_j) = T(\alpha\varphi_j) \to 0$ in $\mathbb{C}$ as $j \to \infty$.

Moreover, the definition of T_2 is independent of α satisfying the stated conditions. Indeed, if $\alpha, \beta \in \mathcal{D}$ are equal to one, each on an open neighborhood of $supp\,(T)$, we can say that $supp\,(T) \bigcap supp\,((\alpha - \beta)\varphi) = \varnothing$, for $\varphi \in \mathcal{E}$. Therefore, Lemma 41 implies that $T((\alpha - \beta)\varphi) = 0$ or $T(\alpha\varphi) = T(\beta\varphi)$.

So, the map e is well defined. It should be clear that e is linear and that $r \circ e$ is the identity map on $\mathcal{K}$.

This completes the proof of the theorem. $\square$

Remark 63. When we refer to $T \in \mathcal{E}'$ as being a distribution with compact support, we will always mean in the sense of Theorem 62.

Example 64.

1. According to 1) in Example 40, the distributions δ_a and T_a^m belong to $\mathcal{E}'$.

   Moreover, if a function f has compact support and belongs to L^p_{loc} for some $1 \leq p \leq \infty$, T_f belongs to $\mathcal{E}'$.

2. Given a sequence $\{a_j\}_{j\geq 1}$ converging to zero in $\mathbb{R}$ as $j \to \infty$, the series

$$\sum_{j\geq 1} \frac{\delta_{a_j}}{j^2}$$

   defines a distribution T in $\mathcal{E}'(\mathbb{R})$.

Indeed, if $\varphi \in \mathcal{E}(\mathbb{R})$ and $K \subseteq \mathbb{R}$ is compact,

$$\left| \left(\sum_{j\geq 1} \frac{\delta_{a_j}}{j^2}, \varphi \right) \right| = \left| \sum_{j\geq 1} \frac{\varphi(a_j)}{j^2} \right| \leq \left(\sum_{j\geq 1} \frac{1}{j^2} \right) \sup_{x\in K} |\varphi(x)|,$$

which shows that $T \in \mathcal{E}'(\mathbb{R})$.

Moreover, if $a = \sup_{j\geq 1} \{|a_j|\}$ and $supp(\varphi) \subseteq \{x \in \mathbb{R} : |x| > a+1\}$, it should be clear that $(T, \varphi) = 0$. Therefore, $supp(T) \subseteq [-a-1, a+1]$.

Lemma 65. *Let $T \in \mathcal{D}'$ and let $\varphi \in \mathcal{E}$. If $supp(T) \bigcap supp(\varphi)$ is a compact set K, the pairing (T, φ) can be defined.*

Proof. Given $\varepsilon > 0$, let $\alpha \in \mathcal{D}$ be so that $\alpha(x) = 1$ for x in the ε-neighborhood(K) and $supp(\alpha) \subseteq 2\varepsilon$-neighborhood$(K)$.

Then, $\alpha\varphi \in \mathcal{D}$ and we can write $(T, \alpha\varphi)_{\mathcal{D}'\mathcal{D}}$.

We only need to show that the pairing does not depend on the function α satisfying the stated conditions. In fact, if $\beta \in \mathcal{D}$ is another function like α, we claim that $supp(T) \bigcap supp((\alpha-\beta)\varphi) = \varnothing$.

Indeed,

$$\begin{aligned} supp((\alpha-\beta)\varphi) &\subseteq (\mathbb{R}^n \backslash \varepsilon\text{-neighborhood}(K)) \bigcap supp(\varphi) \\ &\subseteq \mathbb{R}^n \backslash \left(supp(T) \bigcap supp(\varphi) \right) \bigcap supp(\varphi) \\ &= \left((\mathbb{R}^n \backslash supp(T)) \bigcup \mathbb{R}^n \backslash supp(\varphi) \right) \bigcap supp(\varphi) \\ &\subseteq (\mathbb{R}^n \backslash supp(T)). \end{aligned}$$

Thus, according to Lemma 41,

$$(T, (\alpha-\beta)\varphi)_{\mathcal{D}'\mathcal{D}} = 0.$$

This completes the proof of the lemma. $\square$

Remark 66. In Lemma 65 we can use, for instance, open balls centered at zero, as neighborhoods of the intersection.

Remark 67. A subset $\mathcal{B}$ of $\mathcal{E}$ is bounded when, for each $K \subseteq \mathbb{R}^n$ compact and for each $\alpha \in \mathbb{N}^n$,

$$\sup_{\varphi\in\mathcal{B}} \sup_{x\in K} |(\partial^\alpha \varphi)(x)| < \infty.$$

In the space $\mathcal{E}'$ we consider the dual topology, also called strong topology, induced by the uncountable family of semi-norms

$$\|T\|_{\mathcal{E}',\mathcal{B}} = \sup_{\varphi\in\mathcal{B}} |T(\varphi)|, \tag{4.11}$$

for all the bounded subsets $\mathcal{B}$ of $\mathcal{E}$. With this topology, $\mathcal{E}'$ is a complete, locally convex and non-metrizable topological linear space ([25], p. 89). We will always use this topology in $\mathcal{E}'$.

Let us observe that, although Theorem 62 identifies $\mathcal{E}'$, as a linear space, with the linear subspace $\mathcal{K}$ of $\mathcal{D}'$, the topology of $\mathcal{E}'$ is different, in fact it is finer, than the topology induced by $\mathcal{D}'$ in $\mathcal{K}$ ([25], p. 89).

Still, according to Remark 22, the bounded subsets of $\mathcal{D}$ are also bounded in $\mathcal{E}$. So, if $\{T_j\}_{j\geq 1}$ converges to zero in $\mathcal{E}'$ as $j \to \infty$, the sequence $\{r(T_j)\}_{j\geq 1}$ converges to zero in $\mathcal{D}'$ as $j \to \infty$.

From (4.11), it should be clear that a sequence $\{T_j\}_{j\geq 1}$ converges to zero in the strong topology of $\mathcal{E}'$ as $j \to \infty$ if, and only if, $T_j(\varphi) \to 0$ as $j \to \infty$, uniformly on $\mathcal{B}$, for each bounded subset $\mathcal{B}$ of $\mathcal{E}$. That is to say if, and only if,

$$\sup_{\varphi \in \mathcal{B}} |T_j(\varphi)| \underset{j\to\infty}{\to} 0$$

for each $\mathcal{B} \subseteq \mathcal{E}$ bounded. This form of convergence is called the strong convergence. Moreover, the space $\mathcal{E}$ with the topology induced by the semi-norms (4.1), is the topological dual of the space $\mathcal{E}'$ with the strong topology ([25], p. 90). Thus, $\mathcal{E}$ and $\mathcal{E}'$ are reflexive. Let us point out that the strong topology in $\mathcal{E}'$ coincides with the inductive limit topology induced by the spaces

$$\mathcal{E}'_K = \left\{ T \in \mathcal{E}' : supp\,(T) \subseteq K \right\},$$

for $K \subseteq \mathbb{R}^n$ compact ([25], p. 90).

The strong, or dual, topology of the space $\mathcal{S}'$ is defined in a similar manner. First, we observe that a subset $\mathcal{B}$ of $\mathcal{S}$ is bounded exactly when, for each $\alpha, \beta \in \mathbb{N}^n$,

$$\sup_{\varphi \in \mathcal{B}} \sup_{x \in \mathbb{R}^n} \left| x^\alpha \left(\partial^\beta \varphi\right)(x) \right| < \infty.$$

We consider in $\mathcal{S}'$ the topology induced by the uncountable family of semi-norms

$$\|T\|_{\mathcal{S}',\mathcal{B}} = \sup_{\varphi \in \mathcal{B}} |T(\varphi)|,$$

for all the bounded subsets $\mathcal{B}$ of $\mathcal{S}$. We will always use this topology in $\mathcal{S}'$. With this topology, $\mathcal{S}'$ becomes a complete, locally convex and non-metrizable topological linear space, for which $\mathcal{S}$ is the topological dual ([25], p. 238). Thus, $\mathcal{S}$ and $\mathcal{S}'$ are reflexive.

A sequence $\{T_j\}_{j\geq 1}$converges to zero in the strong topology of $\mathcal{S}'$ as $j \to \infty$ if, and only if, $T_j(\varphi) \to 0$ as $j \to \infty$, uniformly on $\mathcal{B}$, for each bounded subset $\mathcal{B}$ of $\mathcal{S}$. That is to say if, and only if,

$$\sup_{\varphi \in \mathcal{B}} |T_j(\varphi)| \underset{j\to\infty}{\to} 0$$

for each $\mathcal{B} \subseteq \mathcal{S}$ bounded. Once again, this form of convergence is called the strong convergence.

This brief account of the strong topologies in $\mathcal{E}'$ and $\mathcal{S}'$ will suffice for our purpose. For more on these matters, we refer to ([25], Chapter III, Section 7) and Chapter VII, Sections 3 and 4).

Theorem 68. *If ρ_j is the function defined in the proof of Theorem 25, the sequence $\{T_{\rho_j}\}_{j\geq 1}$ converges to δ_0 in $\mathcal{E}'$ as $j \to \infty$.*

Proof. Let us fix a bounded subset $\mathcal{B}$ of $\mathcal{E}$. Then, for $\varphi \in \mathcal{B}$,

$$\begin{aligned}
\left| \int_{\mathbb{R}^n} \rho_j(x)\, \varphi(x)\, dx - \varphi(0) \right| &\leq \int_{|x|\leq 1/j} \rho_j(x)\, |\varphi(x) - \varphi(0)|\, dx \\
&\leq C_n \sup_{\varphi\in\mathcal{B}} \sup_{1\leq l\leq n} \sup_{x\in K} |(\partial_{x_l}\varphi)(x)| \int_{|x|\leq 1/j} \rho_j(x)\, |x|\, dx \\
&= \frac{C_n}{j} \sup_{\varphi\in\mathcal{B}} \sup_{1\leq l\leq n} \sup_{x\in K} |(\partial_{x_l}\varphi)(x)| \underset{j\to\infty}{\to} 0,
\end{aligned}$$

where $K \subseteq \mathbb{R}^n$ is compact.

This completes the proof of the theorem. $\square$

Theorem 69. *Given $1 \leq p \leq \infty$, the map*

$$\begin{array}{ccc} L^p & \to & \mathcal{S}' \\ f & \to & T_f \end{array}$$

is well defined, linear, and continuous.

Proof. If $\varphi \in \mathcal{S}$, we have, for $p = \infty$,

$$\int_{\mathbb{R}^n} |f(x)|\, |\varphi(x)|\, dx \leq \|f\|_{L^\infty} \sup_{x\in\mathbb{R}^n} \left| \left(1+|x|^2\right)^m \varphi(x) \right| \int_{\mathbb{R}^n} \left(1+|x|^2\right)^{-m} dx < \infty$$

for $m > n/2$.

If $1 \leq p < \infty$,

$$\begin{aligned}
\int_{\mathbb{R}^n} |f(x)|\, |\varphi(x)|\, dx &\leq \sup_{x\in\mathbb{R}^n} \left| \left(1+|x|^2\right)^m \varphi(x) \right| \int_{\mathbb{R}^n} |f(x)| \left(1+|x|^2\right)^{-m} dx \\
&\leq \sup_{x\in\mathbb{R}^n} \left| \left(1+|x|^2\right)^m \varphi(x) \right| \|f\|_{L^p} \left\| \left(1+|x|^2\right)^{-m} \right\|_{L^q} < \infty,
\end{aligned}$$

for $mq > n/2$, where q is the conjugate exponent of p.

Therefore, the map is well defined. It should be clear that it is linear.

As for the continuity, if we fix a bounded subset $\mathcal{B}$ of $\mathcal{S}$, the calculations done above and Corollary 2, imply that

$$\|T_f\|_{\mathcal{S}',\mathcal{B}} \leq C_{m,p,n} \|f\|_{L^p} \sup_{\varphi\in\mathcal{B}} \sup_{|\alpha|\leq 2m} \sup_{x\in\mathbb{R}^n} |x^\alpha \varphi(x)|.$$

So, the map is continuous.

This completes the proof of the theorem. □

Before going onto the spaces $\mathcal{D}^{(m)}$ and $\mathcal{D}^{(m)\prime}$, we prove the following characterization of $\mathcal{D}'$:

Theorem 70. *Given a linear functional $T : \mathcal{D} \to \mathbb{C}$, the following statements are equivalent:*

1. *T is a distribution.*
2. *For each compact subset K of $\mathbb{R}^n$, there is $m_K \in \mathbb{N}$ and $C_{K,m} > 0$ so that*

$$|T(\varphi)| \leq C_{K,m} \sup_{|\alpha|\leq m_K} \sup_{x\in\mathbb{R}^n} |\partial^\alpha \varphi(x)|, \tag{4.12}$$

for all $\varphi \in \mathcal{D}_K$.

Proof. Let us assume that the linear functional $T : \mathcal{D} \to \mathbb{C}$ does not satisfy (4.12). This means that there is a compact subset K of $\mathbb{R}^n$ such that for each $C > 0$ and $m = 1, 2, ...$, there is $\varphi \in \mathcal{D}_K$ so that

$$|T(\varphi)| > C \sup_{|\alpha|\leq m} \sup_{x\in\mathbb{R}^n} |\partial^\alpha \varphi(x)|.$$

In particular, if $C = m$, there is $\psi_m \in \mathcal{D}_K$ satisfying

$$|T(\psi_m)| > m \sup_{|\alpha|\leq m} \sup_{x\in\mathbb{R}^n} |\partial^\alpha \psi_m(x)|.$$

Let

$$\varphi_m = \frac{\psi_m}{|T(\psi_m)|}.$$

Given $\beta \in \mathbb{N}^n$, if $m > |\beta|$,

$$\sup_{x\in\mathbb{R}^n} \left|\partial^\beta \varphi_m(x)\right| \leq \sup_{|\alpha|\leq m} \sup_{x\in\mathbb{R}^n} |\partial^\alpha \varphi_m(x)| < \frac{1}{m} \underset{j\to\infty}{\to} 0.$$

That is to say, the sequence $\{\varphi_m\}_{m\geq 1}$ converges to zero in $\mathcal{D}$ as $m \to \infty$. However, by construction, $T(\varphi_m) = 1$ for all m. This shows that T is not a distribution.

So, we have proved 1) ⇒ 2).

Conversely, if $T : \mathcal{D} \to \mathbb{C}$ is a linear functional satisfying (4.12), let us see that $T \in \mathcal{D}'$. That is to say, let us see that T is continuous in the strong topology of $\mathcal{D}$.

In fact, if $\{\varphi_j\}_{j\geq 1}$ converges to zero in $\mathcal{D}$ as $j \to \infty$, then $\varphi_j \in \mathcal{D}_K$ for some $K \subseteq \mathbb{R}^n$ compact. Moreover, for each $\alpha \in \mathbb{N}^n$, the sequence $\{(\partial^\alpha \varphi_j)(x)\}_{j\geq 1}$ converges to zero uniformly for $x \in \mathbb{R}^n$ as $j \to \infty$. Thus, according to (4.12),

$$|T(\varphi_j)| \leq C_{K,m} \sup_{|\alpha|\leq m_K} \sup_{x\in\mathbb{R}^n} |(\partial^\alpha \varphi_j)(x)| \underset{j\to\infty}{\to} 0.$$

So, T is a distribution.

This completes the proof of the theorem. □

Definition 71. *When m_K in (4.12) can be chosen independently of K, we say that the distribution T has finite order $\leq m$. If this is not the case, we say that the distribution has infinite order.*

Example 72. According to Example 23, the distribution T_f defined by a locally integrable function f, and the Dirac distribution δ_a, both have finite order equal to zero.

According to (3.5), the distribution T_a^m has finite order $\leq m$. Actually, the order is m.

Indeed, let $P(x_1)$ be the polynomial function x_1^m and let α be a function in $\mathcal{D}$ satisfying the following conditions: $0 \leq \alpha(x) \leq 1$ for all $x \in \mathbb{R}^n$, $\alpha(x) = 1$ when $|x| \leq 1/3$ and $\alpha(x) = 0$ when $|x| \geq 1/2$. Then, for $j \geq 1$,

$$(T_a^m, P(j(\cdot_1 - a_1))\,\alpha(j(\cdot - a))) = j^m m!.$$

If we assume that the order is m' for some $0 \leq m' < m$, we can write,

$$\begin{aligned} j^m m! &\leq C_{m'} \sup_{|\alpha|\leq m'} \sup_{x\in\mathbb{R}^n} |\partial^\alpha (P(j(\cdot_1 - a_1))\,\alpha(j(\cdot - a)))(x)| \\ &\leq C_{m'} j^{m'} \sup_{|\alpha|\leq m'} \sup_{|x-a|\leq 1/2} |\partial^\alpha (P(\cdot_1 - a_1)\,\alpha(\cdot - a))| \end{aligned}$$

for all $j \geq 1$, which is not possible.

Estimates (3.11) and (3.14) show, respectively, that the distribution $pv\,1/x$ has finite order ≤ 1, while the distribution $fp\,1/x^2$ has finite order ≤ 2. Actually, $pv\,1/x$ does not belong to $\mathcal{D}'^{(0)}$ and $fp\,1/x^2$ does not belong to $\mathcal{D}'^{(1)}$.

Nevertheless, it is quite satisfactory to have just an upper bound for the order of a distribution of finite order.

The distribution T_a^m shows that there are distributions in $\mathcal{D}'$ of arbitrarily large finite order. There are also distributions of infinite order.

Indeed, for $\varphi \in \mathcal{D}(\mathbb{R})$, let us consider the formal series

$$\sum_{k\geq 1} \varphi^{(k)}(k). \tag{4.13}$$

We observe that if $\varphi \in \mathcal{D}_K(\mathbb{R})$ for a fixed compact set $K \subseteq \mathbb{R}$, there is $k_K \geq 1$ so that

$$\sum_{k\geq 1} \varphi^{(k)}(k) = \sum_{1\leq k\leq k_K} \varphi^{(k)}(k).$$

Therefore, the series is well defined in $\mathcal{D}_K(\mathbb{R})$. Moreover,

$$\left|\sum_{k\geq 1} \varphi^{(k)}(k)\right| \leq \sup_{1\leq k\leq k_K} \sup_{x\in\mathbb{R}} \left|\varphi^{(k)}(x)\right|,$$

for every $\varphi \in \mathcal{D}_K(\mathbb{R})$. Hence, (4.13) defines a distribution T in $\mathcal{D}'(\mathbb{R})$.

Let us assume that T has finite order $\leq m_0$. We fix $k_0 > m_0$ and we consider again the test function $P(j(\cdot_1 - a))\,\alpha(j(\cdot - a))$ for $m = k_0$, $a = k_0$ and $j \geq 1$. Then,

$$\begin{aligned} j^{k_0} k_0! &= (T, P(j(\cdot_1 - k_0))\,\alpha(j(\cdot - k_0))) \\ &\leq C_{m_0} j^{m_0} \sup_{|\alpha|\leq m_0} \sup_{|x-k_0|\leq 1/2} \left|\partial^\alpha (P(\cdot_1 - k_0)\,\alpha(\cdot - k_0))\right|, \end{aligned}$$

which it is not possible for all $j \geq 1$.

Definition 73. *We denote $\mathcal{D}'^{(m)}$ the linear subspace of $\mathcal{D}'$ consisting of those distributions that have finite order $\leq m$.*

Definition 74. *Let $\mathcal{D}^{(m)}$ be the linear space of those functions $\varphi : \mathbb{R}^n \to \mathbb{C}$ with compact support, that are continuous and have continuous derivatives of order $\leq m$. For $m = 0$, the functions in $\mathcal{D}^{(0)}$ are continuous functions with compact support.*

If $K \subseteq \mathbb{R}^n$ is compact, let

$$\mathcal{D}_K^{(m)} = \left\{\varphi \in \mathcal{D}^{(m)} : supp(\varphi) \subseteq K\right\}.$$

Then,

$$\mathcal{D}^{(m)} = \bigcup \left\{\mathcal{D}_K^{(m)} : K \subseteq \mathbb{R}^n \text{ compact}\right\}.$$

The linear space $\mathcal{D}_K^{(m)}$ becomes a Banach space with the norm

$$\|\varphi\|_{m,K} = \sup_{|\alpha|\leq m} \sup_{x\in\mathbb{R}^n} |(\partial^\alpha \varphi)(x)|.$$

As in the case of $\mathcal{D}$, the space $\mathcal{D}^{(m)}$ becomes a topological linear space with the inductive limit topology of the spaces $\mathcal{D}_K^{(m)}$ for $K \subseteq \mathbb{R}^n$ compact. As usual, the topological dual space $\mathcal{D}^{(m)\prime}$ is given the dual or strong topology.

However, the spaces $\mathcal{D}^{(m)}$ and $\mathcal{D}^{(m)\prime}$ are not reflexives ([25], p. 75).

Lemma 75. *The space $\mathcal{D}$ is continuously and densely included in $\mathcal{D}^{(m)}$ for all $m = 0, 1, ...$*

Proof. That the inclusion is continuous, follows from Definition 74.

As for the density, given $\varphi \in \mathcal{D}^{(m)}$ let $\varphi_j = \varphi * \rho_j$, where ρ_j is the function used in the proof of Theorem 25.

It should be clear, from Lemma 24, that $\varphi_j \in \mathcal{D}$. Moreover, if $m \geq 1$, $\partial^\alpha \varphi_j = (\partial^\alpha \varphi) * \rho_j$, for $|\alpha| \leq m$. Furthermore, we can see, as in the proof of Theorem 28, that *supp* (φ_j) is contained in an open neighborhood of *supp* (φ) for all $j \geq 1$.

Thus, to show that the sequence $\{\varphi_j\}_{j\geq 1}$ converges to φ in $\mathcal{D}^{(m)}$, it suffices to show that $\varphi_j(x) \to \varphi(x)$ uniformly with respect to x in $\mathbb{R}^n$ as $j \to \infty$.

$$\begin{aligned} |\varphi_j(x) - \varphi(x)| &= \left| \int_{|y|\leq 1/j} \varphi(x-y)\,\rho_j(y)\,dy - \varphi(x) \int_{|y|\leq 1/j} \rho_j(y)\,dy \right| \\ &\leq \int_{|y|\leq 1/j} |\varphi(x-y) - \varphi(x)|\,\rho_j(y)\,dy. \end{aligned}$$

Since φ is uniformly continuous, because it is continuous and it has compact support, we can say that given $\varepsilon > 0$, there is $j_0 = j_0(\varepsilon) \geq 1$ so that

$$|\varphi(x-y) - \varphi(x)| \leq \varepsilon,$$

when $|y| \leq 1/j$ for $j \geq j_0$, for all $x \in \mathbb{R}^n$. So, $|\varphi_j(x) - \varphi(x)| \leq \varepsilon$ for all $x \in \mathbb{R}^n$.

This completes the proof of the lemma. □

Lemma 76. *For each $K \subseteq \mathbb{R}^n$ compact, the operator ∂^α is linear and continuous from $\mathcal{D}_K^{(m)}$ into $\mathcal{D}_K^{(m-|\alpha|)}$, for $m = 1, 2, ...,$ and $|\alpha| \leq m$.*

Proof. According to Lemma 54, it should be clear that the operator is well defined and it is linear.

As for the continuity, if we fix $K \subseteq \mathbb{R}^n$ compact,

$$\sup_{|\gamma|\leq m-|\alpha|} \sup_{x\in\mathbb{R}^n} \left| \left(\partial^{\alpha+\gamma}\varphi\right)(x) \right| \leq \sup_{|\gamma|\leq m} \sup_{x\in\mathbb{R}^n} |(\partial^\gamma \varphi)(x)|,$$

for all $\varphi \in \mathcal{D}_K^{(m)}$. Thus, the operator is continuous.

This completes the proof of the lemma. □

Theorem 77. *The space $\mathcal{D}^{(m)\prime}$ can be identified, as a linear space, with the linear subspace $\mathcal{D}'^{(m)}$ of $\mathcal{D}'$.*

Proof. According to Lemma 75, given $T \in \mathcal{D}^{(m)\prime}$, the restriction $r(T) = T\,|\mathcal{D}$ defines a linear map from $\mathcal{D}^{(m)\prime}$ into $\mathcal{D}'$. We will show that r is a bijection between $\mathcal{D}^{(m)\prime}$ and $\mathcal{D}'^{(m)}$.

To begin with, we claim that given $T \in \mathcal{D}^{(m)\prime}$, the linear map $T_1 = T\,|\mathcal{D}$ belongs to $\mathcal{D}'^{(m)}$. That is to say, T_1 satisfies 2) in Theorem 70. Indeed, if this is not true, as in the proof of Theorem 70, we could find a compact subset K of $\mathbb{R}^n$ and a sequence $\{\varphi_j\}_{j\geq 1}$ in $\mathcal{D}_K$ so that

$$\sup_{|\alpha|\leq m}\sup_{x\in\mathbb{R}^n}|(\partial^\alpha\varphi_j)(x)| < \frac{1}{j},$$

while $(T_1, \varphi_j) = 1$ for all $j \geq 1$, which is a contradiction. Thus, $T_1 \in \mathcal{D}'^{(m)}$.

Let us see that the map r is a bijection. That r is injective, follows from the dense inclusion of $\mathcal{D}$ into $\mathcal{D}^{(m)}$. In fact, given $\varphi \in \mathcal{D}^{(m)}$, there is sequence $\{\varphi_j\}_{j\geq 1}$ in $\mathcal{D}$ that converges to φ in $\mathcal{D}^{(m)}$ as $j \to \infty$. Then, if $T \in \mathcal{D}^{(m)\prime}$, $T(\varphi_j)$ converges to $T(\varphi)$ in $\mathbb{C}$ as $j \to \infty$. Moreover, if $T_1 = 0$, it means that $T(\varphi_j) = T_1(\varphi_j) = 0$ for all $j \geq 1$. Thus, $T(\varphi) = 0$. Since φ is arbitrary, we conclude that $T = 0$.

To prove that the map r is surjective, it suffices to show that it has a right inverse.

Let e be the extension map defined on $\mathcal{D}'^{(m)}$ as follows:

Given $T \in \mathcal{D}'^{(m)}$, by definition, T satisfies (4.12) with the same m for every compact subset K of $\mathbb{R}^n$.

Now, given $\varphi \in \mathcal{D}^{(m)}$, let $\{\varphi_j\}_{j\geq 1}$ be a sequence in $\mathcal{D}$ converging to φ in $\mathcal{D}^{(m)}$ as $j \to \infty$. Then, $\varphi_j, \varphi \in \mathcal{D}_K^{(m)}$ for some $K \subseteq \mathbb{R}^n$ compact and $\varphi_j \to \varphi$ in $\mathcal{D}_K^{(m)}$ as $j \to \infty$. It follows from (4.12) that the sequence $\{T(\varphi_j)\}_{j\geq 1}$ is a Cauchy sequence in $\mathbb{C}$, since

$$|T(\varphi_j) - T(\varphi_k)| \leq C_m \sup_{|\alpha|\leq m}\sup_{x\in\mathbb{R}^n}|(\partial^\alpha(\varphi_j - \varphi_k))(x)| \underset{j\to\infty}{\to} 0.$$

Let $e(T)(\varphi) = \lim_{j\to\infty} T(\varphi_j)$. We claim that this limit does not depend on the approximating sequence. Indeed, if $\{\psi_j\}_{j\geq 1}$ is another sequence in $\mathcal{D}$ converging to φ in $\mathcal{D}^{(m)}$ as $j \to \infty$,

$$|T(\varphi_j) - T(\psi_j)| \leq C_m \sup_{|\alpha|\leq m}\sup_{x\in\mathbb{R}^n}|(\partial^\alpha(\varphi_j - \psi_j))(x)| \underset{j\to\infty}{\to} 0.$$

Moreover, $e(T)$ belongs to $\mathcal{D}^{(m)\prime}$. In fact, it should be clear, by the definition, that $e(T)$ is a linear functional. As for the continuity, since

$$|T(\varphi_j)| \leq C_m \sup_{|\alpha|\leq m}\sup_{x\in\mathbb{R}^n}|\partial^\alpha(\varphi_j)(x)|,$$

taking the limit on both sides as $j \to \infty$, we get

$$|(e(T))(\varphi)| \leq C_{m,K} \sup_{|\alpha|\leq m}\sup_{x\in\mathbb{R}^n}|\partial^\alpha(\varphi)(x)|$$

for each $\varphi \in \mathcal{D}_K^{(m)}$. So, $e(T) \in \mathcal{D}^{(m)\prime}$.

Finally, let us see that $r \circ e$ is the identity on $\mathcal{D}'^{(m)}$.

Given $T \in \mathcal{D}'^{(m)}$ and $\varphi \in \mathcal{D}$, we take $\varphi_j = \varphi$ for all $j \geq 1$ as approximating sequence. Thus,

$$((r \circ e)(T))(\varphi) = r(T)(\varphi) = T(\varphi).$$

This completes the proof of the theorem. □

The following result is an immediate consequence of Theorem 77.

Corollary 78. *A distribution T has order zero exactly, when it can be extended to a linear and continuous functional from $\mathcal{D}^{(0)}$ into $\mathbb{C}$.*

Remark 79. According to Definition 74 and Theorem 70, a linear functional $T : \mathcal{D}^{(m)} \to \mathbb{C}$ is continuous exactly when, for each $K \subseteq \mathbb{R}^n$ compact, there is $C_{m,K} > 0$ so that

$$|T(\varphi)| \leq C_{m,K} \sup_{|\alpha| \leq m} \sup_{x \in \mathbb{R}^n} |\partial^\alpha(\varphi)(x)|,$$

for each $\varphi \in \mathcal{D}_K^{(m)}$.

Remark 80. As in Remark 67, let us observe that the bounded subsets of $\mathcal{D}$ are also bounded in $\mathcal{D}^{(m)}$. Therefore, if $\{T_j\}_{j\geq 1}$ converges to zero in $\mathcal{D}^{(m)\prime}$ as $j \to \infty$, the sequence $\{r(T_j)\}_{j\geq 1}$ converges to zero in $\mathcal{D}'$ as $j \to \infty$.

There is a very interesting connection between distributions of order zero and measures. Indeed, certain linear and continuous functionals on $\mathcal{D}^{(0)}$ can be represented as integrals with respect to certain measures. The result that guarantees this connection between functionals and measures is the Riesz-Markov-Kakutani theorem, named after the mathematicians Frigyes Riesz (1880-1956), Andrei Markov (1856-1922), and Shizuo Kakutani (1911-2004). Starting with Riesz's version, set on $\mathbb{R}$, Markov and Kakutani formulated increasingly general versions of the theorem, which is, in the end, true on a locally compact Hausdorff topological space. For our purposes, it will suffice to work on $\mathbb{R}^n$, and we will do so.

We begin with a few definitions.

Definition 81. *Let $(\mathbb{R}^n, \mathcal{B})$ be the Borel measurable space. A measure $\mu : \mathcal{B} \to [0, \infty]$ is called a Borel measure if $\mu(K)$ is finite for every $K \subseteq \mathbb{R}^n$ compact.*

Definition 82. *A Borel measure μ is regular if given $A \in \mathcal{B}$,*

$$\begin{aligned} \mu(A) &= \sup\{\mu(K) : K \subseteq A \ \textit{compact}\} \\ &= \inf\{\mu(U) : A \subseteq U \ \textit{open}\}. \end{aligned}$$

Remark 83. According to ([7], p. 206, Proposition 7.2.3.), a finite Borel measure is regular.

Definition 84. *Given $T \in \mathcal{D}^{(0)\prime}$, we say that T is positive if $T(\varphi)$ is real and non-negative for every $\varphi \in \mathcal{D}^{(0)}$ that is real and non-negative.*

Now, we state, on $\mathbb{R}^n$, the Riesz-Markov-Kakutani theorem.

Theorem 85. *Given $T \in \mathcal{D}^{(0)\prime}$ positive, there is a unique regular Borel measure μ so that*

$$T(\varphi) = \int_{\mathbb{R}^n} \varphi d\mu,$$

for every $\varphi \in \mathcal{D}^{(0)}$.

For a detailed and self-contained exposition of this theorem in its general version see, for instance, ([19]; [7], p. 209, Theorem 7.2.8).

The result published by Riesz in 1909, became the first of many Riesz representation theorems, that although could differ greatly from one to the other, all had as ultimate goal, to provide a concrete description of a topological dual. An excellent reference for the historical evolution of this matter is [9].

Example 86.

1. In general, the distribution T_f with $f \in L^1_{loc}$, is given, by definition, by the signed measure $f dx$. Thus, by extension, we can say that T_f is a signed measure that is absolutely continuous with respect to the Lebesgue measure. When f is real and non-negative, $f dx : \mathcal{B} \to [0, \infty]$ is a regular Borel measure.

2. The Dirac distribution δ_a is associated with the finite measure $\mu_a : \mathcal{B} \to [0, \infty)$ defined as

$$\mu_a(A) = \begin{cases} 1 & \text{if} \quad a \in A \\ 0 & \text{if} \quad a \notin A \end{cases}.$$

   This means that

$$\delta_a(\varphi) = \int_{\mathbb{R}^n} \varphi d\mu_a,$$

   According to Remark 83, μ_a is a regular Borel measure. The fact that δ_a is not defined by a locally integrable function, means that μ_a is not absolutely continuous with respect to the Lebesgue measure.

   In the sense of Theorem 85, we say that measures generalize the concept of function.

3. Let us consider the linear functional T defined on $\mathcal{D}^{(0)}(\mathbb{R})$ as

$$T(\varphi) = \mathcal{R} \int_{\mathbb{R}} \varphi(x)\, dx,$$

where $\mathcal{R}\int_{\mathbb{R}}$ is the Riemann integral. It should be clear that T is positive. Furthermore, we have the estimate

$$|T(\varphi)| \leq \sup_{x\in\mathbb{R}} |\varphi(x)|\, l(I),$$

where $l(I)$ is the length of an interval in $\mathbb{R}$ containing the support of φ. So, we conclude that $T \in \mathcal{D}^{(0)\prime}(\mathbb{R})$. As it happens, the measure associated with T is the Lebesgue measure on $\mathbb{R}$. Thus, Theorem 85 provides a way of constructing the Lebesgue measure without using outer measures.

Remark 87. The positivity assumption in Theorem 85 is necessary. Indeed, if we consider, on $\mathbb{R}$,

$$T(\varphi) = \mathcal{R}\int_0^\infty \varphi dx - \mathcal{R}\int_{-\infty}^0 \varphi dx,$$

we have that $T \in \mathcal{D}^{(0)\prime}(\mathbb{R})$. However, it should be clear that there is no measure μ associated with T in the sense of Theorem 85.

Based on Example 86, we state the following definition:

Definition 88. *The functionals in $\mathcal{D}^{(0)\prime}$ are called Radon measures.*

Although the use of the word measure in this definition might seem a little far fetched, Schwartz goes actually farther, stating that ([25], p. 17) "Today, it has become indispensable to define a measure μ as a linear and continuous functional on $\mathcal{D}^{(0)}$. It is from this functional $\mu(\varphi)$ that we will identify, when necessary, the countably additive function of sets $\mu(A)$." Let us add that this approach had been advocated by the Bourbaki school.

Remark 89. Definition 88 explains why the Dirac distribution is usually called Dirac measure.

We end this section with a density result.

Theorem 90. *The linear space $\mathbb{C}[x]$ of complex polynomial functions in $x = (x_1, ..., x_n)$ is dense in $\mathcal{E}$.*

The proof of this theorem necessitates some preliminary work, which we take up now.

The function

$$W(x,t) = \frac{1}{(4\pi ct)^{n/2}} e^{-|x|^2/4ct},$$

where c is a positive real number, is a solution of the heat equation, or diffusion equation,

$$(\partial_t u)(x,t) - c(\Delta u)(x,t) = 0,$$

for $(x,t) \in \mathbb{R}^n \times \mathbb{R}_+$. Moreover, it satisfies the distributional initial condition

$$\lim_{t \to 0^+} \int_{\mathbb{R}^n} W(x,t)\,\varphi(x)\,dx = \varphi(0), \tag{4.14}$$

for each $\varphi \in \mathcal{D}^{(0)}$. The first statement can be verified by a straightforward calculation. The second statement uses the properties listed in the lemma that follows.

Lemma 91. *The following statements are true:*

1. $W(x,t) > 0$ *for every* $(x,t) \in \mathbb{R}^n \times \mathbb{R}_+$.
2. $\int_{\mathbb{R}^n} W(x,t)\,dx = 1$ *for all* $t > 0$.
3. *For each* $\delta > 0$ *there is*

$$\lim_{t \to 0^+} \int_{|x|>\delta} W(x,t)\,dx = 0.$$

Proof. It is fairly obvious that 1) is true. As for 2), let us observe that, according to Example 49, the function $W(\cdot,t) \in \mathcal{S}$ for each $t > 0$. Thus, the integral exists, according to Theorem 57.

Since

$$\int_{\mathbb{R}^n} W(x,t)\,\varphi(x)\,dx = \prod_{l=1}^{n} \frac{1}{(4\pi ct)^{1/2}} \int_{\mathbb{R}} e^{-x_l^2/4ct}\,dx_l,$$

to prove 2) it will suffice to show that

$$\frac{1}{(4\pi ct)^{1/2}} \int_{\mathbb{R}} e^{-x^2/4ct}\,dx = 1, \tag{4.15}$$

which we will do by a well known argument, that uses polar coordinates in $\mathbb{R}^2$. In fact, if we denote I the integral in (4.15),

$$\begin{aligned} I^2 &= \left(\frac{1}{(4\pi ct)^{1/2}} \int_{\mathbb{R}} e^{-x^2/4ct}\,dx\right)\left(\frac{1}{(4\pi ct)^{1/2}} \int_{\mathbb{R}} e^{-y^2/4ct}\,dy\right) \\ &= \frac{2\pi}{4\pi ct} \int_0^\infty r e^{-r^2/4ct}\,dr = \frac{1}{2ct} \int_0^\infty (-2ct)\,\frac{d}{dr} e^{-r^2/4ct} = 1. \end{aligned}$$

Since $I > 0$, we conclude that $I = 1$.

Finally, let us prove 3). By the change of variables $x/\sqrt{t} = y$, we can write

$$\begin{aligned} \frac{1}{(4\pi ct)^{n/2}} \int_{|x|>\delta} e^{-|x|^2/4ct}\,dx &= \frac{1}{(4\pi c)^{n/2}} \int_{|x|>\delta/\sqrt{t}} e^{-|x|^2/4c}\,dx \\ &= 1 - \frac{1}{(4\pi c)^{n/2}} \int_{|x|\le\delta/\sqrt{t}} e^{-|x|^2/4ct}\,dx. \end{aligned} \tag{4.16}$$

Let $\{t_k\}_{k\geq 1}$ be a non-increasing sequence in $\mathbb{R}_+$ with $t_k \to 0$ as $k \to \infty$ and let χ_k be the characteristic function of $\left\{y \in \mathbb{R} : |y| \leq \delta/\sqrt{t_k}\right\}$. Then, we can apply the monotone convergence theorem to the sequence

$$f_k(x) = \chi_k(x)\frac{1}{(4\pi c)^{n/2}}e^{-|x|^2/4ct}.$$

Therefore, (4.16) goes to zero as $t \to 0^+$.

This completes the proof of the lemma. □

Remark 92. The distributional initial condition (4.14) can be verified as follows: Using 1) and 2) in Lemma 91,

$$\int_{\mathbb{R}^n} W(x,t)\,\varphi(x)\,dx - \varphi(0) = \int_{\mathbb{R}^n} W(x,t)\,(\varphi(x) - \varphi(0))\,dx.$$

By continuity, given $\varepsilon > 0$ there is $\delta_\varepsilon > 0$, so that $|x| \leq \delta_\varepsilon$ implies $|\varphi(x) - \varphi(0)| \leq \varepsilon$. Hence,

$$\left|\int_{\mathbb{R}^n} W(x,t)\,\varphi(x)\,dx - \varphi(0)\right| \leq 2\sup_{x\in\mathbb{R}^n}|\varphi(x)|\int_{|x|>\delta_\varepsilon} W(x,t)\,dx + \varepsilon,$$

for all $t \in \mathbb{R}^+$. Using 3) in Lemma 91, we can write

$$\left|\int_{\mathbb{R}^n} W(x,t)\,\varphi(x)\,dx - \varphi(0)\right| \leq \left(2\sup_{x\in\mathbb{R}^n}|\varphi(x)| + 1\right)\varepsilon,$$

for $t \geq t_\varepsilon$.

Remark 93. The function $W(x,t)$ defines on $\mathcal{D}^{(0)}$ a linear integral operator by means of the convolution

$$\int_{\mathbb{R}^n} W(x-y,t)\,\varphi(y)\,dy.$$

The function $W(x-y,t)$ is the kernel of the integral operator and it is called the Gauss-Weierstrass kernel, or the heat kernel. It is name after the mathematicians Carl Friedrich Gauss (1777-1855) and Karl Weierstrass (1815-1897). Since we will use $W(x,t)$ as an approximation tool, we assume that the parameter c is one. Moreover, we take $t = 1/j$ for $j \geq 1$. Then, the function $W(x,t)$ becomes

$$W_j(x) = \left(\frac{j}{4\pi}\right)^{n/2} e^{-j|x|^2/4}.$$

We are now ready to prove Theorem 90.

Proof. We fix $\varphi \in \mathcal{E}$, $K \subseteq \mathbb{R}^n$ compact and $m = 0, 1, \ldots$. We select a function $\alpha \in \mathcal{D}$ such that $0 \leq \alpha(x) \leq 1$ for all $x \in \mathbb{R}^n$, $\alpha(x) = 1$ for x in the $\theta/2$-neighborhood of K for some $0 < \theta < 1$, and $supp\,(\alpha) \subseteq \theta$-neighborhood of K. Let $\psi = \alpha\varphi$ and let us fix $\beta \in \mathbb{N}^n$ with $|\beta| \leq m$.

$$\left|\left(W_j * \partial^\beta \psi\right)(x) - \left(\partial^\beta \psi\right)(x)\right| \leq \int_{\mathbb{R}^n} W_j(y) \left|\left(\partial^\beta \psi\right)(x-y) - \left(\partial^\beta \psi\right)(x)\right| dy.$$

Since $\partial^\beta \psi$ is continuous on $\mathbb{R}^n$ and it has compact support, it is uniformly continuous. Thus, for each $\varepsilon > 0$ there is $\delta_{\varepsilon,\beta} > 0$ depending on ε and β, so that $\left|\left(\partial^\beta \psi\right)(x-y) - \left(\partial^\beta \psi\right)(x)\right| \leq \varepsilon$ for $|y| \leq \delta_{\varepsilon,\beta}$.

By taking $\delta_{\varepsilon,m} = \inf\{\delta_{\varepsilon,\beta} : |\beta| \leq m\}$ and using Lemma 91, we have

$$\begin{aligned}\left|\left(W_j * \partial^\beta \psi\right)(x) - \left(\partial^\beta \psi\right)(x)\right| &\leq \varepsilon + 2 \sup_{x \in \mathbb{R}^n} \left|\left(\partial^\beta \psi\right)(x)\right| \int_{|y| > \delta_{\varepsilon,m}} W_j(y)\, dy \\ &\leq \left(1 + 2 \sup_{x \in \mathbb{R}^n} \left|\left(\partial^\beta \psi\right)(x)\right|\right) \varepsilon\end{aligned}$$

if $j \geq j_{\varepsilon,m}$.

Let us observe that $\left(\partial^\beta \psi\right)(x) = \left(\partial^\beta \varphi\right)(x)$ when $x \in \theta/2$-neighborhood(K). So, we can write

$$\sup_{|\beta| \leq m} \sup_{x \in K} \left|\left(W_j * \partial^\beta \psi\right)(x) - \left(\partial^\beta \varphi\right)(x)\right| \leq \left(1 + 2 \sup_{|\beta| \leq m} \sup_{x \in \mathbb{R}^n} \left|\left(\partial^\beta \psi\right)(x)\right|\right) \varepsilon,$$

for $j \geq j_{\varepsilon,m}$. Let us fix $j_0 \geq j_{\varepsilon,m}$.

The kernel $W_{j_0}(x-y)$ can be represented by the series

$$W_{j_0}(x-y) = \left(\frac{j_0}{4\pi}\right)^{n/2} \sum_{k \geq 0} \frac{\left(-\frac{j_0}{4}|x-y|^2\right)^k}{k!}$$

that converges absolutely for every $x, y \in \mathbb{R}^n$.

Let K_θ be the closure of the θ-neighborhood of K. If $x \in K$ and $y \in K_\theta$, then $x - y$ belongs to a compact subset K' of $\mathbb{R}^n$. Therefore, the series converges absolutely and uniformly on $K \times K_\theta$, as well as each of its term-by-term partial derivatives. So, given $m \in \mathbb{N}$, there is $N_{j_0,m} \geq 1$ so that

$$\left|\left(\partial^\beta W_{j_0}\right)(x-y) - \left(\frac{j_0}{4\pi}\right)^{n/2} \sum_{k=0}^{N} \frac{\left(-\frac{j_0}{4}\right)^k}{k!} \partial_x^\beta \left(|x-y|^{2k}\right)\right| \leq \varepsilon$$

for $|\beta| \leq m$ and $N \geq N_{j_0,m}$. Since

$$\left(\partial^\beta W_{j_0}\right) * \psi = W_{j_0} * \left(\partial^\beta \psi\right)$$

$$
\begin{aligned}
&\left| \left(W_{j_0} * \partial^{\beta}\psi\right)(x) - \int_{K_\theta} \left(\frac{j_0}{4\pi}\right)^{n/2} \sum_{k=0}^{N} \frac{\left(-\frac{j_0}{4}\right)^k}{k!} \partial_x^{\beta}\left(|x-y|^{2k}\right)\psi(y)\,dy \right| \\
\leq\ & \int_{K_\theta} \left| W_{j_0}(x-y) - \left(\frac{j_0}{4\pi}\right)^{n/2} \sum_{k=0}^{N} \frac{\left(-\frac{j_0}{4}\right)^k}{k!} \partial_x^{\beta}\left(|x-y|^{2k}\right) \right| |\psi(y)|\,dy \\
\leq\ & meas\,(K_\theta)\left(\sup_{x\in\mathbb{R}^n}|\psi(x)|\right)\varepsilon,
\end{aligned}
$$

for $N > N_{j_0,m}$. We fix $N \geq \max(N_{j_0,m}, m)$

According to Lemma 1,

$$
|x-y|^{2k} = \sum_{|\gamma|=k} \frac{k!}{\gamma_1!...\gamma_n!}(x-y)^{2\gamma} = \sum_{|\gamma|=k} \frac{k!}{\gamma_1!...\gamma_n!} \prod_{s=1}^{n} \sum_{l_s=0}^{2\gamma_s} \binom{2\gamma_s}{l_s} x_s^{l_s} y_s^{2\gamma_s - l_s}.
$$

Therefore,

$$
\begin{aligned}
&\int_{K_\theta} \left(\frac{j_0}{4\pi}\right)^{n/2} \sum_{k=0}^{N} \frac{\left(-\frac{j_0}{4}\right)^k}{k!} \partial_x^{\beta}\left(|x-y|^{2k}\right)\psi(y)\,dy \\
=\ & \left(\frac{j_0}{4\pi}\right)^{n/2} \sum_{k=0}^{N} \frac{-\left(\frac{j_0}{4}\right)^k}{k!} \sum_{|\gamma|=k} \frac{k!}{\gamma_1!...\gamma_n!} \prod_{s=1}^{n} \sum_{l_s=0}^{2\gamma_s} \binom{2\gamma_s}{l_s} \left(\int_{K_\theta} y_s^{2\gamma_s - l_s}\psi(y)\,dy\right) \partial^{\beta}\left(x_s^{l_s}\right),
\end{aligned}
$$

which is a polynomial in x of degree $2N - m$.

This completes the proof of the theorem. □

5 The derivative of a distribution

Given $T \in \mathcal{D}'$, we want to define n distributions, called the n partial derivatives of T with respect to the variables $x_1, ..., x_n$. We will do it in a way that extends to distributions the notion of partial derivative of a function. In other words, if $T = T_f$ for f continuous with continuous partial derivatives, the distribution j-th partial derivative of T will be $T_{\partial_{x_j} f}$.

So, let us begin with this case and let us assume, to simplify the notation, that $j = 1$. Then, given $\varphi \in \mathcal{D}$,

$$
\begin{aligned}
\left(T_{\partial_{x_1} f}, \varphi\right) &= \int_{\mathbb{R}^n} (\partial_{x_1} f)(x)\,\varphi(x)\,dx \\
& \int_{\mathbb{R}^{n-1}} \left(\int_{-\infty}^{\infty} (\partial_{x_1} f)(x_1, x')\,\varphi(x_1, x')\,dx_1\right) dx', \qquad (5.1)
\end{aligned}
$$

where we have written, as usual, $x = (x_1, x')$. Let us observe that for x' fixed, the function $\varphi(x_1, x') = 0$ for x_1 outside a finite interval. Thus, using integration by parts on x_1, the integral in (5.1) becomes

$$-\int_{\mathbb{R}^n} f(x)\left(\partial_{x_1}\varphi\right)(x)\,dx.$$

So, we can write

$$\left(T_{\partial_{x_1} f}, \varphi\right) = -\left(T_f, \partial_{x_1}\varphi\right).$$

If instead of T_f we consider a distribution $T \in \mathcal{D}'$, the functional

$$\begin{array}{ccc} \mathcal{D} & \to & \mathbb{C} \\ \varphi & \to & -\left(T, \partial_{x_1}\varphi\right) \end{array} \tag{5.2}$$

is well defined and it is linear. Moreover, if $\{\varphi_j\}_{j\geq 1}$ converges to zero in $\mathcal{D}$ as $j \to \infty$, it should be clear that $\{\partial_{x_1}\varphi_j\}_{j\geq 1}$ also converges to zero in $\mathcal{D}$ as $j \to \infty$. Thus, $\left(T_f, \partial_{x_1}\varphi_j\right) \to 0$ in $\mathbb{C}$ as $j \to \infty$. Hence, the functional (5.2) belongs to $\mathcal{D}'$. In general,

Definition 94. *Given $T \in \mathcal{D}'$ the distribution*

$$\begin{array}{ccc} \mathcal{D} & \to & \mathbb{C} \\ \varphi & \to & -\left(T, \partial_{x_j}\varphi\right) \end{array}$$

is called the j-th partial derivative of T and it is denoted $\partial_{x_j} T$, for $1 \leq j \leq n$.

Remark 95. The process of computing partial derivatives of a distribution can go on indefinitely. Thus, in this sense, distributions are smooth. Moreover, since given $\varphi \in \mathcal{D}$,

$$\partial_{x_k}\left(\partial_{x_j}\varphi\right) = \partial_{x_j}\left(\partial_{x_k}\varphi\right)$$

for $1 \leq j, k \leq n$, it follows that the result of differentiating a distribution repeatedly, does not depend on the order in which we take the partial derivatives. Therefore, given $\alpha \in \mathbb{N}^n$, we can write

$$\left(\partial^\alpha T, \varphi\right) = (-1)^{|\alpha|}\left(T, \varphi\right),$$

for all $\varphi \in \mathcal{D}$.

So, we can see the usefulness of the notation ∂^α, when differentiating a distribution.

Let us observe that given a continuous function f with enough continuous partial derivatives, we have, integrating by parts,

$$\partial^\alpha T_f = T_{\partial^\alpha f},$$

where, on the left-hand side, ∂^α indicates the partial derivative in the sense of distributions, while, on the right-hand side, ∂^α indicates the partial derivative in the classical sense.

Example 96. 1. The distribution T_a^m defined in 3) of Example 23 is equal to $(-1)^m \, \partial_{x_1}^m \delta_a$.

2. Given the locally integrable function $\ln|x|$ defined on $\mathbb{R}$, we claim that

$$\frac{d}{dx} T_{\ln|x|} = pv\frac{1}{x}. \tag{5.3}$$

Indeed, given $\varphi \in \mathcal{D}(\mathbb{R})$,

$$\left(\frac{d}{dx} T_{\ln|x|}, \varphi\right) = -\int_{\mathbb{R}} (\ln|x|)\, \varphi'(x)\, dx = -\lim_{j\to\infty} \int_{|x|>1/j} (\ln|x|)\, \varphi'(x)\, dx.$$

For $j \geq 1$ fixed, we write

$$\int_{|x|>1/j} (\ln|x|)\, \varphi'(x)\, dx = \underbrace{\int_{-\infty}^{-1/j} (\ln(-x))\, \varphi'(x)\, dx}_{(1)} + \underbrace{\int_{1/j}^{\infty} (\ln x)\, \varphi'(x)\, dx}_{(2)}$$

and we integrate by parts. Then,

$$(1) = \ln\left(\frac{1}{j}\right) \varphi\left(-\frac{1}{j}\right) - \int_{-\infty}^{-1/j} \frac{\varphi(x)}{x} dx,$$

while

$$(2) = -\ln\left(\frac{1}{j}\right) \varphi\left(\frac{1}{j}\right) - \int_{1/j}^{\infty} \frac{\varphi(x)}{x} dx.$$

Thus,

$$\begin{aligned} \int_{|x|>1/j} (\ln|x|)\, \varphi'(x)\, dx &= -\ln\left(\frac{1}{j}\right)\left(\varphi\left(\frac{1}{j}\right) - \varphi\left(-\frac{1}{j}\right)\right) - \int_{|x|>1/j} \frac{\varphi(x)}{x} dx \\ &= -\frac{2}{j} \ln\left(\frac{1}{j}\right) \varphi'(\theta_j) - \int_{|x|>1/j} \frac{\varphi(x)}{x} dx, \end{aligned}$$

for some θ_j between $1/j$ and $-1/j$.

So,

$$\left(\frac{d}{dx} T_{\ln|x|}, \varphi\right) = \lim_{j\to\infty} \int_{|x|>1/j} \frac{\varphi(x)}{x} dx = \left(pv\frac{1}{x}, \varphi\right).$$

3. Next, we claim that

$$\frac{d}{dx}\left(pv\frac{1}{x}\right) = -fp\frac{1}{x^2}. \tag{5.4}$$

In fact, given $\varphi \in \mathcal{D}(\mathbb{R})$,

$$\left(\frac{d}{dx} \left(pv\frac{1}{x} \right), \varphi \right) = -\lim_{j\to\infty} \int_{|x|>1/j} \frac{\varphi'(x)}{x} dx.$$

For $j \geq 1$ fixed, we write

$$\int_{|x|>1/j} \frac{\varphi'(x)}{x} dx = \underbrace{\int_{-\infty}^{-1/j} \frac{\varphi'(x)}{x} dx}_{(1)} + \underbrace{\int_{1/j}^{\infty} \frac{\varphi'(x)}{x} dx}_{(2)}$$

and we integrate by parts. Then,

$$(1) = -j\varphi\left(-\frac{1}{j}\right) + \int_{-\infty}^{-1/j} \frac{\varphi(x)}{x^2} dx$$

and

$$(2) = -j\varphi\left(\frac{1}{j}\right) + \int_{1/j}^{\infty} \frac{\varphi(x)}{x^2} dx.$$

Therefore,

$$(1) + (2) = -j\left(\varphi\left(\frac{1}{j}\right) + \varphi\left(-\frac{1}{j}\right)\right) + \int_{r>|x|>1/j} \frac{\varphi(x)}{x^2} dx,$$

where $r \geq 1$ is such that $supp(\varphi) \subseteq \{x \in \mathbb{R} : |x| < r\}$.

Following Example 43, we write

$$\begin{aligned} \int_{r>|x|>1/j} \frac{\varphi(x)}{x^2} dx &= \int_{r>|x|>1/j} \frac{\varphi(x) - \varphi(0) - \varphi'(0)x}{x^2} dx \\ &+\varphi(0) \int_{r>|x|>1/j} \frac{dx}{x^2} + \varphi'(0) \int_{r>|x|>1/j} \frac{dx}{x}. \end{aligned}$$

As in Example 43,

$$\int_{r>|x|>1/j} \frac{dx}{x} = 0,$$

while

$$\varphi(0) \int_{r>|x|>1/j} \frac{dx}{x^2} = \varphi(0)\left(\frac{1}{r} + j - \frac{1}{r} + j\right) = 2j\varphi(0).$$

Then,

$$\begin{aligned} \int_{|x|>1/j} \frac{\varphi'(x)}{x} dx &= -j\left(\varphi\left(\frac{1}{j}\right) - \varphi(0)\right) - j\left(\varphi\left(-\frac{1}{j}\right) - \varphi(0)\right) \\ &+ \int_{r>|x|>1/j} \frac{\varphi(x) - \varphi(0) - \varphi'(0)x}{x^2} dx \\ &= -j\frac{1}{j}\varphi'(\theta_j) - j\left(-\frac{1}{j}\right)\varphi'(\eta_j) + \int_{r>|x|>1/j} \frac{\varphi(x) - \varphi(0) - \varphi'(0)x}{x^2} dx, \end{aligned}$$

where $0 < \theta_j, \eta_j < 1/j$.

So, taking the limit on both sides as $j \to \infty$ we get, according to Example 43

$$\left(pv\frac{1}{x}, \varphi' \right) = -\varphi'(0) + \varphi'(0) + \left(fp\frac{1}{x^2}, \varphi \right).$$

Thus, we have proved (5.4).

4. Let us consider the function $|x|$ defined on $\mathbb{R}$. Integrating by parts,

$$\begin{aligned}\left(\frac{d}{dx} T_{|x|}, \varphi \right) &= \int_{-\infty}^{0} x\varphi'(x)\, dx - \int_{0}^{\infty} x\varphi'(x)\, dx \\ &= -\int_{-\infty}^{0} \varphi(x)\, dx + \int_{0}^{\infty} \varphi(x)\, dx.\end{aligned}$$

That is to say,

$$\frac{d}{dx} T_{|x|} = T_{sgn(x)}, \tag{5.5}$$

where sgn denotes the sign function, defined as

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

Let us observe that the choice of value at $x = 0$ does not matter. Moreover, the derivative, in the distributional sense, of $T_{|x|}$, is the distribution associated with the classical derivative of $|x|$, which exists for $x \neq 0$.

5. Next, we calculate $\frac{d}{dx} T_{sgn(x)}$. Since

$$\int_{-\infty}^{0} \varphi'(x)\, dx - \int_{0}^{\infty} \varphi'(x)\, dx = 2\varphi(0),$$

we have

$$\frac{d}{dx} T_{sgn(x)} = 2\delta_0. \tag{5.6}$$

That is to say, the jump at zero of the function sgn results in a derivative equal to a multiple of the Dirac measure δ_0, where the factor is equal to the height of the jump.

A similar behavior can be observed in the next example.

6. Let H_n be the n-dimensional Heaviside function,

$$H_n(x) = \begin{cases} 1 & \text{if } x_1, ..., x_n \geq 0 \\ 0 & \text{otherwise} \end{cases}.$$

Then,

$$\partial_{x_1} ... \partial_{x_n} T_{H_n} = \delta_0. \tag{5.7}$$

Indeed, given $\varphi \in \mathcal{D}$, if we integrate on x_1, then x_2, etc., we have

$$(-1)^n \int_0^\infty .... \int_0^\infty (\partial_{x_1} ... \partial_{x_n} \varphi)(x)\, dx_1 ... dx_n = \varphi(0).$$

7. Given $k = 1, 2, ...$, we calculate $d_{-1}\left(\delta_0^{(k)}\right)$, where δ_0 is the Dirac distribution concentrated on $\{0\} \subseteq \mathbb{R}$ and $\delta_0^{(k)}$ is the derivative of δ_0, of order k.

$$\begin{aligned} \left(d_{-1}\left(\delta_0^{(k)}\right), \varphi\right) &= (-1)^k (\varphi(-x))^{(k)}(0) \\ &= \varphi^{(k)}(0) = \left(\delta_0, \varphi^{(k)}\right) \\ &= (-1)^k \left(\delta_0^{(k)}, \varphi\right). \end{aligned}$$

Or,

$$d_{-1}\left(\delta_0^{(k)}\right) = (-1)^k \delta_0^{(k)}.$$

Therefore, the distribution $\delta_0^{(k)}$ is odd when k is odd, and it is even when k is even.

8. In general, if T is an odd distribution in $\mathcal{D}'$, $\partial_{x_j} T$ is even. If T is an even distribution in $\mathcal{D}'$, $\partial_{x_j} T$ is odd.

In fact, let us assume that T is odd. Given $\varphi \in \mathcal{D}$,

$$\begin{aligned} \left(d_{-1}\left(\partial_{x_j} T\right), \varphi\right) &= -\left(T, \partial_{x_j}(\varphi(-\cdot))\right) = \left(T, \left(\partial_{x_j}\varphi\right)(-\cdot)\right) \\ &= \left(d_{-1}T, \partial_{x_j}\varphi\right) \underset{(1)}{=} -\left(T, \partial_{x_j}\varphi\right) \underset{(2)}{=} \left(\partial_{x_j} T, \varphi\right). \end{aligned}$$

When T is even, we only need to move the negative sign from the right-hand side of (1) to the right-hand side of (2), to conclude that $\partial_{x_j} T$ is odd.

Remark 97. It should be clear that $T_{\ln|x|}$ and $T_{|x|}$ are even distributions, while $T_{sgn x}$ is odd.

According to 2) in Remark 44, the distribution $vp\, 1/x$, is odd, while the distribution $fp\, 1/x^2$is even.

Therefore, (5.3), (5.4), (5.5), and (5.6), are all illustrations of 8) in Example 96.

Theorem 98. *If $\mathcal{A}$ denotes one of the spaces $\mathcal{D}'$, $\mathcal{E}'$, or $\mathcal{S}'$, the operator ∂^α is linear and continuous from $\mathcal{A}$ into itself. Moreover, given $T \in \mathcal{A}$,*

$$supp\left(\partial^\alpha T\right) \subseteq supp\left(T\right).$$

Proof. For instance, we will work in the case $\mathcal{A} = \mathcal{S}'$, the other cases being similar. According to Lemma 54, it should be clear that ∂^α is well defined and linear from $\mathcal{S}'$ into $\mathcal{S}'$. So, given $T \in \mathcal{S}'$ and given a bounded subset $\mathcal{B}_1$ of $\mathcal{S}$,

$$\begin{aligned} \left\|\partial^\alpha T\right\|_{\mathcal{S}',\mathcal{B}_1} &= \sup_{\varphi\in\mathcal{B}_1}\left|\left(\partial^\alpha T\right)(\varphi)\right| = \sup_{\varphi\in\mathcal{B}_1}\left|T\left(\partial^\alpha\varphi\right)\right| \\ &\leq \sup_{\varphi\in\mathcal{B}_2}\left|T\left(\varphi\right)\right| = \left\|\partial^\alpha T\right\|_{\mathcal{S}',\mathcal{B}_2}, \end{aligned}$$

because, according to Lemma 54, the operator ∂^α maps bounded subsets of $\mathcal{S}$ into bounded subsets of $\mathcal{S}$. Thus, the continuity is proved.

As for the support of $\partial^\alpha T$, if $x_0 \notin supp\left(T\right)$, by definition, there is an open neighborhood V of x_0 such that $T\left(\varphi\right) = 0$ if $supp\left(\varphi\right) \subseteq V$. According to Lemma 54, $supp\left(\partial^\alpha\varphi\right) \subseteq V$, so

$$T\left(\partial^\alpha\varphi\right) = 0 = (-1)^{|\alpha|}\left(\partial^\alpha T\right)(\varphi).$$

Therefore, $x_0 \notin supp\left(\partial^\alpha T\right)$.

This completes the proof of the lemma. □

The following result is an immediate consequence of Theorem 98.

Corollary 99. *If $\mathcal{A}$ denotes one of the spaces $\mathcal{D}'$, $\mathcal{E}'$, or $\mathcal{S}'$, a convergent sequence $\{T_j\}_{j\geq1} \subseteq \mathcal{A}$ can be differentiated term by term without altering its convergence.*

Theorem 100. *The operator ∂^α is linear and continuous from $\mathcal{D}^{(m)\prime}$ into $\mathcal{D}^{(m+|\alpha|)\prime}$, for $m = 1, 2, ...,$ and $|\alpha| \leq m$.*

Proof. By Lemma 76, it should be clear that the operator is well defined and linear from $\mathcal{D}^{(m)\prime}$ into $\mathcal{D}^{(m+|\alpha|)\prime}$. If $T \in \mathcal{D}^{(m)\prime}$ and φ belongs to a bounded subset $\mathcal{B}_1$ of $\mathcal{D}_K^{(m+|\alpha|)}$,

$$\begin{aligned} \left\|\partial^\alpha T\right\|_{\mathcal{D}^{(m+|\alpha|)\prime},\mathcal{B}_1} &= \sup_{\varphi\in\mathcal{B}_1}\left|\left(\partial^\alpha T,\varphi\right)_{\mathcal{D}^{(m+|\alpha|)\prime},\mathcal{D}^{(m+|\alpha|)}}\right| \\ &= \sup_{\varphi\in\mathcal{B}_1}\left|\left(T,\partial^\alpha\varphi\right)_{\mathcal{D}^{(m)\prime},\mathcal{D}^{(m)}}\right|. \end{aligned}$$

Lemma 76 tells us that $\partial^\alpha\varphi$ belongs to a bounded subset $\mathcal{B}_2$ of $\mathcal{D}^{(m)}$ when φ belongs to a bounded subset $\mathcal{B}_1$ of $\mathcal{D}_K^{(m+|\alpha|)}$. Thus,

$$\sup_{\varphi\in\mathcal{B}_1}\left|\left(T,\partial^\alpha\varphi\right)_{\mathcal{D}^{(m)\prime},\mathcal{D}^{(m)}}\right| \leq \sup_{\psi\in\mathcal{B}_2}\left|\left(T,\psi\right)_{\mathcal{D}^{(m)\prime},\mathcal{D}^{(m)}}\right| = \left\|T\right\|_{\mathcal{D}^{(m)\prime},\mathcal{B}_2}.$$

This completes the proof of the theorem. □

That a function f has a partial derivative with respect to, say x_1, at a point x, means that there is

$$\lim_{h_1\to 0}\frac{f(x_1+h_1,x')-f(x)}{h_1}=(\partial_{x_1}f)(x).$$

Now, if $h=(h_1,0)$ and τ_{-h} is the translation operator

$$\tau_{-h}(f)(x)=f(x+h),$$

we can write

$$\lim_{h\to 0}\frac{\tau_{-h}(f)(x)-f(x)}{h_1}=(\partial_{x_1}f)(x).$$

With the obvious meaning, we have, in general,

$$\lim_{h\to 0}\frac{\tau_{-h}(f)(x)-f(x)}{h_j}=\left(\partial_{x_j}f\right)(x).$$

It is possible to obtain the same result for a distribution, if we extend the above notation. In fact, it should be clear that the operator τ_h defined as

$$\tau_h(\varphi)(x)=\varphi(x-h)$$

is linear and continuous from $\mathcal{D}$ into itself. Thus, we can define the translation operator τ_{-h} on $\mathcal{D}'$ as

$$(\tau_{-h}(T),\varphi)=(T,\tau_h(\varphi)), \tag{5.8}$$

which results from a simple change of variables when T is defined by a locally integrable function. Since τ_h maps bounded subsets of $\mathcal{D}$ into bounded subsets of $\mathcal{D}$, the operator τ_{-h} is continuous from $\mathcal{D}'$ into $\mathcal{D}'$.

Theorem 101. *There is*

$$\lim_{h\to 0}\frac{\tau_{-h}(T)-T}{h_j}=\partial_{x_j}T,$$

where the limit is taken with respect to the strong topology of $\mathcal{D}'$.

Proof. For $T\in\mathcal{D}'$, let $\mathcal{B}'$ be the subset of $\mathcal{D}'$ defined as

$$\left\{\frac{\tau_{-h}(T)-T}{h_j}\right\}_{0<|h_j|<1}. \tag{5.9}$$

Now, for $\varphi\in\mathcal{D}$ fixed and $0<|h_j|<1$, we write

$$\left(\frac{\tau_{-h}(T)-T}{h_j},\varphi\right)=\left(T,\frac{\tau_h(\varphi)-\varphi}{h_j}\right). \tag{5.10}$$

Let us observe that $\frac{\tau_h(\varphi)-\varphi}{h_j} \in \mathcal{D}_K$ for some $K \subseteq \mathbb{R}^n$ compact independent of h. Moreover,

$$\begin{aligned} \partial^{\alpha} \frac{\tau_h(\varphi) - \varphi}{h_j}(x) &= \frac{\tau_h(\partial^{\alpha}\varphi)(x) - \partial^{\alpha}\varphi(x)}{h_j} \\ &= -\left(\partial_{x_j}\partial^{\alpha}\varphi\right)(\theta), \end{aligned}$$

for some θ between $x-h$ and x. So,

$$\sup_{0<|h_j|<1} \sup_{x\in\mathbb{R}^n} \left|\partial_{x_j}\partial^{\alpha}\varphi(x)\right| < \infty.$$

Thus, according to Remark 22, the subset of $\mathcal{D}$

$$\left\{\frac{\tau_h(\varphi)-\varphi}{h_j}\right\}_{0<|h_j|<1}$$

is bounded in $\mathcal{D}$. Therefore,

$$\sup_{0<|h_j|<1} \left|\left(\frac{\tau_{-h}(T)-T}{h_j}, \varphi\right)\right| = \sup_{0<|h_j|<1} \left|\left(T, \frac{\tau_h(\varphi)-\varphi}{h_j}\right)\right| < \infty.$$

According, again, to Remark 22, $\mathcal{B}'$ is a bounded subset of $\mathcal{D}'$. Therefore, the strong and the weak topology coincide on $\mathcal{B}'$. Thus, we are left to show that

$$\left(\frac{\tau_{-h}(T)-T}{h_j}, \varphi\right) \underset{h\to 0}{\to} \left(\partial_{x_j}T, \varphi\right).$$

for each $\varphi \in \mathcal{D}$.

Since $\partial^{\alpha}\varphi$ has the same properties as φ, it will be enough to prove that

$$\frac{\tau_h(\varphi)-\varphi}{h_j}(x) \underset{h\to 0}{\to} -\partial_{x_j}\varphi(x)$$

in $\mathbb{C}$, uniformly with respect to $x \in \mathbb{R}^n$.

$$\begin{aligned} \frac{\tau_{-h}(\varphi)-\varphi}{h_j}(x) + \partial_{x_j}\varphi(x) &= -\left(\partial_{x_j}\varphi\right)(\theta) + \partial_{x_j}\varphi(x) \\ &= \left(\partial^2_{x_j}\varphi\right)(\xi)\,|x-\xi|, \end{aligned}$$

for some ξ between x and θ. Finally,

$$\sup_{x\in\mathbb{R}^n} \left|\frac{\tau_{-h}(\varphi)-\varphi}{h_j}(x) + \partial_{x_j}\varphi(x)\right| \leq \sup_{x\in\mathbb{R}^n} \left|\left(\partial^2_{x_j}\varphi\right)(x)\right| |h| \underset{h\to 0}{\to} 0.$$

This completes the proof of the theorem. □

A similar result can be proved for $T \in \mathcal{S}'$ and for $T \in \mathcal{E}'$.

One of the early goals of the theory of distributions was to be able to take any number of derivatives of a continuous function. The following result tells us that, in the end, distributions fit pretty tightly into that initial goal.

Theorem 102. *(for the proof, see [25], p. 82, Theorem XXI) Given $T \in \mathcal{D}'$ and given an open and bounded subset Ω of $\mathbb{R}^n$, there is a continuous function $f : \Omega \to \mathbb{C}$ and a number $m \in \mathbb{N}$ such that*

$$T\,|\Omega = \frac{\partial^{nm}}{\partial_{x_1^m} \ldots \partial_{x_n^m}} T_f$$

on $\mathcal{D}(\Omega)$.

Remark 103. Theorem 102 shows that every distribution restricted to an open and bounded subset of $\mathbb{R}^n$ has finite order. Since there are distributions of infinite order, the assumption "Ω is bounded" cannot be dropped.

Theorem 104. *(for the proof, see [25], p. 239, Theorem VI) Given $T \in \mathcal{S}'$, there is an slowly increasing and continuous function $f : \mathbb{R}^n \to \mathbb{C}$ and an n-tuple $\alpha \in \mathbb{N}^n$ so that*

$$T = \partial^\alpha T_f$$

on $\mathcal{S}$.

Remark 105. In view of 2) in Example 59, Theorem 104 justifies the name tempered, given to the distributions in $\mathcal{S}'$. Let us observe that Theorem 104 and Theorem 100 imply that every tempered distribution has finite order.

Remark 106. According to Theorem 77 and Theorem 100, the equality (5.3) allows us to say that $pv\frac{1}{x}$ has order ≤ 1, which is the optimal result. Likewise, (5.4) implies that $fp\frac{1}{x^2}$ has order ≤ 2, which is also optimal. However, (5.5), (5.6) and (5.7) do not yield optimal results.

Theorem 107. *Let $T \in \mathcal{E}'$ and let $U \subseteq \mathbb{R}^n$ be an open and bounded neighborhood of supp(T). Then, there is a finite number of n-tuples α and a finite family $\{f_\alpha\}_\alpha$ of continuous functions $f_\alpha : \mathbb{R}^n \to \mathbb{C}$, so that supp$(f_\alpha) \subseteq U$ and*

$$T = \sum_\alpha \partial^\alpha T_{f_\alpha}$$

on $\mathcal{E}$.

Proof. According to Theorem 62, if $\theta \in \mathcal{D}(U)$ is equal to one on an open neighborhood of *supp*(T), by definition we have

$$(T, \varphi) = (T, \theta\varphi),$$

for all $\varphi \in \mathcal{E}$.

Furthermore, in view of Theorem 102, there is a continuous function $f : U \to \mathbb{C}$ and an n-tuple β so that

$$T = \partial^{\beta} T_f$$

on $\mathcal{D}(U)$.

Since $supp\,(\theta\varphi) \subseteq U$,

$$(T, \theta\varphi) = \left(\partial^{\beta} T_f, \theta\varphi\right) = (-1)^{|\beta|} \left(T_f, \partial^{\beta}(\theta\varphi)\right).$$

Theorem 4 implies that

$$\begin{aligned}
(T, \theta\varphi) &= (-1)^{|\beta|} \sum_{\gamma=0}^{\beta} \binom{\beta}{\gamma} \left(T_f, \left(\partial^{\beta-\gamma}\theta\right)(\partial^{\gamma}\varphi)\right) \\
&= (-1)^{|\beta|} \sum_{\gamma=0}^{\beta} \binom{\beta}{\gamma} \int_U f(x) \left(\partial^{\beta-\gamma}\theta\right)(x)\,(\partial^{\gamma}\varphi)(x)\,dx \\
&= \left(\sum_{\gamma=0}^{\beta} (-1)^{|\beta+\gamma|} \binom{\beta}{\gamma} \partial^{\gamma} T_{(\partial^{\beta-\gamma}\theta)f}, \varphi\right),
\end{aligned}$$

for every $\varphi \in \mathcal{E}$.

This completes the proof of the theorem. □

Remark 108. Theorem 107 shows that every distribution with compact support has finite order. However, 1) in Example 96 shows that there is not $m \in \mathbb{N}$ so that $\mathcal{E}' \subseteq \mathcal{D}'^{(m)}$.

Theorem 109. *(for the proof, see [25], p. 93, Theorem XXVIII) Let $T \in \mathcal{E}'$ and let us assume that T has order $\leq m$. Then, $T(\varphi) = 0$ if $\varphi \in \mathcal{E}$ and $(\partial^{\alpha}\varphi)\,|supp\,(T) = 0$ for $|\alpha| \leq m$.*

Using Theorem 109, the result that follows characterizes those distributions in $\mathcal{E}'$ that are concentrated on a point.

Theorem 110. *Given $T \in \mathcal{E}'$, the following statements are equivalent:*

1. *The distribution T is concentrated on $\{a\}$.*

2. $$T = \sum_{|\alpha| \leq m} c_{\alpha} \partial^{\alpha} \delta_a,$$

   for some $m \in \mathbb{N}$ and $c_{\alpha} \in \mathbb{C}$.

Proof. It should be clear that 2) $\Rightarrow$ 1). Conversely, if T has order $\leq m$, given $\varphi \in \mathcal{E}$ we write its Taylor expansion about a as

$$\begin{aligned}\varphi(x) &= \sum_{|\alpha|\leq m} \frac{(\partial^\alpha \varphi)(a)}{\alpha!}(x-a)^\alpha \\ &+ \sum_{|\alpha|=m+1} \frac{m+1}{\alpha!}\left(\int_0^1 (\partial^\alpha \varphi)(a+s(x-a))\,ds\right)(x-a)^\alpha .\end{aligned}$$

Let

$$\varphi_\alpha(x) = \left(\int_0^1 (\partial^\alpha \varphi)(a+s(x-a))\,ds\right),$$

for $|\alpha| = m+1$ fixed. It should be clear that $\varphi_\alpha \in \mathcal{E}$. Given $\beta \in \mathbb{N}^n$ with $|\beta| \leq m$,

$$\begin{aligned}\left(\partial^\beta\left((x-a)^\alpha \varphi_\alpha(x)\right)\right)_{x=a} &= \sum_{\gamma\leq\beta}\binom{\beta}{\gamma}\partial_x^\gamma\left((x-a)^\alpha\right)_{x=a}\left(\partial^{\beta-\gamma}\varphi_\alpha\right)_{x=a} \\ &= 0.\end{aligned}$$

Therefore, Theorem 109 implies that $(T,(x-a)^\alpha \varphi_\alpha(x)) = 0$. Then,

$$\begin{aligned}(T,\varphi) &= \sum_{|\alpha|\leq m}\frac{(\partial^\alpha\varphi)(a)}{\alpha!}(T,(x-a)^\alpha) \\ &= \left(\sum_{|\alpha|\leq m}\frac{(-1)^{|\alpha|}}{\alpha!}(T,(x-a)^\alpha)\,\partial^\alpha\delta_a,\varphi\right).\end{aligned}$$

This completes the proof of the theorem. $\square$

Theorem 111. *(see [25], p. 54, Theorem III; for the proof, see [25], p. 53, Theorem II) Let $f:\mathbb{R}\to\mathbb{C}$ be a continuous function. Let us assume that f' exists a.e. and that it is equal to a locally integrable function g. Then, if f is the indefinite integral of g,*

$$\frac{d}{dx}T_f = T_g.$$

Corollary 112. *The distribution $T_{\ln|x|}$ is tempered.*

Proof.

$$\frac{d}{dx}(x\ln|x|-x) = \ln|x|$$

for $x \neq 0$, where $x\ln|x|-x$, defined as zero for $x=0$, is a slowly increasing continuous function. Therefore, $x\ln|x|-x$ defines a tempered distribution.

In view of Theorem 98, $\frac{d}{dx}T_{x\ln|x|-x}$ belongs to $\mathcal{S}'$.

According to Theorem 111,

$$\frac{d}{dx}T_{x\ln|x|-x} = T_{\ln|x|},$$

so the distribution $T_{\ln|x|}$ is tempered.

This completes the proof of the corollary. □

The following theorem extends to distributions a result well known for functions.

Theorem 113. *Given $T \in \mathcal{D}'(\mathbb{R})$, $\frac{dT}{dx} = 0$ if, and only if, T is the distribution defined by a constant function.*

Proof. If $T = T_c$ for some constant function c,

$$\left(\frac{dT_c}{dx}, \varphi\right) = -c\int_{\mathbb{R}} \frac{d\varphi}{dx}(x)\,dx = 0,$$

since the function φ has compact support.

Conversely, given a fixed function α in $\mathcal{D}(\mathbb{R})$ with integral one, we can write $\varphi \in \mathcal{D}(\mathbb{R})$ as

$$\begin{aligned}
\varphi &= \varphi - \alpha\int_{\mathbb{R}} \varphi(x)\,dx + \alpha\int_{\mathbb{R}} \varphi(x)\,dx \\
&= \varphi_1 + \alpha\int_{\mathbb{R}} \varphi(x)\,dx,
\end{aligned}$$

where

$$\int_{\mathbb{R}} \varphi_1(x)\,dx = 0.$$

We claim that

$$\begin{aligned}
&\left\{\varphi \in \mathcal{D}(\mathbb{R}) : \int_{\mathbb{R}} \varphi(x)\,dx = 0\right\} \\
= &\left\{\varphi \in \mathcal{D}(\mathbb{R}) : \int_{-\infty}^{x} \varphi(t)\,dt \in \mathcal{D}(\mathbb{R})\right\}.
\end{aligned} \tag{5.11}$$

If $\varphi \in \mathcal{D}(\mathbb{R})$, then $supp(\varphi) \subseteq [a,b]$ for some $a, b \in \mathbb{R}$. Therefore, for all $x > b$,

$$\int_a^x \varphi(t)\,dt = \int_a^b \varphi(t)\,dt.$$

If $\int_{\mathbb{R}} \varphi(x)\,dx = 0$,

$$0 = \int_{\mathbb{R}} \varphi(x)\,dx = \int_a^b \varphi(t)\,dt.$$

So,

$$\int_{-\infty}^{x} \varphi(t)\, dt = \int_{a}^{x} \varphi(t)\, dt = 0$$

for $x > b$. Likewise, it should be clear that

$$\int_{-\infty}^{x} \varphi(t)\, dt = 0,$$

for every $x < a$.

Conversely, if $\int_{-\infty}^{x} \varphi(t)\, dt \in \mathcal{D}(\mathbb{R})$, there is $b \in \mathbb{R}$ so that

$$\int_{-\infty}^{x} \varphi(t)\, dt = 0,$$

for every $x > b$. As a consequence, $\int_{\mathbb{R}} \varphi(x)\, dx = 0$.

Therefore, we have proved the equality (5.11).

Then,

$$\begin{aligned}
(T, \varphi) &= (T, \varphi_1) + (T, \alpha) \int_{\mathbb{R}} \varphi(x)\, dx \\
&= \left(T, \frac{d}{dx}\left(\int_{-\infty}^{x} \varphi_1(t)\, dt\right)\right) + \left(T_{(T,\alpha)}, \varphi\right) \\
&\quad - \left(T', \int_{-\infty}^{x} \varphi_1(t)\, dt\right) + \left(T_{(T,\alpha)}, \varphi\right),
\end{aligned}$$

where $T_{(T,\alpha)}$ is the distribution defined by the function identically equal to the complex number (T, α). Therefore,

$$T = T_{(T,\alpha)}.$$

This completes the proof of the theorem. □

The search for primitives of a distribution goes hand in hand with the calculation of derivatives. Schwartz's book dedicates quite a bit of space to this subject (see [25], Chapter II, Sections 4-6).

6 Tensor product, convolution product, and multiplicative product, of distributions

As we did in the previous section with the derivative of a distribution, we will define each of these products in such a manner as to agree with the usual definition, when considering functions.

6.1 Tensor product

We begin with the tensor product of two functions.

Definition 114. *Given $f : \mathbb{R}^n \to \mathbb{C}$ and $g : \mathbb{R}^m \to \mathbb{C}$, the pointwise multiplication $f(x)g(y)$ defines a new function from $\mathbb{R}^n \times \mathbb{R}^m$ into $\mathbb{C}$, denoted $f \otimes g$ and called the tensor product of f and g.*

Lemma 115. *Given $f : \mathbb{R}^n \to \mathbb{C}$ and $g : \mathbb{R}^m \to \mathbb{C}$,*

$$supp\,(f \otimes g) = supp\,(f) \times supp\,(g)\,.$$

Proof. If $(x_0, y_0) \in supp\,(f \otimes g)$, by definition, there is a sequence $\{(x_j, y_j)\}_{j\geq 1}$ in $\mathbb{R}^{n+m}$ so that $(x_j, y_j) \to (x_0, y_0)$ as $j \to \infty$ and $f(x_j)\,g(y_j) \neq 0$ for all $j \geq 1$. Then, it should be clear, again by definition, that $x_0 \in supp\,(f)$ and $y_0 \in supp\,(g)$.

Conversely, if $(x_0, y_0) \in supp\,(f) \times supp\,(g)$, there are sequences $\{x_j\}_{j\geq 1}$ and $\{y_j\}_{j\geq 1}$ converging to x_0 and y_0, respectively, as $j \to \infty$, so that $f(x_j) \neq 0$ and $g(y_j) \neq 0$ for all $j \geq 1$.

Therefore, the sequence $\{(x_j, y_j)\}_{j\geq 1}$ converges to (x_0, y_0) as $j \to \infty$ and $(f \otimes g)(x_j, y_j) \neq 0$ for all $j \geq 1$. So, $(x_0, y_0) \in supp\,(f \otimes g)$.

This completes the proof of the lemma. □

Remark 116. Definition 114 can be extended, associatively, to functions $f_l : \mathbb{R}^{n_l} \to \mathbb{C}$ for $l = 3, 4, ...$ as

$$(f_1 \otimes f_2 \otimes ... \otimes f_{l-1}) \otimes f_l = (f_1(a)\,f_2(b)\,...f_{l-1}(y))\,f_l(z)\,.$$

Example 117. The n-dimensional Heaviside function H_n defined in 6) of Example 96, is the tensor product of n one-dimensional Heaviside functions $H_{x_1}, ..., H_{x_n}$ (see Definition 11).

$$H_n = H_{x_1} \otimes ... \otimes H_{x_n}.$$

In order to extend Definition 114 to distributions, we need to work simultaneously with test functions having different domains and acting on different variables. Therefore, it is convenient to indicate $\mathcal{D}_x$, $\mathcal{D}_y$, and $\mathcal{D}_{x,y}$, the space of test functions defined on $\mathbb{R}^n$, $\mathbb{R}^m$, and $\mathbb{R}^n \times \mathbb{R}^m$, respectively.

For the purpose of defining $\mathcal{D}(\mathbb{R}^n \times \mathbb{R}^m)$, $L^1_{loc}(\mathbb{R}^n \times \mathbb{R}^m)$, etc., or whenever it is convenient, we may identify $\mathbb{R}^n \times \mathbb{R}^m$ with $\mathbb{R}^{n+m}$.

Definition 118. *Let X, Y, Z be topological linear spaces. A bilinear map $F : X \times Y \to Z$ is hypocontinuous if it satisfies the following conditions:*

1. *It is separately continuous. That is, $F(x, \cdot) : Y \to Z$ is continuous for each $x \in X$ fixed, and $F(\cdot, y) : X \to Z$ is continuous for each $y \in Y$ fixed.*

2. *For every bounded set $\mathcal{X} \subseteq X$, the linear map $F(x,\cdot): Y \to Z$ is continuous, uniformly on $x \in \mathcal{X}$.*

3. *For every bounded set $\mathcal{Y} \subseteq Y$, the linear map $F(\cdot,y): X \to Z$ is continuous, uniformly on $y \in \mathcal{Y}$.*

Remark 119. Condition 2) in Definition 118 says that the family of linear maps $\{F(x,\cdot)\}_{x\in\mathcal{X}}$ is equicontinuous.

Likewise, condition 3) in Definition 118 says that the family of linear maps $\{F(\cdot,y)\}_{y\in\mathcal{Y}}$ is equicontinuous.

The notion of hypocontinuity is stronger than separate continuity and weaker that continuity.

For more on hypocontinuity, see ([26], pp. 107 and 300).

Definition 120. *Given $k = 0,1,...,$ let $\mathcal{E}^{(k)}$ be the complex linear space consisting of those functions $\varphi : \mathbb{R}^n \to \mathbb{C}$ that are continuous and have continuous derivatives of order $\leq k$. The space $\mathcal{E}^{(k)}$ is a Fréchet space with the topology defined by the countable family of norms*

$$\|\varphi\|_{k,B_l} = \sup_{|\alpha|\leq k} \sup_{x\in B_l} |(\partial^\alpha \varphi)(x)|,$$

where $B_l = \{x \in \mathbb{R}^n : |x| \leq l\}$ for $l = 1, 2, \ldots$ and k is fixed.

Theorem 121. *([25], p. 112) The map*

$$(f,g) \to f \otimes g$$

is bilinear and continuous from $\mathcal{E}_x \times \mathcal{E}_y$ into $\mathcal{E}_{x,y}$ and it is bilinear and hypocontinuous from $\mathcal{D}_x \times \mathcal{D}_y$ into $\mathcal{D}_{x,y}$ and from $\mathcal{E}_x^{(k)} \times \mathcal{D}_y^{(k)}$ into $\mathcal{D}_{x,y}^{(k)}$.

Let us observe that when the functions f and g in Definition 114 are locally integrable on $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, the tensor product $f \otimes g$ is locally integrable on $\mathbb{R}^n \times \mathbb{R}^m$.

Thus, given $\varphi \in \mathcal{D}_{x,y}$, Fubini's theorem tells us that

$$\begin{aligned}
(T_{f\otimes g}, \varphi) &= \int_{\mathbb{R}^n\times\mathbb{R}^m} f(x)\, g(y)\, \varphi(x,y)\, dxdy \\
&= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} g(y)\, \varphi(x,y)\, dy \right) f(x)\, dx \\
&= \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^n} f(x)\, \varphi(x,y)\, dx \right) g(y)\, dy.
\end{aligned}$$

We claim that these iterated integrals can be written, in the sense of distributions, as

$$\begin{aligned}\left(T_{(f\otimes g)},\varphi\right)_{\mathcal{D}'_{x,y},\mathcal{D}_{x,y}} &= \left(T_{f(x)},\left(T_{g(y)},\varphi\left(x,y\right)\right)_{\mathcal{D}'_y\mathcal{D}_y}\right)_{\mathcal{D}'_x\mathcal{D}_x}\\ &= \left(T_{g(y)},\left(T_{f(x)},\varphi\left(x,y\right)\right)_{\mathcal{D}'_x\mathcal{D}_x}\right)_{\mathcal{D}'_y\mathcal{D}_y}.\end{aligned}$$

First of all, for each $x\in\mathbb{R}^n$, the function $\varphi\left(x,\cdot\right)$ is smooth and *supp* $\left(\varphi\left(x,\cdot\right)\right)$ is contained in the projection of *supp* (φ) onto $\mathbb{R}^m$. Thus, $\varphi\left(x,\cdot\right)\in\mathcal{D}_y$, so $\left(T_g,\varphi\left(x,\cdot\right)\right)_{\mathcal{D}'_y\mathcal{D}_y}$ is well defined for each $x\in\mathbb{R}^n$. Likewise, $\left(T_f,\varphi\left(\cdot,y\right)\right)_{\mathcal{D}'_x\mathcal{D}_x}$ is well defined for each $y\in\mathbb{R}^m$.

According to Theorem 6 and Theorem 8, the functions

$$\lambda\left(x\right)=\left(T_g,\varphi\left(x,\cdot\right)\right)_{\mathcal{D}'_y\mathcal{D}_y}$$

and

$$\psi\left(y\right)=\left(T_f,\varphi\left(\cdot,y\right)\right)_{\mathcal{D}'_x\mathcal{D}_x}$$

are smooth. Moreover, if *supp* $(\varphi)\subseteq K$, compact subset of $\mathbb{R}^n\times\mathbb{R}^m$, the support of λ is included in K_1, the projection of K onto $\mathbb{R}^n$. In fact, if $x\in\mathbb{R}^n\backslash K_1$, then $(x,y)\in(\mathbb{R}^n\times\mathbb{R}^m)\backslash K$ for every $y\in\mathbb{R}^m$. Therefore, $\lambda\left(x\right)=0$. Since K_1 is a compact subset of $\mathbb{R}^n$, we conclude that $\lambda\in\mathcal{D}_x$. Likewise, $\psi\in\mathcal{D}_y$.

Let us observe that if φ has separated variables, that is $\varphi\left(x,y\right)=\alpha\left(x\right)\beta\left(y\right)$, $\alpha\in\mathcal{D}_x$ and $\beta\in\mathcal{D}_y$, we have

$$\left(T_{f\otimes g},\varphi\right)_{\mathcal{D}'_{x,y},\mathcal{D}_{x,y}}=\left(T_f,\alpha\right)_{\mathcal{D}'_x\mathcal{D}_x}\left(T_g,\beta\right)_{\mathcal{D}'_y\mathcal{D}_y}.$$

Our goal is to prove that all these statements remain valid when, instead of T_f and T_g, we consider a pair of arbitrary distributions. To this end, we need the following result:

Theorem 122. *The linear subspace of $\mathcal{D}_{x,y}$ consisting of finite linear combinations of functions with separated variables is dense in $\mathcal{D}_{x,y}$.*

Proof. Let us fix $\varphi\in\mathcal{D}_{x,y}$. We can assume that there are compact sets $K_1\subseteq\mathbb{R}^n$ and $K_2\subseteq\mathbb{R}^m$ so that *supp* $(\varphi)\subseteq K_1\times K_2$. Moreover, let $k,k'\geq1$ be so that, for some $\varepsilon>0$,

$$\begin{aligned}&\varepsilon\text{-neigbordhood}\left(K_1\right)\times\varepsilon\text{-neigbordhood}\left(K_2\right)\\ \subseteq\ &\left\{x\in\mathbb{R}^n:|x|\leq k\right\}\times\left\{y\in\mathbb{R}^m:|y|\leq k\right\}\\ \subseteq\ &\left\{(x,y)\in\mathbb{R}^n\times\mathbb{R}^m:|(x,y)|\leq k'\right\}.\end{aligned}$$

According to (4.1) and Theorem 90, given $l=0,1,...$, there is a sequence $\left\{P_j\right\}_{j\geq1}$ of complex polynomial functions such that $\left\{\partial^\alpha P_j\right\}_{j\geq1}$ converges to $\partial^\alpha\varphi$ as $j\to\infty$, uniformly with respect to $(x,y)\in B_{k'}$ and $\alpha\in\mathbb{N}^n$ with $|\alpha|\leq l$.

Now, let $\psi_1\in\mathcal{D}_x$ and $\psi_2\in\mathcal{D}_y$ be test functions satisfying the following properties:

1. $0 \leq \psi_1(x) \leq 1$ for all $x \in \mathbb{R}^n$, $\psi_1(x) = 1$ for $x \in \varepsilon/2$-neigbordhood(K_1), $supp(\psi_1) \subseteq \varepsilon$-neigbordhood$(K_1)$.

2. $0 \leq \psi_2(y) \leq 1$ for all $y \in \mathbb{R}^m$, $\psi_2(y) = 1$ for $x \in \varepsilon/2$-neigbordhood(K_2), $supp(\psi_2) \subseteq \varepsilon$-neigbordhood$(K_2)$.

Then, the sequence $\{\psi_1\psi_2 P_j\}_{j\geq 1}$ converges to $\partial^\alpha\varphi$ as $j \to \infty$, uniformly with respect to $(x, y) \in K_1 \times K_2$ and $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq l$.

Furthermore, it should be clear that, for each $j \geq 1$, the function $\psi_1(x)\psi_2(y)P(x, y)$ is a finite linear combination of functions with separated variables.

This completes the proof of the theorem. □

Remark 123. According to Theorem 53, $\mathcal{D}_{x,y}$ is dense in $\mathcal{S}_{x,y}$ and in $\mathcal{E}_{x,y}$. Therefore, Theorem 122 implies that the linear subspace of $\mathcal{D}_{x,y}$ consisting of finite linear combinations of functions with separated variables is dense in $\mathcal{S}_{x,y}$ and in $\mathcal{E}_{x,y}$.

Theorem 124. *Let $T \in \mathcal{D}'_x$ and $S \in \mathcal{D}'_y$.*

1. *If $\varphi \in \mathcal{D}_{x,y}$, the function $\lambda(x) = (S_y, \varphi(x, y))_{\mathcal{D}'_y\mathcal{D}_y}$ is well defined and belongs to $\mathcal{D}_x$. Likewise, the function $\psi(y) = (T_x, \varphi(x, y))_{\mathcal{D}'_x\mathcal{D}_x}$ is well defined and belongs to $\mathcal{D}_y$.*

2. *The maps*
$$\varphi \to \left(T_x, (S_y, \varphi(x, y))_{\mathcal{D}'_y\mathcal{D}_y}\right)_{\mathcal{D}'_x\mathcal{D}_x} \tag{6.1}$$
and
$$\varphi \to \left(S_y, (T_x, \varphi(x, y))_{\mathcal{D}'_x\mathcal{D}_x}\right)_{\mathcal{D}'_y\mathcal{D}_y} \tag{6.2}$$
belong to $\mathcal{D}'_{x,y}$.

3. $$\left(T_x, (S_y, \varphi(x, y))_{\mathcal{D}'_y\mathcal{D}_y}\right)_{\mathcal{D}'_x\mathcal{D}_x} = \left(S_y, (T_x, \varphi(x, y))_{\mathcal{D}'_x\mathcal{D}_x}\right)_{\mathcal{D}'_y\mathcal{D}_y}, \tag{6.3}$$
for all $\varphi \in \mathcal{D}_{x,y}$.

Proof. For $x \in \mathbb{R}^n$ fixed, the function $y \to \varphi(x, y)$ has compact support contained in the projection onto $\mathbb{R}^m$ of *supp* (φ). Moreover, it should be clear that it is smooth. Therefore, the function λ is well defined. If $\{x_j\}_{j\geq 1}$ converges to x in $\mathbb{R}^n$ as $j \to \infty$, the sequence $\{\varphi(x_j, \cdot)\}_{j\geq 1}$ converges to $\varphi(x, \cdot)$ in $\mathcal{D}_y$ as $j \to \infty$. Thus, $\lambda(x_j) \to \lambda(x)$ in $\mathbb{C}$ as $j \to \infty$. So, λ is continuous on $\mathbb{R}^n$. With the notation used in Theorem 101,

$$\frac{\lambda(x + h) - \lambda(x)}{h_1} = \left(S_y, \frac{\varphi(x + h, y) - \varphi(x, y)}{h_1}\right)_{\mathcal{D}'_y\mathcal{D}_y}.$$

Let us show that there is

$$\lim_{h\to 0}\frac{\varphi(x+h,\cdot)-\varphi(x,\cdot)}{h_1}=(\partial_{x_1}\varphi)(x,\cdot) \tag{6.4}$$

in $\mathcal{D}_y$ for each $x\in\mathbb{R}^n$. If we fix $k=0,1,...,$ $\beta\in\mathbb{N}^n$ with $|\beta|\le k$ and $h\in\mathbb{R}$ with $|h|\le 1$,

$$\begin{aligned}\frac{\left(\partial_y^\beta\varphi\right)(x+h,y)-\left(\partial_y^\beta\varphi\right)(x,y)}{h_1}-\left(\partial_{x_1}\partial_y^\beta\varphi\right)(x,y) &= \left(\partial_{x_1}\partial_y^\beta\varphi\right)(\theta,x',y)-\left(\partial_{x_1}\partial_y^\beta\varphi\right)(x,y)\\ &= \left(\partial^2_{x_1^2}\partial_y^\beta\varphi\right)(\xi,x',y)(\theta-x_1),\end{aligned}$$

where θ belongs to the segment with end points x_1 and x_1+h_1, while ξ belongs to the segment with end points θ and x_1. Thus,

$$\begin{aligned}&\left|\frac{\left(\partial_y^\beta\varphi\right)(x+h,y)-\left(\partial_y^\beta\varphi\right)(x,y)}{h_1}-\left(\partial_{x_1}\partial_y^\beta\varphi\right)(x,y)\right|\\ \le&\left(\sup_{|\alpha|\le k+2}\ \sup_{(x,y)\in\mathbb{R}^n\times\mathbb{R}^m}\left|\left(\partial^\alpha\varphi\right)(x,y)\right|\right)|\theta-x_1|\underset{h\to 0}{\to}0,\end{aligned}$$

where $\alpha\in\mathbb{N}^{n+m}$. Since

$$supp_y\left(\frac{\left(\partial_y^\beta\varphi\right)(x+h,y)-\left(\partial_y^\beta\varphi\right)(x,y)}{h_1}\right)\subseteq K_2,$$

where K_2 is a compact subset of $\mathbb{R}^m$, we conclude that (6.4) is true.

In the same way, using induction on the order of the partial derivatives of λ, we conclude that $\lambda\in\mathcal{D}_x$.

With the obvious change of notation, we can prove in the exact same manner that the function ψ is well defined and it belongs to $\mathcal{D}_y$. Thus, we have shown that 1) is true.

Moreover, for each $\beta\in\mathbb{N}^n$ and each $\nu\in\mathbb{N}^m$,

$$\left(\partial^\beta\lambda\right)(x)=\left(S_y,\left(\partial_x^\beta\varphi\right)(x,y)\right)_{\mathcal{D}'_y\mathcal{D}_y}$$

and

$$\left(\partial^\gamma\psi\right)(y)=\left(T_x,\left(\partial_y^\gamma\varphi\right)(x,y)\right)_{\mathcal{D}'_x\mathcal{D}_x}.$$

Let us now prove that the map (6.1 belongs to $\mathcal{D}'_{x,y}$. To begin, it should be clear that it is linear. So, it remains to prove that it is continuous on $\mathcal{D}_{x,y}$. According to Theorem 70, it suffices to show that it satisfies (4.12). That is to say, that given $K \subseteq \mathbb{R}^n \times \mathbb{R}^m$ there is $l_K \in \mathbb{N}$ and $C_{K,l} > 0$ so that

$$\begin{aligned}&\left|\left(T_x, (S_y, \varphi(x,y))_{\mathcal{D}'_y \mathcal{D}_y}\right)_{\mathcal{D}'_x \mathcal{D}_x}\right| \\ \leq \; & C_{K,l} \sup_{|\beta|+|\gamma| \leq l_K} \sup_{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m} \left|\left(\partial_x^\beta \partial_y^\gamma \varphi\right)(x,y)\right|,\end{aligned}$$

for all $\varphi \in \mathcal{D}_{x,y}$.

In fact, let K_1 be the projection of K onto $\mathbb{R}^n$. According to 1), the function $\lambda(x) = (S_y, \varphi(x,y))$ belongs to $\mathcal{D}_{x,K_1}$. Since $T \in \mathcal{D}'_x$, Theorem 70 implies that there is $l_{1,K_1} \in \mathbb{N}$ and $C_{K_1,l_1} > 0$ so that

$$\begin{aligned}\left|\left(T_x, (S_y, \varphi(x,y))_{\mathcal{D}'_y \mathcal{D}_y}\right)_{\mathcal{D}'_x \mathcal{D}_x}\right| &\leq C_{K_1,l_1} \sup_{|\beta| \leq l_{1,K_1}} \sup_{x \in \mathbb{R}^n} \left|\partial_x^\beta \left((S_y, \varphi(x,y))_{\mathcal{D}'_y \mathcal{D}_y}\right)\right| \\ &= C_{K_1,l_1} \sup_{|\beta| \leq l_{1,K_1}} \sup_{x \in \mathbb{R}^n} \left|\left(S_y, \left(\partial_x^\beta \varphi\right)(x,y)\right)_{\mathcal{D}'_y \mathcal{D}_y}\right|.\end{aligned}$$

For every $x \in \mathbb{R}^n$, the function $\left(\partial_x^\beta \varphi\right)(x,y)$ belongs to $\mathcal{D}_{y,K_2}$ where K_2 is the projection of K onto $\mathbb{R}^m$. Since $S \in \mathcal{D}'_y$, there is $l_{2,K_2} \in \mathbb{N}$ and $C_{K_2,l_2} > 0$ so that

$$\begin{aligned}&\left|\left(T_x, (S_y, \varphi(x,y))_{\mathcal{D}'_y \mathcal{D}_y}\right)_{\mathcal{D}'_x \mathcal{D}_x}\right| \\ \leq \; & C_{K_1,l_1} C_{K_2,l_2} \sup_{\substack{|\beta| \leq l_{1,K_1} \\ |\gamma| \leq l_{2,K_2}}} \sup_{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m} \left|\left(\partial_x^\beta \partial_y^\gamma \varphi\right)(x,y)\right|.\end{aligned}$$

Thus, condition (4.12) is satisfied, so the map (6.1) belongs to $\mathcal{D}'_{x,y}$. In exactly the same manner, with the obvious change of notation, we can verify that the map (6.2) also belongs to $\mathcal{D}'_{x,y}$. So, we have proved 2).

As for 3), it should be clear that the equality (6.3) is true when φ is a finite linear combination of functions with separated variables. Then, Theorem 122 implies that (6.3) holds for every $\varphi \in \mathcal{D}'_{x,y}$.

This completes the proof of the theorem. □

Remark 125. Statement 1) in Theorem 124 is a distributional version of Theorem 6 and Theorem 8.

Corollary 126. *Given $T \in \mathcal{D}'_x$ and $S \in \mathcal{D}'_y$, there is a distribution $W \in \mathcal{D}'_{x,y}$ such that*

$$(W, \varphi)_{\mathcal{D}'_{x,y}, \mathcal{D}_{x,y}} = (T, \alpha)_{\mathcal{D}'_x \mathcal{D}_x} (S, \beta)_{\mathcal{D}'_y \mathcal{D}_y},$$

when $\varphi(x,y) = \alpha(x)\beta(y)$ for $\alpha \in \mathcal{D}_x$ and $\beta \in \mathcal{D}_y$. Moreover, W is unique.

Proof. Either one of (6.1) or (6.2) proves that W exists. As for the uniqueness, it follows from Theorem 122.

This completes the proof of the corollary. □

Definition 127. *Given $T \in \mathcal{D}'_x$ and $S \in \mathcal{D}'_y$, the distribution W is called the tensor product of T and S, denoted $T \otimes S$.*

Remark 128. According to Theorem 124 and Corollary 126, the distribution $T \otimes S$ can be defined as the map

$$\varphi \to \left(T_x, (S_y, \varphi(x,y))_{\mathcal{D}'_y\mathcal{D}_y}\right)_{\mathcal{D}'_x\mathcal{D}_x}$$

or the map

$$\varphi \to \left(S_y, (T_x, \varphi(x,y))_{\mathcal{D}'_x\mathcal{D}_x}\right)_{\mathcal{D}'_y\mathcal{D}_y},$$

for every $\varphi \in \mathcal{D}_{x,y}$.

Let us observe that

$$\begin{aligned}(T \otimes S, \varphi) &= \left(T_x, (S_y, \varphi(x,y))_{\mathcal{D}'_y\mathcal{D}_y}\right)_{\mathcal{D}'_x\mathcal{D}_x} \\ &= \left(S_y, (T_x, \varphi(x,y))_{\mathcal{D}'_x\mathcal{D}_x}\right)_{\mathcal{D}'_y\mathcal{D}_y}\end{aligned}$$

for all $\varphi \in \mathcal{D}_{x,y}$, which can be interpreted as a Fubini's theorem for distributions.

In what follows, we will not label the pairings, unless it is necessary for clarity.

Theorem 129. *Given $T \in \mathcal{D}'_x$ and $S \in \mathcal{D}'_y$,*

$$supp\,(T \otimes S) = supp\,(T) \times supp\,(S) \tag{6.5}$$

and

$$\partial_x^\beta \partial_y^\gamma (T \otimes S) = \partial_x^\beta T \otimes \partial_y^\gamma S, \tag{6.6}$$

for all $\beta \in \mathbb{N}^n$ and for all $\gamma \in \mathbb{N}^m$.

Proof. If $x \in supp\,(T)$, and $y \in supp\,(S)$, let $U_1 \subseteq \mathbb{R}^n$ and $U_2 \subseteq \mathbb{R}^m$ be open neighborhoods of x and y, respectively. Then, $U = U_1 \times U_2 \subseteq \mathbb{R}^n \times \mathbb{R}^m$ is an open neighborhood of (x, y).

Furthermore, let $\alpha_1 \in \mathcal{D}_x(U_1)$ and $\alpha_2 \in \mathcal{D}_y(U_2)$ be such that $(T, \alpha_1) \neq 0$ and $(S, \alpha_2) \neq 0$. Thus, $(T \otimes S, \alpha_1 \otimes \alpha_2) \neq 0$. So, $(x, y) \in supp\,(T \otimes S)$.

Conversely, if $(x, y) \in supp\,(T \otimes S)$, there is an open neighborhood $U \subseteq \mathbb{R}^n \times \mathbb{R}^m$ of (x, y) and a function $\varphi \in \mathcal{D}_{x,y}$ so that $(T \otimes S, \varphi) \neq 0$.

Since $(T \otimes S, \varphi) = (T_x, (S_y, \varphi(x,y)))$, we conclude that T is different from zero on the test function $(S_y, \varphi(x,y)) \in \mathcal{D}_x(U_1)$, for some open neighborhood $U_1 \subseteq \mathbb{R}^n$ of x. Thus, $x \in supp\,(T)$.

Using the representation $(T \otimes S, \varphi) = (S_y, (T_x, \varphi(x,y)))$, we conclude, in the same manner, that $y \in supp(T)$. Thus, we have proved (6.5).

As for (6.6), according to Theorem 124, if $\varphi \in \mathcal{D}_{x,y}$,

$$\begin{aligned}
\left(\partial_x^\beta \partial_y^\gamma (T \otimes S), \varphi\right) &= (-1)^{|\beta|+|\gamma|} \left(T \otimes S, \partial_x^\beta \partial_y^\gamma \varphi\right) \\
&= (-1)^{|\beta|+|\gamma|} \left(T_x, \left(S_y, \left(\partial_x^\beta \partial_y^\gamma\right) \varphi(x,y)\right)\right) \\
&= (-1)^{|\beta|} \left(T_x, \left(\partial_y^\gamma S_y, \left(\partial_x^\beta \partial_y^\gamma\right) \varphi(x,y)\right)\right) \\
&= (-1)^{|\beta|} \left(T_x, \partial_x^\beta \left(\partial_y^\gamma S_y, \left(\partial_y^\gamma\right) \varphi(x,y)\right)\right) \\
&= \left(\partial_x^\beta T_x, \left(\partial_y^\gamma S_y, \left(\partial_y^\gamma\right) \varphi(x,y)\right)\right) \\
&= \left(\partial^\beta T \otimes \partial^\gamma S, \varphi\right).
\end{aligned}$$

This completes the proof of the theorem. □

Remark 130. Definition 127 can be extended to distributions $T_l \in \mathcal{D}'(\mathbb{R}^{n_l})$ for $l = 3, 4, ...$. The resulting product is associative. That is, we can insert a parenthesis anywhere. For instance,

$$\begin{aligned}
(T_1 \otimes T_2 \otimes ... \otimes T_{l-1}) \otimes T_l &= T_1 \otimes (T_2 \otimes ... \otimes T_{l-1} \otimes T_l) \\
&= T_1 \otimes T_2 \otimes ... \otimes (T_{l-1} \otimes T_l),
\end{aligned}$$

etc..

The following result is an immediate consequence of (6.5) and Theorem 62.

Corollary 131. *If $T \in \mathcal{E}'_x$ and $S \in \mathcal{E}'_y$, the tensor product $T \otimes S$ belongs to $\mathcal{E}'_{x,y}$.*

Theorem 132. *(for the proof, see [25], p. 110, Theorem VI) The map*

$$(T, S) \to T \otimes S$$

is bilinear and continuous from $\mathcal{D}'_x \times \mathcal{D}'_y$ into $\mathcal{D}'_{x,y}$, from $\mathcal{D}_x^{(k)\prime} \times \mathcal{D}_y^{(k)\prime}$ into $\mathcal{D}_{x,y}^{(k)\prime}$ for $k = 0, 1, ...$, and from $\mathcal{E}'_x \times \mathcal{E}'_y$ into $\mathcal{E}'_{x,y}$.

Remark 133. It should be clear that the construction of the tensor product given in Theorem 124, can be extended to distributions T, S in $\mathcal{S}'$.

Example 134. 1. Given $a = (a_1, ..., a_n) \in \mathbb{R}^n$, the n-dimensional Dirac measure δ_a, as defined in 2) of Example 23, is the tensor product of the n one-dimensional Dirac measures $\delta_{a_1}, ..., \delta_{a_n}$.

$$\delta_a = \delta_{a_1} \otimes ... \otimes \delta_{a_n}.$$

2. If $\beta = (1, ..., 1) \in \mathbb{N}^n$ and $0 = (0_1, ..., 0_n)$,

$$\begin{aligned} \partial_x^\beta T_{H_n} &= \left(\partial_{x_1} T_{H_{x_1}}\right) \otimes ... \otimes \left(\partial_{x_n} T_{H_{x_n}}\right) \\ &= \delta_{0_1} \otimes ... \otimes \delta_{0_n} = \delta_0. \end{aligned}$$

3. The tensor product can be used to extend a distribution defined on a linear subspace. Indeed, let $T \in \mathcal{D}'(\mathbb{R}^n)$ and let δ_0 be the Dirac measure on $\mathbb{R}^k$. Then, it should be clear that the map

$$\begin{array}{ccc} \mathcal{D}\left(\mathbb{R}^{n+k}\right) & \to & \mathbb{C} \\ \varphi & \to & (T \otimes \delta_0)(\varphi) = (T_x, \varphi(x, 0)) \end{array}$$

is a distribution in $\mathcal{D}'\left(\mathbb{R}^{n+k}\right)$.

4. There is a version of 3) for Radon measures.

Indeed, if T is a Radon measure on $\mathbb{R}^n$, let $\mu_T : \mathcal{B}_n \to [0, \infty]$ be the unique regular Borel measure that Theorem 85 associates with T. Furthermore, according to 2) in Example 86, let $\mu_0 : \mathcal{B}_k \to [0, \infty)$ be the measure associated with the Dirac measure δ_0. Remark 83, tells us that μ_0 is regular.

Let $\mu_T \times \mu_0 : \mathcal{B}_n \times \mathcal{B}_k \to [0, \infty]$ be the regular Borel measure that Theorem 85 associates with $T \otimes \delta_0$.

Let us recall that if $\mathcal{B}_n$, $\mathcal{B}_k$, and $\mathcal{B}_{n+k}$ are the Borel σ-algebras on $\mathbb{R}^n$, $\mathbb{R}^k$, and $\mathbb{R}^{n+k}$, respectively, the product σ-algebra $\mathcal{B}_n \times \mathcal{B}_k$ is equal to $\mathcal{B}_{n+k}$.

The measure $\mu_T \times \mu_0$ is defined as

$$(\mu_T \times \mu_0)(A) = \int_{\mathbb{R}^n} \chi_A(x, 0)\, d\mu_T,$$

where χ_A is the characteristic function of $A \in \mathcal{B}_n \times \mathcal{B}_k$.

Therefore, $\mu_T \times \mu_0$, called the regular Borel product of T and δ_0 ([7], p. 245), is an extension to $\mathbb{R}^n \times \mathbb{R}^k$ of μ_T.

Let us observe that $\mu_T \times \mu_0 : \mathcal{B}_{n+k} \to [0, \infty]$ and the Lebesgue measure λ_{n+k} on $\mathbb{R}^{n+k}$ restricted to $\mathcal{B}_{n+k}$, are mutually singular.

Indeed, if $\mathbb{X}^n = \{(x, 0) : x \in \mathbb{R}^n\}$, then $\lambda_{n+k}(\mathbb{X}_n) = 0$ and $(\mu_T \times \mu_0)\left(\mathbb{R}^{n+k} \backslash \mathbb{X}^n\right) = 0$.

Finally, for $\varphi \in \mathcal{D}\left(\mathbb{R}^{n+k}\right)$,

$$(T \otimes \delta_0)(\varphi) = \int_{\mathbb{R}^n} \varphi(x, 0)\, d\mu_T.$$

6.2 Convolution product

According to Young's convolution theorem (see, for instance, [29], p. 146, Theorem 9.2 and p. 145, Theorem 9.1), if $f, g \in L^1$, the convolution $f * g \in L^1$. Thus, given $\varphi \in \mathcal{D}$,

$$\begin{aligned}
(T_{f*g}, \varphi) &= \int_{\mathbb{R}^n} (f * g)(x)\, \varphi(x)\, dx \\
&= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(x-y)\, g(y)\, dy \right) \varphi(x)\, dx \\
& \int_{\mathbb{R}^n \times \mathbb{R}^n} f(x)\, g(y)\, \varphi(x+y)\, dx dy,
\end{aligned} \tag{6.7}$$

Since $f(x)\, g(y) \in L^1(\mathbb{R}^n \times \mathbb{R}^n)$, (6.7) can be interpreted as the $\left(L^1, L^\infty\right)$ pairing

$$(f \otimes g, \varphi(x+y))_{L^1, L^\infty}.$$

Given $T, S \in \mathcal{D}'$, we would want to define the convolution $T * S$ as

$$(T \otimes S, \varphi(x+y)), \tag{6.8}$$

for $\varphi \in \mathcal{D}$. Unfortunately, (6.8) is not defined as a $(\mathcal{D}', \mathcal{D})$ pairing, as shown in 1) of the lemma that follows.

Lemma 135. *Given $\varphi \in \mathcal{D}$,*

1. *If φ is not identically zero,*

$$supp_{x,y}(\varphi(x+y)) \supseteq \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : x+y = a \text{ with } \varphi(a) \neq 0\}. \tag{6.9}$$

   Since the right-hand side of (6.9) consists of affine subspaces of $\mathbb{R}^n \times \mathbb{R}^n$, (6.9) implies $\varphi(x+y)$ has compact support when, and only when, φ is identically zero.

2. $$supp_{x,y}(\varphi(x+y)) = \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : x+y \in supp(\varphi)\}. \tag{6.10}$$

Proof. By definition of support of a function, $supp_{x,y}(\varphi(x+y))$ contains all the points (x,y) where the function is not zero. So, (6.9) is true.

If (x_0, y_0) does not belong to the right-hand side of (6.10), by the continuity of the map $(x,y) \to x+y$, there are open neighborhoods U and V of x_0 and y_0, respectively, such that

$$U + V \subseteq (\mathbb{R}^n \times \mathbb{R}^n) \setminus \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : x+y \in supp(\varphi)\}.$$

Then, $\varphi(x+y)$ is zero on $U \times V$, from which we conclude that $(x_0, y_0) \notin supp_{x,y}(\varphi(x+y))$.

Conversely, if (x_0, y_0) does not belong to the left-hand side of (6.10), there are open neighborhoods U and V of x_0 and y_0, respectively, such that

$$U + V \subseteq (\mathbb{R}^n \times \mathbb{R}^n) \setminus supp_{x,y}(\varphi(x+y)).$$

Therefore, $\varphi(x+y) = 0$ for all $(x,y) \in U \times V$. In other words, $\varphi(z) = 0$ for every z in the open subset $U+V$ of $\mathbb{R}^n$. So, (x_0, y_0) does not belong to the right-hand side of (6.10). Thus, we have proved 2).

This completes the proof of the lemma. □

Remark 136. When the intersection of $supp(T \otimes S)$ and $supp_{x,y}(\varphi(x+y))$ is compact, the pairing $(T \otimes S, \varphi(x+y))$ can be defined as in Lemma 65. The next result provides an example of this situation.

Theorem 137. *Given $T, S \in \mathcal{D}'$, let us assume that at least one of them has compact support.*

1. *We can define the pairing (6.8) as a distribution in $\mathcal{D}'$, which is denoted $T * S$ and it is called the convolution of T and S.*

2. $$supp(T * S) \subseteq supp(T) + supp(S).$$

Proof. Let us say that T has compact support. Then, given $\varphi \in \mathcal{D}$, we claim that $supp(T \otimes S) \bigcap supp_{x,y}(\varphi(x+y))$ is compact.

Indeed, let us recall that $supp(T \otimes S) = supp(T) \times supp(S)$. It should be clear that the intersection is a closed subset of $\mathbb{R}^n \times \mathbb{R}^n$. Let us prove that it is bounded.

There are $C_1 > 0$ and $C_2 > 0$ so that $|x| \leq C_1$ for $x \in supp(T)$ and $|z| \leq C_2$ for $z \in supp(\varphi)$. Then, if $(x,y) \in supp(T \otimes S) \bigcap supp_{x,y}(\varphi(x+y))$, we have

$$\begin{aligned} |(x,y)| &= \left(|x|^2 + |y|^2\right)^{1/2} \leq |x| + |y| \\ &\leq 2|x| + |x+y| \leq 2C_1 + C_2. \end{aligned}$$

Therefore, Lemma 65 implies that the pairing $(T \otimes S, \varphi(x+y))$ can be defined, using an appropriate cut-off function $\alpha \in \mathcal{D}\left(\mathbb{R}^{2n}\right)$.

Let us show that the pairing defines a distribution in $\mathcal{D}'$. It should be clear that it is linear on φ.

As for the continuity, let $\{\varphi_j\}_{j \geq 1}$ be a sequence converging to zero in $\mathcal{D}_K$ for $K \subseteq \mathbb{R}^n$ compact. Let

$$B_l = \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : |(x,y)| < l\}$$

be so that

$$supp\,(T\otimes S)\bigcap\{(x,y)\in\mathbb{R}^n\times\mathbb{R}^n : x+y\in K\}\subseteq B_l.$$

According to Remark 66, we can choose $\alpha\in\mathcal{D}\left(\mathbb{R}^{2n}\right)$ so that $\alpha(x,y)=1$ for $x\in B_{2l}$ and $supp\,(\alpha)\subseteq B_{3l}$.

Then, $\alpha\varphi_j\in\mathcal{D}\left(\mathbb{R}^{2n}\right)$ and we can write $(T\otimes S,\alpha(x,y)\,\varphi_j(x+y))$, for all $j\geq 1$. It should be clear that $\{\alpha\varphi_j\}_{j\geq 1}$ converges to zero in $\mathcal{D}_{\overline{B_{3l}}}\left(\mathbb{R}^{2n}\right)$. Thus,

$$(T\otimes S,\alpha(x,y)\,\varphi_j(x+y))\underset{j\to\infty}{\to}0$$

in $\mathbb{C}$. So, we have proved 1).

To prove 2), let us observe that the set $supp\,(T)+supp\,(S)$ is closed in $\mathbb{R}^n$ because $supp\,(S)$ is closed and $supp\,(T)$ is compact. If

$$U=\mathbb{R}^n\backslash\left(supp\,(T)+supp\,(S)\right),$$

let $\varphi\in\mathcal{D}(U)$ and let $\alpha\in\mathcal{D}\left(\mathbb{R}^{2n}\right)$ be a cut-off function as in 1). We claim that

$$supp\,(T\otimes S)\bigcap supp_{x,y}\left(\alpha(x,y)\,\varphi(x+y)\right)=\varnothing.$$

Indeed, if there is (x_0,y_0) in the intersection, we would have $x_0\in supp\,(T)$, $y_0\in supp\,(S)$ and $x_0+y_0\in supp\,(\varphi)$, which is not possible, according to 2) in Lemma 135.

Then, according to Lemma 41, $(T*S,\varphi)=0$. Thus, we have proved 2).

This completes the proof of the theorem. □

Remark 138.
1. Under the hypotheses of Theorem 137, if we fix a compact set $K\subseteq\mathbb{R}^n$, it should be clear that we can choose the cut-off function $\alpha(x,y)$ independently of $\varphi\in\mathcal{D}_K$.
2. Iterating the process in Theorem 137, we can define the convolution of n distributions, when at least $n-1$ of them have compact support.

Example 139.
1. Given $T\in\mathcal{D}'$,

$$(T*\delta_0,\varphi)=(T_x,(\delta_{0,y},\varphi(x+y)))=(T,\varphi)$$

for $\varphi\in\mathcal{D}$.

2. More generally, given $a\in\mathbb{R}^n$,

$$\begin{aligned}(T*\delta_a,\varphi)&=(T_x,(\delta_{a,y},\varphi(x+y)))=(T,\varphi(\cdot+a))\\ \underset{(1)}{=}(T,\tau_{-a}(\varphi))&=(\tau_a(T),\varphi).\end{aligned}$$

We have used (5.8) in (1).

So, the translation operator, acting on a distribution, can be expressed as a convolution.

3. Since the sum of two compact sets is compact, the convolution product defines a structure of commutative algebra on $\mathcal{E}'$, having δ_0 as the unit.

4. Given $\alpha \in \mathbb{N}^n$ and $T \in \mathcal{D}'$,
$$\partial^\alpha T = (\partial^\alpha T) * \delta_0 = T * (\partial^\alpha \delta_0)\,.$$
Thus, a distributional derivative can be written as a convolution.

5. More generally, if $P(X)$ is a complex polynomial, we can perform the formal substitution
$$X = (X_1, ..., X_n) \to \partial = (\partial_{x_1}, ..., \partial_{x_n})\,.$$
Then, $P(\partial)$ becomes a linear differential operator with constant coefficients, for which
$$P(\partial)\,T = (P(\partial)\,T) * \delta_0 = T * (P(\partial)\,\delta_0)\,.$$

Theorem 140. *(for the proof, see [25], p. 157, Theorem IV and Theorem V)*

1. *The map*
$$(T, S) \to T * S$$
is bilinear and continuous from $\mathcal{E}' \times \mathcal{E}'$ into $\mathcal{E}'$.

2. *The map*
$$(T, S) \to T * S$$
is bilinear and hypocontinuous from $\mathcal{E}' \times \mathcal{D}'$ into $\mathcal{D}'$.

Theorem 141. *Given $T \in \mathcal{E}'$, the map*
$$\begin{array}{ccc} \mathcal{D} & \to & \mathbb{C} \\ \varphi & \to & \left(T, \int_{\mathbb{R}^n} \varphi(\cdot + y)\,dy\right) \end{array} \tag{6.11}$$
defines a distribution in $\mathcal{D}'$. Furthermore,
$$\left(T, \int_{\mathbb{R}^n} \varphi(\cdot + y)\,dy\right) = \int_{\mathbb{R}^n} (T_x, \varphi(x + y))\,dy \tag{6.12}$$
for all $\varphi \in \mathcal{D}$.

Proof. According to Theorem 6 and Theorem 8, the function
$$x \to \int_{\mathbb{R}^n} \varphi(\cdot + y)\,dy$$
is smooth. Thus, (6.11) is well defined. It should be clear that it is linear.

As for the continuity, if $\varphi_j \to 0$ in $\mathcal{D}_K$ as $j \to \infty$ for $K \subseteq \mathbb{R}^n$ compact, let $\theta \in \mathcal{D}$ be a cut-off function that equals one on an open neighborhood of *supp* (T). Moreover, let K_1 be a compact subset of $\mathbb{R}^n$ such that for $x \in supp\,(\theta)$, $supp_y\,(\varphi_j\,(x+y)) \subseteq K_1$ for all $j \geq 1$.

Then, for each $\alpha \in \mathbb{N}^n$, there is $C_{\theta,\alpha} > 0$ so that

$$\begin{aligned}
\left|\partial_x^\alpha \left(\theta\,(x) \int_{\mathbb{R}^n} \varphi_j\,(x+y)\,dy\right)\right| &= \left|\int_{\mathbb{R}^n} \partial_x^\alpha\,(\theta\,(x)\,\varphi_j\,(x+y))\,dy\right| \\
&\leq C_{\theta,\alpha} meas\,(K_1) \sup_{\beta\leq\alpha} \sup_{x\in\mathbb{R}^n} \left|\left(\partial^\beta \varphi_j\right)(x)\right| \underset{j\to\infty}{\to} 0.
\end{aligned}$$

Thus,

$$\left(T, \theta\,(\cdot) \int_{\mathbb{R}^n} \varphi_j\,(\cdot+y)\,dy\right) \underset{j\to\infty}{\to} 0.$$

As for (6.12), according to Theorem 107 and Theorem 8, we can write

$$\begin{aligned}
\left(T, \theta\,(\cdot) \int_{\mathbb{R}^n} \varphi\,(\cdot+y)\,dy\right) &= \left(\sum_\alpha \partial^\alpha T_{f_\alpha}, \theta\,(\cdot) \int_{\mathbb{R}^n} \varphi\,(\cdot+y)\,dy\right) \\
&= \sum_\alpha (-1)^{|\alpha|} \int_{\mathbb{R}^n} f_\alpha\,(x) \left(\int_{\mathbb{R}^n} \partial_x^\alpha\,(\theta\,(x)\,\varphi\,(x+y))\,dy\right) dx \\
&= \sum_\alpha (-1)^{|\alpha|} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f_\alpha\,(x)\,\partial_x^\alpha\,(\theta\,(x)\,\varphi\,(x+y))\,dx\right) dy \\
&= \int_{\mathbb{R}^n} \left(\sum_\alpha (-1)^{|\alpha|}\, T_{f_\alpha(x)}, \partial_x^\alpha\,(\theta\,(x)\,\varphi\,(x+y))\right) dy \\
&= \int_{\mathbb{R}^n} \left(T, \theta\,(\cdot) \int_{\mathbb{R}^n} \varphi\,(\cdot+y)\,dy\right).
\end{aligned}$$

This completes the proof of the theorem. □

So far, we have defined the convolution $T * S$ assuming that at least one of the distributions T and S has compact support. We will see now that $T * S$ can be defined under weaker assumptions.

We begin with the following two results, which are due to János Horváth.

Theorem 142. *([16], p. 383) Let us fix two closed sets $E, F \subseteq \mathbb{R}^n$. Then, the following statements are equivalent:*

1. *For every compact set $K \subseteq \mathbb{R}^n$, the set*

$$(E \times F) \bigcap \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : x+y \in K\} \tag{6.13}$$

is a compact subset of $\mathbb{R}^n \times \mathbb{R}^n$.

2. *For every compact set $K \subseteq \mathbb{R}^n$, the set*

$$E \bigcap (K - F) \tag{6.14}$$

is a compact subset of $\mathbb{R}^n$.

Proof. To prove 1) $\Rightarrow$ 2), let x_0 belong to (6.14). Then, $x_0 \in E$ and there is $y_0 \in F$ so that $x_0 + y_0 \in K$. Therefore (x_0, y_0) belongs to (6.13), which is compact, by hypothesis. There are compact subsets K_1 and K_2 such that the set (6.13) is contained in $K_1 \times K_2$. So, $x_0 \in K_1$, which implies that (6.14) is compact.

Conversely, if (x_0, y_0) belongs to (6.13), we have that $x_0 \in E$, $y_0 \in F$ and $x_0 + y_0 \in K$. Therefore, $x_0 \in E$ and it also belongs to $K - F$ since $x_0 = (x_0 + y_0) - y_0$. On the other hand, $y_0 \in F$ and it also belongs to $K - (E \bigcap (K - F))$ since $y_0 = (x_0 + y_0) - x_0$.

That is to say,

$$\begin{aligned} & (E \times F) \bigcap \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : x + y \in K\} \\ \subseteq \ & \left[E \bigcap (K - F)\right] \times \left[F \bigcap \left(K - \left(E \bigcap (K - F)\right)\right)\right]. \end{aligned} \tag{6.15}$$

Since both sets in the cartesian product (6.15) are compact, we conclude that (6.13) is compact. Thus, 2) $\Rightarrow$ 1).

This completes the proof of the theorem. □

Corollary 143. *([16], p. 384)*

1. *The set $E \bigcap (K - F)$ is the projection of the set (6.13) onto $\mathbb{R}^n_x$.*

2. *The set $F \bigcap (K - E)$ is the projection of the set (6.13) onto $\mathbb{R}^n_y$.*

3. *The set $E \bigcap (K - F)$ is compact if, and only if, the set $F \bigcap (K - E)$ is compact.*

Proof. By definition, if $x_0 \in E \bigcap (K - F)$, there is $y_0 \in F$ so that $(x_0, y_0) \in E \times F$ and $x_0 + y_0 \in K$. Thus, (x_0, y_0) belongs to (6.13).

Conversely, if x_0 belongs to the projection of (6.13) onto $\mathbb{R}^n_x$, there is $y_0 \in \mathbb{R}^n$ so that (x_0, y_0) belongs to (6.13).

Therefore, $x_0 \in E$, $y_0 \in F$ and $x_0 + y_0$ belongs to K. So, $x_0 \in K - F$ also. Thus, we have proved 1).

The proof of 2) is similar, reverting the roles of E and F.

As for 3), if $E \bigcap (K - F)$ is compact, the set (6.13) is compact by Theorem 142, and its projection onto $\mathbb{R}^n_y$ must be compact as well.

Conversely, according to Theorem 142, it suffices to show that if $F \bigcap (K - E)$ is compact, the set (6.13) is compact also.

So, if (x_0, y_0) belongs to (6.28), $y_0 \in F$ and $(x_0 + y_0) \in K$, that is, $y_0 \in K - E$. Moreover, $x_0 \in E$ and it also belongs to $K - (F \bigcap (K - E))$, since $x_0 = (x_0 + y_0) - y_0$. So,

$$\begin{aligned} & (E \times F) \bigcap \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : x + y \in K\} \\ \subseteq \quad & \left[K - \left(F \bigcap (K - E)\right)\right] \times \left[F \bigcap (K - E)\right]. \end{aligned} \tag{6.16}$$

Since both sets in the cartesian product (6.16) are compact, we conclude that (6.13) is compact and therefore, $E \bigcap (K - F)$ is compact.

This completes the proof of the corollary. □

We now use Theorem 142 and Corollary 143 to prove the following result:

Theorem 144. *If T and S belong to $\mathcal{D}'$, let $E = supp\,(T)$, $F = supp\,(S)$.*

Moreover, let us assume that the equivalent conditions 1) and 2) in Theorem 142 are true for these particular choices of E and F.

Then, if we fix $K \subseteq \mathbb{R}^n$ compact, the pairing

$$(T_x \times S_y, \varphi\,(x + y)) \tag{6.17}$$

can be defined, for every $\varphi \in \mathcal{D}_K$, in any of the following three ways: Using an appropriate cut-off function $\alpha\,(x, y) \in \mathcal{D}\left(\mathbb{R}^{2n}\right)$, using an appropriate cut-off function $\theta\,(x) \in \mathcal{D}_x$, or using an appropriate cut-off function $\eta\,(y) \in \mathcal{D}_y$. Moreover, all three definitions coincide.

Proof. By Lemma 65, we know that (6.17) can be defined as

$$(T_x \times S_y, \alpha\,(x, y)\,\varphi\,(x + y))_{\mathcal{D}'_x \times \mathcal{D}'_y, \mathcal{D}_{x,y}}, \tag{6.18}$$

where $\alpha\,(x, y) \in \mathcal{D}\left(\mathbb{R}^{2n}\right)$ is a cut-off function $\alpha\,(x, y) \in \mathcal{D}\left(\mathbb{R}^{2n}\right)$ that equals one on an open neighborhood of the set (6.13).

Moreover, Lemma 65 also shows that the pairing does not depend on the function α satisfying the stated condition.

Next, let us pick a cut-off function $\theta\,(x) \in \mathcal{D}_x$ that equals one on an open neighborhood of the compact set $E \bigcap (K - F)$. We define (6.17) as

$$(T_x \times S_y, \theta\,(x)\,\varphi\,(x + y))_{\mathcal{D}'_x \times \mathcal{D}'_y, \mathcal{D}_{x,y}}. \tag{6.19}$$

Then, (6.19) does not depend on the function $\theta\,(x)$ satisfying the stated condition.

Indeed, if $\theta_1\,(x)$ is another function like $\theta\,(x)$, we claim that

$$(E \times F) \bigcap supp\,[(\theta\,(x) - \theta_1\,(x))\,\varphi\,(x + y)] = \varnothing \tag{6.20}$$

for all $\varphi \in \mathcal{D}_K$.

In fact, if (x_0, y_0) belongs to the set in (6.20), according to 4) in Example 12, x_0 belongs to $supp\,(\theta - \theta_1)$ as well as it belongs to E, while $y_0 \in F$ and $x_0 + y_0 \in K$.

However, since $\theta\,(x) - \theta_1\,(x) = 0$ for x in an open neighborhood of $E \bigcap (K - F)$, x_0 cannot belong to $E \bigcap (K - F)$, which is a contradiction.

Therefore, (6.20) is true and, according to Lemma 41,

$$(T_x \times S_y, (\theta\,(x) - \theta_1\,(x))\,\varphi\,(x+y))_{\mathcal{D}'_x \times \mathcal{D}'_y, \mathcal{D}_{x,y}} = 0$$

for all $\varphi \in \mathcal{D}_K$.

If we pick a cut-off function $\eta\,(y) \in \mathcal{D}_y$ that equals one on an open neighborhood of the compact set $F \bigcap (K - E)$, we define (6.17) as

$$(T_x \times S_y, \eta\,(y)\,\varphi\,(x+y))_{\mathcal{D}'_x \times \mathcal{D}'_y, \mathcal{D}_{x,y}}\,. \tag{6.21}$$

If $\eta_1\,(y)$ is another function like $\eta\,(y)$, we claim that

$$(E \times F) \bigcap supp\,[(\eta\,(y) - \eta_1\,(y))\,\varphi\,(x+y)] = \varnothing. \tag{6.22}$$

for all $\varphi \in \mathcal{D}_K$.

Indeed, if (x_0, y_0) belongs to the set in (6.22), then y_0 belongs to $supp\,(\eta - \eta_1)$ and to F, while $x_0 \in E$ and $x_0 + y_0 \in K$. This means that y_0 belongs to $F \bigcap (K - E)$. However, y_0 cannot belong to $F \bigcap (K - E)$ because $\eta\,(y) - \eta_1\,(y) = 0$ for x in an open neighborhood of $E \bigcap (K - F)$.

Therefore, (6.22) holds and

$$(T_x \times S_y, (\eta\,(y) - \eta_1\,(y))\,\varphi\,(x+y))_{\mathcal{D}'_x \times \mathcal{D}'_y, \mathcal{D}_{x,y}} = 0$$

for all $\varphi \in \mathcal{D}_K$.

Let us see that (6.18) and (6.19) are equal. To that effect, we pick a cut-off function $\alpha\,(x, y) \in \mathcal{D}\left(\mathbb{R}^{2n}\right)$ that equals one on an open neighborhood of the set (6.13). We also pick a cut-off function $\theta\,(x) \in \mathcal{D}_x$ that equals one on an open neighborhood of the compact set $E \bigcap (K - F)$. We claim that

$$(E \times F) \bigcap supp\,[(\alpha\,(x, y) - \theta\,(x))\,\varphi\,(x+y)] = \varnothing. \tag{6.23}$$

Indeed, if (x_0, y_0) belongs to (6.23), then $x_0 \in E$, $y_0 \in F$, $x_0 + y_0 \in K$, and $(x_0, y_0) \in supp\,(\alpha\,(x, y) - \theta\,(x))$. In particular, $(x_0, y_0) \in E \bigcap (K - F)$. But, according to 1) in Corollary 143, $E \bigcap (K - F)$ is the projection of (6.13) onto $\mathbb{R}^n_x$.

Therefore, $\alpha\,(x, y) - \theta\,(x) = 0$ for all (x, y) in an open neighborhood of (x_0, y_0).

So, (x_0, y_0) cannot belong to $supp\,(\alpha\,(x, y) - \theta\,(x))$. Thus, (6.23) holds and

$$(T_x \times S_y, (\alpha\,(x, y) - \theta\,(x))\,\varphi\,(x+y))_{\mathcal{D}'_x \times \mathcal{D}'_y, \mathcal{D}_{x,y}} = 0$$

for all $\varphi \in \mathcal{D}_K$.

Finally,

$$(E \times F) \bigcap supp\,[(\alpha\,(x,y) - \eta\,(y))\,\varphi\,(x+y)] = \varnothing,$$

because, according to 2) in Corollary 143, $F \bigcap (K-E)$ is the projection of (6.13) onto $\mathbb{R}^n_y$.

This completes the proof of the theorem. ☐

Remark 145. Under the hypotheses of Theorem 144, the pairing $(T_x \times S_y, \varphi\,(x+y))$ defined in any of the three equivalent ways, is a distribution in $\mathcal{D}'$, denoted $T * S$.

In fact, an examination of the proof of Theorem 137, will show that it only uses the compactness of the set $supp\,(T \otimes S) \bigcap supp_{x,y}\,(\varphi\,(x+y))$.

Thus, it applies to the present situation, using an appropriate cut-off function $\alpha\,(x,y)$.

Theorem 144 implies that the same conclusion holds, when using any of the other two ways of defining the pairing.

Theorem 146. *Given T and S in $\mathcal{D}'$, let us assume that the equivalent conditions 1) and 2) in Theorem 142 are true for $E = supp\,(T)$, $F = supp\,(S)$. Then,*

1. *The convolution product of T and S is commutative. That is to say,*

$$T * S = S * T.$$

2.
$$supp\,(T * S) \subseteq supp\,(T) + supp\,(S)\,. \tag{6.24}$$

3. *Given $\alpha \in \mathbb{N}^n$,*
$$\partial^\alpha\,(T * S) = (\partial^\alpha T) * S = T * \partial^\alpha S.$$

Proof. To prove 1), we fix a compact subset K of $\mathbb{R}^n$. Let $\theta\,(x)$ be a cut-off function that equals one on an open neighborhood of the compact set $E \bigcap (K-F)$ and let $\eta\,(y)$ be a cut-off function that equals one on an open neighborhood of the compact set $F \bigcap (K-E)$.

Then, if $\varphi \in \mathcal{D}_K$,

$$(T * S, \varphi) \stackrel{def}{=} (T_x \times S_y, \theta\,(x)\,\eta\,(y)\,\varphi\,(x+y))\,.$$

Proof. According to 3) in Theorem 124,

$$\begin{aligned}(T_x \times S_y, \theta\,(x)\,\eta\,(y)\,\varphi\,(x+y)) &= (T_x, \theta\,(x)\,(S_y, \eta\,(y)\,\varphi\,(x+y))) \\ &= (S_y, \eta\,(y)\,(T_x, \theta\,(x)\,\varphi\,(x+y)))\,.\end{aligned}$$

Let us recall that $E \bigcap (K-F)$ is the projection onto $\mathbb{R}^n_x$ of the set (6.13), while $F \bigcap (K-E)$ is the projection onto $\mathbb{R}^n_y$ of the set (6.13).

Thus, $\theta(x)\eta(y)$ is a cut-off function that is equal to one on an open neighborhood of the set (6.13).

Moreover, the set

$$(F \times E) \bigcap \{(y,x) \in \mathbb{R}^n \times \mathbb{R}^n : y + x \in K\}, \tag{6.25}$$

is contained in $(F \bigcap (K - E)) \times E \bigcap (K - F)$.

Therefore, the cut-off function $\eta(y)\theta(x)$ is equal to one on an open neighborhood of the set (6.25). So,

$$\begin{aligned}(S_y, \eta(y)(T_x, \theta(x)\varphi(x+y))) &= (S_y \times T_x, \eta(y)\theta(x)\varphi(y+x)) \\ &\overset{def}{=} (S * T, \varphi)\end{aligned}$$

and we have proved 1).

To prove 2), we begin by showing that the set $E + F$ is closed.

Indeed, if z_0 belongs to the closure of $E + F$, there is a sequence $\{(x_j, y_j)\}_{j\geq 1}$ so that $\{x_j + y_j\}_{j\geq 1}$ converges to z_0 as $j \to \infty$. Let K be the set consisting of z_0 and the points (x_j, y_j) for all $j \geq 1$. The set K is compact.

Therefore, Theorem 142 and Corollary 143 hold for K. In particular, x_j belongs to the compact set $E \bigcap (K - F)$ for all $j \geq 1$. Therefore, there is a subsequence $\{x_{j_k}\}_{k\geq 1}$, converging to some $a \in E$ as $k \to \infty$.

So, $x_{j_k} + y_{j_k}$ converges to z_0 as $j \to \infty$, and consequently, the subsequence $\{y_{j_k}\}_{k\geq 1} \subseteq F$ must converge in F as $k \to \infty$. Since it converges to $z_0 - a$ as $j \to \infty$, we conclude that $z_0 - a \in F$.

That is to say, $z_0 = a + (z_0 - a) \in E + F$. So, we have proved that $E + F$ is closed.

Now we are ready to prove (6.24).

If $x_0 \notin E + F$, there is an open neighborhood U of x_0 so that $U \subseteq \mathbb{R}^n \setminus (E + F)$. If $\varphi \in \mathcal{D}(U)$,

$$(E \times F) \bigcap supp_{x,y}(\varphi(x+y)) = \varnothing.$$

Indeed, if (x_1, y_1) belongs to the intersection, we must have $x_1 \in E$, $y_1 \in F$ and $x_1 + y_1 \in supp(\varphi)$. That is, $x_1 + y_1 \in (E + F) \bigcap (supp(\varphi))$, which is not possible. Thus, we have proved 2).

Finally, according to (6.6),

$$\begin{aligned}((\partial^\alpha T) * S, \varphi) &= ((\partial^\alpha T)_x \otimes S_y, \varphi(x+y)) = \left(T_x \otimes (\partial^\alpha S)_y, \varphi(x+y)\right) \\ &= (\partial^\alpha_x (T_x \otimes S_y), \varphi(x+y)),\end{aligned}$$

for all $\varphi \in \mathcal{D}$. □

This completes the proof of the Theorem. □

Remark 147. As shown in Theorem 137, if at least one of $T, S \in \mathcal{D}'$ has compact support, given $\varphi \in \mathcal{D}$, the condition

$$(supp\,(T) \times supp\,(S)) \bigcap supp_{x,y}\,(\varphi\,(x+y)) \text{ is compact,} \tag{6.26}$$

holds.

However, there are other instances when (6.26) is true. For example, let us assume that $supp\,(T)$ and $supp\,(S)$ are both contained in the region

$$\Gamma_a = \{x \in \mathbb{R}^n : x_1 \geq a_1, ..., x_n \geq a_n\}$$

for $a_1, ..., a_n$ fixed.

Then, for $\varphi \in \mathcal{D}$, condition (6.26) holds. In fact, since the set in (6.26) is always closed, we only need to prove that it is bounded.

If (x, y) belongs to the set in (6.26), then $x_j, y_j \geq a_j$ for all j. Moreover, there is $C = C_{supp(\varphi)} > 0$ so that $|x_j + y_j| \leq C$ for all j. Therefore, $a_j \leq x_j, y_j \leq C - a_j$ for all j.

The distribution T_{H_n} associated with the n-dimensional Heaviside function H_n defined in 6) of Example 96, is an example of this situation.

Remark 148. Theorem 144, Remark 145, Theorem 146, and Remark 147 can be adapted, in an obvious way, to distributions in $\mathcal{D}'^{(m)}$.

Remark 149. An obvious extension of (6.13) to l closed sets, would allow us to extend Theorem 144, Remark 145, Theorem 146, and Remark 147, to l distributions $T_1, ..., T_l \in \mathcal{D}'$, under the appropriate assumption.

The definition

$$(T_1 * ... * T_l) \overset{def}{=} (T_{1,a} \otimes ... \otimes T_{l,z}, \varphi\,(a + ... + z))$$

produces an associative operation ([16], p. 390, Proposition 8).

However, let us consider, on $\mathbb{R}$, the distributions T_1, $\frac{d\delta_0}{dx}$, and T_{H_1}, where T_1 is the distribution defined by the function identically equal to one and H_1 is the one-dimensional Heaviside function. Then

$$\left(T_1 * \frac{d\delta_0}{dx}\right) * T_{H_1} = \frac{dT_1}{dx} * T_{H_1} = 0 * T_{H_1} = 0,$$

while

$$T_1 * \left(\frac{d\delta_0}{dx} * T_{H_1}\right) = T_1 * \delta_0 = T_1.$$

Thus, the fact that in each parenthesis one of the distributions has compact support, is not enough to guarantee associativity.

Let us observe that the natural extension of condition 1) in Theorem 142, to the supports of these three distributions, is that the set

$$(\mathbb{R} \times \{0\} \times [0, \infty)) \bigcap \left\{(x, y, z) \in \mathbb{R}^3 : x + y + z \in K\right\}$$

is compact for each $K \subseteq \mathbb{R}$ compact, which is not true.

6.3 Multiplicative product

It would be natural to try to define the multiplication $T \cdot S$ of two distributions $T, S \in \mathcal{D}'$ in such a way that $T_f \cdot S_g = T_{fg}$, when f and g are locally integrable functions.

Unfortunately, we can see very quickly that this approach does not work, since the multiplicative product of two locally integrable functions is not always locally integrable, so it does not always define a distribution.

However, the multiplicative product of a locally integrable function by, say, a smooth function, will always be locally integrable. So, we can attempt the following approach:

Given a distribution $T \in \mathcal{D}'$ and a distribution T_α defined by a function $\alpha \in \mathcal{E}$, we set

$$(T_\alpha \cdot T, \varphi) \overset{def}{=} (T, \alpha\varphi),$$

for $\varphi \in \mathcal{D}$.

It is customary to say that we multiply T by the function α, denoting the result as αT or, equally, $T\alpha$.

The following result shows that the proposed definition of αT is correct.

Theorem 150. *1. For $\alpha \in \mathcal{E}$ fixed, the map*

$$\begin{array}{ccc} \mathcal{D} & \to & \mathcal{D} \\ \varphi & \to & \alpha\varphi \end{array}$$

is well defined, linear, and continuous.

2. The map

$$\begin{array}{ccc} \mathcal{D} & \to & \mathbb{C} \\ \varphi & \to & (T, \alpha\varphi) \end{array}$$

is a distribution in $\mathcal{D}'$.

Proof. It should be clear that the map in 1) is well defined and it linear.

As for the continuity, if we fix $K \subseteq \mathbb{R}^n$ compact and we fix an n-tuple $\beta \in \mathbb{N}^n$, according to Lemma 4,

$$\sup_{x \in \mathbb{R}^n} \left| \partial^\beta (\alpha\varphi)(x) \right| \leq C_\beta \sup_{\gamma \leq \beta} \sup_{x \in K} \left| (\partial^\gamma \alpha)(x) \right| \sup_{\gamma \leq \beta} \sup_{x \in \mathbb{R}^n} \left| (\partial^\gamma \varphi)(x) \right|, \tag{6.27}$$

for all $\varphi \in \mathcal{D}_K$.

Therefore, 1) is true.

It should be clear that the map in 2) is linear, so let us prove that it is continuous.

If $\{\varphi_j\}_{j \geq 1}$ converges to φ in $\mathcal{D}$ as $j \to \infty$, according to 1), the sequence $\{\alpha\varphi_j\}_{j \geq 1}$ converges to $\alpha\varphi$ in $\mathcal{D}$ as $j \to \infty$. So, we have proved 2).

This completes the proof of the theorem. ☐

Remark 151. Since the map

$$\begin{array}{ccc} \mathcal{E}\times\mathcal{D} & \to & \mathcal{D} \\ (\alpha,\varphi) & \to & \alpha\varphi \end{array} \tag{6.28}$$

is bilinear, an extension of the estimate (6.27) will show that the map (6.28) is hypocontinuous, according to Definition 118.

In fact, if we fix $K\subseteq\mathbb{R}^n$ compact and we fix an n-tuple $\beta\in\mathbb{N}^n$,

$$\begin{aligned} \left|\partial^{\beta}(\alpha\varphi)(x)\right| &\leq C_{\beta}\left|(\partial^{\gamma}\alpha)(x)\right|\left|(\partial^{\gamma}\varphi)(x)\right| \\ &\leq C_{\beta}\sup_{|\gamma|\leq|\beta|}\sup_{x\in B_k}\left|(\partial^{\gamma}\alpha)(x)\right|\sup_{\gamma\leq\beta}\sup_{x\in\mathbb{R}^n}\left|(\partial^{\gamma}\varphi)(x)\right| \\ &= C_{\beta}\left\|\alpha\right\|_{|\beta|,B_k}\sup_{\gamma\leq\beta}\sup_{x\in\mathbb{R}^n}\left|(\partial^{\gamma}\varphi)(x)\right| \end{aligned}$$

for all $\varphi\in\mathcal{D}_K$, where $\left\|\alpha\right\|_{|\beta|,B_k}$ is defined as in (4.1).

Corollary 152. *1. If $\alpha\in\mathcal{D}$, or if $T\in\mathcal{E}'$, the distribution αT belongs to $\mathcal{E}'$ and*

$$supp\,(\alpha T)\subseteq supp\,(\alpha)\bigcap supp\,(T)\,.$$

2. The multiplicative product of l distributions of which at least $l-1$ are functions in $\mathcal{E}$, is well defined and it is associative and commutative.

Proof. To prove 1), we fix x_0 in $\mathbb{R}^n\backslash\left(supp\,(\alpha)\bigcap supp\,(T)\right)$. Then, there is an open neighborhood U of x_0 that does not meet $\left(supp\,(\alpha)\bigcap supp\,(T)\right)$.

Therefore, if $\varphi\in\mathcal{D}(U)$,

$$supp\,(\alpha\varphi)\bigcap supp\,(T)\subseteq supp\,(\varphi)\bigcap supp\,(\alpha)\bigcap supp\,(T)=\varnothing.$$

So, $(T,\alpha\varphi)=0$.

The proof of 2) follows from the fact that the multiplicative product of complex-valued functions is associative and commutative.

This completes the proof of the corollary. □

Remark 153. The crucial point in defining the multiplicative product is that a test function φ, multiplied by α, still has to be a test function.

Therefore, there must be a balance between properties of α and properties of φ.

So, it should be clear that Theorem 150 and Corollary 152 apply, with the obvious modifications, if $T\in\mathcal{D}^{(m)\prime}$, and $\alpha\in\mathcal{E}^{(m)}$ as given by Definition 120.

Theorem 154. *(for the proof, see [25], p. 119, Theorem III) The multiplicative product is a bilinear and hypocontinuous map from $\mathcal{E}\times\mathcal{D}'$ into $\mathcal{D}'$ and from $\mathcal{E}^{(m)}\times\mathcal{D}^{(m)\prime}$ into $\mathcal{D}^{(m)\prime}$.*

Example 155. As an illustration, we collect here a few additional properties of the multiplicative product, plus a few examples.

1. The partial derivatives of a multiplicative product are calculated with the usual rule. That is, in whatever duality the product αT is defined,

$$\partial^{\beta}(\alpha T)=\sum_{\gamma=0}^{\beta}\binom{\beta}{\gamma}(\partial^{\gamma}\alpha)\,\partial^{\beta-\gamma}T. \tag{6.29}$$

   Indeed, if $\partial^{\beta}=\partial_{x_j}$ for some $1\le j\le n$,

$$\begin{aligned}\left(\partial_{x_j}(\alpha T),\varphi\right) &= -\left(T,\alpha\partial_{x_j}\varphi\right)=-\left(T,\partial_{x_j}(\alpha\varphi)-\left(\partial_{x_j}\alpha\right)\varphi\right)\\ &= \left(\alpha\partial_{x_j}T,\varphi\right)+\left(\left(\partial_{x_j}\alpha\right)T,\varphi\right),\end{aligned}$$

   while

$$\begin{aligned}\partial_{x_j}\partial^{\beta}(\alpha T) &= \sum_{\gamma=0}^{\beta}\binom{\beta}{\gamma}\partial_{x_j}\left((\partial^{\gamma}\alpha)\,\partial^{\beta-\gamma}T\right)\\ & \sum_{\gamma=0}^{\beta}\binom{\beta}{\gamma}\left(\left(\partial_{x_j}\partial^{\gamma}\alpha\right)\partial^{\beta-\gamma}T+(\partial^{\gamma}\alpha)\,\partial_{x_j}\partial^{\beta-\gamma}T\right).\end{aligned}$$

   Since the exact same calculation done in Lemma 4) applies here, we obtain (6.29).

2. Given $\alpha\in\mathcal{E}_x$, $\beta\in\mathcal{E}_y$, $T\in\mathcal{D}'_x$ and $S\in\mathcal{D}'_y$,

$$(\alpha(x)\otimes\beta(y))\,(T_x\otimes S_y)=(\alpha T)_x\otimes(\beta S)_y\,.$$

   According to Theorem 122, to verify this identity we only need to check it on test functions of the form $\varphi_1(x)\,\varphi_2(y)$.

$$\begin{aligned}\left((\alpha T)_x\otimes(\beta S)_y\,,\varphi_1(x)\,\varphi_2(y)\right) &= \left((\alpha T)_x\,\varphi_1(x)\right)\left((\beta S)_y\,,\varphi_2(y)\right)\\ & (T_x,\alpha(x)\,\varphi_1(x))\,(S_y,\beta(y)\,\varphi_2(y))\\ &= (T_x\otimes S_y,(\alpha(x)\otimes\beta(y))\,\varphi_1(x)\,\varphi_2(y))\\ &= ((\alpha(x)\otimes\beta(y))\,(T_x\otimes S_y)\,,\varphi_1(x)\,\varphi_2(y))\,.\end{aligned}$$

3. Given $\alpha\in\mathcal{E}$,

$$\alpha\delta_0=\alpha(0)\,\delta_0.$$

   Therefore,

$$\partial^{\beta}(\alpha\delta_0)=\alpha(0)\,\partial^{\beta}(\delta_0)\,,$$

   for all $\beta\in\mathbb{N}^n$.

4. For $l, k \in \mathbb{N}$ and $\varphi \in \mathcal{D}'(\mathbb{R})$,

$$\begin{aligned}\left(x^l \delta_0^{(k)}, \varphi\right) &= (-1)^k \left(x^l \varphi\right)^{(k)}(0) \\ &= (-1)^k \sum_{s=0}^{k} \binom{l}{s} \left(x^l\right)^{(s)}(0)\, \varphi^{(k-s)}(0) \\ &= \begin{cases} 0 & \text{if} \quad l > k \\ (-1)^k\, l! \varphi^{(k-l)}(0) & \text{if} \quad l \leq k \end{cases}.\end{aligned}$$

Therefore,

$$x^l \delta_0^{(k)} = \begin{cases} 0 & \text{if} \quad l > k \\ (-1)^l\, l!\, \delta_0^{(k-l)} & \text{if} \quad l \leq k \end{cases}.$$

5.
$$x\left(pv \frac{1}{x}\right) = 1. \tag{6.30}$$

Indeed, given $\varphi \in \mathcal{D}$,

$$\begin{aligned}\left(x\left(pv \frac{1}{x}\right), \varphi\right) &= \lim_{j \to \infty} \int_{|x| > 1/j} \frac{x \varphi(x)}{x} dx \\ &= \int_{\mathbb{R}^n} \varphi(x)\, dx = (T_1, \varphi(x)),\end{aligned}$$

where T_1 denotes the distribution defined by the function identically equal to one.

Actually, $pv\, 1/x$ is the only odd distribution that satisfies (6.30).

In fact, if there are two distributions, T and S, so that $x(T - S) = 0$, then $T - S$ must be concentrated on $\{0\}$. Indeed, if $\varphi \in \mathcal{D}(\mathbb{R} \backslash \{0\})$,

$$(T - S, \varphi) = \left(x(T - S), \frac{\varphi}{x}\right) = 0.$$

According to Theorem 110,

$$T - S = \sum_{k=0}^{m} c_k \delta_0^{(k)}.$$

Then,

$$0 = x(T - S) = \sum_{k=0}^{m} c_k\, x \delta_0^{(k)} = -\sum_{k=1}^{m} c_k \delta_0^{(k-1)}.$$

If $\varphi \in \mathcal{D}(\mathbb{R})$ is equal to one on a neighborhood of zero

$$\left(-\sum_{k=1}^{m} c_k \delta_0^{(k-1)}, x^l \varphi \right) = c_{l+1},$$

for $l = 0, ..., m-1$. Therefore, $c_k = 0$ for $1 \leq k \leq m$, or $T - S = c_0 \delta_0$. If T and S are both odd, since it should be clear that δ_0 is an even distribution, the constant c_0 must be zero.

Theorem 156. *Given $T \in \mathcal{D}'$ and $\alpha \in \mathcal{E}$, let us assume that $supp\,(T)$ and $supp\,(T_\alpha) = supp\,(\alpha)$ satisfy the equivalent conditions 1) and 2) in Theorem 142.*

Then,

1. *The pairing $(T * T_\alpha, \varphi)$ is well defined as a distribution in $\mathcal{D}'$, for every $\varphi \in \mathcal{D}$.*
2. *The pairing $(T_x, \alpha\,(y - x))$ can be defined for each $y \in \mathbb{R}^n$ fixed.*
3. *The function*
$$y \to \beta\,(y) = (T_x, \alpha\,(y - x))$$
belongs to $\mathcal{E}$.
4. $$T * T_\alpha = T_\beta$$
on $\mathcal{D}$.

Proof. The set

$$supp\,(T) \bigcap (K - supp\,(\alpha))$$

is compact for every $K \subseteq \mathbb{R}^n$. Therefore, Theorem 144 tells us that we can define the pairing $(T_x \otimes S_y, \varphi\,(x + y))$ as

$$\left(T_x \otimes T_{\alpha(y)}, \theta\,(x)\,\varphi\,(x + y)\right)_{\mathcal{D}'_{x,y}, \mathcal{D}_{x,y}},$$

where $\theta \in \mathcal{D}$ is a cut-off function equal to one on an open neighborhood of $supp\,(T) \bigcap (K - supp\,(\alpha))$.

So, 1) is true.

Then,

$$\begin{aligned}
\left(T_x \otimes T_{\alpha(y)}, \theta\,(x)\,\varphi\,(x + y)\right)_{\mathcal{D}'_{x,y}, \mathcal{D}_{x,y}} &= \left(T_x, \theta\,(x) \left(\int_{\mathbb{R}^n} \alpha\,(y - x)\,\varphi\,(y)\,dy\right)\right) \\
&\quad \left((\theta T)_x, \left(\int_{\mathbb{R}^n} \alpha\,(y - x)\,\varphi\,(y)\,dy\right)\right) \\
&\underset{(1)}{=} \int_{\mathbb{R}^n} \left((\theta T)_x, \alpha\,(y - x)\right) \varphi\,(y)\,dy \\
&= \int_{\mathbb{R}^n} (T_x, \theta\,(x)\,\alpha\,(y - x))\,\varphi\,(y)\,dy \qquad (6.31)
\end{aligned}$$

where, in (1), we have used Theorem 141 for the distribution θT.

Thus, 2) is true as well.

To prove 3), let us first observe that given y_0 fixed, the cut-off function $\theta(x)$ can be defined independently of y in an open neighborhood of y_0. Therefore, we can prove that the function $\beta(y)$, defined as

$$\beta(y) = (T_x, \theta(x)\,\alpha(y-x))$$

belongs to $\mathcal{E}$ using the work done in 1) of Theorem 124.

Finally, according to (6.31),

$$\begin{aligned}(T * T_\alpha, \varphi) &= \int_{\mathbb{R}^n} (T_x, \theta(x)\,\alpha(y-x))\,\varphi(y)\,dy \\ &= \int_{\mathbb{R}^n} \beta(y)\,\varphi(y)\,dy.\end{aligned}$$

So, we have 4).

This completes the proof of the theorem. □

Remark 157. Theorem 156 is the distributional version of Lemma 24. The function $\beta(y)$ is called a regularization, or smoothing, of the distribution T.

Let us observe that defining the pairings $(T * T_\alpha, \varphi)$ and $(T_x, \alpha(y-x))$ using a cut-off function that only depends on x, simplifies the proof of Theorem 156.

Remark 158. Given $\psi \in \mathcal{D}(\mathbb{R})$, the convolution $(pv\,1/x) * T_\varphi$ is well defined. Moreover, it can be written as

$$\lim_{j\to\infty} \int_{|x-y|\geq 1/j} \frac{\psi(y)}{x-y}\,dy, \tag{6.32}$$

for each $x \in \mathbb{R}$.

Indeed,

$$\begin{aligned}\left(\left(pv\frac{1}{x}\right) * T_\psi, \varphi\right) &= \left(pv\frac{1}{x}, \int_{\mathbb{R}} \psi(y)\,\varphi(x+y)\,dy\right) \\ &= \left(pv\frac{1}{x}, \int_{\mathbb{R}} \psi(z-x)\,\varphi(z)\,dz\right) \\ &\underset{(1)}{=} \int_{\mathbb{R}} \left(pv\frac{1}{x}, \psi(z-x)\right) \varphi(z)\,dz,\end{aligned}$$

for every $\varphi \in \mathcal{D}(\mathbb{R})$, where we have used in (1) Theorem 156.

The formula (6.32) can be extended to ψ in $L^p(\mathbb{R})$ for $1 < p < \infty$, defining a linear and continuous operator from $L^p(\mathbb{R})$ into itself. This operator is called the Hilbert transform after the mathematician David Hilbert (1886-1943).

The Hilbert transform plays a fundamental role in harmonic analysis. Among other things, it is a foremost example of a class of operators called singular integrals, defined and studied by Alberto P. Calderón (1920-1998) and Antoni Zygmund (1900-1992).

Moreover, (6.32) can be written as

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\varphi(x)\psi(y)}{x-y} dx,$$

when $(supp\,(\varphi(x)\psi(y))) \bigcap \{(x,y) \in \mathbb{R}^2 : x = y\} = \varnothing$.

This formula is just a tiny example of a spectacular result due to Schwartz [23], that basically says that for every linear and continuous operator $\mathcal{T}$ from $\mathcal{D}'_y$ into $\mathcal{D}'_x$, there is a distribution $T \in \mathcal{D}'_{x,y}$ such that

$$(\mathcal{T}(\psi), \varphi)_{\mathcal{D}'_x,\mathcal{D}_x} = (T, \varphi \otimes \psi)_{\mathcal{D}'_{x,y},\mathcal{D}_{x,y}}.$$

Roughly stated, in the sense of distributions, every linear and continuous operator is an "integral" operator.

We saw an example of a classical integral operator in Remark 93.

The article [1] gives an overview of these matters. For an in depth presentation, we refer to the excellent original sources cited therein.

Theorem 159. *We have the following continuous and dense inclusions*

$$\begin{array}{ccc} \mathcal{E} & \hookrightarrow & \mathcal{D}' \\ \mathcal{D} & \hookrightarrow & \mathcal{E}' \end{array}.$$

Proof. It should be clear that the inclusions are well defined by means of identifying α and T_α.

If $\{\alpha_j\}_{j\geq 1}$ converges to zero in $\mathcal{E}$ and $\varphi \in \mathcal{B}$, a bounded subset of $\mathcal{D}_K$ for $K \subseteq \mathbb{R}^n$ compact,

$$\left| \int_{\mathbb{R}^n} \alpha_j(x)\varphi(x)\,dx \right| \leq meas\,(K) \sup_{\varphi\in\mathcal{B}} \sup_{x\in\mathbb{R}^n} |\varphi(x)| \sup_{x\in K} |\alpha_j(x)| \underset{j\to\infty}{\to} 0.$$

As for the density of the inclusion, given $T \in \mathcal{D}'$, we consider the sequence $\{T * T_{\rho_j}\}_{j\geq 1}$.

Let us observe that, for $K \subseteq \mathbb{R}^n$ fixed, the set

$$(supp\,(T) \times supp\,(T_{\rho_j})) \bigcap \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : x + y \in K\}$$

is contained in a compact set that does not depend on j, for all $j \geq 1$. Therefore, the cut-off function θ can be chosen independently of j.

According to Theorem 156, $T * T_{\rho_j} = T_{\beta_j}$ with $\beta_j \in \mathcal{E}$.

Theorem 68 and 2) in Theorem 140, imply that the sequence $\left\{T_{\beta_j}\right\}_{j\geq 1}$ converges to $T * \delta_0$ in $\mathcal{D}'$ as $j \to \infty$.

Finally, $T * \delta_0 = T$, as shown in 1) of Example 139.

Since the space $\mathcal{D}$ is the inductive limit of the spaces $\mathcal{D}_K$, the inclusion $\mathcal{D} \to \mathcal{E}'$ will be continuous if, and only if, ([26], p. 268, Corollary 2) the inclusion $\mathcal{D}_K \to \mathcal{E}'$ is continuous, for all $K \subseteq \mathbb{R}^n$ compact.

Let $\{\alpha_j\}_{j\geq 1}$ be a sequence converging to zero in $\mathcal{D}_K$ as $j \to \infty$. Given $\mathcal{B}$, a bounded subset of $\mathcal{E}$,

$$\left| \int_{\mathbb{R}^n} \alpha_j (x) \varphi (x) \, dx \right| \leq meas (K) \sup_{\varphi \in \mathcal{B}} \sup_{x \in K} |\varphi (x)| \sup_{x \in \mathbb{R}^n} |\alpha_j (x)| \underset{j \to \infty}{\to} 0.$$

Let us show that the inclusion is dense.

As before, we consider the sequence $\left\{T * T_{\rho_j}\right\}_{j\geq 1}$. In this case, the function β defined in 3) of Theorem 156, belongs to $\mathcal{D}$.

Theorem 68 and 1) in Theorem 140, imply that the sequence $\left\{T_{\beta_j}\right\}_{j\geq 1}$ converges to $T * \delta_0 = T$ in $\mathcal{E}'$ as $j \to \infty$.

This completes the proof of the theorem. □

Remark 160. According to Theorem 53, $\mathcal{D}$ is dense in $\mathcal{E}$. Therefore, we can say that $\mathcal{D}$ is dense in $\mathcal{D}'$.

So far, we have worked with factors that are, except for one, smooth functions, or functions that are continuous and have a certain number of continuous partial derivatives. Can we do better than this?

The first answer, for the negative, was given by Schwartz ([24]). Roughly speaking, Schwartz's impossibility result, in one of its versions, says the following:

It is not possible to define on $\mathcal{D}' \times \mathcal{D}'$ an associative and bilinear operation $\cdot$, not necessarily commutative, so that

1. The unit is the distribution T_1 defined by the function identically equal to one.

2. The operation $\cdot$ extends the multiplicative product of distributions T_f and T_g defined by continuous functions f and g. That is,

$$T_f \cdot T_g = T_{fg}.$$

3. The operation $\cdot$ extends the standard multiplicative products

$$\begin{aligned} \left(pv \, \frac{1}{x}\right) x &= 1, \\ x\delta_0 &= 0. \end{aligned}$$

Indeed, if such operation existed, it would have to satisfy, in particular,

$$\begin{aligned} 0 &= \left(pv\,\frac{1}{x}\right)\cdot(x\delta_0) \\ &= \left(\left(pv\,\frac{1}{x}\right)x\right)\cdot\delta_0 \\ &= T_1\cdot\delta_0=\delta_0, \end{aligned}$$

which is not possible.

Schwartz also observed that it is not possible, either, to have an associative bilinear operation $\cdot$ satisfying 1), 2), and Leibniz's rule

$$\partial_{x_j}\left(T\cdot S\right)=\left(\partial_{x_j}T\right)\cdot S+T\cdot\left(\partial_{x_j}S\right).$$

Therefore, a multiplicative product and a notion of partial derivative cannot coexist with the Dirac distribution.

In [6], Jean-François Colombeau defined a multiplicative product on $\mathcal{D}'$ by immersing $\mathcal{D}'$ in an associative quotient algebra. The multiplicative operation on the algebra extends the multiplicative product $T_fT_g = T_{fg}$ when f and g belong to $\mathcal{E}$. However, it does not extend the multiplicative product when f and g are only continuous.

There are also multiplicative products defined, at least in special cases, by means of regularizations [2].

The subject of how to multiply distributions, leads naturally to the subject of how to divide distributions (see [25], Chapter V, Sections 4 and 5). We will briefly touch upon it, later on.

In this section, we have not considered the spaces $\mathcal{S}$ and $\mathcal{S}'$. Indeed, their relationship with the convolution product and the multiplicative product is so special, that it merits to have its own section.

In the meantime, we are pretty much done with the basics of distribution theory which we set to present.

For much of the remainder of this article, we will study the Fourier transform of distributions, as a fundamental application of all we have seen so far.

We begin with a refresher section that has definitions and properties pertaining to the Fourier transform on the spaces L^1 and L^2. The material we are about to discuss, besides being of great interest on its own, will prove useful when we attempt to define the Fourier transform of a distribution.

Let us mention that the Fourier transform is named after the mathematician Joseph Fourier (1768-1830).

7 The Fourier transform on L^1 and L^2 (part I)

We start with a definition.

Definition 161. *Given $f \in L^1$, we define the Fourier transform of f, denoted $\widehat{f}$ or $\mathcal{F}[f]$, as*

$$\mathcal{F}[f](\xi) = \int_{\mathbb{R}^n} e^{2\pi i x\xi} f(x)\, dx, \tag{7.1}$$

where $x\xi = x_1\xi_1 + ... + x_n\xi_n$.

It should be clear that the integral in (7.1) is well defined.

Let us recall that given $h \in \mathbb{R}^n$, we defined the translation operator τ_h as $\tau_h(f)(x) = f(x-h)$.

The proof of the following result reduces to changing variables in (7.1).

Theorem 162. *Given $f \in L^1$, $h \in \mathbb{R}^n$ and $k \in \mathbb{R}$, $k \neq 0$,*

$$\begin{aligned} \mathcal{F}[\tau_h(f)](\xi) &= e^{2\pi i h\xi}\mathcal{F}[f](\xi), \\ \mathcal{F}[f(k\cdot)](\xi) &= \frac{1}{|k|^n}\mathcal{F}[f]\left(\frac{\xi}{k}\right). \end{aligned} \tag{7.2}$$

Remark 163. As an immediate consequence of (7.2), the Fourier transform of an odd integrable function is odd, while the Fourier transform of an even integrable function is even.

Moreover, if $f \in L^1$ is odd,

$$\begin{aligned} \mathcal{F}[f](\xi) &= \int_{\mathbb{R}^n} e^{2\pi i x\xi} f(x)\, dx = \int_{\mathbb{R}^n} \frac{e^{2\pi i x\xi} - e^{-2\pi i x\xi}}{2} f(x)\, dx \\ &= i\int_{\mathbb{R}^n} f(x) \sin 2\pi x\xi\, dx. \end{aligned}$$

while, if $f \in L^1$ is even,

$$\begin{aligned} \mathcal{F}[f](\xi) &= \int_{\mathbb{R}^n} e^{2\pi i x\xi} f(x)\, dx = \int_{\mathbb{R}^n} \frac{e^{2\pi i x\xi} + e^{-2\pi i x\xi}}{2} f(x)\, dx \\ &= \int_{\mathbb{R}^n} f(x) \cos 2\pi x\xi\, dx. \end{aligned}$$

Definition 164. *We say that a function $f : \mathbb{R}^n \to \mathbb{C}$ vanishes at infinity if there is*

$$\lim_{|x|\to\infty} f(x) = 0. \tag{7.3}$$

It should be clear that all the functions in $\mathcal{D}$ or in $\mathcal{S}$, vanish at infinity.

Definition 165.

$$\mathcal{C}_0 = \{f : \mathbb{R}^n \to \mathbb{C} : f \text{ is continuous and vanishes at infinity}\}.$$

The space C_0 is a complex linear space, that becomes a Banach space with the sup norm.

Lemma 166. *The following statements hold:*

1. *Every function in C_0 is uniformly continuous.*
2. *The space $\mathcal{D}$ is dense in C_0.*

Proof. We begin with 1).

If $f \in \mathcal{C}_0$, let us fix $\varepsilon > 0$. Then, there is $r = r_\varepsilon > 0$ so that $|f(x)| \leq \varepsilon$ for $|x| > r$. On the other hand, the function f is uniformly continuous on the compact set $B_r = \{x \in \mathbb{R}^n : |x| \leq r\}$. Therefore, if $x, y \in B_r$, there is $\delta = \delta_\varepsilon > 0$ such that

$$|f(x) - f(y)| \leq \varepsilon$$

if $|x - y| < \delta$.

Furthermore,

$$|f(x) - f(y)| \leq |f(x)| + |f(y)| \leq 2\varepsilon$$

if $|x|, |y| > r$.

So, we have proved 1).

To prove 2), we fix $f \in \mathcal{C}_0$. Given $\varepsilon > 0$ there is $r > 0$ so that $|f(x)| \leq \varepsilon$ for $|x| > r$.

Let $\alpha \in \mathcal{D}$ be a cut-off function that is equal to one on $B_r = \{x \in \mathbb{R}^n : |x| \leq r\}$, it is zero for $|x| \geq 2r$, and it is between 0 and 1 everywhere. Then,

$$|(1 - \alpha(x)) f(x)| \leq \varepsilon$$

for all $x \in \mathbb{R}^n$.

Let $f_j = (\alpha f) * \rho_j$, where ρ_j is the function used in the proof of Theorem 25. It should be clear, from Lemma 24, that $f_j \in \mathcal{D}$.

Thus, to prove that the sequence $\{f_j\}_{j\geq 1}$ converges to f in $\mathcal{C}_0$, it suffices to show that $f_j(x) \to \alpha(x) f(x)$ uniformly with respect to x in $\mathbb{R}^n$ as $j \to \infty$.

$$\begin{aligned} |f_j(x) - \alpha(x) f(x)| &= \left| \int_{|y|\leq 1/j} \alpha(x-y) f(x-y) \rho_j(y)\, dy - \alpha(x) f(x) \int_{|y|\leq 1/j} \rho_j(y)\, dy \right| \\ &\leq \int_{|y|\leq 1/j} |\alpha(x-y) f(x-y) - \varphi(x)| \rho_j(y)\, dy. \end{aligned}$$

Since αf is uniformly continuous, we can say that given $\varepsilon > 0$, there is $j_0 = j_{0,\varepsilon} \geq 1$ so that

$$\left|\alpha\left(x-y\right) f\left(x-y\right)-\alpha\left(x\right) f\left(x\right)\right| \leq \varepsilon,$$

when $|y| \leq 1/j$ for $j \geq j_0$, for all $x \in \mathbb{R}^n$.

So, $\left|f_j\left(x\right)-\alpha\left(x\right) f\left(x\right)\right| \leq \varepsilon$ for all $x \in \mathbb{R}^n$. Finally,

$$\left|f_j\left(x\right)-f\left(x\right)\right| \leq \left|f_j\left(x\right)-\alpha\left(x\right) f\left(x\right)\right|+\left|\left(1-\alpha\left(x\right)\right) f\left(x\right)\right| \leq 2\varepsilon$$

for $j \geq j_0$ and for all $x \in \mathbb{R}^n$.

This completes the proof of the lemma. □

Theorem 167. 1. *The Fourier transform $\mathcal{F}$ is a linear and continuous operator from L^1 into $\mathcal{C}_0$.*

2.
$$\|\mathcal{F}\|_{L^1,\mathcal{C}_0}=1,$$
where $\|\cdot\|_{L^1,\mathcal{C}_0}$ indicates the operator norm.

3. *The image $\mathcal{F}\left[L^1\right]$ is a proper linear subspace of $\mathcal{C}_0$.*

4. *The image $\mathcal{F}\left[L^1\right]$ is a dense linear subspace of $\mathcal{C}_0$.*

Proof. To prove 1), let us begin by proving that given $f \in L^1$, $\mathcal{F}[f]$ is a continuous function. We fix $\xi_0 \in \mathbb{R}^n$ and let $\{\xi_j\}_{j\geq 1}$ be a sequence converging to ξ_0 in $\mathbb{R}^n$ as $j \to \infty$.

Then, the sequence $\left\{\int_{\mathbb{R}^n} e^{2\pi i x\xi_j} f\left(x\right)\right\}_{j\geq 1}$ converges pointwise to $e^{2\pi i x\xi_0} f\left(x\right)$ in $\mathbb{C}$ as $j \to \infty$ and $\left|e^{2\pi i x\xi_j} f\left(x\right)\right| = \left|f\left(x\right)\right|$ for every $j \geq 1$.

Therefore, Lebesgue's dominated convergence theorem implies that $\left\{\mathcal{F}\left[f\right]\left(\xi_j\right)\right\}_{j\geq 1}$ converges to $\mathcal{F}\left[f\right]\left(\xi\right)$ in $\mathbb{C}$ as $j \to \infty$. So, $\mathcal{F}[f]$ is continuous on $\mathbb{R}^n$.

To prove (7.3), we begin with a few observations.

First of all,

$$\left|\mathcal{F}\left[f\right]\left(\xi\right)\right| \leq \|f\|_{L^1} \tag{7.4}$$

for all $\xi \in \mathbb{R}^n$, so $\mathcal{F}[f]$ is bounded on $\mathbb{R}^n$.

Moreover, if $\{f_j\}_{j\geq 1}$ converges to f in L^1 as $j \to \infty$, the sequence $\left\{\mathcal{F}\left[f_j\right]\left(\xi\right)\right\}_{j\geq 1}$ converges to $\mathcal{F}\left[f\right]\left(\xi\right)$ in $\mathbb{C}$, uniformly on ξ, as $j \to \infty$.

Furthermore, suppose that, for each $j \geq 1$, there is $\lim_{|\xi|\to\infty} \mathcal{F}\left[f_j\right]\left(\xi\right) = 0$.

We write

$$\left|\mathcal{F}\left[f\right]\left(\xi\right)\right| \leq \underbrace{\left|\mathcal{F}\left[f_j\right]\left(\xi\right)-\mathcal{F}\left[f\right]\left(\xi\right)\right|}_{(1)}+\underbrace{\left|\mathcal{F}\left[f_j\right]\left(\xi\right)\right|}_{(2)}.$$

Therefore, given $\varepsilon > 0$,

$$(1) \leq \|f_j - f\|_{L^1} \leq \varepsilon$$

for $j \geq j_{0,\varepsilon}$.

If we fix $j > j_{0,\varepsilon}$ in (2), there is $M_{j,\varepsilon} > 0$ so that

$$(2) \leq \varepsilon$$

for $|\xi| \geq M_{j,\varepsilon}$.

Finally,

$$|\mathcal{F}[f](\xi)| \leq 2\varepsilon$$

for $|\xi| \geq M_{j,\varepsilon}$.

So, it only remains to find such a sequence $\{f_j\}_{j\geq 1}$.

We select linear combinations of characteristic functions of rectangles, which are known to be dense in L^1.

If $R = [a_1, b_1] \times ... \times [a_n, b_n]$, with $a_l < b_l$ for every $1 \leq l \leq n$,

$$\mathcal{F}\left[\chi_{[a_l,b_l]}\right](\xi_l) = \begin{cases} \int_{a_l}^{b_l} e^{2\pi i x_l \xi_l} dx_l = \frac{e^{2\pi i b_l \xi_l} - e^{2\pi i a_l \xi_l}}{2\pi i \xi_l} & \text{if} \quad \xi_l \neq 0 \\ b_l - a_l & \text{if} \quad \xi_l = 0 \end{cases}.$$

Therefore,

$$\mathcal{F}[\chi_R](\xi) = \prod_{l=1}^{n} \mathcal{F}\left[\chi_{[a_l,b_l]}\right](\xi_l) \underset{|\xi|\to\infty}{\rightarrow} 0.$$

It should be clear that the map $f \to \mathcal{F}[f]$ is linear. As for the continuity, it follows from (7.4). So we have proved 1).

To prove 2), we consider the characteristic function of the cube $Q = [-1, 1]^n$. We have

$$\|\chi_Q\|_{L^1} = 2^n.$$

Moreover,

$$\mathcal{F}\left[\frac{\chi_{[-1,1]}}{2}\right](\xi) = \begin{cases} \int_{-1}^{1} \frac{e^{2\pi i u v}}{2} du = \frac{\sin 2\pi v}{2\pi v} & \text{if} \quad v \neq 0 \\ 1 & \text{if} \quad v = 0 \end{cases}.$$

Therefore,

$$\left\|\frac{\chi_Q}{2^n}\right\|_{L^1} = 1$$

and

$$\sup_{\xi\in\mathbb{R}^n} \left|\mathcal{F}\left[\frac{\chi_Q}{2^n}\right](\xi)\right| = 1.$$

In view of (7.4), 2) is proved.

We postpone the proof of 3) and 4) until Section 9. □

Remark 168. It is not known how to characterize explicitly the complex linear space $\mathcal{F}\left[L^1\right]$.

Theorem 169. *1. Given $f \in \mathcal{E}^{(m)}$ for some $m = 1, 2, ...$, we assume that $\partial^{\alpha} f$ belongs to L^1 for all $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq m$. Then,*

$$\mathcal{F}\left[\partial^{\alpha} f\right](\xi) = (-2\pi i \xi)^{\alpha} \mathcal{F}[f](\xi). \tag{7.5}$$

2. Suppose that the functions f and $|x|^m f$ are integrable, for some $m \geq 1$. Then, $\mathcal{F}[f]$ belongs to $\mathcal{E}^{(m)}$ and

$$\left(\partial^{\alpha} \mathcal{F}[f]\right)(\xi) = \mathcal{F}\left[(2\pi i x)^{\alpha} f\right](\xi) \tag{7.6}$$

for $|\alpha| \leq m$.

Proof. To prove 1), let us begin by showing that (7.5) holds for $\partial^{\alpha} = \partial_{x_j}$. Moreover, to simplify the notation, let us assume that $j = 1$.

Let $\{a_l\}_{l \geq 1}$ be an increasing sequence in $\mathbb{R}$ going to ∞ as $l \to \infty$, and let $\{b_l\}_{l \geq 1}$ be a decreasing sequence in $\mathbb{R}$ going to $-\infty$ as $l \to \infty$. So, $a_l < b_l$ for l large enough.

With the usual notation, we first integrate by parts on the variable x_1.

$$\begin{aligned} & \int_{a_l}^{b_l} e^{2\pi i x_1 \xi_1} \partial_{x_1} f\left(x_1, x'\right) dx_1 \\ = \; & f\left(b_l, x'\right) - f\left(a_l, x'\right) - 2\pi i \xi_1 \int_{a_l}^{b_l} e^{2\pi i x_1 \xi_1} f\left(x_1, x'\right) dx_1. \end{aligned}$$

Therefore,

$$\begin{aligned} & \int_{\mathbb{R}^{n-1}} e^{2\pi i x' \xi'} \left(\int_{a_l}^{b_l} e^{2\pi i x_1 \xi_1} \partial_{x_1} f\left(x_1, x'\right) dx_1 \right) dx' \\ = \; & \underbrace{\int_{\mathbb{R}^{n-1}} e^{2\pi i x' \xi'} f\left(b_l, x'\right) dx'}_{(1)} - \underbrace{\int_{\mathbb{R}^{n-1}} e^{2\pi i x' \xi'} f\left(a_l, x'\right)}_{(2)} \\ & + \left(-2\pi i \xi_1\right) \mathcal{F}[f](\xi). \end{aligned}$$

We claim that each of the terms (1) and (2) converges to zero as $l \to \infty$, for particular sequences $\{a_l\}_{l \geq 1}$ and $\{b_l\}_{l \geq 1}$.

Let

$$F\left(x_1\right) = \int_{\mathbb{R}^{n-1}} \left| f\left(x_1, x'\right) \right| dx'.$$

According to Fubini's theorem, the function F is defined a.e. and integrable on $\mathbb{R}$. Therefore, for each $l = 1, 2, ...$ fixed, the set

$$\left\{ x_1 > 0 : F\left(x_1\right) > \frac{1}{l} \right\}$$

has finite measure.

So, we can find an increasing sequence $\{b_l\}_{l\geq 1}$ going to ∞ as $l \to \infty$, so that $F(b_l) \leq 1/l$ for every l.

Since the set

$$\left\{x_1 < 0 : F(x_1) > \frac{1}{l}\right\}$$

has also finite measure, there is a decreasing sequence $\{a_l\}_{l\geq 1}$ going to $-\infty$ as $l \to \infty$, so that $F(a_l) \leq 1/l$. Moreover, $a_l < b_l$ for l large enough.

Finally, for these particular sequences,

$$|(1)| + |(2)| \leq F(b_l) + F(a_l) \leq \frac{2}{l} \underset{l\to\infty}{\to} 0.$$

If we assume that $m \geq 2$, since $\partial^\alpha f$ belongs to L^1 for all $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq m-1$, we can repeat what we just did, with $\partial^\alpha f$ instead of f, for $|\alpha| \leq m-1$.

This completes the proof of 1).

To prove 2), we observe that given $s \in \mathbb{N}$, $s \leq m$,

$$|x|^s |f(x)| \leq \left\{ \begin{array}{ccc} |f(x)| & \text{if} & |x| \leq 1 \\ |x|^m |f(x)| & \text{if} & |x| \geq 1 \end{array} \right. .$$

So, the function $|x|^s |f(x)|$ is integrable as well. If $\partial_{\xi_{l_1}} ... \partial_{\xi_{l_h}}$ is any partial derivative of order $s \leq m$,

$$\left|\partial_{\xi_{l_1}} ... \partial_{\xi_{l_h}} \left(e^{2\pi i x\xi}\right) f(x)\right| = \left|(2\pi i x_1)^{l_1} ... (2\pi i x_n)^{l_n}\right| |f(x)|$$

for $l_1 + ... + l_n = s$.

Since

$$\begin{aligned} \left|(2\pi i x_1)^{l_1} ... (2\pi i x_n)^{l_n}\right| &= (2\pi)^s \left(x_1^{2l_1} ... x_n^{2l_n}\right)^{1/2} \\ &\leq C_{n,s} |x|^s \end{aligned}$$

for some $C_{n,s} > 0$, we conclude that $\partial_{\xi_{l_1}} ... \partial_{\xi_{l_h}} \left(e^{2\pi i x\xi}\right) f(x)$ is integrable as a function of x.

According to Theorem 6 and Theorem 8, $\mathcal{F}[f]$ belongs to $\mathcal{E}^{(m)}$ and (7.6) is true.

So, we have proved 2).

This completes the proof of the theorem. □

Remark 170. Theorem 169 says that derivability of f translates into $\mathcal{F}[f]$ going to zero faster at infinity. It also says that integrability of $|x|^m f(x)$ for increasing values of m, implies that $\mathcal{F}[f]$ has continuous partial derivatives of higher order.

Definition 171. *Given $f \in L^1$, we define the conjugate Fourier transform of f, denoted $\overline{\mathcal{F}}[f]$, as*

$$\overline{\mathcal{F}}[f](\xi) = \int_{\mathbb{R}^n} e^{-2\pi i x\xi} f(x)\, dx. \tag{7.7}$$

Remark 172.
$$\overline{\mathcal{F}}[f] = \overline{\mathcal{F}\left[\overline{f}\right]}, \tag{7.8}$$
where $\bar{\cdot}$ denotes the complex conjugate of the function $\cdot$.

Therefore, $\overline{\mathcal{F}}$ has the properties stated in Theorem 167. Moreover,
$$\overline{\mathcal{F}}[\partial^\alpha f](\xi) = (2\pi i\xi)^\alpha \overline{\mathcal{F}}[f](\xi)$$
and
$$\left(\partial^\alpha \overline{\mathcal{F}}[f]\right)(\xi) = \overline{\mathcal{F}}[(-2\pi ix)^\alpha f](\xi),$$
under the hypotheses of Theorem 169.

Let us observe that when we talk about the Fourier transform, or the conjugate Fourier transform, we mean the transform of a fixed function as well as the operator from one space into another.

We now enounce two classical theorems, which of course have classical proofs. However, we will postpone their proofs until we define the Fourier transform in the context of distributions. At that time we will able to prove them, using results from the theory of distributions.

Theorem 173. *If the functions f and $\widehat{f}$ are both integrable,*
$$\overline{\mathcal{F}}\left[\widehat{f}\right] = f \text{ a.e.}.$$

Theorem 174. *The linear and continuous operator*
$$\mathcal{F} : L^1 \to \mathcal{F}\left[L^1\right] \subseteq \mathcal{C}_0$$
is injective, in the sense that given $g \in \mathcal{F}\left[L^1\right]$ there is $f \in L^1$, uniquely determined a.e., so that $\mathcal{F}[f] = g$.

Remark 175. 1. If the integrable function f does not belong to $\mathcal{C}_0$, Theorem 173 implies that $\widehat{f}$ is not integrable. Therefore, $\mathcal{F}\left[L^1\right] \subsetneqq L^1$.

2. According to 1), the operator $\mathcal{F}^{-1} : \mathcal{F}\left[L^1\right] :\to L^1$ coincides with $\overline{\mathcal{F}}$ only on $\mathcal{F}\left[L^1\right] \bigcap L^1$.

We now move on to the Fourier transform on L^2. For clarity, we will denote $\mathcal{F}_2$ the operator we are about to define. Since L^2 is not contained in L^1, we can only use (7.1) when $f \in L^1 \bigcap L^2$.

Theorem 176. *The linear operator*
$$L^1 \bigcap L^2 \hookleftarrow \mathcal{S} \quad \xrightarrow{\mathcal{F}} \quad L^2 \tag{7.9}$$
is well defined and
$$\|\mathcal{F}[\varphi]\|_{L^2} = \|\varphi\|_{L^2}$$
for every $\varphi \in \mathcal{S}$.

As we did with Theorem 173 and Theorem 174, we postpone the proof of Theorem 176.

With Theorem 176 in hand, the definition of $\mathcal{F}_2$ is just a straightforward application of the following, well known, extension principle:

Theorem 177. *Let X be a normed linear space, let Y be a Banach space, and let X_1 be a dense and linear subspace of X. Then, given a linear and continuous operator*

$$X \hookleftarrow X_1 \quad \overset{T}{\to} \quad Y$$

there is one, and only one, linear and continuous operator $T : X \to Y$ such that

1. *$T\,|X_1 = T_1$.*
2. *$\|T\|_{X,Y} = \|T_1\|_{X_1,Y}$.*

Theorem 178. *There is a unique, linear, and continuous operator*

$$L^2 \quad \overset{\mathcal{F}_2}{\to} \quad L^2$$

that extends (7.9). Moreover,

$$\|\mathcal{F}_2\,[f]\|_{L^2} = \|f\|_{L^2} \tag{7.10}$$

for every $f \in L^2$.

Remark 179.
1. The equality (7.10) is named Parseval's identity after the mathematician Marc Antoine Parseval (1755-1836).

   Parseval's identity shows that $\mathcal{F}_2$ is an isometry of L^2 into itself.
2. In the same way as with $\mathcal{F}_2$, we can define the extension $\overline{\mathcal{F}}_2$ of the conjugate Fourier transform, which also satisfies

   $$\overline{\mathcal{F}_2}\,[f] = \overline{\mathcal{F}_2\,\left[\overline{f}\right]},$$

   by virtue of the extension principle.

The proof of the following two results will be given in Section 9.

Theorem 180. *The operators $\mathcal{F}_2$ and $\overline{\mathcal{F}_2}$ are isomorphisms of L^2 onto itself, and they are inverse of each other. Moreover, they preserve the scalar product in L^2.*

That is to say,

$$\begin{aligned}(f,g)_{L^2} &\overset{def}{=} \int_{\mathbb{R}^n} f\,(x)\,\overline{g\,(x)}dx \underset{(1)}{=} \int_{\mathbb{R}^n} \mathcal{F}_2\,[f]\,(\xi)\,\overline{\mathcal{F}_2\,[g]\,(\xi)}d\xi \\ &\overset{def}{=} (\mathcal{F}_2\,[f]\,, \mathcal{F}_2\,[g])_{L^2}\,. \end{aligned} \tag{7.11}$$

Equality (1) is named Plancherel's identity, after the mathematician Michel Plancherel (1885-1967).

Theorem 181. *Given* $f \in L^1 \bigcap L^2$,

$$\mathcal{F}[f] = \mathcal{F}_2[f] \ \ a.e..$$

Remark 182. 1. Strictly speaking, $\mathcal{F}_2$ and $\overline{\mathcal{F}_2}$ operate between two different copies of L^2, which are L^2_x and L^2_ξ.

That is to say, the graph of $\mathcal{F}_2$, and the graph of $\overline{\mathcal{F}_2}$, are subsets of $L^2_x \times L^2_\xi$, which, in physics terms, is called the phase space, where x is associated with position and ξ is associated with momentum. From the mathematics point of view, we will always identify the variables x and ξ.

2. The definition of the operator $\mathcal{F} : L^1 \to \mathcal{C}_0$, by means of an integral, is straightforward. However, it presents two difficulties: How to characterize the image and how to compute the inverse.

   On the other hand, the definition of the operator $\mathcal{F}_2$ requires more labor, by means of the extension principle. However, once that is done, none of the difficulties described before exist. In that respect, we could say that the L^2-theory of the Fourier transform is more symmetrical, perhaps more elegant, than the L^1-theory.

3. Once Theorem 181 is proved, we will be able to remove the sub-index from $\mathcal{F}_2$.

We begin to study these matters in the next section.

8 The Fourier transform on $\mathcal{S}$ and $\mathcal{S}'$

If we fix $f \in L^1$, the function $\widehat{f}$, being continuous, defines a distribution in $\mathcal{D}'$.
If $\varphi \in \mathcal{D}$,

$$\begin{aligned} \left(T_{\widehat{f}}, \varphi\right)_{\mathcal{D}',\mathcal{D}} &= \int_{\mathbb{R}^n} \widehat{f}(x)\,\varphi(x)\,dx = \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} e^{2\pi i y x} f(y)\,dy\right)\varphi(x)\,dx \\ &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} e^{2\pi i y x}\varphi(x)\,dx\right) f(y)\,dy \\ &= \underbrace{\int_{\mathbb{R}^n} \widehat{\varphi}(y)\, f(y)\,dy}_{(1)}. \end{aligned}$$

Although the integral (1) is well defined, it does not provide a pairing $(T_f, \widehat{\varphi})_{\mathcal{D}',\mathcal{D}}$. To see why, we begin with the following result:

Lemma 183. *The statements that follow are true:*

1.
$$e^{2\pi ixz}=\sum_{\alpha\in\mathbb{N}^n}\frac{(2\pi ix)^{\alpha}}{\alpha!}z^{\alpha}, \tag{8.1}$$
for all $x\in\mathbb{R}^n$, $z\in\mathbb{C}^n$.

2. *The series (8.1) converges, absolutely and uniformly, for x in each compact subset of $\mathbb{R}^n$ and z in each compact subset of $\mathbb{C}^n$.*

Proof. To prove 1) we write

$$e^{2\pi ixz}=\sum_{j\geq 0}\frac{(2\pi i)^j}{j!}(xz)^j,$$

for $x\in\mathbb{R}^n$and $z\in\mathbb{C}^n$, fixed.

According to Lemma 1,

$$(xz)^j=(x_1z_1+...+x_nz_n)^j=\sum_{|\alpha|=j}\frac{j!}{\alpha!}x^{\alpha}z^{\alpha}.$$

Therefore,

$$\sum_{j\geq 0}\frac{(2\pi i)^j}{j!}(xz)^j=\sum_{j\geq 0}\frac{(2\pi i)^j}{j!}\sum_{|\alpha|=j}\frac{j!}{\alpha!}x^{\alpha}z^{\alpha}=\sum_{\alpha\in\mathbb{N}^n}\frac{(2\pi ix)^{\alpha}}{\alpha!}z^{\alpha}.$$

So, 1) is proved.

For 2), let us assume that $|x_j|\leq R$ and $|z_j|\leq R'$, for $R,R'>0$ and $1\leq j\leq n$.

$$\begin{aligned}\sum_{\alpha\in\mathbb{N}^n}\left|\frac{(2\pi ix)^{\alpha}}{\alpha!}z^{\alpha}\right| &= \sum_{\alpha_1,\dots,\alpha_n\geq 0}\prod_{l=1}^{n}\left|\frac{(2\pi ix_l)^{\alpha_l}}{\alpha_l!}z_l^{\alpha_l}\right|\\ &= \prod_{l=1}^{n}\sum_{\alpha_l\geq 0}\left|\frac{(2\pi ix_l)^{\alpha_l}}{\alpha_l!}z_l^{\alpha_l}\right|\leq\prod_{l=1}^{n}\sum_{\alpha_l\geq 0}\frac{(2\pi R)^{\alpha_l}}{\alpha_l!}R'^{\alpha_l}.\end{aligned} \tag{8.2}$$

Since each of the n series in (8.2) converges, we have proved 2).

This completes the proof of the lemma. □

Lemma 184. *If $\varphi\in\mathcal{D}$, the function $\widehat{\varphi}$ has an extension Φ to $\mathbb{C}^n$ that is an entire function, that is to say, it is a holomorphic function on $\mathbb{C}^n$.*

Proof. The integral
$$\int_{supp(\varphi)} e^{2\pi ix\xi}\varphi(x)\,dx$$
is still defined if we substitute ξ for the variable $z \in \mathbb{C}^n$.

According to Lemma 183,
$$\Phi(z) = \int_{supp(\varphi)} e^{2\pi ixz}\varphi(x)\,dx = \sum_{\alpha\geq 0}\frac{1}{\alpha!}\left(\int_{supp(\varphi)} (2\pi ix)^{\alpha}\varphi(x)\,dx\right) z^{\alpha}.$$

This completes the proof of the lemma. □

As an immediate consequence of Lemma 184, the smooth function $\Phi(\xi) = \widehat{\varphi}(\xi)$ cannot have compact support, unless it is the identically zero function. Therefore, although the integral (1) suggests that given $T \in \mathcal{D}'$, we might attempt to state the definition
$$\left(\widehat{T},\varphi\right)_{\mathcal{D}',\mathcal{D}} = (T,\widehat{\varphi})_{\mathcal{D}',\mathcal{D}},$$
this is not possible.

One of Schwart's great achievements was the realization that the space $\mathcal{S}$, which he defined, is a natural domain for $\mathcal{F}$ and $\mathcal{F}_2$. On $\mathcal{S}$, both operators have all the good features and none of the inconveniences. Schwartz also observed that the L^1 and the L^2 theories, as well as, in general, the L^p-theory for $1 \leq p \leq \infty$, can be seen as part of the theory of the Fourier transform on the space $\mathcal{S}'$ of tempered distributions.

Theorem 185. *The operator*
$$\mathcal{S} \overset{\mathcal{F}}{\to} \mathcal{S}$$
is well defined, linear, and continuous.

Proof. Since $\mathcal{S} \subseteq L^1$, given $\varphi \in \mathcal{S}$, $\mathcal{F}[\varphi]$ is defined by means of an integral, as in (7.1).

It should be clear that the map $\varphi \to \mathcal{F}[\varphi]$ is linear.

Moreover, according to Corollary 58 and Theorem 169,
$$\begin{aligned}\left|\xi^{\beta}\left(\partial^{\alpha}\mathcal{F}[\varphi]\right)(\xi)\right| &= \left|\xi^{\beta}\mathcal{F}\left[(2\pi ix)^{\alpha} f\right](\xi)\right| \\ &= (2\pi)^{-|\beta|}\left|\mathcal{F}\left[\partial^{\beta}\left((2\pi ix)^{\alpha} f\right)\right](\xi)\right|.\end{aligned}$$

Therefore, after using Leibniz's rule, we have, as in the proof of Theorem 57,
$$\begin{aligned}&\left|\xi^{\beta}\left(\partial^{\alpha}\mathcal{F}[\varphi]\right)(\xi)\right| \\ \leq\ & C_{\alpha,\beta,m,n}\sup_{|\gamma|\leq a,|\eta|\leq b}\ \sup_{x\in\mathbb{R}^n}\left|x^{\gamma}\left(\partial^{\eta}\varphi\right)(x)\right|\int_{\mathbb{R}^n}\left(1+|x|^2\right)^{-m}dx,\end{aligned}$$

for $m > n/2$ and for appropriate $a, b \in \mathbb{N}$.

Thus, we have proved that $\mathcal{F}[\varphi] \in \mathcal{S}$ and that the Fourier transform is continuous from $\mathcal{S}$ into itself.

This completes the proof of the theorem. □

Remark 186. From (7.8), it follows that Theorem 185 also applies to the conjugate Fourier transform.

Theorem 187.

$$\mathcal{F}\left[\overline{\mathcal{F}}[\varphi]\right] = \varphi \tag{8.3}$$

and

$$\overline{\mathcal{F}}[\mathcal{F}[\varphi]] = \varphi, \tag{8.4}$$

for every $\varphi \in \mathcal{S}$.

Proof. If we assume that (8.4) is true, using (7.8),

$$\overline{\mathcal{F}\left[\overline{\mathcal{F}}[\varphi]\right]} = \overline{\mathcal{F}}\left[\overline{\overline{\mathcal{F}}[\varphi]}\right] = \overline{\mathcal{F}}[\mathcal{F}[\overline{\varphi}]] = \overline{\varphi}.$$

So, let us prove (8.4). What we need to show is that

$$\begin{aligned} \varphi(y) &= \int_{\mathbb{R}^n} e^{-2\pi i y\xi} \widehat{\varphi}(\xi)\, d\xi \\ &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} e^{2\pi i (x-y)\xi} \varphi(x)\, dx \right) d\xi, \end{aligned} \tag{8.5}$$

for every $\varphi \in \mathcal{S}$ and for all $y \in \mathbb{R}^n$.

The difficulty is, of course, that we cannot change the order of integration in (8.5). However, given $\psi \in \mathcal{S}$, the double integral

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} e^{2\pi i (x-y)\xi} \psi\left(\frac{\xi}{j}\right) \varphi(x)\, dx d\xi \tag{8.6}$$

exists for each $j \geq 1$ and equals

$$\int_{\mathbb{R}^n} e^{-2\pi i y\xi} \psi\left(\frac{\xi}{j}\right) \widehat{\varphi}(\xi). \tag{8.7}$$

With the change of variables $x \to u/(j+y)$ and $\xi \to jv$, the integral (8.6) becomes

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} e^{2\pi i uv} \psi(v)\, \varphi\left(\frac{u}{j} + y\right) du dv = \int_{\mathbb{R}^n} \widehat{\psi}(u)\, \varphi\left(\frac{u}{j} + y\right) du.$$

That is to say, we have the identity

$$\int_{\mathbb{R}^n} e^{-2\pi i y\xi} \psi\left(\frac{\xi}{j}\right) \widehat{\varphi}(\xi) = \int_{\mathbb{R}^n} \widehat{\psi}(u)\, \varphi\left(\frac{u}{j} + y\right) du, \tag{8.8}$$

for all $\varphi, \psi \in \mathcal{S}$.

Using Lebesgue's dominated convergence theorem, we can take the limit on both sides of (8.8 as $j \to \infty$, to get

$$\psi(0)\overline{\mathcal{F}}[\mathcal{F}[\varphi]](y) = \varphi(y)\int_{\mathbb{R}^n} \widehat{\psi}(u)\,du.$$

Therefore, we will have proved (8.4), if we find a function $\psi \in \mathcal{S}$ so that $\psi(0) = 1$ and

$$\int_{\mathbb{R}^n} \widehat{\psi}(u)\,du = 1. \tag{8.9}$$

We claim that all the requirements are fulfilled by the function

$$\psi(x) = e^{-\pi|x|^2}.$$

According to Example 49, the function ψ belongs to $\mathcal{S}$. Moreover, $\psi(0) = 1$.

If we were to prove that $\widehat{\psi} = \psi$, then taking $t = 1$ and making a change of variables in 2) of Lemma 91, we would have (8.9).

So, we set

$$g(\xi) = e^{\pi|\xi|^2}\widehat{\psi}(\xi).$$

Since $g(0) = 1$, to prove that $\widehat{\psi} = \psi$ we only need to verify that $\partial_{\xi_j} g$ is identically zero, for all $1 \le j \le n$.

$$\begin{aligned} \partial_{\xi_j} g(\xi) &= 2\pi\xi_j e^{\pi|\xi|^2}\widehat{\psi}(\xi) + e^{\pi|\xi|^2}\partial_{\xi_j}\widehat{\psi}(\xi) \\ \underset{(1)}{=} e^{\pi|\xi|^2}\widehat{\psi}(\xi)(2\pi\xi_j - 2\pi\xi_j) &= 0, \end{aligned}$$

where we have used, in (1), Theorem 8.

This completes the proof of the theorem. □

Given $T \in \mathcal{S}'$, Theorem 185 implies that the pairing $(T, \mathcal{F}[\varphi])_{\mathcal{S}',\mathcal{S}}$ is well defined, and it is a distribution in $\mathcal{S}'$.

Definition 188. *Given $T \in \mathcal{S}'$,*

$$(\mathcal{F}[T], \varphi) \overset{def}{=} (T, \mathcal{F}[\varphi])_{\mathcal{S}',\mathcal{S}}. \tag{8.10}$$

Likewise,

$$\left(\overline{\mathcal{F}}[T], \varphi\right) \overset{def}{=} \left(T, \overline{\mathcal{F}}[\varphi]\right)_{\mathcal{S}',\mathcal{S}}. \tag{8.11}$$

Theorem 189. *1. The operators $\mathcal{F}$ and $\overline{\mathcal{F}}$ are linear and continuous from $\mathcal{S}'$ into itself.*

2.
$$\overline{\mathcal{F}}\left[\mathcal{F}\left[T\right]\right]=T$$

and
$$\mathcal{F}\left[\overline{\mathcal{F}}\left[T\right]\right]=T,$$

for every $T\in\mathcal{S}'$.

Proof. It should be clear that the maps (8.10) and (8.11) are well defined, and are linear from $\mathcal{S}'$ into itself.

Let us prove that $\mathcal{F}$ is continuous from $\mathcal{S}'$ into itself.

If $\mathcal{B}$ is a bounded subset of $\mathcal{S}$, according to Theorem 185, the image $\mathcal{F}\left[\mathcal{B}\right]$ is a bounded subset of $\mathcal{S}$.

Therefore,

$$\begin{aligned}\left\|\mathcal{F}\left[T\right]\right\|_{\mathcal{S}',\mathcal{B}} &= \sup_{\varphi\in\mathcal{B}}\left|\left(\mathcal{F}\left[T\right],\varphi\right)\right|=\sup_{\varphi\in\mathcal{B}}\left|\left(T,\mathcal{F}\left[\varphi\right]\right)\right| \\ &= \sup_{\varphi\in\mathcal{B}}\left\|T\right\|_{\mathcal{S}',\mathcal{F}\left[\mathcal{B}\right]}.\end{aligned}$$

A similar estimate holds for $\overline{\mathcal{F}}$.

So, 1) is proved.

The proof of 2) follows immediately from Theorem 187.

$$\left(\overline{\mathcal{F}}\left[\mathcal{F}\left[T\right]\right],\varphi\right)=\left(T,\mathcal{F}\left[\overline{\mathcal{F}}\left[\varphi\right]\right]\right)=\left(T,\varphi\right),$$

for all $\varphi\in\mathcal{S}$, and similarly for $\mathcal{F}\overline{\mathcal{F}}$.

This completes the proof of the theorem. □

Theorem 190. *Given $f\in L^1$, there is*

$$\lim_{j\to\infty}\int_{|\xi|\leq j}e^{-2\pi ix\xi}\widehat{f}\left(\xi\right)d\xi=f,$$

in the sense of $\mathcal{S}'$.

That is, for $\varphi\in\mathcal{S}$,

$$\left(\overline{\mathcal{F}}\left[T_{\chi_j\widehat{f}(\xi)}\right],\varphi\right)\underset{j\to\infty}{\rightarrow}\int_{\mathbb{R}^n}f\left(x\right)\varphi\left(x\right)dx,$$

where χ_j denotes the characteristic function of $\{\xi\in\mathbb{R}:|\xi|\leq j\}$.

Proof. For $j\geq 1$ fixed,

$$\begin{aligned}\left(\overline{\mathcal{F}}\left[T_{\chi_j\widehat{f}(\xi)}\right],\varphi\right) &= \int_{|\xi|\leq j}\widehat{f}\left(\xi\right)\overline{\mathcal{F}}\left[\varphi\right]\left(\xi\right)d\xi=\int_{|\xi|\leq j}\widehat{f}\left(\xi\right)\left(\int_{\mathbb{R}^n}e^{-2\pi ix\xi}\varphi\left(x\right)dx\right)d\xi \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}e^{-2\pi ix\xi}\chi_j\left(\xi\right)\widehat{f}\left(\xi\right)\varphi\left(x\right)dxd\xi=\int_{\mathbb{R}^n}\chi_j\left(\xi\right)\widehat{f}\left(\xi\right)\overline{\mathcal{F}}\left[\varphi\right]\left(\xi\right)d\xi.\end{aligned}$$

Lebesgue's dominated convergence theorem tells us that there is

$$\begin{aligned}\lim_{j\to\infty}\int_{\mathbb{R}^n}\chi_j(\xi)\,\widehat{f}(\xi)\,\overline{\mathcal{F}}[\varphi](\xi)\,d\xi &= \int_{\mathbb{R}^n}\widehat{f}(\xi)\,\overline{\mathcal{F}}[\varphi](\xi)\,d\xi\\ &= \left(T_{\widehat{f}},\overline{\mathcal{F}}[\varphi]\right)\\ &= \left(\overline{\mathcal{F}}\mathcal{F}[T_f],\varphi\right)=(T_f,\varphi).\end{aligned}$$

This completes the proof of the theorem. □

Remark 191. Theorem 190 expounds the statement made in 2) of Remark 175.

9 The Fourier transform on L^1 and L^2 (part II)

We begin with the following result:

Lemma 192. *Given $f \in L^1$,*

$$\mathcal{F}[T_f] = T_{\widehat{f}}, \tag{9.1}$$

in the sense of $\mathcal{S}'$.

Proof. First of all, each side of (9.1) is a tempered distribution.

So, for $\varphi \in \mathcal{S}$,

$$\left(T_{\widehat{f}},\varphi\right)=\int_{\mathbb{R}^n}\widehat{f}(\xi)\,\varphi(\xi)\,d\xi=\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}e^{2\pi ix\xi}f(x)\,dx\right)\varphi(\xi)\,d\xi.$$

Since the double integral

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}e^{2\pi ix\xi}f(x)\,\varphi(\xi)\,dx d\xi$$

exists, Fubini's theorem implies that the iterated integrals are equal.

That is to say,

$$\begin{aligned}\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}e^{2\pi ix\xi}f(x)\,dx\right)\varphi(\xi)\,d\xi &= \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}e^{2\pi ix\xi}\varphi(\xi)\,d\xi\right)f(x)\,dx\\ &= (T_f,\mathcal{F}[\varphi])=(\mathcal{F}[T_f],\varphi).\end{aligned}$$

This completes the proof of the lemma. □

Remark 193. Similarly, we can show that, given $f \in L^1$,

$$\overline{\mathcal{F}}[T_f]=T_{\overline{\mathcal{F}}[f]},$$

in the sense of $\mathcal{S}'$.

We are now ready to prove Theorem 173.

Proof. According to Lemma 192 and 2) in Theorem 189,

$$T_{\overline{\mathcal{F}}[\widehat{f}]} = \overline{\mathcal{F}}\left[T_{\widehat{f}}\right] = \overline{\mathcal{F}}\left[\mathcal{F}\left[T_f\right]\right] = T_f,$$

in the sense of $\mathcal{S}'$.

In particular,

$$\left(T_{\overline{\mathcal{F}}[\widehat{f}]}, \varphi\right) = \left(T_f, \varphi\right),$$

for $\varphi \in \mathcal{D}$.

Therefore, Theorem 30 implies that

$$\overline{\mathcal{F}}\left[\widehat{f}\right] = f \text{ a.e.}.$$

This completes the proof of the theorem. □

We go onto proving 3) and 4) in Theorem 167

Proof. To prove 3), we will exhibit a function in $\mathcal{C}_0$ that cannot be the Fourier transform of any function in L^1. The example is taken from ([4], pp. 7-8).

We begin with a general observation, that follows from Remark 163.

Given $f \in L^1(\mathbb{R})$ odd, and given $0 < r < R$ fixed,

$$\begin{aligned}
\int_r^R \frac{\widehat{f}(\xi)}{\xi} d\xi &= 2i \int_r^R \left(\int_0^\infty f(x) \frac{\sin 2\pi x\xi}{\xi} dx\right) d\xi \\
&= 2i \int_0^\infty f(x) \left(\int_r^R \frac{\sin 2\pi x\xi}{\xi} d\xi\right) dx \\
&= 2i \int_0^\infty f(x) \left(\int_{2\pi x r}^{2\pi x R} \frac{\sin y}{y} dy\right) dx.
\end{aligned}$$

It is a classical application of the theory of residues (see, for instance, [11], p.160), that the integral $\int_0^\infty \frac{\sin y}{y} dy$ converges. Indeed,

$$\int_0^\infty \frac{\sin y}{y} dy = \frac{\pi}{2}.$$

Therefore, there is

$$\lim_{r \to 0^+, R \to \infty} \left| \int_r^R \frac{\sin y}{y} dy \right|,$$

which implies that there is $C > 0$ independent of r and R, so that

$$\left| \int_r^R \frac{\sin y}{y} dy \right| \leq C.$$

As a consequence,

$$\left|\int_r^R \frac{\widehat{f}(\xi)}{\xi} d\xi\right| \leq C \|f\|_{L^1}.$$

Let us now consider the function $G(t)$ that is the odd extension to $\mathbb{R}$ of the function

$$g(t) = \begin{cases} \frac{t}{e} & \text{if} \quad 0 \leq t \leq e \\ \frac{1}{\ln t} & \text{if} \quad t \geq e \end{cases}.$$

It should be clear that the function G belongs to $\mathcal{C}_0(\mathbb{R})$.

According to 4) in Example 23, Remark 163, Lemma 192, and Theorem 173, if there is a function $f \in L^1(\mathbb{R})$ so that $\mathcal{F}[f] = G$, the function f must be odd.

Indeed, f is odd if, and only if, T_f is odd. Likewise, $\mathcal{F}[f]$ is odd if, and only if, $T_{\widehat{f}}$ is odd. Finally, $T_f = \overline{\mathcal{F}}\left[T_{\widehat{f}}\right]$.

Given $0 < r < e < R$, we claim that

$$\int_r^R \frac{G(t)}{t} dt \underset{r\to 0^+, R\to\infty}{\to} \infty.$$

Indeed,

$$\int_r^R \frac{G(t)}{t} dt = \int_r^e \frac{dt}{e} + \int_e^R \frac{dt}{t\ln t} = \frac{e-r}{e} + \ln(\ln R).$$

Therefore, there is no odd function f in $L^1(\mathbb{R})$ such that $\widehat{f} = G$.

Given $f \in L^1\left(\mathbb{R}^{n-1}\right)$ odd, the function $\widehat{f}(\xi) G(t)$, where $\widehat{f}$ denotes the Fourier transform on $\mathbb{R}^{n-1}$, provides an example on $\mathbb{R}^n$.

Now, we are ready to prove 4).

According to 2) in Lemma 166, given $f \in \mathcal{C}_0$, there is a sequence $\{f_j\}_{j\geq 1} \subseteq \mathcal{D}$, converging to f in $\mathcal{C}_0$ as $j \to \infty$.

Theorem 187 implies that

$$\mathcal{F}\left[\overline{\mathcal{F}}[f_j]\right] = f_j$$

for all $j \geq 1$.

Hence, f_j belongs to $\mathcal{F}\left[L^1\right]$ for all $j \geq 1$.

So, we have proved 4) in Theorem 167.

The proof of the theorem is, therefore, complete. □

Next, we prove Theorem 174.

Proof. Suppose that there are $f_1, f_2 \in L^1$ so that $\mathcal{F}[f_1] = \mathcal{F}[f_2]$ in $\mathcal{C}_0$.

Then,

$$0 = \overline{\mathcal{F}}\left[T_{\widehat{f_1 - f_2}}\right] = \overline{\mathcal{F}}\mathcal{F}[T_{f_1-f_2}] = T_{f_1-f_2}.$$

Again, Theorem 30 implies that $f_1 = f_2$ a.e..

This completes the proof of the theorem. □

We move to the Fourier transform on L^2. We begin with the following result:

Lemma 194. *Given* $\varphi, \psi \in \mathcal{S}$,

$$(\widehat{\varphi}, \psi)_{L^2} = \left(\varphi, \overline{\mathcal{F}}[\psi]\right)_{L^2}.$$

Proof.

$$\begin{aligned}\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} e^{2\pi i x\xi}\varphi(x)\,dx\right)\overline{\psi(\xi)}d\xi &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{2\pi i x\xi}\varphi(x)\,\overline{\psi(\xi)}dxd\xi \\ &= \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}\overline{e^{-2\pi i x\xi}\psi(\xi)d\xi}\right)\varphi(x)\,dx.\end{aligned}$$

This completes the proof of the lemma. □

We prove now Theorem 176.

Proof. It should be clear that the linear operator

$$L^1 \bigcap L^2 \hookleftarrow \mathcal{S} \quad \overset{\mathcal{F}}{\to} \quad L^2$$

is well defined.

Moreover, if we use Lemma 194 with $\psi = \widehat{\varphi}$, and Theorem 187,

$$\|\mathcal{F}[\varphi]\|_{L^2}^2 = (\widehat{\varphi}, \widehat{\varphi})_{L^2} = \left(\varphi, \overline{\mathcal{F}}[\widehat{\varphi}]\right)_{L^2} = (\varphi, \varphi)_{L^2}^2.$$

This completes the proof of the theorem. □

Let us prove Theorem 178.

Proof. According to Theorem 177, it only remains to prove (7.10).

Given $f \in L^2$, let $\{\varphi_j\}_{j\geq 1}$ be a sequence in $\mathcal{S}$ converging to f in L^2 as $j \to \infty$. Theorem 176 implies that

$$\|\mathcal{F}[\varphi_j]\|_{L^2} = \|\varphi_j\|_{L^2}$$

for every $j \geq 1$.

Therefore,

$$\|\mathcal{F}_2[f]\|_{L^2} = \lim_{j\to\infty}\|\mathcal{F}[\varphi_j]\|_{L^2} = \lim_{j\to\infty}\|\varphi_j\|_{L^2} = \|f\|_{L^2}.$$

This completes the proof of the theorem. □

Next, we prove Theorem 180.

Proof. Given $f \in L^2$ let $\{\varphi_j\}_{j\geq 1}$ be sequence in $\mathcal{S}$ converging to f in L^2 as $j \to \infty$.

Then
$$\overline{\mathcal{F}_2}\left[\mathcal{F}_2\left[f\right]\right] = \lim_{j\to\infty} \overline{\mathcal{F}}\left[\mathcal{F}\left[\varphi_j\right]\right] = \lim_{j\to\infty} \varphi_j = f.$$

Likewise,
$$\mathcal{F}_2\left[\overline{\mathcal{F}_2}\left[f\right]\right] = f$$
for every $f \in L^2$.

As for the preservation of the scalar product, given $f, g \in L^2$, let $\{\varphi_j\}_{j\geq 1}$ and $\{\psi_j\}_{j\geq 1}$ be sequences in $\mathcal{S}$ converging in L^2 to f and g, respectively as $j \to \infty$.

Then, if we use Lemma 194 with $\varphi = \varphi_j$ and $\psi = \widehat{\psi_j}$, and the continuity of the scalar product,

$$\begin{aligned}\left(\mathcal{F}_2\left[f\right], \mathcal{F}_2\left[g\right]\right)_{L^2} &= \lim_{j\to\infty} \left(\mathcal{F}_2\left[\varphi_j\right], \mathcal{F}_2\left[\psi_j\right]\right)_{L^2} \\ &= \lim_{j\to\infty} \left(\varphi_j, \psi_j\right)_{L^2} = \left(f, g\right).\end{aligned}$$

This completes the proof of the theorem. □

Finally, we prove Theorem 181.

Proof. Given $f \in L^1 \bigcap L^2$, let $\{\varphi_j\}_{j\geq 1}$ be a sequence in $\mathcal{S}$ converging to f in L^2 as $j \to \infty$. According to Theorem 69, there is $\lim_{j\to\infty} T_{\mathcal{F}[\varphi_j]} = T_{\mathcal{F}_2[f]}$ in $\mathcal{S}'$. Therefore,
$$T_{\mathcal{F}_2[f]} = \lim_{j\to\infty} \mathcal{F}\left[T_{\varphi_j}\right] = \mathcal{F}\left[T_f\right] = T_{\mathcal{F}[f]}$$
in $\mathcal{S}'$.

Thus,
$$\mathcal{F}_2\left[f\right] = \mathcal{F}\left[f\right] \text{ a.e..}$$

This completes the proof of the theorem. □

According to Theorem 69, if $1 \leq p \leq \infty$, given $f \in L^p$, the distribution T_f belongs to $\mathcal{S}'$. Therefore, $\mathcal{F}\left[T_f\right]$ is well defined in the sense of $\mathcal{S}'$.

Theorem 195. *Given $1 \leq p \leq 2$ fixed, if $f \in L^p$, there is $g \in L^q$, where q is the conjugate exponent of p, so that $\mathcal{F}\left[T_f\right] = T_g$.*

Moreover, there is $C_p > 0$ so that
$$\left\|g\right\|_{L^q} \leq C_p \left\|f\right\|_{L^p}. \tag{9.2}$$

Remark 196. 1. Theorem 195 is one of several theorems known as Hausdorff-Young theorem. They are named after the mathematicians Felix Hausdorff (1868-1942) and William H. Young (1863-1942).

2. Theorem 195 is optimal in the sense that if $p > 2$, there is a function $f \in L^p$ for which $\mathcal{F}[T_f]$ is not defined by a function g in L^q ([8], p. 263).

3. The optimal constant in (9.2) is
$$C_p = p^{1/(2p)} q^{-1/(2q)}$$
([5], p. 94).

4. For a thorough discussion of the Hausdorff-Young theorems, from a historical point of view, see [3].

10 The interaction of the Fourier transform with the tensor product, the convolution product, and the multiplicative product

We begin with the following result:

Theorem 197. *Let $T \in \mathcal{S}'_x$ and $S \in \mathcal{S}'_y$.*

1. *The tensor product $T \otimes S$ belongs to $\mathcal{S}'_{x,y}$.*

2. $$\mathcal{F}_{x,y}[T_x \otimes S_y] = \mathcal{F}_x[T_x] \otimes \mathcal{F}_y[S_y], \tag{10.1}$$
in the sense of $\mathcal{S}'_{x,y}$.

Proof. According to Theorem 124, $T \otimes S$ belongs to $\mathcal{D}'_{x,y}$. To prove that it is a tempered distribution, we use Theorem 104, and various properties of the tensor product.
$$T \otimes S = (\partial^\alpha T_f) \otimes \left(\partial^\beta T_g\right) = \partial_x^\alpha \partial_y^\beta \left(T_{f(x)g(y)}\right). \tag{10.2}$$

Since $f : \mathbb{R}^n \to \mathbb{C}$ and $g : \mathbb{R}^m \to \mathbb{C}$ are slowly increasing and continuous functions, $f(x)\, g(y)$ is an slowly increasing and continuous function from $\mathbb{R}^{n+m}$ into $\mathbb{C}$. So, 1) is proved.

The function $e^{2\pi i x \xi}$ has separated variables, so, 2) should be quite expected.

Indeed, to prove 2), it suffices to test each side of (10.1) on functions of the form $\varphi(\xi)\, \psi(\eta)$, for $\varphi \in \mathcal{D}_\xi$ and $\psi \in \mathcal{D}_\eta$.

$$\begin{aligned} \left(\mathcal{F}_{x,y}[T_x \otimes S_y], \varphi(\xi)\, \psi(\eta)\right) &= \left(T_x \otimes S_y, \widehat{\varphi}(x)\, \widehat{\psi}(y)\right) \\ &= \left(T_x, \widehat{\varphi}(x)\right) \left(S_y, \widehat{\psi}(y)\right) \\ &= \left(\mathcal{F}_x[T_x] \otimes \mathcal{F}_y[S_y], \varphi(\xi)\, \psi(\eta)\right). \end{aligned}$$

This completes the proof of the theorem. □

Next, we prove a result known as Peetre's lemma, named after the mathematician Jaak Peetre (b. 1935).

Lemma 198. *Given $\xi, \eta \in \mathbb{R}^n$ and $r \in \mathbb{R}$,*

$$\left(1+|\xi|^2\right)^r \leq 2^{|r|}\left(1+|\eta|^2\right)^r\left(1+|\xi-\eta|^2\right)^{|r|}. \tag{10.3}$$

Proof. The change of variables $x=\eta$ and $y=\xi-\eta$ reduces (10.3) to the equivalent form

$$\left(1+|x+y|^2\right)^r \leq 2^{|r|}\left(1+|x|^2\right)^r\left(1+|y|^2\right)^{|r|}. \tag{10.4}$$

We have

$$\begin{aligned}\left(1+|x+y|^2\right) &= 1+\langle x+y, x+y\rangle \\ &= 1+|x|^2+|y|^2+2\langle x,y\rangle,\end{aligned}$$

where $\langle\cdot,\cdot\rangle$ is the usual scalar product in $\mathbb{R}^n$.

Since

$$2\langle x,y\rangle \leq 2\,|x|\,|y| \leq |x|^2+|y|^2,$$

we can write

$$\begin{aligned}1+|x|^2+|y|^2+2\langle x,y\rangle &\leq 1+2\,|x|^2+2\,|y|^2 \leq 2\left(1+|x|^2+|y|^2+|x|^2\,|y|^2\right) \\ &= 2\left(1+|x|^2\right)\left(1+|y|^2\right).\end{aligned}$$

Therefore, we have (10.3) for $r \geq 0$.

We can write (10.3), for $r>0$, as

$$\left(1+|\eta|^2\right)^{-r} \leq 2^{|r|}\left(1+|\xi|^2\right)^{-r}\left(1+|\xi-\eta|^2\right)^{|r|}. \tag{10.5}$$

So, with the exchange $\xi \to \eta$ and $\eta \to \xi$, (10.5) is (10.3), for $r<0$.

This completes the proof of the lemma. □

Theorem 199. *The following statements hold:*

1. *The map*

$$\begin{array}{ccc}\mathcal{S}\times\mathcal{S} & \to & \mathcal{S} \\ (\varphi,\psi) & \to & \varphi*\psi\end{array}$$

is well defined, bilinear, and continuous.

2. *The multiplicative product*

$$\begin{array}{ccc}\mathcal{S}\times\mathcal{S} & \to & \mathcal{S} \\ (\varphi,\psi) & \to & \varphi\psi\end{array}$$

is well defined, bilinear, and continuous.

Proof. Given φ, ψ in $\mathcal{S}$ and given $m \in \mathbb{N}$, if $|\alpha| \leq 2m$, and $\beta \in \mathbb{N}^n$, Theorem 6, Theorem 8, Lemma 4 and Lemma 10.3, imply

$$\begin{aligned}
&\left|x^{\alpha}\left(\partial^{\beta}\left(\varphi * \psi\right)\right)(x)\right| \\
\leq\ & C_{m,n,l} \sup_{x\in\mathbb{R}^n}\left(\left(1+|x|^2\right)^m\left|\left(\partial^{\beta}\varphi\right)(x)\right|\right)\sup_{x\in\mathbb{R}^n}\left(1+|x-y|^2\right)^{m+l}|\psi(x)| \\
&\times\int_{\mathbb{R}^n}\left(1+|x|^2\right)^{-l}dy \\
\leq\ & C_{m,n,l}\sup_{|\alpha|\leq 2m}\sup_{x\in\mathbb{R}^n}\left|x^{\alpha}\left(\partial^{\beta}\varphi\right)(x)\right|\sup_{|\alpha|\leq 2m+2l}\sup_{x\in\mathbb{R}^n}\left|x^{\alpha}(\psi)(x)\right|,
\end{aligned}$$

for $l > n/2$.

Therefore, $\varphi * \psi \in \mathcal{S}$, and the map is continuous. It should be clear that it is bilinear.

Thus, we have proved 1).

To prove 2), we invoke again Lemma 4.

$$\left|x^{\alpha}\left(\partial^{\beta}\left(\varphi\psi\right)\right)(x)\right| \leq C_{\beta,n}\sup_{\gamma\leq\beta}\sup_{x\in\mathbb{R}^n}\left|x^{\alpha}\partial^{\gamma}\varphi(x)\right|\sup_{\gamma\leq\beta}\sup_{x\in\mathbb{R}^n}\left|\partial^{\gamma}\psi(x)\right|,$$

which shows that $\varphi\psi \in \mathcal{S}$ and that the map is continuous. It should be clear that the map is bilinear.

This completes the proof of the theorem. □

Lemma 200. *Given φ, ψ in $\mathcal{S}$,*

$$\mathcal{F}\left[\varphi\overline{\mathcal{F}}[\psi]\right] = \mathcal{F}[\varphi] * \psi, \tag{10.6}$$
$$\overline{\mathcal{F}}\left[\varphi\mathcal{F}[\psi]\right] = \overline{\mathcal{F}}[\varphi] * \psi. \tag{10.7}$$

Proof. To prove the first identity we observe that

$$\begin{aligned}
\mathcal{F}\left[\varphi\overline{\mathcal{F}}[\psi]\right](y) &= \int_{\mathbb{R}^n} e^{2\pi i y\xi}\varphi(\xi)\left(\int_{\mathbb{R}^n} e^{-2\pi i x\xi}\psi(x)\,dx\right)d\xi \\
&= \int_{\mathbb{R}^n}\varphi(\xi)\left(\int_{\mathbb{R}^n} e^{2\pi i(y-x)\xi}\psi(x)\,dx\right)d\xi.
\end{aligned}$$

Since $\varphi(\xi)\,\psi(x)$ is integrable on $\mathbb{R}^{2n}$, we can exchange the order of integration, to obtain

$$\begin{aligned}
\int_{\mathbb{R}^n}\psi(x)\left(\int_{\mathbb{R}^n} e^{2\pi i(y-x)\xi}\varphi(\xi)\,dx\right)d\xi &= \int_{\mathbb{R}^n}\psi(x)\,\mathcal{F}[\varphi](y-x)\,dx \\
&= \left(\mathcal{F}[\varphi] * \psi\right)(y).
\end{aligned}$$

The proof of the second identity is similar, if we replace $\mathcal{F}$ with $\overline{\mathcal{F}}$.

This concludes the proof of the lemma. □

Remark 201. 1. If we substitute $\mathcal{F}[\psi]$ for ψ in (10.6),

$$\mathcal{F}[\varphi\psi] = \mathcal{F}[\varphi] * \mathcal{F}[\psi],$$

while substituting $\overline{\mathcal{F}}[\psi]$ for ψ in (10.7),

$$\overline{\mathcal{F}}[\varphi\psi] = \overline{\mathcal{F}}[\varphi] * \overline{\mathcal{F}}[\psi].$$

2. If we replace φ with $\mathcal{F}[\varphi]$ in (10.7), we obtain

$$\overline{\mathcal{F}}[\mathcal{F}[\varphi]\,\mathcal{F}[\psi]] = \varphi * \psi. \tag{10.8}$$

Then, taking the Fourier transform on both sides of (10.8), gives us

$$\mathcal{F}[\varphi]\,\mathcal{F}[\psi] = \mathcal{F}[\varphi * \psi].$$

3. If we substitute $\overline{\mathcal{F}}[\varphi]$ for φ in (10.6),

$$\mathcal{F}\left[\overline{\mathcal{F}}[\varphi]\,\overline{\mathcal{F}}[\psi]\right] = \varphi * \psi. \tag{10.9}$$

Therefore, taking the conjugate Fourier transform on both sides of (10.9),

$$\overline{\mathcal{F}}[\varphi]\,\overline{\mathcal{F}}[\psi] = \overline{\mathcal{F}}[\varphi * \psi].$$

That is to say, $\mathcal{F}$ and $\overline{\mathcal{F}}$, as operators on $\mathcal{S}$, act between the algebras $(\mathcal{S}, *)$ and $(\mathcal{S}, \cdot)$, where $\cdot$ denotes the multiplicative product.

Theorem 202. *Given f, g in L^1,*

$$\begin{aligned}
\mathcal{F}[f * g] &= \mathcal{F}[f]\,\mathcal{F}[g], \\
\overline{\mathcal{F}}[f * g] &= \overline{\mathcal{F}}[f]\,\overline{\mathcal{F}}[g].
\end{aligned}$$

Proof. Let $\{\varphi_j\}_{j\geq 1}$ and $\{\psi_j\}_{j\geq 1}$ be sequences in $\mathcal{S}$ converging, respectively, to f and g in L^1 as $j \to \infty$.

Then, Young's convolution theorem (see, for instance, [29], p. 146, Theorem 9.2; p. 145, Theorem 9.1), implies that there is $\lim_{j\to\infty} \varphi_j * \psi_j = f * g$ in L^1.

According to Theorem 167 and Remark 201,

$$\mathcal{F}[f * g] = \lim_{j\to\infty} \mathcal{F}[\varphi_j * \psi_j] = \lim_{j\to\infty} \left(\mathcal{F}[\varphi_j]\,\mathcal{F}[\psi_j]\right) = \mathcal{F}[f]\,\mathcal{F}[g]$$

in $\mathcal{C}_0$.

The proof of the other identity is similar, replacing $\mathcal{F}$ with $\overline{\mathcal{F}}$.

This completes the proof of the theorem. □

Remark 203. Theorem 202 shows that the Fourier transform is a homomorphism from the algebra $\left(L^1, *\right)$ into the algebra $(\mathcal{C}_0, \cdot)$.

We have said numerous times that properties of the Fourier transform hold, perhaps with very obvious changes, for the conjugate Fourier transform. We will say it no more.

Theorem 204. *If $T \in \mathcal{E}'$ and $S \in \mathcal{S}'$, the convolution $T * S$ belongs to $\mathcal{S}'$.*

Proof. We know already that $T * S$ is well defined as a distribution in $\mathcal{D}'$. To prove that it can be extended to a distribution in $\mathcal{S}'$, we will use Theorem 104 and Theorem 107.

Therefore, in the sense of $\mathcal{D}'$,

$$T * S = \left(\sum_{\alpha} \partial^{\alpha} T_{f_\alpha}\right) * \partial^{\beta} T_f,$$

where the functions f_α are continuous with compact support, and the function f is continuous and slowly increasing.

So,

$$T * S = \sum_{\alpha} \partial^{\alpha+\beta} \left(T_{f_\alpha} * T_f\right).$$

Given $\varphi \in \mathcal{D}$,

$$\begin{aligned}
\left(T_{f_\alpha} * T_f, \varphi\right) &= \left(T_{f_\alpha(x)} \otimes T_{f(y)}, \varphi(x+y)\right) = \left(T_{f_\alpha(x)}, \left(T_{f(y)}, \varphi(x+y)\right)\right) \\
&= \int_K f_\alpha(x) \left(\int_{\mathbb{R}^n} f(y) \varphi(x+y)\, dy\right) dx = \int_K f_\alpha(x) \left(\int_{supp(\varphi)} f(z-x) \varphi(z)\, dz\right) dx,
\end{aligned}$$

where K is a compact subset of $\mathbb{R}^n$ containing the support of f_α for all α. Since the double integral exists, we can write

$$\begin{aligned}
\left(T_{f_\alpha} * T_f, \varphi\right) &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f_\alpha(x) f(z-x)\, dx\right) \varphi(z)\, dz \\
&= \int_{\mathbb{R}^n} (f_\alpha * f)(z) \varphi(z)\, dz.
\end{aligned}$$

Theorem 6 implies that the function $f_\alpha * f$ is continuous. Moreover, we claim that it is slowly increasing.

Indeed, since the function f is slowly increasing, according to 3) in Example 59, there is $m \in \mathbb{N}$ and $C_m > 0$ so that

$$|f(z-x)| \leq C_m \left(1 + |z-x|^2\right)^m$$

for all $z \in \mathbb{R}^n$.

Therefore, Lemma 198 implies

$$\begin{aligned}\left(1+|z|^2\right)^m |(f_\alpha * f)(z)| &\leq 2^m \int_K \left(1+|x|^2\right)^m f_\alpha(x) \left(1+|z-x|^2\right)^m f(z-x)\,dx \\ &\leq 2^m C_m \left\| \left(1+|\cdot|^2\right)^m f_\alpha \right\|_{L^1}.\end{aligned}$$

So, $f_\alpha * f$ defines a tempered distribution.

Finally,

$$T * S = \sum_\alpha \partial^{\alpha+\beta}\left(T_{f_\alpha * f}\right) \tag{10.10}$$

in $\mathcal{S}'$. □

Definition 205.

$$\mathcal{O}_M = \{g \in \mathcal{E} : \partial^\alpha g \text{ is slowly increasing for each } \alpha \in \mathbb{N}^n\}.$$

Example 206. Every polynomial function $P(x)$ belongs to $\mathcal{O}_M$.

Moreover, for each $x \in \mathbb{R}^n$, the function $e^{2\pi i x\xi}$ also belongs to $\mathcal{O}_M$.

Remark 207. Let $f \in \mathcal{O}_M$.

1. The map
$$\begin{array}{ccc}\mathcal{S} & \to & \mathcal{S} \\ \varphi & \to & f\varphi\end{array}$$
is well defined, linear, and continuous.

2. The map
$$\begin{array}{ccc}\mathcal{S}' & \to & \mathcal{S}' \\ T & \to & fT\end{array}$$
is well defined, linear, and continuous.

The verification of these statements uses estimates we have performed quite a few times, so we will omit it.

Let us recall that, according to (5.8),

$$(\tau_{-h}(T), \varphi) = (T, \tau_h(\varphi)),$$

where

$$\tau_h(\varphi)(x) = \varphi(x-h).$$

Moreover, according to 4) in Example 23,

$$(d_k(T), \varphi) = \left(T, \frac{d_{1/k}(\varphi)}{|k|^n}\right).$$

In particular, if $k = -1$,

$$(d_{-1}(T), \varphi) = (T, \varphi(-\cdot)).$$

Theorem 208. *1. Given $T \in \mathcal{S}'$ and $h \in \mathbb{R}^n$,*

$$\mathcal{F}[\tau_{-h}(T)] = e^{2\pi i h\xi}\mathcal{F}[T],$$

where $e^{2\pi i h\xi}\mathcal{F}[T]$ signifies the multiplicative product of the tempered distribution $\mathcal{F}[T]$ by the function $e^{2\pi i h\xi}$ in $\mathcal{O}_M$.

2. Given $T \in \mathcal{S}'$ and $k \in \mathbb{R}$, $k \neq 0$,

$$\mathcal{F}[d_k(T)] = \frac{1}{|k|^n} d_{1/k}(\mathcal{F}[T]),$$

In particular,

$$(\mathcal{F}[d_{-1}(T)], \varphi) = (T, \mathcal{F}[\varphi](-\cdot)) = \left(\overline{\mathcal{F}}[T], \varphi\right) \tag{10.11}$$

3. If $T \in \mathcal{S}'$ is odd, $\mathcal{F}[T]$ is also odd. If T is even, $\mathcal{F}[T]$ is also even.

4. Given $T \in \mathcal{S}'$,

$$\begin{aligned} \mathcal{F}[\partial^\alpha T] &= (-2\pi i\xi)^\alpha \mathcal{F}[T], \\ \partial^\alpha \mathcal{F}[T] &= \mathcal{F}[(2\pi i x)^\alpha T]. \end{aligned} \tag{10.12}$$

Proof. The proof of these statements is straightforward and it will be omitted. However, the interpretation of (10.11) merits a few words.

If the distribution T were defined by an odd function f, say in L^1, then, according to Lemma 192, Remark 193, and Remark 163,

$$\overline{\mathcal{F}}[T_f] = T_{\overline{\mathcal{F}}[f]} = T_{\mathcal{F}[f](-\xi)} = -T_{\mathcal{F}[f](\xi)}.$$

Therefore, if T is an odd distribution, the identity

$$-\mathcal{F}[(T)] = \mathcal{F}[d_{-1}(T)] = \overline{\mathcal{F}}[T]$$

should be interpreted as saying that $\mathcal{F}[(T)]$ is odd.

Likewise when T is even.

This concludes the proof of the theorem. □

Lemma 209. *Given $T \in \mathcal{E}'$, the pairing $\left(T_x, e^{2\pi i x\xi}\right)_{\mathcal{E}',\mathcal{E}}$ defines a function $g(\xi)$ in $\mathcal{O}_M$.*

Proof. It should be clear that the pairing defines a function $g(\xi)$, for $\xi \in \mathbb{R}^n$.

To prove that it belongs to $\mathcal{O}_M$, we use Theorem 107, obtaining

$$\begin{aligned} \left(\sum_\alpha \partial^\alpha T_{f_\alpha(x)}, e^{2\pi i x\xi}\right) &= \sum_\alpha (-1)^{|\alpha|}(2\pi i\xi)^\alpha \int_{\mathbb{R}^n} e^{2\pi i x\xi} f_\alpha(x) \\ &= \sum_\alpha (-1)^{|\alpha|}(2\pi i\xi)^\alpha \widehat{f_\alpha}(\xi). \end{aligned} \tag{10.13}$$

According to 2) in Theorem 169 and 1) in Theorem 167, the function $\widehat{f_\alpha}$ belongs to $\mathcal{E}\bigcap\mathcal{C}_0$.

Moreover, given $\gamma\in\mathbb{N}^n$, $\partial^\gamma\left(\widehat{f_\alpha}\right)=\mathcal{F}\left[(2\pi ix)^\gamma f_\alpha\right]$ belongs to $\mathcal{C}_0$.

Finally, there is $m\in\mathbb{N}$ so that, given $\beta\in\mathbb{N}^n$,

$$\left|\partial^\beta\left((2\pi i\xi)^\alpha\widehat{f_\alpha}(\xi)\right)\right|\leq C_{m,n,\beta}\sup_{\gamma\leq\beta}\left\|(2\pi ix)^\gamma f_\alpha\right\|_{L^1}\left(1+|\xi|^2\right)^m.$$

Therefore, the function (10.13) belongs to $\mathcal{O}_M$.

This completes the proof of the lemma. □

Theorem 210. *Given $T\in\mathcal{E}'$, $\mathcal{F}[T]$ is the tempered distribution defined by the function* $\left(T_\xi,e^{2\pi ix\xi}\right)_{\mathcal{E}',\mathcal{E}}$.

Proof. It should be clear that $\mathcal{F}[T]\in\mathcal{S}'$.

Then, given $\varphi\in\mathcal{S}$,

$$\begin{aligned}
(\mathcal{F}[T],\varphi) &= \left(T,\int_{\mathbb{R}^n}e^{2\pi ix\xi}\varphi(x)\,dx\right)\\
&= \left(\sum_\alpha\partial^\alpha T_{f_\alpha(\xi)},\int_{\mathbb{R}^n}e^{2\pi ix\xi}\varphi(x)\,dx\right)\\
&= \sum_\alpha(-1)^{|\alpha|}\int_K f_\alpha(\xi)\left(\int_{\mathbb{R}^n}\partial_\xi^\alpha e^{2\pi ix\xi}\varphi(x)\,dx\right)d\xi\\
&= \sum_\alpha(-1)^{|\alpha|}\int_{\mathbb{R}^n}\varphi(x)\left(\int_{\mathbb{R}^n}f_\alpha(\xi)\,\partial_\xi^\alpha e^{2\pi ix\xi}d\xi\right)dx\\
&= \sum_\alpha\int_{\mathbb{R}^n}\left(\partial^\alpha T_{f_\alpha(\xi)},e^{2\pi ix\xi}\right)\varphi(x)\,dx\\
&= \left(\left(\sum_\alpha\partial^\alpha T_{f_\alpha(\xi)},e^{2\pi ix\xi}\right),\varphi\right)=\left(T_{\left(T_\xi,e^{2\pi ix\xi}\right)},\varphi\right).
\end{aligned}$$

This completes the proof of the theorem. □

Remark 211. Lemma 209 and Theorem 210 give a small part of the so-called Paley-Wiener theorem, named after the mathematicians Raymond E. A. C. Payley (1907-1933) and Norbert Wiener (1894-1964).

The theorem is as follows (see [13], p. 21, Theorem 1.7.7.):

1. Given $T\in\mathcal{S}'$, the following statements are equivalent:

   (a) $T\in\mathcal{E}'$ and $supp\,(T)\subseteq\{x\in\mathbb{R}^n:|x|\leq R\}$.

(b) $\mathcal{F}[T]$ is defined by a function $g \in \mathcal{O}_M$.
Moreover g extends to $\mathbb{C}^n$ as an entire function G that satisfies

$$|G(z)| \leq C(1+|z|)^N e^{2\pi R|\mathrm{Im}(z)|},$$

for some $C, N > 0$, where $\mathrm{Im}(z)$ denotes the imaginary part of z.

2. Given $T \in \mathcal{S}'$, the following statements are equivalent:

(a) T is defined by a function $\varphi \in \mathcal{D}$ with $supp(\varphi) \subseteq \{x \in \mathbb{R}^n : |x| \leq R\}$.

(b) $\mathcal{F}[T]$ is defined by a function $g \in \mathcal{S}$.
Moreover, g extends to $\mathbb{C}^n$ as an entire function G.
Furthermore, for each $N > 0$ there is $C_N > 0$ such that

$$|G(z)| \leq C_N (1+|z|)^{-N} e^{2\pi R|\mathrm{Im}(z)|}.$$

Example 212. 1.

$$\mathcal{F}[1] = \delta_0,$$

since $\overline{\mathcal{F}}[\delta_0] = \left(\delta_{0,x}, e^{2\pi i x\xi}\right) = 1$.

2. According to 5) in Example 59, $pv\,1/x$ is a tempered distribution.

Therefore, we can compute its Fourier transform.

Let us recall that we showed in 5) of Example 155 that $pv\,1/x$ is the only odd distribution satisfying the equation $xT = 1$.

So, according to 1) and (10.12),

$$\delta_0 = \mathcal{F}[1] = \mathcal{F}\left[x\left(pv\frac{1}{x}\right)\right] = (2\pi i)^{-1}\frac{d\mathcal{F}\left[pv\frac{1}{x}\right]}{dx}.$$

Using 5) in Example 96,

$$(2\pi i)^{-1}\frac{d\mathcal{F}\left[pv\frac{1}{x}\right]}{dx} = \frac{1}{2}\frac{d}{dx}T_{sgn(\xi)}.$$

Therefore, Theorem 113,implies

$$\mathcal{F}\left[pv\frac{1}{x}\right] = \pi i T_{sgn(\xi)} + C,$$

for some $C \in \mathbb{C}$.

Since $\mathcal{F}[pv\,1/x]$ is odd according to 3) in Theorem 201, and $\pi i T_{sgn(x)}$ is also odd, the constant C must be equal to zero.

So,

$$\mathcal{F}\left[pv\frac{1}{x}\right] = \pi i T_{sgn(\xi)}.$$

3. It is not always possible to calculate the Fourier transform of a distribution, even of a simple one, by naive methods. Here is an example:

   According to Corollary 112, $T_{\ln|x|}$ is a tempered distribution, so, we can consider its Fourier transform.

   Using 2) in Example 96,

$$-2\pi i\xi\mathcal{F}\left[T_{\ln|x|}\right]=\mathcal{F}\left[\frac{d}{dx}T_{\ln|x|}\right]=\mathcal{F}\left[pv\frac{1}{x}\right]=\pi iT_{sgn(\xi)}.$$

   Therefore,

$$-2\xi\mathcal{F}\left[T_{\ln|x|}\right]=T_{sgn(\xi)}$$

   and

$$\mathcal{F}\left[T_{\ln|x|}\right]=-\frac{1}{2}T_{\frac{sgn(\xi)}{\xi}}=-\frac{1}{2}T_{\frac{1}{|\xi|}},$$

   on $\mathbb{R}\setminus\{0\}$.

   According to 3) in Example 43,

$$\mathcal{F}\left[T_{\ln|x|}\right]=-\frac{1}{2}fp\frac{1}{|\xi|}+T_0,$$

   where T_0 denotes a distribution concentrated on $\{0\}$.

   Theorem 110 implies that

$$\mathcal{F}\left[T_{\ln|x|}\right]+\frac{1}{2}fp\frac{1}{|\xi|}=\sum_{k=0}^{m}c_k\delta_0^{(k)}.$$

   Since $\mathcal{F}\left[T_{\ln|x|}\right]$ and $fp\frac{1}{|\xi|}$ are even distributions, 7) in Example 96 implies that

$$\mathcal{F}\left[T_{\ln|x|}\right]=-\frac{1}{2}fp\frac{1}{|x|}+\sum_{k\ even}c_k\delta_0^{(k)}.$$

   Schwartz arrives at the result

$$\mathcal{F}\left[T_{\ln|x|}\right]=-\frac{1}{2}fp\frac{1}{|x|}-(\mathcal{C}+\ln 2\pi)\,\delta_0,$$

   where $\mathcal{C}$ is the Euler's constant, as a particular case of fairly technical calculations ([25], p. 258).

Theorem 213. *Given $T\in\mathcal{E}'$ and $S\in\mathcal{S}'$,*

$$\mathcal{F}\left[T*S\right]=\mathcal{F}\left[T\right]\mathcal{F}\left[S\right], \tag{10.14}$$

in the sense of $\mathcal{S}'$.

Proof. Remark 207, Lemma 209 and Theorem 210 imply that the right-hand side of (10.14) is well defined as a tempered distribution.

Furthermore, 1) in Theorem 189 and Theorem 204, imply that the left-hand side of (10.14) is a distribution in $\mathcal{S}'$.

According to (10.10), we can write

$$\mathcal{F}[T*S]=\sum_{\alpha}\mathcal{F}\left[\partial^{\alpha+\beta}\left(T_{f_\alpha * f}\right)\right]=\sum_{\alpha}\left(-2\pi i\xi\right)^{\alpha}\mathcal{F}\left[T_{f_\alpha}*T_g\right].$$

Given $\varphi\in\mathcal{S}$,

$$\left(\mathcal{F}\left[T_{f_\alpha}*T_g\right],\varphi\right)=\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}f_\alpha\left(y\right)g\left(\xi-y\right)dy\right)\widehat{\varphi}\left(\xi\right)d\xi.$$

Since the double integral exists, we can write

$$\begin{aligned}\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}f_\alpha\left(y\right)g\left(\xi-y\right)dy\right)\widehat{\varphi}\left(\xi\right)d\xi &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}f_\alpha\left(y\right)g\left(\xi-y\right)dy\widehat{\varphi}\left(\xi\right)dyd\xi\\ &\underset{(1)}{=}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}f_\alpha\left(y\right)g\left(z\right)\widehat{\varphi}\left(y+z\right)dydz\\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}f_\alpha\left(y\right)g\left(z\right)\mathcal{F}\left[e^{2\pi ixy}\varphi\right]\left(z\right)dydz\\ &= \int_{\mathbb{R}^n}f_\alpha\left(y\right)\left(\mathcal{F}\left[T_g\right]_x,e^{2\pi ixy}\varphi\left(x\right)\right)dy\\ &= \left(\mathcal{F}\left[T_g\right]_x,\left(\int_{\mathbb{R}^n}f_\alpha\left(y\right)e^{2\pi ixy}dy\right)\varphi\left(x\right)\right)\\ &= \left(\mathcal{F}\left[f_\alpha\right]\mathcal{F}\left[T_g\right],\varphi\right),\end{aligned}$$

where we used in (1) the change of variables $y\to y$ and $\xi\to z=\xi-y$.

This completes the proof of the theorem. □

Remark 214. Theorem 213 is a particular case of a beautiful result, due to Schwartz ([25], Chapter VII, Section 5; Chapter VII, Section 8):

There is a space $\mathcal{O}_C'$, of distributions called rapidly decreasing distributions, of which $\mathcal{E}'$ is a linear subspace.

The space $\mathcal{O}_C'$ consists of those distributions that can be convolved with all the tempered distributions, while $\mathcal{O}_M$ is the space of functions that can be multiplied by all the tempered distributions.

With suitable topologies placed on $\mathcal{O}_C'$ and $\mathcal{O}_M$, the Fourier transform $\mathcal{F}$ and the conjugate Fourier transform $\overline{\mathcal{F}}$, become isomorphisms between $\mathcal{O}_C'$ and $\mathcal{O}_M$ that are inverses of each other.

Moreover, given $T\in\mathcal{O}_C'$ and $S\in\mathcal{S}'$,

$$\mathcal{F}[T*S]=\mathcal{F}[T]\,\mathcal{F}[S]$$

in the sense of $\mathcal{S}'$, while given $f \in \mathcal{O}_M$ and $S \in \mathcal{S}'$,

$$\mathcal{F}[fS] = \mathcal{F}[T_f * S],$$

in the sense of $\mathcal{S}'$.

Remark 215. According to 1) in Example 139,

$$T * \delta_0 = T$$

for $T \in \mathcal{D}'$.

That is to say, δ_0 is a unit for the convolution product.

Let us observe that the convolution product on $\mathcal{S}$ does not have a unit.

Indeed, if there is $\Phi \in \mathcal{S}$ so that

$$\varphi * \Phi = \varphi$$

for all $\varphi \in \mathcal{S}$, taking the Fourier transform on both sides, we will have

$$\widehat{\varphi}\,\widehat{\Phi} = \widehat{\varphi}.$$

If $\varphi = e^{-\pi|x|^2}$, we conclude that the function $\widehat{\Phi}$ must be identically one. This is a contradiction to Theorem 185.

Remark 216. With the notation introduced in 4) of Example 139, let us fix a linear differential operator $P(\partial)$ with constant coefficients.

Given $T \in \mathcal{S}'$, we can write

$$\mathcal{F}[P(\partial)T] = P(-2\pi i\xi)\,\mathcal{F}[T]$$

in the sense of $\mathcal{S}'$.

Then, the formal expression

$$\overline{\mathcal{F}}\left[\frac{\widehat{S}}{P(-2\pi i\xi)}\right]$$

becomes a formal solution of the equation $P(\partial)T = S$.

Schwartz conjectured [22] that every tempered distribution can be divided by a non identically zero polynomial, and that the division should have at least one solution that is a tempered distribution.

This conjecture was confirmed by Lars Hörmander (1931-2012) in [12], and by Stanisław Łojasiewicz (1926-2002) in [17].

As a consequence, every linear differential operator $P(\partial)$ with constant coefficients has a tempered fundamental solution. That is, a solution of the equation

$$P(\partial)T = \delta_0.$$

Therefore, if T is a tempered fundamental solution for $P(\partial)$, a solution of the equation $P(\partial)T = S$, for suitable distributions S, is given as $T * S$.

Remark 217. There are functions, such as $e^{|x|}$, which do not define tempered distributions because they grow too fast at infinity.

However, using a different space of test functions, it is possible to define, in the sense of appropriate distributions, the Fourier transform of such functions [10].

We end with the words of George B. Dantzig (1914-2002): "The final test of a theory is its capacity to solve the problems which originated it." (from the preface of his book *Linear Programming and Extensions*, Princeton University Press 1963).

Undoubtedly, Schwartz's theory of distributions, with its generalizations and extensions, passes Dantzig's test.

References

[1] J. Álvarez and M. Guzmán-Partida, *The T(1) theorem revisited*, Surveys in Mathematics and its Applications, **13** (2018), 41-94. MR3794523. Zbl 1413.42018.

[2] P. Antosik, J. Mikusiński and F. Sikorski, *Theory of Distributions, the Sequential Approach*, Elsevier 1973. MR0365130. Zbl 0267.46028.

[3] P. L. Butzer, *The Hausdorff-Young theorems of Fourier analysis and their impact*, J. Fourier Anal. Appl., **1** (2) (1994), 113-130. MR1348739. Zbl 0978.42500.

[4] K. Chandrasekharan, *Classical Fourier Transforms*, Springer-Verlag, 1989. MR978387. Zbl 0681.42001.

[5] P. Cifuentes, *Harmonic analysis and partial differential equations*, Contemporary Mathematics **505**, Amer. Math. Society 2010. MR2604214. Zbl 1181.35003.

[6] J. F. Colombeau, *A multiplication of distributions*, J. Math. Anal. Appl., **94** (1) (1983), 96-115. MR701451. Zbl 0519.46045.

[7] D. L. Cohn, *Measure Theory*, Birkhäuser, 1980. MR578344. Zbl 0436.28001.

[8] W. F. Donoghue, *Distributions and Fourier Transforms*, Academic Press, 1969. MR3363413. Zbl 0188.18102.

[9] J. D. Gray, *The shaping of the Riesz representation theorem: a chapter in the history of analysis*, Archive of History of Exact Sciences, **31** (2) (1984), 127-187. MR753703. Zbl 0549.01010.

[10] I. Guelfand and G. Chilov, *Les Distributions*, Dunod, 1962. MR0132390. Zbl 0115.10102.

[11] L.-S. Hahn and B. Epstein, *Classical Complex Analysis*, Jones and Bartlett, 1996. Zbl 0877.30001.

[12] L. Hörmander, *On the division of distributions by polynomials*, Arkiv för Mat., **3** (6) (1958), 555-568. MR124734. Zbl 0131.11903.

[13] L. Hörmander, *Linear Partial Differential Operators (Third revised printing)*, Springer-Verlag, 1969. MR0248435. Zbl 0175.39201.

[14] L. Hörmander, *On local integrability of fundamental solutions*, Ark. för Mat., **37** (1) (1999), 121-140. MR1673428. Zbl 1036.35004.

[15] J. Horváth, *An introduction to distributions*, Amer. Math. Monthly, **77** (1970), 227-240. Zbl 0188.43904.

[16] J. Horváth, *Topological Vector Spaces and Distributions*, Dover 2012. Republication of the edition published by Addison-Wesley in 1966. MR0205028. Zbl 0143.15101.

[17] S. Łojasiewicz, *Division d'une distribution par une function analytique de variables réelles*, C. R. Ac. R. Sc. Paris, **246** (1958), 683-686. MR96120. Zbl 0115.10202.

[18] J. Lützen, *The Prehistory of the Theory of Distributions*, Springer-Verlag, 1982. MR667854. Zbl 0494.46038.

[19] J. Norqvist, *The Riesz representation theorem for positive linear functionals*, http://www.diva-portal.org/smash/get/diva2:953904/FULLTEXT01.pdf.

[20] J. Rauch, *Book review*, Bull. Amer. Math. Soc., **37** (3) (2000), 363-367.

[21] W. Rudin, *Real and Complex Analysis*, Third Edition, McGraw-Hill, 1987. MR924157. Zbl 0925.00005.

[22] L. Schwartz, *Théorie des distributions et transformation de Fourier*, Ann. Univ. Grenoble, Sect. Sci. Math. Phys. (N.S.) **23** (1948), 7-24. MR0025615. Zbl 0030.12601.

[23] L. Schwartz, *Théorie des noyaux*, Proc. of the ICM (Cambridge, Mass. 1950) **1**, Providence, R. I. Amer. Math. Soc. (1952), 220-230. MR0045307. Zbl 0048.35102.

[24] L. Schwartz, *Sur l'impossibilité de la multiplication des distributions*, C. R. Aca. Sci. Paris, **239** (1954), 847-848. MR64324. Zbl 0056.10602.

[25] L. Schwartz, *Théorie des Distributions*, Hermann, 1966. Reprinted on February 1973. MR0209834. Zbl 0149.09501.

[26] C. Swartz, *An Introduction to Functional Analysis*, Marcel Dekker Inc. 1992. MR1156078. Zbl 0751.46002.

[27] C. Swartz, *Measure, Integration and Function Spaces*,World Scientific, 1994. MR1337502. Zbl 0814.28001.

[28] J. Synowiec, *Book review*, Amer. Math. Monthly,**103** (5) (1996), 435-440.

[29] R. L. Wheeden and A. Zygmund, *Measure and Integral: An Introduction to Real Analysis*, Marcel Dekker, 1977. MR0492146. Zbl 0362.26004.

Josefina Alvarez
Department of Mathematical Sciences,
New Mexico State University,
Las Cruces, NM 88003, USA.
e-mail: jalvarez@nmsu.edu

License